Topologically driven no-superposing theorem with a tight error bound

Zuzana Gavorová[1,2]

[1]Rachel and Selim Benin School of Computer Science and Engineering, The Hebrew University of Jerusalem, 91904 Jerusalem, Israel

[2]Física Teòrica: Informació i Fenòmens Quàntics, Departament de Física, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain

To better understand quantum computation we can search for its limits or no-gos, especially if analogous limits do not appear in classical computation. Classical computation easily implements and extensively employs the addition of two bit strings, so here we study 'quantum addition': the superposition of two quantum states. We prove the impossibility of superposing two unknown states, no matter how many samples of each state are available. The proof uses topology; a quantum algorithm of any sample complexity corresponds to a continuous function, but the function required by the superposition task cannot be continuous by topological arguments. Our result for the first time quantifies the approximation error and the sample complexity N of the superposition task, and it is tight. We present a trivial algorithm with a large approximation error and $N = 1$, and the matching impossibility of any smaller approximation error for any N. Consequently, our results limit state tomography as a useful subroutine for the superposition. State tomography is useful only in a model that tolerates randomness in the superposed state. The optimal protocol in this random model remains open.

1 Introduction

Quantum computing promises remarkable technological advances, some based on discoveries that were initially seen as limitations of quantum mechanics. The no-cloning theorem [1] is an example of an uncovered limitation that turned into an advantage. It establishes the security of quantum cryptography, starting with quantum key distribution [2, 3], and deepens our understanding of quantum mechanics all the way to its first principles; the first principles are often derived from, or checked against, the no-cloning theorem [4, 5, 6, 7, 8]. Cloning distinguishes between classical and quantum mechanics, being easy in one and impossible in the other. The same is true, for example, for programming [9], the universal-NOT gate [10] and deleting [11].

First, we would like to motivate another such task: the superposition. Given some predetermined weights $\alpha, \beta \in \mathbb{R}_+$; the task is to output a state proportional to

$$\alpha \left|u\right\rangle + \beta \left|v\right\rangle \tag{1}$$

from any two input states $\left|u\right\rangle, \left|v\right\rangle \in \hat{\mathcal{H}}$, where $\mathcal{H}$ is a finite-dimensional Hilbert space and $\hat{\mathcal{H}} \subset \mathcal{H}$ denotes the subset of unit vectors[1]. The superposition can be seen as a quantum version of addition. Graphically in (1), the superposition corresponds to inserting + between the two rescaled vectors. How hard is it to implement this +? This fundamental question is interesting in its own right.

Moreover, if easily implementable, + could serve as an elementary gate for quantum computation. Consider Aaronson's *quantum state tree* model [12, 13] whose elementary operations are + and the trivially implementable tensor product. In ref. [12] the purpose of the quantum state tree model was to capture the hardness of a state's classical description, but suppose for a moment

[1]The precise definition will specify the relative phase between the α and β terms, so no generality is lost by assuming $\alpha, \beta \in \mathbb{R}_+$.

that + was easily implementable. Then, the quantum state tree model could directly instruct a state's preparation process. This way to classically describe a state would also correspond to a new approach to physically preparing it.

The superposition question was studied previously. Algorithms were found for cases when the input states are restricted or partially known [12, 14, 15, 16, 17, 18, 19]. For example, algorithms exist for inputs restricted to orthogonal pairs if these inputs are qubits [14] or have known preparation circuits [12][2]. Another restriction of the inputs that makes an algorithm possible fixes their overlap with a reference state [15]. When the input states are unrestricted and completely unknown, Oszmaniec, Grudka, Horodecki, and Wójcik [15] showed that a quantum circuit cannot create an exact superposition from one copy of each input.

In this work, we define and study the superposition question when the input states $|u\rangle, |v\rangle \in \hat{\mathcal{H}}$ are completely unrestricted and unknown, N copies of each input are available, and approximations are allowed. We find that in the most commonly considered models of quantum computation preparing such a superposition is impossible for any N; some inputs necessarily produce a large approximation error in the output. This lower bound on the error is tight. Fortunately, complicating the computational model overcomes the impossibility, as it allows employing state tomography in a useful way. Therefore, our strengthened superposition impossibility has new consequences. In addition to widening the operational distinction between the quantum and classical theories that follows already from the impossibility of ref. [15], our results yield a new separation between different *quantum* computational models, and a better understanding of the topological method developed in [20] and employed here.

1.1 Precise problem statement

Correctly formulating the problem for completely unknown inputs has a subtlety. Oszmaniec et al. [15] observed that for unknown $|u\rangle, |v\rangle \in \hat{\mathcal{H}}$, preparing (1) is trivially impossible. Suppose an equal-weight superposition protocol succeeds for $|u\rangle = |0\rangle$ and $|v\rangle = |1\rangle$, outputting $|0\rangle + |1\rangle$. If we change $|1\rangle$ to $-|1\rangle$ should we expect $|0\rangle - |1\rangle$? Distinguishing the global phase of a vector is unphysical, and thus trivially impossible for any protocol. Taking this into account, our superposition definition should again accept $|0\rangle + |1\rangle$ as the correct answer. It should accept outputs that ignore the global phases of $|u\rangle, |v\rangle$, and instead depend only on the physically relevant representation of the inputs – as density matrices. Instead of the unnormalized state in (1), we can consider, for example,

$$|s_{+(u,v)}\rangle = \alpha\,|u\rangle\,|\langle v|u\rangle| + \beta\,|v\rangle\langle v|u\rangle\,, \tag{2}$$

which ensures that $|s_{+(u,v)}\rangle\langle s_{+(u,v)}|$ is a function of the density matrices $|u\rangle\langle u|$, $|v\rangle\langle v|$. The renormalization of $|s_{+(u,v)}\rangle$ is well defined on all $|u\rangle$, $|v\rangle$ nonorthogonal, which is a dense subset of $\hat{\mathcal{H}} \times \hat{\mathcal{H}}$. We generalize (2) for our final definition.

Definition 1. A superposition of $|u\rangle, |v\rangle \in \hat{\mathcal{H}}$ is a state proportional to

$$|s_{c(u,v)}\rangle = \alpha\,|u\rangle\,|c_{uv}| + \beta\,|v\rangle\,c_{uv}, \tag{3}$$

for some complex function $c_{uv} = c(|u\rangle, |v\rangle)$ such that $S := c^{-1}(\mathbb{C} \setminus \{0\})$ is a dense subset of $\hat{\mathcal{H}} \times \hat{\mathcal{H}}$.

Compared to Oszmaniec et al. [15][3], this article formalizes the superposition differently: It defines the superposition of $|u\rangle$ and $|v\rangle$ as a *deterministic function* of $|u\rangle$ and $|v\rangle$, but it also considers *sets* of superpositions (each superposition coming from a different c-function). This allows us to postpone the (non)determinism nuance to our later discussion of computational models.

Definition 1 includes (2) and other choices of the form $c_{uv} = \langle v|\,M_{uv}\,|u\rangle$. If the matrix M_{uv} depends only on the inputs' density matrices, then so does the superposition. For $c_{uv} = \langle v|j\rangle\langle i|u\rangle$

[2] See the proof of Theorem 6 (OTree $\subseteq$ ΨP) in [12].

[3] The definition of Oszmaniec et al. is equivalent to the following: $|s\rangle$ is a superposition of $|u\rangle$ and $|v\rangle$ with coefficients α, β iff $|s\rangle\langle s| = N(\alpha\,|u\rangle + e^{i\phi}\beta\,|v\rangle)(\alpha^*\,\langle u| + e^{-i\phi}\beta^*\,\langle v|)$ for some real N, ϕ. Restricting ϕ to be a deterministic function of $|u\rangle$ and $|v\rangle$ falls within our definition with $e^{i\phi} = c_{uv}/|c_{uv}|$. This ϕ-function can be undefined on $|u\rangle, |v\rangle$ outside the dense subset.

with $i, j \in \{0, 1, \ldots, \dim \mathcal{H} - 1\}$, the resulting superposition is (on the corresponding dense subset) proportional to

$$\alpha \left|u\right\rangle \frac{\langle u|i\rangle}{|\langle u|i\rangle|} + \beta \left|v\right\rangle \frac{\langle v|j\rangle}{|\langle v|j\rangle|} = \alpha \, \mathrm{v}_i(|u\rangle\langle u|) + \beta \, \mathrm{v}_j(|v\rangle\langle v|). \tag{4}$$

In this particular choice of superposition we can treat each input separately; each input has its own function $|u\rangle\langle u|$ has v_i, $|v\rangle\langle v|$ has v_j) that maps the rank-1 density matrix to a corresponding[4] vector. The function v_i chooses the vector by fixing its i-th entry, $\langle i| \, \mathrm{v}_i(\rho)$, to be real positive.

In this work we study superposition protocols, considering their error and sample complexity. The error is quantified by the trace-norm distance from the target output, the superposition. We are interested in the worst-case performance of the protocols; the greatest error obtained for any valid input $(\left|u\right\rangle, \left|v\right\rangle) \in S$, i.e., any input such that the target output is renormalizable. Demanding worst-case guarantees is common in complexity theory and the reason is practical. An unwarranted error on some inputs is undesirable, especially if the solution is to be used as an elementary operation, as is the case for $+$. Worst-case error bounds are satisfied, for example, by state tomography protocols [21, 22], which also set the expectation for our second performance measure: sample complexity, or the number of input copies used by the protocol. The following superposition protocol seems possible with $N = \exp(n)$ samples of each n-qubit input ($n = \lceil \log_2 \dim \mathcal{H} \rceil$): Perform state tomography to get an approximate classical description of each of the two input states, then classically describe their superposition, then build a new circuit that transforms the all-zero input into this described state[5]. We might, therefore, expect the optimal protocol's sample complexity to exceed 1, due to the previous no-go result [15], and not to exceed $\exp(n)$, due to state tomography. Our results below defy this expectation, revealing the importance of distinguishing between different computational models.

1.2 Overview of the results in different computational models

The answer to our superposition question depends on which model of computation we consider: the trace-preserving model, the postselection (probabilistic) model, or the random model. For each we give a definition, an example protocol, and our result.

Trace-preserving model. The trace-preserving model with sample complexity N is a completely positive trace-preserving map (CPTP) applied to $\rho_{\text{in}}^{\otimes N}$, for example, to $(|u\rangle\langle u| \otimes |v\rangle\langle v|)^{\otimes N}$. The output approximates a given target output, for example, a given superposition $\alpha \left|u\right\rangle \left|c_{uv}\right| + \beta \left|v\right\rangle c_{uv}$ determined by a particular c-function.

