Quantum channel tomography: optimal bounds and a Heisenberg-to-classical phase transition*

Kean Chen† Filippo Girardi‡ Aadil Oufkir§ Nengkun Yu¶ Zhicheng Zhang‖

Abstract

How many black-box queries to a quantum channel are needed to learn its full classical description? This question lies at the heart of quantum channel tomography (also known as quantum process tomography), a fundamental task in the characterization and validation of quantum hardware. Despite extensive prior work, the optimal query complexity for quantum channel tomography is far from fully understood.

In this paper, we study tomography of an unknown quantum channel with input dimension d_1, output dimension d_2, and Kraus rank at most r, to within error ε. We identify the dilation rate $\tau = rd_2/d_1$ (which always satisfies $\tau \geq 1$ due to the trace preservation of quantum channels) as a key parameter, and establish that the optimal query complexity of channel tomography exhibits distinct scaling laws across three regimes of τ.

- In the boundary regime ($\tau = 1$): we show that the query complexity is $\Theta(rd_1d_2/\varepsilon)$ for Choi trace norm error ε, and is upper bounded by $O(\min\{rd_1^{1.5}d_2/\varepsilon, rd_1d_2/\varepsilon^2\})$ and lower bounded by $\Omega(rd_1d_2/\varepsilon)$ for diamond norm error ε.
- In the away-from-boundary regime ($\tau \geq 1 + \Omega(1)$): we show that the query complexity is $\Theta(rd_1d_2/\varepsilon^2)$ for both Choi trace norm and diamond norm errors ε.

Our results uncover a sharp Heisenberg-to-classical phase transition in the query complexity of quantum channel tomography: at $\tau = 1$, the optimal query complexity exhibits Heisenberg scaling $1/\varepsilon$, whereas for $\tau \geq 1 + \Omega(1)$, it exhibits classical scaling $1/\varepsilon^2$. In addition, we show that in the near-boundary regime ($1 < \tau < 1 + o(1)$), the query complexity exhibits a mixture of Heisenberg and classical scaling behaviors.

Contents

*This paper subsumes prior papers [CYZ25, OG26, CZY26].

†University of Pennsylvania, USA. Email: keanchen.gan@gmail.com
‡Scuola Normale Superiore, Italy. Email: filippo.girardi@sns.it
§University Mohammed VI Polytechnic, Morocco. Email: aadil.oufkir@gmail.com
¶Stony Brook University, USA. Email: nengkun.yu@stonybrook.edu
‖University of Technology Sydney, Australia. Email: iszczhang@gmail.com

1 Introduction

Quantum channel tomography (also known as quantum process tomography) asks how to reconstruct an unknown quantum process from experimental data. Specifically, given query access to a quantum channel $\mathcal{E}$, the goal is to learn a full classical description of $\mathcal{E}$ to within some prescribed error, with high probability (say, at least 2/3). This task is central to the characterization and validation of quantum devices, and it has been studied extensively for nearly three decades [CN97, PCZ97, Leu00, DP01, MRL08, KBLG18, KKEG19, BHK+19, SSKKG22, Ouf23b, Ouf23a, CWLY23, HCP23, CLO+23, FFGO23, RCEK24, VH25, FQR24, Car24, RAS+24, ZLK+24, COZ+24, WD25, CGdW25, CG25, CdH+25, WLKD25, ZRCK25, Ang25, YMM25].

A key question is the optimal query complexity of quantum channel tomography. We consider an unknown channel with input dimension d_1, output dimension d_2, and Kraus rank at most r. Several important special cases are by now well understood:

- When the input dimension $d_1 = 1$, the task reduces to *quantum state tomography.* The optimal query complexity for pure-state tomography was established in [Hay98, BM99, KW99]. The mixed-state case was settled later by Haah, Harrow, Ji, Wu, and Yu [HHJ+17] and by O'Donnell and Wright [OW16]. Further refinements and extensions appear in [OW17, GKKT20, Yue23, CHL+23, CGZ24, SSW25, PSW25, PSTW25].
- When the unknown channel is a *unitary channel* (i.e., $d_1 = d_2 = d$ and $r = 1$), Haah, Kothari, O'Donnell, and Tang [HKOT23] established that the optimal query complexity is $\Theta(d^2/\varepsilon)$, where ε is the diamond norm error. Notably, in this case, Heisenberg scaling $1/\varepsilon$ is achievable.
- When the unknown channel is an *isometry channel*, Yoshida, Miyazaki, and Murao [YMM25] established a lower bound of $\widetilde{\Omega}((d_2 - d_1)d_1/\varepsilon^2)$ for Choi trace norm error ε, where $\widetilde{\Omega}$ suppress the logarithmic factors.[1]
- When only *non-adaptive incoherent queries* are allowed, Oufkir [Ouf23b, Ouf23a] established near-optimal bounds of $\widetilde{\Theta}(d_1^3 d_2^3/\varepsilon^2)$ for full-Kraus-rank (i.e., $r = d_1 d_2$) channel tomography under diamond norm error ε, generalizing the algorithm in [SSKKG22].
- A folklore approach, based on optimal mixed-state tomography in trace norm applied to the Choi state of the unknown channel, yields a query upper bound of $O(r d_1^3 d_2/\varepsilon^2)$ for quantum channel tomography under diamond norm error.

Despite this progress, the optimal query complexity of *general* quantum channel tomography is still far from fully understood. In particular, it remains unclear how the dimension parameters d_1, d_2, r interact with the error parameter ε in the optimal query complexity. A central question is to understand when this optimal query complexity has Heisenberg scaling $1/\varepsilon$ (compared to classical scaling $1/\varepsilon^2$), while simultaneously maintaining the optimal dependence on the dimension parameters. In other words:

Question 1.1. *When and how does the optimal query complexity of quantum channel tomography exhibit Heisenberg scaling?*

[1] Since we consider the tomography with success probability at least 2/3, the lower bound in [YMM25] applies to our setting, provided that the success probability is amplified to $1 - O(\varepsilon^2)$, which incurs an additional logarithmic factor in ε.

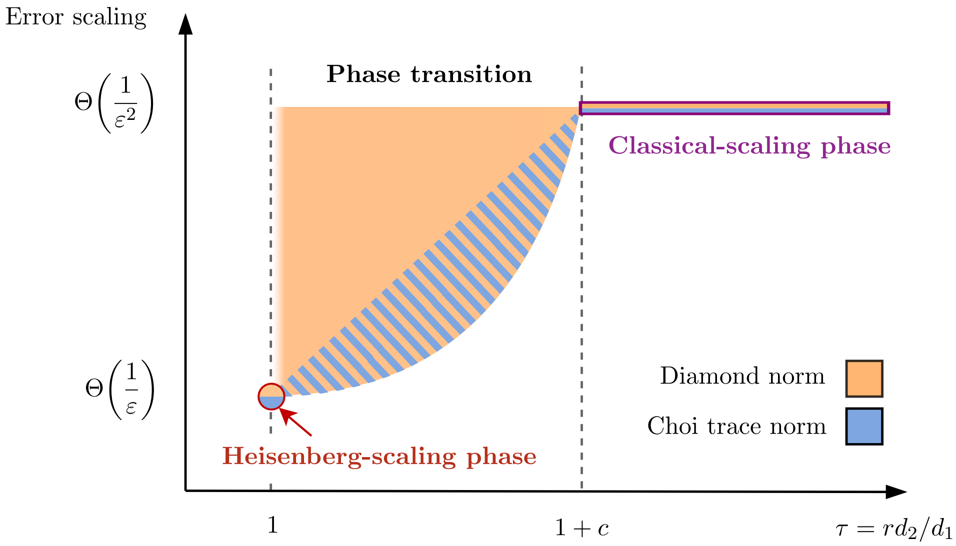


Figure 1: Heisenberg-to-classical phase transition in the query complexity of quantum channel tomography. Here, $c \in (0, 1]$ can be an arbitrary constant. For simplicity of presentation, we focus on the scaling of ε, by taking d_1 to be a constant and τ to be an independent parameter. Orange (resp. blue) regions indicate where the optimal query complexity of channel tomography should lie, according to our upper and lower bounds, under diamond norm error (resp. Choi trace norm error). The optimal query complexity exhibits Heisenberg scaling $1/\varepsilon$ in the boundary regime and classical scaling $1/\varepsilon^2$ in the away-from-boundary regime. In the near-boundary regime, it exhibits a mixture of Heisenberg and classical scalings.

1.1 Our results

To answer the above question, we prove new upper and lower bounds on the query complexity of quantum channel tomography. In particular, we identify the *dilation rate* $\tau = rd_2/d_1$ as a key parameter. Note that $\tau \geq 1$ always holds since quantum channels are trace-preserving. Depending on the value of τ, this leads to three parameter regimes: boundary regime ($\tau = 1$), near-boundary regime ($1 < \tau < 1 + o(1)$), and away-from-boundary regime ($\tau \geq 1 + \Omega(1)$). We establish multiple optimal bounds and identify when the optimal query complexity exhibits Heisenberg scaling: the dependence $1/\varepsilon$ is achievable in the boundary regime, but is no longer possible once one moves beyond the boundary regime; and in the away-from-boundary regime, only the classical scaling $1/\varepsilon^2$ is possible. An illustration of such Heisenberg-to-classical phase transition is shown in Figure 1.

In the following, we characterize in detail how the query complexity behaves across these regimes. Our results cover both the Choi trace norm[2] and the diamond norm, which are standard metrics in quantum information theory for quantifying average-case and worst-case errors, respectively.

1.1.1 Boundary to away-from-boundary: a phase transition

For boundary and away-from-boundary regimes, we establish the following query complexity bounds for quantum channel tomography and observe a Heisenberg-to-classical phase transition.

Theorem 1.2 (Boundary and away-from-boundary regimes, Theorems 3.4, 3.5, 6.1 and 6.2 restated)**.** *Consider an unknown quantum channel $\mathcal{E}$ with input dimension d_1, output dimension d_2,*

[2] By the Choi trace norm, we mean the trace norm of normalized Choi states; see Definition 2.4.

and Kraus rank at most r. Let $\tau = rd_2/d_1 \geq 1$ be the dilation rate. Then, the query complexity of the tomography of $\mathcal{E}$ to within error ε has the following bounds.

- *In the boundary regime $\tau = 1$: we establish matching upper and lower bounds of $\Theta\left(\frac{rd_1d_2}{\varepsilon}\right)$ for Choi trace norm error ε; and also establish an upper bound of $O\left(\min\left\{\frac{rd_1^{1.5}d_2}{\varepsilon}, \frac{rd_1d_2}{\varepsilon^2}\right\}\right)$ and a lower bound of $\Omega\left(\frac{rd_1d_2}{\varepsilon}\right)$ for diamond norm error ε.*
- *In the away-from-boundary regime $\tau \geq 1 + c$, for $c \in (0, 1]$ an arbitrary (but fixed) constant: we establish matching upper and lower bounds of $\Theta(\frac{rd_1d_2}{\varepsilon^2})$ for both diamond norm and Choi trace norm errors ε.*

	Boundary $\tau = 1$	Away-from-boundary $\tau \geq 1 + c$
Upper bounds	$O\left(\frac{rd_1d_2}{\varepsilon}\right)^{\dagger}$, $O\left(\min\left\{\frac{rd_1^{1.5}d_2}{\varepsilon}, \frac{rd_1d_2}{\varepsilon^2}\right\}\right)^{\ddagger}$ Theorem 3.5	$O\left(\frac{rd_1d_2}{\varepsilon^2}\right)$ Theorem 3.4
Lower bounds	$\Omega\left(\frac{rd_1d_2}{\varepsilon}\right)$ Theorem 6.1	$\Omega\left(\frac{rd_1d_2}{\varepsilon^2}\right)$ Theorem 6.2

Table 1: Boundary and away-from-boundary regimes. Here $c \in (0, 1]$ can be an arbitrary constant. All bounds hold for both diamond norm and Choi trace norm errors except: †holds for Choi trace norm error; ‡holds for diamond norm error.

Table 1 summarizes the results in Theorem 1.2. We note that Theorem 1.2 is optimal in the following senses:

- The dependence on all parameters d_1, d_2, r and ε is optimal in the away-from-boundary regime, for both diamond norm and Choi trace norm errors.
- The dependence on all parameters d_1, d_2, r and ε is optimal in the boundary regime for Choi trace norm error.
- The dependence on d_1, d_2, and r is optimal in all regimes for both constant diamond norm and Choi trace norm errors.

It also shows, from the boundary regime to the away-from-boundary regime, the optimal query complexity exhibits a phase transition from Heisenberg scaling $1/\varepsilon$ to classical scaling $1/\varepsilon^2$.

As a special case of Theorem 1.2, we consider quantum channels with equal input and output dimensions, i.e., $d_1 = d_2 = d$, which are simply called d-dimensional quantum channels. In this case, $\tau = r$ is always a positive integer and $\tau = 1$ if and only if $\mathcal{E}$ is a unitary channel. Using Theorem 1.2 in conjunction with the unitary channel tomography results by Haah, Kothari, O'Donnell, and Tang [HKOT23], we obtain a complete characterization of the query complexity for tomography of d-dimensional quantum channels.

Corollary 1.3 (Tomography of d-dimensional quantum channels)**.** *The query complexity of tomography of d-dimensional quantum channels with Kraus rank at most r, and to within either diamond norm error ε or Choi trace norm error ε, is*

$$\Theta\left(\frac{rd^2}{\varepsilon^{\min\{r,2\}}}\right).$$

As another illuminating and well-studied special case of Theorem 1.2, we consider tomography of quantum channels with input dimension $d_1 = 1$, which reduces to quantum state tomography. In this regime, our result recovers the recent optimal sample lower bound [SSW25] (without logarithmic factors) for quantum state tomography, matching the upper bound established in [OW16].

Corollary 1.4 (Optimal lower bound for state tomography)**.** *Tomography of a d-dimensional mixed state with rank at most r, to within trace norm error ε, requires $\Omega\big(\frac{dr}{\varepsilon^2}\big)$ samples.*

Our proof of this lower bound proceeds via a different line of reasoning than that in [SSW25] and may offer complementary insights.

The following result underpins our upper bounds in Theorem 1.2 by establishing a connection between the query complexities in two different access models: given access to an unknown channel itself, or given access to one of its Stinespring dilations. We say a quantum algorithm is a tester if it makes queries to an unknown quantum channel and outputs a classical outcome. A tester is a parallel tester if its queries to the unknown channel can be made in parallel. Formal definitions can be found in Section 2.3.

Theorem 1.5 (Dilation does not help for parallel testers, Theorem 3.3 restated)**.** *If there exists a parallel (possibly coherent) tester that solves a channel estimation task using n queries to an arbitrary dilation of an unknown quantum channel $\mathcal{E}$, then there exists a parallel tester that solves this task using n queries to $\mathcal{E}$ itself.*

We note that Theorem 1.5 partially answers a conjecture from Tang, Wright, and Zhandry [TWZ25], which asserts that access to channel dilations does not help. In particular, if we can extend the statement in Theorem 1.2 to general (sequential) testers, then this may be viewed as a dual version of the conjecture in [TWZ25], in the Heisenberg and Schrödinger pictures, respectively. Further discussion can be found in Section 1.4.

Moreover, we also obtain the following Heisenberg-scaling query upper bound for quantum state tomography with state-preparation channels, by applying Theorem 1.5 to the pure-state tomography algorithm due to Chen [Che25], which achieves the Heisenberg scaling using parallel queries without inverses.

Corollary 1.6 (State tomography with state-preparation channels, Theorem 3.6 restated)**.** *When $\tau = 1$, tomography of the mixed state $\mathcal{E}(|0\rangle\langle 0|)$ to within trace norm error ε can be done using $O\Big(\min\Big\{\frac{d_1^{1.5}}{\varepsilon}, \frac{d_1}{\varepsilon^2}\Big\}\Big)$ queries to $\mathcal{E}$.*

1.1.2 Near-boundary: new upper and lower bounds

In the near-boundary regime, we prove new upper and lower bounds on the query complexity of quantum channel tomography, which exhibits a mixture of Heisenberg and classical scalings.

Theorem 1.7 (Near-boundary regime, Theorems 3.4, 3.7, 6.1 and 6.3 and Corollary 6.4 restated)**.** *Consider an unknown quantum channel $\mathcal{E}$ with input dimension d_1, output dimension d_2, and Kraus rank at most r. Let $\tau = rd_2/d_1 \geq 1$ be the dilation rate. Then, for $\tau \in (1, \frac{4}{3})$ where τ can be arbitrarily close to 1, the query complexity of the tomography of $\mathcal{E}$ to within error ε has the following bounds.*

- *For Choi trace norm error ε: we establish an upper bound of $O\Big(\frac{d_1^2}{\varepsilon} + \frac{(\tau-1)d_1^2}{\varepsilon^2}\Big)$ and a lower bound of $\Omega\Big(\frac{d_1^2}{\varepsilon} + \frac{(\tau-1)^2}{\varepsilon^2}\Big)$.*

- *For diamond norm error ε: we establish an upper bound of $O\left(\frac{d_1^2}{\varepsilon^2}\right)$ and a lower bound of $\Omega\left(\frac{d_1^2}{\varepsilon} + \frac{(\tau-1)^2 d_1^2}{\varepsilon^2}\right)$.*

	Choi trace norm	Diamond norm
Upper bounds	$O\left(\frac{d_1^2}{\varepsilon} + \frac{(\tau-1)d_1^2}{\varepsilon^2}\right)$ Theorem 3.7	$O\left(\frac{d_1^2}{\varepsilon^2}\right)$ Theorem 3.4
Lower bounds	$\Omega\left(\frac{d_1^2}{\varepsilon} + \frac{(\tau-1)^2}{\varepsilon^2}\right)$ Theorem 6.1 and Corollary 6.4	$\Omega\left(\frac{d_1^2}{\varepsilon} + \frac{(\tau-1)^2 d_1^2}{\varepsilon^2}\right)$ Theorems 6.1 and 6.3

Table 2: Near-boundary regime, i.e., $\tau \in (1, \frac{4}{3})$ can be arbitrarily close to 1.

Table 2 summarizes the results in Theorem 1.7. We note that the $1/\varepsilon^2$ terms in the lower bounds in the near-boundary regimes imply that the query complexity of quantum channel tomography no longer exhibits Heisenberg scaling $1/\varepsilon$ once it leaves the boundary regime, i.e., even for τ arbitrarily close to 1, as long as it is independent of ε. This transition was illustrated in Figure 1.

1.2 Overview of techniques

Upper bounds. Our upper bounds for quantum channel tomography are directly inspired by the recent work of Pelecanos, Spilecki, Tang, and Wright [PSTW25], who showed that mixed-state tomography can be reduced to pure-state tomography while still achieving optimal performance. In this paper, we prove in Theorem 3.1 that any channel tomography algorithm that can make parallel queries to a Stinespring dilation of an unknown channel $\mathcal{E}$ can be faithfully simulated by an algorithm that queries only $\mathcal{E}$ itself. This enables a reduction from general quantum channel tomography to isometry channel tomography, which is a more tractable problem. As a consequence, by combining this reduction with existing parallel-query algorithms for isometry tomography [YRC20, YMM25] and with a generalization of the base tomography algorithm from [HKOT23] (see Section 8), we obtain the upper bounds presented in this paper.

Theorem 3.1 is proved by constructing a "local" tester (see Equations (9) and (10)) that can faithfully simulate the "global" tester with access to a random Stinespring dilation, using representation-theoretic tools together with the quantum comb formalism [CDP09, BMQ22]. This result can be viewed as a generalization of the earlier result by [CWZ24], which studies the local test for quantum states and shows that access to purifications does not help for mixed-state testing. Related results trace back to [SW22, Theorem 35], and were recently strengthened in an algorithmic sense by [TWZ25], which explicitly constructs an algorithm for generating random purifications of a mixed state. Intuitively, local test and random purification can be interpreted as dual concepts in the Heisenberg and Schrödinger pictures, respectively. This line of work is further extended in [GML25, WW25, MGC+25, GMZ+25, YNM25]. In particular, [GMZ+25, YNM25] explicitly construct random Stinespring dilation superchannels that convert parallel queries to a quantum channel into parallel queries to its random Stinespring dilation. These results can also be viewed as dual to our local test techniques with respect to the Schrödinger and Heisenberg pictures. Moreover, from an algorithmic sense, they additionally provide explicit circuit implementations of this conversion.

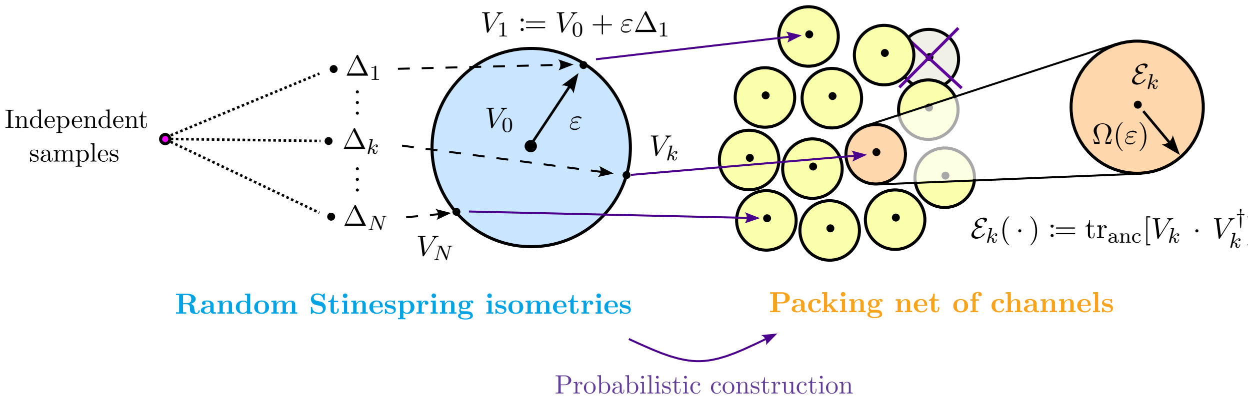


Figure 2: A pictorial representation of our constructions of channel packing nets.

Lower bounds. To prove our lower bounds, we proceed using a two-stage method. The first stage consists of constructing sufficiently large packing nets of quantum channels with specific structures and desired properties. The second stage consists of proving lower bounds on the query complexity of algorithms that can learn, and thus discriminate between, the quantum channels in the packing nets.

Existing results on nets of the set of isometries [Sza97] lead to packing nets of the set of quantum channels of input dimension d_1, output dimension d_2 and Kraus rank r of cardinality satisfying $\ln|\mathcal{N}| = \Omega(rd_1d_2)$. This along with a polynomial method can be used to obtain the general lower bound $\Omega(rd_1d_2)$ [GMZ+25], which is tight in the dimensions and the Kraus rank but does not capture the dependency on error parameter ε. To make the dependence on ε precise and optimal, we opt for carefully designed packing nets and a new proof strategy. We distinguish between the boundary and non-boundary regimes.

Roughly speaking, our constructions in the non-boundary regimes share the following idea (see Figure 2). To ensure complete-positivity, we construct the channels via their Stinespring isometries. The isometries are chosen of the following form

$$V = V_0 + \varepsilon\Delta, \tag{1}$$

where V_0 corresponds to a center map and Δ is a noisy perturbation. The center supermap $\mathcal{V}_0 = V_0(\cdot)V_0^\dagger$ admits Kraus operators that are orthogonal and have almost the same Hilbert–Schmidt norm. The noisy perturbation is constructed using the Haar randomness applied on certain output space of Δ. The concentration phenomenon in high-dimensions permits us to prove that two independent channels constructed in this way will be ε-far apart in the Choi trace norm (or diamond norm) with a probability that is at least $1-\exp(-Crd_1d_2)$ with C a universal constant. A probabilistic argument then concludes existence of the packing net of cardinality satisfying $\ln|\mathcal{N}| = \Omega(rd_1d_2)$. A crucial feature of our construction that leads to classical scaling is that the center and perturbation components have orthogonal images. This can give insights to the phase transition behavior, since at the boundary such construction with the orthogonality feature is not possible.

In the boundary regime, the construction has a similar form as in Equation (1). The center map is close to identity. The perturbation, unlike in the non-boundary regime, is a Haar conjugated anti-Hermitian traceless diagonal matrix with entries $\pm\mathrm{i}$. This construction is inspired by Paninski distribution [Pan08] that was used to prove optimality of the mixedness testing problem. As stated before, this construction does not exhibit the orthogonality feature. Nevertheless, it has the property that the center and perturbation matrices are orthogonal in Hilbert-Schmidt inner product. This is sufficient to establish the Heisenberg-scaling lower bound.

We now proceed to the second stage of the lower bound proof. We study two structured families of isometries, referred to as type I and type II instances, corresponding to the boundary and non-boundary regimes, respectively. Both families take the perturbative form in Equation (1), but with different choices of V_0 and Δ. Then, we show that discriminating a set of isometries of type I or type II requires at least $\Omega(f/\varepsilon)$ or $\Omega(f/\varepsilon^2)$ queries, respectively, where f depends on both the cardinality of the set and the dimensions of the isometries. By construction, the channels in the packing nets described above admit Stinespring dilations that are exactly of type I or type II. It follows that the hardness of isometry discrimination transfers directly to channel discrimination: any algorithm that can distinguish among the channels in the packing net can also distinguish among the corresponding dilation isometries, simply by discarding the ancilla system. Since a tomography algorithm with sufficiently high precision can solve the discrimination problem, these query lower bounds also apply to quantum channel tomography.

We now elaborate on the hardness of the isometry discrimination problem. We formulate the task of discriminating among a set $\mathcal{N}$ of isometries within the quantum comb framework [CDP08, CDP09, BMQ22]. In this framework, any discrimination algorithm making n adaptive and coherent queries to an unknown channel can be represented by a sequential tester $\{T_V\}_{V\in\mathcal{N}}$, where $\sum_{V\in\mathcal{N}} T_V$ forms an $(n+1)$-comb with 1-dimensional input and output spaces. Then, the success probability of the discrimination problem is $\frac{1}{|\mathcal{N}|}\sum_{V\in\mathcal{N}} T_V \star |V\rangle\!\rangle\langle\!\langle V|^{\otimes n}$, where $\star$ denotes the link product (see Definition 2.8), representing the contraction between combs and $|V\rangle\!\rangle\langle\!\langle V|$ is the Choi operator of the isometry channel $\mathcal{V}$. The technical heart of the proof is a decomposition of $|V\rangle\!\rangle^{\otimes n}$ into orthogonal sectors indexed by either where (for type I isometry V) or how many (for type II isometry V) perturbative components (i.e., Δ) appear. By collecting the Haar average over the orbit of the perturbation for each sector, we can construct an operator $\lambda\cdot\Gamma$ that can universally upper bound (w.r.t. the Löwner order) $|V\rangle\!\rangle\langle\!\langle V|^{\otimes n}$ for all $V\in\mathcal{N}$, where $\lambda>0$ is a scalar depending on n, d, ε and Γ is an n-comb. Since $\sum_{V\in\mathcal{N}} T_V$ is an $(n+1)$-comb, its contraction with the n-comb Γ evaluates to 1. It follows that the success probability, which must be at least $2/3$, is upper bounded by

$$\frac{1}{|\mathcal{N}|}\sum_{V\in\mathcal{N}} T_V \star (\lambda\cdot\Gamma) = \frac{\lambda}{|\mathcal{N}|}.$$

Solving the inequality $2/3 \le \lambda/|\mathcal{N}|$ yields the desired query lower bounds.

We remark that, when the orthogonal decomposition of $|V\rangle\!\rangle^{\otimes n}$ is based on how many perturbative components appear (the type II case), the property that Γ is a comb relies on the orthogonality between the image of the perturbation Δ and that of the center map V_0. However, as mentioned above, this is not possible in the boundary regime. Thus for the type I isometries, the decomposition is based on where perturbative components appear, which is less "compact" and ultimately yields a Heisenberg-scaling lower bound of $\Omega(1/\varepsilon)$ rather than the classical-scaling bound $\Omega(1/\varepsilon^2)$.

1.3 Related work

Quantum channel tomography. The problem of quantum channel tomography (also known as quantum process tomography) has been studied under a variety of models and error metrics. Early works [CN97, PCZ97] consider protocols that prepare input states from a complete set of basis states, perform state tomography on the corresponding output states, and reconstruct the channel using an inversion protocol. Later, [Leu00, DP01] use the Choi-Jamiołkowski isomorphism [Cho75, Jam72] to reduce channel tomography to Choi state tomography. These approaches, together with the direct scheme of [ML06], are compared in terms of resource requirements in [MRL08].

Subsequent works [Kah07, KKEG19, SSKKG22, HKOT23, Ouf23b, RAS$^+$24, GMZ$^+$25] focus on obtaining rigorous bounds on the query complexity with explicit dependence on the estimation error.

For example, [KKEG19, SSKKG22] propose randomized protocols for channel tomography and establish query complexity bounds under the Choi-Frobenius norm and the Choi trace norm errors. One of the protocols in [SSKKG22] is generalized by [Ouf23b], which shows that it achieves near-optimal query complexity for full-Kraus-rank channel tomography under the diamond norm error in the non-adaptive single-copy setting. Complementing these upper bounds, [RAS+24] establishes a query lower bound of $\Omega(d_1^2 d_2^2/\log(d_1 d_2))$ for full-Kraus-rank (i.e., $r = d_1 d_2$) channel tomography under constant Choi trace norm error. Moreover, [GMZ+25] proves a query lower bound of $\Omega(rd_1d_2)$ for channel tomography under constant diamond norm error.

Unitary channel tomography. Unitary channel tomography is a special case of general channel tomography where the input and output dimensions are equal and the Kraus rank is 1. It is studied under different assumptions, models, and error metrics [AJV01, PS02, BBMT04, CDPS04, Hay06, CDS05, Kah07, YRC20, HKOT23, GL25, YMM25]. In particular, [YRC20] and [YMM25] study tomography of unitary and isometry channels under average-case channel fidelity; [HKOT23] establishes the optimal query complexity $\Theta(d^2/\varepsilon)$ for unitary channel tomography under the diamond norm. Moreover, the lower bound in [YMM25] with the upper bound in [HKOT23] establishes that pure Heisenberg scaling is impossible for isometry channel tomography, except in the unitary case.

On the upper bound side, our result is based on a reduction from general channel tomography to isometry channel tomography, building on the base unitary tomography algorithm in [HKOT23] that makes parallel queries. Through a reduction to the algorithm in [YRC20], we further show that Heisenberg scaling can be achieved even for non-unitary channels. We also leverage a reduction to the algorithm in [YMM25] to obtain upper bounds for channel tomography under the Choi trace norm error in the near-boundary regime. On the lower bound side, new packing-net constructions are needed to capture the full dependence on the channel dimensions, Kraus rank, and error. In contrast to [HKOT23], our packing nets have additional structure and leverage symmetry arguments.

Quantum state tomography. Quantum state tomography (or more generally, quantum state learning [AdW17, AA23]) is a special case of quantum channel tomography where the input dimension is 1. The optimal sample complexity of d-dimensional rank-r quantum state tomography under the trace norm error is established as $\tilde{\Theta}(dr/\varepsilon^2)$ in [OW16, HHJ+17], and the logarithmic factor is removed only recently by [SSW25]. As noted above, quantum state tomography algorithms can be used to perform channel tomography via various reductions. In particular, pure state tomography [CL14, KRT17, GKKT20] is a key ingredient in the optimal algorithm for unitary channel tomography in [HKOT23]. Moreover, using a different proof strategy, our results recover the upper bound of [OW16, HHJ+17] and the recently established lower bound of [SSW25] as special cases.

Independent work. Mele and Bittel [MB25] established the same upper bound $O(rd_1d_2/\varepsilon^2)$ for quantum channel tomography under diamond norm error, which is concurrent and independent to our Theorem 3.4. They also derived an explicit and non-trivial dependence on the failure probability. We note that their method and ours differ substantially. Specifically, they obtain the upper bound through an analysis of the tomography of Choi states, which leads to an explicit tomography algorithm. In contrast, our approach uses the local test technique to simulate access to dilations of quantum channels, allowing us to reduce general channel tomography to a more tractable problem of isometry channel tomography. Moreover, we establish additional upper bounds achieving Heisenberg scaling $1/\varepsilon$ for (possibly non-unitary) channels in the boundary regime $\tau = 1$.

They also established a query lower bound. When the failure probability is constant, their lower bound is $\Omega\left(\frac{rd_1d_2}{\varepsilon^\beta} + \frac{m\log(3m)}{\varepsilon^2}\right)$ for diamond norm error ε, where $\beta = \frac{2rd_1d_2 - d_1^2 - r^2}{2(rd_1d_2-1)}$ and

$$m = \max\left\{\min\left\{d_1, \left\lfloor \frac{r - \lceil d_1/d_2 \rceil}{2} \right\rfloor\right\},\ \mathbb{1}_{r \geq \lceil d_1/d_2 \rceil + 1},\ d_1 \cdot \mathbb{1}_{r \geq 2d_1}\right\}.$$

We can see that $\beta \in (\frac{3}{8}, 1]$ and $m \in [0, d_1]$. Moreover, when $r \geq \lceil d_1/d_2 \rceil + 1$,[3] it follows that $m \geq 1$, thereby showing that pure Heisenberg scaling is unattainable for any quantum channel whose Kraus rank is not minimal. However, when d_1 is not a multiple of d_2 and $r = \lceil d_1/d_2 \rceil$ is minimal, their lower bound does not rule out Heisenberg scaling, although this case still lies in the non-boundary regime ($\tau > 1$). By comparison, our lower bounds rule out pure Heisenberg scaling across the entire non-boundary regime, and exhibit sharper parameter dependence in all regimes (in particular, our lower bounds are optimal in the away-from-boundary regime).

1.4 Discussion

In this paper, we study the query complexity of quantum channel tomography. Our results answer Question 1.1 as follows. We identify the dilation rate $\tau = rd_2/d_1$ as a key parameter. We show that when $\tau = 1$, the optimal query complexity exhibits Heisenberg scaling $1/\varepsilon$; whereas for $\tau > 1$, it no longer has Heisenberg scaling due to our lower bounds. Moreover, we establish multiple optimal bounds in the boundary regime ($\tau = 1$) and the away-from-boundary regime ($\tau \geq 1 + \Omega(1)$), as well as new upper and lower bounds in the near-boundary regime ($1 < \tau < 1 + o(1)$). We conclude by highlighting several directions for future work.

A natural open problem is to determine the optimal query complexity of quantum channel tomography across all parameter regimes. This includes sharpening the upper and lower bounds in the near-boundary regime, as well as improving the bounds in the boundary regime for diamond norm error.

A second question, closely related to the first, is whether Theorem 1.5 can be extended to sequential testers. More precisely, suppose a tester that makes an arbitrary sequence of queries, not necessarily in parallel, to an arbitrary dilation of an unknown quantum channel $\mathcal{E}$. Can one then construct another tester that makes the same number of queries to $\mathcal{E}$ itself and solves the same estimation task? This may be viewed as a dual formulation, in the Heisenberg and Schrödinger pictures respectively, of a conjecture of Tang, Wright, and Zhandry [TWZ25, Conjecture 1.8]. A positive answer would lead to progress on the first question: by applying such an extended "dilation does not help" theorem to the bootstrapped tomography algorithm proposed by Haah, Kothari, O'Donnell, and Tang [HKOT23], one could obtain a matching upper bound on the query complexity of quantum channel tomography in the boundary regime for diamond norm error.

2 Preliminaries

Notation. Given a Hilbert space $\mathcal{H}$, we denote by $\mathcal{L}(\mathcal{H})$ the set of linear operators on $\mathcal{H}$. Similarly, we denote by $\mathcal{L}(\mathcal{H}_0, \mathcal{H}_1)$ the set of linear operators from $\mathcal{H}_0$ to $\mathcal{H}_1$. Given two orthonormal bases for $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively, we can represent every linear operator in $\mathcal{L}(\mathcal{H}_0, \mathcal{H}_1)$ by a $\dim(\mathcal{H}_1) \times \dim(\mathcal{H}_0)$ matrix, and, for such a matrix X, we use $|X\rangle\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_0$ to denote the vector obtained by flattening the matrix X. It is easy to see the following properties:

$$||\psi\rangle\langle\phi|\rangle\rangle = |\psi\rangle|\phi^*\rangle, \qquad |XYZ\rangle\rangle = X \otimes Z^{\mathrm{T}}|Y\rangle\rangle,$$

[3] That is, just above the minimal Kraus rank $\lceil d_1/d_2 \rceil$ that a valid quantum channel can have, given the input dimension d_1 and output dimension d_2.

Symbol	Description	Introduced in
$C_{\mathcal{E}}$	Choi-Jamiołkowski operator of the channel $\mathcal{E}$, (i.e., unnormalized Choi-Jamiołkowski state).	Equation (2)
$\mathbf{QChan}^r_{d_1,d_2}$	Set of all quantum channels $\mathcal{E} : \mathcal{L}(\mathbb{C}^{d_1}) \to \mathcal{L}(\mathbb{C}^{d_2})$ that have Kraus rank at most r.	Notation 2.1
$\mathbf{ISO}_{d_1,d_2}$	Set of isometry channels with input dimension d_1 and output dimension d_2, which is equivalent to $\mathbf{QChan}^1_{d_1,d_2}$	*Ibid.*
$\mathbf{Dilation}_r(\mathcal{E})$	Set of all dilations of the quantum channel $\mathcal{E}$ with an ancilla system of dimension r.	Notation 2.2
$\mathbf{Contract}_{\mathcal{H}_3}(\mathcal{V})$	Quantum channel defined as the contration $\rho \mapsto \mathrm{tr}_{\mathcal{H}_3}(V\rho V^\dagger)$, where $V : \mathcal{H}_1 \to \mathcal{H}_2 \otimes \mathcal{H}_3$ is an isometry.	*Ibid.*
$\mathcal{V} \sim \mathbf{Dilation}_r(\mathcal{E})$	The dilation $\mathcal{V}$ is sampled from Haar distributions on $\mathbf{Dilation}_r(\mathcal{E})$.	Notation 2.3
$U \sim \mathbb{U}_d$	The unitary U is sampled from Haar distributions on $\mathbb{U}_d$.	*Ibid.*
$X \star Y$	Link product between X and Y.	Definition 2.8
$\{T_i\}_i$	Quantum channel tester.	Section 2.3
$\{A_{\mathcal{E}}\}_{\mathcal{E}\in\mathbf{QChan}^r_{d_1,d_2}}$	Estimation task of quantum channels, i.e., set of classical outcomes that are regarded as correct answers when the unknown channel is $\mathcal{E}$.	Definition 3.2

Table 3: Summary of the notation used in the paper.

where $|\phi^*\rangle$ denotes the entry-wise complex conjugate of $|\phi\rangle$ and Z^{T} denotes the transpose of the matrix Z with respect to a given orthonormal basis. The inner product can be rewritten as $\langle\langle X|Y\rangle\rangle = \mathrm{tr}(X^\dagger Y)$. For two Hermitian operators X, Y, we write $X \leq Y$ to denote that $Y - X$ is positive semidefinite.

Let n, m, d be positive integers such that $n \geq m$. Let $\mathcal{H}_1 \cong \cdots \cong \mathcal{H}_n \cong \mathbb{C}^d$ be n copies of a d-dimensional Hilbert space $\mathcal{H}$. Let $S \subseteq [n] = \{1, 2, \ldots, n\}$ be a set of integers and let $|\psi\rangle \in \mathbb{C}^d$ be a state. We denote by

$$|\psi\rangle^{\otimes S}$$

the state $|\psi\rangle^{\otimes|S|} \in \bigotimes_{i\in S} \mathcal{H}_i$. Therefore, if $|\varphi\rangle \in \mathbb{C}^d$ is another state, then we call

$$|\psi\rangle^{\otimes S} \otimes |\varphi\rangle^{\otimes [n]\setminus S}$$

the state $\bigotimes_{i=1}^n |x_i\rangle \in \bigotimes_{i=1}^n \mathcal{H}_i$, where $|x_i\rangle = |\psi\rangle$ for $i \in S$, and $|x_i\rangle = |\varphi\rangle$ for $i \notin S$.

2.1 Quantum channels

A *quantum channel* with input dimension d_1 and output dimension d_2 is a linear, completely positive and trace-preserving map $\mathcal{E} : \mathcal{L}(\mathbb{C}^{d_1}) \to \mathcal{L}(\mathbb{C}^{d_2})$ (see, e.g., [NC10, Wat18, Hay17]), which is also called a CPTP map.

Any quantum channel $\mathcal{E}$, in the Kraus representation [Kar83], can be written as

$$\mathcal{E}(\rho) = \sum_{i=1}^{r} E_i \rho E_i^\dagger,$$

where $E_i : \mathbb{C}^{d_1} \to \mathbb{C}^{d_2}$ are non-zero linear operators satisfying

$$\sum_{i=1}^{r} E_i^\dagger E_i = I_{d_1} \qquad \text{and} \qquad \mathrm{tr}(E_i^\dagger E_j) = 0 \quad \text{for} \quad i \neq j.$$

The integer r is called *Kraus rank* and $\{E_i\}$ are called *Kraus operators*. Since, for each i, $E_i^\dagger E_i$ has rank at most d_2, and since these operators sum to I_{d_1}, one can easily see that $rd_2 \geq d_1$. In particular, a quantum channel having Kraus rank $r = 1$ is an isometry $\mathcal{V}(\,\cdot\,) = V \,\cdot\, V^\dagger$, where $V : \mathbb{C}^{d_1} \to \mathbb{C}^{d_2}$ is an isometry operator – i.e. it satisfies $V^\dagger V = I_{d_1}$ – and it must hold that $d_2 \geq d_1$.

Notation 2.1. *We denote by* $\mathbf{QChan}^r_{d_1,d_2}$ *the set of all quantum channels* $\mathcal{E} : \mathcal{L}(\mathbb{C}^{d_1}) \to \mathcal{L}(\mathbb{C}^{d_2})$ *having Kraus rank at most* r*. In particular, we denote by* $\mathbf{ISO}_{d_1,d_2}$ *the set of all isometry channels with input dimension* d_1 *and output dimension* d_2*, which is corresponds to* $\mathbf{QChan}^1_{d_1,d_2}$*.*

In the Choi-Jamiołkowski representation [Cho75, Jam72], any channel $\mathcal{E}$ can be identified by its Choi-Jamiołkowski operator as follows:

$$C_{\mathcal{E}} = (\mathcal{E} \otimes \mathcal{I})(|I\rangle\!\rangle\langle\!\langle I|) \in \mathcal{L}(\mathbb{C}^{d_2} \otimes \mathbb{C}^{d_1}), \tag{2}$$

where we have denoted by $|I\rangle\!\rangle = \sum_i |i\rangle|i\rangle \in \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_1}$ an unnormalized maximally entangled state. We may simply call $C_{\mathcal{E}}$ Choi operator or (unnormalized) Choi state. Note that it it possible to write the Choi operator as $C_{\mathcal{E}} = \sum_{i=1}^r |E_i\rangle\!\rangle\langle\!\langle E_i|$, where E_i are orthogonal Kraus operators and thus $|E_i\rangle\!\rangle$ are pairwise orthogonal vectors. Hence, the Kraus rank is equal to the rank of the Choi operator. As a consequence, one can easily see that $r \leq d_1 d_2$.

Stinespring dilation. Given a quantum channel $\mathcal{E}$ with Kraus operators $\{E_i\}_{i=1}^r$, using its Stinespring dilation [Sti55], we can also write $\mathcal{E}$ as

$$\mathcal{E}(\rho) = \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(V \rho V^\dagger), \tag{3}$$

where $\mathcal{H}_{\mathrm{anc}} \cong \mathbb{C}^r$ is an ancilla system and $V = \sum_{i=1}^r |i\rangle_{\mathrm{anc}} \otimes E_i$ is an isometry operator. Any isometry channel $\mathcal{V}(\,\cdot\,) = V \,\cdot\, V^\dagger$ satisfying Equation (3) is called a dilation of $\mathcal{E}$. Two isometries $\mathcal{V}_1$ and $\mathcal{V}_2$ are two dilations of the same channel $\mathcal{E}$ if and only if they differ by a unitary on $\mathcal{H}_{\mathrm{anc}}$, namely, $V_2 = (U \otimes I_{d_2})V_1$ for a unitary $U : \mathcal{H}_{\mathrm{anc}} \to \mathcal{H}_{\mathrm{anc}}$.

Notation 2.2. *Given a quantum channel* $\mathcal{E}$ *with Kraus rank at most* r*, we denote by* $\mathbf{Dilation}_r(\mathcal{E})$ *the set of all dilations of* $\mathcal{E}$ *with an ancilla system of dimension* r*. Given an isometry channel* $\mathcal{V} : \mathcal{L}(\mathcal{H}_1) \to \mathcal{L}(\mathcal{H}_2 \otimes \mathcal{H}_3)$*, we denote by* $\mathbf{Contract}_{\mathcal{H}_3}(\mathcal{V})$ *the quantum channel*

$$\rho \mapsto \mathrm{tr}_{\mathcal{H}_3}(V \rho V^\dagger).$$

Haar distribution. Given any arbitrary quantum channel $\mathcal{E} \in \mathbf{QChan}^r_{d_1,d_2}$, we define the Haar distribution on $\mathbf{Dilation}_r(\mathcal{E})$ as follows: choose an arbitrary dilation $\mathcal{V} \in \mathbf{Dilation}_r(\mathcal{E})$, and output $(\mathcal{U} \otimes \mathcal{I}_{d_2}) \circ \mathcal{V}$, where $U \in \mathbb{U}_r$ is a Haar random unitary. This procedure is well defined since the resulting distribution does not depend on the choice of the initial dilation $\mathcal{V}$. It is easy to see that such distribution is invariant under $\mathbb{U}_r$, i.e.

$$\mathbb{P}[A] = \mathbb{P}[\{(\mathcal{U} \otimes \mathcal{I}_{d_2}) \circ \mathcal{V} \,|\, \mathcal{V} \in A\}],$$

for any unitary $U \in \mathbb{U}_r$ and any measurable set $A \subseteq \mathbf{Dilation}_r(\mathcal{E})$.

Notation 2.3. *We denote by* $\mathcal{V} \sim \mathbf{Dilation}_r(\mathcal{E})$ *and* $U \sim \mathbb{U}_d$ *two random variables* $\mathcal{V}$ *and* U *that are sampled from the Haar distributions on* $\mathbf{Dilation}_r(\mathcal{E})$ *and* $\mathbb{U}_d$*, respectively.*

Distance measures. In this paper, we are going to use the Choi trace norm and the diamond norm [AKN98, Wat18] as the *average-case* and *worst-case* measures between quantum channels, respectively.

Definition 2.4. *Let $\mathcal{E}, \mathcal{F} \in \mathbf{QChan}^r_{d_1,d_2}$ be two quantum channels. We define the Choi trace norm between $\mathcal{E}$ and $\mathcal{F}$ as*

$$\frac{1}{d_1}\|C_{\mathcal{E}} - C_{\mathcal{F}}\|_1,$$

where $\|\cdot\|_1$ denotes the Schatten 1-norm (i.e. trace norm) and $C_{\mathcal{E}}$ is the (unnormalized) Choi state of $\mathcal{E}$. We define the diamond norm between $\mathcal{E}$ and $\mathcal{F}$ as

$$\|\mathcal{E} - \mathcal{F}\|_\diamond := \sup_\rho \|(\mathcal{E} \otimes \mathcal{I}_{d_1})(\rho) - (\mathcal{F} \otimes \mathcal{I}_{d_1})(\rho)\|_1,$$

where the supremum is taken over all quantum states $\rho \in \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_1}$.

A simple relation between the diamond norm of quantum channels and trace norm of the corresponding Choi-Jamiołkowski operators is the following: given two quantum channels $\mathcal{E}, \mathcal{F} \in \mathbf{QChan}^r_{d_1,d_2}$, we have (see, e.g., [KR21, Proposition 50]):

$$\frac{1}{d_1}\|C_{\mathcal{E}} - C_{\mathcal{F}}\|_1 \le \|\mathcal{E} - \mathcal{F}\|_\diamond \le \|C_{\mathcal{E}} - C_{\mathcal{F}}\|_1. \tag{4}$$

As an average-case measure, the Choi trace norm is widely used in learning of quantum channels [KKEG19, SSKKG22] and also in other tasks in quantum information theory [MdW16, KR21, RAS+24]. Moreover, we will also consider the *channel fidelity* [Rag01] (or *entanglement fidelity* [Nie02]). Specifically, given two quantum channels $\mathcal{E}, \mathcal{F} \in \mathbf{QChan}^r_{d_1,d_2}$, their channel fidelity F_{ch} is defined as the fidelity of their normalized Choi states:

$$\mathrm{F}_{\mathrm{ch}}(\mathcal{E}, \mathcal{F}) := \mathrm{F}\left(\frac{C_{\mathcal{E}}}{d_1}, \frac{C_{\mathcal{F}}}{d_1}\right),$$

where F denotes the fidelity between quantum states. Note that channel fidelity is closely related to the Choi trace norm by a straightforward lifting of the relation between fidelity and trace norm for quantum states.

