

# Inverse-free quantum state estimation with Heisenberg scaling

Kean Chen\*

## Abstract

We present an inverse-free algorithm for pure quantum state estimation that achieves Heisenberg scaling. Specifically, let  $|\psi\rangle$  be an unknown  $d$ -dimensional state and suppose we are given black-box query access to a unitary process  $U$  that prepares it. Our algorithm estimates the state  $|\psi\rangle$  to within trace distance error  $\varepsilon$  using  $O(\min\{d^{3/2}/\varepsilon, d/\varepsilon^2\})$  forward queries to the state-preparation unitary  $U$ . This surpasses the classical scaling  $O(d/\varepsilon^2)$  in the sample-access model, and complements the prior best Heisenberg scaling result  $O(d \log(d)/\varepsilon)$  due to van Apeldoorn, Cornelissen, Gilyén, and Nannicini (SODA 2023), which, by contrast, requires both forward and inverse queries. Moreover, our result implies a query upper bound  $O(\min\{d^{3/2}/\varepsilon, 1/\varepsilon^2\})$  for inverse-free amplitude estimation, improving the prior best upper bound  $O(\min\{d^2/\varepsilon, 1/\varepsilon^2\})$  based on optimal unitary estimation by Haah, Kothari, O'Donnell, and Tang (FOCS 2023), and disproving a conjecture posed in Tang and Wright (2025).

## 1 Introduction

Given access to an unknown  $d$ -dimensional quantum state  $|\psi\rangle$ , how efficiently can we estimate  $|\psi\rangle$  to within trace distance error  $\varepsilon$ ? This state estimation problem has been extensively studied in quantum information and quantum learning theory, and can be categorized into two settings. In the sample-access setting, one assumes access to multiple copies of the unknown state  $|\psi\rangle$  and the sample complexity has been shown to be  $\Theta(d/\varepsilon^2)$  [Hay98, Wer98, BM99, KW99, SSW25] (see also [Key06, HHJ<sup>+</sup>16, OW16, OW17, GKKT20, Yue23, SSW25, PSW25] for the mixed state case). In contrast, the query-access setting assumes access to a state-preparation unitary  $U$  of  $|\psi\rangle$ . In this model, van Apeldoorn, Cornelissen, Gilyén and Nannicini [vACGN23] showed that  $\tilde{O}(d/\varepsilon)^1$  queries suffice if both *forward* and *inverse* queries to  $U$  are available. The key distinction between these two models lies in their error scaling: the query model achieves the Heisenberg scaling  $\Theta(1/\varepsilon)$ , whereas the sample model is limited to the classical scaling  $\Theta(1/\varepsilon^2)$ .

Many studies have explored and established the theoretical advantages of inverse queries in quantum learning and metrology tasks (e.g., [FK18, CSM23, SHH25, TW25]). Recently, Tang and Wright [TW25] established that, in the absence of inverse query access, quantum amplitude estimation requires a query complexity of  $\Omega(\min\{d, 1/\varepsilon^2\})$ . This stands in sharp contrast to the well-known upper bound of  $O(1/\varepsilon)$  [BHMT02], which assumes access to both forward and inverse queries. It therefore reveals that disallowing inverse access incurs an additional  $\Omega(d)$  overhead in order to go beyond the classical scaling  $\Theta(1/\varepsilon^2)$ .

Conversely, for the unitary estimation task, Haah, Kothari, O'Donnell, and Tang [HKOT23] showed that inverse access is not necessary to achieve the optimal query complexity  $\Theta(d^2/\varepsilon)$  (see also the work by Yang, Renner, and Chiribella [YRC20] for unitary estimation under an average-case distance). Based on this, [TW25] conjectured that the query complexity of inverse-free amplitude

---

\*University of Pennsylvania, Philadelphia, USA. Email: [keanchen.gan@gmail.com](mailto:keanchen.gan@gmail.com)

<sup>1</sup> $\tilde{O}$  hides logarithmic factors.

estimation is  $\Theta(\min\{d^2/\varepsilon, 1/\varepsilon^2\})$ , suggesting that achieving inverse-free Heisenberg scaling requires estimating the entire unitary. The intuition behind this conjecture is a two-regime characterization of the inverse-free setting:

- *Estimating all  $\Theta(d^2)$  features of a unitary* can be done coherently with a  $\Theta(1/\varepsilon)$  scaling;
- *Estimating one desired feature of the unitary* cannot be done better than the naive incoherent  $\Theta(1/\varepsilon^2)$  algorithm.

