Notes on distinguishability of postselected computations

Zuzana Gavorová

School of Computer Science and Engineering, The Hebrew University of Jerusalem, Jerusalem, Israel

Abstract

The framework of postselection is becoming more and more important in various recent directions in Quantum Computation research. Postselection renders simple computational models able to perform general quantum computation. This was first observed for the linear optics model [E. Knill, R. Laflamme, G. J. Milburn, Nature 409, 46 (2001)], and has since provided us with many near-term candidates for the quantum advantage, commuting computations [M. J. Bremner, R. Jozsa, D. J. Shepherd, Proc. R. Soc. A 467, 459 (2011)] being the first. To facilitate the discussion of errors in the presence of postselection, we define and characterize trace-induced distance and diamond distance of postselected computations. We show counterexamples to simple properties that one would expect of any distance measure; the properties of convexity (when considering only the pure-state inputs would suffice), contractivity, and subadditivity of errors. On the positive side, we prove that certain weaker versions of contractivity and subadditivity and a number of other properties are preserved in the postselected setting. We achieve this via a "conversion lemma" that translates any inequality from the standard to the postselected setting.

1 Introduction

The postselected setting has recently drawn a lot of attention. Postselection renders linear optics universal for quantum computation [1] and larger gate errors tolerable in fault-tolerant quantum computation [2, 3, 4]. Arguments using postselection [5] provide the evidence for quantum supremacy, i.e. that we cannot simulate classically certain quantum computations: commuting quantum computations [6, 7], boson sampling [8], one clean qubit DQC1 computations [9], random circuits [10]. These computations are, moreover, quite simple and believed to be achievable in the near future - by noisy intermediate term quantum (NISQ) devices.

In this note we define postselection equivalents of trace-induced distance and diamond distance. We pick some known inequalities between the standard distance measures, also adding a new one, and prove that equivalent relations hold in the postselected setting.

The difficulty is that in the postselected setting it only makes sense to compare computations after their output has been renormalized. Let $L(\mathcal{H})$ be the set of linear operators from the finite-dimensional$^1$ Hilbert space $\mathcal{H}$ to itself. We say that a map $\Phi : L(\mathcal{H}) \to L(\mathcal{H}')$ is a postselection superoperator if (1) it is linear, completely positive (CP), trace-nonincreasing and (2) its postselection probability $\text{tr}[\Phi(\rho)]$ is nonzero for all $\rho \in \mathcal{D}(\mathcal{H})$, where $\mathcal{D}(\mathcal{H}) \subset L(\mathcal{H})$ is the set of

---

$^1$All Hilbert spaces throughout this note are finite-dimensional.
normalized density operators. A postselection superoperator \( \Phi \) followed by the renormalisation of the output corresponds to the map

\[
\rho \mapsto \frac{\Phi(\rho)}{\text{tr}[\Phi(\rho)]},
\]

which introduces a non-linearity in \( \rho \). The distances of such maps break certain relations: we will see counterexamples to convexity, contractivity and subadditivity. However, under certain conditions translation of some standard inequalities to the postselected setting is possible. We achieve this via our "conversion lemma", which states that under these conditions the postselection probability is almost constant in \( \rho \), so that the nonlinear map (1) can be replaced by a linear one.

Errors in the postselected setting were previously investigated [11, 12, 13, 14] for some specific applications requiring a version of subadditivity (also called union bound). We take the more general approach of deriving postselection equivalents of the following well-established results on diamond distance.

1.1 Preceding work: Two inequalities to carry over to the postselected setting

Kitaev [15] defined diamond distance \( d_{\diamond} \) and Aharonov, Kitaev and Nisan proved [16, Theorem 4] the following property of \( d_{\diamond} \), useful when evaluating how well a computation composed of subroutines simulates an ideal computation composed of ideal subroutines:

**Theorem 1.1 (due to [16]: Subadditivity of \( d_{\diamond} \)).** Let the simulating computation use \( N \) subroutines, each at most \( \epsilon \)-far in \( d_{\diamond} \) from its ideal. Then the simulating computation is at most \( N\epsilon \)-far in \( d_{\diamond} \) from the ideal computation.

The next result from Watrous [17] relates \( d_{\diamond} \) to what we call the *operational* trace-induced distance \( d^D_{tr} \) of two linear trace-nonincreasing CP maps \( \Phi, \Psi : L(H) \to L(H') \):

\[
d^D_{tr}(\Psi, \Phi) := \sup_{\rho \in D(H)} \|\Psi(\rho) - \Phi(\rho)\|_{tr},
\]

where \( \|\cdot\|_{tr} \) is the trace norm. This is not the same as the usual trace-induced distance \( d_{tr} \), but \( d_{\diamond} \) is the stabilized version of both \( d_{tr} \) and \( d^D_{tr} \) (see Section 2). Watrous [17, Theorem 3.56] proved that under a certain condition upper bounding the stabilized version is 'for free':

**Theorem 1.2 (due to [17]: \( d_{tr} \) small \( \Rightarrow \) \( d_{\diamond} \) small).** Let \( U : H \to H' \) be an isometry, \( \Psi : L(H) \to L(H') \) a linear CP trace-preserving map. If \( d^D_{tr}(\Psi, U \cdot U^\dagger) \leq \epsilon \), then \( d_{\diamond}(\Psi, U \cdot U^\dagger) \leq \sqrt{2\epsilon} \).

The condition is that one of the superoperators corresponds to an isometry. In Theorem 1.2 and throughout we indicate this condition by the superscript \( U \) on the implication symbol.

1.2 Overview of the results

A new inequality for the standard setting

We add an inequality that uses Stinespring dilation to relate \( d^D_{tr} \) to the operator norm \( \|\cdot\|_{op} \). (The fact that \( d^D_{tr}(A \cdot A^\dagger, B \cdot B^\dagger) \leq 2 \|A - B\|_{op} \leq 1 \) follows from Lemma 12.6 of [16], here we prove an opposite bound.) Recall that by Stinespring dilation any linear CP map \( \Psi : L(H) \to L(H') \) can be written as \( \Psi(\cdot) = \text{tr}_{K'}[A \cdot A^\dagger] \) for some (non-unique) linear operator \( A : H \to H' \otimes K' \).
Theorem 1.3 (d_{tr}^P small \implies \| \cdot \|_{op} small) (roughly). Let U : \mathcal{H} \to \mathcal{H}' be an isometry and \Psi : L(\mathcal{H}) \to L(\mathcal{H}') a linear CP trace-nonincreasing map. Denote by A : \mathcal{H} \to \mathcal{H}' \otimes K' some Stinespring-dilation operator of \Psi. If d_{tr}^P(\Psi, U \cdot U^\dagger) \leq \epsilon, then there exists \langle g \rangle \in K' such that \| A - U \otimes \langle g \rangle \|_{op} \leq 2\sqrt{\epsilon}.