2.2 Quantum combs

The framework of quantum combs [CDP08, CDP09] provide a powerful description of higher-order transformations of quantum processes. More specifically, the Choi-Jamiołkowski representation of quantum channels, which describes transformations of quantum states, can be extended to the higher-order setting of transformations of quantum processes. Such higher-order objects are called *quantum combs*. More precisely, deterministic and probabilistic quantum combs are defined as follows.

Definition 2.5 (Deterministic comb [CDP09])**.** *Let $n \ge 1$ be an integer. A deterministic n-comb, defined on a sequence of $2n$ Hilbert spaces $(\mathcal{H}_0, \mathcal{H}_1, \ldots, \mathcal{H}_{2n-1})$, is a positive semidefinite operator X on $\bigotimes_{j=0}^{2n-1} \mathcal{H}_j$ such that there exists a sequence of operators $X^{(n)}, X^{(n-1)}, \ldots, X^{(1)}, X^{(0)}$ satisfying*

$$\mathrm{tr}_{\mathcal{H}_{2j-1}}\left(X^{(j)}\right) = I_{\mathcal{H}_{2j-2}} \otimes X^{(j-1)}, \quad 1 \le j \le n, \tag{5}$$

where $X^{(n)} = X$ and $X^{(0)} = 1$.

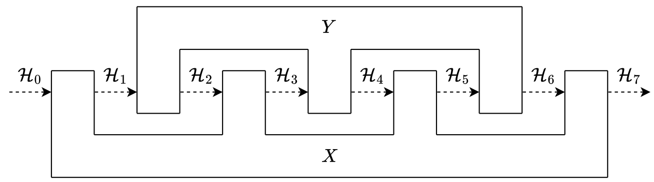


Figure 3: The combination of a 4-comb X with a 3-comb Y, yielding a 1-comb $X \star Y$ on $(\mathcal{H}_0, \mathcal{H}_7)$.

Definition 2.6 (Probabilistic comb [CDP09]). *Let $n \geq 1$ be an integer. A probabilistic n-comb, defined on a sequence of $2n$ Hilbert spaces $(\mathcal{H}_0, \mathcal{H}_1, \ldots, \mathcal{H}_{2n-1})$, is a positive semidefinite operator X on $\bigotimes_{j=0}^{2n-1} \mathcal{H}_j$ such that $X \leq Y$ for some deterministic n-comb Y on $(\mathcal{H}_0, \mathcal{H}_1, \ldots, \mathcal{H}_{2n-1})$.*

Remark 2.7. *In this paper, quantum comb refers to a deterministic comb by default, and the term probabilistic comb will be used explicitly when needed.*

It is easy to see verify following facts:

- a quantum 1-comb is the Choi-Jamiołkowski operator of a quantum channel;
- any convex combination of quantum n-combs is also a quantum n-comb.

Now, let us introduce the link product "$\star$".

Definition 2.8 (Link product "$\star$" [CDP08, CDP09]). *Let X be a linear operator on $\mathcal{H}_{\boldsymbol{i}} = \mathcal{H}_{i_1} \otimes \mathcal{H}_{i_2} \otimes \cdots \otimes \mathcal{H}_{i_n}$ and let Y be a linear operator on $\mathcal{H}_{\boldsymbol{j}} = \mathcal{H}_{j_1} \otimes \mathcal{H}_{j_2} \otimes \cdots \otimes \mathcal{H}_{j_m}$, where $\boldsymbol{i} = (i_1, \ldots, i_n)$ is a sequence of pairwise distinct indices, and likewise for $\boldsymbol{j} = (j_1, \ldots, j_m)$. Let $\boldsymbol{a} = \boldsymbol{i} \cap \boldsymbol{j}$ be the set of indices appearing in both $\boldsymbol{i}$ and $\boldsymbol{j}$ and $\boldsymbol{b} = \boldsymbol{i} \cup \boldsymbol{j}$ be the set of indices appearing in either $\boldsymbol{i}$ or $\boldsymbol{j}$. Then, the combination of X and Y is*

$$X \star Y \coloneqq \operatorname{tr}_{\mathcal{H}_{\boldsymbol{a}}}\left(X^{\mathrm{T}_{\mathcal{H}_{\boldsymbol{a}}}} \cdot Y\right) = \operatorname{tr}_{\mathcal{H}_{\boldsymbol{a}}}\left(X \cdot Y^{\mathrm{T}_{\mathcal{H}_{\boldsymbol{a}}}}\right),$$

where $\mathcal{H}_{\boldsymbol{a}}$ denotes the tensor product of subsystems labeled by the indices in $\boldsymbol{a}$, $\mathrm{T}_{\mathcal{H}_{\boldsymbol{a}}}$ denotes the partial transpose on $\mathcal{H}_{\boldsymbol{a}}$, both X and Y are treated as linear operators on $\mathcal{H}_{\boldsymbol{b}}$, extended by tensoring with the identity operator as needed.

The link product provides the mathematical description of the combination of quantum combs. For instance, suppose X is an n-comb on $(\mathcal{H}_0, \mathcal{H}_1, \ldots, \mathcal{H}_{2n-1})$ and Y is an $(n-1)$-comb on $(\mathcal{H}_1, \mathcal{H}_2 \ldots, \mathcal{H}_{2n-2})$. Then,

$$X \star Y = \operatorname{tr}_{\mathcal{H}_{1:2n-2}}\left(X^{\mathrm{T}_{\mathcal{H}_{1:2n-2}}} \cdot (I_{\mathcal{H}_{2n-1}} \otimes Y \otimes I_{\mathcal{H}_0})\right) = \operatorname{tr}_{\mathcal{H}_{1:2n-2}}\left(X \cdot (I_{\mathcal{H}_{2n-1}} \otimes Y^{\mathrm{T}} \otimes I_{\mathcal{H}_0})\right)$$

turns out to be a 1-comb on $(\mathcal{H}_0, \mathcal{H}_{2n-1})$, as in the example illustrated in Figure 3. The link product also has many useful properties:

- it preserves the Löwner order: if $X, Y \geq 0$ then $X \star Y \geq 0$ [CDP09, Theorem 2];
- it is commutative, namely $X \star Y = Y \star X$, and associative, that is, $(X \star Y) \star Z = X \star (Y \star Z)$, whenever X, Y, Z do not share a common subsystem (i.e., there is no subsystem that is a subsystem of all three).

Moreover, the link product characterizes the channel concatenation under the Choi representation. Namely, given two quantum channels $\mathcal{E}_1 : \mathcal{L}(\mathcal{H}_1) \to \mathcal{L}(\mathcal{H}_2)$ and $\mathcal{E}_2 : \mathcal{L}(\mathcal{H}_2) \to \mathcal{L}(\mathcal{H}_3)$, we have $C_{\mathcal{E}_2 \circ \mathcal{E}_1} = C_{\mathcal{E}_2} \star C_{\mathcal{E}_1}$, where $C_{\mathcal{E}}$ is the Choi operator of $\mathcal{E}$.

2.3 Formalism of quantum channel testers

A *quantum channel tester* is any quantum algorithm that can make multiple queries to an unknown quantum channel in order to produce a classical output. We are going to use the quantum tester formalism based on Choi-Jamiołkowski representation (see, e.g., [CDP09, BMQ21, BMQ22]), which provides a practical framework in order to study various classes of quantum testers, such as parallel and sequential ones.

Consider a quantum channel tester using n queries to an unknown quantum channel $\mathcal{E}$. Let us label the input and output systems of the i-th query to $\mathcal{E}$ as $\mathcal{H}_{\mathrm{A},i}$ and $\mathcal{H}_{\mathrm{B},i}$, i.e. the i-th copy of the unknown channel is a linear map from $\mathcal{L}(\mathcal{H}_{\mathrm{A},i})$ to $\mathcal{L}(\mathcal{H}_{\mathrm{B},i})$.

Parallel testers. A parallel tester can only make parallel queries. Specifically, it prepares a multipartite input state, possibly including auxiliary systems, and applies the unknown channel in parallel to its subsystems, ensuring that the output of any use never interacts with the inputs of the others. Eventually, after all channel uses, a single joint measurement is performed on the combined output state.

Definition 2.9 (Parallel tester). *A parallel tester is defined as any set of linear operators $\{T_i\}_i$, with $T_i \in \mathcal{L}(\bigotimes_{j=1}^n \mathcal{H}_{\mathrm{A},j} \otimes \mathcal{H}_{\mathrm{B},j})$, such that $T_i \geq 0$ and $\sum_i T_i = \rho_{\mathrm{A}} \otimes I_{\mathrm{B}}$, where I_{B} is the identity operator on $\bigotimes_{j=1}^n \mathcal{H}_{\mathrm{B},j}$, and ρ_{A} is a positive semidefinite operator on $\bigotimes_{j=1}^n \mathcal{H}_{\mathrm{A},j}$ with* $\mathrm{tr}(\rho_{\mathrm{A}}) = 1$.

When applying a parallel tester $\{T_i\}_i$ to a quantum channel $\mathcal{E}$, we get the classical outcome i with probability

$$p_i = T_i \star C_{\mathcal{E}}^{\otimes n} = \mathrm{tr}(T_i (C_{\mathcal{E}}^{\otimes n})^{\mathrm{T}}) = \mathrm{tr}(T_i^{\mathrm{T}} C_{\mathcal{E}}^{\otimes n}), \tag{6}$$

where $C_{\mathcal{E}}^{\otimes n} \in \mathcal{L}(\bigotimes_{j=1}^n \mathcal{H}_{\mathrm{A},j} \otimes \mathcal{H}_{\mathrm{B},j})$ denotes the Choi operator of all n queries to the channel $\mathcal{E}$ and $(\,\cdot\,)^{\mathrm{T}}$ is matrix transposition.

To see how the parallel tester $\{T_i\}_i$ can be realized by an algorithm that makes queries in parallel, we consider the following procedure.

- Assume $\sum_i T_i = \rho_{\mathrm{A}} \otimes I_{\mathrm{B}}$; prepare a quantum state $(\sqrt{\rho_{\mathrm{A}}}^{\mathrm{T}} \otimes I_{\mathrm{A}})|I_{\mathrm{A}}\rangle\rangle$ in $\bigotimes_{j=1}^n \mathcal{H}_{\mathrm{A},j} \otimes \bigotimes_{j=1}^n \mathcal{H}_{\mathrm{A},j}$. This is indeed a valid quantum state, since $\langle\langle I_{\mathrm{A}}|(\rho_{\mathrm{A}}^{\mathrm{T}} \otimes I_{\mathrm{A}})|I_{\mathrm{A}}\rangle\rangle = \mathrm{tr}(\rho_{\mathrm{A}}^{\mathrm{T}}) = 1$.
- Then, apply the quantum channel $\mathcal{I}_{\mathrm{A}} \otimes \mathcal{E}^{\otimes n}$ to the prepared state, obtaining the mixed state $(\sqrt{\rho_{\mathrm{A}}}^{\mathrm{T}} \otimes I_{\mathrm{B}}) C_{\mathcal{E}}^{\otimes n} (\sqrt{\rho_{\mathrm{A}}}^{\mathrm{T}} \otimes I_{\mathrm{B}})$.
- Finally, perform the POVM $\left\{(\sqrt{\rho}_{\mathrm{A}}^{\mathrm{T}} \otimes I_{\mathrm{B}})^{-1}\, T_i^{\mathrm{T}}\, (\sqrt{\rho}_{\mathrm{A}}^{\mathrm{T}} \otimes I_{\mathrm{B}})^{-1}\right\}_i$, obtaining the result i, where $(\cdot)^{-1}$ is the pseudo-inverse. Then, it is easy to see that the probability of getting result i is exactly that in Equation (6).

Conversely, all algorithms that make queries in parallel can be described by a parallel tester. To see this, assume that the algorithm first prepares a state ρ on $(\bigotimes_{j=1}^n \mathcal{H}_{\mathrm{A},j}) \otimes \mathcal{H}_{\mathrm{anc}}$, where $\mathcal{H}_{\mathrm{anc}}$ is an ancilla system, and then apply the channel $\mathcal{E}^{\otimes n} \otimes \mathcal{I}_{\mathrm{anc}}$ on ρ followed by a POVM $\{E_i\}_i$, where each $E_i \in \mathcal{L}((\bigotimes_{j=1}^n \mathcal{H}_{\mathrm{B},j}) \otimes \mathcal{H}_{\mathrm{anc}})$ is positive semidefinite. Then, let $T_i = E_i^{\mathrm{T}} \star \rho$. We see that $\{T_i\}_i$ is a parallel tester, and the probability of obtain outcome i is given by

$$\mathrm{tr}(E_i \cdot (\mathcal{E}^{\otimes n} \otimes \mathcal{I}_{\mathrm{anc}})(\rho)) = \mathrm{tr}(E_i \cdot (C_{\mathcal{E}}^{\otimes n} \star \rho)) = E_i^{\mathrm{T}} \star C_{\mathcal{E}}^{\otimes n} \star \rho = T_i \star C_{\mathcal{E}}^{\otimes n},$$

which is exactly the same as Equation (6).

Sequential testers. A sequential tester can make queries sequentially, adaptively and coherently. Specifically, it sends a quantum system through the first use of the channel $\mathcal{E}$ and then it feeds the resulting (quantum) output into subsequent uses, potentially along with auxiliary systems, while allowing arbitrary CPTP maps to act between uses of $\mathcal{E}$. After all n uses of the channel $\mathcal{E}$, a POVM is performed on the final state. In other words, sequential testers represent coherent and adaptive query-access algorithms.

Definition 2.10 (Sequential tester)**.** *A sequential tester that uses n queries to an unknown channel is defined as any set of linear operators $\{T_i\}_i$, with $T_i \in \mathcal{L}(\bigotimes_{j=1}^n \mathcal{H}_{\mathrm{A},j} \otimes \mathcal{H}_{\mathrm{B},j})$, such that $T_i \geq 0$ and $\sum_i T_i$ is a quantum $(n+1)$-comb on $(\mathcal{H}_0, \mathcal{H}_{\mathrm{A},1}, \mathcal{H}_{\mathrm{B},1}, \ldots, \mathcal{H}_{\mathrm{A},n}, \mathcal{H}_{\mathrm{B},n}, \mathcal{H}_{n+1})$, where $\mathcal{H}_0 \cong \mathcal{H}_{n+1} \cong \mathbb{C}$ are one-dimensional.*

It is known that any sequential tester can be obtained by a sequential query-access algorithm and any sequential query-access algorithm can be described by a sequential tester [CDP09, BMQ22]. When applying a sequential tester $\{T_i\}_i$ to n queries to a quantum channel $\mathcal{E}$, we get the classical outcome i with probability

$$p_i = T_i \star C_{\mathcal{E}}^{\otimes n} = \mathrm{tr}(T_i (C_{\mathcal{E}}^{\otimes n})^{\mathrm{T}}) = \mathrm{tr}(T_i^{\mathrm{T}} C_{\mathcal{E}}^{\otimes n}),$$

where $C_{\mathcal{E}}^{\otimes n}$ denotes the Choi operator of all n queries to the channel $\mathcal{E}$ and $(\,\cdot\,)^{\mathrm{T}}$ represents matrix transposition.

Quantum channel discrimination. Let $\mathcal{N}$ be a finite set of quantum channels. Then, the discrimination problem for channels in $\mathcal{N}$ can be defined as follows.

Problem 2.11. *Suppose the channel $\mathcal{E}$ is uniformly randomly chosen from the set $\mathcal{N}$. The algorithm (or tester) can make n queries to the channel $\mathcal{E}$ in order to identify $\mathcal{E}$.*

Consider a sequential tester $\{T_{\mathcal{E}}\}_{\mathcal{E}\in\mathcal{N}}$ for this discrimination task, where $T_{\mathcal{E}}$ corresponds to outputting the label $\mathcal{E}$. Then, the success probability can be computed as

$$\mathbb{P}[\mathrm{success}] = \frac{1}{|\mathcal{N}|} \sum_{\mathcal{E}\in\mathcal{N}} T_{\mathcal{E}} \star C_{\mathcal{E}}^{\otimes n},$$

where $C_{\mathcal{E}}$ is the (unnormalized) Choi state of $\mathcal{E}$. We say an algorithm solves the discrimination problem if the success probability is larger than $2/3$.

2.4 Schur-Weyl duality on bipartite systems

Let us consider a sequence of Hilbert spaces $\mathcal{H}_1, \mathcal{H}_2, \ldots, \mathcal{H}_n$ such that $\mathcal{H}_i \cong \mathbb{C}^d$ for $1 \leq i \leq n$. The Hilbert space $\bigotimes_{i=1}^n \mathcal{H}_i$ admits representations of the symmetric group $\mathfrak{S}_n$ (i.e. the group of all permutations on the set $\{1, 2, \ldots, n\}$) and unitary group $\mathbb{U}_d$ (i.e. the group of unitaries on d-dimensional Hilbert space). The unitary group acts by simultaneous "rotation" as $U^{\otimes n}$ for any $U \in \mathbb{U}_d$, while the symmetric group acts by permuting tensor factors:

$$\mathtt{p}(\pi)|\psi_1\rangle \cdots |\psi_n\rangle = |\psi_{\pi^{-1}(1)}\rangle \cdots |\psi_{\pi^{-1}(n)}\rangle, \tag{7}$$

where $\pi \in \mathfrak{S}_n$. The two actions $U^{\otimes n}$ and $\mathtt{p}(\pi)$ commute with each other, hence $\bigotimes_{i=1}^n \mathcal{H}_i$ admits a representation of group $\mathbb{U}_d \times \mathfrak{S}_n$. In particular, the Schur-Weyl duality (see, e.g., [FH13]) states that

$$\bigotimes_{i=1}^n \mathcal{H}_i \overset{\mathfrak{S}_n \times \mathbb{U}_d}{\cong} \bigoplus_{\lambda \vdash_d n} \mathcal{P}_\lambda \otimes \mathcal{Q}_\lambda^d, \tag{8}$$

where $\mathcal{P}_\lambda$ and $\mathcal{Q}_\lambda^d$ are irreducible representations of $\mathfrak{S}_n$ and $\mathbb{U}_d$ labeled by Young diagram λ, respectively. We denote by $\mathtt{p}_\lambda(\pi)$ and $\mathtt{q}_\lambda(U)$ the actions of $\pi \in \mathfrak{S}_n$ and $U \in \mathbb{U}_d$ on $\mathcal{P}_\lambda$ and $\mathcal{Q}_\lambda^d$, respectively.

Now, let us consider two sequences of Hilbert spaces $(\mathcal{H}_{\mathrm{A},1}, \dots, \mathcal{H}_{\mathrm{A},n})$ and $(\mathcal{H}_{\mathrm{B},1}, \dots, \mathcal{H}_{\mathrm{B},n})$, where $\mathcal{H}_{\mathrm{A},i} \cong \mathbb{C}^{d_1}$ and $\mathcal{H}_{\mathrm{B},j} \cong \mathbb{C}^{d_2}$. We define the action of group $\mathfrak{S}_n \times \mathfrak{S}_n$ on $\bigotimes_{i=1}^n \mathcal{H}_{\mathrm{A},i} \otimes \mathcal{H}_{\mathrm{B},i}$ as $\mathtt{p}_\mathrm{A}(\pi_1) \otimes \mathtt{p}_\mathrm{B}(\pi_2)$ for $(\pi_1, \pi_2) \in \mathfrak{S}_n \times \mathfrak{S}_n$, where $\mathtt{p}_\mathrm{A}(\cdot)$ denotes the permutation action on $\bigotimes_{i=1}^n \mathcal{H}_{\mathrm{A},i}$, and similarly for $\mathtt{p}_\mathrm{B}(\,\cdot\,)$. We define the action of $\mathbb{U}_{d_1} \times \mathbb{U}_{d_2}$ on $\bigotimes_{i=1}^n \mathcal{H}_{\mathrm{A},i} \otimes \mathcal{H}_{\mathrm{B},i}$ as $(U_\mathrm{A} \otimes U_\mathrm{B})^{\otimes n}$ for $(U_\mathrm{A}, U_\mathrm{B}) \in \mathbb{U}_{d_1} \times \mathbb{U}_{d_2}$. Since the tensor product of irreducible representations of two groups G_1, G_2 is still irreducible with respect to $G_1 \times G_2$, we can see the Schur-Weyl duality on the bipartite system:

$$\bigotimes_{i=1}^n \mathcal{H}_{\mathrm{A},i} \otimes \mathcal{H}_{\mathrm{B},i} \overset{\mathfrak{S}_n \times \mathfrak{S}_n \times \mathbb{U}_{d_1} \times \mathbb{U}_{d_2}}{\cong} \bigoplus_{\substack{\lambda \vdash_{d_1} n \\ \mu \vdash_{d_2} n}} \mathcal{P}_\lambda \otimes \mathcal{P}_\mu \otimes \mathcal{Q}_\lambda^{d_1} \otimes \mathcal{Q}_\mu^{d_2},$$

where $\mathcal{P}_\lambda \otimes \mathcal{P}_\mu \otimes \mathcal{Q}_\lambda^{d_1} \otimes \mathcal{Q}_\mu^{d_2}$ is an irreducible representation of $\mathfrak{S}_n \times \mathfrak{S}_n \times \mathbb{U}_{d_1} \times \mathbb{U}_{d_2}$.

3 Local test of quantum channels

Here, we first prove the local test of quantum channels and use it to derive our upper bounds for quantum channel tomography.

Theorem 3.1. *Suppose d_1, d_2 and r are positive integers such that $rd_2 \geq d_1$. If there exists a parallel tester $\{T_i\}_i$ that makes n queries to an unknown isometry $\mathcal{V} \in \mathbf{ISO}_{d_1, rd_2}$ and produces a classical outcome i with probability $P_i(\mathcal{V})$, then there exists another parallel tester $\{\widetilde{T}_i\}_i \cup \{\widetilde{T}_\perp\}$ (where $\perp$ is an extra irrelevant label outside the range of i) that makes n queries to an unknown channel $\mathcal{E} \in \mathbf{QChan}_{d_1,d_2}^r$, and produces a classical outcome i with probability $\mathop{\mathbb{E}}_{\mathcal{V} \sim \mathbf{Dilation}_r(\mathcal{E})}[P_i(\mathcal{V})]$.*

The proof of Theorem 3.1 is deferred to Section 3.2. Then, for a more convenient use of Theorem 3.1, we need the following definition.

Definition 3.2 (Estimation tasks of quantum channels)**.** *An estimation task of quantum channels in $\mathbf{QChan}_{d_1,d_2}^r$ is described by a set $\{A_\mathcal{E}\}_{\mathcal{E} \in \mathbf{QChan}_{d_1,d_2}^r}$, where $A_\mathcal{E}$ denotes the set of classical outcomes considered to be the "correct answer" w.r.t. the unknown channel $\mathcal{E}$.*

We can easily obtain the following result, by using Theorem 3.1.

Theorem 3.3. *Suppose d_1, d_2 and r are positive integers such that $rd_2 \geq d_1$, and $\{A_\mathcal{E}\}_{\mathcal{E} \in \mathbf{QChan}_{d_1,d_2}^r}$ is an estimation task of quantum channels. If there exists a parallel tester that makes n queries to an arbitrary dilation $\mathcal{V} \in \mathbf{Dilation}_r(\mathcal{E})$ of an unknown channel $\mathcal{E} \in \mathbf{QChan}_{d_1,d_2}^r$ and produces a classical outcome $i \in A_\mathcal{E}$ with probability at least $1-\delta$, then there exists another parallel tester that makes n queries to $\mathcal{E}$ itself and produces a classical outcome $i \in A_\mathcal{E}$ with probability at least $1-\delta$.*

Proof. Suppose $P_i(\mathcal{V})$ is the probability of the parallel tester outputting i when making queries to $\mathcal{V}$. Therefore, we can see that $\sum_{i \in A_\mathcal{E}} P_i(\mathcal{V}) \geq 1 - \delta$ for any $\mathcal{V} \in \mathbf{Dilation}_r(\mathcal{E})$. Due to Theorem 3.1, there exists a parallel tester that makes n queries to $\mathcal{E}$ and outputs i with probability

$$\widetilde{P}_i(\mathcal{E}) = \mathop{\mathbb{E}}_{\mathcal{V} \sim \mathbf{Dilation}_r(\mathcal{E})}[P_i(\mathcal{V})].$$

Therefore, we can see that the probability of this tester outputting $i \in A_{\mathcal{E}}$ is lower bounded as

$$\sum_{i \in A_{\mathcal{E}}} \widetilde{P}_i(\mathcal{E}) = \mathop{\mathbb{E}}_{\mathcal{V} \sim \mathbf{Dilation}_r(\mathcal{E})} \left[\sum_{i \in A_{\mathcal{E}}} P_i(\mathcal{V}) \right] \geq 1 - \delta.$$

□

3.1 Quantum channel tomography and estimation

We can use Theorem 3.3 and the isometry tomography algorithm under diamond norm shown in Section 8.1 to give the following result.

Theorem 3.4. *There exists a parallel tester that makes $O(rd_1d_2/\varepsilon^2)$ queries to an unknown channel $\mathcal{E} \in \mathbf{QChan}^r_{d_1,d_2}$ and produces an estimate $\mathcal{F}$ satisfying $\|\mathcal{F} - \mathcal{E}\|_\diamond \leq \varepsilon$ with probability at least $2/3$. Here, $\|\cdot\|_\diamond$ denotes the diamond norm.*

Proof. We consider the estimation task $\{A_{\mathcal{E}}\}_{\mathcal{E} \in \mathbf{QChan}^r_{d_1,d_2}}$ where

$$A_{\mathcal{E}} = \left\{ \mathcal{F} \in \mathbf{QChan}^r_{d_1,d_2} \;\middle|\; \|\mathcal{F} - \mathcal{E}\|_\diamond \leq \varepsilon \right\}.$$

Note that, for $\mathcal{E} \in \mathbf{QChan}^r_{d_1,d_2}$, any isometry in $\mathbf{Dilation}_r(\mathcal{E})$ is in $\mathbf{ISO}_{d_1,rd_2}$. Then, due to Lemma 8.1, we have a parallel tester that makes $n = O(rd_1d_2/\varepsilon^2)$ queries to a dilation $\mathcal{V} \in \mathbf{Dilation}_r(\mathcal{E})$ and produces $\mathcal{W}$ satisfying

$$\|\mathbf{Contract}_r(\mathcal{W}) - \mathcal{E}\|_\diamond \leq \|\mathcal{W} - \mathcal{V}\|_\diamond \leq \varepsilon,$$

with probability at least $2/3$, where the first inequality is by the contractivity of the diamond norm. We set the tester to finally output $\mathbf{Contract}_r(\mathcal{W})$ when producing $\mathcal{W}$. Then, it solves the task $\{A_{\mathcal{E}}\}_{\mathcal{E} \in \mathbf{QChan}^r_{d_1,d_2}}$ by making n queries to a dilation of $\mathcal{E}$. Due to Theorem 3.3, there exists another parallel tester that solves this task by making n queries to $\mathcal{E}$ it self. □

We also provide upper bounds with Heisenberg scaling $O(1/\varepsilon)$ for (possibly non-unitary) channel tomography in the boundary regime ($\tau = 1$), by using Theorem 3.3 and the unitary tomography algorithm shown in [YRC20].

Theorem 3.5. *Suppose $\tau = rd_2/d_1 = 1$. There exists a parallel tester that makes $O(rd_1d_2/\varepsilon)$ queries to an unknown channel $\mathcal{E} \in \mathbf{QChan}^r_{d_1,d_2}$ and produces an estimate $\mathcal{F}$ satisfying $\|\frac{1}{d}C_{\mathcal{F}} - \frac{1}{d}C_{\mathcal{E}}\|_1 \leq \varepsilon$ with probability at least $2/3$. Here, $C_{\mathcal{E}}$ is the (unnormalized) Choi state of $\mathcal{E}$ and $\|\cdot\|_1$ denotes the trace norm.*

Moreover, there exists a parallel tester that makes $O(\min\{rd_1^{1.5}d_2/\varepsilon, rd_1d_2/\varepsilon^2\})$ queries to an unknown channel $\mathcal{E} \in \mathbf{QChan}^r_{d_1,d_2}$ and produces an estimate $\mathcal{F}$ satisfying $\|\mathcal{F} - \mathcal{E}\|_\diamond \leq \varepsilon$ with probability at least $2/3$, where $\|\cdot\|_\diamond$ denotes the diamond norm.

Proof. Consider the task $\{A_{\mathcal{E}}\}_{\mathcal{E} \in \mathbf{QChan}^r_{d_1,d_2}}$ where

$$A_{\mathcal{E}} = \left\{ \mathcal{F} \in \mathbf{QChan}^r_{d_1,d_2} \;\middle|\; \left\| \frac{1}{d}C_{\mathcal{F}} - \frac{1}{d}C_{\mathcal{E}} \right\|_1 \leq \varepsilon \right\}.$$

Note that for $\mathcal{E} \in \mathbf{QChan}^r_{d_1,d_2}$, any isometry in $\mathbf{Dilation}_r(\mathcal{E})$ is a d_1-dimensional unitary. Using Yang-Renner-Chiribella algorithm [YRC20], we have a parallel tester that makes $n = O(d_1^2/\varepsilon) = O(rd_1d_2/\varepsilon)$ queries to a unitary $\mathcal{U} \in \mathbf{Dilation}_r(\mathcal{E})$ and produces $\mathcal{W}$ satisfying

$$\left\| \frac{1}{d}C_{\mathbf{Contract}_r(\mathcal{W})} - \frac{1}{d}C_{\mathcal{E}} \right\|_1 \leq \left\| \frac{1}{d}C_{\mathcal{W}} - \frac{1}{d}C_{\mathcal{U}} \right\|_1 = \sqrt{1 - \mathrm{F}_{\mathrm{ch}}(\mathcal{W}, \mathcal{U})} \leq \varepsilon,$$

with probability at least $2/3$, where the first inequality is due to the contractivity of trace norm, and the last inequality is because the Yang-Renner-Chiribella algorithm produces an estimate $\mathcal{W}$ satisfying $\mathrm{F_{ch}}(\mathcal{W},\mathcal{U}) \geq 1-\varepsilon^2$, with probability at least $2/3$. Here, $\mathrm{F_{ch}}$ denotes the channel fidelity (or equivalently, entanglement fidelity). We set the tester to finally output $\mathbf{Contract}_r(\mathcal{W})$ upon producing $\mathcal{W}$. Then it solves the task $\{A_\mathcal{E}\}_{\mathcal{E}\in\mathbf{QChan}^r_{d_1,d_2}}$ by making n queries to a dilation of $\mathcal{E}$. Due to Theorem 3.3, there exists a parallel tester that solves this task by making n queries to $\mathcal{E}$ itself.

Similarly, consider the task $\{A_\mathcal{E}\}_{\mathcal{E}\in\mathbf{QChan}^r_{d_1,d_2}}$ where

$$A_\mathcal{E} = \left\{\mathcal{F}\in\mathbf{QChan}^r_{d_1,d_2} \;\middle|\; \|\mathcal{F}-\mathcal{E}\|_\diamond \leq \varepsilon\right\}.$$

Using to Yang-Renner-Chiribella algorithm [YRC20], we have a parallel tester that makes $n = O(rd_1d_2/(\varepsilon/\sqrt{d_1})) = O(rd_1^{1.5}d_2/\varepsilon)$ queries to a unitary $\mathcal{U}\in\mathbf{Dilation}_r(\mathcal{E})$ and produces $\mathcal{W}$ satisfying

$$\|\mathbf{Contract}_r(\mathcal{W})-\mathcal{E}\|_\diamond \leq \|\mathcal{W}-\mathcal{U}\|_\diamond \leq \sqrt{2d_1}\sqrt{1-\mathrm{F_{ch}}(\mathcal{W},\mathcal{U})} \leq \varepsilon,$$

with probability at least $2/3$, where the second inequality is by [HKOT23, Proposition 1.9]. We set the tester to finally output $\mathbf{Contract}_r(\mathcal{W})$ upon producing $\mathcal{W}$. Then it solves the task $\{A_\mathcal{E}\}_{\mathcal{E}\in\mathbf{QChan}^r_{d_1,d_2}}$ by making n queries to a dilation of $\mathcal{E}$. Due to Theorem 3.3, there exists a parallel tester that solves this task by making n queries to $\mathcal{E}$ itself. Combining this with Theorem 3.4, we can see that $O(\min\{rd_1^{1.5}d_2/\varepsilon, rd_1d_2/\varepsilon^2\})$ queries suffice. ☐

We also demonstrate (mixed) state tomography with Heisenberg scaling, when state-preparation channels are available. This is obtained by combining our Theorem 3.3 with the pure state tomography algorithm given in [Che25].

Theorem 3.6. *Suppose $\tau = rd_2/d_1 = 1$. There exists a parallel tester that makes $O(\min\{d_1^{1.5}/\varepsilon, d_1/\varepsilon^2\})$ queries to a channel $\mathcal{E}\in\mathbf{QChan}_{d_1,d_2}$ and produces an estimate ρ satisfying $\|\rho-\mathcal{E}(|0\rangle\langle 0|)\|_1 \leq \varepsilon$ with probability at least $2/3$, where $\|\cdot\|_1$ is the trace norm.*

Proof. Consider the task $\{A_\mathcal{E}\}_{\mathcal{E}\in\mathbf{QChan}^r_{d_1,d_2}}$ where

$$A_\mathcal{E} = \left\{\rho\in\mathcal{L}(\mathbb{C}^{d_2}) \;\middle|\; \|\rho-\mathcal{E}(|0\rangle\langle 0|)\|_1 \leq \varepsilon\right\}.$$

Note that, for $\mathcal{E}\in\mathbf{QChan}^r_{d_1,d_2}$, any isometry in $\mathbf{Dilation}_r(\mathcal{E})$ is a d_1-dimensional unitary. Using Chen's algorithm [Che25], we have a parallel tester that makes $n = O(\min\{d_1^{1.5}/\varepsilon, d_1/\varepsilon^2\})$ queries to a unitary $\mathcal{U}\in\mathbf{Dilation}_r(\mathcal{E})$ and produces $|\psi\rangle$ satisfying

$$\|\operatorname{tr}_r(|\psi\rangle\langle\psi|)-\mathcal{E}(|0\rangle\langle 0|)\|_1 \leq \||\psi\rangle\langle\psi| - U|0\rangle\langle 0|U^\dagger\|_1 \leq \varepsilon,$$

with probability at least $2/3$, where the second inequality is because Chen's algorithm produces an estimate $|\psi\rangle$ for $U|0\rangle$[4] to within trace norm error ε with probability at least $2/3$. We set the tester to finally output $\operatorname{tr}_r(|\psi\rangle\langle\psi|)$ upon producing $|\psi\rangle$. Then, it solves the task $\{A_\mathcal{E}\}_{\mathcal{E}\in\mathbf{QChan}^r_{d_1,d_2}}$ by making n queries to a dilation of $\mathcal{E}$. Due to Theorem 3.3, there exists a parallel tester that solves this task by making n queries to $\mathcal{E}$ itself. ☐

Note that Theorem 3.5 and Theorem 3.6 provide Heisenberg-scaling upper bounds in the boundary regime $\tau = rd_2/d_1 = 1$. We also provide upper bounds with mixing Heisenberg- and classical-scalings in the non-boundary regime $\tau > 1$. This is obtained by combining our Theorem 3.3 with the isometry tomography algorithm provided in [YMM25].

[4] In [Che25], the author considered estimating $U|d\rangle$ for notation convenience, here we consider estimating $U|0\rangle$.

Theorem 3.7. *Suppose $\tau = rd_2/d_1 > 1$. There exists a parallel tester that makes $O((rd_2 - d_1)d_1/\varepsilon^2 + d_1^2/\varepsilon)$ queries to an unknown channel $\mathcal{E} \in \mathbf{QChan}^r_{d_1,d_2}$ and produces an estimate $\mathcal{F}$ satisfying $\|\frac{1}{d}C_{\mathcal{F}} - \frac{1}{d}C_{\mathcal{E}}\|_1 \le \varepsilon$ with probability at least $2/3$. Here, $C_{\mathcal{E}}$ is the unnormalized Choi state of $\mathcal{E}$ and $\|\cdot\|_1$ denotes the trace norm.*

Proof. Consider the estimation task $\{A_{\mathcal{E}}\}_{\mathcal{E}\in\mathbf{QChan}^r_{d_1,d_2}}$ where

$$A_{\mathcal{E}} = \left\{\mathcal{F} \in \mathbf{QChan}^r_{d_1,d_2} \;\middle|\; \left\|\frac{1}{d}C_{\mathcal{F}} - \frac{1}{d}C_{\mathcal{E}}\right\|_1 \le \varepsilon\right\}.$$

Note that, for $\mathcal{E} \in \mathbf{QChan}^r_{d_1,d_2}$, any isometry in $\mathbf{Dilation}_r(\mathcal{E})$ is in $\mathbf{ISO}_{d_1,rd_2}$. Using Yoshida-Miyazaki-Murao algorithm [YMM25] (see also Lemma 8.4), we have a parallel tester that makes $n = O((rd_2 - d_1)d_1/\varepsilon^2 + d_1^2/\varepsilon)$ queries to an isometry $\mathcal{V} \in \mathbf{Dilation}_r(\mathcal{E})$ and produces an estimate $\mathcal{W}$ satisfying

$$\left\|\frac{1}{d}C_{\mathbf{Contract}_r(\mathcal{W})} - \frac{1}{d}C_{\mathcal{E}}\right\|_1 \le \left\|\frac{1}{d}C_{\mathcal{W}} - \frac{1}{d}C_{\mathcal{V}}\right\|_1 \le \varepsilon,$$

with probability at least $2/3$, where the first inequality is by the contractivity of the trace norm. We set the tester to finally output $\mathbf{Contract}_r(\mathcal{W})$ upon producing $\mathcal{W}$. Then it solves the task $\{A_{\mathcal{E}}\}_{\mathcal{E}\in\mathbf{QChan}^r_{d_1,d_2}}$ by making n queries to a dilation of $\mathcal{E}$. Due to Theorem 3.3, there exists a parallel tester that solves this task by making n queries to $\mathcal{E}$ itself. □

3.2 Construction of the local testers

Now, we prove Theorem 3.1. Our proof draws on an idea similar to the construction of local testers for quantum states in [CWZ24], and extends it to local testers for quantum channels.

We introduce some notation that will be used in the proof later.

Notation 3.8. *For each $i \in [n]$, we define $\mathcal{H}_{\mathrm{A},i} \cong \mathbb{C}^{d_1}$ and $\mathcal{H}_{\mathrm{B},i} \otimes \mathcal{H}_{\mathrm{anc},i} \cong \mathbb{C}^{d_2} \otimes \mathbb{C}^r$ to be the input and output systems of the i-th query to an isometry channel $\mathcal{V} \in \mathbf{ISO}_{d_1,rd_2}$. The following decompositions are due to Schur-Weyl duality:*

$$\bigotimes_{i=1}^n \mathcal{H}_{\mathrm{A},i} \otimes \mathcal{H}_{\mathrm{B},i} \overset{\mathfrak{S}_n\times\mathbb{U}_{d_1d_2}}{\cong} \bigoplus_{\lambda\vdash_{d_1d_2}n} \mathcal{P}_\lambda \otimes \mathcal{Q}^{d_1d_2}_\lambda \quad \text{and} \quad \bigotimes_{i=1}^n \mathcal{H}_{\mathrm{anc},i} \overset{\mathfrak{S}_n\times\mathbb{U}_r}{\cong} \bigoplus_{\lambda\vdash_r n} \mathcal{P}_\lambda \otimes \mathcal{Q}^r_\lambda.$$

Hence, we can see that

$$\bigotimes_{i=1}^n \mathcal{H}_{\mathrm{A},i} \otimes \mathcal{H}_{\mathrm{B},i} \otimes \mathcal{H}_{\mathrm{anc},i} \overset{\mathfrak{S}_n\times\mathfrak{S}_n\times\mathbb{U}_{d_1d_2}\times\mathbb{U}_r}{\cong} \bigoplus_{\substack{\lambda\vdash_{d_1d_2}n\\ \mu\vdash_r n}} \mathcal{P}_{\mathrm{AB},\lambda} \otimes \mathcal{P}_{\mathrm{anc},\mu} \otimes \mathcal{Q}^{d_1d_2}_{\mathrm{AB},\lambda} \otimes \mathcal{Q}^r_{\mathrm{anc},\mu},$$

where we use $\mathcal{P}_{\mathrm{AB},\lambda} \otimes \mathcal{Q}^{d_1d_2}_{\mathrm{AB},\lambda}$ to denote the subspace $\mathcal{P}_\lambda \otimes \mathcal{Q}^{d_1d_2}_\lambda$ in $\bigotimes_{i=1}^n \mathcal{H}_{\mathrm{A},i} \otimes \mathcal{H}_{\mathrm{B},i}$, and use $\mathcal{P}_{\mathrm{anc},\mu} \otimes \mathcal{Q}^r_{\mathrm{anc},\mu}$ to denote the subspace $\mathcal{P}_\mu \otimes \mathcal{Q}^r_\mu$ in $\bigotimes_{i=1}^n \mathcal{H}_{\mathrm{anc},i}$.

We give the proof of Theorem 3.1 as follows.

Proof of Theorem 3.1. Define $s \coloneqq \min\{d_1d_2, r\}$. Note that here, we do not assume $r \le d_1d_2$. The construction of the parallel tester $\{\widetilde{T}_i\}_i \cup \{\widetilde{T}_\perp\}$ is as follows.

- First, we define a new tester $\{\overline{T}_i\}_i$ such that

$$\overline{T}_i \coloneqq \mathop{\mathbb{E}}_{U\sim\mathbb{U}_r}[U^{\otimes n} T_i U^{\dagger\otimes n}], \tag{9}$$

  where $U^{\otimes n}$ acts on $\bigotimes_{j=1}^n \mathcal{H}_{\mathrm{anc},j}$.

- We then define the tester $\{\widetilde{T}_i\}_i \cup \{\widetilde{T}_\perp\}$ as follows. For each i,

$$\widetilde{T}_i \coloneqq \bigoplus_{\lambda\vdash_s n} \frac{1}{\dim(\mathcal{P}_\lambda)\dim(\mathcal{Q}_\lambda^r)} \cdot I_{\mathcal{P}_{\mathrm{AB},\lambda}} \otimes \operatorname{tr}_{\mathcal{Q}^r_{\mathrm{anc},\lambda}}\left(\langle\!\langle I_{\mathcal{P}_\lambda}|\overline{T}_i|I_{\mathcal{P}_\lambda}\rangle\!\rangle\right), \tag{10}$$

  where $|I_{\mathcal{P}_\lambda}\rangle\!\rangle \in \mathcal{P}_{\mathrm{AB},\lambda}\otimes\mathcal{P}_{\mathrm{anc},\lambda}$ is the (unnormalized) maximally entangled state defined w.r.t. the Young's orthogonal basis (or Young-Yamanouchi basis) on which $\pi\in\mathfrak{S}_n$ acts as a real matrix [CSST10]. Note that $\widetilde{T}_i$ is a linear operator on $\bigotimes_{j=1}^n \mathcal{H}_{\mathrm{A},j}\otimes\mathcal{H}_{\mathrm{B},j}$. As we will not explicitly use $\widetilde{T}_\perp$, its definition is deferred to Equation (21) for clarity.

In order to verify our construction, we prove in Lemma 3.9 that $\{\overline{T}_i\}_i$ is a parallel tester which makes n queries to an isometry $\mathcal{V}\in\mathbf{ISO}_{d_1,rd_2}$ and produces a classical outcome i with probability

$$\mathop{\mathbb{E}}_{\mathcal{W}\sim\mathbf{Dilation}_r(\mathbf{Contract}_r(\mathcal{V}))}[T_i \star C_{\mathcal{W}}^{\otimes n}],$$

and we also give an explicit expression of this probability. Then, by using Lemma 3.9, we prove in Lemma 3.10 that $\{\widetilde{T}_i\}_i\cup\{\widetilde{T}_\perp\}$ is indeed a parallel tester that makes n queries to a quantum channel $\mathcal{E}\in\mathbf{QChan}^r_{d_1,d_2}$ and produces outcome i with probability $\mathop{\mathbb{E}}_{\mathcal{W}\sim\mathbf{Dilation}_r(\mathcal{E})}[T_i\star C_{\mathcal{W}}^{\otimes n}]$, as desired. ☐

In the following lemma, we prove some properties of the tester $\{\overline{T}_i\}_i$.

Lemma 3.9. *The tester $\{\overline{T}_i\}_i$ defined in Equation (9) satisfies:*

1. *$\{\overline{T}_i\}_i$ is a parallel tester. Moreover, for $\mathcal{V}\in\mathbf{ISO}_{d_1,rd_2}$, the tester produces i with probability*

$$\overline{T}_i\star C_{\mathcal{V}}^{\otimes n} = \mathop{\mathbb{E}}_{\mathcal{W}\sim\mathbf{Dilation}_r(\mathbf{Contract}_r(\mathcal{V}))}[T_i\star C_{\mathcal{W}}^{\otimes n}]. \tag{11}$$

2. *The probability in Equation (11) can also be expressed as*

$$\overline{T}_i\star C_{\mathcal{V}}^{\otimes n} = \sum_{\lambda\vdash_s n}\frac{1}{\dim(\mathcal{Q}_\lambda^r)}\operatorname{tr}\left(\operatorname{tr}_{\mathcal{Q}^r_{\mathrm{anc},\lambda}}\left(\langle\!\langle I_{\mathcal{P}_\lambda}|\overline{T}_i^{\mathrm{T}}|I_{\mathcal{P}_\lambda}\rangle\!\rangle\right)\cdot\operatorname{tr}_{\mathcal{Q}^r_{\mathrm{anc},\lambda}}\left(|V_\lambda\rangle\langle V_\lambda|\right)\right),$$

   where $s=\min\{d_1d_2,r\}$, and $|V_\lambda\rangle\in\mathcal{Q}^{d_1d_2}_{\mathrm{AB},\lambda}\otimes\mathcal{Q}^r_{\mathrm{anc},\lambda}$ is a vector appearing in the decomposition $|V\rangle\!\rangle^{\otimes n}=\bigoplus_{\lambda\vdash_s n}|I_{\mathcal{P}_\lambda}\rangle\!\rangle\otimes|V_\lambda\rangle$ according to Lemma 3.11.

Proof. **Item 1**. We have

$$\sum_i \overline{T}_i = \mathop{\mathbb{E}}_{U\sim\mathbb{U}_r}\left[U^{\otimes n}\sum_i T_i U^{\dagger\otimes n}\right] = \mathop{\mathbb{E}}_{U\sim\mathbb{U}_r}\left[U^{\otimes n}(\rho_{\mathrm{A}}\otimes I_{\mathrm{B,anc}})U^{\dagger\otimes n}\right] = \rho_{\mathrm{A}}\otimes I_{\mathrm{B,anc}}, \tag{12}$$

where we used the fact that $\{T_i\}_i$ is a parallel tester so that $\sum_i T_i = \rho_{\mathrm{A}}\otimes I_{\mathrm{B,anc}}$ for ρ_{A} a density operator on $\bigotimes_{j=1}^n\mathcal{H}_{\mathrm{A},j}$ and $I_{\mathrm{B,anc}}$ the identity operator on $\bigotimes_{j=1}^n\mathcal{H}_{\mathrm{B},j}\otimes\mathcal{H}_{\mathrm{anc},j}$. We further note that $U^{\otimes n}$ acts only non-trivially on $\bigotimes_{j=1}^n\mathcal{H}_{\mathrm{anc},j}$. Therefore, $\{\overline{T}_i\}_i$ is indeed a parallel tester.

On the other hand, we have

$$
\begin{aligned}
\overline{T}_i \star C_{\mathcal{V}}^{\otimes n} &= \operatorname{tr}\left(\overline{T}_i^{\mathrm{T}} C_{\mathcal{V}}^{\otimes n}\right) \\
&= \mathop{\mathbb{E}}_{U\sim \mathbb{U}_r}\left[\operatorname{tr}\left(T_i^{\mathrm{T}} U^{\otimes n} C_{\mathcal{V}}^{\otimes n} U^{\dagger \otimes n}\right)\right] \qquad (13)\\
&= \mathop{\mathbb{E}}_{U\sim \mathbb{U}_r}\left[\operatorname{tr}\left(T_i^{\mathrm{T}} C_{\mathcal{U}\circ\mathcal{V}}^{\otimes n}\right)\right] \\
&= \mathop{\mathbb{E}}_{\mathcal{W}\sim \mathbf{Dilation}_r(\mathbf{Contract}_r(\mathcal{V}))}\left[\operatorname{tr}\left(T_i^{\mathrm{T}} C_{\mathcal{W}}^{\otimes n}\right)\right] \qquad (14)\\
&= \mathop{\mathbb{E}}_{\mathcal{W}\sim \mathbf{Dilation}_r(\mathbf{Contract}_r(\mathcal{V}))}\left[T_i \star C_{\mathcal{W}}^{\otimes n}\right]
\end{aligned}
$$

where Equation (13) is due to the definition of $\overline{T}_i$ (see Equation (9)), and Equation (14) uses the definition of the Haar distribution on $\mathbf{Dilation}_r(\cdot)$ (see Notation 2.3).