## 1.1 Our results

In this paper, we study the quantum state estimation problem in the query access model. We present an inverse-free state estimation algorithm with query complexity  $O(\min\{d^{3/2}/\varepsilon, d/\varepsilon^2\})$ . This further implies a query upper bound  $O(\min\{d^{3/2}/\varepsilon, 1/\varepsilon^2\})$  for inverse-free amplitude estimation by combining it with the naive  $O(1/\varepsilon^2)$  amplitude estimation. Our results demonstrate that, for both state and amplitude estimation, Heisenberg scaling can be achieved without incurring a  $\Theta(d^2)$  overhead in the inverse-free setting, and thereby disproving the conjecture proposed in [TW25] (see Section 1.4 for further discussions). We formalize the problem and our main results as follows.

**Problem 1.1** (State estimation). *Let  $\mathcal{H} \cong \mathbb{C}^d$  be a  $d$ -dimensional Hilbert space with an orthonormal basis  $\{|1\rangle, |2\rangle, \dots, |d\rangle\}$ . Suppose  $U$  is an unknown unitary on  $\mathcal{H}$ . The goal is to estimate the state  $U|d\rangle$ <sup>2</sup> to within trace distance error  $\varepsilon$  using only forward queries to  $U$ .*

**Theorem 1.2.** *There is a quantum algorithm that uses  $n$  forward queries to  $U$  and outputs a classical description of a unit vector  $|\psi\rangle$ . The output  $|\psi\rangle$  is random and satisfies*

$$1 - \mathbf{E}_{|\psi\rangle} [|\langle\psi|U|d\rangle|^2] \leq O\left(\frac{d^3}{n(n+d^2)}\right). \quad (1)$$

By Theorem 1.2, we can easily obtain the following corollaries.

**Corollary 1.3.** *The inverse-free state estimation problem can be solved using  $O(\min\{d^{3/2}/\varepsilon, d/\varepsilon^2\})$  queries.*

*Proof.* Using Theorem 1.2 and letting  $n = O(\min\{d^{3/2}/\varepsilon, d/\varepsilon^2\})$ , we can obtain a random unit vector  $|\psi\rangle$  such that  $1 - \mathbf{E}_{|\psi\rangle} [|\langle\psi|U|d\rangle|^2] \leq \varepsilon^2/100$ , which means  $\Pr[1 - |\langle\psi|U|d\rangle|^2 \leq \varepsilon^2/10] \geq 2/3$ . Note that  $\|\langle\psi|\psi\rangle - U|d\rangle\langle d|U^\dagger\|_1 = 2\sqrt{1 - |\langle\psi|U|d\rangle|^2}$  where  $\|\cdot\|_1$  denotes the trace norm. Thus the trace distance error is at most  $\varepsilon$  with probability at least 2/3.  $\square$

Then, we introduce our result for amplitude estimation.

**Problem 1.4** (Amplitude estimation). *Let  $\mathcal{H} \cong \mathbb{C}^d$  be a  $d$ -dimensional Hilbert space with an orthonormal basis  $\{|1\rangle, |2\rangle, \dots, |d\rangle\}$ . Suppose  $U$  is an unknown unitary on  $\mathcal{H}$  and  $\Pi$  is an orthogonal projector onto a subspace of  $\mathcal{H}$ . The goal is to estimate the amplitude of  $U|d\rangle$  on  $\Pi$ , i.e.  $\sqrt{\langle d|U^\dagger\Pi U|d\rangle}$ , to within error  $\varepsilon$  using only forward queries to  $U$ .*

**Corollary 1.5.** *The inverse-free amplitude estimation problem can be solved using  $O(\min\{d^{3/2}/\varepsilon, 1/\varepsilon^2\})$  queries.*

---

<sup>2</sup>We consider the estimation of  $U|d\rangle$  rather than other  $U|i\rangle$ , simply because this choice is more convenient within the conventional notation of the Gelfand–Tsetlin basis used later in this work.