Next, we give an example protocol in the trace-preserving model for the target superposition $|s_{+(u,v)}\rangle$. The protocol Ψ accepts $N = 1$ samples of each state and outputs the mixed state $\cos^2\theta \, |u\rangle\langle u| + \sin^2\theta \, |v\rangle\langle v|$, where $\cos\theta = \alpha/\sqrt{\alpha^2 + \beta^2}$. A probabilistic swap (Figure 1a) implements this Ψ. For $\left|u\right\rangle, \left|v\right\rangle \in S$ (nonorthogonal pairs from $\hat{\mathcal{H}}$) it satisfies

$$\left|\left|\Psi(|u\rangle\langle u| \otimes |v\rangle\langle v|) - \tfrac{|s_+(u,v)\rangle\langle s_+(u,v)|}{\langle s_+(u,v)|s_+(u,v)\rangle}\right|\right|_{\text{tr}} = r\sqrt{\tfrac{1-r|\langle v|u\rangle|}{1+r|\langle v|u\rangle|}}\sqrt{1 - |\,\langle v|u\rangle\,|^2}, \tag{5}$$

$$\text{where } r = \sin(2\theta) = \tfrac{2\alpha\beta}{\alpha^2+\beta^2}, \tag{6}$$

and $||O||_{\text{tr}} := \text{tr}\sqrt{OO^\dagger}$ so the maximal trace distance of two density states is 2. Above, the output state is the furthest from the desired state as $\langle v|u\rangle \to 0$, when the trace distance equals r (see Figure 2). We say the worst-case error of this protocol is r. Our first result discusses the worst-case performance of *any* protocol in the trace-preserving model.

Result 1. For any choice of target superposition and any sample complexity N, no trace-preserving protocol achieves a worst-case error smaller than r. In light of the above protocol, this impossibility is tight.

Postselection model. The postselection model with sample complexity N corresponds to a completely positive trace-nonvanishing map (CPTNV) applied to $\rho_{\text{in}}^{\otimes N}$. Only trace-nonincreasing

[4] A vector $|w\rangle \in \hat{\mathcal{H}}$ corresponds to $\rho \in \mathcal{D}_{\text{pure}}(\mathcal{H})$ iff $|w\rangle\langle w| = \rho$.

[5] This is always possible by the universality of quantum circuits.

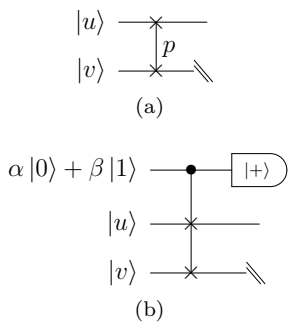


Figure 1: (a) The probabilistic swap protocol swaps the two registers with probability $p = \sin^2\theta$ and does not swap them with probability $1 - p = \cos^2\theta$. Then, the second register is traced out. (b) The controlled swap protocol has an additional control qubit to control the controlled swap and a control measurement to postselect the $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ outcome.

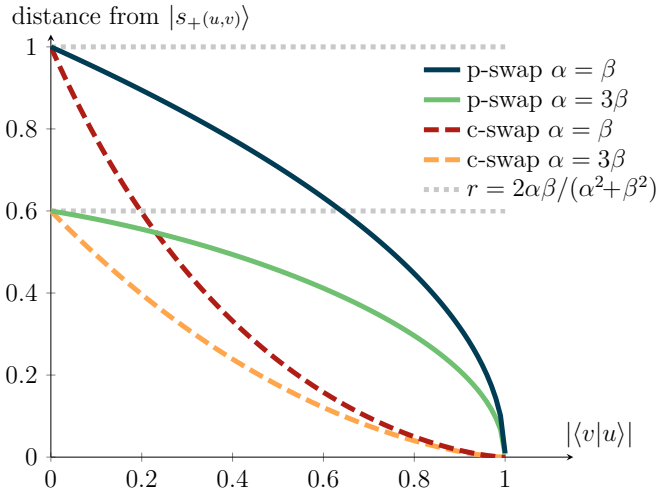


Figure 2: The error as a function of the inputs for different protocols and $\alpha : \beta$ ratios. The controlled swap protocol (c-swap) outperforms the probabilistic swap protocol (p-swap) on almost all inputs, but the worst-case error of both is r.

maps are physical. Trace-nonvanishing means that for any input state, the map outputs a nonzero[6] positive semidefinite matrix so that it can be renormalized and then compared to a given target output. The postselection model generalizes the trace-preserving model.

Our example protocol Φ is again for the $|s_{+(u,v)}\rangle$ superposition and $N = 1$. It initializes a control qubit to $\cos\theta\,|0\rangle + \sin\theta\,|1\rangle$, then controlled-swaps $|u\rangle$ and $|v\rangle$, then projects the control onto the $|+\rangle$ state (Figure 1b). For nonorthogonal $|u\rangle, |v\rangle$ this CPTNV Φ has the error

$$\left|\left| \frac{\Phi(|u\rangle\langle u| \otimes |v\rangle\langle v|)}{\mathrm{tr}\left[\Phi(|u\rangle\langle u| \otimes |v\rangle\langle v|)\right]} - \frac{|s_{+(u,v)}\rangle\langle s_{+(u,v)}|}{\langle s_{+(u,v)}|s_{+(u,v)}\rangle} \right|\right|_{\mathrm{tr}} = r\sqrt{\frac{1 - r|\langle v|u\rangle|}{1 + r|\langle v|u\rangle|}}\sqrt{1 - |\langle v|u\rangle|^2}\,\frac{1 - |\langle v|u\rangle|}{1 + r|\langle v|u\rangle|^2}, \tag{7}$$

an improvement for $|\langle v|u\rangle| \in (0, 1)$, but not in the worst case, when the error again equals r (Figure 2). Indeed, our matching lower bound extends to the postselection model.

Result 1.* The worst-case trace distance r is optimal for any N and any choice of superposition also in the postselection model.

Random model. The random model with sample complexity N corresponds to a quantum instrument $\{\Psi_i\}_{i\in I}$ (an indexed set of CP maps that sum up to a CPTP) applied to $\rho_{\mathrm{in}}^{\otimes N}$. We obtain the label i with probability $\mathrm{tr}\left[\Psi_i(\rho_{\mathrm{in}}^{\otimes N})\right]$, in which case the algorithm's quantum output $\Psi_i(\rho_{\mathrm{in}}^{\otimes N})$ approximates (up to the renormalizations) the i-th target output, for example, the i-th superposition $\alpha\,|u\rangle\,|c_{i\,uv}| + \beta\,|v\rangle\,c_{i\,uv}$. Thus, in this model we request an indexed *set* of $|I|$ targets. Which target's approximation is the output of a particular run of the protocol is random, but revealed by the classical outcome i. This model includes the trace-preserving model (obtained by setting $|I| = 1$).

An example protocol in the random model is the protocol for qubit inputs $|u\rangle, |v\rangle \in \mathbb{C}^2$ of Doosti, Kianvash and Karimipour [14], which has $N = 1$ and zero error if the inputs are restricted to orthogonal states, $\langle v|u\rangle = 0$. An example protocol with unrestricted inputs is the state tomography subroutine suggested above as a building block of a possible superposition protocol. The naïve suggestion is to run a state tomography for both $|u\rangle$ and $|v\rangle$, to use the resulting classical estimates to calculate a classical description of a state close to

$$(\alpha\,|u\rangle + \beta\,|v\rangle)\,(\alpha\,\langle u| + \beta\,\langle v|)\,, \tag{8}$$

[6]The nonzero requirement, also appearing in the definition of the complexity class PostBQP [23], allows the model to remain worst-case.

and then to build a new quantum circuit that outputs a state that is arbitrarily close to that classical description (see footnote 5). To estimate (8), classical estimates of the *density matrices* $|u\rangle\langle u|$, $|v\rangle\langle v|$ do not suffice, because eq. (8) contains cross-terms. Thus, we ask for *vector tomography*, a state tomography that outputs a classical estimate of a vector $\mathrm{v}(|u\rangle\langle u|) \in \hat{\mathcal{H}}$.

Result 2. Vector tomography is not possible in the trace-preserving and postselection models.

Result 3. Vector tomography is possible in the random model.

In the above definitions, randomness enters protocols in two crucially distinct ways. A postselection protocol randomly (sometimes) generates an output, but the output itself is deterministic. A random protocol deterministically (always) generates an output, but the output itself is random. The results reveal that this distinction is important; the random model is strictly stronger than the postselection model. As a corollary of Result 3, vector tomography gives a superposition protocol in the random model, with sample complexity exponential in $n = \lceil \log_2 \dim \mathcal{H} \rceil$. The optimality of this protocol remains open.

1.3 Topology as a tool to prove no-go results

Results 1, 1*, and 2 are all no-go results. They use the method developed in [20], which combines two observations:

1. any quantum circuit corresponds to a continuous function,
2. the task requires implementing a function that cannot be continuous by topological arguments.

In ref. [20], the function's input is a unitary oracle. Here, it is a pair of states. Specifically, the first observation corresponds to:

Proposition 1. *If a postselection algorithm takes as the input any number of copies of $|u\rangle\langle u|$ and $|v\rangle\langle v|$, then its output is a continuous function of $|u\rangle\langle u|, |v\rangle\langle v|$.*

This is a weaker version of the well-known observation that any quantum circuit corresponds to a polynomial [24]. A lower bound on the polynomial degree of the required function implies a lower bound on the sample complexity of any implementing algorithm. Continuity is more general. Disproving the continuity of the required function excludes algorithms of *any* sample complexity. Can we relate our problem to some continuous functions forbidden by topology? Comparing to [20] we suggest two functions.

The first function is a *section*. Topology forbids a continuous function, also called *section*, that maps a pure density operator to a corresponding vector. In other words, no continuous $\mathrm{v} : \mathcal{D}_{\mathrm{pure}}(\mathcal{H}) \to \hat{\mathcal{H}}$ satisfies $\mathrm{v}(\rho)\,\mathrm{v}(\rho)^\dagger = \rho$, i.e., right-inverts the outer product map[7]. This is analogous to no-section for unitaries [20], where no continuous function maps a unitary superoperator $\mathcal{U}(\cdot) = U \cdot U^\dagger$ to a "canonical" unitary $\mathcal{U} \mapsto \widetilde{U}$, such that $\widetilde{U} \cdot \widetilde{U}^\dagger = \mathcal{U}$. For states, the nonexistence of a section yields a no-go for the exact superposition when $\dim \mathcal{H} \geq 3$ (see Section A for the proofs). To deal with approximations, we need to "quantify" the discontinuity in any v. Intuitively, the desired maps $\mathrm{v} : |u\rangle\langle u| \mapsto |\widetilde{u}\rangle$ and $U \cdot U^\dagger \mapsto \widetilde{U}$ have something in common; as opposed to the domains, the ranges distinguish the global phase $\lambda \in \mathrm{U}(1) \approx \mathrm{S}^1$, where we note the homeomorphism between $\mathrm{U}(1)$ and the circle S^1 (implicit from now on). The next function extracts this S^1 structure to contradict continuity, exploiting the nontrivial topology of the circle S^1: different numbers of loops on S^1 cannot be continuously contracted to each other.