Item 2. We treat $|V\rangle\rangle$ as a bipartite vector in $(\mathbb{C}^{d_1}\otimes\mathbb{C}^{d_2})\otimes\mathbb{C}^r$. Using Lemma 3.11, we can write $|V\rangle\rangle^{\otimes n} = \bigoplus_{\lambda\vdash_s n} |I_{\mathcal{P}_\lambda}\rangle\rangle \otimes |V_\lambda\rangle$, in which $|V_\lambda\rangle \in \mathcal{Q}_{\mathrm{AB},\lambda}^{d_1 d_2}\otimes \mathcal{Q}_{\mathrm{anc},\lambda}^r$ and $|I_{\mathcal{P}_\lambda}\rangle\rangle \in \mathcal{P}_{\mathrm{AB},\lambda}\otimes\mathcal{P}_{\mathrm{anc},\lambda}$ is an (unnormalized) maximally entangled state. Also, we note that $U^{\otimes n}$ commutes with $\overline{T}_i$ for any $U\in\mathbb{U}_r$, in which $U^{\otimes n}$ acts non-trivially on $\bigotimes_{j=1}^n \mathcal{H}_{\mathrm{anc},j}$. Thus,

$$
\begin{aligned}
\overline{T}_i \star C_{\mathcal{V}}^{\otimes n} &= \operatorname{tr}\left(\overline{T}_i^{\mathrm{T}} |V\rangle\rangle\langle\langle V|^{\otimes n}\right) \\
&= \operatorname{tr}\left(\overline{T}_i^{\mathrm{T}} \mathop{\mathbb{E}}_{U\sim\mathbb{U}_r}[U^{\otimes n}|V\rangle\rangle\langle\langle V|^{\otimes n} U^{\dagger\otimes n}]\right) \\
&= \operatorname{tr}\left(\overline{T}_i^{\mathrm{T}} \mathop{\mathbb{E}}_{U\sim\mathbb{U}_r}\left[\bigoplus_{\lambda,\mu\vdash_s n} |I_{\mathcal{P}_\lambda}\rangle\rangle\langle\langle I_{\mathcal{P}_\mu}| \otimes \mathsf{q}_\lambda(U)|V_\lambda\rangle\langle V_\mu|\mathsf{q}_\mu(U)^\dagger\right]\right) \qquad (15)\\
&= \operatorname{tr}\left(\overline{T}_i^{\mathrm{T}} \cdot \left(\bigoplus_{\lambda\vdash_s n} |I_{\mathcal{P}_\lambda}\rangle\rangle\langle\langle I_{\mathcal{P}_\lambda}| \otimes \operatorname{tr}_{\mathcal{Q}_{\mathrm{anc},\lambda}^r}\left(|V_\lambda\rangle\langle V_\lambda|\right) \otimes \frac{1}{\dim(\mathcal{Q}_\lambda^r)} I_{\mathcal{Q}_{\mathrm{anc},\lambda}^r}\right)\right) \qquad (16)\\
&= \sum_{\lambda\vdash_s n} \frac{1}{\dim(\mathcal{Q}_\lambda^r)} \operatorname{tr}\left(\operatorname{tr}_{\mathcal{Q}_{\mathrm{anc},\lambda}^r}\left(\langle\langle I_{\mathcal{P}_\lambda}|\overline{T}_i^{\mathrm{T}}|I_{\mathcal{P}_\lambda}\rangle\rangle\right)\cdot \operatorname{tr}_{\mathcal{Q}_{\mathrm{anc},\lambda}^r}\left(|V_\lambda\rangle\langle V_\lambda|\right)\right),
\end{aligned}
$$

where Equation (16) uses Schur's lemma [FH13], and in Equation (15) $\mathsf{q}_\lambda(U)$ acts on $\mathcal{Q}_{\mathrm{anc},\lambda}^r$. □

Next, we prove some properties of $\{\widetilde{T}_i\}_i$ in the following lemma.

Lemma 3.10. *The operators $\{\widetilde{T}_i\}_i$ defined in Equation (10) satisfies:*

1. *There exists a positive semidefinite operator $\widetilde{T}_\perp$ such that $\{\widetilde{T}_i\}_i\cup\{\widetilde{T}_\perp\}$ is a parallel tester. In fact, $\widetilde{T}_\perp$ can be explicitly constructed, as shown in Equation (21).*
2. *For any quantum channel $\mathcal{E}\in\mathbf{QChan}_{d_1,d_2}^r$, the tester produces outcome i with probability*

$$
\widetilde{T}_i \star C_{\mathcal{E}}^{\otimes n} = \mathop{\mathbb{E}}_{\mathcal{W}\sim\mathbf{Dilation}_r(\mathcal{E})}[T_i\star C_{\mathcal{W}}^{\otimes n}].
$$

Proof. **Item 1**. According to the definition in Equation (10), we have

$$
\sum_i \widetilde{T}_i = \bigoplus_{\lambda\vdash_s n} \frac{1}{\dim(\mathcal{P}_\lambda)\dim(\mathcal{Q}_\lambda^r)} \cdot I_{\mathcal{P}_{\mathrm{AB},\lambda}} \otimes \operatorname{tr}_{\mathcal{Q}_{\mathrm{anc},\lambda}^r}\left(\langle\langle I_{\mathcal{P}_\lambda}|\sum_i \overline{T}_i|I_{\mathcal{P}_\lambda}\rangle\rangle\right)
$$

$$= \bigoplus_{\lambda \vdash_s n} \frac{1}{\dim(\mathcal{P}_\lambda)\dim(\mathcal{Q}^r_\lambda)} \cdot I_{\mathcal{P}_{\mathrm{AB},\lambda}} \otimes \mathrm{tr}_{\mathcal{Q}^r_{\mathrm{anc},\lambda}} \left(\langle\langle I_{\mathcal{P}_\lambda} | \rho_{\mathrm{A}} \otimes I_{\mathrm{B,anc}} | I_{\mathcal{P}_\lambda} \rangle\rangle \right), \tag{17}$$

where in Equation (17) we use Equation (12). Then, we write $\rho_{\mathrm{A}} \otimes I_{\mathrm{B,anc}} = (\rho_{\mathrm{A}} \otimes I_{\mathrm{B}}) \otimes I_{\mathrm{anc}}$ and write $\rho_{\mathrm{A}} \otimes I_{\mathrm{B}}$ according to the subspace decomposition in the Schur-Weyl duality:

$$\rho_{\mathrm{A}} \otimes I_{\mathrm{B}} = \bigoplus_{\lambda,\mu \vdash_{d_1 d_2} n} M_{\lambda\to\mu},$$

where $M_{\lambda\to\mu}$ is a linear operator from the subspace $\mathcal{P}_{\mathrm{AB},\lambda} \otimes \mathcal{Q}^{d_1 d_2}_{\mathrm{AB},\lambda}$ to the subspace $\mathcal{P}_{\mathrm{AB},\mu} \otimes \mathcal{Q}^{d_1 d_2}_{\mathrm{AB},\mu}$, and $M_{\lambda\to\lambda}$ is positive semidefinite for each λ since $\rho_{\mathrm{A}} \otimes I_{\mathrm{B}}$ is positive semidefinite. Moreoever, we can see that

$$(\rho_{\mathrm{A}} \otimes I_{\mathrm{B}}) \otimes I_{\mathrm{anc}} = \bigoplus_{\substack{\lambda,\mu \vdash_{d_1 d_2} n \\ \nu \vdash_r n}} M_{\lambda\to\mu} \otimes I_{\mathcal{P}_{\mathrm{anc},\nu}} \otimes I_{\mathcal{Q}^r_{\mathrm{anc},\nu}}.$$

This means, Equation (17) is equal to

$$\begin{aligned}
&\bigoplus_{\lambda \vdash_s n} \frac{1}{\dim(\mathcal{P}_\lambda)\dim(\mathcal{Q}^r_\lambda)} \cdot I_{\mathcal{P}_{\mathrm{AB},\lambda}} \otimes \mathrm{tr}_{\mathcal{Q}^r_{\mathrm{anc},\lambda}} \left(\langle\langle I_{\mathcal{P}_\lambda} | \bigoplus_{\substack{\kappa,\mu \vdash_{d_1 d_2} n \\ \nu \vdash_r n}} M_{\kappa\to\mu} \otimes I_{\mathcal{P}_{\mathrm{anc},\nu}} \otimes I_{\mathcal{Q}^r_{\mathrm{anc},\nu}} | I_{\mathcal{P}_\lambda} \rangle\rangle \right) \\
&= \bigoplus_{\lambda \vdash_s n} \frac{1}{\dim(\mathcal{P}_\lambda)\dim(\mathcal{Q}^r_\lambda)} \cdot I_{\mathcal{P}_{\mathrm{AB},\lambda}} \otimes \mathrm{tr}_{\mathcal{Q}^r_{\mathrm{anc},\lambda}} \left(\mathrm{tr}_{\mathcal{P}_{\mathrm{AB},\lambda}}(M_{\lambda\to\lambda}) \otimes I_{\mathcal{Q}^r_{\mathrm{anc},\lambda}} \right) \qquad (18) \\
&= \bigoplus_{\lambda \vdash_s n} \frac{1}{\dim(\mathcal{P}_\lambda)} \cdot I_{\mathcal{P}_{\mathrm{AB},\lambda}} \otimes \mathrm{tr}_{\mathcal{P}_{\mathrm{AB},\lambda}}(M_{\lambda\to\lambda}) \\
&\le \bigoplus_{\lambda \vdash_{d_1 d_2} n} \frac{1}{\dim(\mathcal{P}_\lambda)} \cdot I_{\mathcal{P}_{\mathrm{AB},\lambda}} \otimes \mathrm{tr}_{\mathcal{P}_{\mathrm{AB},\lambda}}(M_{\lambda\to\lambda}), \qquad (19)
\end{aligned}$$

where Equation (18) uses that $|I_{\mathcal{P}_\lambda}\rangle\rangle$ is an (unnormalized) maximally entangled state on $\mathcal{P}_{\mathrm{AB},\lambda} \otimes \mathcal{P}_{\mathrm{anc},\lambda}$, and Equation (19) uses that $s \le d_1 d_2$ and $M_{\lambda\to\lambda} \ge 0$ for any λ. Next, we have

$$\begin{aligned}
\frac{1}{n!} \sum_{\pi \in \mathfrak{S}_n} \mathtt{p}_{\mathrm{A}}(\pi)\, \rho_{\mathrm{A}}\, \mathtt{p}_{\mathrm{A}}(\pi)^\dagger \otimes I_{\mathrm{B}} &= \frac{1}{n!} \sum_{\pi \in \mathfrak{S}_n} \mathtt{p}_{\mathrm{AB}}(\pi)(\rho_{\mathrm{A}} \otimes I_{\mathrm{B}})\mathtt{p}_{\mathrm{AB}}(\pi)^\dagger \\
&= \bigoplus_{\lambda \vdash_{d_1 d_2} n} \frac{1}{\dim(\mathcal{P}_\lambda)} \cdot I_{\mathcal{P}_{\mathrm{AB},\lambda}} \otimes \mathrm{tr}_{\mathcal{P}_{\mathrm{AB},\lambda}}(M_{\lambda\to\lambda}), \qquad (20)
\end{aligned}$$

where $\mathtt{p}_{\mathrm{A}}(\pi)$ and $\mathtt{p}_{\mathrm{AB}}(\pi)$ denote the actions of π (i.e., permuting tensor factors) on $\bigotimes_{j=1}^n \mathcal{H}_{\mathrm{A},j}$ and $\bigotimes_{j=1}^n \mathcal{H}_{\mathrm{A},j} \otimes \mathcal{H}_{\mathrm{B},j}$, respectively, and Equation (20) uses Schur's lemma. We can see that Equation (20) equals exactly Equation (19). This means

$$\sum_i \widetilde{T}_i \le \rho_{\mathrm{A}}^{\mathrm{sym}} \otimes I_{\mathrm{B}},$$

where $\rho_{\mathrm{A}}^{\mathrm{sym}} := \frac{1}{n!} \sum_{\pi \in \mathfrak{S}_n} \mathtt{p}_{\mathrm{A}}(\pi)\, \rho_{\mathrm{A}}\, \mathtt{p}_{\mathrm{A}}(\pi)^\dagger$ is a mixed quantum state. Thus there exists a positive semidefinite linear operator $\widetilde{T}_\perp$ satisfying $\sum_i \widetilde{T}_i + \widetilde{T}_\perp = \rho_{\mathrm{A}}^{\mathrm{sym}} \otimes I_{\mathrm{B}}$. In fact, we can explicitly define

$$\widetilde{T}_\perp := \sum_{\substack{\lambda \vdash_{d_1 d_2} n \\ l(\lambda) > s}} P_\lambda (\rho_{\mathrm{A}}^{\mathrm{sym}} \otimes I_{\mathrm{B}}) P_\lambda = \frac{1}{r^n} \sum_i \sum_{\substack{\lambda \vdash_{d_1 d_2} n \\ l(\lambda) > s}} P_\lambda\, \mathrm{tr}_{\mathrm{anc}}(T_i)^{\mathrm{sym}}\, P_\lambda, \tag{21}$$

where $l(\lambda)$ denotes the number of rows of λ, P_λ denotes the orghotonal projector onto the subspace $\mathcal{P}_{\mathrm{AB},\lambda} \otimes \mathcal{Q}^{d_1 d_2}_{\mathrm{AB},\lambda}$, and $\mathrm{tr}_{\mathrm{anc}}(T_i)^{\mathrm{sym}} = \frac{1}{n!}\sum_{\pi\in\mathfrak{S}_n} \mathtt{p}_{\mathrm{AB}}(\pi)\ \mathrm{tr}_{\mathrm{anc}}(T_i)\ \mathtt{p}_{\mathrm{AB}}(\pi)^\dagger$. Therefore, $\{\widetilde{T}_i\}_i \cup \{\widetilde{T}_\perp\}$ is a parallel tester.

Item 2. Suppose $\mathcal{E} \in \mathbf{QChan}^r_{d_1,d_2}$ is a quantum channel. Let $\mathcal{V} \in \mathbf{Dilation}_r(\mathcal{E})$ be an arbitrary dilation, where $V : \mathcal{H}_\mathrm{A} \to \mathcal{H}_\mathrm{B} \otimes \mathcal{H}_{\mathrm{anc}}$, $\mathcal{H}_\mathrm{A} \cong \mathbb{C}^{d_1}$, $\mathcal{H}_\mathrm{B} \cong \mathbb{C}^{d_2}$, and $\mathcal{H}_{\mathrm{anc}} \cong \mathbb{C}^r$. We can see that

$$\mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(C_\mathcal{V}) = \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(|V\rangle\!\rangle\langle\!\langle V|) = C_\mathcal{E}.$$

Treating $|V\rangle\!\rangle$ as a bipartite vector in $(\mathcal{H}_\mathrm{A} \otimes \mathcal{H}_\mathrm{B}) \otimes \mathcal{H}_{\mathrm{anc}}$ and according to Lemma 3.11, we have $|V\rangle\!\rangle^{\otimes n} = \bigoplus_{\lambda\vdash_s n} |I_{\mathcal{P}_\lambda}\rangle\!\rangle \otimes |V_\lambda\rangle$ for $|I_{\mathcal{P}_\lambda}\rangle\!\rangle \in \mathcal{P}_{\mathrm{AB},\lambda} \otimes \mathcal{P}_{\mathrm{anc},\lambda}$ and $|V_\lambda\rangle \in \mathcal{Q}^{d_1d_2}_{\mathrm{AB},\lambda} \otimes \mathcal{Q}^r_{\mathrm{anc},\lambda}$. Thus,

$$\begin{aligned}\mathrm{tr}_{\mathrm{anc}}(|V\rangle\!\rangle\langle\!\langle V|^{\otimes n}) &= \mathrm{tr}_{\mathrm{anc}}\left(\bigoplus_{\lambda,\mu\vdash_s n} |I_{\mathcal{P}_\lambda}\rangle\!\rangle\langle\!\langle I_{\mathcal{P}_\mu}| \otimes |V_\lambda\rangle\langle V_\mu|\right)\\ &= \bigoplus_{\lambda\vdash_s n} \mathrm{tr}_{\mathcal{P}_{\mathrm{anc},\lambda}}\Big(|I_{\mathcal{P}_\lambda}\rangle\!\rangle\langle\!\langle I_{\mathcal{P}_\lambda}|\Big) \otimes \mathrm{tr}_{\mathcal{Q}^r_{\mathrm{anc},\lambda}}\Big(|V_\lambda\rangle\langle V_\lambda|\Big)\\ &= \bigoplus_{\lambda\vdash_s n} I_{\mathcal{P}_{\mathrm{AB},\lambda}} \otimes \mathrm{tr}_{\mathcal{Q}^r_{\mathrm{anc},\lambda}}\Big(|V_\lambda\rangle\langle V_\lambda|\Big), \end{aligned} \tag{22}$$

in which $\mathrm{tr}_{\mathrm{anc}}(\cdot)$ denotes the partial trace on all ancilla systems $\bigotimes_{j=1}^n \mathcal{H}_{\mathrm{anc},j}$. Also, we can see

$$C_\mathcal{E}^{\otimes n} = \bigoplus_{\lambda\vdash_{d_1d_2} n} I_{\mathcal{P}_{\mathrm{AB},\lambda}} \otimes C_{\mathcal{E},\lambda}, \tag{23}$$

for certain $\mathcal{C}_{\mathcal{E},\lambda} \in \mathcal{L}(\mathcal{Q}^{d_1d_2}_{\mathrm{AB},\lambda})$. Comparing Equation (22) with Equation (23), we find that $\mathrm{tr}_{\mathcal{Q}^r_{\mathrm{anc},\lambda}}(|V_\lambda\rangle\langle V_\lambda|) = C_{\mathcal{E},\lambda}$ for $\lambda \vdash_s n$, and also $C_{\mathcal{E},\lambda} = 0$ for those λ with more than s rows. This means,

$$\begin{aligned}\widetilde{T}_i \star C_\mathcal{E}^{\otimes n} &= \mathrm{tr}\left(\left(\bigoplus_{\lambda\vdash_s n} \frac{1}{\dim(\mathcal{P}_\lambda)\dim(\mathcal{Q}^r_\lambda)} \cdot I_{\mathcal{P}_{\mathrm{AB},\lambda}} \otimes \mathrm{tr}_{\mathcal{Q}^r_{\mathrm{anc},\lambda}}\Big(\langle\!\langle I_{\mathcal{P}_\lambda}|\overline{T}_i|I_{\mathcal{P}_\lambda}\rangle\!\rangle\Big)\right)^{\mathrm{T}} \cdot C_\mathcal{E}^{\otimes n}\right)\\ &= \sum_{\lambda\vdash_s n} \frac{1}{\dim(\mathcal{Q}^r_\lambda)} \mathrm{tr}\left(\mathrm{tr}_{\mathcal{Q}^r_{\mathrm{anc},\lambda}}\Big(\langle\!\langle I_{\mathcal{P}_\lambda}|\overline{T}_i|I_{\mathcal{P}_\lambda}\rangle\!\rangle\Big)^{\mathrm{T}} \cdot C_{\mathcal{E},\lambda}\right)\\ &= \sum_{\lambda\vdash_s n} \frac{1}{\dim(\mathcal{Q}^r_\lambda)} \mathrm{tr}\left(\mathrm{tr}_{\mathcal{Q}^r_{\mathrm{anc},\lambda}}\Big(\langle\!\langle I_{\mathcal{P}_\lambda}|\overline{T}_i^{\mathrm{T}}|I_{\mathcal{P}_\lambda}\rangle\!\rangle\Big) \cdot \mathrm{tr}_{\mathcal{Q}^r_{\mathrm{anc},\lambda}}\Big(|V_\lambda\rangle\langle V_\lambda|\Big)\right)\\ &= \overline{T}_i \star C_\mathcal{V}^{\otimes n} \qquad (24)\\ &= \mathop{\mathbb{E}}_{\mathcal{W}\sim\mathbf{Dilation}_r(\mathcal{E})}[T_i \star C_\mathcal{W}^{\otimes n}], \qquad (25)\end{aligned}$$

where Equation (24) is due to item 2 of Lemma 3.9 and Equation (25) is due to item 1 of Lemma 3.9. □

The following result about bipartite pure states is widely used in quantum information theory (see, e.g., [MH07]).

Lemma 3.11. *Suppose $|\psi\rangle \in \mathcal{H}_\mathrm{A} \otimes \mathcal{H}_\mathrm{B} \cong \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$ is a vector and define $s = \min\{d_1, d_2\}$. Then, $|\psi\rangle^{\otimes n}$ has the form*

$$|\psi\rangle^{\otimes n} = \bigoplus_{\lambda\vdash_s n} |I_{\mathcal{P}_\lambda}\rangle\!\rangle \otimes |\psi_\lambda\rangle,$$

where $|\psi_\lambda\rangle \in \mathcal{Q}^{d_1}_{\mathrm{A},\lambda} \otimes \mathcal{Q}^{d_2}_{\mathrm{B},\lambda}$, and $|I_{\mathcal{P}_\lambda}\rangle\!\rangle$ is the (unnormalized) maximally entangled state on the bipartite system $\mathcal{P}_{\mathrm{A},\lambda} \otimes \mathcal{P}_{\mathrm{B},\lambda}$ w.r.t. the Young's orthogonal basis.

Proof. First, the state $|\psi\rangle^{\otimes n}$ is invariant under the action of $\mathtt{p}_{\mathrm{A}}(\pi)\otimes\mathtt{p}_{\mathrm{B}}(\pi)$ for any $\pi\in\mathfrak{S}_n$. Due to the Schur-Weyl duality, we can see that

$$\begin{aligned}\frac{1}{n!}\sum_{\pi\in\mathfrak{S}_n}\mathtt{p}_{\mathrm{A}}(\pi)\otimes\mathtt{p}_{\mathrm{B}}(\pi) &= \frac{1}{n!}\sum_{\pi\in\mathfrak{S}_n}\bigoplus_{\substack{\lambda\vdash_{d_1}n\\ \mu\vdash_{d_2}n}}\mathtt{p}_{\mathrm{A},\lambda}(\pi)\otimes\mathtt{p}_{\mathrm{B},\mu}(\pi)\otimes I_{\mathcal{Q}^{d_1}_{\mathrm{A},\lambda}}\otimes I_{\mathcal{Q}^{d_2}_{\mathrm{B},\mu}}\\ &= \bigoplus_{\substack{\lambda\vdash_{d_1}n\\ \mu\vdash_{d_2}n}}\frac{1}{n!}\sum_{\pi\in\mathfrak{S}_n}\mathtt{p}^*_{\mathrm{A},\lambda}(\pi)\otimes\mathtt{p}_{\mathrm{B},\mu}(\pi)\otimes I_{\mathcal{Q}^{d_1}_{\mathrm{A},\lambda}}\otimes I_{\mathcal{Q}^{d_2}_{\mathrm{B},\mu}} \qquad (26)\\ &= \bigoplus_{\lambda\vdash_s n}\frac{1}{\dim(\mathcal{P}_\lambda)}|I_{\mathcal{P}_\lambda}\rangle\!\rangle\langle\!\langle I_{\mathcal{P}_\lambda}|\otimes I_{\mathcal{Q}^{d_1}_{\mathrm{A},\lambda}}\otimes I_{\mathcal{Q}^{d_2}_{\mathrm{B},\lambda}}, \qquad (27)\end{aligned}$$

where in Equation (26) the $(\cdot)^*$ denotes complex conjugate (w.r.t. the Young's orthogonal basis so that $\mathtt{p}_\lambda(\pi)$ is a real matrix [CSST10]), Equation (27) uses the fact that the only subspace invariant under $\mathtt{p}^*_{\mathrm{A},\lambda}(\pi)\otimes\mathtt{p}_{\mathrm{B},\mu}(\pi)$ for all $\pi\in\mathfrak{S}_n$ is spanned by $|I_{\mathcal{P}_\lambda}\rangle\!\rangle\in\mathcal{P}_{\mathrm{A},\lambda}\otimes\mathcal{P}_{\mathrm{B},\mu}$ when $\lambda=\mu$, and is the trivial space $\{0\}$ otherwise.[5] Thus, the vector $|\psi\rangle^{\otimes n}$ lies in the support of the projector $\frac{1}{n!}\sum_{\pi\in\mathfrak{S}_n}\mathtt{p}_{\mathrm{A}}(\pi)\otimes\mathtt{p}_{\mathrm{B}}(\pi)$, which, according to Equation (27), means $|\psi\rangle^{\otimes n}=\bigoplus_{\lambda\vdash_s n}|I_{\mathcal{P}_\lambda}\rangle\!\rangle\otimes|\psi_\lambda\rangle$. ☐

4 Hardness of isometry channel discrimination

In this section, we establish the difficulty of distinguishing between isometry channels with particular structures. In Section 4.1, we present a construction of a family of channels that are ε-close and prove that $\Omega(1/\varepsilon)$ queries are necessary to distinguish between the corresponding isometries. Moreover, in Section 4.2, we present a different family of channels, that are also ε-close, but $\Omega(1/\varepsilon^2)$ queries are needed to discriminate between the corresponding isometries. For ease of reference, we denote these constructions as Type I and Type II hard instances, respectively.

4.1 Type I: Heisenberg-scaling hardness

4.1.1 Definition

Let $D\geq d$ be two positive integers. We consider isometries with input dimension d and output dimension D.

Let $\varepsilon\in(0,1/2)$, $d\geq 2$ and $d':=2\left\lfloor\frac{d}{2}\right\rfloor$. Let $\{|1\rangle_{\mathrm{A}},\ldots,|d\rangle_{\mathrm{A}}\}$ and $\{|1\rangle_{\mathrm{B}},\ldots,|D\rangle_{\mathrm{B}}\}$ be two orthonormal bases of the Hilbert spaces $\mathcal{H}_{\mathrm{A}}\cong\mathbb{C}^d$ and $\mathcal{H}_{\mathrm{B}}\cong\mathbb{C}^D$, respectively. Define the linear operator $V_0:\mathcal{H}_{\mathrm{A}}\to\mathcal{H}_{\mathrm{B}}$ by

$$V_0:=\begin{cases}\sqrt{1-\varepsilon^2}\sum_{i=1}^{d'}|i\rangle_{\mathrm{B}}\langle i|_{\mathrm{A}}+|d\rangle_{\mathrm{B}}\langle d|_{\mathrm{A}}, & \text{if } d \text{ is odd}\\ \sqrt{1-\varepsilon^2}\sum_{i=1}^{d'}|i\rangle_{\mathrm{B}}\langle i|_{\mathrm{A}}, & \text{if } d \text{ is even.}\end{cases}$$

Define the linear operator $\Delta:\mathcal{H}_{\mathrm{A}}\to\mathcal{H}_{\mathrm{B}}$ by

$$\Delta:=\mathrm{i}\left(\sum_{i=1}^{\lfloor d/2\rfloor}|i\rangle_{\mathrm{B}}\langle i|_{\mathrm{A}}-\sum_{i=\lfloor d/2\rfloor+1}^{d'}|i\rangle_{\mathrm{B}}\langle i|_{\mathrm{A}}\right),$$

[5]We can see this by considering the representation isomorphism $\mathcal{P}^*_{\mathrm{A},\lambda}\otimes\mathcal{P}_{\mathrm{B},\mu}\overset{\mathfrak{S}_n}{\cong}\mathcal{L}(\mathcal{P}_{\mathrm{A},\lambda},\mathcal{P}_{\mathrm{B},\mu})$. Due to Schur's lemma, any linear operator in $\mathcal{L}(\mathcal{P}_{\mathrm{A},\lambda},\mathcal{P}_{\mathrm{B},\mu})$ that commutes with the action of π must be proportional to the identity operator when $\lambda=\mu$ and 0 otherwise.

where i denotes the imaginary unit.

Let $U \in \mathbb{U}_{d'}$ be a unitary. We define the isometry $V_{\varepsilon,U} : \mathcal{H}_A \to \mathcal{H}_B$ by

$$V_{\varepsilon,U} \coloneqq V_0 + \varepsilon U \Delta U^\dagger. \tag{28}$$

Here, we abuse the notation and assume that U can act either on $\mathcal{H}_A$ or $\mathcal{H}_B$ (we identify $|i\rangle_A$ with $|i\rangle_B$ for $i \in [d]$ and U acts trivially on $|d'+1\rangle_B, \ldots, |D\rangle_B$).

Note that $V_{\varepsilon,U}$ is an isometry as one can verify:

$$\begin{aligned} V_{\varepsilon,U}^\dagger V_{\varepsilon,U} &= V_0^\dagger V_0 + \varepsilon \left(V_0^\dagger U \Delta U^\dagger + U \Delta^\dagger U^\dagger V_0 \right) + \varepsilon^2 U \Delta^\dagger \Delta U^\dagger \\ &= V_0^\dagger V_0 + \varepsilon^2 U \Delta^\dagger \Delta U^\dagger \\ &= I_A. \end{aligned}$$

Furthermore, we observe that $\langle\!\langle V_0 | \Delta \rangle\!\rangle = \mathrm{tr}(V_0^\dagger \Delta) = 0$.

4.1.2 Hardness of the discrimination problem

We then obtain the following main result of this section.

Theorem 4.1. *Let $C > 0$ be a constant. Suppose $d \geq 8/C$. Let $\mathcal{N} \subseteq \{V_{\varepsilon,U} \,|\, U \in \mathbb{U}_{2\lfloor \frac{d}{2} \rfloor}\}$ with $V_{\varepsilon,U}$ defined in Equation (28) of cardinality $|\mathcal{N}| \geq \exp(Cd^2)$. Then, at least $n \geq \min\{\frac{C^{3/2}}{21}, \frac{1}{\sqrt{2e^5}}\} \cdot \frac{d^2}{\varepsilon}$ queries are needed for any algorithm that can distinguish between the isometries in $\mathcal{N}$ with success probability at least $2/3$.*

Proof. Consider an algorithm that distinguishes between the isometries in $\mathcal{N}$ using n queries. We may assume $n \leq \frac{1}{\sqrt{2e^5}} d^2/\varepsilon$ as otherwise the theorem is proven.

Recall that each isometry V in $\mathcal{N}$ can be written as

$$V = V_0 + \varepsilon U \Delta U^\dagger$$

for some unitary $U \in \mathbb{U}_{d'}$.

Let $\{T_V\}_{V \in \mathcal{N}}$ be a sequential tester that describes the algorithm for distinguishing between the isometries in the set $\mathcal{N}$. Using the bound on the cardinality of the family $\mathcal{N}$, we can bound the success probability as follows

$$\begin{aligned} \mathbb{P}[\text{success}] &= \frac{1}{|\mathcal{N}|} \cdot \sum_{V \in \mathcal{N}} T_V \star |V\rangle\!\rangle\langle\!\langle V|^{\otimes n} \\ &\leq \exp(-Cd^2) \cdot \sum_{V \in \mathcal{N}} T_V \star |V\rangle\!\rangle\langle\!\langle V|^{\otimes n}, \end{aligned}$$

where $|V\rangle\!\rangle\langle\!\langle V|^{\otimes n}$ is an n-comb on $(\mathcal{H}_{A,1}, \mathcal{H}_{B,1}, \ldots, \mathcal{H}_{A,n}, \mathcal{H}_{B,n})$, and $\mathcal{H}_{A,i}$ and $\mathcal{H}_{B,i}$ denote the input and output spaces of the i-th query to V, respectively.

Observe that for $V \in \mathcal{N}$, we can express $|V\rangle\!\rangle^{\otimes n}$ as follows

$$\begin{aligned} |V\rangle\!\rangle^{\otimes n} &= \left(|V_0\rangle\!\rangle + \varepsilon |U \Delta U^\dagger\rangle\!\rangle \right)^{\otimes n} \\ &= \sum_{i=0}^{n} \varepsilon^i \sum_{\substack{S \subseteq [n] \\ |S| = i}} |V_0\rangle\!\rangle^{\otimes [n] \setminus S} \otimes |U \Delta U^\dagger\rangle\!\rangle^{\otimes S} \end{aligned}$$

$$= \sum_{i=0}^{n} \varepsilon^i \left(U \otimes U^*\right)^{\otimes n} \sum_{\substack{S \subseteq [n] \\ |S| = i}} |\gamma_S\rangle, \tag{29}$$

where U^* acts on on $\mathcal{H}_{\mathrm{A},i}$ and U acts on $\mathcal{H}_{\mathrm{B},i}$ and for any $S \subseteq [n]$, we define the vector

$$|\gamma_S\rangle := |V_0\rangle\!\rangle^{\otimes [n] \setminus S} \otimes |\Delta\rangle\!\rangle^{\otimes S}. \tag{30}$$

Note that, since $\langle\!\langle V_0|\Delta\rangle\!\rangle = \mathrm{tr}(V_0^\dagger \Delta) = 0$, the vectors $\{|\gamma_S\rangle\}_{S \subseteq [n]}$ are pairwise orthogonal.

Next, for $S \subseteq [n]$, we introduce the operator Γ_S

$$\Gamma_S := \mathop{\mathbb{E}}_{U \sim \mathbb{U}_{d'}} \left[\left(U \otimes U^*\right)^{\otimes n} |\gamma_S\rangle\langle\gamma_S| \left(U^\dagger \otimes U^{\mathrm{T}}\right)^{\otimes n} \right]. \tag{31}$$

Similarly, $\{\mathrm{supp}(\Gamma_S)\}_{S \subseteq [n]}$ are pairwise orthogonal. To see this, we observe that $U \otimes U^*$ fixes $|V_0\rangle\!\rangle$ and $U \otimes U^* |\Delta\rangle\!\rangle$ remains orthogonal to $|V_0\rangle\!\rangle$ for any $U \in \mathbb{U}_{d'}$.

By Lemma 4.3, there exist a set of positive numbers $\{\lambda_S \,|\, S \subseteq [n]\}$ satisfying for any $V \in \mathcal{N}$,

$$\sum_{S \subseteq [n]} \lambda_S \le 3d^4 \exp\left(\sqrt[3]{54 n^2 \varepsilon^2 d^2}\right),$$
$$|V\rangle\!\rangle\langle\!\langle V|^{\otimes n} \le \sum_{S \subseteq [n]} \lambda_S \Gamma_S.$$

Hence the success probability can be bounded as

$$\begin{aligned}
\mathbb{P}[\text{success}] &\le \exp(-Cd^2) \cdot \sum_{V \in \mathcal{N}} T_V \star |V\rangle\!\rangle\langle\!\langle V|^{\otimes n} \\
&\le \exp(-Cd^2) \cdot \sum_{V \in \mathcal{N}} T_V \star \sum_{S \subseteq [n]} \lambda_S \Gamma_S \\
&= \exp(-Cd^2) \cdot \sum_{T \subseteq [n]} \lambda_T \cdot \sum_{V \in \mathcal{N}} T_V \star \sum_{S \subseteq [n]} \frac{\lambda_S}{\sum_{T \subseteq [n]} \lambda_T} \Gamma_S \\
&\le \exp(-Cd^2) \cdot \sum_{S \subseteq [n]} \lambda_S, \qquad (32) \\
&\le \exp(-Cd^2) \cdot 3d^4 \exp\left(\sqrt[3]{54 n^2 \varepsilon^2 d^2}\right),
\end{aligned}$$

where Equation (32) uses the fact that

- $\sum_{V \in \mathcal{N}} T_V$ forms an $(n+1)$-comb with input and output dimensions 1, and
- $\sum_{S \subseteq [n]} \frac{\lambda_S}{\sum_{T \subseteq [n]} \lambda_T} \Gamma_S$ forms a probabilistic n-comb. Here, we use the fact that each Γ_S is a probabilistic comb (see Lemma 4.2) and their convex combination is also a probabilistic comb,

and the contraction of an $(n+1)$-comb with input and output dimensions 1 and a probabilistic n-comb is at most 1.

Since the algorithm succeeds with probability $\mathbb{P}[\text{success}] \ge 2/3$, we obtain that

$$\begin{aligned}
54 n^2 \varepsilon^2 d^2 &\ge \left(Cd^2 - 4\ln(d) + \ln(2/9)\right)^3 \\
&\ge \left(Cd^2 - 4d\right)^3 \ge \frac{C^3}{8} d^6,
\end{aligned}$$

where we used $\ln(d) \le d-1$, $\ln(9/2) \le 4$ and $d \ge 8/C$. Finally, we deduce

$$n \ge \frac{C^{3/2}}{12\sqrt{3}} \cdot \frac{d^2}{\varepsilon} > \frac{C^{3/2}}{21} \cdot \frac{d^2}{\varepsilon}.$$

$\square$

Technical lemmas. For $S \subseteq [n]$, recall the definition of $|\gamma_S\rangle$ (see Equation (30)) and Γ_S (see Equation (31)).

Lemma 4.2. *$|\gamma_S\rangle\langle\gamma_S|$ and Γ_S are probabilistic n-combs.*

Proof. Observe that

$$\mathrm{tr}_{\mathcal{H}_\mathrm{B}}(|V_0\rangle\rangle\langle\langle V_0|) = \begin{cases} (1-\varepsilon^2)\sum_{i=1}^{d'}|i\rangle\langle i|_\mathrm{A} + |d\rangle\langle d|_\mathrm{A}, & \text{if } d \text{ is odd} \\ (1-\varepsilon^2)\sum_{i=1}^{d'}|i\rangle\langle i|_\mathrm{A}, & \text{if } d \text{ is even} \end{cases},$$

so $\mathrm{tr}_{\mathcal{H}_\mathrm{B}}(|V_0\rangle\rangle\langle\langle V_0|) \le I_\mathrm{A}$. Moreover, we have that

$$\mathrm{tr}_{\mathcal{H}_\mathrm{B}}(|\Delta\rangle\langle\Delta|) = \sum_{i=1}^{d'}|i\rangle\langle i|_\mathrm{A} \le I_\mathrm{A}.$$

hence $|\gamma_S\rangle$ is a probabilistic n-comb.

The similar reasoning shows that $(U\otimes U^*)^{\otimes n}|\gamma_S\rangle\langle\gamma_S|(U^\dagger\otimes U^\mathrm{T})^{\otimes n}$ is also a probabilistic n-comb for any $U \in \mathbb{U}_{d'}$. Finally, we deduce that Γ_S is also a probabilistic n-comb since it is a convex combination of probabilistic combs. $\square$

Lemma 4.3. *Suppose $d \ge 2$ and $n \le \frac{1}{\sqrt{2e^5}}d^2/\varepsilon$. There exists a set of positive numbers $\{\lambda_S \mid S \subseteq [n]\}$ such that for any $V \in \mathcal{N}$,*

$$\sum_{S\subseteq[n]} \lambda_S \le 3d^4 \exp\left(\sqrt[3]{54n^2\varepsilon^2d^2}\right),$$

$$|V\rangle\rangle\langle\langle V|^{\otimes n} \le \sum_{S\subseteq[n]} \lambda_S\Gamma_S. \tag{33}$$

Proof. We remark that, because of Equation (29), $|V\rangle\rangle^{\otimes n}$ is contained in $\bigoplus_{S\subseteq[n]}\mathrm{supp}(\Gamma_S)$. Moreover, by Fact 7.7, to establish Equation (33), it is sufficient to show that

$$\sum_{S\subseteq[n]} \frac{1}{\lambda_S}\mathrm{tr}\left(\Gamma_S^{-1}|V\rangle\rangle\langle\langle V|^{\otimes n}\right) \le 1. \tag{34}$$

To this end, we bound the left hand side of Equation (34) as

$$\sum_{S\subseteq[n]} \frac{1}{\lambda_S}\mathrm{tr}\left(\Gamma_S^{-1}|V\rangle\rangle\langle\langle V|^{\otimes n}\right) = \sum_{S\subseteq[n]} \frac{1}{\lambda_S}\varepsilon^{2|S|}\,\mathrm{tr}\left(\Gamma_S^{-1}(U\otimes U^*)^{\otimes n}|\gamma_S\rangle\langle\gamma_S|(U^\dagger\otimes U^\mathrm{T})^{\otimes n}\right) \tag{35}$$

$$= \sum_{S\subseteq[n]} \frac{1}{\lambda_S}\varepsilon^{2|S|}\,\mathrm{tr}\left(\Gamma_S^{-1}|\gamma_S\rangle\langle\gamma_S|\right) \tag{36}$$

$$\le \sum_{S\subseteq[n]} \frac{1}{\lambda_S}\varepsilon^{2|S|}\binom{d^2+|S|-2}{|S|}. \tag{37}$$

Here, Equation (35) follows from Equation (29) and the fact that $(U \otimes U^*)^{\otimes n}|\gamma_S\rangle \in \mathrm{supp}(\Gamma_S)$, Equation (36) uses the fact that Γ_S commutes with $(U \otimes U^*)^{\otimes n}$, and Equation (37) is an application of Lemma 7.1 where $|\gamma_S\rangle$ is viewed as a vector in the space

$$\mathrm{span}\left(\left\{|V_0\rangle\!\rangle^{\otimes [n]\setminus S} \otimes |\psi\rangle^{\otimes S} \;\middle|\; |\psi\rangle \in \mathrm{span}\{|1\rangle_{\mathrm{A}}, \dots, |d\rangle_{\mathrm{A}}\} \otimes \mathrm{span}\{|1\rangle_{\mathrm{B}}, \dots, |d\rangle_{\mathrm{B}}\},\ |\psi\rangle \in |V_0\rangle\!\rangle^{\perp}\right\}\right),$$

which is invariant under the action $(U \otimes U^*)^{\otimes n}$. This space has dimension $\binom{d^2+|S|-2}{|S|}$ since it is isomorphic to the symmetric space $\vee^{|S|}\mathbb{C}^{d^2-1}$ [Har13].

For $S \subseteq [n]$, we set $\lambda_S = \mu_i / \binom{n}{i}$ where $i = |S|$ for some positive number μ_i. Then, we observe $\sum_{S \subseteq [n]} \lambda_S = \sum_{i=0}^n \mu_i$ and Equation (37) can be written as

$$\sum_{i=0}^{n} \frac{1}{\mu_i} \binom{n}{i}^2 \varepsilon^{2i} \binom{d^2+i-2}{i}. \tag{38}$$

Our goal is to look for positive numbers $\mu_0, \dots, \mu_n$ such that $\sum_{i=0}^n \mu_i$ is small and Equation (38) is upper bounded by 1. For this, Equation (38) can be bounded using Fact 7.6 as

$$\begin{aligned}
(38) &\le \sum_{i=0}^{n} \frac{1}{\mu_i} \exp\left(2nH\left(\frac{i}{n}\right) + 2i\ln(\varepsilon) + (d^2+i)H\left(\frac{i}{d^2+i}\right)\right) \\
&= \sum_{i=0}^{n} \frac{1}{\mu_i} \exp\left(-2i\ln\left(\frac{i}{n\varepsilon}\right) - 2(n-i)\ln\left(\frac{n-i}{n}\right) + i\ln\left(1+\frac{d^2}{i}\right) + d^2\ln\left(1+\frac{i}{d^2}\right)\right) \\
&\le \sum_{i=0}^{n} \frac{1}{\mu_i} \exp\left(-i\ln\left(\frac{i^2}{n^2\varepsilon^2}\right) + i\ln\left(1+\frac{d^2}{i}\right) + 3i\right),
\end{aligned} \tag{39}$$

where the last inequality uses

$$2(n-i)\ln\left(\frac{n}{n-i}\right) \le 2(n-i)\left(\frac{n}{n-i} - 1\right) = 2i \quad \text{and} \quad d^2\ln\left(1+\frac{i}{d^2}\right) \le d^2\frac{i}{d^2} = i.$$

To further bound Equation (39), we consider two cases depending on i:

- When $i < d^2$, we have

$$\begin{aligned}
-i\ln\left(\frac{i^2}{n^2\varepsilon^2}\right) + i\ln\left(1+\frac{d^2}{i}\right) + 3i &\le -i\ln\left(\frac{i^2}{n^2\varepsilon^2}\right) + i\ln\left(\frac{2d^2}{i}\right) + 3i \\
&= 3i\ln\left(\frac{\sqrt[3]{2e^3n^2\varepsilon^2d^2}}{i}\right) \\
&\le \sqrt[3]{54n^2\varepsilon^2d^2},
\end{aligned}$$

  where we used Fact 7.8 in the last inequality.

- When $i \ge d^2$, since we assumed that $n \le \frac{1}{\sqrt{2e^5}} d^2/\varepsilon$ we have $i \ge \sqrt{2e^5}n\varepsilon$, thus

$$-i\ln\left(\frac{i^2}{n^2\varepsilon^2}\right) + i\ln\left(1+\frac{d^2}{i}\right) + 3i \le -i\ln(2e^5) + i\ln(2) + 3i = -i\ln\left(2e^5/2e^3\right) = -2i.$$

Hence, the choice

$$\mu_i = \begin{cases} 2d^2 \exp\left(\sqrt[3]{54n^2\varepsilon^2 d^2}\right), & \text{if } i < d^2 \\ \exp(-i), & \text{if } i \geq d^2, \end{cases}$$

ensures that Equation (39) can be upper bounded by

$$\begin{aligned} (39) &\leq \sum_{i<d^2} \frac{1}{\mu_i} \exp\left(\sqrt[3]{54n^2\varepsilon^2 d^2}\right) + \sum_{i\geq d^2} \frac{1}{\mu_i} \exp(-2i) \\ &\leq \frac{1}{2} + \sum_{i\geq d^2} \exp(-i) \leq \frac{1}{2} + \exp(-d^2)\frac{e}{e-1} < 1, \end{aligned}$$

where we used that $d^2 \geq 2$. Finally, we have that

$$\begin{aligned} \sum_{i=0}^{n} \mu_i &\leq 2d^4 \exp\left(\sqrt[3]{54n^2\varepsilon^2 d^2}\right) + \exp(-d^2)\frac{e}{e-1} \\ &< 2d^4 \exp\left(\sqrt[3]{54n^2\varepsilon^2 d^2}\right) + \frac{1}{2} < 3d^4 \exp\left(\sqrt[3]{54n^2\varepsilon^2 d^2}\right), \end{aligned}$$

which concludes the proof. □

4.2 Type II: classical-scaling hardness

4.2.1 Definition

Let $D > d$ be two positive integers. We consider isometries with input dimension d and output dimension D.

Let $\varepsilon \in (0, 1/2)$ and $d' \leq \min\{d, D-d\}$ be a positive number. Let $\{|1\rangle_{\mathrm{A}}, \ldots, |d\rangle_{\mathrm{A}}\}$ and $\{|1\rangle_{\mathrm{B}}, \ldots, |D\rangle_{\mathrm{B}}\}$ be two (arbitrary) orthonormal bases of the Hilbert spaces $\mathcal{H}_{\mathrm{A}} \cong \mathbb{C}^d$ and $\mathcal{H}_{\mathrm{B}} \cong \mathbb{C}^D$, respectively. Define the linear operator $V_0 : \mathcal{H}_{\mathrm{A}} \to \mathcal{H}_{\mathrm{B}}$ by

$$V_0 \coloneqq \sqrt{1-\varepsilon^2} \sum_{i=1}^{d'} |i\rangle_{\mathrm{B}}\langle i|_{\mathrm{A}} + \sum_{i=d'+1}^{d} |i\rangle_{\mathrm{B}}\langle i|_{\mathrm{A}},$$

and the linear operator $\Delta : \mathcal{H}_{\mathrm{A}} \to \mathcal{H}_{\mathrm{B}}$ by

$$\Delta \coloneqq \sum_{i=1}^{d'} |d+i\rangle_{\mathrm{B}}\langle i|_{\mathrm{A}}.$$

Let $U \in \mathbb{U}_{D-d}$ be a unitary. We define the isometry $V_{\varepsilon,U} : \mathcal{H}_{\mathrm{A}} \to \mathcal{H}_{\mathrm{B}}$ as

$$V_{\varepsilon,U} \coloneqq V_0 + \varepsilon U \Delta. \tag{40}$$

Here, $U \in \mathbb{U}_{D-d}$ acts on the subspace spanned by $\{|d+1\rangle_{\mathrm{B}}, \ldots, |D\rangle_{\mathrm{B}}\}$.

It can be easily checked that $V_{\varepsilon,U}$ is an isometry:

$$V_{\varepsilon,U}^\dagger V_{\varepsilon,U} = (1-\varepsilon^2) \sum_{i=1}^{d'} |i\rangle_{\mathrm{A}}\langle i|_{\mathrm{A}} + \sum_{i=d'+1}^{d} |i\rangle_{\mathrm{A}}\langle i|_{\mathrm{A}} + \varepsilon^2 \sum_{i=1}^{d'} |i\rangle_{\mathrm{A}}\langle i|_{\mathrm{A}} = I_{\mathrm{A}}.$$

Observe that the images of V_0 and $U\Delta$ are orthogonal.

4.2.2 Hardness of the discrimination problem

We then obtain the following main result of this section.