*Proof.* By Fact 5.1, any state estimation algorithm with trace distance error  $\varepsilon$  is also an amplitude estimation algorithm with error  $O(\varepsilon)$ . Therefore, Corollary 1.3 combined with the naive  $O(1/\varepsilon^2)$  amplitude estimation can imply an  $O(\min\{d^{3/2}/\varepsilon, 1/\varepsilon^2\})$  query upper bound for the inverse-free amplitude estimation.  $\square$

## 1.2 Related work

**State estimation with Heisenberg scaling.** Van Apeldoorn, Cornelissen, Gilyén and Nannicini [vACGN23] developed a family of quantum state estimation algorithms using state-preparation unitaries, based on block encoding techniques [LC19, CGJ19, vAG19] and quantum singular value transformation [GSLW19]. They showed that  $O(d \log(d)/\varepsilon)$  queries to the state-preparation unitary and its inverse suffice, and established a nearly tight lower bound  $\tilde{\Omega}(d/\varepsilon)$ . This result achieves a quadratic improvement in the dependence on the error parameter, reducing the classical scaling of  $\Theta(1/\varepsilon^2)$  (in the sample access model) to  $\Theta(1/\varepsilon)$ , thereby achieving what is known as Heisenberg scaling. Moreover, they also provided algorithms applicable to estimating a mixed state  $\rho$  of rank  $r$ , assuming that the state-preparation unitary produces a purification of  $\rho$ . In this case, they showed that  $O(dr \text{polylog}(d)/\varepsilon)$  queries are sufficient. However, since their methods are based on quantum singular value transformation, the use of inverse queries appear to be an inherent limitation of their approach.

**Power of inverse queries.** The power of inverse queries has been extensively investigated across a wide range of quantum information tasks. A particularly relevant and directly related question, known as unitary time-reversal, asks how to simulate a single call to the inverse unitary  $U^{-1}$  using only forward queries to  $U$ . It was recently established that this task has query complexity  $\Theta(d^2)$  [CML<sup>+</sup>24, OYM25, CYZ25], implying that an exponential overhead is unavoidable when attempting to directly simulate an algorithm involving inverses in an inverse-free setting. Exponential separations between the inverse-ful and inverse-free models have also been demonstrated in several other contexts, including quantum learning [CSM23, SHH25] and preimage finding with an in-place permutation oracle [FK18]. Notably, amplitude estimation—a fundamental problem that serves as a key subroutine in many quantum algorithms—has been shown to require inverse access. Tang and Wright [TW25] established a query lower bound  $\Omega(\min\{d, 1/\varepsilon^2\})$  in the inverse-free model, thereby separating it from the  $O(1/\varepsilon)$  scaling achievable when inverse queries are available [BHMT02].

In contrast, for the unitary estimation task [Hay06, Kah07, Hay11, Hay16, YRC20, HKOT23, GL25, YYM25], one can achieve optimal query complexity without using inverses. Kahn [Kah07] showed an inverse-free unitary estimation algorithm with query complexity  $O(f(d)/\varepsilon)$ , where the estimate achieves an entanglement infidelity error  $\varepsilon^2$  and  $f(d)$  was left unspecified. Notably, in 2020, Yang, Renner, and Chiribella [YRC20] were able to fully determine the scaling to be  $O(d^2/\varepsilon)$  (see also [YYM25] for an improved characterization of  $f(d)$ ). More recently, Haah, Kothari, O'Donnell, and Tang [HKOT23] proved that the scaling  $O(d^2/\varepsilon)$  even holds for diamond-norm error  $\varepsilon$ , along with a matching lower bound  $\Omega(d^2/\varepsilon)$ , even when inverses are allowed. These findings highlight that the power of inverse queries varies substantially across tasks. A general understanding of when and how an advantage exists remains an open question.