The second function is a *continuous homogeneous function into* S^1. For $m \in \mathbb{Z}$, a function f is m-homogeneous, if for all scalars λ and inputs x, it satisfies $f(\lambda x) = \lambda^m f(x)$. If the domain of f is $\hat{\mathcal{H}}$, we only consider $|\lambda| = 1$. For example, on $\hat{\mathcal{H}}$ the outer product map $|u\rangle \mapsto |u\rangle\langle u|$ is 0-homogeneous. We find excluding *continuous homogeneous functions into* S^1 more useful for quantifying approximations. Specifically, denoting by $\hat{\mathbb{C}}^2$ the space of norm-1 vectors in $\mathbb{C}^2$, our no-go results rely on the following lemma:

[7] The outer product map $\pi : \hat{\mathcal{H}} \to \mathcal{D}_{\mathrm{pure}}(\mathcal{H})$, $|u\rangle \mapsto |u\rangle\langle u|$ is the familiar Hopf fibration. When the input $|u\rangle = (u_0, u_1)^T \in \hat{\mathbb{C}}^2$ is a qubit, the Hopf fibration $\pi(|u\rangle) = |u\rangle\langle u| = (I + r \cdot \vec{\sigma})/2$ with $r = (\mathrm{Re}[2u_0^* u_1], \mathrm{Im}[2u_0^* u_1], |u_0|^2 - |u_1|^2)^T$ lets us represent the qubit on the Bloch sphere, because $\mathcal{D}_{\mathrm{pure}}(\mathbb{C}^2)$ is homeomorphic to the 2-sphere ($r \in \mathrm{S}^2$).

Lemma 1. *Let $m \neq 0$. There is no continuous m-homogeneous function $\hat{\mathbb{C}}^2 \to \mathrm{S}^1$.*

Next, in Section 2 we prove Results 1 and 1* using Lemma 1, then we prove Lemma 1. In Section 3 we prove Results 2 and 3. In Section 4 we discuss the significance of our results for setting quantum mechanics apart from classical, for harnessing the power of random outputs, and for understanding the applicability of the topological method.

2 The superposition impossibility

Results 1 and 1* are the superposition no-gos in the trace-preserving and the postselection models. Since the trace-preserving model is a special case of the postselection model, we focus on the latter. Denote by $\mathcal{D}_+(\mathcal{H})$ the set of positive semidefinite and nonzero linear operators from $\mathcal{H}$ to itself. By definition, the outputs of a postselection algorithm are in $\mathcal{D}_+(\mathcal{H})$.

Theorem 1. *Let $\alpha, \beta \in \mathbb{R}_+$. Let any function $c : \hat{\mathcal{H}} \times \hat{\mathcal{H}} \to \mathbb{C}$, such that $S := c^{-1}(\mathbb{C} \setminus \{0\})$ is dense in $\hat{\mathcal{H}} \times \hat{\mathcal{H}}$, determine the superposition $|s_{c(u,v)}\rangle = \alpha\,|c(u,v)|\,|u\rangle + \beta\,c(u,v)\,|v\rangle$. There is no continuous function $\mathcal{A} : \mathcal{D}_{pure}(\mathcal{H}) \times \mathcal{D}_{pure}(\mathcal{H}) \to \mathcal{D}_+(\mathcal{H})$ such that*

$$\sup_{(|u\rangle,|v\rangle)\in S} \left|\left| \frac{\mathcal{A}(|u\rangle\langle u|\,,|v\rangle\langle v|)}{\mathrm{tr}\left[\mathcal{A}(|u\rangle\langle u|\,,|v\rangle\langle v|)\right]} - \frac{|s_{c(u,v)}\rangle\langle s_{c(u,v)}|}{\langle s_{c(u,v)}|s_{c(u,v)}\rangle} \right|\right|_{\mathrm{tr}} < \frac{2\alpha\beta}{\alpha^2+\beta^2}.$$

Combining with Proposition 1 gives the no-gos.

Proof. Assume towards contradiction that a function $\mathcal{A}$ of Theorem 1 exists. We will use it to build a function excluded by Lemma 1. Let $u = (u_0, u_1)^T \in \hat{\mathbb{C}}^2$. We define vectors in $\hat{\mathcal{H}}$ that are homogeneous functions of u by adding trailing zeros, namely

$$|u\rangle := \begin{pmatrix} u_0 \\ u_1 \\ 0 \\ \vdots \end{pmatrix}, \quad |u^\perp\rangle := \begin{pmatrix} u_1^* \\ -u_0^* \\ 0 \\ \vdots \end{pmatrix}, \quad \text{and } \langle u^\perp| = \begin{pmatrix} u_1 & -u_0 & 0 & \dots \end{pmatrix}$$

are respectively 1-homogeneous, (-1)-homogeneous, and 1-homogeneous. Then the following function

$$g(u) := \langle u^\perp| \frac{\mathcal{A}\big(\,|u\rangle\langle u|\,,|u^\perp\rangle\langle u^\perp|\,\big)}{\mathrm{tr}\left[\mathcal{A}(|u\rangle\langle u|\,,|u^\perp\rangle\langle u^\perp|)\right]} |u\rangle\,, \tag{9}$$

is a continuous, 2-homogeneous function $\hat{\mathbb{C}}^2 \to \mathbb{C}$. Since the matrix inside the contractive map $\langle u^\perp| \cdot |u\rangle$ is positive semidefinite and trace-1, we have $|g(u)| \leq 1$. In the remaining part of the proof we show that $|g(u)| > 0$ for all $u \in \hat{\mathbb{C}}^2$, which allows us to define $\hat{g}(u) := g(u)/|g(u)|$. The map $\hat{g}$ is also continuous and 2-homogeneous and, moreover, maps into S^1. The existence of such $\hat{g}$ contradicts Lemma 1.

To prove $|g(u)| > 0$ we first observe that even if $c(|u\rangle\,,|u^\perp\rangle) = 0$, the set $S := c^{-1}(\mathbb{C} \setminus \{0\})$ is dense in $\hat{\mathcal{H}} \times \hat{\mathcal{H}}$, so inside S there exists a converging subsequence $(|u_n\rangle\,,|u_n^\perp\rangle)$ such that

$$\lim_{n\to\infty} |u_n\rangle = |u\rangle$$
$$\lim_{n\to\infty} |u_n^\perp\rangle = |u^\perp\rangle\,.$$

We choose this notation because $\lim_{n\to\infty} \langle u_n^\perp|u_n\rangle = 0$, but note that for a given n the inner product $\langle u_n^\perp|u_n\rangle$ might be nonzero. Our assumption towards contradiction implies

$$\frac{2\alpha\beta}{\alpha^2+\beta^2} > \lim_{n\to\infty} \left|\left| \frac{\mathcal{A}(|u_n\rangle\langle u_n|\,,|u_n^\perp\rangle\langle u_n^\perp|)}{\mathrm{tr}\left[\mathcal{A}(|u_n\rangle\langle u_n|\,,|u_n^\perp\rangle\langle u_n^\perp|)\right]} - \frac{|s_{c(u_n,u_n^\perp)}\rangle\langle s_{c(u_n,u_n^\perp)}|}{\langle s_{c(u_n,u_n^\perp)}|s_{c(u_n,u_n^\perp)}\rangle} \right|\right|_{\mathrm{tr}}. \tag{10}$$

Denote by X_n the entire expression inside the trace norm above. X_n is Hermitian. For any positive trace-nonincreasing superoperator Ψ, $||X_n||_{\mathrm{tr}} \geq ||\Psi(X_n)||_{\mathrm{tr}}$. Choose $\Psi(X_n) = |0\rangle\langle u^\perp|\,X_n\,|u\rangle\langle 0| + |1\rangle\langle u|\,X_n\,|u^\perp\rangle\langle 1|$. This Ψ is (completely) positive trace-nonincreasing, because it corresponds to

the conjugation by the operator $|00\rangle\langle u^\perp| + |11\rangle\langle u|$ followed by a partial trace. Since X_n is Hermitian, we get $||X_n||_{\mathrm{tr}} \geq ||\Psi(X_n)||_{\mathrm{tr}} = 2\,|\langle u^\perp|\,X_n\,|u\rangle|$. Substituting this into (10), cancelling the factors of 2, and using the triangle inequality, we get

$$\frac{\alpha\beta}{\alpha^2+\beta^2} > \lim_{n\to\infty}\left[-\left|\langle u^\perp|\,\frac{\mathcal{A}(|u_n\rangle\langle u_n|\,,|u_n^\perp\rangle\langle u_n^\perp|)}{\mathrm{tr}\left[\mathcal{A}(|u_n\rangle\langle u_n|\,,|u_n^\perp\rangle\langle u_n^\perp|)\right]}\,|u\rangle\right| + \left|\frac{\langle u^\perp|s_{\mathrm{c}(u_n,u_n^\perp)}\rangle\,\langle s_{\mathrm{c}(u_n,u_n^\perp)}|u\rangle}{\langle s_{\mathrm{c}(u_n,u_n^\perp)}|s_{\mathrm{c}(u_n,u_n^\perp)}\rangle}\right|\right] \tag{11}$$

Inside the limit, the first term converges to $|g(u)| \in [0,1]$. Denoting $c(|u_n\rangle\,,|u_n^\perp\rangle) =: c_n = |c_n|e^{i\gamma_n}$ the second term converges to

$$\frac{|c_n|^2\left|\langle u^\perp|\left(\alpha\,|u_n\rangle + e^{i\gamma_n}\beta\,|u_n^\perp\rangle\right)\left(\alpha\,\langle u_n| + e^{-i\gamma_n}\beta\,\langle u_n^\perp|\right)|u\rangle\right|}{|c_n|^2\left(\alpha^2+\beta^2+2\alpha\beta\,\mathrm{Re}\left[e^{i\gamma_n}\,\langle u_n^\perp|u_n\rangle\right]\right)} \xrightarrow{n\to\infty} \frac{|e^{i\gamma_n}\beta\alpha|}{\alpha^2+\beta^2}.$$

Inequality (11) becomes

$$\frac{\alpha\beta}{\alpha^2+\beta^2} > -\,|g(u)| + \frac{|e^{i\gamma_n}\beta\alpha|}{\alpha^2+\beta^2},$$

which implies $|g(u)| > 0$, the missing step to establish the contradiction with Lemma 1. □

Proof of Lemma 1. We parametrize paths on $\hat{\mathbb{C}}^2$ by the parameter $t \in [0,1]$. The following paths are loops on $\hat{\mathbb{C}}^2$ based at a fixed point $|u_b\rangle \in \hat{\mathbb{C}}^2$:

$$f(t) := |u_b\rangle\,, \qquad f'(t) := e^{i2\pi t}\,|u_b\rangle\,.$$

Namely, f is constant and f' loops in its global phase. The paths are continuously contractible to each other, or homotopic, because on $\hat{\mathbb{C}}^2 \approx \mathrm{S}^3$ any loop is continuously shrinkable to a point – we say that S^3 is simply connected. See Fig. 3 for an intuition.