Theorem 4.4. *Let $C > 0$ be a constant. Suppose $D - d \geq 64/C^2$. Let $\mathcal{N} \subseteq \{V_{\varepsilon,U} \,|\, U \in \mathbb{U}_{D-d}\}$ with $V_{\varepsilon,U}$ defined in Equation (40) of cardinality $|\mathcal{N}| \geq \exp(C(D-d)\min\{d, D-d\})$. Then, at least $n \geq \min\{\frac{C^2}{32}, \frac{1}{2e^4}\} \cdot \frac{(D-d)\min\{d,D-d\}}{\varepsilon^2}$ queries are needed for any algorithm that can distinguish between the isometries in $\mathcal{N}$ with success probability at least $2/3$.*

Proof. Let $d_{\min} \coloneqq \min\{d, D-d\}$. Consider an algorithm that distinguishes between the isometries in $\mathcal{N}$ using n queries. We may assume $n \leq \frac{1}{2e^4}(D-d)d_{\min}/\varepsilon^2$ as otherwise the theorem is proven.

Recall that each isometry V in $\mathcal{N}$ is of the form

$$V = V_0 + \varepsilon U \Delta,$$

for some unitary $U \in \mathbb{U}_{D-d}$.

Let $\{T_V\}_{V \in \mathcal{N}}$ be a sequential tester that describes the algorithm for distinguishing between the isometries in the set $\mathcal{N}$. Using the bound on the cardinality of the family $\mathcal{N}$, we can bound the success probability as follows

$$\begin{aligned}\mathbb{P}[\text{success}] &= \frac{1}{|\mathcal{N}|} \cdot \sum_{V \in \mathcal{N}} T_V \star |V\rangle\!\rangle\langle\!\langle V|^{\otimes n} \\ &\leq \exp(-C(D-d)d_{\min}) \cdot \sum_{V \in \mathcal{N}} T_V \star |V\rangle\!\rangle\langle\!\langle V|^{\otimes n}.\end{aligned}$$

where $|V\rangle\!\rangle\langle\!\langle V|^{\otimes n}$ is an n-comb on $(\mathcal{H}_{\mathrm{A},1}, \mathcal{H}_{\mathrm{B},1}, \ldots, \mathcal{H}_{\mathrm{A},n}, \mathcal{H}_{\mathrm{B},n})$, and $\mathcal{H}_{\mathrm{A},i}$ and $\mathcal{H}_{\mathrm{B},i}$ denote the input and output spaces of the i-th query to V, respectively.

Observe that for $V \in \mathcal{N}$, we can express $|V\rangle\!\rangle^{\otimes n}$ as follows

$$\begin{aligned}|V\rangle\!\rangle^{\otimes n} &= (|V_0\rangle\!\rangle + \varepsilon |U\Delta\rangle\!\rangle)^{\otimes n} \\ &= \sum_{i=0}^{n} \varepsilon^i \sum_{\substack{S \subseteq [n] \\ |S| = i}} |V_0\rangle\!\rangle^{\otimes [n] \setminus S} \otimes |U\Delta\rangle\!\rangle^{\otimes S} \\ &= \sum_{i=0}^{n} \varepsilon^i \sqrt{\binom{n}{i}} U^{\otimes n} |\gamma_i\rangle \end{aligned} \tag{41}$$

where $U^{\otimes n}$ acts as $(I_d \oplus U)^{\otimes n}$ on $\bigotimes_{i=1}^n \mathcal{H}_{\mathrm{B},i}$ and for $i \in \{0, 1, \ldots, n\}$, we define the vector

$$|\gamma_i\rangle \coloneqq \frac{1}{\sqrt{\binom{n}{i}}} \sum_{\substack{S \subseteq [n] \\ |S| = i}} |V_0\rangle\!\rangle^{\otimes [n] \setminus S} \otimes |\Delta\rangle\!\rangle^{\otimes S}. \tag{42}$$

Observe that, since $\langle\!\langle V_0 | \Delta \rangle\!\rangle = \mathrm{tr}(V_0^\dagger \Delta) = 0$, the vectors $\{|\gamma_i\rangle\}_{i \in [n]}$ are pairwise orthogonal. Next, for $i \in [n]$, we introduce the operator Γ_i

$$\Gamma_i \coloneqq \mathop{\mathbb{E}}_{U \sim \mathbb{U}_{D-d}} \left[U^{\otimes n} |\gamma_i\rangle\langle\gamma_i| U^{\dagger \otimes n} \right]. \tag{43}$$

We observe that $\{\mathrm{supp}(\Gamma_i)\}_{i \in [n]}$ are also pairwise orthogonal. To see this, we note that the number of $|V_0\rangle\!\rangle$ in $|\gamma_i\rangle$ is exactly $n - i$, and $|V_0\rangle\!\rangle$ is orthogonal to $U|\Delta\rangle\!\rangle$ for any $U \in \mathbb{U}_{D-d}$.

By Lemma 4.6, there exists a set of positive numbers $\{\lambda_i\}_{i=0}^n$ satisfying for any $V \in \mathcal{N}$,

$$\sum_{i=0}^{n} \lambda_i \leq 4(D-d)^2 d_{\min}^2 \exp\left(\sqrt{8n\varepsilon^2(D-d)d_{\min}}\right),$$
$$|V\rangle\!\rangle\langle\!\langle V|^{\otimes n} \leq \sum_{i=0}^{n} \lambda_i \Gamma_i.$$

Hence, the success probability can be bounded as

$$\begin{aligned}
\mathbb{P}[\text{success}] &\leq \exp(-C(D-d)d_{\min}) \cdot \sum_{V\in\mathcal{N}} T_V \star |V\rangle\!\rangle\langle\!\langle V|^{\otimes n} \\
&\leq \exp(-C(D-d)d_{\min}) \cdot \sum_{V\in\mathcal{N}} T_V \star \sum_{j=0}^{n} \lambda_j \Gamma_j \\
&= \exp(-C(D-d)d_{\min}) \cdot \sum_{k=0}^{n} \lambda_k \cdot \sum_{V\in\mathcal{N}} T_V \star \sum_{j=0}^{n} \frac{\lambda_j}{\sum_{k=0}^{n}\lambda_k} \Gamma_j \\
&\leq \exp(-C(D-d)d_{\min}) \cdot \sum_{i=0}^{n} \lambda_i \qquad (44) \\
&\leq \exp(-C(D-d)d_{\min}) \cdot 4(D-d)^2 d_{\min}^2 \exp\left(\sqrt{8n\varepsilon^2(D-d)d_{\min}}\right),
\end{aligned}$$

where Equation (44) uses the fact that

- $\sum_{V\in\mathcal{N}} T_V$ forms an $(n+1)$-comb with input and output dimensions 1, and
- $\sum_{j=0}^{n} \frac{\lambda_j}{\sum_{k=0}^{n}\lambda_k} \Gamma_j$ forms a probabilistic n-comb. Here, we use the facts that each Γ_i are probabilistic combs (see Lemma 4.5) and their convex combination is also a probabilistic comb,

and the contraction of an $(n+1)$-comb with input and output dimensions 1 and a probabilistic n-comb is at most 1.

For ease of notation, we denote by $Q = (D-d)d_{\min}$. Since the algorithm succeeds with probability $\mathbb{P}[\text{success}] \geq 2/3$, we obtain that

$$\begin{aligned}
\sqrt{8n\varepsilon^2 Q} &\geq CQ - 2\ln(Q) + \ln(1/6) \\
&\geq CQ - 4\sqrt{Q} \geq \frac{C}{2} Q,
\end{aligned}$$

where we used $\ln(\sqrt{Q}) \leq \sqrt{Q} - 1$, $\ln(6) \leq 2$ and $Q = (D-d)d_{\min} \geq D - d \geq 64/C^2$. Finally, we conclude

$$n \geq \frac{C^2}{32} \cdot \frac{Q}{\varepsilon^2} = \frac{C^2}{32} \cdot \frac{(D-d)d_{\min}}{\varepsilon^2}.$$

□

Technical lemmas. For $i \in [n]$, recall the definition of $|\gamma_i\rangle$ (see Equation (42)) and Γ_i (see Equation (43)).

Lemma 4.5. *$|\gamma_i\rangle\langle\gamma_i|$ and Γ_i are probabilistic n-combs.*

Proof. Denote by $|\gamma_i^n\rangle$ the state $|\gamma_i\rangle$ defined in Equation (42). We prove, by induction on n, that for all $i \le n$, $|\gamma_i^n\rangle\!\langle\gamma_i^n|$ is a probabilistic n-comb.

First, we observe that

$$\begin{aligned} \mathrm{tr}_{\mathcal{H}_\mathrm{B}}(|V_0\rangle\!\rangle\!\langle\!\langle V_0|) &= V_0^\mathrm{T} V_0 \le I_\mathrm{A}, \qquad & \mathrm{tr}_{\mathcal{H}_\mathrm{B}}(|\Delta\rangle\!\rangle\!\langle\!\langle V_0|) &= \Delta^\mathrm{T} V_0 = 0, \\ \mathrm{tr}_{\mathcal{H}_\mathrm{B}}(|V_0\rangle\!\rangle\!\langle\!\langle \Delta|) &= V_0^\mathrm{T} \Delta = 0, \qquad & \mathrm{tr}_{\mathcal{H}_\mathrm{B}}(|\Delta\rangle\!\rangle\!\langle\!\langle \Delta|) &= \Delta^\mathrm{T} \Delta \le I_\mathrm{A}. \end{aligned} \tag{45}$$

For $i = n$, we have that $|\gamma_n^n\rangle\!\langle\gamma_n^n| = |\Delta\rangle\!\rangle\!\langle\!\langle\Delta|^{\otimes n}$ is a probabilistic n-comb because $\mathrm{tr}_{\mathcal{H}_\mathrm{B}}(|\Delta\rangle\!\rangle\!\langle\!\langle\Delta|) \le I_\mathrm{A}$. Similarly, for $i = 0$, we have that $|\gamma_0^n\rangle\!\langle\gamma_0^n| = |V_0\rangle\!\rangle\!\langle\!\langle V_0|^{\otimes n}$ is a probabilistic n-comb because $\mathrm{tr}_{\mathcal{H}_\mathrm{B}}(|V_0\rangle\!\rangle\!\langle\!\langle V_0|) \le I_\mathrm{A}$. Thus the hypothesis holds automatically for the case $n = 1$.

Let $n \ge 2$ and suppose that $|\gamma_i^{n-1}\rangle\!\langle\gamma_i^{n-1}|$ is a probabilistic $(n-1)$-comb for all $i \le n-1$. Now we prove $|\gamma_i^n\rangle\!\langle\gamma_i^n|$ is a probabilistic n-comb for any $i \le n$.

Let $0 < i < n$ (since the cases $i = 0$ and $i = n$ are already proved). We can express $|\gamma_i^n\rangle$ as follows

$$|\gamma_i^n\rangle = \sqrt{\frac{\binom{n-1}{i}}{\binom{n}{i}}} |V_0\rangle\!\rangle \otimes |\gamma_i^{n-1}\rangle + \sqrt{\frac{\binom{n-1}{i-1}}{\binom{n}{i}}} |\Delta\rangle\!\rangle \otimes |\gamma_{i-1}^{n-1}\rangle.$$

Hence,

$$\begin{aligned} \mathrm{tr}_{\mathcal{H}_{\mathrm{B},n}}(|\gamma_i^n\rangle\!\langle\gamma_i^n|) = {} & \frac{\binom{n-1}{i}}{\binom{n}{i}} \mathrm{tr}_{\mathcal{H}_\mathrm{B}}(|V_0\rangle\!\rangle\!\langle\!\langle V_0|) \otimes |\gamma_i^{n-1}\rangle\!\langle\gamma_i^{n-1}| + \frac{\binom{n-1}{i-1}}{\binom{n}{i}} \mathrm{tr}_{\mathcal{H}_\mathrm{B}}(|\Delta\rangle\!\rangle\!\langle\!\langle \Delta|) \otimes |\gamma_{i-1}^{n-1}\rangle\!\langle\gamma_{i-1}^{n-1}| \\ & + \frac{\sqrt{\binom{n-1}{i}\binom{n-1}{i-1}}}{\binom{n}{i}} \Big(\mathrm{tr}_{\mathcal{H}_\mathrm{B}}(|V_0\rangle\!\rangle\!\langle\!\langle \Delta|) \otimes |\gamma_i^{n-1}\rangle\!\langle\gamma_{i-1}^{n-1}| + \mathrm{tr}_{\mathcal{H}_\mathrm{B}}(|\Delta\rangle\!\rangle\!\langle\!\langle V_0|) \otimes |\gamma_{i-1}^{n-1}\rangle\!\langle\gamma_i^{n-1}| \Big) \\ \le {} & \frac{\binom{n-1}{i}}{\binom{n}{i}} I_\mathrm{A} \otimes |\gamma_i^{n-1}\rangle\!\langle\gamma_i^{n-1}| + \frac{\binom{n-1}{i-1}}{\binom{n}{i}} I_\mathrm{A} \otimes |\gamma_{i-1}^{n-1}\rangle\!\langle\gamma_{i-1}^{n-1}|, \end{aligned}$$

where we use Equation (45) in the last inequality. By induction hypothesis, we have that $|\gamma_i^{n-1}\rangle\!\langle\gamma_i^{n-1}|$ and $|\gamma_{i-1}^{n-1}\rangle\!\langle\gamma_{i-1}^{n-1}|$ are probabilistic $(n-1)$-combs. Since $\binom{n-1}{i} + \binom{n-1}{i-1} = \binom{n}{i}$, we deduce that

$$\frac{\binom{n-1}{i}}{\binom{n}{i}} |\gamma_i^{n-1}\rangle\!\langle\gamma_i^{n-1}| + \frac{\binom{n-1}{i-1}}{\binom{n}{i}} |\gamma_{i-1}^{n-1}\rangle\!\langle\gamma_{i-1}^{n-1}|$$

is a probabilistic $(n-1)$-comb. Therefore, $|\gamma_i^n\rangle\!\langle\gamma_i^n|$ is a probabilistic n-comb.

Since U is applied only on $\mathcal{H}_\mathrm{B}$, it is easy to see that $U^{\otimes n}|\gamma_i\rangle\!\langle\gamma_i|U^{\dagger\otimes n}$ is also a probabilistic n-comb. Finally, Γ_i is also a probabilistic n-comb because it is a convex combination of probabilistic n-combs. ◻

Lemma 4.6. *Recall that $d_{\min} = \min\{d, D-d\}$. Suppose $n \le \frac{1}{2e^4}(D-d)d_{\min}/\varepsilon^2$. There exists positive numbers $\lambda_0, \ldots, \lambda_n$ such that for any $V \in \mathcal{N}$*

$$\begin{aligned} \sum_{i=0}^{n} \lambda_i &\le 4(D-d)^2 d_{\min}^2 \exp\left(\sqrt{8n\varepsilon^2(D-d)d_{\min}}\right), \\ |V\rangle\!\rangle\!\langle\!\langle V|^{\otimes n} &\le \sum_{i=0}^{n} \lambda_i \Gamma_i. \end{aligned} \tag{46}$$

Proof. We remark that, because of Equation (41), $|V\rangle\!\rangle^{\otimes n}$ is contained in $\bigoplus_{j=0}^{n} \mathrm{supp}(\Gamma_j)$. Moreover, by Fact 7.7, to establish Equation (46), it is sufficient to show that

$$\sum_{i=0}^{n} \frac{1}{\lambda_i} \mathrm{tr}\big(\Gamma_i^{-1} |V\rangle\!\rangle\langle\!\langle V|^{\otimes n}\big) \leq 1. \tag{47}$$

To this end, we bound the left hand side of Equation (47) as

$$\sum_{i=0}^{n} \frac{1}{\lambda_i} \mathrm{tr}\big(\Gamma_i^{-1} |V\rangle\!\rangle\langle\!\langle V|^{\otimes n}\big) = \sum_{i=0}^{n} \frac{1}{\lambda_i} \binom{n}{i} \varepsilon^{2i} \, \mathrm{tr}\Big(\Gamma_i^{-1} U^{\otimes n} |\gamma_i\rangle\langle\gamma_i| U^{\dagger\otimes n}\Big) \tag{48}$$

$$= \sum_{i=0}^{n} \frac{1}{\lambda_i} \binom{n}{i} \varepsilon^{2i} \, \mathrm{tr}\big(\Gamma_i^{-1} |\gamma_i\rangle\langle\gamma_i|\big) \tag{49}$$

$$\leq \sum_{i=0}^{n} \frac{1}{\lambda_i} \binom{n}{i} \varepsilon^{2i} \binom{(D-d)d' + i - 1}{i} \tag{50}$$

$$\leq \sum_{i=0}^{n} \frac{1}{\lambda_i} \binom{n}{i} \varepsilon^{2i} \binom{(D-d)d_{\min} + i - 1}{i}. \tag{51}$$

Here Equation (48) follows from Equation (41) and the fact that $U^{\otimes n}|\gamma_i\rangle \in \mathrm{supp}(\Gamma_i)$, Equation (49) uses the fact that Γ_i commutes with $U^{\otimes n}$, and Equation (50) is an application of Lemma 7.1 where $|\gamma_i\rangle$ is viewed as a vector in the space

$$\mathrm{span}\left(\left\{ \sum_{\substack{S\subseteq[n] \\ |S|=i}} |\psi\rangle^{\otimes S} \otimes |V_0\rangle\!\rangle^{\otimes [n]\setminus S} \;\middle|\; |\psi\rangle \in \Pi \right\}\right), \tag{52}$$

where

$$\Pi = \mathrm{span}\{|d+1\rangle_{\mathrm{B}}, \ldots, |D\rangle_{\mathrm{B}}\} \otimes \mathrm{span}\{|1\rangle_{\mathrm{A}}, \ldots, |d'\rangle_{\mathrm{A}}\}.$$

The space defined in Equation (52) is invariant under the action $U^{\otimes n}$ for any $U \in \mathbb{U}_{D-d}$ since U fixes $|V_0\rangle\!\rangle$ and Π is an invariant space under U (recall that U acts as $I_d \oplus U$ on $\mathcal{H}_{\mathrm{B}}$). By Lemma 7.2, it has dimension $\binom{(D-d)d'+i-1}{i}$. Therefore, we obtain Equation (50). Equation (51) uses $d' \leq \min\{d, D-d\} = d_{\min}$.

Our goal is look for positive numbers $\lambda_0, \ldots, \lambda_n$ such that $\sum_{i=0}^{n} \lambda_i$ is small and Equation (51) is upper bounded by 1. For ease of notation, we denote $Q = (D-d)d_{\min}$. Equation (51) can be upper bounded using Fact 7.6 as follows

$$(51) \leq \sum_{i=0}^{n} \frac{1}{\lambda_i} \exp\left(nH\left(\frac{i}{n}\right) + 2i\ln(\varepsilon) + (Q+i)H\left(\frac{i}{Q+i}\right)\right)$$

$$= \sum_{i=0}^{n} \frac{1}{\lambda_i} \exp\left(-i\ln\left(\frac{i}{n\varepsilon^2}\right) - (n-i)\ln\left(\frac{n-i}{n}\right) + i\ln\left(1+\frac{Q}{i}\right) + Q\ln\left(1+\frac{i}{Q}\right)\right)$$

$$\leq \sum_{i=0}^{n} \frac{1}{\lambda_i} \exp\left(-i\ln\left(\frac{i}{n\varepsilon^2}\right) + i\ln\left(1+\frac{Q}{i}\right) + 2i\right), \tag{53}$$

where the last inequality uses

$$(n-i)\ln\left(\frac{n}{n-i}\right) \leq (n-i)\left(\frac{n}{n-i} - 1\right) = i \quad \text{and} \quad Q\ln\left(1+\frac{i}{Q}\right) \leq Q\frac{i}{Q} = i.$$

To further bound Equation (53), we consider two cases depending on the value of i:

- When $i < Q$, we have

$$
\begin{aligned}
-i\ln\left(\frac{i}{n\varepsilon^2}\right)+i\ln\left(1+\frac{Q}{i}\right)+2i &\le -i\ln\left(\frac{i}{n\varepsilon^2}\right)+i\ln\left(\frac{2Q}{i}\right)+2i \\
&= 2i\ln\left(\frac{\sqrt{2e^2n\varepsilon^2Q}}{i}\right) \\
&\le \sqrt{8n\varepsilon^2 Q},
\end{aligned} \tag{54}
$$

  where the last inequality follows from Fact 7.8.

- When $i \ge Q$, since we assumed that $n \le \frac{1}{2e^4}(D-d)d_{\min}/\varepsilon^2 = \frac{1}{2e^4}Q/\varepsilon^2$ we have $i \ge 2e^4 n\varepsilon^2$, thus

$$
-i\ln\left(\frac{i}{n\varepsilon^2}\right)+i\ln\left(1+\frac{Q}{i}\right)+2i \le -i\ln(2e^4)+i\ln(2)+2i = -2i.
$$

Therefore, the choice

$$
\lambda_i = \begin{cases} 3Q\exp\left(\sqrt{8n\varepsilon^2 Q}\right), & \text{if } i < Q \\ \exp(-i), & \text{if } i \ge Q, \end{cases}
$$

ensures that Equation (53) can be upper bounded by

$$
\begin{aligned}
(53) &\le \sum_{i<Q}\frac{1}{\lambda_i}\exp\left(\sqrt{8n\varepsilon^2Q}\right)+\sum_{i\ge Q}\frac{1}{\lambda_i}\exp(-2i) \\
&\le \frac{1}{3}+\sum_{i\ge Q}\exp(-i) \le \frac{1}{3}+\exp(-Q)\frac{e}{e-1} < 1,
\end{aligned}
$$

where we use that $Q = (D-d)d_{\min} \ge 1$. Finally, we have that

$$
\begin{aligned}
\sum_{i=0}^{n}\lambda_i &\le 3Q^2\exp\left(\sqrt{8n\varepsilon^2Q}\right)+\exp(-Q)\frac{e}{e-1} \\
&< 3Q^2\exp\left(\sqrt{8n\varepsilon^2Q}\right)+\frac{1}{2} < 4Q^2\exp\left(\sqrt{8n\varepsilon^2Q}\right) \\
&= 4(D-d)^2d_{\min}^2\exp\left(\sqrt{8n\varepsilon^2(D-d)d_{\min}}\right),
\end{aligned}
$$

which concludes the proof. □

5 Packing nets of quantum channels

In this section, we construct packing nets of quantum channels that correspond to the type I and type II hard instances defined in Section 4. The packing nets have different distances and cardinality depending on the parameter regimes. We summarize the results in Table 4.

Let d_1, d_2 and r denote the input dimension, output dimension and upper bound of Kraus rank of the quantum channels. In this section, we further assume

$$
r \le d_1 d_2/2, \tag{55}
$$

Type of hardness	Assumption	Distance	Logarithm of cardinality
Type I Section 5.1	$d_1 \le rd_2 \le \frac{4}{3}d_1$	Choi: $\Omega(\varepsilon)$	$\Omega(d_1^2)$
Type II Section 5.2	$rd_2 > d_1$ $rd_2 < d_1 + r$	Choi: $\Omega(\kappa^{3/2}\varepsilon)$	$\Omega(\kappa^2(rd_2 - d_1)^2)$
		Diamond: $\Omega(\varepsilon)$	$\Omega((rd_2 - d_1)^2)$
Type II Section 5.3	$rd_2 > d_1$ $d_1 + r \le rd_2$, $r \le d_1$	Choi: $\Omega(\kappa^{3/2}\varepsilon)$	$\Omega(\kappa^3 d_1(rd_2 - d_1))$
		Diamond: $\Omega(\varepsilon)$	$\Omega(\kappa d_1(rd_2 - d_1))$
Type II Section 5.4	$rd_2 > d_1$ $d_1 + r \le rd_2$, $d_1 < r$	Choi: $\Omega(\varepsilon)$	$\Omega(rd_1d_2)$

Table 4: Construction of packing nets. Here, $\kappa := \min\{(rd_2 - d_1)/d_1, 1\} = \min\{\tau - 1, 1\}$. Note that for the non-boundary case ($rd_2 > d_1$), κ can be close to 0 while $\kappa d_1 = \min\{rd_2 - d_1, d_1\} \ge 1$.

instead of the more general $r \le d_1 d_2$. Note that adding this constraint here (in the packing net construction) still allows us to derive general lower bounds for quantum channel tomography, since a larger Kraus rank only makes the tomography task more difficult. Then, the following notation will be used in this section.

Notation 5.1. *Let $\mathcal{H}_{\mathrm{A}} \cong \mathbb{C}^{d_1}$, $\mathcal{H}_{\mathrm{B}} \cong \mathbb{C}^{d_2}$ and $\mathcal{H}_{\mathrm{anc}} \cong \mathbb{C}^{r}$ be the input, output and ancilla systems. Let $\{|1\rangle_{\mathrm{A}}, \ldots, |d_1\rangle_{\mathrm{A}}\}$ be an orthonormal basis of $\mathcal{H}_{\mathrm{A}}$, and similarly let $\{|1\rangle_{\mathrm{B}}, \ldots, |d_2\rangle_{\mathrm{B}}\}$ and $\{|1\rangle_{\mathrm{anc}}, \ldots, |r\rangle_{\mathrm{anc}}\}$ be orthonormal bases of $\mathcal{H}_{\mathrm{B}}$ and $\mathcal{H}_{\mathrm{anc}}$, respectively. Moreover, for $\mathrm{X} \in \{\mathrm{A}, \mathrm{B}, \mathrm{anc}\}$, we use $\mathcal{H}_{\mathrm{X}}[i:j]$ to denote the subspace of a Hilbert space $\mathcal{H}_{\mathrm{X}}$ spanned by $\{|i\rangle_X, \ldots, |j\rangle_X\}$, and we use $\mathcal{H}[i]$ to denote $\mathcal{H}[i:i]$.*

5.1 Type I instance: $d_1 \le rd_2 \le \frac{4}{3}d_1$

5.1.1 Construction

In this subsection, we will use the definition in Section 4.1.1, where the parameter d, D in Section 4.1.1 correspond to d_1 and rd_2 here.

We define $d_1' = d_1$ when d_1 is even and $d_1' = d_1 - 1$ when d_1 is odd. Let $g : [d_1] \to [d_2] \times [r]$ be the function $g(i) = (\lfloor \frac{i-1}{r} \rfloor + 1, i - r\lfloor \frac{i-1}{r} \rfloor)$, i.e., g maps the integers in $[d_1]$ to $[d_2] \times [r]$ in row-major order. Let $\mathcal{H}_{\mathrm{B}'}$ be a rd_2-dimensional space with an orthonormal basis $\{|1\rangle_{\mathrm{B}'}, \ldots, |d_2 r\rangle_{\mathrm{B}'}\}$. We identify $\mathcal{H}_{\mathrm{B}'}$ with $\mathcal{H}_{\mathrm{B}} \otimes \mathcal{H}_{\mathrm{anc}}$ by the map $|i\rangle_{\mathrm{B}'} \mapsto |g(i)_1\rangle_{\mathrm{B}} \otimes |g(i)_2\rangle_{\mathrm{anc}}$. In other words, the basis vectors are identified as follows:

$$|1\rangle_{\mathrm{B}'},\, |2\rangle_{\mathrm{B}'},\, \ldots,\, |d_2 r\rangle_{\mathrm{B}'} \longleftrightarrow |1\rangle_{\mathrm{B}} \otimes |1\rangle_{\mathrm{anc}},\, |1\rangle_{\mathrm{B}} \otimes |2\rangle_{\mathrm{anc}},\, \ldots, |d_2\rangle_{\mathrm{B}} \otimes |r\rangle_{\mathrm{anc}}.$$

We also identify $\mathcal{H}_{\mathrm{A}}$ with the subspace $\mathcal{H}_{\mathrm{B}'}[1:d_1]$ of $\mathcal{H}_{\mathrm{B}'}$ using the map $|i\rangle_{\mathrm{A}} \mapsto |i\rangle_{\mathrm{B}'}$ for $i \in [d_1]$. In other words, we have the following correspondence:

$$|1\rangle_{\mathrm{A}},\, |2\rangle_{\mathrm{A}},\, \ldots,\, |d_1\rangle_{\mathrm{A}} \longleftrightarrow |1\rangle_{\mathrm{B}'},\, |2\rangle_{\mathrm{B}'},\, \ldots,\, |d_1\rangle_{\mathrm{B}'}. \tag{56}$$

Then, define the linear operator $V_0 : \mathcal{H}_{\mathrm{A}} \to \mathcal{H}_{\mathrm{B}} \otimes \mathcal{H}_{\mathrm{anc}}$ as

$$V_0 := \begin{cases} \sqrt{1-\varepsilon^2} \sum_{i=1}^{d'} |i\rangle_{\mathrm{B}'}\langle i|_{\mathrm{A}}, & \text{if } d_1 \text{ is even} \\ \sqrt{1-\varepsilon^2} \sum_{i=1}^{d'} |i\rangle_{\mathrm{B}'}\langle i|_{\mathrm{A}} + |d\rangle_{\mathrm{B}'}\langle d|_{\mathrm{A}}, & \text{if } d_1 \text{ is odd.} \end{cases}$$

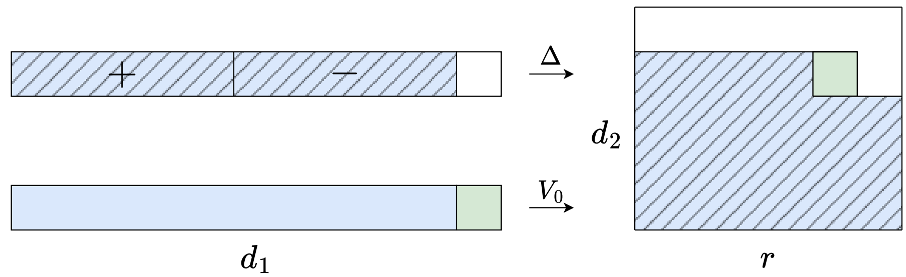


Figure 4: Illustration of our construction. We define linear operators V_0 and Δ from the d_1-dimensional space $\mathcal{H}_{\mathrm{A}}$ to the $d_2 r$-dimensional space $\mathcal{H}_{\mathrm{B}} \otimes \mathcal{H}_{\mathrm{anc}}$. The Haar randomness is applied on the hatched area.

Define the linear operator $\Delta : \mathcal{H}_{\mathrm{A}} \to \mathcal{H}_{\mathrm{B}'}$ as

$$\Delta \coloneqq \mathrm{i} \left(\sum_{i=1}^{\lfloor d_1/2 \rfloor} |i\rangle_{\mathrm{B}'} \langle i|_{\mathrm{A}} - \sum_{i=\lfloor d_1/2 \rfloor + 1}^{d_1'} |i\rangle_{\mathrm{B}'} \langle i|_{\mathrm{A}} \right), \tag{57}$$

where i is the imaginary unit. Note that under the identification in Equation (56), we can simply discard the subscript A and B′ in the definition of V_0 and Δ.

Suppose $U \in \mathbb{U}_{d_1'}$ is a unitary acting on $\mathrm{span}\{|1\rangle, \ldots, |d_1'\rangle\}$, and U fixes $|d_1\rangle$ if $d_1 > d_1'$. Define the isometry $V_{\varepsilon,U} : \mathcal{H}_{\mathrm{A}} \to \mathcal{H}_{\mathrm{B}'}$ as

$$V_{\varepsilon,U} \coloneqq U \left(V_0 + \varepsilon \Delta \right) U^\dagger = V_0 + \varepsilon U \Delta U^\dagger, \tag{58}$$

where with a little abuse of notation, U can act either on $\mathcal{H}_{\mathrm{A}}$ or $\mathcal{H}_{\mathrm{B}'}$ (using the identification in Equation (56), and we assume that U acts trivially on $\mathrm{span}\{|d_1+1\rangle_{\mathrm{B}'}, \ldots, |d_2 r\rangle_{\mathrm{B}'}\}$), which should not cause confusion given the context. For clarity, we illustrate our construction in Figure 4.

5.1.2 Existence

Theorem 5.2. *Suppose $\varepsilon \leq 1/39660$. There exists a finite subset $\mathcal{N}$ of $\{V_{\varepsilon,U} \,|\, U \in \mathbb{U}_{d_1'}\}$ for $V_{\varepsilon,U}$ defined in Equation (58) with cardinality $|\mathcal{N}| \geq \exp(d_1^2/750197760001)$, such that for any $V_x \neq V_y \in \mathcal{N}$, if we set $\mathcal{E}_x = \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(V_x(\cdot)V_x^\dagger)$ and $\mathcal{E}_y = \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(V_y(\cdot)V_y^\dagger)$, then*

$$\frac{1}{d_1} \left\| C_{\mathcal{E}_x} - C_{\mathcal{E}_y} \right\|_1 \geq \frac{\varepsilon}{1983000}. \tag{59}$$

Proof. First, we need the following lemma.

Lemma 5.3. *Suppose $\varepsilon \leq 1/39660$. Given any unitary $U \in \mathbb{U}_{d_1'}$, let $\mathcal{E}_U = \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(V_{\varepsilon,U}(\cdot)V_{\varepsilon,U}^\dagger)$ where $V_{\varepsilon,U}$ is the isometry defined in Equation (58). Then, the function*

$$f(U_x, U_y) = \frac{1}{d_1} \left\| C_{\mathcal{E}_{U_x}} - C_{\mathcal{E}_{U_y}} \right\|_1$$

is $\varepsilon\sqrt{\frac{32}{d_1}}$-Lipschitz with respect to the ℓ_2-sum of the 2-norms (Frobenius norm), namely

$$|f(U_x, U_y) - f(U_x', U_y')| \leq \varepsilon \sqrt{\frac{32}{d_1}} \|(U_x, U_y) - (U_x', U_y')\|_F$$

for all $U_x, U_x', U_y, U_y' \in \mathbb{U}_{d'}$*, where* $\|(A,B)\|_F := \sqrt{\|A\|_F^2 + \|B\|_F^2}$ *and* $\|\cdot\|_F$ *denotes the Frobenius norm. Furthermore, if we consider independent random unitaries* $U_x, U_y \sim \mathbb{U}_{d_1'}$*, we have*

$$\mathbb{E}[f(U_x, U_y)] \geq \frac{\varepsilon}{19830}.$$

The proof of Lemma 5.3 is deferred to Section 5.1.3.

We now have all the ingredients to prove Theorem 5.2. Since the function $f(U_x, U_y)$ is $\varepsilon\sqrt{\frac{32}{d_1}}$-Lipschitz, when we sample two independent unitaries $U_x, U_u \sim \mathbb{U}_{d_1'}$, by Theorem 7.10, we have[6]

$$\mathbb{P}\Big[f(U_i, U_j) \leq \frac{\varepsilon}{1983000}\Big] \leq \exp\left(-\frac{d_1'\varepsilon^2}{12 \cdot 22100^2} \cdot \frac{d_1}{32\varepsilon^2}\right) \leq \exp\left(-\frac{d_1^2}{768 \cdot 22100^2}\right).$$

Then, we independently sample $\exp(d_1^2/(1536 \cdot 22100^2 + 1)) = \exp(d_1^2/750197760001)$ Haar random unitaries in $\mathbb{U}_{d_1'}$ and the union bound shows that there exists a non-zero probability that for any pair U_x, U_y, we have $f(U_x, U_y) \geq \varepsilon/1983000$. Therefore, there exists a subset $\mathcal{M} \subseteq \mathbb{U}_{d_1'}$ with cardinality $|\mathcal{M}| \geq \exp(d_1^2/750197760001)$ such that for any pair $U_x, U_y \in \mathcal{M}$, we have $f(U_x, U_y) \geq \varepsilon/1983000$. Then, the set $\{V_{\varepsilon,U} | U \in \mathcal{M}\}$ is as desired. □

5.1.3 Proof of Lemma 5.3

Proof that f is Lipschitz. Recall that $V_{\varepsilon,U} = V_0 + \varepsilon U \Delta U^\dagger$. We have that by the triangle inequality:

$$\begin{aligned}
&|f(U_x, U_y) - f(U_x', U_y')| \\
=&\frac{1}{d_1}\left|\left\| \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(|V_{\varepsilon,U_x}\rangle\!\rangle\langle\!\langle V_{\varepsilon,U_x}| - |V_{\varepsilon,U_y}\rangle\!\rangle\langle\!\langle V_{\varepsilon,U_y}|)\right\|_1 - \left\|\mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(|V_{\varepsilon,U_x'}\rangle\!\rangle\langle\!\langle V_{\varepsilon,U_x'}| - |V_{\varepsilon,U_y'}\rangle\!\rangle\langle\!\langle V_{\varepsilon,U_y'}|)\right\|_1\right| \\
\leq&\frac{1}{d_1}\left\|\mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(|V_{\varepsilon,U_x}\rangle\!\rangle\langle\!\langle V_{\varepsilon,U_x}| - |V_{\varepsilon,U_x'}\rangle\!\rangle\langle\!\langle V_{\varepsilon,U_x'}|) - \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(|V_{\varepsilon,U_y}\rangle\!\rangle\langle\!\langle V_{\varepsilon,U_y}| - |V_{\varepsilon,U_y'}\rangle\!\rangle\langle\!\langle V_{\varepsilon,U_y'}|)\right\|_1 \\
\leq&\frac{1}{d_1}\left\| \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(|V_{\varepsilon,U_x}\rangle\!\rangle\langle\!\langle V_{\varepsilon,U_x}| - |V_{\varepsilon,U_x'}\rangle\!\rangle\langle\!\langle V_{\varepsilon,U_x'}|)\right\|_1 + \frac{1}{d_1}\left\|\mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(|V_{\varepsilon,U_y}\rangle\!\rangle\langle\!\langle V_{\varepsilon,U_y}| - |V_{\varepsilon,U_y'}\rangle\!\rangle\langle\!\langle V_{\varepsilon,U_y'}|)\right\|_1.
\end{aligned} \tag{60}$$

Note that

$$\begin{aligned}
&\left\| \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\left((|V_{\varepsilon,U_x}\rangle\!\rangle - |V_{\varepsilon,U_x'}\rangle\!\rangle)\langle\!\langle V_{\varepsilon,U_x}|\right)\right\|_1 \leq \varepsilon\left\|\left(|U_x\Delta U_x^\dagger\rangle\!\rangle - |U_x'\Delta U_x'^\dagger\rangle\!\rangle\right)\langle\!\langle V_{\varepsilon,U_x}|\right\|_1 \\
=&\varepsilon\sqrt{d_1}\left\||U_x\Delta U_x^\dagger\rangle\!\rangle - |U_x'\Delta U_x'^\dagger\rangle\!\rangle\right\| \leq \varepsilon\sqrt{d_1}\left(\left\||(U_x - U_x')\Delta U_x^\dagger\rangle\!\rangle\right\| + \left\||U_x'\Delta(U_x - U_x')^\dagger\rangle\!\rangle\right\|\right) \\
\leq&\varepsilon\sqrt{d_1}\sqrt{\mathrm{tr}((U_x - U_x')^\dagger(U_x - U_x'))} + \varepsilon\sqrt{d_1}\sqrt{\mathrm{tr}((U_x - U_x')^\dagger(U_x - U_x'))} \qquad (61) \\
=&2\varepsilon\sqrt{d_1}\|U_x - U_x'\|_F,
\end{aligned}$$

where Equation (61) is by using $\Delta^\dagger\Delta \leq I$. Hence by the triangle inequality

$$\begin{aligned}
&\left\| \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(|V_{\varepsilon,U_x}\rangle\!\rangle\langle\!\langle V_{\varepsilon,U_x}| - |V_{\varepsilon,U_x'}\rangle\!\rangle\langle\!\langle V_{\varepsilon,U_x'}|)\right\|_1 \\
\leq&\left\| \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\left((|V_{\varepsilon,U_x}\rangle\!\rangle - |V_{\varepsilon,U_x'}\rangle\!\rangle)\langle\!\langle V_{\varepsilon,U_x}|\right)\right\|_1 + \left\| \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\left(|V_{\varepsilon,U_x'}\rangle\!\rangle(\langle\!\langle V_{\varepsilon,U_x}| - \langle\!\langle V_{\varepsilon,U_x'}|)\right)\right\|_1
\end{aligned}$$

[6]To be precise, we are applying Theorem 7.10 to $-f$.

$$\leq 4\varepsilon\sqrt{d_1}\|U_x - U'_x\|_F. \tag{62}$$

Therefore, combining Equation (60) and Equation (62), we have

$$\begin{aligned}|f(U_x,U_y) - f(U'_x,U'_y)| &\leq 4\varepsilon\sqrt{\frac{1}{d_1}}\|U_x - U'_x\|_F + 4\varepsilon\sqrt{\frac{1}{d_1}}\|U_y - U'_y\|_F \\ &\leq 4\varepsilon\sqrt{\frac{2}{d_1}}\sqrt{\|U_x - U'_x\|_F^2 + \|U_y - U'_y\|_F^2}.\end{aligned}$$

Proof of the lower bound on the the expectation of f. Recall the definition of the $d_1 \times d_1$ matrix Δ in Equation (57). This construction ensures the following properties, which will be used later:

- $\mathrm{tr}(\Delta) = 0$,
- $\|\Delta\|_{\mathrm{op}} = 1$,
- $\mathrm{tr}(\Delta\Delta^\dagger\Delta\Delta^\dagger) = \mathrm{tr}(\Delta\Delta^\dagger) = d'_1$.

Note that, for unitaries $U_x, U_y \in \mathbb{U}_{d'_1}$, $f(U_x,U_y)$ equals

$$\begin{aligned}\frac{1}{d_1}\left\|C_{\mathcal{E}_{U_x}} - C_{\mathcal{E}_{U_y}}\right\|_1 &\geq \frac{\varepsilon}{d_1}\left\|\mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big(|U_x\Delta U_x^\dagger\rangle\!\rangle\langle\!\langle V_0| + |V_0\rangle\!\rangle\langle\!\langle U_x\Delta U_x^\dagger| - |U_y\Delta U_y^\dagger\rangle\!\rangle\langle\!\langle V_0| - |V_0\rangle\!\rangle\langle\!\langle U_y\Delta U_y^\dagger|\Big)\right\|_1 \\ &\quad - \frac{\varepsilon^2}{d_1}\left\|\mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big(|U_x\Delta U_x^\dagger\rangle\!\rangle\langle\!\langle U_x\Delta U_x^\dagger| - |U_y\Delta U_y^\dagger\rangle\!\rangle\langle\!\langle U_y\Delta U_y^\dagger|\Big)\right\|_1 \\ &\geq \frac{\varepsilon}{d_1}\|D(U_x,U_y)\|_1 - 2\varepsilon^2, \end{aligned} \tag{63}$$

where we define

$$D(U_x,U_y) \coloneqq A_x + A_x^\dagger - A_y - A_y^\dagger, \quad A_x \coloneqq \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\left(|U_x\Delta U_x^\dagger\rangle\!\rangle\langle\!\langle V_0|\right), \quad A_y \coloneqq \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\left(|U_y\Delta U_y^\dagger\rangle\!\rangle\langle\!\langle V_0|\right),$$

and used that $\left\|\mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big(|U_x\Delta U_x^\dagger\rangle\!\rangle\langle\!\langle U_x\Delta U_x^\dagger|\Big)\right\|_1 = \mathrm{tr}\left(|U_x\Delta U_x^\dagger\rangle\!\rangle\langle\!\langle U_x\Delta U_x^\dagger|\right) = \mathrm{tr}\big(\Delta\Delta^\dagger\big) \leq d_1$.

Now, we are interested in proving the bound

$$\mathbb{E}\big[\mathrm{tr}\big(|D(U_x,U_y)|^2\big)\big] \geq \frac{d_1^2}{20r}, \tag{64}$$

where the expectation is over the Haar random $U_x, U_y \sim \mathbb{U}_{d'_1}$. We have that

$$\mathbb{E}\big[\mathrm{tr}\big(|D(U_x,U_y)|^2\big)\big] = 2\,\mathbb{E}\Big[\mathrm{tr}(A_x^2) + 2\,\mathrm{tr}(A_x^\dagger A_x) + \mathrm{tr}(A_x^{\dagger 2})\Big] - 2\Re\,\mathbb{E}\Big[\mathrm{tr}(A_xA_y) + \mathrm{tr}(A_x^\dagger A_y)\Big]. \tag{65}$$

Since $\mathrm{tr}(\Delta) = 0$, we know that $\mathbb{E}\Big[U_x\Delta U_x^\dagger\Big] = \mathrm{tr}(\Delta)\frac{I}{d_1} = 0$ by Schur's lemma. This means

$$\mathbb{E}[\mathrm{tr}(A_xA_y)] = \mathbb{E}\Big[\mathrm{tr}(A_x^\dagger A_y)\Big] = 0. \tag{66}$$

Moreover, we have that

$$\mathbb{E}\Big[\mathrm{tr}(A_xA_x^\dagger)\Big] = \mathbb{E}\Big[\mathrm{tr}\Big(\mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big(|U_x\Delta U_x^\dagger\rangle\!\rangle\langle\!\langle V_0|\Big)\,\mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big(|V_0\rangle\!\rangle\langle\!\langle U_x\Delta U_x^\dagger|\Big)\Big)\Big]$$

$$
\begin{aligned}
&= \sum_{i,j=1}^{r} \mathbb{E}\Big[\operatorname{tr}\Big(|U_x\Delta U_x^\dagger\rangle\!\rangle\langle\!\langle V_0| \cdot |i\rangle\langle j|_{\text{anc}} \cdot |V_0\rangle\!\rangle\langle\!\langle U_x\Delta U_x^\dagger| \cdot |j\rangle\langle i|_{\text{anc}}\Big)\Big] \\
&= \sum_{i,j=1}^{r} \langle\!\langle V_0| \cdot |i\rangle\langle j|_{\text{anc}} \cdot |V_0\rangle\!\rangle \cdot \mathbb{E}\Big[\operatorname{tr}\Big(|U_x\Delta U_x^\dagger\rangle\!\rangle\langle\!\langle U_x\Delta U_x^\dagger| \cdot |j\rangle\langle i|_{\text{anc}}\Big)\Big] \\
&= \sum_{i=1}^{r} \langle\!\langle V_0| \cdot |i\rangle\langle i|_{\text{anc}} \cdot |V_0\rangle\!\rangle \cdot \mathbb{E}\Big[\operatorname{tr}\Big(|U_x\Delta U_x^\dagger\rangle\!\rangle\langle\!\langle U_x\Delta U_x^\dagger| \cdot |i\rangle\langle i|_{\text{anc}}\Big)\Big] \\
&\geq (1-\varepsilon^2)\left\lfloor \frac{d_1}{r} \right\rfloor \cdot \sum_{i=1}^{r} \mathbb{E}\Big[\operatorname{tr}\Big(|U_x\Delta U_x^\dagger\rangle\!\rangle\langle\!\langle U_x\Delta^\dagger U_x^\dagger| \cdot |i\rangle\langle i|_{\text{anc}}\Big)\Big] && (67) \\
&= (1-\varepsilon^2)\left\lfloor \frac{d_1}{r} \right\rfloor \cdot \mathbb{E}\Big[\operatorname{tr}\Big(|U_x\Delta U_x^\dagger\rangle\!\rangle\langle\!\langle U_x\Delta U_x^\dagger|\Big)\Big] \\
&= (1-\varepsilon^2)\left\lfloor \frac{d_1}{r} \right\rfloor \cdot \mathbb{E}\Big[\operatorname{tr}\Big(\Delta^\dagger\Delta\Big)\Big] \\
&= (1-\varepsilon^2)\left\lfloor \frac{d_1}{r} \right\rfloor d_1', && (68)
\end{aligned}
$$

where in Equation (67) the inequality we used that $\langle\!\langle V_0| \cdot |i\rangle\langle i|_{\text{anc}} \cdot |V_0\rangle\!\rangle \geq (1-\varepsilon^2)\sum_{j=1}^{d_1} \mathbb{1}_{g_2(j)=i} \geq (1-\varepsilon^2)\lfloor \frac{d_1}{r} \rfloor$. We also have

$$
\begin{aligned}
&\mathbb{E}\big[\operatorname{tr}(A_x^2)\big] \\
&= \mathbb{E}\Big[\operatorname{tr}\Big(\operatorname{tr}_{\mathcal{H}_{\text{anc}}}\Big(|U_x\Delta U_x^\dagger\rangle\!\rangle\langle\!\langle V_0|\Big) \operatorname{tr}_{\mathcal{H}_{\text{anc}}}\Big(|U_x\Delta U_x^\dagger\rangle\!\rangle\langle\!\langle V_0|\Big)\Big)\Big] \\
&= \sum_{i,j=1}^{r} \mathbb{E}\Big[\operatorname{tr}\Big(|U_x\Delta U_x^\dagger\rangle\!\rangle\langle\!\langle V_0| \cdot |i\rangle\langle j|_{\text{anc}} \cdot |U_x\Delta U_x^\dagger\rangle\!\rangle\langle\!\langle V_0| \cdot |j\rangle\langle i|_{\text{anc}}\Big)\Big] \\
&= \sum_{i,j=1}^{r} \mathbb{E}\Big[\operatorname{tr}\Big(V_0^\dagger \cdot |i\rangle\langle j|_{\text{anc}} \cdot U_x\Delta U_x^\dagger\Big) \operatorname{tr}\Big(V_0^\dagger \cdot |j\rangle\langle i|_{\text{anc}} \cdot U_x\Delta U_x^\dagger\Big)\Big] \\
&= \sum_{i,j=1}^{r}\sum_{a,b=1}^{d_1} \mathbb{E}\Big[\operatorname{tr}\Big(V_0^\dagger \cdot |i\rangle\langle j|_{\text{anc}} \cdot U_x\Delta U_x^\dagger \cdot |a\rangle\langle b| \cdot V_0^\dagger \cdot |j\rangle\langle i|_{\text{anc}} \cdot U_x\Delta U_x^\dagger \cdot |b\rangle\langle a|\Big)\Big] \\
&= \sum_{i,j=1}^{r}\sum_{a,b=1}^{d_1} \mathbb{E}\Big[\operatorname{tr}\Big(U_x\Delta U_x^\dagger P \cdot |a\rangle\langle b| \cdot V_0^\dagger \cdot |j\rangle\langle i|_{\text{anc}} \cdot PU_x\Delta U_x^\dagger P \cdot |b\rangle\langle a| \cdot V_0^\dagger \cdot |i\rangle\langle j|_{\text{anc}} \cdot P\Big)\Big] && (69) \\
&= \sum_{i,j=1}^{r}\sum_{a,b=1}^{d_1'} \Bigg[\frac{1}{d_1'^2-1}\Big(\operatorname{tr}(\Delta^2)\operatorname{tr}(P|a\rangle\langle b|V_0^\dagger|j\rangle\langle i|_{\text{anc}}P)\operatorname{tr}(P|b\rangle\langle a|V_0^\dagger|i\rangle\langle j|_{\text{anc}}P)\Big) \\
&\qquad - \frac{1}{d_1'(d_1'^2-1)}\Big(\operatorname{tr}(\Delta^2)\operatorname{tr}\big(P|a\rangle\langle b|V_0^\dagger|j\rangle\langle i|_{\text{anc}}P|b\rangle\langle a|V_0^\dagger|i\rangle\langle j|_{\text{anc}}P\big)\Big)\Bigg] && (70) \\
&= \sum_{i,j=1}^{r}\sum_{a,b=1}^{d_1'} \Bigg[-\frac{d_1'(1-\varepsilon^2)}{d_1'^2-1}\big|\operatorname{tr}(|a\rangle\langle b| \cdot |j\rangle\langle i|_{\text{anc}})\big|^2 + \frac{1-\varepsilon^2}{d_1'^2-1}\operatorname{tr}\big(|a\rangle\langle b| \cdot |j\rangle\langle i|_{\text{anc}} \cdot |b\rangle\langle a| \cdot |i\rangle\langle j|_{\text{anc}}\big)\Bigg] \\
&\geq -\frac{d_1'(1-\varepsilon^2)}{d_1'^2-1}\cdot\left(r^2\cdot\left\lceil\frac{d_1'}{r}\right\rceil\right) + \frac{1-\varepsilon^2}{d_1'^2-1}\cdot\left(r\cdot\left\lfloor\frac{d_1'}{r}\right\rfloor\right) && (71) \\
&\geq -\frac{(1-\varepsilon^2)r^2 d_1'}{d_1'^2-1}\left\lceil\frac{d_1'}{r}\right\rceil, && (72)
\end{aligned}
$$

where in Equation (69) we introduce the projector $P \coloneqq \sum_{i=1}^{d_1'} |i\rangle\langle i|$ and note that $PU_x\Delta U_x^\dagger P = U_x\Delta U_x^\dagger$, and Equation (70) is due to Corollary 7.13 and the fact that $\mathrm{tr}(\Delta) = 0$, Equation (71) is because $\sum_{a,b=1}^{d_1'} |\,\mathrm{tr}(|a\rangle\langle b| \cdot |j\rangle\langle i|_{\mathrm{anc}})|^2$ equals either $\lfloor \frac{d_1'}{r} \rfloor$ or $\lceil \frac{d_1'}{r} \rceil$ depending on i, j and thus is no more than $\lceil \frac{d_1'}{r} \rceil$; and $\sum_{a,b=1}^{d_1'} \mathrm{tr}(|a\rangle\langle b||j\rangle\langle i|_{\mathrm{anc}}|b\rangle\langle a||i\rangle\langle j|_{\mathrm{anc}})$ is non-zero only when $i = j$ and equals either $\lfloor \frac{d_1'}{r} \rfloor$ or $\lceil \frac{d_1'}{r} \rceil$ depending on i (which equals j) and thus is no less than $\lfloor \frac{d_1'}{r} \rfloor$, and Equation (72) is obtained by simply discarding the positive term. Therefore, combining Equation (65) with Equation (66), Equation (68) and Equation (72), we have

$$
\begin{aligned}
\mathbb{E}\big[\mathrm{tr}\big(|D(U_x,U_y)|^2\big)\big] &\geq 4(1-\varepsilon^2)\left(d_1'\left\lfloor\frac{d_1}{r}\right\rfloor - \frac{r^2 d_1'}{d_1'^2-1}\left\lceil\frac{d_1'}{r}\right\rceil\right) = 4(1-\varepsilon^2)d_1'\left\lfloor\frac{d_1}{r}\right\rfloor\left(1-\frac{r^2}{d_1'^2-1}\cdot\frac{\lceil d_1'/r\rceil}{\lfloor d_1/r\rfloor}\right) \\
&\geq 4(1-\varepsilon^2)d_1'\left\lfloor\frac{d_1}{r}\right\rfloor\cdot\frac{1}{10} \qquad (73) \\
&\geq (1-\varepsilon^2)\frac{d_1'd_1}{5r} \geq \frac{d_1'd_1}{10r} \geq \frac{d_1^2}{20r},
\end{aligned}
$$

where in Equation (73) we used $\lceil d_1'/r\rceil \leq \lceil d_1/r\rceil \leq 2\lfloor d_1/r\rfloor$ and

$$
2r^2 \leq \frac{8}{9}d_1^2 \leq \frac{9}{10}(d_1'^2 - 1),
$$

since $2r \leq rd_2 \leq \frac{4}{3}\cdot d_1$ and $d_1 \geq 162$ by assumption. Therefore, Equation (64) is proved.