To prove Theorem 1.3 we first prove that d_{tr}(\Psi, \Phi) \leq 2d_{tr}^P(\Psi, \Phi) for any pair of linear maps \Psi, \Phi : L(\mathcal{H}) \to L(\mathcal{H}') (Lemma 1). A corollary of Theorem 1.3 is a new version of Theorem 1.2 with the trace-preserving assumption on \Psi relaxed (Corollary 1).

The postselected setting

We define postselection equivalents of d_{tr} and d_{\diamond} in the following way: Let \Psi, \Phi : L(\mathcal{H}) \to L(\mathcal{H}') be postselection superoperators, then

\[ \hat{d}_{tr}(\Psi, \Phi) := \sup_{\rho \in \mathcal{D}(\mathcal{H})} \| \frac{\Psi(\rho)}{\text{tr}[\Psi(\rho)]} - \frac{\Phi(\rho)}{\text{tr}[\Phi(\rho)]} \|_{tr} \]

\[ \hat{d}_{\diamond}(\Psi, \Phi) := \sup_{K} \hat{d}_{tr}(\Psi \otimes I_K, \Phi \otimes I_K), \]

where I_K : L(K) \to L(K) is the identity superoperator. Note that \hat{d}_{tr} and \hat{d}_{\diamond} are pseudometrics: they are zero if \Psi = \Phi (but not only if!), they are symmetric and obey the triangle inequality. Unfortunately, the function of \rho that \hat{d}_{tr} maximizes could be not convex (see Claim 1). We will still show that both suprema are achieved and that the pseudometrics obey inequalities equivalent to Theorems 1.1 to 1.3:

Theorem (roughly). Let U : \mathcal{H} \to \mathcal{H}' be an isometry, \Psi : L(\mathcal{H}) \to L(\mathcal{H}') a postselection superoperator. Denote by A : \mathcal{H} \to \mathcal{H}' \otimes K' a Stinespring-dilation operator of \Psi.

2.1 (Weak subadditivity of \hat{d}_{\diamond}). If the ideal subroutines are trace-preserving, Theorem 1.1 holds also when the distances are measured by \hat{d}_{\diamond}.

2.2 (\hat{d}_{tr} small \implies \hat{d}_{\diamond} small). Theorem 1.2 has a postselection equivalent.

2.3 (\hat{d}_{tr} small \implies \| \cdot \|_{op} small). Theorem 1.3 has a postselection equivalent.

Remarkably, the trace-preserving requirement in Theorem 2.1 is not superfluous; once dropped, a counterexample exists (see Section 4.1). Theorems 2.2 and 2.3 are proven using our main result - the Conversion lemma:

Lemma 2 (Main: Conversion lemma) (roughly). Let \Psi, \Phi be postselection superoperators, \Phi trace-preserving. There exists a positive scalar k such that \hat{d}_{tr}(\Psi, \Phi) and d_{tr}^P(\frac{\Phi}{k}, \Phi) are equivalent up to constant factors.

As a (zero-distance) example take \Phi : L(\mathcal{H}) \to L(\mathcal{H}') to be the identity superoperator between two spatially separated registers \mathcal{H} and \mathcal{H}' of the same dimension. The quantum teleportation with postselection on the appropriate Bell-measurement outcome is a local \Psi that simulates the nonlocal \Phi exactly. By Lemma 2 we have \frac{\Phi}{k} = \Phi, so \Psi’s postselection probability must be k, i.e constant over all input states - indeed we know that tr[\Psi_{telep}(\rho)] = k_{telep} = \text{dim}(\mathcal{H})^{-2}.
2 Preliminaries

In this section we review some standard distance measures on linear trace-nonincreasing CP superoperators. Let $\Psi, \Phi : L(H) \to L(H')$ be such superoperators. The operational trace induced distance is induced from a norm $\| \|_D$, which, in turn, is induced from the trace norm on operators in the following way

$$d_D(\Psi, \Phi) = \| \Psi - \Phi \|_D := \sup_{\rho \in D(H)} \| \Psi(\rho) - \Phi(\rho) \|_{tr}.$$  \hspace{1cm} (2)

The trace-induced distance is similar, the only difference being that the supremum is now over all linear operators $L(H)$, not only density matrices

$$d_{tr}(\Psi, \Phi) = \| \Psi - \Phi \|_{tr} := \sup_{X \in L(H), \| X \|_{tr} \leq 1} \| \Psi(X) - \Phi(X) \|_{tr}.$$  \hspace{1cm} (3)

By definition, this distance upper bounds the first one, $d_D(\Psi, \Phi) \leq d_{tr}(\Psi, \Phi)$. On one hand, Watrous [18] found $\Psi, \Phi$ such that this inequality is strict. On the other hand, the diamond distance, $d_\Diamond(\Psi, \Phi) = \| \Psi - \Phi \|_\Diamond$, is the stabilized version of both

$$d_\Diamond(\Psi, \Phi) := \sup_{K} d_{\text{tr}}(\Psi \otimes I_K, \Phi \otimes I_K)$$

$$= \sup_{K} d_D(\Psi \otimes I_K, \Phi \otimes I_K),$$  \hspace{1cm} (4)

where $I_K : L(K) \to L(K)$ is the identity superoperator. The last equality in (4) follows from the following result of Watrous [17, Lemma 3.45 and Theorem 3.51] by setting $f = \Psi - \Phi$:

**Fact 1 (due to [17]).** For any Hermitian-preserving linear map $f : L(H) \to L(H')$, there exists a normalized vector $|u\rangle \in H \otimes H$ such that setting $K = H$ and $X = |u\rangle \langle u| \in D(H \otimes H)$ achieves the suprema in definitions (4) and (3), i.e. $\|f\|_\Diamond = \|f \otimes I_H(|u\rangle \langle u|)\|_{tr}$.

We have mentioned the Stinespring dilation[19], which states that any linear CP superoperator $\Psi : L(H) \to L(H')$ can be written (non-uniquely) as $\Psi(\cdot) = \text{tr}_{K'}[A \cdot A^\dagger]$ for some extension Hilbert space $K'$ and some linear operator $A : H \to H' \otimes K'$. For linear operators $O : H \to H'$ we will use the operator norm defined as $\|O\|_{op} := \sup_{\|v\| = 1} \|O|v\rangle\|$. We will find the following fact useful:

**Fact 2.** If $\Psi$ is a linear completely positive superoperator and $A$ corresponds to its Stinespring dilation, then $\|\Psi\|_\Diamond = \|\Psi\|_{tr} = \|\Psi\|_D = \|A\|_{op}^2$.