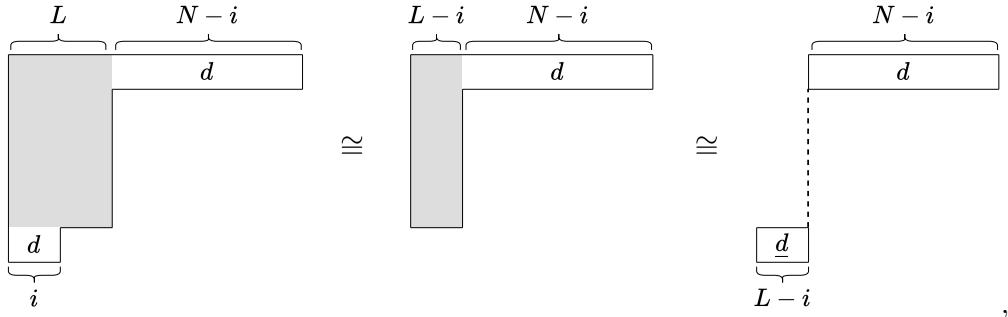
## 1.3 Overview of techniques

The estimation algorithm is conceptually simple. We apply  $U^{\otimes n}$  in parallel to a suitably chosen probe state  $|p\rangle \in \mathcal{H}^{\otimes n}$  and then perform an appropriate measurement on the resulting state to extract the desired information. The design of the probe state and the measurement is based on

representation-theoretic techniques, inspired by the works [Kah07, YRC20]. However, different from them, we are only interested in the last column of  $U$  (i.e.,  $U|d\rangle$ ), thus the probe state should be tailored to extract information exclusively from that column. Guided by this intuition, we construct a family of semistandard Young tableaux  $\Gamma_i$  defined in Figure 1a, whose corresponding Gelfand–Tsetlin basis vectors  $|\Gamma_i\rangle$  possess the desired properties. The probe state  $|p\rangle$  is then constructed as a superposition of these vectors with carefully chosen coefficients, enabling a partial yet efficient estimation of the target unitary  $U$ .

We provide here some intuition about the state  $|\Gamma_i\rangle$ . Consider the subgroup  $\mathbb{U}_{d-1} \subset \mathbb{U}_d$  consisting of unitaries that fix the vector  $|d\rangle$ . By the definition of Gelfand–Tsetlin basis, the action of any  $V \in \mathbb{U}_{d-1}$  on  $|\Gamma_i\rangle$  corresponds to the action of  $V$  in the 1-dimensional  $\mathbb{U}_{d-1}$ -representation labeled by the Young diagram of  $d-1$  rows and  $L$  columns, which simply multiplies  $|\Gamma_i\rangle$  by a global phase  $\det(V)^L$ . This means, for any  $U, V \in \mathbb{U}_d$  such that  $U|d\rangle = V|d\rangle$ , we have  $U^\dagger V \in \mathbb{U}_{d-1}$  and so the actions of  $U$  and  $V$  on  $|\Gamma_i\rangle$  differ only by the phase  $\det(U^\dagger V)^L$ . Therefore,  $|\Gamma_i\rangle$  probes information solely from the last column.

One can also interpret it using dual representations. For convenience, let us assume the unitary  $U$  has determinant 1 (i.e.,  $U \in \mathbb{SU}_d$ ). When restricting to  $\mathbb{SU}_d$ -representations, we can discard the first  $i$  full-length columns of  $\Gamma_i$ . Then, let  $\lambda$  be the Young diagram with  $d-1$  rows and 1 column and let  $S_i$  be the semistandard Young tableau obtained from  $\lambda$  by filling it with the integers  $\{1, \dots, d\} \setminus i$ . There is a natural isomorphism between  $\mathbb{SU}_d$ -representations  $\mathcal{Q}_\lambda^d \rightarrow \mathcal{Q}_{-\square}^d : |S_i\rangle \mapsto (-1)^{d-i} |\underline{i}\rangle$ , where  $\mathcal{Q}_{-\square}^d$  is the dual representation of  $\mathcal{Q}_{\square}^d \cong \mathcal{H}$ ,  $|\underline{i}\rangle \in \mathcal{Q}_{-\square}^d$  is the covector corresponding to  $|\underline{i}\rangle \in \mathcal{Q}_{\square}^d$ . Thus, under these isomorphisms, we can interpret  $\Gamma_i$  as