Finally, let us prove that

$$
\mathbb{E}\big[\mathrm{tr}(|D(U_x,U_y)|^4)\big] \leq 12288\cdot\frac{d_1^4}{r^3}. \qquad (74)
$$

Recall that $D(U_x,U_y) = A_x + A_x^\dagger - A_y - A_y^\dagger$ so by the triangle inequality and Hölder inequality:

$$
\begin{aligned}
\mathbb{E}\big[\mathrm{tr}(|D(U_x,U_y)|^4)\big] &= \mathbb{E}\Big[\|A_x + A_x^\dagger - A_y - A_y^\dagger\|_4^4\Big] \leq \mathbb{E}\Big[\Big(\|A_x\|_4 + \|A_x^\dagger\|_4 + \|A_y\|_4 + \|A_y^\dagger\|_4\Big)^4\Big] \\
&\leq 4^3\cdot\mathbb{E}\Big[\|A_x\|_4^4 + \|A_x^\dagger\|_4^4 + \|A_y\|_4^4 + \|A_y^\dagger\|_4^4\Big] \\
&= 4^4\cdot\mathbb{E}\left[\|A_x\|_4^4\right], \qquad (75)
\end{aligned}
$$

where $\|\cdot\|_4$ denotes the Schatten 4-norm. Moreover,

$$
\begin{aligned}
&\mathbb{E}\big[\|A_x\|_4^4\big] = \mathbb{E}\Big[\mathrm{tr}(A_xA_x^\dagger A_xA_x^\dagger)\Big] \\
&= \mathbb{E}\Big[\mathrm{tr}\Big(\mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big(|U_x\Delta U_x^\dagger\rangle\!\rangle\langle\!\langle V_0|\Big)\,\mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big(|V_0\rangle\!\rangle\langle\!\langle U_x\Delta U_x^\dagger|\Big)\,\mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big(|U_x\Delta U_x^\dagger\rangle\!\rangle\langle\!\langle V_0|\Big)\,\mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big(|V_0\rangle\!\rangle\langle\!\langle U_x\Delta U_x^\dagger|\Big)\Big)\Big] \\
&= \sum_{i,j,k,l=1}^{r}\mathbb{E}\bigg[\mathrm{tr}\Big(|U_x\Delta U_x^\dagger\rangle\!\rangle\langle\!\langle V_0|\cdot|i\rangle\langle j|_{\mathrm{anc}}\cdot|V_0\rangle\!\rangle\langle\!\langle U_x\Delta U_x^\dagger|\cdot|j\rangle\langle k|_{\mathrm{anc}} \\
&\qquad\qquad\cdot|U_x\Delta U_x^\dagger\rangle\!\rangle\langle\!\langle V_0|\cdot|k\rangle\langle l|_{\mathrm{anc}}\cdot|V_0\rangle\!\rangle\langle\!\langle U_x\Delta U_x^\dagger|\cdot|l\rangle\langle i|_{\mathrm{anc}}\Big)\bigg] \\
&\leq \sum_{i,k=1}^{r}\left\lceil\frac{d_1}{r}\right\rceil^2\cdot\mathbb{E}\Big[\mathrm{tr}\Big(|U_x\Delta U_x^\dagger\rangle\!\rangle\langle\!\langle U_x\Delta U_x^\dagger|\cdot|i\rangle\langle k|_{\mathrm{anc}}\cdot|U_x\Delta U_x^\dagger\rangle\!\rangle\langle\!\langle U_x\Delta U_x^\dagger|\cdot|k\rangle\langle i|_{\mathrm{anc}}\Big)\Big] \\
&= \sum_{i,k=1}^{r}\left\lceil\frac{d_1}{r}\right\rceil^2\cdot\mathbb{E}\Big[\mathrm{tr}\Big(U_x\Delta^\dagger U_x^\dagger\cdot|i\rangle\langle k|_{\mathrm{anc}}\cdot U_x\Delta U_x^\dagger\Big)\,\mathrm{tr}\Big(U_x\Delta^\dagger U_x^\dagger\cdot|k\rangle\langle i|_{\mathrm{anc}}\cdot U_x\Delta U_x^\dagger\Big)\Big]
\end{aligned}
$$

$$= \sum_{i,k=1}^{r} \left\lceil \frac{d_1}{r} \right\rceil^2 \cdot \sum_{a,b=1}^{rd_2} \mathbb{E}\Big[\mathrm{tr}\Big(U_x \Delta\Delta^\dagger U_x^\dagger \cdot |i\rangle\langle k|_{\mathrm{anc}} \cdot |a\rangle\langle b| \cdot U_x \Delta\Delta^\dagger U_x^\dagger \cdot |k\rangle\langle i|_{\mathrm{anc}} \cdot |b\rangle\langle a|\Big)\Big]$$

$$= \sum_{i,k=1}^{r} \left\lceil \frac{d_1}{r} \right\rceil^2 \cdot \sum_{a,b=1}^{d_1'} \mathbb{E}\Big[\mathrm{tr}\Big(U_x \Delta\Delta^\dagger U_x^\dagger P \cdot |i\rangle\langle k|_{\mathrm{anc}} \cdot |a\rangle\langle b| \cdot P U_x \Delta\Delta^\dagger U_x^\dagger P \cdot |k\rangle\langle i|_{\mathrm{anc}} \cdot |b\rangle\langle a| \cdot P\Big)\Big] \tag{76}$$

$$= \sum_{i,k=1}^{r} \left\lceil \frac{d_1}{r} \right\rceil^2 \cdot \sum_{a,b=1}^{d_1'} \Bigg(\frac{1}{d_1'^2 - 1} \Big[\mathrm{tr}(\Delta\Delta^\dagger)^2 \cdot \langle a||i\rangle\langle k|_{\mathrm{anc}}|a\rangle \cdot \langle b||k\rangle\langle i|_{\mathrm{anc}}|b\rangle + \mathrm{tr}(\Delta\Delta^\dagger\Delta\Delta^\dagger) \cdot |\langle b||i\rangle\langle k|_{\mathrm{anc}}|a\rangle|^2 \Big]$$
$$- \frac{1}{d_1'(d_1'^2 - 1)} \Big[\mathrm{tr}(\Delta\Delta^\dagger)^2 \cdot |\langle b||i\rangle\langle k|_{\mathrm{anc}}|a\rangle|^2 + \mathrm{tr}(\Delta\Delta^\dagger\Delta\Delta^\dagger) \cdot \langle a||i\rangle\langle k|_{\mathrm{anc}}|a\rangle \cdot \langle b||k\rangle\langle i|_{\mathrm{anc}}|b\rangle \Big] \Bigg) \tag{77}$$

$$\leq \sum_{i,k=1}^{r} \left\lceil \frac{d_1}{r} \right\rceil^2 \cdot \frac{1}{d_1'^2 - 1} \cdot \left(d_1'^2 \cdot \mathbb{1}_{i=k} \cdot \left\lceil \frac{d_1'}{r} \right\rceil^2 + d_1' \cdot \left\lceil \frac{d_1'}{r} \right\rceil \right) \tag{78}$$

$$= \left\lceil \frac{d_1}{r} \right\rceil^2 \cdot \frac{1}{d_1'^2 - 1} \cdot \left(d_1'^2 r \left\lceil \frac{d_1'}{r} \right\rceil^2 + r^2 d_1' \left\lceil \frac{d_1'}{r} \right\rceil \right) \leq \frac{48 d_1^4}{r^3}, \tag{79}$$

where in Equation (76) we recall the projector $P = \sum_{i=1}^{d'}|i\rangle\langle i|$ and note that $PU_x\Delta\Delta^\dagger U_x^\dagger P = U_x\Delta\Delta^\dagger U_x^\dagger$, Equation (77) is due to Corollary 7.13, Equation (78) is obtained by directly discarding the negative terms, and in Equation (79) we used that $d_1' \leq d_1$, $\lceil d_1/r \rceil \leq 2d_1/r$, $d_1'^2 - 1 \geq d_1^2/2$. Therefore, combining Equation (75) with Equation (79), Equation (74) is proved.

By Hölder's inequality, we have

$$\mathbb{E}\big[\mathrm{tr}\big(|D(U_x, U_y)|\big)\big]^{2/3} \mathbb{E}\big[\mathrm{tr}\left(|D(U_x, U_y)|^4\right)\big]^{1/3} \geq \mathbb{E}\big[\mathrm{tr}\big(|D(U_x, U_y)|^2\big)\big]\,.$$

Then, using Equation (64) and Equation (74), we know that

$$\mathbb{E}\big[\mathrm{tr}\big(|D(U_x, U_y)|\big)\big] \geq \sqrt{\frac{d_1^6}{8000 \cdot r^3} \cdot \frac{r^3}{12288 \cdot d_1^4}} \geq \frac{d_1}{9915}.$$

Finally, using Equation (63), we conclude that

$$\frac{1}{d_1} \mathbb{E}\Big[\Big\|C_{\mathcal{E}_{U_x}} - C_{\mathcal{E}_{U_y}}\Big\|_1\Big] \geq \frac{\varepsilon}{d_1} \mathbb{E}\big[\|D(U_x, U_y)\|_1\big] - 2\varepsilon^2 \geq \frac{\varepsilon}{9915} - 2\varepsilon^2 \geq \frac{\varepsilon}{19830},$$

where we used $\varepsilon \leq 1/39660$.

5.2 Type II instance: $d_1 < rd_2$ with $rd_2 < d_1 + r$

5.2.1 Construction

In this subsection, we will use the definition in Section 4.2.1, where the parameter d, D in Section 4.2.1 correspond to d_1 and rd_2 here.

Suppose $d_2 > 1$ and $rd_2 < d_1 + r$. Let $\kappa := \min\{(rd_2 - d_1)/d_1, 1\} > 0$ and $\eta := d_1 - r\lfloor \frac{d_1}{r} \rfloor$. One can see that:

- $d_1 + r > rd_2 \geq 2r$ and thus $d_1 > r$, which means $rd_2 < d_1 + r < 2d_1$. Thus $\kappa = (rd_2 - d_1)/d_1 < 1$.

- $d_2 - 1 < d_1/r < d_2$, thus $\lfloor \frac{d_1}{r} \rfloor = d_2 - 1 \geq 1$ and $1 \leq \eta \leq r - 1$.

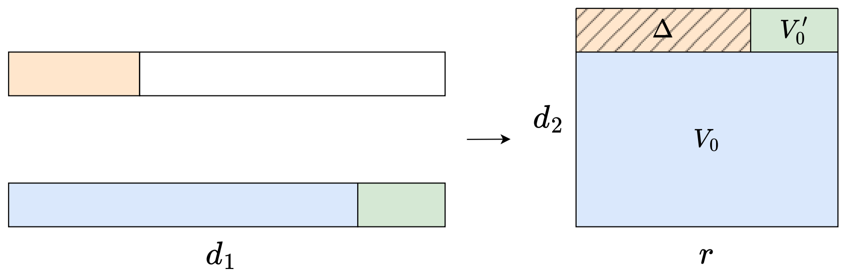


Figure 5: Illustration of our construction. We define linear operators V_0, V_0' and Δ from the d_1-dimensional space $\mathcal{H}_{\mathrm{A}}$ to the d_2r-dimensional space $\mathcal{H}_{\mathrm{B}} \otimes \mathcal{H}_{\mathrm{anc}}$. The Haar randomness is applied on the hatched area.

- $\kappa d_1 = rd_2 - d_1 \leq r$.

Let $\varepsilon \in (0, 1/2)$ and $\Sigma_\varepsilon := \sqrt{1-\varepsilon^2}\sum_{i=1}^{r-\eta}|i\rangle\langle i|_{\mathrm{A}} + \sum_{i=r-\eta+1}^{r(d_2-1)}|i\rangle\langle i|_{\mathrm{A}}$ be a diagonal matrix. Let $g : [d_1] \to [d_2] \times [r]$ be the function $g(i) = (\lfloor \frac{i-1}{r} \rfloor + 1, i - r\lfloor \frac{i-1}{r} \rfloor)$, i.e., g maps the integers in $[d_1]$ to $[d_2] \times [r]$ in row-major order. Then, we define the linear operator $V_0 : \mathcal{H}_{\mathrm{A}}[1 : r(d_2-1)] \to \mathcal{H}_{\mathrm{B}}[1 : d_2 - 1] \otimes \mathcal{H}_{\mathrm{anc}}$ as

$$V_0 := \sum_{i=1}^{r(d_2-1)} \Big(|g(i)_1\rangle_{\mathrm{B}} \otimes |g(i)_2\rangle_{\mathrm{anc}}\langle i|_{\mathrm{A}}\Big) \cdot \Sigma_\varepsilon. \tag{80}$$

One can easily check that V_0 can be written as

$$V_0 = \sum_{i=1}^{r} |i\rangle_{\mathrm{anc}} \otimes K_i, \tag{81}$$

for $K_i : \mathcal{H}_{\mathrm{A}}[1 : r(d_2-1)] \to \mathcal{H}_{\mathrm{B}}[1 : d_2 - 1]$ satisfying

$$\sum_{i=1}^{r} K_i^\dagger K_i = \Sigma_\varepsilon^2 \quad \text{and} \quad \frac{d_2-1}{2} \cdot \mathbb{1}_{i=j} \leq \left|\mathrm{tr}\Big(K_i^\dagger K_j\Big)\right| \leq (d_2-1) \cdot \mathbb{1}_{i=j}. \tag{82}$$

Define isometry $V_0' : \mathcal{H}_{\mathrm{A}}[r(d_2-1)+1 : d_1] \to \mathcal{H}_{\mathrm{B}}[d_2] \otimes \mathcal{H}_{\mathrm{anc}}[r - \eta + 1 : r]$ as

$$V_0' := |d_2\rangle_{\mathrm{B}} \otimes \sum_{i=1}^{\eta} |i + r - \eta\rangle_{\mathrm{anc}}\langle i + r(d_2-1)|_{\mathrm{A}}.$$

Define isometry $\Delta : \mathcal{H}_{\mathrm{A}}[1 : r - \eta] \to \mathcal{H}_{\mathrm{B}}[d_2] \otimes \mathcal{H}_{\mathrm{anc}}[1 : r - \eta]$ as

$$\Delta := |d_2\rangle_{\mathrm{B}} \otimes \sum_{i=1}^{r-\eta} |i\rangle_{\mathrm{anc}}\langle i|_{\mathrm{A}}$$

Then, for $U \in \mathbb{U}_{r-\eta}$, we define the isometry $V_{\varepsilon,U} : \mathcal{H}_{\mathrm{A}} \to \mathcal{H}_{\mathrm{B}} \otimes \mathcal{H}_{\mathrm{anc}}$ as

$$V_{\varepsilon,U} := V_0 + V_0' + \varepsilon U\Delta, \tag{83}$$

where U acts on $\mathcal{H}_{\mathrm{B}}[d_2] \otimes \mathcal{H}_{\mathrm{anc}}[1 : r - \eta]$. We can verify that $V_{\varepsilon,U}$ is indeed an isometry from Equation (80) and the definition of $U\Delta$. Furthermore, we note that the image of $U\Delta$ is orthogonal to the image of $V_0 + V_0'$. For clarity, we also illustrate our construction in Figure 5.

5.2.2 Existence for Choi trace norm

Then, we prove that there exists a large set of isometries $V_{\varepsilon,U}$ with good separation property.

Theorem 5.4. *Recall that $\kappa = (rd_2 - d_1)/d_1$ in this regime. Suppose $\varepsilon \leq \kappa^{3/2}/9200$. There exists a finite subset $\mathcal{N}$ of $\{V_{\varepsilon,U} \,|\, U \in \mathbb{U}_{r-\eta}\}$ for $V_{\varepsilon,U}$ defined in Equation (83) with cardinality $|\mathcal{N}| \geq \exp(\kappa^2(rd_2 - d_1)^2/480001)$, such that for any $V_x \neq V_y \in \mathcal{N}$, if we set $\mathcal{E}_x = \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(V_x(\cdot)V_x^\dagger)$ and $\mathcal{E}_y = \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(V_y(\cdot)V_y^\dagger)$, then*

$$\frac{1}{d_1}\|C_{\mathcal{E}_x} - C_{\mathcal{E}_y}\|_1 \geq \frac{\kappa^{3/2}}{4600}\varepsilon.$$

Proof. First, we need the following lemma.

Lemma 5.5. *There exists a finite subset $\mathcal{M} \subseteq \mathbb{U}_{r-\eta}$ with cardinality $|\mathcal{M}| \geq \exp(\kappa^2(rd_2-d_1)^2/480001)$ such that for any $U_x \neq U_y \in \mathcal{M}$,*

$$\frac{1}{d_1}\left\|\mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big((|V_0\rangle\rangle + |V_0'\rangle\rangle)\Big(\langle\langle U_x\Delta| - \langle\langle U_y\Delta|\Big)\Big)\right\|_1 \geq \frac{\kappa^{3/2}}{4600}. \tag{84}$$

The proof of Lemma 5.5 is deferred to Section 5.2.3.

Now, we are able to prove Theorem 5.4. Let $\mathcal{M}$ be the set given in Lemma 5.5. For any $U_x \neq U_y \in \mathcal{M}$, if we set $\mathcal{E}_x = \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(V_{\varepsilon,U_x}(\cdot)V_{\varepsilon,U_x}^\dagger)$ and $\mathcal{E}_y = \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(V_{\varepsilon,U_y}(\cdot)V_{\varepsilon,U_y}^\dagger)$, then we have

$$\begin{aligned}
&\|C_{\mathcal{E}_x} - C_{\mathcal{E}_y}\|_1 \\
=&\Big\| \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(|V_{\varepsilon,U_x}\rangle\rangle\langle\langle V_{\varepsilon,U_x}|) - \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(|V_{\varepsilon,U_y}\rangle\rangle\langle\langle V_{\varepsilon,U_y}|) \Big\|_1 \\
=&\Big\|\varepsilon^2\Big(\mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(|U_x\Delta\rangle\rangle\langle\langle U_x\Delta|) - \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(|U_y\Delta\rangle\rangle\langle\langle U_y\Delta|) \Big) \\
&\qquad + \varepsilon\, \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big(|V_0 + V_0'\rangle\rangle\langle\langle (U_x - U_y)\Delta|\Big) + \varepsilon\, \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big(|(U_x - U_y)\Delta\rangle\rangle\langle\langle V_0 + V_0'|\Big)\Big\|_1 \\
\geq&\varepsilon\Big\| \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big(|V_0 + V_0'\rangle\rangle\langle\langle (U_x - U_y)\Delta|\Big) + \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big(|(U_x - U_y)\Delta\rangle\rangle\langle\langle V_0 + V_0'|\Big)\Big\|_1 - 2\varepsilon^2 d_1 &(85)\\
=&2\varepsilon\Big\| \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big(|V_0 + V_0'\rangle\rangle\langle\langle (U_x - U_y)\Delta|\Big)\Big\|_1 - 2\varepsilon^2 d_1 &(86)\\
\geq&\frac{2\kappa^{3/2}}{4600}\varepsilon d_1 - 2\varepsilon^2 d_1 &(87)\\
\geq&\frac{\kappa^{3/2}}{4600}\varepsilon d_1. &(88)
\end{aligned}$$

In Equation (85) we used

$$\|\mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(|U\Delta\rangle\rangle\langle\langle U\Delta|)\|_1 = \mathrm{tr}(|U\Delta\rangle\rangle\langle\langle U\Delta|) = r - \eta \leq r < d_1.$$

In Equation (86) we used the fact that $\mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big(|V_0 + V_0'\rangle\rangle\langle\langle (U_x - U_y)\Delta|\Big)$ is a linear operator with support in $\mathcal{H}_{\mathrm{A}}[1 : r - \eta] \otimes \mathcal{H}_{\mathrm{B}}[d_2]$ and image in

$$\big(\mathcal{H}_{\mathrm{A}}[1 : r(d_2 - 1)] \otimes \mathcal{H}_{\mathrm{B}}[1 : d_2 - 1]\big) \oplus \big(\mathcal{H}_{\mathrm{A}}[r(d_2 - 1) + 1 : d_1] \otimes \mathcal{H}_{\mathrm{B}}[d_2]\big),$$

which is orthogonal to its support (noting that $r-\eta \leq r(d_2-1)$); and then we used Fact 7.9. In Equation (87) we used Equation (84). In Equation (88) we used that $\varepsilon \leq \kappa^{3/2}/9200$. Therefore, we can lower bound the Choi trace norm

$$\frac{1}{d_1}\left\|C_{\mathcal{E}_x}-C_{\mathcal{E}_y}\right\|_1 \geq \frac{\kappa^{3/2}}{4600}\varepsilon.$$

Thus, the set $\mathcal{N}=\{V_{\varepsilon,U}\,|\,U\in\mathcal{M}\}$ is the desired set. □

5.2.3 Proof of Lemma 5.5

Proof. Let $\hat{V}_0 \coloneqq V_0+V_0'$. Then, we need the following lemma:

Lemma 5.6. *For $U_x, U_y \in \mathbb{U}_{r-\eta}$, let us define*

$$F(U_x,U_y)=\frac{1}{d_1}\operatorname{tr}_{\mathcal{H}_{\text{anc}}}\Big(|\hat{V}_0\rangle\!\rangle\Big(\langle\!\langle U_x\Delta|-\langle\!\langle U_y\Delta|\Big)\Big),$$

then the function $f(U_x,U_y)=\|F(U_x,U_y)\|_1=\operatorname{tr}(|F(U_x,U_y)|)$ is $\sqrt{\frac{2}{d_1}}$-Lipschitz with respect to the ℓ_2-sum of the 2-norms (Frobenius norm). Furthermore, for independent random $U_x,U_y\sim\mathbb{U}_{r-\eta}$, we have $\mathbb{E}\big[\operatorname{tr}(|F(U_x,U_y)|^2)\big]\geq\frac{\kappa}{2r}$, and $\mathbb{E}\big[\operatorname{tr}(|F(U_x,U_y)|^4)\big]\leq\frac{256}{r^3}$.

By the Hölder's inequality we have

$$\mathbb{E}\big[\operatorname{tr}\big(|F(U_x,U_y)|^2\big)\big]\leq\mathbb{E}\big[\operatorname{tr}\big(|F(U_x,U_y)|^4\big)\big]^{1/3}\,\mathbb{E}[\operatorname{tr}(|F(U_x,U_y)|)]^{2/3},$$

which, combined with Lemma 5.6, implies

$$\mathbb{E}[\operatorname{tr}(|F(U_x,U_y)|)]^2\geq\frac{\kappa^3}{2048}.$$

Thus $\mathbb{E}[\operatorname{tr}(|F(U_x,U_y)|)]>\kappa^{3/2}/46$. Then, since the function $f(U_x,U_y)=\operatorname{tr}(|F(U_x,U_y)|)$ is $\sqrt{\frac{2}{d_1}}$-Lipschitz, we can use Theorem 7.10 to prove the concentration result:

$$\mathbb{P}\left[\operatorname{tr}(|F(U_x,U_y)|)\leq\frac{\kappa^{3/2}}{4600}\right]\leq\exp\!\left(-\frac{d_1(r-\eta)}{2}\cdot\frac{\kappa^3}{10000\cdot 12}\right)\leq\exp\!\left(-\frac{\kappa^2(rd_2-d_1)^2}{240000}\right),$$

where we used $r-\eta\geq\kappa d_1=rd_2-d_1$. Then, we independently sample $\exp(\kappa^2(rd_2-d_1)^2/480001)$ Haar random unitaries in $\mathbb{U}_{r-\eta}$ and the union bound shows that there exists a non-zero probability that for any pair U_x,U_y, we have $\operatorname{tr}(|F(U_x,U_y)|)\geq\kappa^{3/2}/4600$. Thus, there exists a set with cardinality $\geq\exp(\kappa^2(rd_2-d_1)^2/480001)$ such that Equation (84) holds. □

Then, we give the proof of Lemma 5.6.

Proof of Lemma 5.6. The Lipschitz continuity can be seen from Lemma 7.3.

Note that $\hat{V}_0$ can also be written as

$$\hat{V}_0=\sum_{i=1}^{r}|i\rangle_{\text{anc}}\otimes\big(K_i\oplus|d_2\rangle_{\text{B}}\langle z_i|_{\text{A}}\big),$$

where K_i are defined in Equation (81), and we set $|z_i\rangle_{\text{A}}=|i-r+\eta+r(d_2-1)\rangle_{\text{A}}$ for $r-\eta+1\leq i\leq r$ and $|z_i\rangle_{\text{A}}=0$ for $1\leq i\leq r-\eta$. Let us define

$$K_i'=K_i\oplus|d_2\rangle_{\text{B}}\langle z_i|_{\text{A}},$$

then K'_i satisfy

$$\left|\mathrm{tr}\left(K_i'^\dagger K_j'\right)\right| = \left|\mathrm{tr}\left(K_i^\dagger K_j\right) + \mathbb{1}_{r-\eta+1\le i=j\le r}\right| \le d_2 \cdot \mathbb{1}_{i=j}, \tag{89}$$

and

$$\mathrm{tr}\left(K_i'^\dagger K_i'\right) = \mathrm{tr}\left(K_i^\dagger K_i\right) + \mathbb{1}_{r-\eta+1\le i\le r} \ge \mathrm{tr}\left(K_i^\dagger K_i\right) \ge \frac{d_2-1}{2}, \tag{90}$$

due to Equation (82).

Define $K_{x,i} = \langle i|_{\mathrm{anc}} U_x \Delta$, $K_{y,i} = \langle i|_{\mathrm{anc}} U_y \Delta$. This means

$$F(U_x, U_y) = \frac{1}{d_1}\sum_{i=1}^r |K_i'\rangle\rangle\left(\langle\langle K_{x,i}| - \langle\langle K_{y,i}|\right).$$

Then, we note that

$$\begin{aligned}
\mathbb{E}\left[\mathrm{tr}(|F(U_x,U_y)|^2)\right] &= \frac{1}{d_1^2}\,\mathbb{E}\left[\mathrm{tr}\left(\sum_{i,j=1}^r |K_i'\rangle\rangle\left(\langle\langle K_{x,i}| - \langle\langle K_{y,i}|\right)\left(|K_{x,j}\rangle\rangle - |K_{y,j}\rangle\rangle\right)\langle\langle K_j'|\right)\right] \\
&= \frac{1}{d_1^2}\,\mathbb{E}\left[\sum_{i=1}^r \langle\langle K_i'|K_i'\rangle\rangle\left(\langle\langle K_{x,i}| - \langle\langle K_{y,i}|\right)\left(|K_{x,i}\rangle\rangle - |K_{y,i}\rangle\rangle\right)\right] \\
&= \frac{2}{d_1^2}\sum_{i=1}^{r-\eta}\langle\langle K_i'|K_i'\rangle\rangle \qquad (91)\\
&\ge \frac{2}{d_1^2}(r-\eta)\frac{d_2-1}{2} \qquad (92)\\
&\ge \frac{\kappa(d_2-1)}{d_1} \ge \frac{\kappa}{2r} \qquad (93)
\end{aligned}$$

where Equation (91) is because for $t_1, t_2 \in \{x, y\}$, we have

$$\begin{aligned}
\mathbb{E}[\langle\langle K_{t_1,i}|K_{t_2,i}\rangle\rangle] = \mathbb{E}\left[\mathrm{tr}\left(K_{t_1,i}^\dagger K_{t_2,i}\right)\right] &= \mathrm{tr}\left(\Delta^\dagger\,\mathbb{E}\left[U_{t_1}^\dagger |i\rangle\langle i|_{\mathrm{anc}} U_{t_2}\right]\Delta\right) \\
&= \mathbb{1}_{t_1=t_2}\mathbb{1}_{1\le i\le r-\eta}\cdot\frac{1}{r-\eta}\,\mathrm{tr}(\Delta^\dagger\Delta) \qquad (94)\\
&= \mathbb{1}_{t_1=t_2}\mathbb{1}_{1\le i\le r-\eta},
\end{aligned}$$

where Equation (94) is due to Schur's lemma (noting that U_{z_1} and U_{z_2} are unitaries acting on the $(r-\eta)$-dimensional space $\mathcal{H}_{\mathrm{anc}}[1:r-\eta]\otimes\mathcal{H}_{\mathrm{B}}[d_2]$) and also the fact that $\langle i|_{\mathrm{anc}}U_x\Delta = 0$ for $i \ge r-\eta+1$, Equation (92) is by Equation (90), and Equation (93) is because $r-\eta \ge \kappa d_1$ and $d_2 - 1 \ge d_2/2 \ge d_1/(2r)$.

Then, we have

$$\begin{aligned}
\mathbb{E}\left[\mathrm{tr}(|F(U_x,U_y)|^4)\right] &= \frac{1}{d_1^4}\,\mathbb{E}\Bigg[\sum_{i,j,k,l=1}^r \mathrm{tr}\Big(|K_i'\rangle\rangle\left(\langle\langle K_{x,i}| - \langle\langle K_{y,i}|\right)\left(|K_{x,j}\rangle\rangle - |K_{y,j}\rangle\rangle\right)\langle\langle K_j'| \\
&\qquad |K_k'\rangle\rangle\left(\langle\langle K_{x,k}| - \langle\langle K_{y,k}|\right)\left(|K_{x,l}\rangle\rangle - |K_{y,l}\rangle\rangle\right)\langle\langle K_l'|\Big)\Bigg] \\
&\le \frac{d_2^2}{d_1^4}\sum_{i,j=1}^r \mathbb{E}\left[\left|\left(\langle\langle K_{x,i}| - \langle\langle K_{y,i}|\right)\left(|K_{x,j}\rangle\rangle - |K_{y,j}\rangle\rangle\right)\right|^2\right] \qquad (95)
\end{aligned}$$

$$\leq \frac{4d_2^2}{d_1^4} \sum_{i,j=1}^{r} \mathbb{E}\big[|\langle\langle K_{x,i}|K_{x,j}\rangle\rangle|^2 + |\langle\langle K_{y,i}|K_{y,j}\rangle\rangle|^2 + |\langle\langle K_{x,i}|K_{y,j}\rangle\rangle|^2 + |\langle\langle K_{y,i}|K_{x,j}\rangle\rangle|^2\big]$$

$$\leq \frac{8d_2^2}{d_1^4} \sum_{i,j=1}^{r} \mathbb{E}\big[|\langle\langle K_{x,i}|K_{x,j}\rangle\rangle|^2 + |\langle\langle K_{y,i}|K_{y,j}\rangle\rangle|^2\big] \tag{96}$$

$$= \frac{16d_2^2}{d_1^4} \sum_{i,j=1}^{r} \mathbb{E}\big[|\langle\langle K_{x,i}|K_{x,j}\rangle\rangle|^2\big] = \frac{16d_2^2}{d_1^4} \sum_{i,j=1}^{r} \mathbb{E}\left[\left|\mathrm{tr}(K_{x,i}^\dagger K_{x,j})\right|^2\right],$$

$$\leq \frac{16d_2^2}{d_1^4} \cdot \frac{4(r-\eta)^2}{r-\eta} \leq \frac{256}{rd_1^2} \leq \frac{256}{r^3}, \tag{97}$$

where Equation (95) is because Equation (89), Equation (96) is because

$$\sum_{i,j=1}^{r} |\langle\langle K_{x,i}|K_{y,j}\rangle\rangle|^2 = \|K_x^\dagger K_y\|_F^2 = \mathrm{tr}\Big(K_x^\dagger K_y K_y^\dagger K_x\Big) \leq \|K_x K_x^\dagger\|_F \cdot \|K_y K_y^\dagger\|_F$$

$$\leq \frac{1}{2}\left(\|K_x^\dagger K_x\|_F^2 + \|K_y^\dagger K_y\|_F^2\right) = \frac{1}{2}\left(\sum_{i,j=1}^{r} |\langle\langle K_{x,i}|K_{x,j}\rangle\rangle|^2 + \sum_{i,j=1}^{r} |\langle\langle K_{y,i}|K_{y,j}\rangle\rangle|^2\right), \tag{98}$$

where K_x denotes the matrix with columns $|K_{x,i}\rangle\rangle$, and in Equation (97), the first inequality is due to Lemma 7.4 and that Δ is an isometry from $\mathcal{H}_{\mathrm{A}}[1:r-\eta]$ to $\mathcal{H}_{\mathrm{B}}[d_2] \otimes \mathcal{H}_{\mathrm{anc}}[1:r-\eta]$ (further note that $\dim(\mathcal{H}_{\mathrm{anc}}[1:r-\eta]) \leq \dim(\mathcal{H}_{\mathrm{A}}[1:r-\eta])\dim(\mathcal{H}_{\mathrm{B}}[d_2])$ so we choose $k=1$ in Lemma 7.4), the second and third inequalities is due to $d_2 \leq d_1/r + 1 \leq 2d_1/r$ and $r \leq d_1$. □

5.2.4 Existence for diamond norm

Note that in Theorem 5.4, the cardinality and distance both depend on κ, which can be arbitrarily close to 0. Here, we give another packing nets w.r.t. diamond norm that does not depend on κ. We will use the same construction as that given in Section 5.2.1. Specifically, recall that $\eta \coloneqq d_1 - r\lfloor \frac{d_1}{r} \rfloor = d_1 - rd_2 + r$. Note that $r - \eta = rd_2 - d_1$.

Theorem 5.7. *Suppose $\varepsilon \leq 1/160$. There exists a finite subset $\mathcal{N}$ of $\{V_{\varepsilon,U} \,|\, U \in \mathbb{U}_{r-\eta}\}$ for $V_{\varepsilon,U}$ defined in Equation (83) with cardinality $|\mathcal{N}| \geq \exp((rd_2 - d_1)^2/4801)$, such that for any $V_x \neq V_y \in \mathcal{N}$, if we set $\mathcal{E}_x = \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(V_x(\cdot)V_x^\dagger)$ and $\mathcal{E}_y = \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(V_y(\cdot)V_y^\dagger)$, then*

$$\|\mathcal{E}_x - \mathcal{E}_y\|_\diamond \geq \frac{1}{80}\varepsilon.$$

Proof. We define $|\Psi\rangle \coloneqq \frac{1}{\sqrt{r-\eta}}\sum_{i=1}^{r-\eta}|i\rangle_{\mathrm{A}}|i\rangle_{\mathrm{A}}$ be an entangled state on $\mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{A}}$ and $\Psi = |\Psi\rangle\langle\Psi|$. We also define $W_0 : \mathcal{H}_{\mathrm{A}}[1:r-\eta] \to \mathcal{H}_{\mathrm{B}}[1] \otimes \mathcal{H}_{\mathrm{anc}}[1:r-\eta]$ as:

$$W_0 \coloneqq \sqrt{1-\varepsilon^2}\sum_{i=1}^{r-\eta}|i\rangle_{\mathrm{anc}} \otimes |1\rangle_{\mathrm{B}}\langle i|_{\mathrm{A}}.$$

We can easily see that

$$V_0|\Psi\rangle = W_0|\Psi\rangle = \frac{1}{\sqrt{r-\eta}}|W_0\rangle\rangle, \quad V_0'|\Psi\rangle = 0. \tag{99}$$

Then, we need the following lemma.

Lemma 5.8. *There exists a finite subset $\mathcal{M} \subseteq \mathbb{U}_{r-\eta}$ with cardinality $|\mathcal{M}| \geq \exp((rd_2 - d_1)^2/4801)$ such that for any $U_x \neq U_y \in \mathcal{M}$,*

$$\frac{1}{r-\eta} \left\| \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}} \Big(|W_0\rangle\!\rangle \Big(\langle\!\langle U_x \Delta| - \langle\!\langle U_y \Delta| \Big) \Big) \right\|_1 \geq \frac{1}{80}. \tag{100}$$

The proof of Lemma 5.8 is deferred to Section 5.2.5.

Now, we are able to prove Theorem 5.7. Let $\mathcal{M}$ be the set given in Lemma 5.8. For any $U_x \neq U_y \in \mathcal{M}$, if we set $\mathcal{E}_x = \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(V_{\varepsilon,U_x}(\cdot)V_{\varepsilon,U_x}^\dagger)$ and $\mathcal{E}_y = \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(V_{\varepsilon,U_y}(\cdot)V_{\varepsilon,U_y}^\dagger)$, then we have

$$\begin{aligned}
&\|\mathcal{E}_x - \mathcal{E}_y\|_\diamond \geq \|\mathcal{E}_x(\Psi) - \mathcal{E}_y(\Psi)\|_1 \\
=&\left\| \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}} \Big((W_0 + \varepsilon U_x \Delta)\Psi(W_0 + \varepsilon U_x \Delta)^\dagger \Big) - \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}} \Big((W_0 + \varepsilon U_y \Delta)\Psi(W_0 + \varepsilon U_y \Delta)^\dagger \Big) \right\|_1 &(101)\\
=&\frac{1}{r-\eta} \Bigg\| \varepsilon^2 \Big(\mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(|U_x\Delta\rangle\!\rangle\langle\!\langle U_x\Delta|) - \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(|U_y\Delta\rangle\!\rangle\langle\!\langle U_y\Delta|) \Big) \\
&\qquad + \varepsilon\, \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big(|W_0\rangle\!\rangle\langle\!\langle (U_x - U_y)\Delta| \Big) + \varepsilon\, \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big(|(U_x - U_y)\Delta\rangle\!\rangle\langle\!\langle W_0| \Big) \Bigg\|_1 &(102)\\
\geq&\frac{\varepsilon}{r-\eta} \left\| \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big(|W_0\rangle\!\rangle\langle\!\langle (U_x - U_y)\Delta| \Big) + \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big(|(U_x - U_y)\Delta\rangle\!\rangle\langle\!\langle W_0| \Big) \right\|_1 - 2\varepsilon^2 &(103)\\
=&\frac{2\varepsilon}{r-\eta} \left\| \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big(|W_0\rangle\!\rangle\langle\!\langle (U_x - U_y)\Delta| \Big) \right\|_1 - 2\varepsilon^2 &(104)\\
\geq&\frac{1}{40}\varepsilon - 2\varepsilon^2 &(105)\\
\geq&\frac{1}{80}\varepsilon. &(106)
\end{aligned}$$

In Equation (101), we used Equation (99). In Equation (102), we used $U\Delta|\Psi\rangle = \frac{1}{\sqrt{r-\eta}}|U\Delta\rangle\!\rangle$, $W_0|\Psi\rangle = \frac{1}{\sqrt{r-\eta}}|W_0\rangle\!\rangle$. In Equation (103), we used

$$\|\,\mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(|U\Delta\rangle\!\rangle\langle\!\langle U\Delta|)\|_1 = \mathrm{tr}(|U\Delta\rangle\!\rangle\langle\!\langle U\Delta|) = r - \eta.$$

In Equation (104) we used the fact that $\mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big(|W_0\rangle\!\rangle\langle\!\langle (U_x - U_y)\Delta|\Big)$ is a linear operator with support in $\mathcal{H}_{\mathrm{A}}[1 : r - \eta] \otimes \mathcal{H}_{\mathrm{B}}[d_2]$ and image in $\mathcal{H}_{\mathrm{A}}[1 : r - \eta] \otimes \mathcal{H}_{\mathrm{B}}[1]$, which is orthogonal to its support; and then we used Fact 7.9. In Equation (105) we used Equation (100). In Equation (106) we used that $\varepsilon \leq 1/160$. Therefore, we can lower bound the diamond norm

$$\|\mathcal{E}_x - \mathcal{E}_y\|_\diamond \geq \frac{1}{80}\varepsilon.$$

Thus, the set $\mathcal{N} = \{V_{\varepsilon,U} \,|\, U \in \mathcal{M}\}$ is the desired set. □

5.2.5 Proof of Lemma 5.8

Proof. We need the following lemma:

Lemma 5.9. *For $U_x, U_y \in \mathbb{U}_{r-\eta}$, let us define*

$$F(U_x, U_y) = \frac{1}{r-\eta}\, \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big(|W_0\rangle\!\rangle \Big(\langle\!\langle U_x\Delta| - \langle\!\langle U_y\Delta| \Big) \Big),$$

then the function $f(U_x, U_y) = \|F(U_x, U_y)\|_1 = \mathrm{tr}(|F(U_x, U_y)|)$ is $\sqrt{\frac{2}{r-\eta}}$-Lipschitz with respect to the ℓ_2-sum of the 2-norms (Frobenius norm). Furthermore, for independent random $U_x, U_y \sim \mathbb{U}_{r-\eta}$, we have $\mathbb{E}\big[\mathrm{tr}(|F(U_x, U_y)|^2)\big] \geq \frac{1}{r-\eta}$, and $\mathbb{E}\big[\mathrm{tr}(|F(U_x, U_y)|^4)\big] \leq \frac{64}{(r-\eta)^3}$.

By the Hölder's inequality we have

$$\mathbb{E}\big[\mathrm{tr}\big(|F(U_x,U_y)|^2\big)\big] \le \mathbb{E}\big[\mathrm{tr}\big(|F(U_x,U_y)|^4\big)\big]^{1/3}\,\mathbb{E}[\mathrm{tr}(|F(U_x,U_y)|)]^{2/3},$$

which, combined with Lemma 5.9, implies

$$\mathbb{E}[\mathrm{tr}(|F(U_x,U_y)|)]^2 \ge \frac{1}{64}.$$

Thus $\mathbb{E}[\mathrm{tr}(|F(U_x,U_y)|)] > 1/8$. Then, since the function $f(U_x,U_y) = \mathrm{tr}(|F(U_x,U_y)|)$ is $\sqrt{\frac{2}{r-\eta}}$-Lipschitz, we can use Theorem 7.10 to prove the concentration result:

$$\mathbb{P}\left[\mathrm{tr}(|F(U_x,U_y)|) \le \frac{1}{80}\right] \le \exp\left(-\frac{(r-\eta)^2}{2}\cdot\frac{1}{100\cdot 12}\right) \le \exp\left(-\frac{(r-\eta)^2}{2400}\right).$$

Then, we independently sample $\exp((r-\eta)^2/4801)$ Haar random unitaries in $\mathbb{U}_{r-\eta}$ and the union bound shows that there exists a non-zero probability that for any pair U_x, U_y, we have $\mathrm{tr}(|F(U_x,U_y)|) \ge 1/80$. Thus, there exists a set with cardinality $\ge \exp((r-\eta)^2/4801) = \exp((rd_2-d_1)^2/4801)$ such that Equation (100) holds. □

Then, we give the proof of Lemma 5.9.

Proof of Lemma 5.9. The Lipschitz continuity can be seen from Lemma 7.3, where we treat W_0 and $U\Delta$ as linear operators acting on input space with dimension $r-\eta$.