**Proof.** Note that $\|\Psi\|_\Diamond \geq \|\Psi\|_{tr} \geq \|\Psi\|_D \geq \|A\|_{op}^2$, the first two inequalities following immediately from the definitions and the last from the fact that for a positive semidefinite operator $X$ we have $\|X\|_{tr} = \text{tr}[X]$ and so $\|\Psi\|_D = \sup_{\rho \in D(H)} \text{tr}[\rho A A^\dagger]$, which is bigger or equal to $\|A\|_{op}^2 = \sup_{\|v\| = 1} \text{tr}[A|v\rangle\langle v|A^\dagger]$. It remains to show $\|\Psi\|_\Diamond \leq \|A\|_{op}^2$. Apply Fact 1 to get $\|\Psi\|_\Diamond = \text{tr}[\Psi \otimes I_H(|u\rangle \langle u|)] \leq \|A \otimes 1\|_{op}^2 = \|A\|_{op}^2$, which completes the proof. \hfill $\square$

If $\Psi$ is also trace-nonincreasing, we have $\|\Psi(\rho)\|_{tr} = \text{tr}[\Psi(\rho)] \leq 1$ for all $\rho \in D(H)$ so that by definition (2) $\|\Psi\|_{tr} \leq 1$. By definition (3) $\|\Psi(X)\|_{tr} \leq \|\Psi\|_{tr} \|X\|_{tr}$, which with Fact 2 gives:

**Fact 3.** A trace non-increasing CP map $\Psi$ contracts the trace norm: $\|\Psi(\cdot)\|_{tr} \leq \|\cdot\|_{tr}$.
3 Results: Standard setting

In this section we prove Theorem 1.3 and introduce a new version of Theorem 1.2.

**Theorem 1.3** (\(d_{tr}^2\) small \(\Rightarrow\) \(\|\cdot\|_{op}\) small). Let \(U: \mathcal{H} \rightarrow \mathcal{H}'\) be an isometry and \(\Psi: L(\mathcal{H}) \rightarrow L(\mathcal{H}')\) a linear CP trace-nonincreasing superoperator. Denote by \(A: \mathcal{H} \rightarrow \mathcal{H}' \otimes \mathcal{K}'\) a Stinespring-dilation operator of \(\Psi\). If \(d_{tr}^2(\Psi, U \cdot U^\dagger) \leq \epsilon\), then there exists \(|g\rangle \in \mathcal{K}',\) with the norm \(1 - \epsilon \leq \|g\|^2 \leq \|A\|^2_{op} \leq 1,\) such that \(\|A - U \otimes |g\rangle\|_{op} \leq 2\sqrt{\epsilon}\).

As a corollary we get a version of Theorem 1.2 with a worse bound but working also for \(\Psi\)s that decrease the trace of some inputs. This we need for the next section.

**Corollary 1.** Let \(U: \mathcal{H} \rightarrow \mathcal{H}'\) be an isometry and \(\Psi: L(\mathcal{H}) \rightarrow L(\mathcal{H}')\) a linear CP trace-nonincreasing superoperator. If \(d_{tr}^2(\Psi, U \cdot U^\dagger) \leq \epsilon\), then \(d_\diamond(\Psi, U \cdot U^\dagger) \leq 4\sqrt{\epsilon} + \epsilon\).

The corollary follows from \(d_\diamond(\Psi, U \cdot U^\dagger) \leq d_\diamond(\Psi, \|g\|^2 U \cdot U^\dagger) + \|\|g\|^2 - 1\|\) (an application of the triangle inequality for \(d_\diamond\), from the fact that partial trace contracts \(d_\diamond\) and from Lemma 12.6 of [16], according to which \(d_\diamond(A \cdot A^\dagger, B \cdot B^\dagger) \leq 2\|A - B\|_{op}\) if \(\|A\|_{op}, \|B\|_{op} \leq 1\).

To prove Theorem 1.3 we first prove the following:

**Lemma 1.** Let \(\Psi, \Phi: L(\mathcal{H}) \rightarrow L(\mathcal{H}')\) be two linear maps. \(d_{tr}(\Psi, \Phi) \leq 2 d_{tr}^2(\Psi, \Phi)\).

**Proof.** For any pair \(|u\rangle, |v\rangle \in \mathcal{H}\) of unit vectors define the unnormalized \(|w_k\rangle := |u\rangle + i^k |v\rangle\) and observe that

\[
\sum_{k=0}^{3} \|w_k\|^2 = 8 \quad \text{and} \quad |u\rangle \langle v| = \frac{1}{4} \sum_{k=0}^{3} i^k |w_k\rangle \langle w_k|
\]

because the roots of unity sum to zero, \(\sum_{k=0}^{3} i^k = 0\). For any linear \(f: L(\mathcal{H}) \rightarrow L(\mathcal{H}')\) we get by the triangle inequality and the absolute homogeneity of trace norm

\[
\|f(|u\rangle \langle v|)\|_{tr} \leq \frac{1}{4} \sum_{k=0}^{3} \|f(|w_k\rangle \langle w_k|)\|_{tr}
\]

\[
\leq \frac{1}{4} \left( \sum_{k=0}^{3} \|w_k\|^2 \right) \|f\|^2_{tr} = 2 \|f\|^2_{tr}.
\]

By a convexity argument\(^\dagger\) there exist \(|u^*\rangle, |v^*\rangle \in \mathcal{H}\) that achieve the supremum of Eq. (3). We get \(\|f\|_{tr} = \|f(|u^*\rangle \langle v^*|)|_{tr} \leq 2 \|f\|^2_{tr}\). Setting \(f = \Psi - \Phi\) completes the proof.