Assume towards contradiction that there is a function $g : \hat{\mathbb{C}}^2 \to \mathrm{S}^1$ which is continuous and m-homogeneous for $m \neq 0$, $m \in \mathbb{Z}$. Composing g onto our loops on $\hat{\mathbb{C}}^2$ gives us loops on the circle, S^1:

$$\begin{aligned}(g\circ f)(t) &= g(|u_b\rangle) =: \lambda_b\\ (g\circ f')(t) &= g(e^{i2\pi t}\,|u_b\rangle) = e^{i2\pi t m}\lambda_b.\end{aligned}$$

The function $g\circ f$ is constant – it corresponds to zero loops on S^1, while $g\circ f'$ by the m-homogeneity of g makes $|m|$ loops on S^1. By the continuity of g the loops $g\circ f$ and $g\circ f'$ must also be continuously deformable to each other. But on the circle S^1 this is impossible for different numbers of loops. □

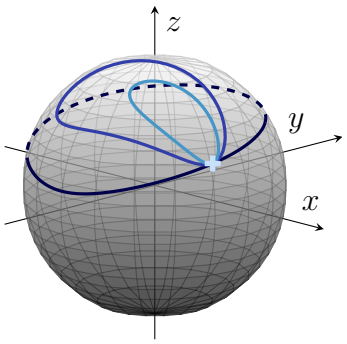


Figure 3: Any closed loop on S^2 is continuously contractible to a point. The same is true for S^3 which is homeomorphic to $\hat{\mathbb{C}}^2$. We say that these spaces are simply connected.

3 State tomography for superposition

Here we prove Results 2 and 3 about vector tomography. Instead of outputting a classical description of a pure density state in $\mathcal{D}_{\text{pure}}(\mathcal{H})$, vector tomography outputs a classical description of a vector in $\hat{\mathcal{H}}$. However, different vectors in $\hat{\mathcal{H}}$ match the same rank-1 matrix in $\mathcal{D}_{\text{pure}}(\mathcal{H})$. Choosing one (canonical) vector for each matrix is the purpose of the function $\mathrm{v} : T \to \hat{\mathcal{H}}$, where T is dense in $\mathcal{D}_{\text{pure}}(\mathcal{H})$. The requirement $\mathrm{v}(\rho)\,\mathrm{v}(\rho)^\dagger = \rho$ ensures that the vector matches the matrix. An example is

$$\mathrm{v}_0(\rho) = \frac{\rho\,|0\rangle}{\sqrt{\langle 0|\,\rho\,|0\rangle}}, \tag{12}$$

which accepts inputs with a nonzero first matrix element, $T = \{\rho \in \mathcal{D}_{\text{pure}}(\mathcal{H}); \langle 0|\,\rho\,|0\rangle \neq 0\}$, and chooses a vector with the first element $\langle 0|\,\mathrm{v}_0(\rho)$ real positive.

Vector tomography should satisfy a precision guarantee similar to that of state tomography. State tomography protocols [21, 22] satisfy a guarantee in terms of trace norm $||\cdot||_{\text{tr}}$; given N samples of ρ they output a classical description x from a set of values T_N according to a probability distribution $p_{N,\rho}(x)$ such that

$$Pr_{x\sim p_{N,\rho}(\cdot)}\left[||x-\rho||_{\text{tr}} \leq \epsilon_N^{\text{tr}}\right] \geq 1-\delta_N^{\text{tr}}, \tag{13}$$

where $\lim_{N\to\infty}\epsilon_N^{\text{tr}} = 0$ and $\lim_{N\to\infty}\delta_N^{\text{tr}} = 0$. The classical output x is a result of quantum measurements and a classical postprocessing, which can ensure $x \in \mathcal{D}_{\text{pure}}(\mathcal{H})$ (as in our case $\rho \in \mathcal{D}_{\text{pure}}(\mathcal{H})$). The procedure is in principle convertible to a fully quantum algorithm with a deferred measurement, so in (13) we can take $x \in T_N \subset \mathcal{D}_{\text{pure}}(\mathcal{H})$. This equation-(13) characterization of state tomography inspires our following, equation-(14) characterization of vector tomography.

Definition 2 (Vector tomography)**.** A sequence of T_N-algorithms is a *vector tomography* if $T := \bigcup_{N=0}^{\infty} T_N$ is dense in $\mathcal{D}_{\text{pure}}(\mathcal{H})$, if there exists $\mathrm{v} : T \to \hat{\mathcal{H}}$ such that $\mathrm{v}(\rho)\,\mathrm{v}(\rho)^\dagger = \rho$, and if the N-th algorithm works as follows: Its quantum part on input $\rho^{\otimes N}$ where $\rho \in \mathcal{D}_{\text{pure}}(\mathcal{H})$ outputs a measurement outcome $x \in T_N$ with probability $p_{N,\rho}(x)$, its classical part then calculates $\mathrm{v}(x)$, and its probability distribution $p_{N,\rho}$ for all $\rho \in T$ satisfies

$$Pr_{x\sim p_{N,\rho}(\cdot)}\left[||\mathrm{v}(x)-\mathrm{v}(\rho)||_2 \leq \epsilon_N\right] \geq 1-\delta_N, \tag{14}$$

with $\lim_{N\to\infty}\epsilon_N = 0$ and $\lim_{N\to\infty}\delta_N = 0$.

Vector tomography of Definition 2 gives a trace-preserving protocol. The N-th algorithm, including calculating $\mathrm{v}(x)$ and writing it onto a classical register, can be made fully quantum, with a deferred measurement. Just before the deferred measurement, the reduced state on that classical register is a probabilistic mixture of basis states, each basis state indexed by $\mathrm{v}(x) \in \mathrm{v}(T_N) \subset \hat{\mathcal{H}}$:

$$\Psi(\rho^{\otimes N}) := \sum_{x\in T_N} p_{N,\rho}(x)\left|e_{\mathrm{v}(x)}\right\rangle\!\left\langle e_{\mathrm{v}(x)}\right|, \tag{15}$$

where Ψ is the N-th algorithm's CPTP. We measure the output's closeness to the target $\mathrm{v}(\rho)$ by using the linear map $h : e_i e_i^T \mapsto i$ from diagonal matrices to $\mathrm{v}(T_N)$. The triangle inequality and condition (14) imply that for all $\rho \in T$

$$\begin{aligned}
\left|\left|h\left(\Psi(\rho^{\otimes N})\right)-\mathrm{v}(\rho)\right|\right|_2 &= \left|\left|\sum_{x\in T_N} p_{N,\rho}(x)\,\mathrm{v}(x)-\mathrm{v}(\rho)\right|\right|_2 \\
&\leq \sum_{x\in T_N} p_{N,\rho}(x)\,||\mathrm{v}(x)-\mathrm{v}(\rho)||_2 \\
&\leq \sum_{x\text{ good}} p_{N,\rho}(x)\epsilon_N + \sum_{x\text{ bad}} p_{N,\rho}(x)\times 2 \\
&\leq \epsilon_N + 2\delta_N.
\end{aligned} \tag{16}$$

In this sense, Definition 2 implies a trace-preserving protocol. Since such a protocol would contradict the superposition impossibility, it cannot exist. Next, we show this directly.

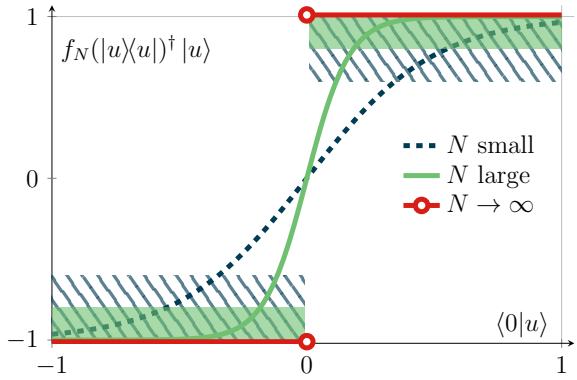


Figure 4: The contradiction illustrated on $\mathrm{v} = \lim_{N\to\infty} f_N$ of eq. (12). This $\mathrm{v} = \mathrm{v}_0$ for a given matrix $\rho \in S \subset \mathcal{D}_{\mathsf{pure}}(\mathcal{H})$ outputs that corresponding vector whose first entry is real positive, $\langle 0|\,\mathrm{v}_0(\rho)\rangle > 0$. The contradiction: functions f_N should be both continuous and $(\epsilon_N + 2\delta_N)$-close to v (i.e., confined to the marked regions).

Theorem 2. *The trace-preserving protocol Ψ is impossible. Consequently, vector tomography of Definition 2 is impossible.*

Proof. The output state $\Psi(\rho^{\otimes N})$ is a continuous function of $\rho \in \mathcal{D}_{\mathrm{pure}}(\mathcal{H})$. Since h is linear, the function

$$f_N(\rho) := h \circ \Psi(\rho^{\otimes N}) = \mathbb{E}_{p_{N,\rho}}\left[\mathrm{v}(x)\right] \tag{17}$$

is also continuous in $\rho \in \mathcal{D}_{\mathrm{pure}}(\mathcal{H})$. Moreover, by (16) it is close to $\mathrm{v}(\rho)$ for all $\rho \in T$. The continuity of $f_N(\rho)$ and its closeness to $\mathrm{v}(\rho)$ lead to a contradiction illustrated in Figure 4 and proven next. The function $g_N(|u\rangle) := f_N(|u\rangle\langle u|)^\dagger |u\rangle$ is continuous and 1-homogeneous in $|u\rangle$. From (16) we get that whenever $|u\rangle\langle u| \in T$

$$\begin{aligned} |g_N(|u\rangle)| &\geq |\mathrm{v}(|u\rangle\langle u|)\,|u\rangle| - |f_N(|u\rangle\langle u|)\,|u\rangle - \mathrm{v}(|u\rangle\langle u|)\,|u\rangle| \\ &\geq 1 - (\epsilon_N + 2\delta_N). \end{aligned}$$

By the continuity of g_N the above inequality actually holds for all $|u\rangle\langle u| \in \mathcal{D}_{\mathrm{pure}}(\mathcal{H})$. Choosing N large enough so that $\epsilon_N + 2\delta_N < 1$, we define $\hat{g}_N(|u\rangle) := g_N(|u\rangle)/|g_N(|u\rangle)|$ and restrict it to input vectors with all elements other than the first two fixed to zero. We get a continuous 1-homogeneous function $\hat{\mathbb{C}}^2 \to \mathrm{S}^1$, which contradicts Lemma 1. □

Intuitively, the problem is the impossibility of a continuous v. To see the impossibility, consider the above proof for a v continuous (i.e., for $\epsilon_N = \delta_N = 0$). To see that this is a problem, consider v_0 of eq. (12). For the N-th state tomography algorithm of eq. (13) take any $\epsilon \leq \epsilon_N^{\mathrm{tr}}/4$ and

$$\rho = \begin{pmatrix} \epsilon^2 & -\epsilon\sqrt{1-\epsilon^2} \\ -\epsilon\sqrt{1-\epsilon^2} & 1-\epsilon^2 \end{pmatrix}, \qquad x = \begin{pmatrix} \epsilon^2 & \epsilon\sqrt{1-\epsilon^2} \\ \epsilon\sqrt{1-\epsilon^2} & 1-\epsilon^2 \end{pmatrix},$$

and suppose that on input ρ the algorithm outputs x with high probability. This supposition is consistent with state tomography, because $||x-\rho||_{\mathrm{tr}} \leq \epsilon_N^{\mathrm{tr}}$. However, we have $||\,\mathrm{v}_0(x) - \mathrm{v}_0(\rho)||_2 = 2\sqrt{1-\epsilon^2} \geq 2 - \epsilon_N^{\mathrm{tr}}/2$, because $\rho = |u\rangle\langle u|$ is too close to the discontinuity of v_0 (corresponding to $|\,\langle 0|u\rangle\,| = \epsilon$ in Figure 4).