Define $K_{x,i} = \langle i|_{\mathrm{anc}} U_x\Delta$, $K_{y,i} = \langle i|_{\mathrm{anc}} U_y\Delta$, and $K'_i = |1\rangle_{\mathrm{B}}\langle i|_{\mathrm{A}} = \frac{1}{\sqrt{1-\varepsilon^2}}\langle i|_{\mathrm{anc}} W_0$. This means

$$F(U_x,U_y) = \frac{\sqrt{1-\varepsilon^2}}{r-\eta}\sum_{i=1}^{r-\eta}|K'_i\rangle\!\rangle\Big(\langle\!\langle K_{x,i}| - \langle\!\langle K_{y,i}|\Big).$$

Then, we note that

$$\begin{aligned}
\mathbb{E}\big[\mathrm{tr}\big(|F(U_x,U_y)|^2\big)\big] &= \frac{1-\varepsilon^2}{(r-\eta)^2}\,\mathbb{E}\left[\mathrm{tr}\left(\sum_{i,j=1}^{r-\eta}|K'_i\rangle\!\rangle\Big(\langle\!\langle K_{x,i}| - \langle\!\langle K_{y,i}|\Big)\Big(|K_{x,j}\rangle\!\rangle - |K_{y,j}\rangle\!\rangle\Big)\langle\!\langle K'_j|\right)\right] \\
&= \frac{1-\varepsilon^2}{(r-\eta)^2}\,\mathbb{E}\left[\sum_{i=1}^{r-\eta}\Big(\langle\!\langle K_{x,i}| - \langle\!\langle K_{y,i}|\Big)\Big(|K_{x,i}\rangle\!\rangle - |K_{y,i}\rangle\!\rangle\Big)\right] \\
&= \frac{2(1-\varepsilon^2)}{(r-\eta)^2}(r-\eta) = \frac{2(1-\varepsilon^2)}{r-\eta} \ge \frac{1}{r-\eta} \qquad (107)
\end{aligned}$$

where Equation (107) is because for $z_1, z_2 \in \{x,y\}$, we have

$$\begin{aligned}
\mathbb{E}[\langle\!\langle K_{z_1,i}|K_{z_2,i}\rangle\!\rangle] = \mathbb{E}\left[\mathrm{tr}\left(K^\dagger_{z_1,i}K_{z_2,i}\right)\right] &= \mathrm{tr}\left(\Delta^\dagger\,\mathbb{E}\left[U^\dagger_{z_1}|i\rangle\langle i|_{\mathrm{anc}}U_{z_2}\right]\Delta\right) \\
&= \mathbb{1}_{z_1=z_2}\mathbb{1}_{1\le i\le r-\eta}\cdot\frac{1}{r-\eta}\,\mathrm{tr}(\Delta^\dagger\Delta) \qquad (108) \\
&= \mathbb{1}_{z_1=z_2}\mathbb{1}_{1\le i\le r-\eta},
\end{aligned}$$

where Equation (108) is due to Schur's lemma and also the fact that $\langle i|_{\mathrm{anc}}U_x\Delta = 0$ for $i > r-\eta$.

Then, we also have

$$\begin{aligned}\mathbb{E}\big[\mathrm{tr}(|F(U_x,U_y)|^4)\big] &= \frac{1}{(r-\eta)^4}\,\mathbb{E}\Bigg[\sum_{i,j,k,l=1}^{r-\eta}\mathrm{tr}\Big(|K_i'\rangle\!\rangle\Big(\langle\!\langle K_{x,i}|-\langle\!\langle K_{y,i}|\Big)\Big(|K_{x,j}\rangle\!\rangle-|K_{y,j}\rangle\!\rangle\Big)\langle\!\langle K_j'| \\ &\qquad |K_k'\rangle\!\rangle\Big(\langle\!\langle K_{x,k}|-\langle\!\langle K_{y,k}|\Big)\Big(|K_{x,l}\rangle\!\rangle-|K_{y,l}\rangle\!\rangle\Big)\langle\!\langle K_l'|\Big)\Bigg] \\ &\le \frac{1}{(r-\eta)^4}\sum_{i,j=1}^{r-\eta}\mathbb{E}\Big[\Big|\Big(\langle\!\langle K_{x,i}|-\langle\!\langle K_{y,i}|\Big)\Big(|K_{x,j}\rangle\!\rangle-|K_{y,j}\rangle\!\rangle\Big)\Big|^2\Big] \\ &\le \frac{4}{(r-\eta)^4}\sum_{i,j=1}^{r-\eta}\mathbb{E}\big[|\langle\!\langle K_{x,i}|K_{x,j}\rangle\!\rangle|^2+|\langle\!\langle K_{y,i}|K_{y,j}\rangle\!\rangle|^2+|\langle\!\langle K_{x,i}|K_{y,j}\rangle\!\rangle|^2+|\langle\!\langle K_{y,i}|K_{x,j}\rangle\!\rangle|^2\big] \\ &\le \frac{8}{(r-\eta)^4}\sum_{i,j=1}^{r-\eta}\mathbb{E}\big[|\langle\!\langle K_{x,i}|K_{x,j}\rangle\!\rangle|^2+|\langle\!\langle K_{y,i}|K_{y,j}\rangle\!\rangle|^2\big] \qquad (109) \\ &= \frac{16}{(r-\eta)^4}\sum_{i,j=1}^{r-\eta}\mathbb{E}\big[|\langle\!\langle K_{x,i}|K_{x,j}\rangle\!\rangle|^2\big] = \frac{16}{(r-\eta)^4}\sum_{i,j=1}^{r}\mathbb{E}\Big[\Big|\mathrm{tr}(K_{x,i}^\dagger K_{x,j})\Big|^2\Big], \\ &\le \frac{16}{(r-\eta)^4}\cdot\frac{4(r-\eta)^2}{r-\eta}\le\frac{64}{(r-\eta)^3}, \qquad (110)\end{aligned}$$

where Equation (109) is due to a similar argument as that in Equation (98), and in Equation (110), the first inequality is due to Lemma 7.4 and that Δ is an isometry from $\mathcal{H}_{\mathrm{A}}[1:r-\eta]$ to $\mathcal{H}_{\mathrm{B}}[d_2]\otimes\mathcal{H}_{\mathrm{anc}}[1:r-\eta]$ (further note that $\dim(\mathcal{H}_{\mathrm{anc}}[1:r-\eta])\le\dim(\mathcal{H}_{\mathrm{A}}[1:r-\eta])\dim(\mathcal{H}_{\mathrm{B}}[d_2])$ so we choose $k=1$ in Lemma 7.4). □

5.3 Type II instance: $d_1 < rd_2$ with $d_1 + r \le rd_2$, $r \le d_1$

5.3.1 Construction

In this subsection, we will use the definition in Section 4.2.1, where the parameter d, D in Section 4.2.1 correspond to d_1 and rd_2 here.

Suppose $d_2 > 1$, $d_1 + r \le rd_2$ and $r \le d_1$. Let $\kappa \coloneqq \min\{(rd_2-d_1)/d_1, 1\}\in(0,1]$ and let

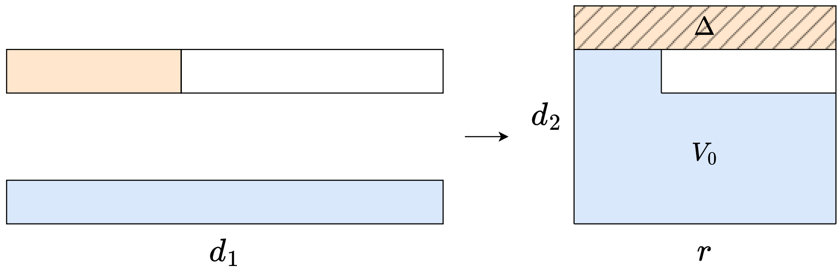
$$\underline{\chi}\coloneqq\lfloor\tfrac{d_1}{r}\rfloor,\quad \overline{\chi}\coloneqq\lceil\tfrac{d_1}{r}\rceil,\quad \zeta\coloneqq\min\{\underline{\chi},d_2-\overline{\chi}\}.$$

We note the following relations can be easily verified:

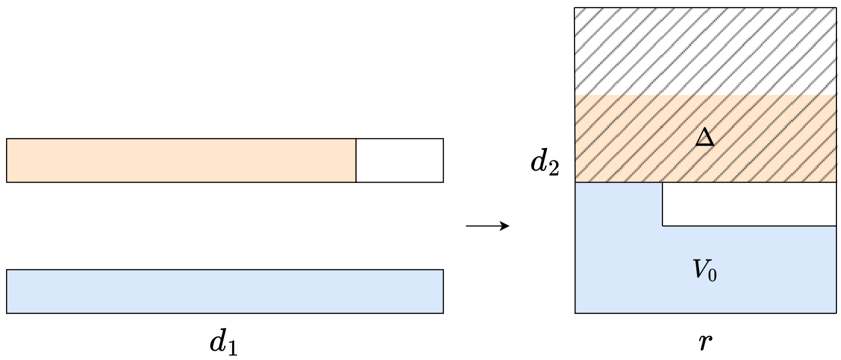
- $\underline{\chi}\ge 1$, and $d_2-\overline{\chi}\ge 1$ (since $d_1/r+1\le d_2$),
- $r(d_2-\overline{\chi})\ge rd_2-d_1-r\ge\frac{\kappa}{1+\kappa}rd_2-r$, thus $r(d_2-\overline{\chi})\ge\max\{r,\frac{\kappa}{1+\kappa}rd_2-r\}\ge\frac{\kappa}{2+2\kappa}rd_2\ge\frac{\kappa}{4}rd_2$,
- $r\underline{\chi}\ge d_1/2$ and $r(d_2-\overline{\chi})\ge\frac{\kappa}{4}rd_2>\frac{\kappa}{4}d_1$, and thus $r\zeta\ge\frac{\kappa}{4}d_1$.

Let $\varepsilon\in(0,1/2)$ and $\Sigma_\varepsilon\coloneqq\sqrt{1-\varepsilon^2}\sum_{i=1}^{r\zeta}|i\rangle\langle i|_{\mathrm{A}}+\sum_{i=r\zeta+1}^{d_1}|i\rangle\langle i|_{\mathrm{A}}$ be a diagonal matrix. Let $g:[d_1]\to[d_2]\times[r]$ be a bijective function $g(i)=(\lfloor\frac{i-1}{r}\rfloor+1, i-r\lfloor\frac{i-1}{r}\rfloor)$, i.e., g maps the integers in $[d_1]$ to $[d_2]\times[r]$ in row-major order. Then, we define the linear operator $V_0:\mathcal{H}_{\mathrm{A}}\to\mathcal{H}_{\mathrm{B}}[1:\overline{\chi}]\otimes\mathcal{H}_{\mathrm{anc}}$ as

$$V_0\coloneqq\sum_{i=1}^{d_1}\Big(|g(i)_1\rangle_{\mathrm{B}}\otimes|g(i)_2\rangle_{\mathrm{anc}}\langle i|_{\mathrm{A}}\Big)\cdot\Sigma_\varepsilon. \qquad (111)$$



(a) Case 1



(b) Case 2

Figure 6: Illustration of our construction. There are two cases depending on whether Case 1 : $\lfloor \frac{d_1}{r} \rfloor \geq d_2 - \lceil \frac{d_1}{r} \rceil$ or Case 2 : $\lfloor \frac{d_1}{r} \rfloor < d_2 - \lceil \frac{d_1}{r} \rceil$. We define linear operators V_0 and Δ from the d_1-dimensional space $\mathcal{H}_\mathrm{A}$ to the $d_2 r$-dimensional space $\mathcal{H}_\mathrm{B} \otimes \mathcal{H}_\mathrm{anc}$. The Haar randomness is applied on the hatched area.

One can easily check that V_0 can be written as

$$V_0 = \sum_{i=1}^{r} |i\rangle_\mathrm{anc} \otimes K_i, \tag{112}$$

for $K_i : \mathcal{H}_\mathrm{A} \to \mathcal{H}_\mathrm{B}[1 : \overline{\chi}]$ satisfying

$$\sum_{i=1}^{r} K_i^\dagger K_i = \Sigma_\varepsilon^2 \quad \text{and} \quad \frac{1}{2}\underline{\chi} \cdot \mathbb{1}_{i=j} \leq \left| \mathrm{tr}\left(K_i^\dagger K_j \right) \right| \leq \overline{\chi} \cdot \mathbb{1}_{i=j}. \tag{113}$$

Define $\Delta : \mathcal{H}_\mathrm{A}[1 : r\zeta] \to \mathcal{H}_\mathrm{B}[\overline{\chi} + 1 : d_2] \otimes \mathcal{H}_\mathrm{anc}$ be an arbitrary (but fixed) isometry. Then, for $U \in \mathbb{U}_{r(d_2 - \overline{\chi})}$, we define the isometry $V_{\varepsilon,U} : \mathcal{H}_\mathrm{A} \to \mathcal{H}_\mathrm{B} \otimes \mathcal{H}_\mathrm{anc}$ as

$$V_{\varepsilon,U} \coloneqq V_0 + \varepsilon U \Delta, \tag{114}$$

where U acts on $\mathcal{H}_\mathrm{B}[\overline{\chi} + 1 : d_2] \otimes \mathcal{H}_\mathrm{anc}$. We can verify that $V_{\varepsilon,U}$ is indeed an isometry from Equation (111) and the definition of $U\Delta$. Note that the image of $U\Delta$ is orthogonal to the image of V_0. For clarity, we also illustrate our construction in Figure 6.

5.3.2 Existence for Choi trace norm

Then, we prove that there exists a large set of isometries $V_{\varepsilon,U}$ with good separation property.

Theorem 5.10. *Suppose $\varepsilon \le \kappa^{3/2}/25600$. There exists a finite subset $\mathcal{N}$ of $\{V_{\varepsilon,U} \,|\, U \in \mathbb{U}_{r(d_2-\underline{\chi})}\}$ for $V_{\varepsilon,U}$ defined in Equation (114) with cardinality*

$$|\mathcal{N}| \ge \exp\left(\frac{\kappa^2(rd_2-d_1)\min\{d_1, rd_2-d_1\}}{3840001}\right),$$

such that for any $V_x \neq V_y \in \mathcal{N}$, if we set $\mathcal{E}_x = \operatorname{tr}_{\mathcal{H}_{\text{anc}}}(V_x(\cdot)V_x^\dagger)$ and $\mathcal{E}_y = \operatorname{tr}_{\mathcal{H}_{\text{anc}}}(V_y(\cdot)V_y^\dagger)$, then

$$\frac{1}{d_1}\|C_{\mathcal{E}_x} - C_{\mathcal{E}_y}\|_1 \ge \frac{\kappa^{3/2}}{12800}\varepsilon.$$

Proof. First, we need the following lemma.

Lemma 5.11. *There exists a finite subset $\mathcal{M} \subseteq \mathbb{U}_{r(d_2-\underline{\chi})}$ with cardinality $|\mathcal{M}| \ge \exp(\kappa^3 d_1(rd_2 - d_1)/3840001)$ such that for any $U_x \neq U_y \in \mathcal{M}$,*

$$\frac{1}{d_1}\left\|\operatorname{tr}_{\mathcal{H}_{\text{anc}}}\Big(|V_0\rangle\!\rangle\Big(\langle\!\langle U_x\Delta| - \langle\!\langle U_y\Delta|\Big)\Big)\right\|_1 \ge \frac{\kappa^{3/2}}{12800}. \tag{115}$$

The proof of Lemma 5.11 is deferred to Section 5.3.3.

Now, we are able to prove Theorem 5.10. Let $\mathcal{M}$ be the set given in Lemma 5.11. Then, for any $U_x \neq U_y \in \mathcal{M}$, if we set $\mathcal{E}_x = \operatorname{tr}_{\mathcal{H}_{\text{anc}}}(V_{\varepsilon,U_x}(\cdot)V_{\varepsilon,U_x}^\dagger)$ and $\mathcal{E}_y = \operatorname{tr}_{\mathcal{H}_{\text{anc}}}(V_{\varepsilon,U_y}(\cdot)V_{\varepsilon,U_y}^\dagger)$, we have

$$\begin{aligned}
&\|C_{\mathcal{E}_x} - C_{\mathcal{E}_y}\|_1 \\
=&\left\| \operatorname{tr}_{\mathcal{H}_{\text{anc}}}(|V_{\varepsilon,U_x}\rangle\!\rangle\langle\!\langle V_{\varepsilon,U_x}|) - \operatorname{tr}_{\mathcal{H}_{\text{anc}}}\big(|V_{\varepsilon,U_y}\rangle\!\rangle\langle\!\langle V_{\varepsilon,U_y}|\big) \right\|_1 \\
=&\left\|\varepsilon^2\Big(\operatorname{tr}_{\mathcal{H}_{\text{anc}}}(|U_x\Delta\rangle\!\rangle\langle\!\langle U_x\Delta|) - \operatorname{tr}_{\mathcal{H}_{\text{anc}}}(|U_y\Delta\rangle\!\rangle\langle\!\langle U_y\Delta|) \Big)\right. \\
&\qquad\left. + \varepsilon \operatorname{tr}_{\mathcal{H}_{\text{anc}}}\Big(|V_0\rangle\!\rangle\langle\!\langle (U_x-U_y)\Delta|\Big) + \varepsilon \operatorname{tr}_{\mathcal{H}_{\text{anc}}}\Big(|(U_x-U_y)\Delta\rangle\!\rangle\langle\!\langle V_0|\Big)\right\|_1
\end{aligned}$$

$$\ge\varepsilon\left\| \operatorname{tr}_{\mathcal{H}_{\text{anc}}}\Big(|V_0\rangle\!\rangle\langle\!\langle (U_x-U_y)\Delta|\Big) + \operatorname{tr}_{\mathcal{H}_{\text{anc}}}\Big(|(U_x-U_y)\Delta\rangle\!\rangle\langle\!\langle V_0|\Big)\right\|_1 - 2\varepsilon^2 d_1 \tag{116}$$

$$=2\varepsilon\left\| \operatorname{tr}_{\mathcal{H}_{\text{anc}}}\Big(|V_0\rangle\!\rangle\langle\!\langle (U_x-U_y)\Delta|\Big)\right\|_1 - 2\varepsilon^2 d_1 \tag{117}$$

$$\ge\frac{2\kappa^{3/2}}{12800}\varepsilon d_1 - 2\varepsilon^2 d_1 \tag{118}$$

$$\ge\frac{\kappa^{3/2}}{12800}\varepsilon d_1. \tag{119}$$

In Equation (116), we used $\|\operatorname{tr}_{\mathcal{H}_{\text{anc}}}(|U\Delta\rangle\!\rangle\langle\!\langle U\Delta|)\|_1 = \operatorname{tr}(|U\Delta\rangle\!\rangle\langle\!\langle U\Delta|) = r\zeta \le r\underline{\chi} \le d_1$. In Equation (117) we used the fact that $\operatorname{tr}_{\mathcal{H}_{\text{anc}}}\Big(|V_0\rangle\!\rangle\langle\!\langle (U_x-U_y)\Delta|\Big)$ is a linear operator with support in $\mathcal{H}_{\text{A}}[1:r\zeta] \otimes \mathcal{H}_{\text{B}}[\overline{\chi}+1:d_2]$ and image in $\mathcal{H}_{\text{A}} \otimes \mathcal{H}_{\text{B}}[1:\overline{\chi}]$ which is orthogonal to its support; and then we used Fact 7.9. In Equation (118) we used Equation (115). In Equation (119) we used that $\varepsilon \le \kappa^{3/2}/25600$. Therefore, we can lower bound the Choi trace norm

$$\frac{1}{d_1}\left\|C_{\mathcal{E}_x} - C_{\mathcal{E}_y}\right\|_1 \ge \frac{\kappa^{3/2}}{12800}\varepsilon.$$

Thus, the set $\mathcal{N} = \{V_{\varepsilon,U} \,|\, U \in \mathcal{M}\}$ is the desired set, and the cardinality is $\exp(\kappa^3 d_1(rd_2 - d_1)/3840001) = \exp(\kappa^2(rd_2-d_1)\min\{d_1, rd_2-d_1\}/3840001)$. ☐

5.3.3 Proof of Lemma 5.11

Proof. First, we need the following lemma:

Lemma 5.12. *For $U_x, U_y \in \mathbb{U}_{r(d_2-\overline{\chi})}$, let us define*

$$F(U_x, U_y) = \frac{1}{d_1} \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big(|V_0\rangle\!\rangle\Big(\langle\!\langle U_x\Delta| - \langle\!\langle U_y\Delta|\Big)\Big),$$

then the function $f(U_x, U_y) = \|F(U_x, U_y)\|_1 = \mathrm{tr}(|F(U_x, U_y)|)$ is $\sqrt{\frac{2}{d_1}}$-Lipschitz with respect to the ℓ_2-sum of the 2-norms (Frobenius norm). Furthermore, for independent random $U_x, U_y \sim \mathbb{U}_{r(d_2-\overline{\chi})}$, we have $\mathbb{E}\big[\mathrm{tr}(|F(U_x, U_y)|^2)\big] \geq \frac{\kappa}{4r}$, and $\mathbb{E}\big[\mathrm{tr}(|F(U_x, U_y)|^4)\big] \leq \frac{256}{r^3}$.

By the Hölder's inequality we have

$$\mathbb{E}\big[\mathrm{tr}\big(|F(U_x, U_y)|^2\big)\big] \leq \mathbb{E}\big[\mathrm{tr}\big(|F(U_x, U_y)|^4\big)\big]^{1/3}\, \mathbb{E}[\mathrm{tr}(|F(U_x, U_y)|)]^{2/3},$$

which, combined with Lemma 5.12, implies

$$\mathbb{E}[\mathrm{tr}(|F(U_x, U_y)|)]^2 \geq \frac{\kappa^3}{16384}.$$

Thus $\mathbb{E}[\mathrm{tr}(|F(U_x, U_y)|)] > \kappa^{3/2}/128$. Then, since the fucntion $f(U_x, U_y) = \mathrm{tr}(|F(U_x, U_y)|)$ is $\sqrt{\frac{2}{d_1}}$-Lipschitz, we can use Theorem 7.10 to prove the concentration result:

$$\mathbb{P}\left[\mathrm{tr}(|F(U_x, U_y)|) \leq \frac{\kappa^{3/2}}{12800}\right] \leq \exp\left(-\frac{d_1 r(d_2-\overline{\chi})}{2} \cdot \frac{\kappa^3}{40000\cdot 12}\right) \leq \exp\left(-\frac{\kappa^3 d_1(rd_2-d_1)}{1920000}\right),$$

where we used $r(d_2-\overline{\chi}) = r\lfloor\frac{rd_2-d_1}{r}\rfloor \geq \frac{rd_2-d_1}{2}$ (since $rd_2 - d_1 \geq r$). Then, we independently sample $\exp(\kappa^3 d_1(rd_2-d_1)/3840001)$ Haar random unitaries in $\mathbb{U}_{r(d_2-\overline{\chi})}$ and the union bound shows that there exists a non-zero probability that for any pair U_x, U_y, we have $\mathrm{tr}(|F(U_x, U_y)|) \geq \kappa^{3/2}/12800$. Thus, there exists a set with cardinality $\geq \exp(\kappa^3 d_1(rd_2-d_1)/3840001)$ such that Equation (115) holds. □

Then, we give the proof of Lemma 5.12.

Proof of Lemma 5.12. The Lipschitz continuity can be seen from Lemma 7.3.

Define $K_{x,i} = \langle i|_{\mathrm{anc}} U_x\Delta$ and $K_{y,i} = \langle i|_{\mathrm{anc}} U_y\Delta$. Thus $F(U_x, U_y) = \frac{1}{d_1}\sum_{i=1}^r |K_i\rangle\!\rangle\Big(\langle\!\langle K_{x,i}| - \langle\!\langle K_{y,i}|\Big)$. Then, we note that

$$\begin{aligned}
\mathbb{E}\big[\mathrm{tr}(|F(U_x, U_y)|^2)\big] &= \frac{1}{d_1^2}\mathbb{E}\left[\mathrm{tr}\left(\sum_{i,j=1}^r |K_i\rangle\!\rangle\Big(\langle\!\langle K_{x,i}| - \langle\!\langle K_{y,i}|\Big)\Big(|K_{x,j}\rangle\!\rangle - |K_{y,j}\rangle\!\rangle\Big)\langle\!\langle K_j|\right)\right] \\
&= \frac{1}{d_1^2}\mathbb{E}\left[\sum_{i=1}^r \langle\!\langle K_i|K_i\rangle\!\rangle\Big(\langle\!\langle K_{x,i}| - \langle\!\langle K_{y,i}|\Big)\Big(|K_{x,i}\rangle\!\rangle - |K_{y,i}\rangle\!\rangle\Big)\right] \\
&= \frac{2\zeta}{d_1^2}\sum_{i=1}^r \langle\!\langle K_i|K_i\rangle\!\rangle && (120) \\
&\geq \frac{\zeta}{d_1} \geq \frac{\kappa}{4r}, && (121)
\end{aligned}$$

where Equation (120) is because for $t_1, t_2 \in \{x, y\}$, we have

$$\begin{aligned}\mathbb{E}[\langle\langle K_{t_1,i}|K_{t_2,i}\rangle\rangle] = \mathbb{E}\Big[\operatorname{tr}\Big(K_{t_1,i}^\dagger K_{t_2,i}\Big)\Big] &= \operatorname{tr}\Big(\Delta^\dagger\,\mathbb{E}\Big[U_{t_1}^\dagger |i\rangle\langle i|_{\text{anc}} U_{t_2}\Big]\,\Delta\Big)\\ &= \mathbb{1}_{t_1=t_2}\cdot\frac{1}{r}\operatorname{tr}(\Delta^\dagger\Delta) \qquad (122)\\ &= \mathbb{1}_{t_1=t_2}\cdot\zeta,\end{aligned}$$

where Equation (122) is due to Schur's lemma, Equation (121) is because $\sum_{i=1}^r \operatorname{tr}(K_i^\dagger K_i) = \operatorname{tr}(V_0^\dagger V_0) \ge \frac{1}{2}\operatorname{tr}(I_{\mathcal{H}_\text{A}}) = d_1/2$.

Then, we also have

$$\begin{aligned}\mathbb{E}\big[\operatorname{tr}(|F(U_x,U_y)|^4)\big] &= \frac{1}{d_1^4}\,\mathbb{E}\Bigg[\sum_{i,j,k,l=1}^r \operatorname{tr}\Big(|K_i\rangle\rangle\Big(\langle\langle K_{x,i}| - \langle\langle K_{y,i}|\Big)\Big(|K_{x,j}\rangle\rangle - |K_{y,j}\rangle\rangle\Big)\langle\langle K_j| \\ &\qquad |K_k\rangle\rangle\Big(\langle\langle K_{x,k}| - \langle\langle K_{y,k}|\Big)\Big(|K_{x,l}\rangle\rangle - |K_{y,l}\rangle\rangle\Big)\langle\langle K_l|\Big)\Bigg]\\ &\le \frac{\overline{\chi}^2}{d_1^4}\sum_{i,j=1}^r \mathbb{E}\Big[\Big|\Big(\langle\langle K_{x,i}| - \langle\langle K_{y,i}|\Big)\Big(|K_{x,j}\rangle\rangle - |K_{y,j}\rangle\rangle\Big)\Big|^2\Big] \qquad (123)\\ &\le \frac{4\overline{\chi}^2}{d_1^4}\sum_{i,j=1}^r \mathbb{E}\big[|\langle\langle K_{x,i}|K_{x,j}\rangle\rangle|^2 + |\langle\langle K_{y,i}|K_{y,j}\rangle\rangle|^2 + |\langle\langle K_{x,i}|K_{y,j}\rangle\rangle|^2 + |\langle\langle K_{y,i}|K_{x,j}\rangle\rangle|^2\big]\\ &\le \frac{8\overline{\chi}^2}{d_1^4}\sum_{i,j=1}^r \mathbb{E}\big[|\langle\langle K_{x,i}|K_{x,j}\rangle\rangle|^2 + |\langle\langle K_{y,i}|K_{y,j}\rangle\rangle|^2\big] \qquad (124)\\ &= \frac{16\overline{\chi}^2}{d_1^4}\sum_{i,j=1}^r \mathbb{E}\big[|\langle\langle K_{x,i}|K_{x,j}\rangle\rangle|^2\big] = \frac{16\overline{\chi}^2}{d_1^4}\sum_{i,j=1}^r \mathbb{E}\Big[\Big|\operatorname{tr}(K_{x,i}^\dagger K_{x,j})\Big|^2\Big]\\ &\le \frac{16\overline{\chi}^2}{d_1^4}\cdot\frac{4r^2\zeta^2}{r} = \frac{64\cdot r\overline{\chi}^2\zeta^2}{d_1^4} \le \frac{64\cdot r\overline{\chi}^2\underline{\chi}^2}{d_1^4} \le \frac{256}{r^3}, \qquad (125)\end{aligned}$$

where Equation (123) is due to Equation (113), Equation (124) due to the similar argument as that in Equation (98), and in Equation (125), the first inequality is due to Lemma 7.4 and that Δ is an isometry from $\mathcal{H}_\text{A}[1:r\zeta]$ to $\mathcal{H}_\text{B}[\overline{\chi}+1:d_2]\otimes\mathcal{H}_\text{anc}$ (further note that $\dim(\mathcal{H}_\text{anc}) \le \dim(\mathcal{H}_\text{A}[1:r\zeta])\dim(\mathcal{H}_\text{B}[\overline{\chi}+1:d_2])$ so we choose $k=1$ in Lemma 7.4). □

5.3.4 Existence for diamond norm

Note that in Theorem 5.10, the cardinality and distance both depend on κ, which can be arbitrarily close to 0. Here, we give another packing nets w.r.t. diamond norm that does not depend on κ. We will use the same construction as that given in Section 5.3.1. Specifically, recall that $\underline{\chi} \coloneqq \lfloor\frac{d_1}{r}\rfloor$, $\overline{\chi} \coloneqq \lceil\frac{d_1}{r}\rceil$, and $\zeta \coloneqq \min\{\underline{\chi}, d_2-\overline{\chi}\}$.

Theorem 5.13. *Suppose $\varepsilon \le 1/160$. There exists a finite subset $\mathcal{N}$ of $\{V_{\varepsilon,U}\,|\,U \in \mathbb{U}_{r(d_2-\overline{\chi})}\}$ for $V_{\varepsilon,U}$ defined in Equation (114) with cardinality*

$$|\mathcal{N}| \ge \exp\!\left(\frac{(rd_2-d_1)\min\{d_1, rd_2-d_1\}}{38404}\right),$$

such that for any $V_x \neq V_y \in \mathcal{N}$*, if we set* $\mathcal{E}_x = \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(V_x(\cdot)V_x^\dagger)$ *and* $\mathcal{E}_y = \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(V_y(\cdot)V_y^\dagger)$*, then*

$$\|\mathcal{E}_x - \mathcal{E}_y\|_\diamond \geq \frac{1}{80}\varepsilon.$$

Proof. We define $|\Psi\rangle \coloneqq \frac{1}{\sqrt{r\zeta}}\sum_{i=1}^{r\zeta}|i\rangle_{\mathrm{A}}|i\rangle_{\mathrm{A}}$ be an entangled state on $\mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{A}}$ and $\Psi = |\Psi\rangle\langle\Psi|$. We also define $W_0 : \mathcal{H}_{\mathrm{A}}[1 : r\zeta] \to \mathcal{H}_{\mathrm{B}}[1 : \zeta] \otimes \mathcal{H}_{\mathrm{anc}}$ as:

$$W_0 \coloneqq \sqrt{1-\varepsilon^2}\sum_{i=1}^{r\zeta}|g(i)_1\rangle_{\mathrm{B}} \otimes |g(i)_2\rangle_{\mathrm{anc}}\langle i|_{\mathrm{A}}.$$

We can easily see that (c.f. Equation (111)):

$$V_0|\Psi\rangle = W_0|\Psi\rangle = \frac{1}{\sqrt{r\zeta}}|W_0\rangle\!\rangle. \tag{126}$$

Then, we need the following lemma.

Lemma 5.14. *There exists a finite subset* $\mathcal{M} \subseteq \mathbb{U}_{r(d_2-\overline{\chi})}$ *with cardinality* $|\mathcal{M}| \geq \exp(r\zeta(rd_2 - d_1)/9601)$ *such that for any* $U_x \neq U_y \in \mathcal{M}$*,*

$$\frac{1}{r\zeta}\left\|\mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big(|W_0\rangle\!\rangle\Big(\langle\!\langle U_x\Delta| - \langle\!\langle U_y\Delta|\Big)\Big)\right\|_1 \geq \frac{1}{80}. \tag{127}$$

The proof of Lemma 5.14 is deferred to Section 5.3.5.

Now, we are able to prove Theorem 5.7. Let $\mathcal{M}$ be the set given in Lemma 5.8. For any $U_x \neq U_y \in \mathcal{M}$, if we set $\mathcal{E}_x = \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(V_{\varepsilon,U_x}(\cdot)V_{\varepsilon,U_x}^\dagger)$ and $\mathcal{E}_y = \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(V_{\varepsilon,U_y}(\cdot)V_{\varepsilon,U_y}^\dagger)$, then we have

$$
\begin{aligned}
&\|\mathcal{E}_x - \mathcal{E}_y\|_\diamond \geq \|\mathcal{E}_x(\Psi) - \mathcal{E}_y(\Psi)\|_1 \\
=&\left\| \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big((W_0 + \varepsilon U_x\Delta)\Psi(W_0 + \varepsilon U_x\Delta)^\dagger\Big) - \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big((W_0 + \varepsilon U_y\Delta)\Psi(W_0 + \varepsilon U_y\Delta)^\dagger\Big) \right\|_1 &(128)\\
=&\frac{1}{r\zeta}\bigg\|\varepsilon^2\Big(\mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(|U_x\Delta\rangle\!\rangle\langle\!\langle U_x\Delta|) - \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(|U_y\Delta\rangle\!\rangle\langle\!\langle U_y\Delta|) \Big) \\
&\qquad + \varepsilon\, \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big(|W_0\rangle\!\rangle\langle\!\langle (U_x - U_y)\Delta|\Big) + \varepsilon\, \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big(|(U_x - U_y)\Delta\rangle\!\rangle\langle\!\langle W_0|\Big)\bigg\|_1 &(129)\\
\geq&\frac{\varepsilon}{r\zeta}\left\| \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big(|W_0\rangle\!\rangle\langle\!\langle (U_x - U_y)\Delta|\Big) + \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big(|(U_x - U_y)\Delta\rangle\!\rangle\langle\!\langle W_0|\Big)\right\|_1 - 2\varepsilon^2 &(130)\\
=&\frac{2\varepsilon}{r\zeta}\left\| \mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big(|W_0\rangle\!\rangle\langle\!\langle (U_x - U_y)\Delta|\Big)\right\|_1 - 2\varepsilon^2 &(131)\\
\geq&\frac{1}{40}\varepsilon - 2\varepsilon^2 &(132)\\
\geq&\frac{1}{80}\varepsilon. &(133)
\end{aligned}
$$

In Equation (128), we used Equation (126). In Equation (129), we used $U\Delta|\Psi\rangle = \frac{1}{\sqrt{r\zeta}}|U\Delta\rangle\!\rangle$, $W_0|\Psi\rangle = \frac{1}{\sqrt{r\zeta}}|W_0\rangle\!\rangle$. In Equation (130), we used $\|\mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}(|U\Delta\rangle\!\rangle\langle\!\langle U\Delta|)\|_1 = \mathrm{tr}(|U\Delta\rangle\!\rangle\langle\!\langle U\Delta|) = r\zeta$. In Equation (131) we used the fact that $\mathrm{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big(|W_0\rangle\!\rangle\langle\!\langle (U_x - U_y)\Delta|\Big)$ is a linear operator with support in $\mathcal{H}_{\mathrm{A}}[1 : r\zeta] \otimes \mathcal{H}_{\mathrm{B}}[\overline{\chi} + 1 : d_2]$ and image in $\mathcal{H}_{\mathrm{A}}[1 : r\zeta] \otimes \mathcal{H}_{\mathrm{B}}[1 : \zeta]$, which is orthogonal to its support; and then we used Fact 7.9. In Equation (132) we used Equation (127). In Equation (133) we used that $\varepsilon \leq 1/160$. Therefore, we can lower bound the diamond norm

$$\|\mathcal{E}_x - \mathcal{E}_y\|_\diamond \geq \frac{1}{80}\varepsilon.$$

Thus, the set $\mathcal{N} = \{V_{\varepsilon,U} \,|\, U \in \mathcal{M}\}$ is the desired set, and the cardinality is at least $\exp\left(\frac{r\zeta(rd_2-d_1)}{9601}\right) \geq \exp\left(\frac{(rd_2-d_1)\min\{d_1,rd_2-d_1\}}{38404}\right)$, since $r\zeta \geq \frac{\kappa}{4}d_1 = \frac{1}{4}\min\{d_1, rd_2-d_1\}$. □

5.3.5 Proof of Lemma 5.14

Proof. We need the following lemma:

Lemma 5.15. *For $U_x, U_y \in \mathbb{U}_{r(d_2-\overline{\chi})}$, let us define*

$$F(U_x,U_y) = \frac{1}{r\zeta}\operatorname{tr}_{\mathcal{H}_{\text{anc}}}\Big(|W_0\rangle\!\rangle\Big(\langle\!\langle U_x\Delta| - \langle\!\langle U_y\Delta|\Big)\Big),$$

then the function $f(U_x,U_y) = \|F(U_x,U_y)\|_1 = \operatorname{tr}(|F(U_x,U_y)|)$ is $\sqrt{\frac{2}{r\zeta}}$-Lipschitz with respect to the ℓ_2-sum of the 2-norms (Frobenius norm). Furthermore, for independent random $U_x, U_y \sim \mathbb{U}_{r(d_2-\overline{\chi})}$, we have $\mathbb{E}\big[\operatorname{tr}(|F(U_x,U_y)|^2)\big] \geq \frac{1}{r}$, and $\mathbb{E}\big[\operatorname{tr}(|F(U_x,U_y)|^4)\big] \leq \frac{64}{r^3}$.

By the Hölder's inequality we have

$$\mathbb{E}\big[\operatorname{tr}\big(|F(U_x,U_y)|^2\big)\big] \leq \mathbb{E}\big[\operatorname{tr}\big(|F(U_x,U_y)|^4\big)\big]^{1/3}\,\mathbb{E}[\operatorname{tr}(|F(U_x,U_y)|)]^{2/3},$$

which, combined with Lemma 5.15, implies

$$\mathbb{E}[\operatorname{tr}(|F(U_x,U_y)|)]^2 \geq \frac{1}{64}.$$

Thus $\mathbb{E}[\operatorname{tr}(|F(U_x,U_y)|)] > 1/8$. Then, since the function $f(U_x,U_y) = \operatorname{tr}(|F(U_x,U_y)|)$ is $\sqrt{\frac{2}{r\zeta}}$-Lipschitz, we can use Theorem 7.10 to prove the concentration result:

$$\mathbb{P}\left[\operatorname{tr}(|F(U_x,U_y)|) \leq \frac{1}{80}\right] \leq \exp\left(-\frac{r\zeta\cdot r(d_2-\overline{\chi})}{2}\cdot\frac{1}{100\cdot 12}\right) \leq \exp\left(-\frac{r\zeta(rd_2-d_1)}{4800}\right),$$

where in the last inequality we used $r(d_2-\overline{\chi}) = r\lfloor\frac{rd_2-d_1}{r}\rfloor \geq \frac{rd_2-d_1}{2}$ (since $rd_2 - d_1 \geq r$). Then, we independently sample $\exp(r\zeta(rd_2-d_1)/9601)$ Haar random unitaries in $\mathbb{U}_{r(d_2-\overline{\chi})}$ and the union bound shows that there exists a non-zero probability that for any pair U_x, U_y, we have $\operatorname{tr}(|F(U_x,U_y)|) \geq 1/80$. Thus, there exists a set with cardinality $\geq \exp(r\zeta(rd_2-d_1)/9601)$ such that Equation (127) holds. □

Then, we give the proof of Lemma 5.15.

Proof of Lemma 5.15. The Lipschitz continuity can be seen from Lemma 7.3, where we treat W_0 and $U\Delta$ as linear operators acting on the input space with dimension $r\zeta$.

Define $K_{x,i} = \langle i|_{\text{anc}}U_x\Delta$, $K_{y,i} = \langle i|_{\text{anc}}U_y\Delta$, and $K'_i = \langle i|_{\text{anc}}W_0$. We can easily see that

$$\frac{1}{2}\zeta\cdot\mathbb{1}_{i=j} \leq \left|\operatorname{tr}(K_i'^{\dagger}K'_j)\right| \leq \zeta\cdot\mathbb{1}_{i=j}, \tag{134}$$

and also

$$F(U_x,U_y)=\frac{1}{r\zeta}\sum_{i=1}^{r}|K_i'\rangle\!\rangle\Big(\langle\!\langle K_{x,i}|-\langle\!\langle K_{y,i}|\Big).$$

Then, we note that

$$\begin{aligned}
\mathbb{E}\big[\operatorname{tr}(|F(U_x,U_y)|^2)\big]&=\frac{1}{r^2\zeta^2}\,\mathbb{E}\left[\operatorname{tr}\left(\sum_{i,j=1}^{r}|K_i'\rangle\!\rangle\Big(\langle\!\langle K_{x,i}|-\langle\!\langle K_{y,i}|\Big)\Big(|K_{x,j}\rangle\!\rangle-|K_{y,j}\rangle\!\rangle\Big)\langle\!\langle K_j'|\right)\right]\\
&\geq\frac{1}{2r^2\zeta}\,\mathbb{E}\left[\sum_{i=1}^{r}\Big(\langle\!\langle K_{x,i}|-\langle\!\langle K_{y,i}|\Big)\Big(|K_{x,i}\rangle\!\rangle-|K_{y,i}\rangle\!\rangle\Big)\right] \qquad (135)\\
&=\frac{1}{2r^2\zeta}\cdot 2r\zeta=\frac{1}{r} \qquad (136)
\end{aligned}$$

where Equation (135) is due to Equation (134), and Equation (136) is because for $t_1,t_2\in\{x,y\}$, we have

$$\begin{aligned}
\mathbb{E}[\langle\!\langle K_{t_1,i}|K_{t_2,i}\rangle\!\rangle]=\mathbb{E}\Big[\operatorname{tr}\Big(K_{t_1,i}^{\dagger}K_{t_2,i}\Big)\Big]&=\operatorname{tr}\Big(\Delta^{\dagger}\,\mathbb{E}\Big[U_{t_1}^{\dagger}|i\rangle\langle i|_{\mathrm{anc}}U_{t_2}\Big]\,\Delta\Big)\\
&=\mathbb{1}_{t_1=t_2}\cdot\frac{1}{r}\operatorname{tr}(\Delta^{\dagger}\Delta) \qquad (137)\\
&=\mathbb{1}_{t_1=t_2}\zeta,
\end{aligned}$$

where Equation (137) is due to Schur's lemma..

Then, we also have

$$\begin{aligned}
\mathbb{E}\big[\operatorname{tr}(|F(U_x,U_y)|^4)\big]&=\frac{1}{r^4\zeta^4}\,\mathbb{E}\Bigg[\sum_{i,j,k,l=1}^{r}\operatorname{tr}\Big(|K_i'\rangle\!\rangle\Big(\langle\!\langle K_{x,i}|-\langle\!\langle K_{y,i}|\Big)\Big(|K_{x,j}\rangle\!\rangle-|K_{y,j}\rangle\!\rangle\Big)\langle\!\langle K_j'|\\
&\qquad\qquad |K_k'\rangle\!\rangle\Big(\langle\!\langle K_{x,k}|-\langle\!\langle K_{y,k}|\Big)\Big(|K_{x,l}\rangle\!\rangle-|K_{y,l}\rangle\!\rangle\Big)\langle\!\langle K_l'|\Big)\Bigg]\\
&\leq\frac{1}{r^4\zeta^2}\sum_{i,j=1}^{r}\mathbb{E}\left[\left|\Big(\langle\!\langle K_{x,i}|-\langle\!\langle K_{y,i}|\Big)\Big(|K_{x,j}\rangle\!\rangle-|K_{y,j}\rangle\!\rangle\Big)\right|^2\right] \qquad (138)\\
&\leq\frac{4}{r^4\zeta^2}\sum_{i,j=1}^{r}\mathbb{E}\big[|\langle\!\langle K_{x,i}|K_{x,j}\rangle\!\rangle|^2+|\langle\!\langle K_{y,i}|K_{y,j}\rangle\!\rangle|^2+|\langle\!\langle K_{x,i}|K_{y,j}\rangle\!\rangle|^2+|\langle\!\langle K_{y,i}|K_{x,j}\rangle\!\rangle|^2\big]\\
&\leq\frac{8}{r^4\zeta^2}\sum_{i,j=1}^{r}\mathbb{E}\big[|\langle\!\langle K_{x,i}|K_{x,j}\rangle\!\rangle|^2+|\langle\!\langle K_{y,i}|K_{y,j}\rangle\!\rangle|^2\big] \qquad (139)\\
&=\frac{16}{r^4\zeta^2}\sum_{i,j=1}^{r}\mathbb{E}\big[|\langle\!\langle K_{x,i}|K_{x,j}\rangle\!\rangle|^2\big]=\frac{16}{r^4\zeta^2}\sum_{i,j=1}^{r}\mathbb{E}\left[\left|\operatorname{tr}(K_{x,i}^{\dagger}K_{x,j})\right|^2\right],\\
&\leq\frac{16}{r^4\zeta^2}\cdot\frac{4(r\zeta)^2}{r}=\frac{64}{r^3}, \qquad (140)
\end{aligned}$$

where Equation (138) is due to Equation (134), Equation (139) is due to a similar argument as that in Equation (98), and in Equation (140), the first inequality is due to Lemma 7.4 and that Δ is an isometry from $\mathcal{H}_{\mathrm{A}}[1:r\zeta]$ to $\mathcal{H}_{\mathrm{B}}[\overline{\chi}+1:d_2]\otimes\mathcal{H}_{\mathrm{anc}}$ (further note that $\dim(\mathcal{H}_{\mathrm{anc}})\leq\dim(\mathcal{H}_{\mathrm{A}}[1:r\zeta])\dim(\mathcal{H}_{\mathrm{B}}[\overline{\chi}+1:d_2])$ so we choose $k=1$ in Lemma 7.4). □

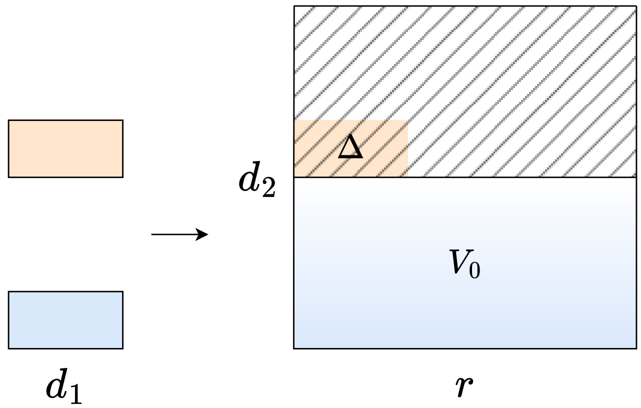


Figure 7: Illustration of our construction. We define linear operators V_0 and Δ from the d_1-dimensional space $\mathcal{H}_{\mathrm{A}}$ to the $d_2 r$-dimensional space $\mathcal{H}_{\mathrm{B}} \otimes \mathcal{H}_{\mathrm{anc}}$. The Haar randomness is applied on the hatched area. The gradient-colored area means the image of V_0 is approximately "uniformly" distributed along the r-axis, using Lemma 7.5.

5.4 Type II instance: $d_1 < rd_2$ with $d_1 + r \leq rd_2$, $d_1 < r$

5.4.1 Construction

In this subsection, we will use the definition in Section 4.2.1, where the parameter d, D in Section 4.2.1 correspond to d_1 and rd_2 here.

Suppose $d_2 > 1$. We further assume $d_1 + r \leq rd_2$ and $d_1 < r$. Let $\chi = \lceil \frac{r}{d_1} \rceil$. Thus $\chi \geq 1$ and $\chi d_1 \geq r$. Moreover, since $d_1 d_2 \geq 2r$ (see Equation (55)), we have $\chi \leq \lceil \frac{d_2}{2} \rceil$ and thus $d_2 - \chi \geq \lfloor \frac{d_2}{2} \rfloor$.

Note that $\chi d_1 \geq r$. Thus by Lemma 7.5, there exists a set of linear operators $\{K_i\}_{i=1}^r$ for $K_i : \mathcal{H}_{\mathrm{A}} \to \mathcal{H}_{\mathrm{B}}[1 : \chi]$ such that

$$\sum_{i=1}^{r} K_i^\dagger K_i = (1 - \varepsilon^2) I_{\mathcal{H}_{\mathrm{A}}} \quad \text{and} \quad \left| \mathrm{tr}\left(K_i^\dagger K_j \right) \right| \leq \frac{2 \dim(\mathcal{H}_{\mathrm{A}})}{r} \cdot \mathbb{1}_{i=j} = \frac{2d_1}{r} \cdot \mathbb{1}_{i=j}. \tag{141}$$

Then we define $V_0 : \mathcal{H}_{\mathrm{A}} \to \mathcal{H}_{\mathrm{B}}[1 : \chi] \otimes \mathcal{H}_{\mathrm{anc}}$ as

$$V_0 = \sum_{i=1}^{r} |i\rangle_{\mathrm{anc}} \otimes K_i.$$

Note that V_0 can also be written as $V_0 = \sqrt{1 - \varepsilon^2} V$ for some isometry $V : \mathcal{H}_{\mathrm{A}} \to \mathcal{H}_{\mathrm{B}}[1 : \chi] \otimes \mathcal{H}_{\mathrm{anc}}$. Define $\Delta : \mathcal{H}_{\mathrm{A}} \to \mathcal{H}_{\mathrm{B}}[\chi + 1 : d_2] \otimes \mathcal{H}_{\mathrm{anc}}$ be an arbitrary (but fixed) isometry. Then, for $\varepsilon \in (0, 1/2)$ and $U \in \mathbb{U}_{r(d_2 - \chi)}$, we define the isometry $V_{\varepsilon,U} : \mathcal{H}_{\mathrm{A}} \to \mathcal{H}_{\mathrm{B}} \otimes \mathcal{H}_{\mathrm{anc}}$ as

$$V_{\varepsilon,U} \coloneqq V_0 + \varepsilon U \Delta, \tag{142}$$

where U acts on $\mathcal{H}_{\mathrm{B}}[\chi + 1 : d_2] \otimes \mathcal{H}_{\mathrm{anc}}$. We can easily verify that $V_{\varepsilon,U}$ is indeed an isometry. We also note that the image of $U\Delta$ is orthogonal to the image of V_0. For clarity, we illustrate our construction in Figure 7.