**Proof of Theorem 1.3.** By the definition of \(d_{tr}\) in (3) we have for all normalized \(|u\rangle, |v\rangle \in \mathcal{H}\)

\[
d_{tr}(\Psi, U \cdot U^\dagger) \geq \left\| \Psi(|u\rangle \langle v|) - U |u\rangle \langle v| U^\dagger \right\|_{tr}
\]

\[
= \left\| U^\dagger \Psi(|u\rangle \langle v|) U - |u\rangle \langle v| \right\|_{tr} \| |v\rangle \langle u| \|_{op}
\]

\[
\geq \left\| U^\dagger \Psi(|u\rangle \langle v|) U |v\rangle \langle u| - |u\rangle \langle u| \right\|_{tr}
\]

\[
\geq \left| \langle u | U^\dagger \Psi (|u\rangle \langle v|) U |v\rangle - 1 \right|
\]

\(^\dagger\)Operators of the form \(|u\rangle \langle v|\) are the extreme points of the convex set \(\{ X \in L(\mathcal{H}); \|X\|_{tr} \leq 1 \}\) and \(\|f(\cdot)\|_{tr}\) is a convex function; it maps convex set to a convex set and extreme points to extreme points.
where we inserted \( \|v\rangle\langle u\|_{op} = 1 \) into the second line. The last two inequalities are \( |X|_{tr} \|O\|_{op} \geq \|XO|_{tr} \) and \( \|X\|_{tr} \geq |\text{tr}[X]| \) that hold for any \( X,O \in L(H) \) (see [16, Lemma 10]). When \( |u\rangle = |v\rangle \) we can write \( d_{tr}^{D} \) instead of \( d_{tr} \) (see definition (2)). Writing out the Stinespring-dilation for \( \Psi \) we get

\[
d_{tr}(\Psi, U \cdot U^\dagger) \geq |\text{tr}(\langle g_u\rangle A_{u}) - 1| \tag{5}
\]

with \( |g_u\rangle := (\langle u| U^\dagger \otimes \mathbb{1}_{K'} A |u\rangle) \tag{6} \)

From Fact 2 and \( \Psi \) being CP trace-nonincreasing we have

\[
\|A |u\rangle\|^2 \leq \|A\|^2_{op} = \|\Psi\|^2_{\diamond} \leq 1 \tag{7}
\]

From the \( d_{tr}^{D} \) equivalent of (5) and from (6) and (7) we get

\[
1 - d_{tr}^{D}(\Psi, U \cdot U^\dagger) \leq ||g_u\rangle|^2 \leq \|A\|^2_{op} \leq 1.
\]

So the norm of \( |g_u\rangle \) is withing the bounds required by Theorem 1.3. For any normalized \( |u\rangle, |v\rangle \in H \) we can use (7) again to upper bound

\[
\|A |v\rangle - U |v\rangle \otimes |g_u\rangle\|^2 = \|A |v\rangle\|^2 + ||g_u\rangle\|^2 - 2 \Re \left( \langle v| U^\dagger \otimes \langle g_u| A |v\rangle \right)
\]

\[
\leq 2 \left( 1 - \Re(g_u|g_v) \right)
\]

\[
\leq 2 d_{tr}(\Psi, U \cdot U^\dagger)
\]

where \( \Re \) stands for the real part of the complex number and the last inequality follows from (5). Combining with Lemma 1 we find that \( \|A - U \otimes |g_u\rangle\|_{op} \leq 2 \sqrt{d_{tr}^{D}(\Psi, U \cdot U^\dagger)} \) with \( |g_u\rangle \) is defined by Eq. (6) for any normalized \( |u\rangle \in H \).

\[\square\]

### 4 Results: Postselected setting

In the introduction we defined the trace-induced distance and the diamond distance for the postselected setting. For any postselection superoperators \( \Psi, \Phi : L(H) \rightarrow L(H') \) we had

\[
\hat{d}_{tr}(\Psi, \Phi) := \sup_{\rho \in \mathcal{D}(H)} \frac{\left\| \frac{\Psi(\rho)}{\text{tr}[\Psi(\rho)]} - \frac{\Phi(\rho)}{\text{tr}[\Phi(\rho)]} \right\|_{tr}}{f_{\Psi, \Phi}(\rho)} \tag{8}
\]

\[
\hat{d}_{\diamond}(\Psi, \Phi) := \sup_{K} \hat{d}_{tr}(\Psi \otimes \mathcal{I}_K, \Phi \otimes \mathcal{I}_K), \tag{9}
\]

where we added the shorthand \( f_{\Psi, \Phi} \) to denote the objective function of the maximisation in (8). In general, this function is not convex.

**Claim 1.** There exist \( \Psi, \Phi \) such that \( f_{\Psi, \Phi} \) is not convex on \( \mathcal{D}(H) \).

**Proof.** Consider the maps \( \Psi_\epsilon, \Phi_\epsilon : L(C^2) \rightarrow L(C^2) \) where \( \Psi_\epsilon(\rho) = (1-\epsilon) |0\rangle\langle 0| \rho |0\rangle\langle 0| + \epsilon |1\rangle\langle 1| \rho |1\rangle\langle 1| \) and \( \Phi_\epsilon(\rho) = (1-\epsilon) |1\rangle\langle 1| \rho |1\rangle\langle 1| + \epsilon |0\rangle\langle 0| \rho |0\rangle\langle 0| \), for some \( \epsilon \in (0, \frac{1}{2}) \). Both are postselection superoperators so \( f_{\Psi_\epsilon, \Phi_\epsilon} \) is well-defined on \( \mathcal{D}(C^2) \). Observe that \( f_{\Psi_\epsilon, \Phi_\epsilon}(|0\rangle\langle 0|) = 0, f_{\Psi_\epsilon, \Phi_\epsilon}(|1\rangle\langle 1|) = 0 \), and \( f_{\Psi_\epsilon, \Phi_\epsilon}(\frac{1}{2} |0\rangle\langle 0| + \frac{1}{2} |1\rangle\langle 1|) = 2 - 4\epsilon \), therefore this \( f_{\Psi_\epsilon, \Phi_\epsilon} \) is not convex. See Fig. 1. \[\square\]
We choose a $\rho$ for its non-postselected predecessor, sup $K_D$ of the convex $D$.

Let $\Phi$ be postselection superoperators on $L(H)$. Then $\hat{d}_\triangledown(\Psi, \Phi) = \hat{d}_{tr}(\Psi \otimes I_H, \Phi \otimes I_H)$.

Proof. The $\geq$ direction follows from definition (9). We prove here the $\leq$ direction.