To fix the problem, we would like to avoid the discontinuity in a way that still accepts completely unrestricted and unknown $\rho \in \mathcal{D}_{\mathrm{pure}}(\mathcal{H})$. The idea is to approximately locate each input ρ and *then* to choose a v function that is continuous there. Instead of a single v_0, we consider multiple v_i, so that for every $\rho \in \mathcal{D}_{\mathrm{pure}}(\mathcal{H})$ there is $i \in \{0, 1, \ldots, d-1\}$ such that ρ is far from the discontinuity of v_i. We suggest

$$\mathrm{v}_i(\rho) = \frac{\rho\,|i\rangle}{\sqrt{\langle i|\,\rho\,|i\rangle}}. \tag{18}$$

Having sampled the estimate x we apply $\mathrm{v}_{r(x)}$, where $r(x) = \min\{i : \langle i|\,x\,|i\rangle \geq 1/d\}$ determines

which v function to use[8]. This way, the chosen vector function is always continuous around the sampled x, circumventing the impossibility and motivating the following modification.

Definition 3 (Vector tomography, revised)**.** A sequence of T_N-algorithms is a *vector tomography* if $T := \bigcup_{N=0}^{\infty} T_N$ is dense in $\mathcal{D}_{\text{pure}}(\mathcal{H})$, if there exists an indexed set $\{\mathrm{v}_i\}_{i\in I}$ of maps $\mathrm{v}_i : T \to \hat{\mathcal{H}}$ such that $\mathrm{v}_i(\rho)\,\mathrm{v}_i(\rho)^\dagger = \rho$, and if the N-th algorithm works as follows: Its quantum part on input $\rho^{\otimes N}$ where $\rho \in \mathcal{D}_{\text{pure}}(\mathcal{H})$ outputs a measurement outcome $x \in T_N$ with probability $p_{N,\rho}(x)$, its classical part then calculates $r(x) \in I$ and $\mathrm{v}_{r(x)}(x)$, and its probability distribution $p_{N,\rho}$ for all $\rho \in T$ satisfies

$$Pr_{x\sim p_{N,\rho}(\cdot)}\left[\left|\left|\mathrm{v}_{r(x)}(x) - \mathrm{v}_{r(x)}(\rho)\right|\right|_2 \leq \epsilon_N\right] \geq 1 - \delta_N, \tag{19}$$

with $\lim_{N\to\infty} \epsilon_N = 0$ and $\lim_{N\to\infty} \delta_N = 0$.

While Definition 2 yields a trace-preserving protocol, Definition 3 yields a random protocol, as we show next. This time the fully quantum version of the N-th algorithm needs an additional register to store $r(x)$. Before the deferred measurement, we have

$$\begin{aligned}\sum_i \Psi_i(\rho^{\otimes N}) \otimes |i\rangle\langle i| &= \sum_{x\in T_N} p_{N,\rho}(x) \left|e_{\mathrm{v}_{r(x)}(x)}\right\rangle\!\left\langle e_{\mathrm{v}_{r(x)}(x)}\right| \otimes |r(x)\rangle\langle r(x)| \\ \Psi_i(\rho^{\otimes N}) &= \sum_{x\in T_N} p_{N,\rho}(x,i) \left|e_{\mathrm{v}_i(x)}\right\rangle\!\left\langle e_{\mathrm{v}_i(x)}\right|, \end{aligned} \tag{20}$$

where the joint probability distribution of the random variables x and i is $p_{N,\rho}(x,i) = \delta_{i,r(x)} p_{N,\rho}(x)$, because the conditional probability of i given x is deterministic and equals the Kronecker delta function $\delta_{i,r(x)}$. The second line follows from the first by measuring the last register in the computational basis (the deferred measurement that makes the protocol random). This time, the error is quantified by

$$\begin{aligned}\sum_i \left|\left|h\left(\Psi_i(\rho^{\otimes N})\right) - p_{N,\rho}(i)\,\mathrm{v}_i(\rho)\right|\right|_2 &= \sum_i \left|\left|\sum_{x\in T_N} p_{N,\rho}(x,i)\,\mathrm{v}_i(x) - p_{N,\rho}(i)\,\mathrm{v}_i(\rho)\right|\right|_2 \\ &\leq \sum_i \sum_{x\in T_N} p_{N,\rho}(x,i)\,||\mathrm{v}_i(x) - \mathrm{v}_i(\rho)||_2 \\ &= \sum_{x\in T_N} p_{N,\rho}(x)\left|\left|\mathrm{v}_{r(x)}(x) - \mathrm{v}_{r(x)}(\rho)\right|\right|_2 \\ &\leq \epsilon_N + 2\delta_N, \end{aligned} \tag{21}$$

where the second line expands the marginal $p_{N,\rho}(i) = \sum_{x\in T_N} p_{N,\rho}(x,i)$ and applies the triangle inequality, the third is due to the Kronecker delta function inside $p_{N,\rho}(x,i)$, and the fourth follows from condition (19) by splitting x to "good" and "bad". Satisfying the inequality (21), the quantum instrument $\{\Psi_i\}_i$ constitutes a random protocol as claimed. Thus, for the existence of a random protocol for vector tomography (Result 3), it suffices to find an instance of Definition 3.

Theorem 3. *Any state tomography (satisfying eq. (13)) gives a vector tomography of Definition 3 for the appropriate choice of* v *functions.*

Section B contains the proof and the application to the superposition task. The proof hinges on the fact that a small trace-norm distance of pure states $|||w\rangle\langle w| - |u\rangle\langle u|||_{\text{tr}}$ implies a small Euclidean-norm distance of vectors $|||w\rangle - \langle u|w\rangle\,|u\rangle||_2$. This is enough, because, as opposed to Definition 2, Definition 3 does not require removing the w-dependence from $\langle u|w\rangle\,|u\rangle$ (i.e., the x-dependence from $\mathrm{v}_{r(x)}(\rho)$). In Section B we use this vector tomography for the superposition task, obtaining a random protocol:

[8]Other $r(x)$ functions are possible, as long as they are deterministic functions into $\{0, 1, \ldots, d-1\}$. For example, $\arg\max_i \langle i|\,x\,|i\rangle$ requires an additional tie-breaking rule, e.g. $r(x) = \min\{\arg\max_i \langle i|\,x\,|i\rangle\}$.

Corollary 1 (Random superposition protocol)**.** *Let* $\alpha, \beta \in \mathbb{R}_+$*. Let* $|s_{ij(\rho,\sigma)}\rangle := \alpha\, \mathrm{v}_i(\rho) + \beta\, \mathrm{v}_j(\sigma)$ *where* $\mathrm{v}_i, \mathrm{v}_j$ *are given by eq. (18). There exists a random protocol for* $|s_{ij(\rho,\sigma)}\rangle$*, i.e., for any* $\epsilon > 0$ *there exists* N *and a quantum instrument* $\{\Phi_{ij}\}_{ij}$ *such that*

$$\sum_{ij} \left|\left| \Phi_{ij}(\rho^{\otimes N} \otimes \sigma^{\otimes N}) - \mathrm{tr}\left[\Phi_{ij}(\rho^{\otimes N} \otimes \sigma^{\otimes N})\right] \frac{|s_{ij(\rho,\sigma)}\rangle\langle s_{ij(\rho,\sigma)}|}{\langle s_{ij(\rho,\sigma)}|s_{ij(\rho,\sigma)}\rangle} \right|\right|_{\mathrm{tr}} \leq \epsilon \tag{22}$$

holds for all pairs $\rho, \sigma \in \mathcal{D}_{pure}(\mathcal{H})$ *such that* $|s_{ij(\rho,\sigma)}\rangle \neq 0$*.*

4 Discussion

The first consequence of the superposition impossibility follows already from the work of Oszmaniec et al. [15] and becomes more pronounced by our generalization: Quantum computation is set apart from classical computation. We can first compare it to *deterministic* classical computation. Such a comparison seems natural, because the quantum state tree [12] strongly resembles the (deterministic) arithmetic formula tree [25]. The resulting contrast is in the availability of their respective building blocks: the quantum + is impossible while the arithmetic + is easy. This is analogous to cloning: quantum cloning is impossible, while deterministic classical cloning is easy.

A more demanding comparison is to *probabilistic* classical computation. Cloning probability vectors – equivalent to cloning diagonal density matrices – is impossible [26][9]. Thus, the task of cloning does not separate quantum and *probabilistic* classical computation. The superposing task, the quantum +, does! Weighted addition of unknown quantum states is impossible by ref. [15] and our generalization, while weighted addition of unknown probability vectors is done by the classical probabilistic protocol in Figure 1a. From this perspective, the superposition impossibility is even more distinctly quantum than the no-cloning.

Furthermore, our work has new consequences: the separations between different *quantum* computational models. The random model is strictly stronger than the trace-preserving or postselection model, when asked to implement certain tasks (functions). In terms of sample complexity, these separations are infinite! The separating task has restricted inputs (a 'partial function') or unrestricted inputs (a 'total function'). Consider the superposition task with inputs restricted to orthogonal pairs. Replace all occurrences of $\hat{\mathcal{H}} \times \hat{\mathcal{H}}$ in Theorem 1 and its proof by the set of orthonormal pairs. The result is a postselection protocol no-go for orthogonal inputs. Comparing this to the one-sample random protocol of [14] reveals the advantage of the random model on restricted inputs. With unrestricted inputs (primarily considered in this work), we have shown superposition and vector tomography no-go results for any sample complexity in the postselection (or trace-preserving) model. However, both tasks have at most exponential sample complexity in the random model. There, their precise sample complexities remain open.

As the second use of the topological method of ref. [20], our superposition no-go elucidates the method's applicability. When no algorithm can implement a certain function (task), the topological method might be able to prove so, if both of its two steps are applicable: 1) the existence of an algorithm imposes continuity on the function, 2) topology shows that the required function cannot be continuous. We can reverse the direction, first asking which functions cannot be continuous to then find the tasks for which the topological method is relevant. The topologically forbidden functions in this work are of the same type as those in [20]: sections and continuous homogeneous functions to S^1. While the latter facilitate quantitative no-gos, the former, sections, provide intuition for the past tasks and a way to recognize future use cases. Specifically, the superposition and the controlled unitary [20] both seem to involve choosing an unphysical representative (vector or operator) for the physical input (density matrix or superoperator). No section means no representative can be chosen continuously. The method may work for other tasks that involve such choices. More generally, the tasks must be quantum. Consistent with the fact that boolean functions are always computable by a complex enough algorithm, the topological method is restricted to tasks (functions) with topologically nontrivial output spaces, e.g. density states or superoperators. In

[9] In addition to their result on broadcasting, ref. [26] shows that a set of density states is clonable only if those state's pairwise fidelities are 0 or 1, i.e. only if the states in the set are pairwise orthogonal or identical. This is not true for the set of diagonal density matrices.

this sense, and as opposed to the polynomial method, the topological method is purely quantum. Quantum outputs are inherent in state complexity [12], quantum state synthesis [27, 28, 29] and unitary synthesis [30, 31, 32].