5.4.2 Existence

Then, we prove that there exists a large set of isometries $V_{\varepsilon,U}$ with good separation property.

Theorem 5.16. *Suppose $\varepsilon \le 1/4000$. There exists a finite subset $\mathcal{N}$ of $\{V_{\varepsilon,U} \,|\, U \in \mathbb{U}_{r(d_2-\chi)}\}$ for $V_{\varepsilon,U}$ defined in Equation (142) with cardinality $|\mathcal{N}| \ge \exp(rd_1d_2/307201)$, such that for any $V_x \ne V_y \in \mathcal{N}$, if we set $\mathcal{E}_x = \operatorname{tr}_{\mathcal{H}_{\text{anc}}}(V_x(\cdot)V_x^\dagger)$ and $\mathcal{E}_y = \operatorname{tr}_{\mathcal{H}_{\text{anc}}}(V_y(\cdot)V_y^\dagger)$, then*

$$\frac{1}{d_1}\|C_{\mathcal{E}_x} - C_{\mathcal{E}_y}\|_1 \ge \frac{1}{2000}\varepsilon.$$

Proof. First, we need the following lemma.

Lemma 5.17. *There exists a finite subset $\mathcal{M} \subseteq \mathbb{U}_{r(d_2-\chi)}$ with cardinality $|\mathcal{M}| \ge \exp(rd_1d_2/307201)$ such that for any $U_x \ne U_y \in \mathcal{M}$,*

$$\frac{1}{d_1}\left\|\operatorname{tr}_{\mathcal{H}_{\text{anc}}}\Big(|V_0\rangle\!\rangle\Big(\langle\!\langle U_x\Delta| - \langle\!\langle U_y\Delta|\Big)\Big)\right\|_1 \ge \frac{1}{2000}. \tag{143}$$

The proof of Lemma 5.17 is deferred to Section 5.4.3.

Now we are able to prove Theorem 5.16. Let $\mathcal{M}$ be the set given in Lemma 5.17. Then, for any $U_x \ne U_y \in \mathcal{M}$, if set $\mathcal{E}_x = \operatorname{tr}_{\mathcal{H}_{\text{anc}}}(V_{\varepsilon,U_x}(\cdot)V_{\varepsilon,U_x}^\dagger)$ and $\mathcal{E}_y = \operatorname{tr}_{\mathcal{H}_{\text{anc}}}(V_{\varepsilon,U_y}(\cdot)V_{\varepsilon,U_y}^\dagger)$, we have

$$\begin{aligned}
&\|C_{\mathcal{E}_x} - C_{\mathcal{E}_y}\|_1 \\
=&\left\|\operatorname{tr}_{\mathcal{H}_{\text{anc}}}(|V_{\varepsilon,U_x}\rangle\!\rangle\langle\!\langle V_{\varepsilon,U_x}|) - \operatorname{tr}_{\mathcal{H}_{\text{anc}}}(|V_{\varepsilon,U_y}\rangle\!\rangle\langle\!\langle V_{\varepsilon,U_y}|)\right\|_1 \\
=&\bigg\|\varepsilon^2\Big(\operatorname{tr}_{\mathcal{H}_{\text{anc}}}(|U_x\Delta\rangle\!\rangle\langle\!\langle U_x\Delta|) - \operatorname{tr}_{\mathcal{H}_{\text{anc}}}(|U_y\Delta\rangle\!\rangle\langle\!\langle U_y\Delta|)\Big) \\
&\qquad + \varepsilon\operatorname{tr}_{\mathcal{H}_{\text{anc}}}\Big(|V_0\rangle\!\rangle\langle\!\langle (U_x - U_y)\Delta|\Big) + \varepsilon\operatorname{tr}_{\mathcal{H}_{\text{anc}}}\Big(|(U_x - U_y)\Delta\rangle\!\rangle\langle\!\langle V_0|\Big)\bigg\|_1 \\
\ge&\varepsilon\left\|\operatorname{tr}_{\mathcal{H}_{\text{anc}}}\Big(|V_0\rangle\!\rangle\langle\!\langle (U_x - U_y)\Delta|\Big) + \operatorname{tr}_{\mathcal{H}_{\text{anc}}}\Big(|(U_x - U_y)\Delta\rangle\!\rangle\langle\!\langle V_0|\Big)\right\|_1 - 2\varepsilon^2 d_1 &(144)\\
=&2\varepsilon\left\|\operatorname{tr}_{\mathcal{H}_{\text{anc}}}\Big(|V_0\rangle\!\rangle\langle\!\langle (U_x - U_y)\Delta|\Big)\right\|_1 - 2\varepsilon^2 d_1 &(145)\\
\ge&\frac{1}{1000}\varepsilon d_1 - 2\varepsilon^2 d_1 &(146)\\
\ge&\frac{1}{2000}\varepsilon d_1. &(147)
\end{aligned}$$

In Equation (144), we used $\|\operatorname{tr}_{\mathcal{H}_{\text{anc}}}(|U\Delta\rangle\!\rangle\langle\!\langle U\Delta|)\|_1 = \operatorname{tr}(|U\Delta\rangle\!\rangle\langle\!\langle U\Delta|) = d_1$. In Equation (145) we used the fact that $\operatorname{tr}_{\mathcal{H}_{\text{anc}}}\Big(|V_0\rangle\!\rangle\langle\!\langle (U_x - U_y)\Delta|\Big)$ is a linear operator with support in $\mathcal{H}_{\text{A}} \otimes \mathcal{H}_{\text{B}}[\chi+1 : d_2]$ and image in $\mathcal{H}_{\text{A}} \otimes \mathcal{H}_{\text{B}}[1 : \chi]$ which is orthogonal to its support; and then we used Fact 7.9. In Equation (146) we used Equation (143). In Equation (147) we used that $\varepsilon \le 1/4000$. Therefore, we can lower bound the Choi trace norm

$$\frac{1}{d_1}\left\|C_{\mathcal{E}_x} - C_{\mathcal{E}_y}\right\|_1 \ge \frac{1}{2000}\varepsilon.$$

Thus, the set $\mathcal{N} = \{V_{\varepsilon,U} \,|\, U \in \mathcal{M}\}$ is the desired set. □

5.4.3 Proof of Lemma 5.17

Proof. First, we need the following lemma:

Lemma 5.18. *For $U_x, U_y \in \mathbb{U}_{r(d_2-\chi)}$, let us define*

$$F(U_x, U_y) = \frac{1}{d_1} \operatorname{tr}_{\mathcal{H}_{\mathrm{anc}}}\Big(|V\rangle\!\rangle\Big(\langle\!\langle U_x\Delta| - \langle\!\langle U_y\Delta|\Big)\Big),$$

then the function $f(U_x, U_y) = \|F(U_x, U_y)\|_1 = \operatorname{tr}(|F(U_x, U_y)|)$ is $\sqrt{\frac{2}{d_1}}$-Lipschitz with respect to the ℓ_2-sum of the 2-norms (Frobenius norm). Furthermore, for independent random $U_x, U_y \sim \mathbb{U}_{r(d_2-\chi)}$, we have $\mathbb{E}\big[\operatorname{tr}(|F(U_x, U_y)|^2)\big] \geq \frac{1}{r}$, and $\mathbb{E}\big[\operatorname{tr}(|F(U_x, U_y)|^4)\big] \leq \frac{384}{r^3}$.

By the Hölder's inequality we have

$$\mathbb{E}\big[\operatorname{tr}\big(|F(U_x, U_y)|^2\big)\big] \leq \mathbb{E}\big[\operatorname{tr}\big(|F(U_x, U_y)|^4\big)\big]^{1/3}\,\mathbb{E}[\operatorname{tr}(|F(U_x, U_y)|)]^{2/3},$$

which, combined with Lemma 5.18, implies

$$\mathbb{E}[\operatorname{tr}(|F(U_x, U_y)|)]^2 \geq \frac{1}{384}.$$

Thus $\mathbb{E}[\operatorname{tr}(|F(U_x, U_y)|)] > 1/20$. Then, since the function $f(U_x, U_y) = \operatorname{tr}(|F(U_x, U_y)|)$ is $\sqrt{\frac{2}{d_1}}$-Lipschitz, we can use Theorem 7.10 to prove the concentration result:

$$\mathbb{P}\left[\operatorname{tr}(|F(U_x, U_y)|) \leq \frac{1}{2000}\right] \leq \exp\left(-\frac{d_1 r(d_2-\chi)}{2}\cdot\frac{1}{1600\cdot 12}\right) \leq \exp\left(-\frac{r d_1 d_2}{153600}\right),$$

where we used $d_2 - \chi \geq \lfloor d_2/2 \rfloor \geq d_2/4$. Then, we independently sample $\exp(r d_1 d_2/307201)$ Haar random unitaries in $\mathbb{U}_{r(d_2-\chi)}$ and the union bound shows that there exists a non-zero probability that for any pair U_x, U_y, we have $\operatorname{tr}(|F(U_x, U_y)|) \geq 1/2000$. Thus, there exists a set with cardinality $\geq \exp(r d_1 d_2/307201)$ such that Equation (143) holds. □

Then, we give the proof of Lemma 5.18.

Proof of Lemma 5.18. The Lipschitz continuity can be seen from Lemma 7.3.

Define $K_{x,i} = \langle i|_{\mathrm{anc}} U_x \Delta$ and $K_{y,i} = \langle i|_{\mathrm{anc}} U_y \Delta$. Thus $F(U_x, U_y) = \frac{1}{d_1}\sum_{i=1}^r |K_i\rangle\!\rangle\Big(\langle\!\langle K_{x,i}| - \langle\!\langle K_{y,i}|\Big)$. Then, we note that

$$\begin{aligned}
\mathbb{E}\big[\operatorname{tr}(|F(U_x, U_y)|^2)\big] &= \frac{1}{d_1^2}\,\mathbb{E}\left[\operatorname{tr}\left(\sum_{i,j=1}^r |K_i\rangle\!\rangle\Big(\langle\!\langle K_{x,i}| - \langle\!\langle K_{y,i}|\Big)\Big(|K_{x,j}\rangle\!\rangle - |K_{y,j}\rangle\!\rangle\Big)\langle\!\langle K_j|\right)\right] \\
&= \frac{1}{d_1^2}\,\mathbb{E}\left[\sum_{i=1}^r \langle\!\langle K_i|K_i\rangle\!\rangle\Big(\langle\!\langle K_{x,i}| - \langle\!\langle K_{y,i}|\Big)\Big(|K_{x,i}\rangle\!\rangle - |K_{y,i}\rangle\!\rangle\Big)\right] \\
&= \frac{2}{r d_1}\sum_{i=1}^r \langle\!\langle K_i|K_i\rangle\!\rangle && (148) \\
&\geq \frac{1}{r}, && (149)
\end{aligned}$$

where Equation (148) is because for $z_1, z_2 \in \{x, y\}$, we have

$$\begin{aligned}\mathbb{E}[\langle\langle K_{z_1,i}|K_{z_2,i}\rangle\rangle] = \mathbb{E}\Big[\operatorname{tr}\Big(K_{z_1,i}^\dagger K_{z_2,i}\Big)\Big] &= \operatorname{tr}\Big(\Delta^\dagger\,\mathbb{E}\Big[U_{z_1}^\dagger |i\rangle\langle i|_{\mathrm{anc}} U_{z_2}\Big]\,\Delta\Big) \\ &= \mathbb{1}_{z_1=z_2}\cdot\frac{1}{r}\operatorname{tr}(\Delta^\dagger\Delta) \qquad (150) \\ &= \mathbb{1}_{z_1=z_2}\cdot\frac{d_1}{r},\end{aligned}$$

where Equation (150) is due to Schur's lemma, Equation (149) is because $\sum_{i=1}^r \operatorname{tr}(K_i^\dagger K_i) = \operatorname{tr}(V_0^\dagger V_0) \geq \frac{1}{2}\operatorname{tr}(I_{\mathcal{H}_{\mathrm{A}}}) = d_1/2$.

Then, we also have

$$\begin{aligned}\mathbb{E}\big[\operatorname{tr}(|F(U_x,U_y)|^4)\big] &= \frac{1}{d_1^4}\,\mathbb{E}\Bigg[\sum_{i,j,k,l=1}^r \operatorname{tr}\Big(|K_i\rangle\rangle\Big(\langle\langle K_{x,i}| - \langle\langle K_{y,i}|\Big)\Big(|K_{x,j}\rangle\rangle - |K_{y,j}\rangle\rangle\Big)\langle\langle K_j| \\ &\qquad |K_k\rangle\rangle\Big(\langle\langle K_{x,k}| - \langle\langle K_{y,k}|\Big)\Big(|K_{x,l}\rangle\rangle - |K_{y,l}\rangle\rangle\Big)\langle\langle K_l|\Big)\Bigg] \\ &\leq \frac{4}{r^2 d_1^2}\sum_{i,j=1}^r \mathbb{E}\Big[\Big|\Big(\langle\langle K_{x,i}| - \langle\langle K_{y,i}|\Big)\Big(|K_{x,j}\rangle\rangle - |K_{y,j}\rangle\rangle\Big)\Big|^2\Big] \qquad (151) \\ &\leq \frac{16}{r^2 d_1^2}\sum_{i,j=1}^r \mathbb{E}\big[|\langle\langle K_{x,i}|K_{x,j}\rangle\rangle|^2 + |\langle\langle K_{y,i}|K_{y,j}\rangle\rangle|^2 + |\langle\langle K_{x,i}|K_{y,j}\rangle\rangle|^2 + |\langle\langle K_{y,i}|K_{x,j}\rangle\rangle|^2\big] \\ &\leq \frac{32}{r^2 d_1^2}\sum_{i,j=1}^r \mathbb{E}\big[|\langle\langle K_{x,i}|K_{x,j}\rangle\rangle|^2 + |\langle\langle K_{y,i}|K_{y,j}\rangle\rangle|^2\big] \qquad (152) \\ &= \frac{64}{r^2 d_1^2}\sum_{i,j=1}^r \mathbb{E}\big[|\langle\langle K_{x,i}|K_{x,j}\rangle\rangle|^2\big] = \frac{64}{r^2 d_1^2}\sum_{i,j=1}^r \mathbb{E}\Big[\Big|\operatorname{tr}(K_{x,i}^\dagger K_{x,j})\Big|^2\Big], \\ &\leq \frac{384}{r^3}, \qquad (153)\end{aligned}$$

where Equation (151) is due to Equation (141), Equation (152) due to the similar argument as that in Equation (98), and in Equation (153), the first inequality is due to Lemma 7.4 and that Δ is an isometry from $\mathcal{H}_{\mathrm{A}}$ to $\mathcal{H}_{\mathrm{B}}[\chi+1:d_2]\otimes\mathcal{H}_{\mathrm{anc}}$ (further note that $\dim(\mathcal{H}_{\mathrm{anc}}) = r \leq d_1 d_2/2 \leq 2d_1\lfloor\frac{d_2}{2}\rfloor \leq 2d_1(d_2-\chi) = 2\dim(\mathcal{H}_{\mathrm{A}})\dim(\mathcal{H}_{\mathrm{B}}[\chi+1:d_2])$ so we choose $k=2$ in Lemma 7.4). □

6 Query lower bounds

Now, we combine the results from Section 4 and Section 5 to give the lower bounds for quantum channel tomography.

Theorem 6.1. *Let d_1, d_2, r be positive integers such that $d_1 \leq rd_2 \leq \frac{4}{3}d_1$. Suppose $\varepsilon \leq \frac{1}{39660}$ and $rd_1d_2 \geq 7\cdot 10^{25}$. Then, tomography of an unknown channel $\mathcal{E} \in \mathbf{QChan}_{d_1,d_2}^r$ to within Choi trace norm or diamond norm error $\frac{\varepsilon}{3966000}$ with probability at least $2/3$ must use at least $n \geq \frac{1}{2\cdot 10^{19}}\cdot\frac{d_1^2}{\varepsilon}$ queries to $\mathcal{E}$.*

Proof. This can be obtained by combining Theorem 4.1 with Theorem 5.2. Specifically, a tomography algorithm to within Choi trace norm error $\frac{\varepsilon}{2\cdot 1983000} = \frac{\varepsilon}{3966000}$ can discriminate between the quantum channels constructed by Theorem 5.2. Therefore, this algorithm can be used to discriminate between the corresponding Stinespring dilation isometries (simply discard the ancilla system and then apply this algorithm). Note that these isometries coincide with those defined in Section 4.1.1 (up to a change of basis). Furthermore, note that $\frac{4}{3}d_1^2 \geq rd_1d_2 \geq 7 \cdot 10^{25}$ implies $d_1 \geq 7 \cdot 10^{12}$. Therefore, Theorem 4.1 applies and shows that any such algorithm must use at least $n \geq \frac{1}{2 \cdot 10^{19}} \cdot \frac{d_1^2}{\varepsilon}$ queries. Since diamond norm is always larger than or equal to the Choi trace norm, this lower bound also applies to diamond norm tomography algorithms. □

Theorem 6.2. *Let $c \in (0, 1]$ be constant, and d_1, d_2, r be positive integers such that $(1+c)d_1 \leq rd_2$. Suppose $\varepsilon \leq \frac{c^{3/2}}{25600}$ and $rd_1d_2 \geq \frac{2 \cdot 10^{30}}{c^{10}}$. Then, tomography of an unknown channel $\mathcal{E} \in \mathbf{QChan}^r_{d_1,d_2}$ to within Choi trace norm or diamond norm error $\frac{c^{3/2}}{25600}\varepsilon$ with probability at least $2/3$ must use at least $n \geq \frac{c^6}{10^{16}} \cdot \frac{rd_1d_2}{\varepsilon^2}$ queries to $\mathcal{E}$.*

Proof. This can be obtained by combining Theorem 4.4 with Theorem 5.4, Theorem 5.10 and Theorem 5.16. Specifically, the regime $d_1 < rd_2$ can be divided into three regime:

(i) $rd_2 < d_1 + r$,

(ii) $d_1 + r \leq rd_2$ and $r \leq d_1$,

(iii) $d_1 + r \leq rd_2$ and $d_1 < r$,

which are covered by Theorem 5.4, Theorem 5.10 and Theorem 5.16, respectively. We define $\kappa = \min\{\frac{rd_2 - d_1}{d_1}, 1\}$. By assumption, we know that $\kappa \geq c$. Note that $rd_2 \leq \frac{2}{c}(rd_2 - d_1)$, $d_1 \leq \frac{1}{c}(rd_2 - d_1)$ and $rd_1d_2 \leq \frac{2}{c^2}(rd_2 - d_1)^2$. Since $rd_1d_2 \geq \frac{2 \cdot 10^{30}}{c^{10}}$, we know $rd_2 - d_1 \geq \frac{10^{15}}{c^4}$.

In regime (i), first note that $2r \leq rd_2 < d_1 + r$ (since $d_2 = 1$ is trivial), which means $d_1 > r$. Thus $rd_2 - d_1 \leq r < d_1$. A tomography algorithm to within Choi trace norm error $\frac{c^{3/2}}{9200}\varepsilon$ can discriminate between $\exp(\frac{c^2}{480001}(rd_2 - d_1)^2)$ quantum channels constructed by Theorem 5.4. Therefore, this algorithm can be used to discriminate between the corresponding Stinespring dilation isometries (simply discard the ancilla system and then apply this algorithm). Note that these isometries coincide with those defined in Section 4.2.1 (up to a change of basis). Furthermore, note that $rd_2 - d_1 \geq \frac{2 \cdot 10^{13}}{c^4}$. Therefore, Theorem 4.4 applies and shows that any such algorithm must use n queries such that

$$n \geq \frac{c^4}{8 \cdot 10^{12}} \cdot \frac{(rd_2 - d_1)^2}{\varepsilon^2} \geq \frac{c^6}{16 \cdot 10^{12}} \cdot \frac{rd_2d_1}{\varepsilon^2}.$$

In regime (ii), we define $d' = \min\{d_1, rd_2 - d_1\}$. A tomography algorithm to within Choi trace norm error $\frac{c^{3/2}}{25600}\varepsilon$ can discriminate between $\exp(\frac{c^2}{3840001}(rd_2 - d_1)d')$ quantum channels constructed by Theorem 5.10. Therefore, this algorithm can be used to discriminate between the corresponding Stinespring dilation isometries (simply discard the ancilla system and then apply this algorithm). Note that these isometries coincide with those defined in Section 4.2.1 (up to a change of basis). Furthermore, note that $rd_2 - d_1 \geq \frac{10^{15}}{c^4}$. Therefore, Theorem 4.4 applies and shows that any such algorithm must use n queries such that

$$n \geq \frac{c^4}{5 \cdot 10^{14}} \cdot \frac{(rd_2 - d_1)d'}{\varepsilon^2} \geq \frac{c^6}{10^{16}} \cdot \frac{rd_1d_2}{\varepsilon^2}.$$

In regime (iii), note that $rd_2 - d_1 \geq r > d_1$. Furthermore, since $rd_2 \geq 2r > 2d_1$ (since $d_2 = 1$ is trivial), we have $rd_2 - d_1 \geq rd_2/2$. A tomography algorithm to within Choi trace norm error $\frac{1}{4000}\varepsilon$ can discriminate between $\exp(\frac{1}{307201}rd_1d_2) \geq \exp(\frac{1}{307201}(rd_2 - d_1)d_1)$ quantum channels constructed by Theorem 5.16. Therefore, this algorithm can be used to discriminate between the corresponding Stinespring dilation isometries (simply discard the ancilla system and then apply this algorithm). Note that these isometries coincide with those defined in Section 4.2.1 (up to a change of basis). Furthermore, note that $rd_2 - d_1 \geq 7 \cdot 10^{12}$. Therefore, Theorem 4.4 applies and shows that any such algorithm must use n queries such that

$$n \geq \frac{1}{4 \cdot 10^{12}} \cdot \frac{(rd_2 - d_1)d_1}{\varepsilon^2} \geq \frac{1}{8 \cdot 10^{12}} \cdot \frac{rd_1 d_2}{\varepsilon^2}.$$

Then, combining the results in these three regimes, we obtain the desired lower bound. Moreover, the Choi trace norm lower bound directly implies the same lower bound for diamond norm. □

Theorem 6.3. *Let d_1, d_2, r be positive integers such that $d_1 < rd_2 \leq 2d_1$. Suppose $\varepsilon \leq \frac{1}{160}$ and $rd_2 - d_1 \geq 64 \cdot 38404^2$. Then, tomography of an unknown channel $\mathcal{E} \in \mathbf{QChan}^r_{d_1,d_2}$ to within diamond norm error $\frac{\varepsilon}{160}$ with probability at least $2/3$ must use at least $n \geq \frac{1}{32 \cdot 38404^2} \cdot \frac{(rd_2 - d_1)^2}{\varepsilon^2}$ queries to $\mathcal{E}$.*

Proof. This can be easily obtained by combining Theorem 4.4 with Theorem 5.7 and Theorem 5.13. Specifically, the regime $d_1 < rd_2$ can be divided into three regime: (i) $rd_2 \leq d_1 + r$, (ii) $rd_2 \geq d_1 + r$ and $d_1 \geq r$, (iii) $rd_2 \geq d_1 + r$ and $d_1 < r$. Since the case (iii) implies $rd_2 > 2d_1$, so we know that the regime $d_1 < rd_2 \leq 2d_1$ is in the union of regime (i) and (ii), which are covered by Theorem 5.7 and Theorem 5.13, respectively.

In regime (i), a tomography algorithm to within diamond norm error $\varepsilon/160$ can discriminate between $\exp((rd_2 - d_1)^2/4801)$ quantum channels constructed by Theorem 5.7. Therefore, this algorithm can be used to discriminate between the corresponding Stinespring dilation isometries (simply discard the ancilla system and then apply this algorithm). Note that these isometries coincide with those defined in Section 4.2.1 (up to a change of basis). Therefore, Theorem 4.4 applies and shows that any such algorithm must use at least $n \geq \frac{1}{32 \cdot 4801^2} \cdot \frac{(rd_2 - d_1)^2}{\varepsilon^2}$ queries (note that $rd_2 - d_1 \leq d_1$).

In regime (ii), we use a similar argument. Note that $rd_2 - d_1 \leq d_1$. Thus Theorem 5.13 provides $\exp((rd_2 - d_1)^2/38404)$ quantum channels, which also have Stinespring dilation isometries that coincide with those defined in Section 4.2.1 (up to a change of basis). Therefore, Theorem 4.4 applies and shows that any such algorithm must use at least $n \geq \frac{1}{32 \cdot 38404^2} \cdot \frac{(rd_2 - d_1)^2}{\varepsilon^2}$ queries.

Then, combining the results in these two regimes, we obtain the desired lower bound. □

Corollary 6.4. *Let d_1, d_2, r be positive integers such that $d_1 < rd_2 \leq 2d_1$. Suppose $\varepsilon \leq \frac{1}{160 d_1}$ and $rd_2 - d_1 \geq 64 \cdot 38404^2$. Then, tomography of an unknown channel $\mathcal{E} \in \mathbf{QChan}^r_{d_1,d_2}$ to within Choi trace norm error $\frac{\varepsilon}{160}$ with probability at least $2/3$ must use at least $n \geq \frac{1}{32 \cdot 38404^2} \cdot \frac{(rd_2/d_1 - 1)^2}{\varepsilon^2}$ queries to $\mathcal{E}$.*

Proof. Let $\varepsilon' = d_1 \cdot \varepsilon$. By Equation (4), a tomography algorithm to within Choi trace norm error $\frac{\varepsilon}{160}$ is also a tomography algorithm to within diamond norm error $\frac{1}{160}\varepsilon'$. Note that $\varepsilon' \leq \frac{1}{160}$. Thus Theorem 6.3 applies and shows that this algorithm must use $n \geq \frac{1}{32 \cdot 38404^2} \cdot \frac{(rd_2 - d_1)^2}{\varepsilon'^2} = \frac{1}{32 \cdot 38404^2} \cdot \frac{(rd_2/d_1 - 1)^2}{\varepsilon^2}$ queries. □

7 Deferred lemmas

7.1 Technical lemmas

Lemma 7.1. *Suppose G is a compact Lie group and $\rho(\cdot)$ is an action of G on a finite-dimensional Hilbert space $\mathcal{H}$. Suppose $X \in \mathcal{L}(\mathcal{H})$ is positive semidefinite. Then,*

$$\operatorname{tr}\left(\left(\mathop{\mathbb{E}}_{g\sim G}\left[\rho(g)X\rho(g)^{-1}\right]\right)^{-1} X\right) \le \dim(\mathcal{H}),$$

where $\mathop{\mathbb{E}}_{g\sim G}$ denotes the expectation over the Haar measure of G and $(\cdot)^{-1}$ denotes pseudo-inverse.

Proof. Define $\overline{X} = \mathop{\mathbb{E}}_{g\sim G}\left[\rho(g)X\rho(g)^{-1}\right]$. Then

$$\begin{aligned}
\operatorname{tr}\left(\overline{X}^{-1}X\right) &= \operatorname{tr}\left(\overline{X}^{-1}X\right) \\
&= \mathop{\mathbb{E}}_{g\sim G}\Big[\operatorname{tr}\Big(\rho(g)\overline{X}^{-1}X\rho(g)^{-1}\Big)\Big] \\
&= \mathop{\mathbb{E}}_{g\sim G}\Big[\operatorname{tr}\Big(\overline{X}^{-1}\rho(g)X\rho(g)^{-1}\Big)\Big] \\
&= \operatorname{tr}\left(\overline{X}^{-1}\overline{X}\right) \\
&\le \dim(\mathcal{H}),
\end{aligned}$$

where we used that $\overline{X}$ commutes with $\rho(g)$. □

Lemma 7.2. *Suppose n, m are positive integers such that $n \ge m$. For $i \in [n]$, let $\mathcal{H}_i \cong \mathbb{C}^d$ be a Hilbert space. We use $\{|0\rangle, \ldots, |d-1\rangle\}$ to denote an orthonormal basis of $\mathbb{C}^d$. We define the following subspace of $\bigotimes_{i=1}^n \mathcal{H}_i$:*

$$A = \operatorname{span}\left(\left\{\sum_{\substack{S\subseteq[n]\\|S|=m}} |\psi\rangle^{\otimes S} \otimes |0\rangle^{\otimes[n]\setminus S} \,\middle|\, |\psi\rangle \in \Pi\right\}\right),$$

where $\Pi \subseteq \operatorname{span}\{|1\rangle, \ldots, |d-1\rangle\}$ is a subspace orthogonal to $|0\rangle$. Then,

$$\dim(A) = \binom{\dim(\Pi)+m-1}{m}.$$

Proof. We define

$$P = \sum_{\pi\in\mathfrak{S}_n} \mathtt{p}(\pi),$$

where $\mathtt{p}(\cdot)$ denotes the action of $\mathfrak{S}_n$ on $\bigotimes_{i=1}^n \mathcal{H}_i$ (that is, for $\pi \in \mathfrak{S}_n$, $\mathtt{p}(\pi)|\psi_1\rangle \otimes \cdots \otimes |\psi_n\rangle = |\psi_{\pi^{-1}(1)}\rangle \otimes \cdots \otimes |\psi_{\pi^{-1}(n)}\rangle$). We can see that, when restricting to the subspace $\operatorname{span}(\{|\psi\rangle^{\otimes m} \otimes |0\rangle^{\otimes n-m} \mid |\psi\rangle \in \Pi\})$, P is injective and A is exactly the image of P on this subspace. Therefore, A has the same dimension as $\operatorname{span}(\{|\psi\rangle^{\otimes m} \mid |\psi\rangle \in \Pi\}) \cong \vee^m\Pi$. The latter is the symmetric subspace of $\Pi^{\otimes m}$, which is of dimension $\binom{\dim(\Pi)+m-1}{m}$ [Har13]. □

Lemma 7.3. *Let $\mathcal{H}_1$, $\mathcal{H}_2$ and $\mathcal{H}_3$ be Hilbert spaces with dimension d_1, d_2 and d_3, respectively. Let $U_x, U_y \in \mathbb{U}_{d_2 d_3}$ be unitaries acting on $\mathcal{H}_2 \otimes \mathcal{H}_3$. Let $V, W : \mathcal{H}_1 \to \mathcal{H}_2 \otimes \mathcal{H}_3$ be sub-isometries (i.e., $V^\dagger V$, $W^\dagger W \leq I_{\mathcal{H}_1}$). Then, the following function is $\sqrt{\frac{2}{d_1}}$-Lipschitz*

$$f(U_x, U_y) = \frac{1}{d_1}\left\| \operatorname{tr}_{\mathcal{H}_3}\big(|V\rangle\!\rangle (\langle\!\langle U_x W| - \langle\!\langle U_y W|)\big) \right\|_1,$$

i.e., $|f(U_x, U_y) - f(U'_x, U'_y)| \leq \sqrt{\frac{2}{d_1}}\sqrt{\|U_x - U'_x\|_F^2 + \|U_y - U'_y\|_F^2}$. Here $\|\cdot\|_1$ denotes the trace norm and $\|\cdot\|_F$ denotes the Frobenius norm.

Proof. First, note that

$$\begin{aligned}
\left\| \operatorname{tr}_{\mathcal{H}_3}\big(|V\rangle\!\rangle (\langle\!\langle U_x W| - \langle\!\langle U_y W|)\big) \right\|_1 &\leq \Big\| |V\rangle\!\rangle (\langle\!\langle U_x W| - \langle\!\langle U_y W|) \Big\|_1 &(154)\\
&= \big\| |V\rangle\!\rangle \big\| \cdot \big\| |U_x W\rangle\!\rangle - |U_y W\rangle\!\rangle \big\| \\
&\leq \sqrt{d_1} \cdot \big\| |U_x W\rangle\!\rangle - |U_y W\rangle\!\rangle \big\| &(155)\\
&= \sqrt{d_1} \cdot \sqrt{\operatorname{tr}(W^\dagger (U_x - U_y)^\dagger (U_x - U_y) W)} \\
&\leq \sqrt{d_1}\sqrt{\operatorname{tr}((U_x - U_y)^\dagger (U_x - U_y))} = \sqrt{d_1} \cdot \|U_x - U_y\|_F, &(156)
\end{aligned}$$

where Equation (154) is because partial trace is contractive in trace norm, Equation (155) is due to $V^\dagger V \leq I_{H_1}$, Equation (156) is because $W W^\dagger \leq I_{H_2} \otimes I_{H_3}$ since $W^\dagger W \leq I_{H_1}$. Then, we can show that

$$\begin{aligned}
&|f(U_x, U_y) - f(U'_x, U'_y)| \\
\leq& \frac{1}{d_1}\left\| \operatorname{tr}_{\mathcal{H}_3}\big(|V\rangle\!\rangle (\langle\!\langle U_x W| - \langle\!\langle U_y W|)\big) - \operatorname{tr}_{\mathcal{H}_3}\big(|V\rangle\!\rangle (\langle\!\langle U'_x W| - \langle\!\langle U'_y W|)\big) \right\|_1 \\
\leq& \frac{1}{d_1}\left\| \operatorname{tr}_{\mathcal{H}_3}\big(|V\rangle\!\rangle (\langle\!\langle U_x W| - \langle\!\langle U'_x W|)\big) \right\|_1 + \left\| \operatorname{tr}_{\mathcal{H}_3}\big(|V\rangle\!\rangle (\langle\!\langle U_y W| - \langle\!\langle U'_y W|)\big) \right\|_1 \\
\leq& \frac{1}{\sqrt{d_1}}(\|U_x - U'_x\|_F + \|U_y - U'_y\|_F) &(157)\\
\leq& \sqrt{\frac{2}{d_1}}\sqrt{\|U_x - U'_x\|_F^2 + \|U_y - U'_y\|_F^2},
\end{aligned}$$

where Equation (157) is due to Equation (156). ◻

Lemma 7.4. *Suppose $\mathcal{H}_{\mathrm{A}}$, $\mathcal{H}_{\mathrm{B}}$ and $\mathcal{H}_{\mathrm{anc}}$ are Hilbert spaces with dimensions d_1, d_2 and r, respectively. Further assume that $d_1/d_2 \leq r \leq k d_1 d_2$ for some $k \geq 1$. Suppose $U \in \mathbb{U}_{r d_2}$ is a Haar-random unitary acting on $\mathcal{H}_{\mathrm{B}} \otimes \mathcal{H}_{\mathrm{anc}}$ and $\Delta : \mathcal{H}_{\mathrm{A}} \to \mathcal{H}_{\mathrm{B}} \otimes \mathcal{H}_{\mathrm{anc}}$ is an isometry. Then*

$$\sum_{i,j=1}^{r} \mathbb{E}\left[\left|\operatorname{tr}\Big(\Delta^\dagger U^\dagger |i\rangle\langle j|_{\mathrm{anc}} U \Delta\Big)\right|^2\right] \leq \frac{2(1+k)d_1^2}{r}.$$

Proof. For any $i, j \in [r]$,

$$\begin{aligned}
&\mathbb{E}\left[\left|\operatorname{tr}\Big(\Delta^\dagger U^\dagger |i\rangle\langle j|_{\mathrm{anc}} U \Delta\Big)\right|^2\right] \\
=& \sum_{k,l=1}^{d_1} \mathbb{E}\Big[\operatorname{tr}\Big(U^\dagger |i\rangle\langle j|_{\mathrm{anc}} U \Delta |l\rangle\langle k|_{\mathrm{A}} \Delta^\dagger U^\dagger |j\rangle\langle i|_{\mathrm{anc}} U \Delta |k\rangle\langle l|_{\mathrm{A}} \Delta^\dagger\Big)\Big]
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{k,l=1}^{d_1} \frac{1}{r^2 d_2^2 - 1}\Big[\mathrm{tr}\Big(\Delta|l\rangle\langle k|_{\mathrm{A}}\Delta^\dagger\Big) \cdot \mathrm{tr}\Big(\Delta|k\rangle\langle l|_{\mathrm{A}}\Delta^\dagger\Big) \cdot \mathrm{tr}\left(|i\rangle\langle j|_{\mathrm{anc}} \otimes I_{\mathrm{B}} \cdot |j\rangle\langle i|_{\mathrm{anc}} \otimes I_{\mathrm{B}}\right) \\
&\qquad + \mathrm{tr}\Big(\Delta|l\rangle\langle k|_{\mathrm{A}}\Delta^\dagger \cdot \Delta|k\rangle\langle l|_{\mathrm{A}}\Delta^\dagger\Big) \cdot \mathrm{tr}\left(|i\rangle\langle j|_{\mathrm{anc}} \otimes I_{\mathrm{B}}\right) \cdot \mathrm{tr}\left(|j\rangle\langle i|_{\mathrm{anc}} \otimes I_{\mathrm{B}}\right)\Big] \\
&\quad - \frac{1}{r d_2 (r^2 d_2^2 - 1)}\Big[\mathrm{tr}\Big(\Delta|l\rangle\langle k|_{\mathrm{A}}\Delta^\dagger\Big) \cdot \mathrm{tr}\Big(\Delta|k\rangle\langle l|_{\mathrm{A}}\Delta^\dagger\Big) \cdot \mathrm{tr}\left(|i\rangle\langle j|_{\mathrm{anc}} \otimes I_{\mathrm{B}}\right) \cdot \mathrm{tr}\left(|j\rangle\langle i|_{\mathrm{anc}} \otimes I_{\mathrm{B}}\right) \\
&\qquad + \mathrm{tr}\Big(\Delta|l\rangle\langle k|_{\mathrm{A}}\Delta^\dagger \cdot \Delta|k\rangle\langle l|_{\mathrm{A}}\Delta^\dagger\Big) \cdot \mathrm{tr}\left(|i\rangle\langle j|_{\mathrm{anc}} \otimes I_{\mathrm{B}} \cdot |j\rangle\langle i|_{\mathrm{anc}} \otimes I_{\mathrm{B}}\right)\Big] \qquad (158) \\
&= \sum_{k,l=1}^{d_1} \left(\frac{1}{r^2 d_2^2 - 1}(\mathbb{1}_{k=l} d_2 + \mathbb{1}_{i=j} d_2^2) - \frac{1}{r d_2 (r^2 d_2^2 - 1)}(\mathbb{1}_{k=l}\mathbb{1}_{i=j} d_2^2 + d_2) \right) \\
&= \frac{d_1 d_2 + \mathbb{1}_{i=j} d_1^2 d_2^2}{r^2 d_2^2 - 1} - \frac{\mathbb{1}_{i=j} d_1 d_2^2 + d_1^2 d_2}{r^3 d_2^3 - r d_2},
\end{aligned}
$$

where Equation (158) is by Corollary 7.13. Hence,

$$
\begin{aligned}
\sum_{i,j=1}^{r} \mathbb{E}\left[\left|\mathrm{tr}\Big(\Delta^\dagger U^\dagger |i\rangle\langle j|_{\mathrm{anc}} U \Delta\Big)\right|^2\right] &= \frac{r^2 d_1 d_2 + r d_1^2 d_2^2}{r^2 d_2^2 - 1} - \frac{r d_1 d_2^2 + r^2 d_1^2 d_2}{r^3 d_2^3 - r d_2} \\
&\leq \frac{r^2 d_1 d_2 + r d_1^2 d_2^2}{r^2 d_2^2 - 1} \\
&\leq 2\left(\frac{d_1}{d_2} + \frac{d_1^2}{r}\right) \\
&\leq \frac{2(1+k) d_1^2}{r}, \qquad (159)
\end{aligned}
$$

where Equation (159) uses $r \leq k d_1 d_2$. □

Lemma 7.5. *Let d_1, d_2 and r be integers such that $d_1/d_2 \leq r \leq d_1 d_2$. Let $\mathcal{H}_{\mathrm{A}}$ be a Hilbert space of dimension d_1 and $\mathcal{H}_{\mathrm{B}}$ be a Hilbert space of dimension d_2. Then, there exists a set of linear operators $\{K_i\}_{i=1}^r$ with $K_i : \mathcal{H}_{\mathrm{A}} \to \mathcal{H}_{\mathrm{B}}$ satisfying*

- $\mathrm{tr}(K_i^\dagger K_j) = 0$ *for any* $i \neq j$,
- $\mathrm{tr}(K_i^\dagger K_i) \leq 2d_1/r$ *for any* $i \in [r]$,
- $\sum_{i=1}^r K_i^\dagger K_i = I_{\mathrm{A}}$.

Proof. We distinguish between two cases depending on whether $d_1 \leq d_2$ or not.

Case 1: $d_1 \leq d_2$, let $k = \left\lfloor \frac{d_2}{d_1} \right\rfloor \geq 1$. We can write $\mathcal{H}_{\mathrm{B}} \cong \left(\bigoplus_{i=1}^k \mathcal{H}_{\mathrm{A}_i}\right) \oplus \mathbb{C}^{d_3}$, with $\mathcal{H}_{\mathrm{A}_i} \cong \mathbb{C}^{d_1}$ for each i and $d_3 = d_2 - k d_1 < d_1$.

For each $i \in [k]$, there exists $l = d_1^2$ orthogonal $d_1 \times d_1$ unitary matrices $\{U_{i,j}\}_{j \in [l]}$ (one can take the generalized Pauli operators). We may find a subset $S \subseteq [k] \times [l]$ with $|S| = \left\lceil \frac{r}{2} \right\rceil$ because $kl = \left\lfloor \frac{d_2}{d_1} \right\rfloor d_1^2 \geq \left\lceil \frac{d_1 d_2}{2} \right\rceil \geq \left\lceil \frac{r}{2} \right\rceil$. For $(i,j) \in S$, we view $U_{i,j} : \mathcal{H}_{\mathrm{A}} \to \mathcal{H}_{\mathrm{A}_i}$ as a linear operator that maps from $\mathcal{H}_{\mathrm{A}}$ to the i-th block $\mathcal{H}_{\mathrm{A}_i} \subseteq \mathcal{H}_{\mathrm{B}}$, and we define the linear operator $K_{i,j}$ as

$$
K_{i,j} = \frac{1}{\sqrt{|S|}} U_{i,j}.
$$

We then check:

(a) For all $(i,j),(i',j') \in S$,

$$\left|\mathrm{tr}\big(K_{i,j}^{\dagger} K_{i',j'}\big)\right| = \frac{d_1}{|S|}\,\mathbb{1}_{i=i'}\mathbb{1}_{j=j'} \le \frac{2d_1}{r}\,\mathbb{1}_{i=i'}\mathbb{1}_{j=j'}.$$

(b)

$$\sum_{(i,j)\in S} K_{i,j}^{\dagger} K_{i,j} = \sum_{(i,j)\in S} \frac{1}{|S|}\,U_{i,j}^{\dagger} U_{i,j} = I_{\mathrm{A}}.$$

(c) The total number of constructed operators is $|S| \le r$.

Case 2: $d_1 > d_2$, let $k = \lfloor \frac{d_1}{d_2} \rfloor \in [1,r]$ and write $d_1 = kd_2 + d_3$ with $0 \le d_3 < d_2$. We can then write $\mathcal{H}_{\mathrm{A}} = \left(\bigoplus_{i=1}^{k} \mathcal{H}_{\mathrm{B}_i}\right) \oplus \mathcal{H}_{\mathrm{C}}$, with each $\mathcal{H}_{\mathrm{B}_i} \cong \mathbb{C}^{d_2}$.

For each $i \in [k]$, there exists a family $\{U_{i,j}\}_{j\in[l]}$ of $l = \lceil \frac{r}{2k} \rceil \in [1, d_2^2]$ orthogonal $d_2 \times d_2$ unitary matrices. For the block $\mathcal{H}_{\mathrm{B}_i}$ $(i = 1,\ldots,k)$, we view $U_{i,j} : \mathcal{H}_{\mathrm{B}_i} \to \mathcal{H}_{\mathrm{B}}$ as a linear operator that acts nontrivially only on the i-th block $\mathcal{H}_{\mathrm{B}_i} \subseteq \mathcal{H}_{\mathrm{A}}$, and we define the linear operator $K_{i,j}$ as

$$K_{i,j} = \frac{1}{\sqrt{l}}\,U_{i,j}.$$

For the remaining block $\mathcal{H}_{\mathrm{C}}$, since $d_3 < d_2$ we can use the construction similar to that in Case 1. Specifically, we split $\mathcal{H}_{\mathrm{B}}$ into $\lfloor \frac{d_2}{d_3} \rfloor$ equal-dimension blocks (possibly leaving a remainder). For each $i \le \lfloor \frac{d_2}{d_3} \rfloor$, there are d_3^2 orthogonal $d_3 \times d_3$ unitary matrices that maps from $\mathcal{H}_{\mathrm{C}}$ to the i-th block in $\mathcal{H}_{\mathrm{B}}$. These unitaries can be viewed as orthogonal isometries from $\mathcal{H}_{\mathrm{C}}$ to $\mathcal{H}_{\mathrm{B}}$ and there are $\lfloor \frac{d_2}{d_3} \rfloor \cdot d_3^2$ such isometries in total. We choose $r' = \lceil \frac{rd_3}{2d_1} \rceil \le \lfloor \frac{d_2}{d_3} \rfloor \cdot d_3^2$ of them, say $V_1', \ldots, V_{r'}'$, and define

$$K_{k+1,i'} = \frac{1}{\sqrt{r'}} V_{i'}',$$

for $i' \in [r']$. We can check

(a) For all $i, i' \in [k]$ and $j, j' \in [l]$,

$$\mathrm{tr}\big(K_{i,j}^{\dagger} K_{i',j'}\big) = \mathbb{1}_{i=i'}\mathbb{1}_{j=j'}\frac{d_2}{l} \le \mathbb{1}_{i=i'}\mathbb{1}_{j=j'}\frac{2d_1}{r},$$

where we used $\frac{d_2}{l} \le \frac{2kd_2}{r} \le \frac{2d_1}{r}$, and for $i', i'' \in [r']$,

$$\mathrm{tr}\big(K_{k+1,i'}^{\dagger} K_{k+1,i''}\big) = \mathbb{1}_{i'=i''}\frac{d_3}{r'} \le \mathbb{1}_{i'=i''}\frac{2d_1}{r},$$

and obviously $\mathrm{tr}(K_{i,j}^{\dagger} K_{k+1,i'}) = 0$ for $i \in [k]$.

(b)

$$\begin{aligned}\sum_{i=1}^{k}\sum_{j=1}^{l} K_{i,j}^{\dagger} K_{i,j} + \sum_{i'=1}^{r'} K_{k+1,i'}^{\dagger} K_{k+1,i'} &= \sum_{i=1}^{k}\sum_{j=1}^{l} \frac{1}{l}\,I_{\mathrm{B}_i} + \sum_{i'=1}^{r'} \frac{1}{r'} I_{\mathrm{C}} \\ &= I_{\mathrm{A}}.\end{aligned}$$

(c) The total number of constructed operators is

$$lk + r' = \left\lfloor \frac{r}{2k} \right\rfloor k + \left\lceil \frac{rd_3}{2d_1} \right\rceil \le \left\lfloor \frac{r}{2} \right\rfloor + \left\lceil \frac{r}{2} - \frac{rkd_2}{2d_1} \right\rceil \le r.$$

□

7.2 Auxiliary facts

We also need the following facts.

Fact 7.6. *Let $p \in [0,1]$, and n and k be two positive integers such that $n \geq k$. We have that*

$$\binom{n}{k} \leq \exp(nH(k/n)),$$

where $H(\cdot)$ denotes the binary entropy function.

Proof. We observe that

$$\binom{n}{k}\left(\frac{k}{n}\right)^k\left(1-\frac{k}{n}\right)^{n-k} \leq \left(\frac{k}{n}+1-\frac{k}{n}\right)^n = 1,$$

so the inequality $\binom{n}{k} \leq \frac{n^n}{k^k(n-k)^{n-k}}$ follows. □

Fact 7.7. *Let $|\psi\rangle$ be a vector in the support of a positive semidefinite operator M. We have that*

$$M \geq |\psi\rangle\langle\psi| \iff 1 \geq \langle\psi|M^{-1}|\psi\rangle,$$

where M^{-1} is the pseudoinverse of M.