First, we claim that for any normalized $|v\rangle \in H \otimes K$, where $K$ is of a higher dimension than $H$, there exists a normalized $|u\rangle \in H \otimes H$ such that

$$f_{\Psi \otimes I_K, \Phi \otimes I_K}(|v\rangle\langle v|) = f_{\Psi \otimes I_H, \Phi \otimes I_H}(|u\rangle\langle u|).$$  \hspace{1cm} (10)

The proof is standard: Since the Schmidt decomposition of $|v\rangle$ has at most $\dim(H)$ terms, there exists $|u\rangle \in H \otimes H$ such that $|v\rangle = (1_H \otimes U)|u\rangle$ with $U : H \to K$ an isometry; $U^\dagger U = 1_H$. Plugging this in for $|v\rangle$, noting that superoperators acting on different registers commute and that $\| (1_H \otimes U) \cdot (1_H \otimes U^\dagger) \|_{tr} = \|\cdot\|_{tr}$ (for example, by Fact 3 applied in both directions) we get Eq. (10).³

Next, we prove that for any postselection superoperators $A, B : L(N) \to L(M)$ and for any $\rho \in D(N)$ there exists $|v\rangle \in N \otimes N$ such that

$$f_{A,B}(\rho) \leq f_{A \otimes I_N, B \otimes I_N}(|v\rangle\langle v|).$$  \hspace{1cm} (11)

We choose a $|v\rangle$ that is a purification of $\rho$, substituting $\rho = tr_N(|v\rangle\langle v|)$ into the left-hand side. The inequality follows by commutation and by $|tr_N(\cdot)|_{tr} \leq \|\cdot\|_{tr}$ (Fact 3).

Finally, we set $A = \Psi \otimes I_{K'}, B = \Phi \otimes I_{K'}$ in (11) and to the right-hand side we apply Eq. (10) with $K = K' \otimes H \otimes K'$. We get $f_{\Psi \otimes I_{K'}, \Phi \otimes I_{K'}}(\rho) \leq \max_{|u\rangle \in H \otimes H} f_{\Psi \otimes I_H, \Phi \otimes I_H}(|u\rangle\langle u|)$ for any $K'$ and $\rho \in H \otimes K'$. Taking suprema over them completes the proof. \hfill $\Box$

³The entire proof would stop here if the objective function was convex: Since pure states are the extreme points of the convex $D(H \otimes K)$, we would have some $|v\rangle\langle v| \in D(H \otimes K)$ achieve each $\hat{d}_{tr}$ in (9), similarly as in the footnote on the previous page.
Interestingly, Corollary 2, and consequently Theorem 2.1, do not hold if \( \hat{s} \) serving postselection superoperator. Then preserving. Consider the following two postselection superoperators, constant on

\[ K \]
trace-preserving with a

Corollary 2 (Weak contractivity).

where \( I \)
be postselection superoperators, \( \Phi \) equal to

Figure 2: As opposed to \( d_\Diamond \), the renormalisation-containing \( \hat{d}_\Diamond \) is not contractive. We can visualize the counterexample in the text by probability vectors \( p \in \mathbb{R}^3 \), because all the density matrices of interest are of the form \( \text{diag}(p) \). The diamond distance is the \( L_1 \) distance of the corresponding vector pair. It grows after the grey pair is mapped by \( \tau (\epsilon \to 0) \) to the dashed light blue pair and renormalized to the dark blue.

4.1 Weak subadditivity and weak contractivity

In this section we discuss how \( \hat{d}_\Diamond \) behaves with respect to composition of superoperators.

Theorem 2.1 (Weak subadditivity). Let \( \Psi, \Phi : L(H) \to L(H') \) and \( \Psi', \Phi' : L(H' \otimes K') \to L(H'') \) be postselection superoperators, \( \Phi' \) being trace-preserving. Then

\[
\hat{d}_\Diamond (\Psi' \circ (\Psi \otimes I_{K'}), \Phi' \circ (\Phi \otimes I_{K'})) \leq \hat{d}_\Diamond (\Psi', \Phi') + \hat{d}_\Diamond (\Psi, \Phi),
\]

(12)

where \( I_{K'} : L(K') \to L(K') \) is the identity superoperator.

Setting \( K' \) trivial and \( \Psi' = \Phi' = : \tau \) we immediately get that \( \hat{d}_\Diamond \) contracts under the composition with a trace-preserving postselection superoperator:

Corollary 2 (Weak contractivity). With \( \Psi, \Phi \) as before, let \( \tau : L(H') \to L(H'') \) be a trace-preserving postselection superoperator. Then \( \hat{d}_\Diamond (\tau \circ \Psi, \tau \circ \Phi) \leq \hat{d}_\Diamond (\Psi, \Phi) \).

Interestingly, Corollary 2, and consequently Theorem 2.1, do not hold if \( \tau, \Phi' \) are not trace-preserving. Consider the following two postselection superoperators, constant on \( \rho \in \mathcal{D}(\mathbb{C}^3) \):

\[
\Psi(\rho) = \frac{1}{2} \ket{0}\bra{0} + \frac{1}{2} \ket{1}\bra{1} \quad \text{and} \quad \Phi(\rho) = \frac{1}{2} \ket{0}\bra{0} + \frac{1}{2} \ket{2}\bra{2} \quad \text{if} \rho \text{ is outside } \mathcal{D}(\mathbb{C}^3) \text{ scale the outputs by } \text{tr}\rho.
\]

Now let \( \tau(\rho) = (1 - \epsilon) \Pi \rho \Pi + \epsilon \rho \) with \( \Pi = \ket{1}\bra{1} + \ket{2}\bra{2} \) and \( \epsilon \in (0,1) \). Observe that

\[
\hat{d}_\Diamond (\Psi, \Phi) = 1 \quad \text{but} \quad \hat{d}_\Diamond (\tau \circ \Psi, \tau \circ \Phi) = \frac{2}{1 + \epsilon} > 2 - 2\epsilon, \quad \text{i.e. this } \tau \text{ increases the distance! See Fig. 2.}
\]

Proof of Theorem 2.1. For shorthand, denote \( \bar{\Psi} := \Psi \otimes I_{K'} \otimes I_{K''} \) and similarly for \( \bar{\Phi} \); denote \( \Psi' := \Psi' \otimes I_{K''} \) and similarly for \( \Phi' \). Then by definition (9) the left-hand side of inequality (12) is equal to

\[
\sup_{\rho \in \mathcal{D}(H \otimes K' \otimes K'')} \frac{\left\| \frac{\bar{\Psi}'(\bar{\Psi}(\rho))}{\text{tr}[\bar{\Psi}'(\bar{\Psi}(\rho))]} - \frac{\bar{\Phi}'(\bar{\Phi}(\rho))}{\text{tr}[\bar{\Phi}'(\bar{\Phi}(\rho))]} \right\|_\tau(h(\rho)),
\]

because \( \Phi' \) is trace-preserving and linear. Let \( \rho' = \bar{\Psi}(\rho)/\text{tr}[\bar{\Psi}(\rho)] \). From the linearity of \( \bar{\Psi}' \), the
argument of the supremum is equal to

\[ h(\rho) = \left\| \frac{\bar{\Psi}'(\rho')}{\operatorname{tr}[\bar{\Psi}'(\rho')]} - \frac{\bar{\Phi}(\rho)}{\operatorname{tr}[\bar{\Phi}(\rho)]} \right\|_{\text{tr}} \]

\[ \leq \left\| \frac{\bar{\Psi}'(\rho')}{\operatorname{tr}[\bar{\Psi}'(\rho')]} - \bar{\Phi}'(\rho') \right\|_{\text{tr}} + \left\| \frac{\bar{\Phi}(\rho) - \bar{\Phi}'(\rho')}{\operatorname{tr}[\bar{\Phi}(\rho)]} \right\|_{\text{tr}} \]

because \( \bar{\Phi}' \) contracts the trace norm (Fact 3). We get (12) by taking the supremum. \( \square \)