Acknowledgments

The author would like to thank Alon Dotan for helpful discussions of algebraic topology, Dorit Aharonov, Michał Oszmaniec, Itai Leigh and Michalis Skotiniotis for useful feedback. This work was supported by the Simons Foundation (Grant No. 385590), by the Israel Science Foundation (Grants No. 2137/19 and No. 1721/17), by the European Commission QuantERA grant ExTRaQT (Spanish MICINN project No. PCI2022-132965), by the Ministry for Digital Transformation and of Civil Service of the Spanish Government through the QUANTUM ENIA project call, Quantum Spain project, by the European Union through the Recovery, Transformation and Resilience Plan - NextGenerationEU within the framework of the Digital Spain 2026 Agenda, by Spanish Agencia Estatal de Investigación (Project No. PID2022-141283NB-I00).

References

[1] W. K. Wootters and W. H. Zurek. "A single quantum cannot be cloned". Nature **299**, 802–803 (1982).

[2] Charles H. Bennett and Gilles Brassard. "Quantum cryptography: Public key distribution and coin tossing". Theoretical Computer Science **560**, 7–11 (2014).

[3] Valerio Scarani, Sofyan Iblisdir, Nicolas Gisin, and Antonio Acín. "Quantum cloning". Reviews of Modern Physics **77**, 1225–1256 (2005).

[4] Giulio Chiribella, Giacomo Mauro D'Ariano, and Paolo Perinotti. "Probabilistic theories with purification". Physical Review A**81** (2010).

[5] Bob Coecke and Ross Duncan. "Interacting quantum observables: categorical algebra and diagrammatics". New Journal of Physics **13**, 043016 (2011).

[6] Adrian Kent. "Quantum tasks in Minkowski space". Classical and Quantum Gravity **29**, 224013 (2012).

[7] Wojciech Hubert Zurek. "Quantum Darwinism". Nature Physics **5**, 181–188 (2009).

[8] Paul M.B. Vitányi. "Quantum Kolmogorov complexity based on classical descriptions". IEEE Transactions on Information Theory **47**, 2464–2479 (2001).

[9] M. A. Nielsen and Isaac L. Chuang. "Programmable quantum gate arrays". Physical Review Letters **79**, 321–324 (1997).

[10] V. Bužek, M. Hillery, and R. F. Werner. "Optimal manipulations with qubits: Universal-NOT gate". Physical Review A **60**, R2626–R2629 (1999).

[11] Arun Kumar Pati and Samuel L. Braunstein. "Impossibility of deleting an unknown quantum state". Nature **404**, 164–165 (2000).

[12] Scott Aaronson. "Multilinear formulas and skepticism of quantum computing". In Proceedings of the thirty-sixth annual ACM symposium on Theory of computing. STOC04. ACM (2004).

[13] Yu Cai, Huy Nguyen Le, and Valerio Scarani. "State complexity and quantum computation". Annalen der Physik **527**, 684–700 (2015).

[14] Mina Doosti, Farzad Kianvash, and Vahid Karimipour. "Universal superposition of orthogonal states". Physical Review A**96** (2017).

[15] Michał Oszmaniec, Andrzej Grudka, Michał Horodecki, and Antoni Wójcik. "Creating a superposition of unknown quantum states". Physical Review Letters**116** (2016).

[16] U. Alvarez-Rodriguez, M. Sanz, L. Lamata, and E. Solano. "The forbidden quantum adder". Scientific Reports**5** (2015).

[17] Rui Li, Unai Alvarez-Rodriguez, Lucas Lamata, and Enrique Solano. "Approximate quantum adders with genetic algorithms: An IBM quantum experience". Quantum Measurements and Quantum Metrology **4**, 1–7 (2017).

[18] Xiao-Min Hu, Meng-Jun Hu, Jiang-Shan Chen, Bi-Heng Liu, Yun-Feng Huang, Chuan-Feng Li, Guang-Can Guo, and Yong-Sheng Zhang. "Experimental creation of superposition of unknown photonic quantum states". Physical Review A**94** (2016).

[19] Keren Li, Guofei Long, Hemant Katiyar, Tao Xin, Guanru Feng, Dawei Lu, and Raymond Laflamme. "Experimentally superposing two pure states with partial prior knowledge". Physical Review A**95** (2017).

[20] Zuzana Gavorová, Matan Seidel, and Yonathan Touati. "Topological obstructions to quantum computation with unitary oracles". Physical Review A**109** (2024).

[21] Takanori Sugiyama, Peter S. Turner, and Mio Murao. "Precision-guaranteed quantum tomography". Physical Review Letters**111** (2013).

[22] Matthias Christandl and Renato Renner. "Reliable quantum state tomography". Physical Review Letters**109** (2012).

[23] Scott Aaronson. "Quantum computing, postselection, and probabilistic polynomial-time". Proceedings of The Royal Society of London. Series A. Mathematical, Physical and Engineering Sciences **461**, 3473–3482 (2005).

[24] Robert Beals, Harry Buhrman, Richard Cleve, Michele Mosca, and Ronald de Wolf. "Quantum lower bounds by polynomials". Journal of the ACM **48**, 778–797 (2001).

[25] Ran Raz. "Tensor-rank and lower bounds for arithmetic formulas". Journal of the ACM **60**, 1–15 (2013).

[26] Howard Barnum, Carlton M. Caves, Christopher A. Fuchs, Richard Jozsa, and Benjamin Schumacher. "Noncommuting mixed states cannot be broadcast". Physical Review Letters **76**, 2818–2821 (1996).

[27] Sandy Irani, Anand Natarajan, Chinmay Nirkhe, Sujit Rao, and Henry Yuen. "Quantum search-to-decision reductions and the state synthesis problem". In 37th Computational Complexity Conference. Volume 234 of LIPIcs. Leibniz Int. Proc. Inform., pages Art. No. 5, 19. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern (2022).

[28] Zhengfeng Ji, Yi-Kai Liu, and Fang Song. "Pseudorandom quantum states". Pages 126–152. Springer International Publishing. (2018).

[29] William Kretschmer. "Quantum pseudorandomness and classical complexity". In 16th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2021). Pages 2:1–2:20. (2021).

[30] Jonas Haferkamp, Philippe Faist, Naga B. T. Kothakonda, Jens Eisert, and Nicole Yunger Halpern. "Linear growth of quantum circuit complexity". Nature Physics **18**, 528–532 (2022).

[31] Leonard Susskind. "Black holes and complexity classes". Preprint (2018) arXiv:1802.02175.

[32] Adam Bouland, Bill Fefferman, and Umesh Vazirani. "Computational pseudorandomness, the wormhole growth paradox, and constraints on the AdS/CFT duality". In 11th Innovations in Theoretical Computer Science Conference (ITCS 2020). Pages 63:1–63:2. (2020).

[33] Allen Hatcher. "Algebraic topology". Cambridge University Press, Cambridge. (2002).

[34] Mark M Wilde. "From classical to quantum Shannon theory". Preprint (2011) arXiv:1106.1445.

A Superposition impossibility via no-section

This section supports the intuition we gave in Section 1.3, that our impossibility relates to the nonexistence of continuous functions mapping a pure density state to some corresponding vector $|u\rangle\langle u| \mapsto |\widetilde{u}\rangle$. Noting that S^{2d-1} is the space of unit vectors in $\mathbb{R}^{2d}$, or, alternatively, unit vectors in $\mathbb{C}^d$, we restate this formally:

Lemma 2. *Let $d \in \mathbb{N}$, $d \geq 2$. Let $\pi : \mathrm{S}^{2d-1} \to \mathcal{D}_{pure}(\mathbb{C}^d)$ be the outer product function $\pi(|u\rangle) = |u\rangle\langle u|$. There is no continuous function $s : \mathcal{D}_{pure}(\mathbb{C}^d) \to \mathrm{S}^{2d-1}$ such that $\pi \circ s = id$.*

The term *section* refers to a continuous function that composes with a projection into the identity function *id*, as above. We use Lemma 2 to prove this special-case version of Theorem 1:

Theorem 4. *Let $\alpha, \beta \in \mathbb{R}_+$ and suppose $\dim \mathcal{H} \geq 3$. Let any function $c : \hat{\mathcal{H}} \times \hat{\mathcal{H}} \to \mathbb{C}$, such that $S := c^{-1}(\mathbb{C} \setminus \{0\})$ is dense in $\hat{\mathcal{H}} \times \hat{\mathcal{H}}$, determine the superposition $|s_{c(u,v)}\rangle = \alpha\,|c(u,v)|\,|u\rangle + \beta\,c(u,v)\,|v\rangle$. There is no continuous function $\mathcal{A} : \mathcal{D}_{pure}(\mathcal{H}) \times \mathcal{D}_{pure}(\mathcal{H}) \to \mathcal{D}_+(\mathcal{H})$ such that for all $(|u\rangle, |v\rangle) \in S$*

$$\frac{\mathcal{A}(|u\rangle\langle u|, |v\rangle\langle v|)}{\mathrm{tr}\left[\mathcal{A}(|u\rangle\langle u|, |v\rangle\langle v|)\right]} = \frac{|s_{c(u,v)}\rangle\langle s_{c(u,v)}|}{\langle s_{c(u,v)}|s_{c(u,v)}\rangle}. \tag{23}$$

Proof. Assume towards contradiction that such $\mathcal{A}$ exists. Let $|\gamma\rangle \in \hat{\mathcal{H}}$ be some fixed vector. Denote by $\mathcal{K}_\gamma$ the subspace of $\mathcal{H}$ orthogonal to $|\gamma\rangle \in \mathcal{H}$. The dimension of $\mathcal{K}_\gamma$ is $d = \dim \mathcal{H} - 1 \geq 2$. Next we define the function $h_\gamma : \mathcal{D}_{\mathrm{pure}}(\mathcal{K}_\gamma) \to \mathcal{K}_\gamma$:

$$h_\gamma(|u\rangle\langle u|) := \frac{\alpha^2 + \beta^2}{\alpha\beta}\,|u\rangle\langle u|\,\frac{\mathcal{A}(|u\rangle\langle u|, |\gamma\rangle\langle\gamma|)}{\mathrm{tr}\left[\mathcal{A}(|u\rangle\langle u|, |\gamma\rangle\langle\gamma|)\right]}\,|\gamma\rangle\,.$$