where the gray region contains multiple  $S_d$  arranged horizontally, the white region labeled by  $d$  ( $\underline{d}$ ) contains boxes filling with  $d$  ( $\underline{d}$ ), the final tableau represents a mixed semistandard Young tableau [GBO24, Ngu24]. This means  $|\Gamma_i\rangle$  can be viewed as an irreducible components in

$$|\underline{d}\rangle^{\otimes L-i} \otimes |d\rangle^{\otimes N-i} \in (\mathcal{Q}_{-\square}^d)^{\otimes L-i} \otimes (\mathcal{Q}_{\square}^d)^{\otimes N-i}.$$

Note that  $U$  acts as  $U^*$  on  $\mathcal{Q}_{-\square}^d$ . Thus, we can see that the action of  $U$  on  $|\Gamma_i\rangle$  depends only on  $U|d\rangle$ .

Then, we use the pretty good measurement (PGM) [HW94, Bel75a, Bel75b, BK02] to extract information from the state  $U^{\otimes n}|p\rangle$ . Specifically, we consider it as a state discrimination problem where our goal is to identify the label  $U$  from the ensemble  $\{(dV, V^{\otimes n}|p\rangle)\}_{V \in \mathbb{U}_d}$  with  $dV$  denoting the Haar measure. Let  $\hat{U}$  denote the measurement outcome. We set  $\hat{U}|d\rangle$  as the estimate of  $U|d\rangle$ , and show that the expected infidelity of this estimator can be upper bounded by  $O(d^3/(n^2 + nd^2))$ , thereby establishing Theorem 1.2.

	Inverse-free	Inverse-ful
Unitary estimation	$\Theta(d^2/\varepsilon)$ [HKOT23]	
State estimation	$O(\min\{d^{3/2}/\varepsilon, d/\varepsilon^2\})$ (This work) $\tilde{\Omega}(d/\varepsilon)$ [vACGN23]	$\tilde{\Theta}(d/\varepsilon)$ [vACGN23]
Amplitude estimation	$O(\min\{d^{3/2}/\varepsilon, 1/\varepsilon^2\})$ (This work) $\Omega(\min\{d, 1/\varepsilon^2\})$ [TW25]	$\Theta(1/\varepsilon)$ [BHMT02]

Table 1: Query upper and lower bounds for unitary, state and amplitude estimation in inverse-free and inverse-ful settings.

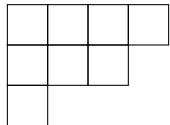
## 1.4 Discussion

Table 1 summarizes the existing results of unitary, state and amplitude estimation in both inverse-free and inverse-ful settings. In this table, we conjecture that our upper bound for inverse-free state estimation is tight. We also conjecture that the query complexity of inverse-free amplitude estimation is  $\Theta(\min\{d/\varepsilon, 1/\varepsilon^2\})$ . Let us consider the additional overhead if one aims to achieve Heisenberg scaling but removes inverse access. If our conjectures hold, then the corresponding overheads are:  $\Theta(1)$  for unitary estimation;  $\Theta(\sqrt{d})$  for state estimation; and  $\Theta(d)$  for amplitude estimation. These might suggest that, rather than exhibiting a sharp transition between estimating all features and estimating a single desired feature, the benefit of inverse access may increase smoothly as the number of features to be estimated decreases.

## 2 Preliminaries

### 2.1 Young diagrams and Young tableaux

A *Young diagram*  $\lambda$  consisting of  $n$  boxes and  $\ell$  rows is a partition  $(\lambda_1, \dots, \lambda_\ell)$  of  $n$  such that  $\sum_{i=1}^\ell \lambda_i = n$  and  $\lambda_1 \geq \dots \geq \lambda_\ell > 0$ . The  $i$ -th row consists of  $\lambda_i$  boxes. By convention, the Young diagram is drawn with left-justified rows, arranged from top to bottom. For example, the Young diagram  $\lambda = (4, 3, 1)$  is drawn as:



We use  $\ell(\lambda)$  to denote the number of rows of  $\lambda$ . We use  $\lambda \vdash n$  to denote that  $\lambda$  is a Young diagram with  $n$  boxes and use  $\lambda \vdash_d n$  to denote that  $\lambda \vdash n$  and  $\ell(\lambda) \leq d$ .