4.2 Inequalities via conversion

To prove inequalities in the postselected setting, sometimes it is possible to go back to the standard setting and use the results that hold there. Here we apply this method to obtain the following:

**Theorem 2.2 (\( \hat{d}_{\text{tr}} \) small \( \Rightarrow \hat{d}_d \) small).** Let \( U : \mathcal{H} \to \mathcal{H}' \) be an isometry and \( \Psi : L(\mathcal{H}) \to L(\mathcal{H}') \) a postselection superoperator. Assume \( \hat{d}_{\text{tr}}(\Psi, U \cdot U^\dagger) \leq \epsilon \). Then \( \hat{d}_d(\Psi, U \cdot U^\dagger) \leq 24\sqrt{\epsilon} + 18\epsilon \).

**Theorem 2.3 (\( \hat{d}_{\text{tr}} \) small \( \Rightarrow \| \cdot \|_{\text{op}} \) small).** Let \( U : \mathcal{H} \to \mathcal{H}' \) be an isometry, \( \Psi : L(\mathcal{H}) \to L(\mathcal{H}') \) a postselection superoperator. Let \( A : \mathcal{H} \to \mathcal{H}' \otimes \mathcal{K}' \) be a Stinespring-dilation operator of \( \Psi \). If \( \hat{d}_{\text{tr}}(\Psi, U \cdot U^\dagger) \leq \epsilon \), then there exists \( |g\rangle \in \mathcal{K}' \), \((1 - 9\epsilon)\|A\|_{\text{op}}^2 \leq |\langle g |\|^2 \leq \|A\|_{\text{op}}^2 \), such that \( \|A - U \otimes |g\rangle\|_{\text{op}} \leq 6\|A\|_{\text{op}} \sqrt{\epsilon} \).

The conversion between the settings is done by the following lemma, which we prove in the next section.

**Lemma 2 (Conversion Lemma).** Let \( \Psi, \Phi : L(\mathcal{H}) \to L(\mathcal{H}') \) be postselection superoperators, with \( \Phi \) trace-preserving. Then \( \left| \operatorname{tr}[\Psi(\rho)] - \|\Psi\|_\diamond \right| \leq \alpha \|\Psi\|_\diamond \hat{d}_{\text{tr}}(\Psi, \Phi) \) for all \( \rho \in \mathcal{D}(\mathcal{H}) \) and

\[ \frac{1}{2} \hat{d}_{\text{tr}}(\Psi, \Phi) \leq d_{\text{tr}}^D \left( \frac{\Psi}{\|\Psi\|_\diamond}, \Phi \right) \leq (\alpha + 1) \hat{d}_{\text{tr}}(\Psi, \Phi), \quad (13) \]

with \( \alpha = 40 / \max_{\rho_0, \rho_1 \in \mathcal{D}(\mathcal{H})} \|\Phi(\rho_0) - \Phi(\rho_1)\|_{\text{tr}} \) in general but with \( \alpha = 8 \) if \( \Phi \) is also a unitary superoperator.

The inequalities of Lemma 2 hold also for diamond distances. This is immediate from the fact that the identity superoperator \( \mathcal{I}_\mathcal{K} \) is unitary, that \( \|\Psi \otimes \mathcal{I}_\mathcal{K}\|_\diamond = \|\Psi\|_\diamond \), and that taking supremum (\( \sup_{\mathcal{K}} \)) preserves inequalities. Thus Lemma 2 allows us to move from postselection distances to the corresponding standard distances of linear trace non-increasing CP maps (since \( \|\Psi\|_\diamond = \max_{\rho \in \mathcal{D}(\mathcal{H})} \operatorname{tr}[\Psi(\rho)] \) the superoperator \( \Psi / \|\Psi\|_\diamond \) is indeed trace non-increasing). We are ready to prove the theorems.

**Proof of Theorem 2.2.** Using the shorthand \( D := d_{\text{tr}}^D \left( \frac{\Psi}{\|\Psi\|_\diamond}, U \cdot U^\dagger \right) \) we claim that

\[ \frac{1}{2} \hat{d}_d(\Psi, U \cdot U^\dagger) \leq d_{\text{tr}}^D \left( \frac{\Psi}{\|\Psi\|_\diamond}, U \cdot U^\dagger \right) \leq 4\sqrt{D} + D \]

The first inequality is from Lemma 2 applied to diamond distances, the second from Corollary 1. Substituting \( D \leq 9\epsilon \), also from Lemma 2, completes the proof. \( \square \)
Proof of Theorem 2.3. Take $D := d^D_{tr} (\Psi/\|\Psi\|_\Diamond, U \cdot U^\dagger)$. Since $\Psi(\cdot) = \text{tr}_\mathcal{K} [A \cdot A^\dagger]$ and $|A|_{op}^2 = \|\Psi\|_\Diamond$ (Fact 2), $A/|A|_{op}$ is a Stinespring-dilation operator of $\Psi/\|\Psi\|_\Diamond$. Then by Theorem 1.3 there exists $|v\rangle \in \mathcal{K}'$ such that

$$1 - 9\epsilon \leq 1 - D \leq \|v\|^2 \leq 1,$$

$$\left\| \frac{A}{|A|_{op}} - U \otimes |v\rangle \right\| \leq 2\sqrt{D} \leq 6\sqrt{\epsilon},$$

where we used $D \leq 9\epsilon$ from Lemma 2. Taking $|g\rangle := |A|_{op}|v\rangle$ completes the proof. \hfill \qed

4.3 Proof of Convergence lemma

In this section we prove our main result, Lemma 2. When $\Phi$ is not promised to be unitary, the inverse selection superoperators $\Phi(\rho) = |0\rangle\langle 0| \text{tr} \rho$ and $\Psi(\rho) = \frac{1}{2} |0\rangle\langle 0| \text{tr} \rho + \frac{1}{2} |0\rangle\langle 0| \rho |0\rangle\langle 0|$ and note that $\Phi$ is indeed trace-preserving and that $\|\Psi\|_\Diamond = 1$. We have $d_{tr}(\Psi, \Phi) = 0$, but $d^D_{tr} (\Psi, \Phi) = \frac{1}{2}$, so for the inequality (13) to hold we need $\alpha \to \infty$.