Since $\mathcal{A}$ is continuous, so is h_γ. Let $(|u_n\rangle, |\gamma_n\rangle)$ be a sequence in S that converges to $(|u\rangle, |\gamma\rangle)$. Then, in particular $|u_n\rangle \xrightarrow{n\to\infty} |u\rangle$ and

$$h_\gamma(|u\rangle\langle u|) = \lim_{n\to\infty} h_{\gamma_n}(|u_n\rangle\langle u_n|) = \lim_{n\to\infty} \frac{c(u_n, \gamma_n)^*}{|c(u_n, \gamma_n)|}\,|u_n\rangle\,, \tag{24}$$

where the second equality follows by substituting eq. (23) and removing the additive terms that converge to zero because $\langle\gamma_n|u_n\rangle \xrightarrow{n\to\infty} 0$. Similarly we get that $\langle u|\,h_\gamma(|u\rangle\langle u|) = \lim_{n\to\infty} \frac{c(u_n,\gamma_n)^*}{|c(u_n,\gamma_n)|}$. In other words, the fractions converge to $\lambda_u := \langle u|\,h_\gamma(|u\rangle\langle u|)$, and since the n-th fraction has norm 1, we have $|\lambda_u| = 1$. Equation (24) is the limit of the product of two converging sequences, therefore it equals the product of the limits, $h_\gamma(|u\rangle\langle u|) = \lambda_u\,|u\rangle$. Since $\mathcal{K}_\gamma \approx \mathbb{C}^d$ we got a function $h_\gamma : \mathcal{D}_{\mathrm{pure}}(\mathbb{C}^d) \to \mathrm{S}^{2d-1}$ that contradicts Lemma 2. □

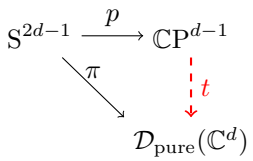
Next we prove Lemma 2 by first showing that $\mathcal{D}_{\mathrm{pure}}(\mathbb{C}^d)$ is homeomorphic to the complex projective space $\mathbb{C}\mathrm{P}^{d-1}$ (Claim 1), then showing that there is no section $\mathbb{C}\mathrm{P}^{d-1} \to \mathrm{S}^{2d-1}$ (Claim 2).

Claim 1. *$\mathcal{D}_{pure}(\mathbb{C}^d)$ is homeomorphic to the complex projective space $\mathbb{C}\mathrm{P}^{d-1}$.*

If we regard the vectors in S^{2d-1} as the d-dimensional complex vectors of norm 1, then complex projective space is the quotient space $\mathbb{C}\mathrm{P}^{d-1} = \mathrm{S}^{2d-1} / \sim$ under the identification $|u\rangle \sim e^{i\alpha}\,|u\rangle$ for any real α.

Claim 2. *Let $d \in \mathbb{N}$, $d \geq 2$. Let $p : \mathrm{S}^{2d-1} \to \mathbb{C}\mathrm{P}^{d-1}$ be the quotient map. There is no continuous function $s : \mathbb{C}\mathrm{P}^{d-1} \to \mathrm{S}^{2d-1}$ such that $p \circ s = id$.*

Proof of Claim 1. Let $\pi : \mathrm{S}^{2d-1} \to \mathcal{D}_{\mathrm{pure}}(\mathbb{C}^d)$ be the outer product function $\pi(|u\rangle) = |u\rangle\langle u|$. Since $\mathcal{D}_{\mathrm{pure}}(\mathbb{C}^d)$ is in fact the image of π, π is surjective. Let p be the quotient map $p : \mathrm{S}^{2d-1} \to \mathbb{C}\mathrm{P}^{d-1}$ as above, $p(|u\rangle) = [|u\rangle]$, where $[|u\rangle]$ denotes the equivalence class under the equivalence relation $|u\rangle \sim e^{i\alpha}\,|u\rangle$ for any real α. Observe that $\pi\left(|u\rangle\right) = \pi\left(|v\rangle\right)$ whenever $p\left(|u\rangle\right) = p\left(|v\rangle\right)$. Then by the universal property of the quotient space in topology there exists $t : \mathbb{C}\mathrm{P}^{d-1} \to \mathcal{D}_{\mathrm{pure}}(\mathbb{C}^d)$ continuous, such that $t \circ p = \pi$:



In particular, t is surjective. Next, we show that $t : [|u\rangle] \mapsto |u\rangle\langle u|$ is injective: Suppose $|u\rangle\langle u| = |v\rangle\langle v| =: M$. Since $|u\rangle$, $|v\rangle$ are unit vectors, there exists a column in matrix M that is nonzero, label this column's index by j. This column is equal to $M_j = u_j^* |u\rangle = v_j^* |v\rangle$. Since it is nonzero we can write $|u\rangle = \frac{v_j^*}{u_j^*} |v\rangle$, and since $|u\rangle$, $|v\rangle$ are unit vectors, we have $\frac{v_j^*}{u_j^*} = e^{i\alpha}$ for some real α and therefore $[|u\rangle] = [|v\rangle]$. We have shown that $t : \mathbb{C}\mathrm{P}^{d-1} \to \mathcal{D}_{\mathrm{pure}}(\mathbb{C}^d)$ is continuous and bijective. The domain of t is compact, because a quotient space of a compact space (here S^{2d-1}) is compact. The range of t is Hausdorff. A bijective continuous map from a compact space to a Hausdorff space is closed and hence a homeomorphism. Thus, t is a homeomorphism. □

Proof of Claim 2. This proof uses some standard facts in homology theory. They can be found, for example, in Chapter 2 in [33]. The reduced homology groups of S^m are

$$\tilde{H}_i\left(\mathrm{S}^m\right) = \begin{cases} \mathbb{Z}, & i = m \\ 0, & i \neq m \end{cases}.$$

The singular homology groups of $\mathbb{C}\mathrm{P}^k$ are

$$H_i\left(\mathbb{C}\mathrm{P}^k\right) = \begin{cases} \mathbb{Z}, & i = 0, 2, 4 \ldots, 2k \\ 0, & \text{otherwise} \end{cases},$$

and by definition $\tilde{H}_i(X) = H_i(X)$ for all $i > 0$ for any space X. Assume towards contradiction that $s : \mathbb{C}\mathrm{P}^{d-1} \to \mathrm{S}^{2d-1}$ continuous exists for some $d \geq 2$ such that $p \circ s = id$. In diagram:

$$\mathbb{C}\mathrm{P}^{d-1} \xrightarrow{s} \mathrm{S}^{2d-1} \xrightarrow{p} \mathbb{C}\mathrm{P}^{d-1}, \quad \mathbb{C}\mathrm{P}^{d-1} \xrightarrow{id} \mathbb{C}\mathrm{P}^{d-1}$$

A continuous map between two spaces $f : X \to Y$ induces for any i a homomorphism $f_* : H_i(X) \to H_i(Y)$ between their i-th singular homology groups. For $i = 2d - 2$ we get

$$H_{2d-2}\left(\mathbb{C}\mathrm{P}^{d-1}\right) \xrightarrow{s_*} H_{2d-2}\left(\mathrm{S}^{2d-1}\right) \xrightarrow{p_*} H_{2d-2}\left(\mathbb{C}\mathrm{P}^{d-1}\right), \quad H_{2d-2}\left(\mathbb{C}\mathrm{P}^{d-1}\right) \xrightarrow{id_*} H_{2d-2}\left(\mathbb{C}\mathrm{P}^{d-1}\right)$$

where id_* is the identity map. However, by the above, $H_{2d-2}(\mathbb{C}\mathrm{P}^{d-1}) = \mathbb{Z}$ and $H_{2d-2}(\mathrm{S}^{2d-1}) = 0$ for $d \geq 2$, which contradicts the last diagram. □

B Possibilites via randomness (proofs)

Here we prove Theorem 3 and Corollary 1.

Proof of Theorem 3. We write $\rho = |u\rangle\langle u|$ and $x = |w\rangle\langle w|$ for the unknown state and the measurement outcome respectively. Suppose

$$||x - \rho||_{\mathrm{tr}} = |||w\rangle\langle w| - |u\rangle\langle u|||_{\mathrm{tr}} \leq \epsilon_N^{\mathrm{tr}} < \frac{1}{d}. \tag{25}$$

We have

$$\begin{aligned} |||w\rangle - \langle u|w\rangle |u\rangle||_2 &= \sqrt{|||w\rangle||_2^2 - 2\,|\langle u|w\rangle\langle w|u\rangle| + |\langle u|w\rangle|^2\,|||u\rangle||_2^2} \\ &= \sqrt{1 - |\langle u|w\rangle|^2} \\ &\leq \frac{1}{2}\epsilon_N^{\mathrm{tr}}, \end{aligned} \tag{26}$$

where the last inequality follows from the identity

$$\left|1-|\langle u|w\rangle|^2\right| = \frac{1}{4}\left\||u\rangle\langle u| - |w\rangle\langle w|\right\|_{\mathrm{tr}}^2 \leq \frac{1}{4}\left(\epsilon_N^{\mathrm{tr}}\right)^2. \tag{27}$$

See, for example, [34, eq. (9.172)].

Next we use (26) to upper bound

$$\left|\left|\mathrm{v}_{r(x)}(x) - \mathrm{v}_{r(x)}(\rho)\right|\right|_2 = \left|\left||w\rangle - \frac{|\langle r|u\rangle|}{|\langle r|w\rangle|}\frac{\langle r|w\rangle}{\langle r|u\rangle}|u\rangle\right|\right|_2. \tag{28}$$

where we define v_i as in (18) and we wrote $r(x) = r$ for readability. The rule for choosing $r(x)$ ensures $|\langle r|w\rangle|^2 \geq 1/d$ which together with (25) ensures $\langle r|u\rangle \neq 0$. Comparing the left-hand side of (26) and the right-hand side of (28), we note that it is enough to compare the scalars next to $|u\rangle$. Using the closeness of the scalars that follows from (26),

$$|\langle r|w\rangle - \langle u|w\rangle\langle r|u\rangle| \leq \frac{1}{2}\epsilon_N^{\mathrm{tr}} \tag{29}$$

for any $r \in \{0, 1, \ldots d-1\}$, we upper bound

$$\begin{aligned}
\left|\langle u|w\rangle - \frac{|\langle r|u\rangle|}{|\langle r|w\rangle|}\frac{\langle r|w\rangle}{\langle r|u\rangle}\right| &\leq \left|\langle u|w\rangle - \tfrac{\langle u|w\rangle}{|\langle u|w\rangle|}\right| + \left|\tfrac{\langle u|w\rangle}{|\langle u|w\rangle|} - \langle u|w\rangle\frac{|\langle r|u\rangle|}{|\langle r|w\rangle|}\right| + \left|\langle u|w\rangle\frac{|\langle r|u\rangle|}{|\langle r|w\rangle|} - \frac{|\langle r|u\rangle|}{|\langle r|w\rangle|}\frac{\langle r|w\rangle}{\langle r|u\rangle}\right| \\
&= \left||\langle u|w\rangle| - \tfrac{|\langle u|w\rangle|}{|\langle u|w\rangle|}\right| + |\langle u|w\rangle|\left|\tfrac{1}{|\langle u|w\rangle|} - \frac{|\langle r|u\rangle|}{|\langle r|w\rangle|}\right| + \frac{|\langle r|u\rangle|}{|\langle r|w\rangle|}\left|\langle u|w\rangle - \frac{\langle r|w\rangle}{\langle r|u\rangle}\right| \\
&\leq \left||\langle u|w\rangle| - 1\right| + |\langle u|w\rangle|\frac{\frac{1}{2}\epsilon_N^{\mathrm{tr}}}{|\langle u|w\rangle|\,|\langle r|w\rangle|} + \frac{|\langle r|u\rangle|}{|\langle r|w\rangle|}\frac{\frac{1}{2}\epsilon_N^{\mathrm{tr}}}{|\langle r|u\rangle|},
\end{aligned} \tag{30}$$

where, in addition to (29), for the middle term in the last line we used also $\left|\,|a| - |b|\,\right| \leq |a-b|$ $a, b \in \mathbb{C}$. Applying to the first term $|1 - |\langle u|w\rangle|| \leq |1 - |\langle u|w\rangle|^2|$ and (27), and simplifying the other terms, we get

$$\left|\langle u|w\rangle - \frac{|\langle r|u\rangle|}{|\langle r|w\rangle|}\frac{\langle r|w\rangle}{\langle r|u\rangle}\right| \leq \frac{1}{4}\left(\epsilon_N^{\mathrm{tr}}\right)^2 + \frac{\epsilon_N^{\mathrm{tr}}}{|\langle r|w\rangle|}. \tag{31}$$