Let  $\mu, \lambda$  be Young diagrams. We write  $\mu \nearrow \lambda$  or  $\lambda \searrow \mu$  if  $\lambda$  can be obtained from  $\mu$  by adding one box. We also write  $\lambda \gtrsim \mu$  or  $\mu \lesssim \lambda$  if  $\ell(\mu) \leq \ell(\lambda) =: \ell$  and

$$\lambda_1 \geq \mu_1 \geq \lambda_2 \geq \mu_2 \geq \dots \geq \lambda_\ell \geq \mu_\ell,$$

where we define  $\mu_i = 0$  if  $i > \ell(\mu)$ . Equivalently, this means the row-lengths of  $\lambda$  and  $\mu$  interlace and we say  $\mu$  interlaces  $\lambda$ . An example is  $(4, 3, 1) \gtrsim (3, 3, 1) \gtrsim (3, 1) \gtrsim (2)$ :

$$\begin{array}{c|c|c|c} \hline & \square & \square & \square \\ \hline \square & \square & \square & \square \\ \hline & \square & \square & \square \\ \hline \end{array} \quad \gtrsim \quad \begin{array}{c|c|c|c} \hline & \square & \square & \square \\ \hline \square & \square & \square & \square \\ \hline & \square & \square & \square \\ \hline \end{array} \quad \gtrsim \quad \begin{array}{c|c|c} \hline & \square & \square \\ \hline \square & \square & \square \\ \hline \end{array} \quad \gtrsim \quad \begin{array}{c|c} \hline \square & \square \\ \hline \end{array}. \tag{2}$$

A *semistandard Young tableau* of shape  $\lambda \vdash n$  and alphabet  $[d]$  is a filling of  $\lambda$  using the integers from  $[d]$ , such that the entries weakly increase from left to right in each row and strictly increase from top to bottom in each column. An example of semistandard Young tableau of shape  $(4, 3, 1)$  and alphabet  $[4]$  is

$$\begin{array}{|c|c|c|c|} \hline 1 & 1 & 2 & 4 \\ \hline 2 & 3 & 3 & \\ \hline 3 & & & \\ \hline \end{array} . \quad (3)$$

We use  $\text{SSYT}(\lambda, d)$  to denote the set of semistandard Young tableaux of shape  $\lambda$  with alphabet  $[d]$ . We also use  $\text{Sh}(S)$  to denote the shape of the semistandard Young tableau  $S$ .

## 2.2 Gelfand-Tsetlin basis

Let  $\mathcal{H} \cong \mathbb{C}^d$  be a  $d$ -dimensional Hilbert space with an orthonormal basis  $\{|1\rangle, |2\rangle, \dots, |d\rangle\}$ . Let  $\mathbb{U}_d$  denote the group of  $d \times d$  unitary matrices on  $\mathcal{H}$  and  $\mathfrak{S}_n$  denote the group of permutations of set  $[n] = \{1, 2, \dots, n\}$ .

The space  $\mathcal{H}^{\otimes n}$  is a representation of the symmetric group  $\mathfrak{S}_n$  and the unitary group  $\mathbb{U}_d$ . The unitary group  $\mathbb{U}_d$  acts by simultaneous “rotation” as  $U^{\otimes n}$  for any  $U \in \mathbb{U}_d$  and the symmetric group  $\mathfrak{S}_n$  acts by permuting tensor factors:

$$P(\pi)|\psi_1\rangle \cdots |\psi_n\rangle = |\psi_{\pi^{-1}(1)}\rangle \cdots |\psi_{\pi^{-1}(n)}\rangle,$$

for any  $\pi \in \mathfrak{S}_n$ . By the Schur-Weyl duality (see, e.g., [FH13]), we have the following decomposition

$$\mathcal{H}^{\otimes n} \stackrel{\mathbb{U}_d \times \mathfrak{S}_n}{\cong} \bigoplus_{\lambda \vdash_d n} \mathcal{Q}_{\lambda}^d \otimes \mathcal{P}_{\lambda}. \quad (4)$$