Proof. We have the following chain of equivalences

$$\begin{aligned} M \geq |\psi\rangle\langle\psi| &\iff I_{\mathrm{supp}(M)} \geq M^{-1/2}|\psi\rangle\langle\psi|M^{-1/2} \\ &\iff 1 \geq \mathrm{tr}(M^{-1/2}|\psi\rangle\langle\psi|M^{-1/2}) \\ &\iff 1 \geq \langle\psi|M^{-1}|\psi\rangle, \end{aligned}$$

where $M^{-1/2}$ is the pseudoinverse of $M^{1/2}$. □

Fact 7.8. *We have that for all positive numbers x and M*

$$x\ln(M/x) \leq M/e.$$

Proof. The derivative of the function $f(x) = x\ln(M/x)$ is $f'(x) = \ln(M/x) - 1$. So f' is strictly monotonically decreasing and $f'(M/e) = 0$. Hence $f(x) \leq f(M/e) = M/e$. □

Fact 7.9. *For any square matrix M such that $M^2 = 0$, we have*

$$\|M + M^\dagger\|_1 = 2\|M\|_1.$$

Proof. Note that

$$\|M + M^\dagger\|_1 = \mathrm{tr}\left(\sqrt{(M+M^\dagger)^2}\right) = \mathrm{tr}\left(\sqrt{MM^\dagger + M^\dagger M}\right) = \mathrm{tr}\left(\sqrt{MM^\dagger}\right) + \mathrm{tr}\left(\sqrt{M^\dagger M}\right),$$

where the last equality is by using $\mathrm{supp}(MM^\dagger) \perp \mathrm{supp}(M^\dagger M)$. □

Theorem 7.10 ([MM13, Corollary 17])**.** *Let $k, d \geq 1$. Suppose that $f : (\mathbb{U}_d)^k \to \mathbb{R}$ is L-Lipschitz with respect to the ℓ_2-sum of the 2-norms (Frobenius norm), i.e.*

$$\left|f(U_1,\ldots,U_k) - f(U_1',\ldots,U_k')\right| \leq L\sqrt{\sum_{i=1}^{k}\|U_i - U_i'\|_F^2} \tag{160}$$

for all $U_i, U_i' \in \mathbb{U}_d$, with $i = 1,\ldots,k$. Then, if we independently sample $U_1,\ldots,U_k$ according to the Haar measure on $\mathbb{U}_d$, the following inequality holds for each $t > 0$:

$$\mathbb{P}[f(U_1,\ldots,U_k) \geq \mathbb{E}[f(U_1,\ldots,U_k)] + t] \leq \exp\left(-\frac{dt^2}{12L^2}\right). \tag{161}$$

The following facts from Weingarten calculus are needed in this work. For $\pi \in \mathfrak{S}_n$ a permutation of $[n]$, we denote by $\mathrm{Wg}(\pi, d)$ the Weingarten function of dimension d.

Lemma 7.11. *Let $U \in \mathbb{U}_d$ be a Haar-random unitary and let $\{A_i, B_i\}_{i=1}^n$ be a sequence of complex $(d \times d)$-matrices. We have the following formula for the expectation value:*

$$\begin{aligned}&\mathbb{E}\left[\mathrm{tr}(UB_1U^\dagger A_1 U \ldots UB_nU^\dagger A_n)\right]\\ &\quad= \sum_{\alpha,\beta\in\mathfrak{S}_n} \mathrm{Wg}(\beta\alpha^{-1}, d)\,\mathrm{tr}_{\beta^{-1}}(B_1,\ldots,B_n)\,\mathrm{tr}_{\alpha\gamma_n}(A_1,\ldots,A_n),\end{aligned} \tag{162}$$

where $\gamma_n = (12\ldots n) \in \mathfrak{S}_n$ and, writing σ in terms of cycles $\{C_j\}$ as $\sigma = \prod_j C_j$,

$$\mathrm{tr}_\sigma(M_1,\ldots,M_n) \coloneqq \prod_j \mathrm{tr}\prod_{i\in C_j} M_i.$$

Proof. For elementary matrices $A_1,\ldots,A_n, B_1,\ldots,B_n$, the lemma is exactly [CŚ06, Corollary 2.4]. To see this, let us assume that $A_1 = |i_1'\rangle\langle i_2|, A_2 = |i_2'\rangle\langle i_3|, \ldots, A_n = |i_n'\rangle\langle i_1|$ and $B_1 = |j_1\rangle\langle j_1'|, B_2 = |j_2\rangle\langle j_2'|, \ldots, B_n = |j_n\rangle\langle j_n'|$ for some $i_1,\ldots,i_n,i_1',\ldots,i_n' \in [d]$ and $j_1,\ldots,j_n,j_1',\ldots,j_n' \in [d]$. The LHS of (162) can be computed using [CŚ06, Corollary 2.4]:

$$\begin{aligned}\mathbb{E}\left[\mathrm{tr}(UB_1U^\dagger A_1U\ldots UB_nU^\dagger A_n)\right] &= \mathbb{E}\left[U_{i_1,j_1}\cdots U_{i_n,j_n}\cdot \bar{U}_{i_1',j_1'}\cdots\bar{U}_{i_n',j_n'}\right]\\ &= \sum_{\alpha,\beta\in\mathfrak{S}_n}\mathrm{Wg}(\beta\alpha^{-1},d)\mathbb{1}_{i_1=i'_{\alpha(1)},\ldots,i_n=i'_{\alpha(n)}}\cdot\mathbb{1}_{j_1=j'_{\beta(1)},\ldots,j_n=j'_{\beta(n)}}.\end{aligned}$$

Since the matrices $A_1,\ldots,A_n$ and $B_1,\ldots,B_n$ are elementary, the RHS of (162) can be expressed as follows:

$$\begin{aligned}&\sum_{\alpha,\beta\in\mathfrak{S}_n}\mathrm{Wg}(\beta\alpha^{-1},d)\,\mathrm{tr}_{\beta^{-1}}(B_1,\ldots,B_n)\,\mathrm{tr}_{\alpha\gamma_n}(A_1,\ldots,A_n)\\ &= \sum_{\alpha,\beta\in\mathfrak{S}_n}\mathrm{Wg}(\beta\alpha^{-1},d)\mathbb{1}_{j_{\beta^{-1}(1)}=j_1',\ldots,j_{\beta^{-1}(n)}=j_n'}\cdot\mathbb{1}_{i_2=i'_{\alpha\gamma_n(1)},\ldots,i_1=i'_{\alpha\gamma_n(n-1)}}\\ &= \sum_{\alpha,\beta\in\mathfrak{S}_n}\mathrm{Wg}(\beta\alpha^{-1},d)\mathbb{1}_{j_1=j'_{\beta(1)},\ldots,j_n=j'_{\beta(n)}}\cdot\mathbb{1}_{i_1=i'_{\alpha(1)},\ldots,i_n=i'_{\alpha(n)}}.\end{aligned}$$

Hence, (162) holds for elementary matrices. The generalization is obtained by linearity. □

The following values of the Weingarten function are known [CŚ06, Section 6].

Lemma 7.12. *The function* $\mathrm{Wg}(\pi, d)$ *has the following values:*

- $\mathrm{Wg}((1), d) = \frac{1}{d}$,
- $\mathrm{Wg}((12), d) = \frac{-1}{d(d^2-1)}$,
- $\mathrm{Wg}((1)(2), d) = \frac{1}{d^2-1}$.

Then, we can easily see the following result.

Corollary 7.13. *Let* $U \in \mathbb{U}_d$ *be a Haar-random unitary and* A_1, B_1, A_2, B_2 *are complex* $(d \times d)$*-matrices. We have:*

$$\mathbb{E}\Big[\mathrm{tr}(UB_1U^\dagger A_1 UB_2U^\dagger A_2)\Big] = \frac{1}{d^2-1}\Big[\mathrm{tr}(B_1)\,\mathrm{tr}(B_2)\,\mathrm{tr}(A_1A_2) + \mathrm{tr}(B_1B_2)\,\mathrm{tr}(A_1)\,\mathrm{tr}(A_2)\Big] - \frac{1}{d(d^2-1)}\Big[\mathrm{tr}(B_1B_2)\,\mathrm{tr}(A_1A_2) + \mathrm{tr}(B_1)\,\mathrm{tr}(B_2)\,\mathrm{tr}(A_1)\,\mathrm{tr}(A_2)\Big].$$

Proof. By Lemma 7.11, we have

$$\begin{aligned}
&\mathbb{E}\Big[\mathrm{tr}(UB_1U^\dagger A_1 UB_2U^\dagger A_2)\Big] \\
=&\mathrm{Wg}((1)(2), d)\Big[\mathrm{tr}_{(1)(2)}(B_1, B_2)\,\mathrm{tr}_{(12)}(A_1, A_2) + \mathrm{tr}_{(12)}(B_1, B_2)\,\mathrm{tr}_{(1)(2)}(A_1, A_2)\Big] \\
&\quad + \mathrm{Wg}((12), d)\Big[\mathrm{tr}_{(12)}(B_1, B_2)\,\mathrm{tr}_{(12)}(A_1, A_2) + \mathrm{tr}_{(1)(2)}(B_1, B_2)\,\mathrm{tr}_{(1)(2)}(A_1, A_2)\Big] \\
=&\frac{1}{d^2-1}\Big[\mathrm{tr}(B_1)\,\mathrm{tr}(B_2)\,\mathrm{tr}(A_1A_2) + \mathrm{tr}(B_1B_2)\,\mathrm{tr}(A_1)\,\mathrm{tr}(A_2)\Big] \\
&\quad - \frac{1}{d(d^2-1)}\Big[\mathrm{tr}(B_1B_2)\,\mathrm{tr}(A_1A_2) + \mathrm{tr}(B_1)\,\mathrm{tr}(B_2)\,\mathrm{tr}(A_1)\,\mathrm{tr}(A_2)\Big],
\end{aligned}$$

where the last equality is due to Lemma 7.12. □

8 Isometry channel tomography

8.1 Diamond norm tomography

In this section, we prove a lemma that extends the unitary channel tomography algorithm that uses $O(d^2/\varepsilon^2)$ queries in [HKOT23] to isometry channel tomography.

Lemma 8.1 (Isometry channel tomography in diamond norm)**.** *Suppose that* $d_1 \le d_2$ *are two positive integers and* $\varepsilon \in (0,1)$*. Let* $V : \mathbb{C}^{d_1} \to \mathbb{C}^{d_2}$ *be an isometry and* $\mathcal{V} = V(\cdot)V^\dagger \in \mathbf{ISO}_{d_1,d_2}$ *be the associated isometry channel. Then, there exists an algorithm using* $O(d_1d_2/\varepsilon^2)$ *queries to channel* $\mathcal{V}$ *and it outputs an isometry channel estimate* $\widehat{\mathcal{V}}$ *such that* $\|\mathcal{V} - \widehat{\mathcal{V}}\|_\diamond \le \varepsilon$ *with probability* $\ge 2/3$*. Further, the algorithhm uses these queries in parallel.*

To prove Lemma 8.1, one of the key techniques is the following lemma [CL14, KRT17, GKKT20, HKOT23] for pure state tomography.

Lemma 8.2 (Pure state tomography, cf. [HKOT23, Proposition 2.2])**.** *Suppose d is a positive integer. Then, there exists an algorithm for pure state tomography using $O(d/\varepsilon_{\max})$ copies of the unknown quantum state $|v\rangle \in \mathbb{C}^d$, and it outputs a pure state estimate (by a classical description) $|\widehat{v}\rangle$ such that*

$$|\widehat{v}\rangle = \phi\sqrt{1-\varepsilon}|v\rangle + \sqrt{\varepsilon}|w\rangle,$$

where ϕ is a random phase, $\varepsilon \in [0,1]$ is a random number with $\mathbb{P}[\varepsilon \leq \varepsilon_{\max}] \geq 1 - \exp(-5d)$, and $|w\rangle$ is a Haar random state orthogonal to $|v\rangle$.

We also need the following lemma that allows us to convert a weak tomography algorithm into a standard tomography algorithm, as required by Lemma 8.1.

Lemma 8.3. *Suppose $d_1 \leq d_2$ are two positive integers. Let $V : \mathbb{C}^{d_1} \to \mathbb{C}^{d_2}$ be an isometry and $\mathcal{V} = V(\cdot)V^\dagger \in \mathbf{ISO}_{d_1,d_2}$ be the associated isometry channel. Let $\mathcal{A}$ be an algorithm for weak isometry channel tomography that satisfies the following condition: using queries to $\mathcal{V}$, it outputs an isometry estimate $\widehat{V} : \mathbb{C}^{d_1} \to \mathbb{C}^{d_2}$ such that*

$$\mathbb{P}\left[\exists \text{ diagonal unitary } \Phi : \mathbb{C}^{d_1} \to \mathbb{C}^{d_1}, \left\|V\Phi - \widehat{V}\right\|_{\mathrm{op}} \leq \varepsilon \leq \frac{1}{8}\right] \geq 1 - \eta, \tag{163}$$

where $\|\cdot\|_{\mathrm{op}}$ denotes the operator norm. Then, there exists an algorithm for isometry channel tomography that uses $\mathcal{A}$ twice in parallel and outputs an isometry estimate $\widehat{\mathcal{V}}'$ such that

$$\mathbb{P}\left[\left\|\mathcal{V} - \widehat{\mathcal{V}}'\right\|_{\diamond} \leq 98\varepsilon\right] \geq 1 - 2\eta.$$

Proof. Our proof extends that of [HKOT23, Proposition 2.3] for unitary channel tomography to the setting of isometry channel tomography. Let $\mathcal{A}$ be an algorithm for weak isometry channel tomography as described in Lemma 8.3. We first apply $\mathcal{A}$ that uses queries to the original channel $\mathcal{V}$ to obtain an isometry estimate $\widehat{V_1} : \mathbb{C}^{d_1} \to \mathbb{C}^{d_2}$. In parallel, we can apply $\mathcal{A}$ that uses queries to the modified channel $\mathcal{V} \circ \mathcal{F}$ to obtain another isometry estimate $\widehat{V_2} : \mathbb{C}^{d_1} \to \mathbb{C}^{d_2}$, where $\mathcal{F}$ is the quantum channel for implementing the quantum Fourier transform $F : \mathbb{C}^{d_1} \to \mathbb{C}^{d_1}$.

Combining our condition of $\mathcal{A}$ and the union bound, we obtain

$$\left\|V\Phi_1 - \widehat{V_1}\right\|_{\mathrm{op}} \leq \varepsilon \quad \text{and} \quad \left\|VF\Phi_2 - \widehat{V_2}\right\|_{\mathrm{op}} \leq \varepsilon \tag{164}$$

for some diagonal unitaries $\Phi_1, \Phi_2 : \mathbb{C}^{d_1} \to \mathbb{C}^{d_1}$, with probability $\geq 1 - 2\eta$. Because V is an isometry, we have $V^\dagger V = \sum_{j=0}^{d_1-1} |j\rangle\langle j| = I_{d_1}$. As a result,

$$\begin{aligned}\left\|\widehat{V_1}^\dagger \widehat{V_2} - \Phi_1^\dagger F \Phi_2\right\|_{\mathrm{op}} &\leq \left\|\left(\widehat{V_1}^\dagger - \Phi_1^\dagger V^\dagger\right)\widehat{V_2}\right\|_{\mathrm{op}} + \left\|\Phi_1^\dagger V^\dagger\left(\widehat{V_2} - VF\Phi_2\right)\right\|_{\mathrm{op}} \\ &\leq \left\|V\Phi_1 - \widehat{V_1}\right\|_{\mathrm{op}} + \left\|VF\Phi_2 - \widehat{V_2}\right\|_{\mathrm{op}} \\ &\leq 2\varepsilon.\end{aligned}$$

Now we define $p(k,j)$ to be the proposition

$$\left|\langle k|\left(\widehat{V_1}^\dagger \widehat{V_2} - \Phi_1^\dagger F \Phi_2\right)|j\rangle\right| \leq \frac{4\varepsilon}{\sqrt{d_1}}. \tag{165}$$

Using the pigeonhole principle, one can obtain

$$\text{for any } j = 0 \text{ to } d_1 - 1, \quad \#\{k : p(k,j)\} \geq \frac{3d_1}{4}, \tag{166}$$

with probability $\geq 1-2\eta$, where $\#\{k : p(k,j)\}$ stands for the number of k such that $p(k,j)$ is satisfied.

Define $\Phi_3 = \sum_{k,j=0}^{d_1-1} \frac{\langle k|\widehat{V_1}^\dagger \widehat{V_2}|j\rangle}{\langle k|F|j\rangle} |k\rangle\langle j|$. If $p(k,j)$ is satisfied, we obtain, from Equation (165),

$$\left|\langle k|\Phi_3|j\rangle - \langle k|\Phi_1^\dagger|k\rangle \cdot \langle j|\Phi_2|j\rangle\right| \leq 4\varepsilon, \tag{167}$$

where we use the fact that $|\langle k|F|j\rangle| = \frac{1}{\sqrt{d_1}}$ for any k, j. Moreover, if both $p(k,0)$ and $p(k,j)$ are satisfied, then from Equation (167) one can derive

$$\left|\frac{\langle k|\Phi_3|j\rangle}{\langle k|\Phi_3|0\rangle} - \frac{\langle j|\Phi_2|j\rangle}{\langle 0|\Phi_2|0\rangle}\right| \leq \frac{2\cdot 4\varepsilon}{1-4\varepsilon} \leq 16\varepsilon. \tag{168}$$

Note that from Equation (166), we have

$$\text{for any } j = 0 \text{ to } d_1 - 1, \quad \#\{k : p(k,j) \wedge p(k,0)\} \geq \frac{d_1}{2}, \tag{169}$$

with probability $\geq 1-2\eta$. For each $j = 0$ to $d_1 - 1$, define a_j and b_j to be the medians of the real parts and the imaginary parts of the set $\left\{\frac{\langle k|\Phi_3|j\rangle}{\langle k|\Phi_3|0\rangle}\right\}_k$, respectively. In this case, Equations (168) and (169) together lead to

$$\left|(a_j + ib_j) - \frac{\langle j|\Phi_2|j\rangle}{\langle 0|\Phi_2|0\rangle}\right| \leq \sqrt{(16\varepsilon)^2 + (16\varepsilon)^2},$$

which further means $\phi_j = \frac{a_j+ib_j}{|a_j+ib_j|}$ satisfies $\left|\phi_j - \frac{\langle j|\Phi_2|j\rangle}{\langle 0|\Phi_2|0\rangle}\right| \leq 48\varepsilon$. Let $\Phi = \sum_{j=0}^{d_1-1} \phi_j |j\rangle\langle j|$, then we have

$$\|\langle 0|\Phi_2|0\rangle \cdot \Phi - \Phi_2\|_{\text{op}} \leq 48\varepsilon \tag{170}$$

with probability $\geq 1-2\eta$. Consequently,

$$\begin{aligned}
\left\|\mathcal{V} - \widehat{\mathcal{V}_2}\Phi^\dagger\mathcal{F}^\dagger\right\|_\diamond &\leq 2\left\|V\langle 0|\Phi_2|0\rangle - \widehat{V_2}\Phi^\dagger F^\dagger\right\|_{\text{op}} && (171)\\
&\leq 2\left\|V\langle 0|\Phi_2|0\rangle - \widehat{V_2}\Phi_2^\dagger F^\dagger\langle 0|\Phi_2|0\rangle\right\|_{\text{op}} + 2\left\|\widehat{V_2}\Phi_2^\dagger F^\dagger\langle 0|\Phi_2|0\rangle - \widehat{V_2}\Phi^\dagger F^\dagger\right\|_{\text{op}}\\
&= 2\left\|V - \widehat{V_2}\Phi_2^\dagger F^\dagger\right\|_{\text{op}} + 2\|\langle 0|\Phi_2|0\rangle\cdot\Phi - \Phi_2\|_{\text{op}} && (172)\\
&\leq 98\varepsilon, && (173)
\end{aligned}$$

with probability $\geq 1-2\eta$. Here, Equation (171) comes from [AKN98, Lemma 12] (see also [KSW08]). Equation (172) exploits the fact that Φ_2 is a diagonal unitary, $\widehat{V_2}$ is an isometry and F is a unitary. Equation (173) is due to Equations (164) and (170). The algorithm can output the isometry channel corresponding to $\widehat{V'} = \widehat{V_2}\Phi^\dagger F^\dagger$ as the final estimate. □

Given the above lemma, one can prove Lemma 8.1 for isometry channel tomography.

Proof of Lemma 8.1. The proof extends that of [HKOT23, Theorem 2.1] to isometry channel tomography. From Lemma 8.3 (with appropriate rescaling of ε), it suffices to construct an algorithm for weak isometry channel tomography that satisfies Equation (163) with $\eta = \frac{1}{6}$. Our algorithm works as follows:

1. Given queries to $\mathcal{V}$, we first use the algorithm in Lemma 8.2 for pure state tomography (taking $d = d_2$ and $\varepsilon_{\max} = \Theta(\varepsilon^2)$ to be determined later) on computational basis input states $|0\rangle, |1\rangle, \ldots, |d_1 - 1\rangle$ to get estimates $|\widetilde{v_j}\rangle$ of $|v_j\rangle = V|j\rangle$ for all j in parallel. Then, the following holds
$$|\widetilde{v_j}\rangle = \phi_j \sqrt{1 - \varepsilon_j}|v_j\rangle + \sqrt{\varepsilon_j}|w_j\rangle, \tag{174}$$
where for each $j = 0$ to $d_1 - 1$, the random variables $\phi_j, \varepsilon_j, |w_j\rangle$ are as in Lemma 8.2.

2. Let $\widetilde{V} = \sum_j |\widetilde{v_j}\rangle\langle j|$. Let $\widetilde{V} = U_2 \Lambda U_1$ be the singular value decomposition of $\widetilde{V}$ with $U_1 \in \mathbb{U}_{d_1}$ and $U_2 \in \mathbb{U}_{d_2}$, respectively. We then output the quantum channel $\widehat{\mathcal{V}}$ corresponding to the isometry $\widehat{V} = U_2 \sum_{j=0}^{d_1-1} |j\rangle\langle j| U_1$.

One can easily compute that the number of queries to $\mathcal{V}$ in the algorithm above is $O\big(d_1 d_2/\varepsilon^2\big)$. Next, to see that $\widehat{V}$ satisfies Equation (163) in Lemma 8.3, we prove
$$\left\| V\Phi - \widetilde{V} \right\|_{\text{op}} \leq \varepsilon/2 \tag{175}$$
for some diagonal unitary $\Phi : \mathbb{C}^{d_1} \to \mathbb{C}^{d_1}$, with probability $\geq 0.97 \geq \frac{5}{6}$. Whenever Equation (175) holds, the estimate $\widehat{V}$ in the algorithm above satisfies
$$\left\| V\Phi - \widehat{V} \right\|_{\text{op}} \leq \left\| V\Phi - \widetilde{V} \right\|_{\text{op}} + \left\| \widetilde{V} - \widehat{V} \right\|_{\text{op}} \leq \varepsilon.$$
Here, we use $\|\widetilde{V} - \widehat{V}\|_{\text{op}} \leq \varepsilon/2$, because once Equation (175) holds, the operator norm between $\widetilde{V}$ and an isometry is at most $\varepsilon/2$ and therefore the differences between the singular values of $\widetilde{V}$ and 1 are at most $\varepsilon/2$.

Suppose $W = \sum_{j=0}^{d_1-1} |w_j\rangle\langle j|$, $\Phi = \sum_{j=0}^{d_1-1} \phi_j |j\rangle\langle j|$, $B_1 = \sum_{j=0}^{d_1-1} \sqrt{\varepsilon_j}|j\rangle\langle j|$, and $B_2 = \sum_{j=0}^{d_1-1} \sqrt{1-\varepsilon_j}|j\rangle\langle j|$. Then, $\left|\sqrt{1-\varepsilon_j} - 1\right| \leq \sqrt{\varepsilon_j}$ for all $j = 0$ to $d_1 - 1$ implies $\|B_2 - I_{d_1}\|_{\text{op}} \leq \|B_1\|_{\text{op}}$, where $I_{d_1} = \sum_{j=0}^{d_1-1} |j\rangle\langle j|$. Using Lemma 8.2 (taking $d = d_2$, where d_2 is sufficiently large), we obtain
$$\|B_1\|_{\text{op}} \leq \sqrt{\varepsilon_{\max}},$$
with probability ≥ 0.99. From the triangle inequality, we have
$$\left\| V\Phi - \widetilde{V} \right\|_{\text{op}} = \|V\Phi(B_2 - I_{d_1}) + WB_1\|_{\text{op}} \leq \|V\|_{\text{op}} \cdot \|\Phi\|_{\text{op}} \cdot \|B_2 - I_{d_1}\|_{\text{op}} + \|W\|_{\text{op}} \cdot \|B_1\|_{\text{op}},$$
which implies
$$\left\| V\Phi - \widetilde{V} \right\|_{\text{op}} \leq \sqrt{\varepsilon_{\max}}(1 + \|W\|_{\text{op}}), \tag{176}$$
with probability ≥ 0.99.

Now we prove
$$\|W\|_{\text{op}} \leq c_W \tag{177}$$
for some constant $c_W > 0$, with probability ≥ 0.98. This will imply Equation (175) by combining with Equation (176) and taking $\varepsilon_{\max} = \Theta(\varepsilon^2)$ to be sufficiently small.

We define the following quantum states: for each $j = 0$ to $d_1 - 1$,
$$|y_j\rangle = \sqrt{\delta_j}\psi_j|v_j\rangle + \sqrt{1-\delta_j}|w_j\rangle, \tag{178}$$
where $\sqrt{\delta_j} = |\langle 0|x_j\rangle|$ is the overlap between a Haar random state $|x_j\rangle \sim \mathbb{C}^{d_2}$ and the state $|0\rangle$, and $\psi_j \sim [0, 2\pi)$ is an uniformly random phase. Here, we require that $|x_0\rangle, \ldots, |x_{d_1-1}\rangle$ are independent

and $\psi_0, \dots, \psi_{d_1-1}$ are independent. In this case, $|y_j\rangle \sim \mathbb{C}^{d_2}$ and $|y_0\rangle, \dots, |y_{d_1-1}\rangle$ are independent. Let $Y = \sum_{j=0}^{d_1-1} |y_j\rangle\langle j|$. According to [Ver18, Theorem 3.4.6, complex version], $\sqrt{d_2}Y$ has its column vectors being independent sub-gaussian isotropic random in $\mathbb{C}^{d_2}$, and we can upper bound the maximal singular value of Y with high probability by [Ver18, Theorem 4.6.1, complex version]:

$$\|Y\|_{\mathrm{op}} \leq c_Y \tag{179}$$

for some constant $c_Y > 0$, with probability ≥ 0.99. Suppose $E_1 = \sum_{j=0}^{d_1-1} \sqrt{\delta_j}\psi_j |j\rangle\langle j|$ and $E_2 = \sum_{j=0}^{d_1-1} \sqrt{1-\delta_j}|j\rangle\langle j|$. As

$$\|W\|_{\mathrm{op}} = \left\|(Y - VE_1)E_2^{-1}\right\|_{\mathrm{op}} \leq \left(\|Y\|_{\mathrm{op}} + \|V\|_{\mathrm{op}} \cdot \|E_1\|_{\mathrm{op}}\right) \cdot \left\|E_2^{-1}\right\|_{\mathrm{op}},$$

by combining with Equations (178) and (179), we obtain

$$\|W\|_{\mathrm{op}} \leq (c_Y + 1) \cdot (1 - \max_j \delta_j)^{-1/2} \tag{180}$$

with probability ≥ 0.99. Since $\sqrt{d_2}|x_j\rangle$ are sub-Gaussian (similar to the case of $\sqrt{d_2}|y_j\rangle$, by [Ver18, Theorem 3.4.6, complex version]), $\sqrt{d_2\delta_j}$ are also sub-gaussian by definition, which yields

$$\mathbb{P}\left[\sqrt{\delta_j} \leq 0.1\right] \geq 1 - e^{-\Theta(d_2)}.$$

Using the union bound and $d_1 \leq d_2$, we have $\mathbb{P}\left[(1 - \max_j \delta_j)^{-1/2} \leq 2\right] \geq 0.99$ for sufficiently large d_2. Then, we can establish Equation (177). □

8.2 Choi trace norm tomography

In this section, we prove the following lemma, which essentially builds on an algorithm for isometry channel tomography in [YMM25], with a slightly different analysis.

Lemma 8.4 (Isometry channel tomography in Choi trace norm, adapted from [YMM25])**.** *Suppose $d \leq D$ are two positive integers and $\varepsilon \in (0, 1)$. Let $V : \mathbb{C}^d \to \mathbb{C}^D$ be an isometry and $\mathcal{V} = V(\cdot)V^\dagger \in \mathbf{ISO}_{d,D}$ be the associated isometry channel. Then, there exists an algorithm that uses $n = O((D-d)d/\varepsilon^2 + d^2/\varepsilon)$ queries to $\mathcal{V}$ and outputs an isometry channel estimate $\widehat{\mathcal{V}}$ such that $\left\|\frac{1}{d}C_{\mathcal{V}} - \frac{1}{d}C_{\widehat{\mathcal{V}}}\right\|_1 \leq \varepsilon$ with probability $\geq 2/3$. Further, the algorithm uses these queries in parallel.*

Proof. Suppose $n \geq 3d^2/2$. Using [YMM25, Lemma S5] together with $g(n) = \left\lfloor \frac{2}{3d^2}n \right\rfloor \geq \frac{1}{3d^2}n$, there exists an algorithm for isometry channel tomography that uses n queries to achieve the average channel fidelity:

$$\mathrm{F} \geq 1 - \frac{\pi^2(d-1)^2}{d^2 g(n)^2} - \frac{D-d}{\frac{n}{d} + \frac{d-1}{2}g(n) + D - d} \geq 1 - \frac{144d^4}{n^2} - \frac{D-d}{\frac{n}{d}} \geq 1 - 144\left(\frac{d^4}{n^2} + \frac{(D-d)d}{n}\right),$$

where the average channel fidelity F is defined as follows:

$$\mathrm{F} := \mathop{\mathbb{E}}_{\mathcal{V}\sim\mathbf{ISO}_{d,D}}\left[\mathop{\mathbb{E}}_{\widehat{\mathcal{V}}}\left[\mathrm{F}_{\mathrm{ch}}(\mathcal{V}, \widehat{\mathcal{V}})\right]\right].$$

Here, $\widehat{\mathcal{V}}$ denotes the output of the tomography algorithm when the input channel is $\mathcal{V}$. It is worth noticing that this algorithm is also covariant. Thus, for any $\mathcal{V} \in \mathbf{ISO}_{d,D}$, we have

$$1 - \mathop{\mathbb{E}}_{\widehat{\mathcal{V}}}\left[\mathrm{F}_{\mathrm{ch}}(\mathcal{V}, \widehat{\mathcal{V}})\right] = 1 - \mathrm{F} \leq 144\left(\frac{d^4}{n^2} + \frac{(D-d)d}{n}\right) \leq \frac{\varepsilon^2}{12},$$

if we set

$$n = \Theta\left(\frac{(D-d)d}{\varepsilon^2} + \frac{d^2}{\varepsilon}\right).$$

Moreover, this implies that with probability at least 2/3, we have

$$\left\|\frac{1}{d}C_{\mathcal{V}} - \frac{1}{d}C_{\widehat{\mathcal{V}}}\right\|_1 = 2\sqrt{1 - \mathrm{F}_{\mathrm{ch}}(\mathcal{V}, \widehat{\mathcal{V}})} \leq 2\sqrt{3 \cdot \frac{\varepsilon^2}{12}} = \varepsilon,$$

where the inequality comes from the Markov's inequality. Further, note that this algorithm uses queries to $\mathcal{V}$ in parallel. □

Acknowledgments

We thank Antonio Anna Mele and Lennart Bittel for helpful discussions regarding the results presented in their paper [MB25]. FG acknowledges financial support from the European Union under the European Research Council (ERC Grant Agreement No. 101165230). The work of ZZ was supported in part by the Australian Research Council under Grant DP250102952.

References

[AA23] Anurag Anshu and Srinivasan Arunachalam. A survey on the complexity of learning quantum states (2023). *arXiv preprint arXiv:2305.20069*, 2023.

[AdW17] Srinivasan Arunachalam and Ronald de Wolf. Guest column: A survey of quantum learning theory. *ACM Sigact News*, 48(2):41–67, 2017.

[AJV01] A. Acín, E. Jané, and G. Vidal. Optimal estimation of quantum dynamics. *Phys. Rev. A*, 64(5):050302, October 2001.

[AKN98] 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.

[Ang25] Armando Angrisani. Learning unitaries with quantum statistical queries. *Quantum*, 9:1817, 2025.

[BBMT04] E. Bagan, M. Baig, and R. Muñoz-Tapia. Quantum reverse engineering and reference-frame alignment without nonlocal correlations. *Phys. Rev. A*, 70(3):030301, September 2004.

[BHK+19] Frédéric Bouchard, Felix Hufnagel, Dominik Koutnỳ, Aazad Abbas, Alicia Sit, Khabat Heshami, Robert Fickler, and Ebrahim Karimi. Quantum process tomography of a high-dimensional quantum communication channel. *Quantum*, 3:138, 2019.

[BM99] Dagmar Bruß and Chiara Macchiavello. Optimal state estimation for d-dimensional quantum systems. *Physics Letters A*, 253(5-6):249–251, 1999.

[BMQ21] Jessica Bavaresco, Mio Murao, and Marco Túlio Quintino. Strict hierarchy between parallel, sequential, and indefinite-causal-order strategies for channel discrimination. *Physical review letters*, 127(20):200504, 2021.

[BMQ22] Jessica Bavaresco, Mio Murao, and Marco Túlio Quintino. Unitary channel discrimination beyond group structures: Advantages of sequential and indefinite-causal-order strategies. *Journal of Mathematical Physics*, 63(4), 2022.

[Car24] Matthias C Caro. Learning quantum processes and hamiltonians via the pauli transfer matrix. *ACM Transactions on Quantum Computing*, 5(2):1–53, 2024.

[CdH+25] Sitan Chen, Jaume de Dios Pont, Jun-Ting Hsieh, Hsin-Yuan Huang, Jane Lange, and Jerry Li. Predicting quantum channels over general product distributions. In *Proceedings of Thirty Eighth Conference on Learning Theory*, volume 291, pages 986–1007. PMLR, 2025.

[CDP08] Giulio Chiribella, G Mauro D'Ariano, and Paolo Perinotti. Quantum circuit architecture. *Physical review letters*, 101(6):060401, 2008.

[CDP09] Giulio Chiribella, Giacomo Mauro D'Ariano, and Paolo Perinotti. Theoretical framework for quantum networks. *Physical Review A—Atomic, Molecular, and Optical Physics*, 80(2):022339, 2009.

[CDPS04] G. Chiribella, G. M. D'Ariano, P. Perinotti, and M. F. Sacchi. Efficient Use of Quantum Resources for the Transmission of a Reference Frame. *Phys. Rev. Lett.*, 93(18):180503, October 2004.

[CDS05] G. Chiribella, G. M. D'Ariano, and M. F. Sacchi. Optimal estimation of group transformations using entanglement. *Phys. Rev. A*, 72(4):042338, October 2005.

[CG25] Sitan Chen and Weiyuan Gong. Efficient pauli channel estimation with logarithmic quantum memory. *PRX Quantum*, 6(2):020323, 2025.

[CGdW25] Yanlin Chen, András Gilyén, and Ronald de Wolf. A quantum speed-up for approximating the top eigenvectors of a matrix. In *Proceedings of the 2025 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 994–1036. SIAM, 2025.

[CGZ24] Sitan Chen, Weiyuan Gong, and Zhihan Zhang. Adaptivity can help exponentially for shadow tomography. *arXiv preprint arXiv:2412.19022*, 2024.

[Che25] Kean Chen. Inverse-free quantum state estimation with heisenberg scaling. *arXiv preprint arXiv:2510.25750*, 2025.

[CHL+23] Sitan Chen, Brice Huang, Jerry Li, Allen Liu, and Mark Sellke. When does adaptivity help for quantum state learning? In *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 391–404. IEEE, 2023.

[Cho75] Man-Duen Choi. Completely positive linear maps on complex matrices. *Linear algebra and its applications*, 10(3):285–290, 1975.

[CL14] Emmanuel J. Candès and Xiaodong Li. Solving quadratic equations via PhaseLift when there are about as many equations as unknowns. *Foundations of Computational Mathematics*, 14(5):1017–1026, 2014.

[CLO+23] Senrui Chen, Yunchao Liu, Matthew Otten, Alireza Seif, Bill Fefferman, and Liang Jiang. The learnability of pauli noise. *Nature Communications*, 14(1):52, 2023.

[CN97] Isaac L. Chuang and Michael A. Nielsen. Prescription for experimental determination of the dynamics of a quantum black box. *Journal of Modern Optics*, 44(11-12):2455–2467, 1997.

[COZ+24] Senrui Chen, Changhun Oh, Sisi Zhou, Hsin-Yuan Huang, and Liang Jiang. Tight bounds on pauli channel learning without entanglement. *Physical Review Letters*, 132(18):180805, 2024.

[CŚ06] Benoît Collins and Piotr Śniady. Integration with respect to the haar measure on unitary, orthogonal and symplectic group. *Communications in Mathematical Physics*, 264(3):773–795, 2006.

[CSST10] Tullio Ceccherini-Silberstein, Fabio Scarabotti, and Filippo Tolli. *Representation theory of the symmetric groups: the Okounkov-Vershik approach, character formulas, and partition algebras*, volume 121. Cambridge University Press, 2010.

[CWLY23] Kean Chen, Qisheng Wang, Peixun Long, and Mingsheng Ying. Unitarity estimation for quantum channels. *IEEE Transactions on Information Theory*, 69(8):5116–5134, 2023.

[CWZ24] Kean Chen, Qisheng Wang, and Zhicheng Zhang. Local test for unitarily invariant properties of bipartite quantum states. *arXiv preprint arXiv:2404.04599*, 2024.

[CYZ25] Kean Chen, Nengkun Yu, and Zhicheng Zhang. Quantum channel tomography and estimation by local test. *arXiv preprint arXiv:2512.13614*, 2025.

[CZY26] Kean Chen, Zhicheng Zhang, and Nengkun Yu. Optimal lower bound for quantum channel tomography in away-from-boundary regime. *arXiv preprint arXiv:2601.10683*, 2026.

[DP01] GM D'Ariano and P Lo Presti. Quantum tomography for measuring experimentally the matrix elements of an arbitrary quantum operation. *Physical review letters*, 86(19):4195, 2001.

[FFGO23] Omar Fawzi, Nicolas Flammarion, Aurélien Garivier, and Aadil Oufkir. Quantum channel certification with incoherent measurements. In *The Thirty Sixth Annual Conference on Learning Theory*, pages 1822–1884. PMLR, 2023.

[FH13] William Fulton and Joe Harris. *Representation Theory: A First Course*, volume 129 of *Graduate Texts in Mathematics*. Springer, 2013.

[FQR24] Marco Fanizza, Yihui Quek, and Matteo Rosati. Learning quantum processes without input control. *PRX Quantum*, 5(2):020367, 2024.

[GKKT20] Madalin Guţă, Jonas Kahn, Richard Kueng, and Joel A. Tropp. Fast state tomography with optimal error bounds. *Journal of Physics A: Mathematical and Theoretical*, 53(20):204001, 2020.

[GL25] Sabee Grewal and Daniel Liang. Efficient learning of structured quantum circuits via Pauli dimensionality and sparsity. *arXiv preprint arXiv:2510.00168*, 2025.

[GML25] Filippo Girardi, Francesco Anna Mele, and Ludovico Lami. Random purification channel made simple. *arXiv preprint arXiv:2511.23451*, 2025.

[GMZ+25] Filippo Girardi, Francesco Anna Mele, Haimeng Zhao, Marco Fanizza, and Ludovico Lami. Random Stinespring superchannel: converting channel queries into dilation isometry queries. *arXiv preprint arXiv:2512.20599*, 2025.

[Har13] Aram W Harrow. The church of the symmetric subspace. *arXiv preprint arXiv:1308.6595*, 2013.

[Hay98] Masahito Hayashi. Asymptotic estimation theory for a finite-dimensional pure state model. *Journal of Physics A: Mathematical and General*, 31(20):4633, 1998.

[Hay06] Masahito Hayashi. Parallel treatment of estimation of SU(2) and phase estimation. *Phys. Lett. A*, 354(3):183–189, May 2006.

[Hay17] Masahito Hayashi. *Quantum information theory.* Springer, 2017.

[HCP23] Hsin-Yuan Huang, Sitan Chen, and John Preskill. Learning to predict arbitrary quantum processes. *PRX Quantum*, 4(4):040337, 2023.

[HHJ+17] Jeongwan Haah, Aram W. Harrow, Zhengfeng Ji, Xiaodi Wu, and Nengkun Yu. Sample-optimal tomography of quantum states. *IEEE Transactions on Information Theory*, page 1–1, 2017.

[HKOT23] Jeongwan Haah, Robin Kothari, Ryan O'Donnell, and Ewin Tang. Query-optimal estimation of unitary channels in diamond distance. In *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 363–390. IEEE, 2023.

[Jam72] Andrzej Jamiołkowski. Linear transformations which preserve trace and positive semidefiniteness of operators. *Reports on mathematical physics*, 3(4):275–278, 1972.

[Kah07] Jonas Kahn. Fast rate estimation of a unitary operation in SU(d). *Physical Review A—Atomic, Molecular, and Optical Physics*, 75(2):022326, 2007.

[Kar83] Kraus Karl. *States, effects, and operations.* Springer, 1983.

[KBLG18] George C Knee, Eliot Bolduc, Jonathan Leach, and Erik M Gauger. Quantum process tomography via completely positive and trace-preserving projection. *Physical Review A*, 98(6):062336, 2018.

[KKEG19] Martin Kliesch, Richard Kueng, Jens Eisert, and David Gross. Guaranteed recovery of quantum processes from few measurements. *Quantum*, 3:171, 2019.

[KR21] Martin Kliesch and Ingo Roth. Theory of quantum system certification. *PRX quantum*, 2(1):010201, 2021.

[KRT17] Richard Kueng, Holger Rauhut, and Ulrich Terstiege. Low rank matrix recovery from rank one measurements. *Applied and Computational Harmonic Analysis*, 42(1):88–116, 2017.

[KSW08] Dennis Kretschmann, Dirk Schlingemann, and Reinhard F Werner. The information-disturbance tradeoff and the continuity of stinespring's representation. *IEEE transactions on information theory*, 54(4):1708–1717, 2008.

[KW99] Michael Keyl and Reinhard F Werner. Optimal cloning of pure states, testing single clones. *Journal of Mathematical Physics*, 40(7):3283–3299, 1999.

[Leu00] Debbie Wun Chi Leung. *Towards robust quantum computation.* Stanford university, 2000.

[MB25] Antonio Anna Mele and Lennart Bittel. Optimal learning of quantum channels in diamond distance. *arXiv preprint arXiv:2512.10214*, 2025.

[MdW16] Ashley Montanaro and Ronald de Wolf. *A Survey of Quantum Property Testing.* Number 7 in Graduate Surveys. Theory of Computing Library, 2016.

[MGC+25] Francesco Anna Mele, Filippo Girardi, Senrui Chen, Marco Fanizza, and Ludovico Lami. Random purification channel for passive gaussian bosons. *arXiv preprint arXiv:2512.16878*, 2025.

[MH07] Keiji Matsumoto and Masahito Hayashi. Universal distortion-free entanglement concentration. *Physical Review A—Atomic, Molecular, and Optical Physics*, 75(6):062338, 2007.

[ML06] M. Mohseni and D. A. Lidar. Direct Characterization of Quantum Dynamics. *Phys. Rev. Lett.*, 97(17):170501, October 2006.

[MM13] Elizabeth Meckes and Mark Meckes. Spectral measures of powers of random matrices. *Electronic Communications in Probability*, 18:1 – 13, 2013.

[MRL08] Masoud Mohseni, Ali T Rezakhani, and Daniel A Lidar. Quantum-process tomography: Resource analysis of different strategies. *Physical Review A—Atomic, Molecular, and Optical Physics*, 77(3):032322, 2008.

[NC10] Michael A. Nielsen and Isaac L. Chuang. *Quantum computation and quantum information.* Cambridge university press, 2010.

[Nie02] Michael A Nielsen. A simple formula for the average gate fidelity of a quantum dynamical operation. *Physics Letters A*, 303(4):249–252, 2002.

[OG26] Aadil Oufkir and Filippo Girardi. Improved lower bounds for learning quantum channels in diamond distance. *arXiv preprint arXiv:2601.04180*, 2026.

[Ouf23a] Aadil Oufkir. *On Adaptivity in Classical and Quantum Learning.* PhD thesis, Ecole normale supérieure de lyon-ENS LYON, 2023.

[Ouf23b] Aadil Oufkir. Sample-optimal quantum process tomography with non-adaptive incoherent measurements. In *2023 IEEE International Symposium on Information Theory (ISIT)*, page 1919–1924. IEEE, June 2023.

[OW16] Ryan O'Donnell and John Wright. Efficient quantum tomography. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pages 899–912, 2016.

[OW17] Ryan O'Donnell and John Wright. Efficient quantum tomography ii. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 962–974, 2017.

[Pan08] Liam Paninski. A Coincidence-Based Test for Uniformity Given Very Sparsely Sampled Discrete Data. *IEEE Trans. Inf. Theory*, 54(10):4750–4755, October 2008.

[PCZ97] J. F. Poyatos, J. I. Cirac, and P. Zoller. Complete characterization of a quantum process: the two-bit quantum gate. *Physical Review Letters*, 78(2):390–393, 1997.

[PS02] Asher Peres and Petra Scudo. Covariant quantum measurements may not be optimal. *J. Mod. Opt.*, July 2002.

[PSTW25] Angelos Pelecanos, Jack Spilecki, Ewin Tang, and John Wright. Mixed state tomography reduces to pure state tomography. *arXiv preprint arXiv:2511.15806*, 2025.

[PSW25] Angelos Pelecanos, Jack Spilecki, and John Wright. The debiased keyl's algorithm: a new unbiased estimator for full state tomography. *arXiv preprint arXiv:2510.07788*, 2025.

[Rag01] Maxim Raginsky. A fidelity measure for quantum channels. *Physics Letters A*, 290(1-2):11–18, 2001.

[RAS+24] Gregory Rosenthal, Hugo Aaronson, Sathyawageeswar Subramanian, Animesh Datta, and Tom Gur. Quantum channel testing in average-case distance. *arXiv preprint arXiv:2409.12566*, 2024.

[RCEK24] Asad Raza, Matthias C Caro, Jens Eisert, and Sumeet Khatri. Online learning of quantum processes. *arXiv preprint arXiv:2406.04250*, 2024.

[SSKKG22] Trystan Surawy-Stepney, Jonas Kahn, Richard Kueng, and Madalin Guta. Projected least-squares quantum process tomography. *Quantum*, 6:844, 2022.

[SSW25] Thilo Scharnhorst, Jack Spilecki, and John Wright. Optimal lower bounds for quantum state tomography. *arXiv preprint arXiv:2510.07699*, 2025.

[Sti55] W Forrest Stinespring. Positive functions on C*-algebras. *Proceedings of the American Mathematical Society*, 6(2):211–216, 1955.

[SW22] Mehdi Soleimanifar and John Wright. Testing matrix product states. In *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1679–1701. SIAM, 2022.

[Sza97] Stanislaw J. Szarek. Metric Entropy of Homogeneous Spaces. *arXiv*, January 1997.

[TWZ25] Ewin Tang, John Wright, and Mark Zhandry. Conjugate queries can help. *arXiv preprint arXiv:2510.07622*, 2025.

[Ver18] Roman Vershynin. *High-dimensional probability: An introduction with applications in data science*, volume 47. Cambridge university press, 2018.

[VH25] Francisca Vasconcelos and Hsin-Yuan Huang. Learning shallow quantum circuits with many-qubit gates. In *Conference on Learning Theory*, 2025.

[Wat18] John Watrous. *The theory of quantum information.* Cambridge university press, 2018.

[WD25] Chirag Wadhwa and Mina Doosti. Learning quantum processes with quantum statistical queries. *Quantum*, 9:1739, 2025.

[WLKD25] Chirag Wadhwa, Laura Lewis, Elham Kashefi, and Mina Doosti. Agnostic process tomography. *PRX Quantum*, 6(4):040371, 2025.

[WW25] Michael Walter and Freek Witteveen. A random purification channel for arbitrary symmetries with applications to fermions and bosons. *arXiv preprint arXiv:2512.15690*, 2025.

[YMM25] Satoshi Yoshida, Jisho Miyazaki, and Mio Murao. Quantum advantage in storage and retrieval of isometry channels. *arXiv preprint arXiv:2507.10784*, 2025.

[YNM25] Satoshi Yoshida, Ryotaro Niwa, and Mio Murao. Random dilation superchannel. *arXiv preprint arXiv:2512.21260*, 2025.

[YRC20] Yuxiang Yang, Renato Renner, and Giulio Chiribella. Optimal universal programming of unitary gates. *Physical review letters*, 125(21):210501, 2020.

[Yue23] Henry Yuen. An improved sample complexity lower bound for (fidelity) quantum state tomography. *Quantum*, 7:890, 2023.

[ZLK+24] Haimeng Zhao, Laura Lewis, Ishaan Kannan, Yihui Quek, Hsin-Yuan Huang, and Matthias C Caro. Learning quantum states and unitaries of bounded gate complexity. *PRX Quantum*, 5(4):040306, 2024.

[ZRCK25] Leonardo Zambrano, Sergi Ramos-Calderer, and Richard Kueng. Fast quantum measurement tomography with dimension-optimal error bounds. *arXiv preprint arXiv:2507.04500*, 2025.