We obtain

$$\left|\left|\mathrm{v}_{r(x)}(x) - \mathrm{v}_{r(x)}(\rho)\right|\right|_2 \leq \left(\frac{1}{|\langle r|w\rangle|} + \frac{1}{2}\right)\epsilon_N^{\mathrm{tr}} + \frac{1}{4}\left(\epsilon_N^{\mathrm{tr}}\right)^2 \tag{32}$$

by combining (26) and (31).

The following parameters then satisfy Definition 3. Set $\delta_N = \delta_N^{\mathrm{tr}}$. Let N_0 be such that $\epsilon_N^{\mathrm{tr}} \leq 1/d$ for all $N \geq N_0$, and set

$$\epsilon_N := \begin{cases} 2 & \text{if } N < N_0 \\ \left(\sqrt{d} + \frac{1}{2}\right)\epsilon_N^{\mathrm{tr}} + \frac{1}{4}\left(\epsilon_N^{\mathrm{tr}}\right)^2 & \text{if } N \geq N_0 \end{cases}. \tag{33}$$

Since the $r(x)$ function ensures $|\,\langle r|w\rangle\,|^2 \geq 1/d$ for $N \geq N_0$, this ϵ_N upper bounds (32). Moreover, $\lim_{N\to\infty}\epsilon_N = 0$ and $\lim_{N\to\infty}\delta_N = 0$, satisfying Definition 3.

□

Proof of Corollary 1. We suggest the following random protocol $\{\Phi_{ij}\}_{ij}$. First, it performs a vector tomography on each of the two input states. The classical estimate of the desired superposition is then $|s_{ij(x,y)}\rangle$ with probability $p_{N,\rho,\sigma}(x,i,y,j)$, where x and y are the estimates of ρ and σ. Assume for now that whenever the probability is nonzero, we have $||\,|s_{ij(x,y)}\rangle\,||_2^2 \geq K$ for some positive constant K. We will justify this later. Having the classical description of $|s_{ij(x,y)}\rangle$, the protocol runs a circuit that prepares the renormalized $|s_{ij(x,y)}\rangle$ within the error $\epsilon/4$. In other words, Φ_{ij} satisfies:

$$\left\|\Phi_{ij}(\rho^{\otimes N}\otimes\sigma^{\otimes N}) - \sum_{x,y\in T_N} p_{N,\rho,\sigma}(x,i,y,j)\frac{|s_{ij(x,y)}\rangle\langle s_{ij(x,y)}|}{\langle s_{ij(x,y)}|s_{ij(x,y)}\rangle}\right\|_{\mathrm{tr}} \leq \epsilon/4 \tag{34}$$

and, consequently

$$\left|\left|\mathrm{tr}\left[\Phi_{ij}(\rho^{\otimes N}\otimes\sigma^{\otimes N})\right]-\sum_{x,y\in T_N}p_{N,\rho,\sigma}(x,i,y,j)\right|\right|_{\mathrm{tr}}\leq\epsilon/4. \tag{35}$$

Thus, to prove Corollary 1 it remains to show that

$$D:=\sum_{i,j}\left|\left|\sum_{x,y\in T_N}p_{N,\rho,\sigma}(x,i,y,j)\left(\frac{|s_{ij(x,y)}\rangle\langle s_{ij(x,y)}|}{\langle s_{ij(x,y)}|s_{ij(x,y)}\rangle}-\frac{|s_{ij(\rho,\sigma)}\rangle\langle s_{ij(\rho,\sigma)}|}{\langle s_{ij(\rho,\sigma)}|s_{ij(\rho,\sigma)}\rangle}\right)\right|\right|_{\mathrm{tr}}\leq\epsilon/2. \tag{36}$$

Towards proving the above, we first show that

$$\begin{aligned}
D&\leq\sum_{i,j}\sum_{x,y}p_{N,\rho,\sigma}(x,i,y,j)\left|\left|\frac{|s_{ij(x,y)}\rangle\langle s_{ij(x,y)}|}{\langle s_{ij(x,y)}|s_{ij(x,y)}\rangle}-\frac{|s_{ij(\rho,\sigma)}\rangle\langle s_{ij(\rho,\sigma)}|}{\langle s_{ij(\rho,\sigma)}|s_{ij(\rho,\sigma)}\rangle}\right|\right|_{\mathrm{tr}}\\
&\leq\sum_{i,j,x,y}p_{N,\rho,\sigma}(x,i,y,j)\frac{2}{||\,|s_{ij(x,y)}\rangle\,||_2^2}\left|\left|\left|s_{ij(x,y)}\rangle\langle s_{ij(x,y)}|-|s_{ij(\rho,\sigma)}\rangle\langle s_{ij(\rho,\sigma)}|\right|\right|\right|_{\mathrm{tr}}\\
&\leq\frac{4\,(\alpha+\beta)}{K}\sum_{i,j,x,y}p_{N,\rho,\sigma}(x,i,y,j)\left|\left|\left|s_{ij(x,y)}\rangle-|s_{ij(\rho,\sigma)}\rangle\right|\right|\right|_2\\
&\leq\frac{4\,(\alpha+\beta)}{K}\left(\alpha\sum_{i,x}p_{N,\rho,\sigma}(x,i)\,||\mathrm{v}_i(x)-\mathrm{v}_i(\rho)||_2+\beta\sum_{j,y}p_{N,\rho,\sigma}(y,j)\,||\mathrm{v}_j(y)-\mathrm{v}_j(\sigma)||_2\right)\\
&\leq\frac{4\,(\alpha+\beta)^2}{K}\left(\epsilon_N+2\delta_N\right),
\end{aligned} \tag{37}$$

where the second inequality holds because of the identity

$$\left|\left|\frac{A}{||A||}-\frac{B}{||B||}\right|\right|\leq\frac{2}{||A||}\,||A-B||\,, \tag{38}$$

which holds for any norm and operators A, B, and can be proven by substracting and adding $B/||A||$, the third inequality follows from our assumption that either $p_{N,\rho}(x,i)\,p_{N,\sigma}(y,j)=0$ or $||\,|s_{ij(x,y)}\rangle\,||_2^2\geq K$ and from the identity

$$||\,|u\rangle\langle u|-|w\rangle\langle w|\,||_{\mathrm{tr}}\leq 2\max(||\,|u\rangle\,||,||\,|w\rangle\,||)||\,|u\rangle-|w\rangle\,||_2, \tag{39}$$

the fourth inequality follows from the triangle inequality, and the fifth inequality holds assuming the two vector tomographies satisfy (21).

Next we ensure that K exists. If $\alpha\neq\beta$ then we can set $K=|\alpha-\beta|^2$, because

$$||\,|s_{ij(x,y)}\rangle\,||_2=||\alpha\,\mathrm{v}_i(x)+\beta\,\mathrm{v}_j(y)||_2\geq|\alpha-\beta|, \tag{40}$$

by the triangle inequality. This K is indeed independent of N, x, y. If $\alpha=\beta$, then we need another way to make sure that a small-norm $|s_{ij(x,y)}\rangle$ has probability zero, $p_{N,\rho,\sigma}(x,i,y,j)=0$. In the vector tomography we have seen so far, for a given x all i-values except one have probability zero, because i, the index of the v-function, is chosen deterministically from x by the rule $i=r(x)=\min\{l:\langle l|\,x\,|l\rangle\geq 1/d\}$. For j, we will use a new rule $j=r'(x,y)$:

$$r'(x,y):=\begin{cases}r(x) & \text{if } ||x-y||_{\mathrm{tr}}<\frac{1}{2d}\\ \min\{l:\langle l|\,y\,|l\rangle\geq 1/d\} & \text{otherwise}\end{cases} \tag{41}$$

Definition 3 is satisfied also with this rule (replace d with $2d$ in the proof of Theorem 3).

Within the case $\alpha=\beta$, suppose that the two state tomographies output x, y such that $||x-y||_{\mathrm{tr}}<\frac{1}{2d}$. The rule r' ensures that $p_{N,\rho,\sigma}(x,i,y,j)$ is nonzero only for $|s_{ij(x,y)}\rangle=|s_{ii(x,y)}\rangle=$

$\alpha(\mathrm{v}_i(x) + \mathrm{v}_i(y))$ and

$$\begin{aligned} |||s_{ii(x,y)}\rangle||_2 = \alpha\,||\mathrm{v}_i(x) + \mathrm{v}_i(y)||_2 &\geq \alpha\,(2 - ||\mathrm{v}_i(x) - \mathrm{v}_i(y)||_2) \\ &\geq \alpha\left(2 - \frac{2}{\sqrt{\langle i|x|i\rangle}}\,||x\,|i\rangle - y\,|i\rangle||_2\right) \\ &\geq \alpha\left(2 - \frac{2}{\sqrt{\langle i|x|i\rangle}}\,||x - y||_{\mathrm{tr}}\right) \\ &\geq \alpha\left(2 - \sqrt{d}\frac{2}{\sqrt{2d}}\right) = \alpha(2 - \sqrt{2}), \end{aligned} \tag{42}$$

where we applied the triangle inequality in the first line, inequality (38) in the second line, the contractivity of the trace norm in the third line, and the r and r' rules in the fourth line. The resulting lower bound is our $\sqrt{K}$ for this case.

Now suppose $\alpha = \beta$ and x, y are such that $||x - y||_{\mathrm{tr}} \geq \frac{1}{2d}$. We have

$$|||s_{ij(x,y)}\rangle||_2 = \alpha\,||\mathrm{v}_i(x) + \mathrm{v}_j(y))||_2 \geq \frac{\alpha}{2}\,||x - y||_{\mathrm{tr}} \geq \frac{\alpha}{4d}, \tag{43}$$

which is an application of (39). This lower bound is our $\sqrt{K}$ for this final case. Taking the minimum K over all the cases, we choose N large enough so that (37) is upper bounded by $\epsilon/2$. This completes the proof. □