Proof of Lemma 2 (the left inequality in (13)). The inequality actually holds also when $\|\Psi\|_\Diamond$ is replaced by any $k \neq 0$. To prove this, observe that for any $\rho \in \mathcal{D}(\mathcal{H})$, $\Psi(\rho)$ is positive semidefinite and we have $\|\Psi(\rho)\|_{tr} = \text{tr}[\Psi(\rho)]$. Therefore for all $\rho \in \mathcal{D}(\mathcal{H})$

$$\left\| \frac{\Psi(\rho)}{\text{tr}[\Psi(\rho)]} - \frac{\Psi(\rho)}{k} \right\|_{tr} = 1 - \frac{\text{tr}[\Psi(\rho)]}{k} \leq \frac{\text{tr}[\Phi(\rho)] - \text{tr}[\Psi(\rho)]}{k} \leq \frac{\text{tr}[\Phi(\rho)]}{k} - \frac{\text{tr}[\Psi(\rho)]}{k},$$

where we used the fact that $\Phi$ is trace-preserving. From the triangle inequality and the above we get

$$\left\| \frac{\Psi(\rho)}{\text{tr}[\Psi(\rho)]} - \Phi(\rho) \right\|_{tr} \leq \left\| \frac{\Psi(\rho)}{\text{tr}[\Psi(\rho)]} - \frac{\Psi(\rho)}{k} \right\|_{tr} + \left\| \frac{\Psi(\rho)}{k} - \Phi(\rho) \right\|_{tr} \leq \frac{2}{k} \left\| \frac{\Psi(\rho)}{k} - \Phi(\rho) \right\|_{tr}.$$

Taking $\sup_{\rho}$ gives $\hat{d}_{tr}(\Psi, \Phi) \leq 2 d^D_{tr}(\Psi, \Phi)$ as required. \hfill \qed

Proof of Lemma 2 (the rest). Use the shorthand $\hat{D} := \hat{d}_{tr}(\Psi, \Phi)$. We will first prove the bound $|\text{tr}[\Psi(\rho)] - \|\Psi\|_\Diamond| \leq \alpha \|\Psi\|_\Diamond \hat{D}$. For this purpose we will use the fact that $\rho \mapsto \Phi(\rho)/\text{tr}[\Phi(\rho)] = \Phi(\rho)$ is linear in $\rho$. For any $\rho_0, \rho_1 \in \mathcal{D}(\mathcal{H})$ denote $\rho_p := (1 - p)\rho_0 + p \rho_1 = \rho_0 + p \Delta$ where $p \in [0, 1]$ and $\Delta := \rho_1 - \rho_0$. Since $\|\Psi\|_\Diamond = \|\Psi\|_{tr} = \max_{\rho \in \mathcal{D}(\mathcal{H})} \text{tr}[\Psi(\rho)]$ (see Fact 2), we actually want to upper bound $|\text{tr}[\Psi(\Delta)]|$. We have

$$\left\| \frac{\Psi(\rho_p)}{\text{tr}[\Psi(\rho_p)]} - \Phi(\rho_p) \right\|_{tr} \leq \hat{D},$$

$$\frac{1}{\text{tr}[\Psi(\rho_p)]} \left\| \Psi(\rho_0) + p \Psi(\Delta) - (\text{tr}[\Psi(\rho_0)] + p \text{tr}[\Psi(\Delta)]) (\Phi(\rho_0) + p \Phi(\Delta)) \right\|_{tr} \leq \hat{D},$$

$$\left\| A_0 + Bp - Cp^2 \right\|_{tr} \leq \|\Psi\|_\Diamond \hat{D},$$

10
where we rewrote the left-hand side as a trace norm of a polynomial in $p \in [0, 1]$ with the coefficients $A_0 := \Psi(\rho_0) - \tr[\Psi(\rho_0)] \Phi(\rho_0)$, $B := \Psi(\Delta) - \tr[\Psi(\rho_0)] \Phi(\Delta) - \tr[\Psi(\Delta)] \Phi(\rho_0)$ and $C := \tr[\Psi(\Delta)] \Phi(\Delta)$.

To show that $|\tr[\Psi(\Delta)]|$ is small we will bound $\|C\|_{tr}$. Define $A_1$ analogously to $A_0$ and note that $\|A_i\|_{tr} \leq \tr[\psi(\rho_i)] \widehat{D} \leq \|\Psi\|_{\diamond} \widehat{D}$. Observe also that $B = A_1 - A_0 + C$. By the triangle inequality

$$
\|Cp - Cp^2\|_{tr} \leq \|A_0 + (A_1 - A_0 + C)p - Cp^2\|_{tr} + \|A_0(1-p)\|_{tr} + \|A_1 p\|_{tr}
$$

Taking $p = \frac{1}{2}$ so that the bound is the tightest and substituting $\Delta = \rho_1 - \rho_0$ into the definition of $C$ we get that for all pairs $\rho_0, \rho_1 \in \mathcal{D}(\mathcal{H})$

$$
|\tr[\Psi(\rho_1)] - \tr[\Psi(\rho_0)]| \leq \frac{8\|\Psi\|_{\diamond} \widehat{D}}{\|\Phi(\rho_1) - \Phi(\rho_0)\|_{tr}}.
$$

(14)

If $\Phi$ is not unitary, denote by $\rho_0^*, \rho_1^*$ the two states that maximize the denominator, i.e. $\Phi(\rho_0^*)$ and $\Phi(\rho_1^*)$ are the furthest away in trace norm. Call their distance $s$. Now for any state $\rho \in \mathcal{D}(\mathcal{H})$, $\Phi(\rho)$ is more than $\frac{s}{2}$-far either from $\Phi(\rho_0^*)$ or from $\Phi(\rho_1^*)$, or both. Therefore, for all pairs $\rho, \rho' \in \mathcal{D}(\mathcal{H})$ there exist $i, j \in \{0, 1\}$ such that

$$
|\tr[\Psi(\rho)] - \tr[\Psi(\rho')]| \leq \frac{8\|\Psi\|_{\diamond} \widehat{D}}{s/2} + \frac{8\|\Psi\|_{\diamond} \widehat{D}}{s/2} + \frac{8\|\Psi\|_{\diamond} \widehat{D}}{s/2} = \frac{40}{s} \|\Psi\|_{\diamond} \widehat{D}.
$$

(15)

If $\Phi$ is a unitary superoperator, we can obtain a tighter bound, because then for any orthogonal pure states $\rho, \rho^\perp \in \mathcal{D}(\mathcal{H})$ we have $\|\Phi(\rho) - \Phi(\rho^\perp)\|_{tr} = \|\rho - \rho^\perp\|_{tr} = 2$ and Eq. (14) gives $|\tr[\Psi(\rho)] - \tr[\Psi(\rho^\perp)]| \leq 4\|\Psi\|_{\diamond} \widehat{D}$. Now note that for any pure state $\rho \in \mathcal{D}(\mathcal{H})$ we can write the completely mixed state as $\frac{1}{d} \mathbb{1} = \frac{1}{d} \rho + \frac{1}{d} \sum_{i=1}^{d-1} \rho_i^\perp$, with $d = \dim(\mathcal{H})$ and with each $\rho_i^\perp$ pure and orthogonal to $\rho$. We get that

$$
|\tr[\Psi(\rho)] - \tr[\Psi\left(\frac{1}{d} \mathbb{1}\right)]| \leq \frac{1}{d} \sum_{i=1}^{d-1} |\tr[\Psi(\rho)] - \tr[\Psi(\rho_i^\perp)]| \leq 4\|\Psi\|_{\diamond} \widehat{D}
$$

(16)

holds for any pure $\rho$. By the convexity of $\mathcal{D}(\mathcal{H})$, it in fact holds for any $\rho \in \mathcal{D}(\mathcal{H})$.

To continue both the unitary and the non-unitary case, choose $\rho'$ that maximizes $\tr[\Psi(\rho')]$. Then for any $\rho \in \mathcal{D}(\mathcal{H})$

$$
|\tr[\Psi(\rho)] - |\Psi\|_{\diamond} = |\tr[\Psi(\rho)] - \tr[\Psi(\rho')]| \leq \alpha \|\Psi\|_{\diamond} \widehat{D}
$$

with $\alpha$ as in the statement of Lemma 2 (the non-unitary case follows from (15), the unitary from (16) and the triangle inequality). This proves the first conclusion of Lemma 2. To prove the
right inequality in (13) note that \( \|\Psi(\rho)\|_{tr} = \text{tr}[\Psi(\rho)] \) from \( \Psi \)'s complete positivity, and therefore
\[
\left\| \frac{\Psi(\rho)}{\|\Psi\|_\infty} - \frac{\Phi(\rho)}{\|\Phi\|_\infty} \right\|_{tr} \leq \alpha \tilde{D}.
\]
We get that for all \( \rho \in \mathcal{D}(\mathcal{H}) \)
\[
\left\| \frac{\Psi(\rho)}{\|\Psi\|_\infty} - \frac{\Phi(\rho)}{\|\Phi\|_\infty} \right\|_{tr} \leq \alpha \tilde{D} + \tilde{D}.
\]
Taking \( \sup_\rho \) completes the proof. \( \Box \)

Acknowledgments

The author would like to thank Dorit Aharonov, Itai Leigh, and Matan Seidel for useful discussions. This research was supported by Simons Foundation (grant 385590) and Israel Science Foundation (grants 2137/19 and 1721/17).

References

[1] Emanuel Knill, Raymond Laflamme, and Gerald J Milburn. A scheme for efficient quantum computation with linear optics. Nature, 409(6816):46–52, 2001.

[2] Emanuel Knill. Quantum computing with realistically noisy devices. Nature, 434(7029):39–44, 2005.

[3] Ben W Reichardt. Error-detection-based quantum fault tolerance against discrete pauli noise. Preprint at https://arxiv.org/abs/quant-ph/0612004, 2006.

[4] Panos Aliferis, Daniel Gottesman, and John Preskill. Accuracy threshold for postselected quantum computation. Quantum Information & Computation, 8(3):181–244, 2008.

[5] Scott Aaronson. Quantum computing, postselection, and probabilistic polynomial-time. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 461(2063):3473–3482, 2005.

[6] Michael J Bremner, Richard Jozsa, and Dan J Shepherd. Classical simulation of commuting quantum computations implies collapse of the polynomial hierarchy. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 467(2126):459–472, 2011.

[7] Michael J Bremner, Ashley Montanaro, and Dan J Shepherd. Achieving quantum supremacy with sparse and noisy commuting quantum computations. Quantum, 1:8, 2017.

[8] Scott Aaronson and Alex Arkhipov. The computational complexity of linear optics. In Proceedings of the forty-third annual ACM symposium on Theory of computing, pages 333–342, 2011.

[9] Tomoyuki Morimae, Keisuke Fujii, and Joseph F Fitzsimons. Hardness of classically simulating the one-clean-qubit model. Physical review letters, 112(13):130502, 2014.
[10] Adam Bouland, Bill Fefferman, Chinmay Nirkhe, and Umesh Vazirani. On the complexity and verification of quantum random circuit sampling. *Nature Physics*, 15(2):159–163, 2019.

[11] Michael E Beverland, Earl Campbell, Mark Howard, and Vadym Kliuchnikov. Lower bounds on the non-clifford resources for quantum computations. *Quantum Science and Technology*, 2020.

[12] Scott Aaronson, Xinyi Chen, Elad Hazan, Satyen Kale, and Ashwin Nayak. Online learning of quantum states. *Journal of Statistical Mechanics: Theory and Experiment*, 2019(12):124019, 2019.

[13] Jingliang Gao. Quantum union bounds for sequential projective measurements. *Physical Review A*, 92(5):052331, 2015.

[14] Scott Aaronson. QMA/qpoly ⊆ PSPACE/poly: de-Merlinizing quantum protocols. In *21st Annual IEEE Conference on Computational Complexity (CCC’06)*, pages 13–pp. IEEE, 2006.

[15] A Yu Kitaev. Quantum computations: algorithms and error correction. *Russian Mathematical Surveys*, 52(6):1191, 1997.

[16] Dorit Aharonov, Alexei Kitaev, and Noam Nisan. Quantum circuits with mixed states. In *Proceedings of the thirtieth annual ACM symposium on Theory of computing*, pages 20–30, 1998.

[17] John Watrous. *The theory of quantum information*. Cambridge University Press, 2018.

[18] John Watrous. Notes on super-operator norms induced by schatten norms. *Quantum Information & Computation*, 5(1):58–68, 2005.

[19] W Forrest Stinespring. Positive functions on C*-algebras. *Proceedings of the American Mathematical Society*, 6(2):211–216, 1955.