Here, we use  $\mathcal{Q}_{\lambda}^d$  and  $\mathcal{P}_{\lambda}$  to denote the irreducible representations of  $\mathbb{U}_d$  and  $\mathfrak{S}_n$  corresponding to the Young diagram  $\lambda$ . We remark that in this paper, without loss of generality, all representations are unitary representations and the homomorphisms between representations preserve inner product. We use  $\mathbf{q}_{\lambda}(U)$  to denote the action of  $U$  on  $\mathcal{Q}_{\lambda}^d$ . The dimension of  $\mathcal{Q}_{\lambda}^d$  can be computed using the Weyl dimension formula (see, e.g., [EGH<sup>+</sup>11])

$$\dim(\mathcal{Q}_{\lambda}^d) = \prod_{1 \leq i < j \leq d} \frac{\lambda_i - i - \lambda_j + j}{j - i}. \quad (5)$$

**Branching rule of  $\mathbb{U}_d$ .** We now look deeper into the structure of  $\mathcal{Q}_{\lambda}^d$ . Consider the following chain of embeddings of groups:

$$\mathbb{U}_1 \rightarrow \mathbb{U}_2 \rightarrow \cdots \rightarrow \mathbb{U}_d,$$

where the embedding  $\mathbb{U}_i \rightarrow \mathbb{U}_{i+1}$  is given by

$$U \mapsto \begin{bmatrix} U & 0 \\ 0 & 1 \end{bmatrix},$$

for any  $U \in \mathbb{U}_i$ . Let  $\lambda \vdash_i n$  be a Young diagram of  $n$  boxes with at most  $i$  rows and  $\mathcal{Q}_{\lambda}^i$  be an irreducible representation of  $\mathbb{U}_i$ . Then,  $\mathcal{Q}_{\lambda}^i$  can be viewed as a representation of  $\mathbb{U}_{i-1}$ , which is called the restriction of  $\mathcal{Q}_{\lambda}^i$  to  $\mathbb{U}_{i-1}$  and is denoted by  $\text{Res}_{\mathbb{U}_{i-1}}^{\mathbb{U}_i}(\mathcal{Q}_{\lambda}^i)$ . Note that  $\text{Res}_{\mathbb{U}_{i-1}}^{\mathbb{U}_i}(\mathcal{Q}_{\lambda}^i)$

also decomposes as a direct sum of irreducible representations of  $\mathbb{U}_{i-1}$ . Specifically, we have the following branching rule (see, e.g., [GW09])

$$\text{Res}_{\mathbb{U}_{i-1}}^{\mathbb{U}_i}(\mathcal{Q}_\lambda^i) \stackrel{\mathbb{U}_{i-1}}{\cong} \bigoplus_{\substack{\mu: \mu \prec \lambda \\ \ell(\mu) \leq i-1}} \mathcal{Q}_\mu^{i-1}. \quad (6)$$

Thus, we can see that  $\mathcal{Q}_\lambda^d$  decomposes into a direct sum of pairwise non-equivalent irreducible representations and they are pairwise orthogonal. By iterating this process (i.e., taking further restrictions on  $\mathcal{Q}_\mu^d$ ), we finally obtain a direct sum of the one-dimensional irreducible representations of  $\mathbb{U}_1$  (since  $\mathbb{U}_1$  is Abelian thus has only 1-dimensional irreducible representations). Therefore, this process determines an orthonormal basis (up to phases) of  $\mathcal{Q}_\lambda^d$  and each basis vector can be parameterized by its branching path:

$$\lambda^{(1)} \rightarrow \lambda^{(2)} \rightarrow \dots \rightarrow \lambda^{(d)}, \quad (7)$$

where  $\lambda^{(d)} = \lambda$ ,  $\lambda^{(i-1)} \precsim \lambda^{(i)}$  and  $\ell(\lambda^{(i)}) \leq i$ . This basis is called the *Gelfand-Tsetlin basis* of  $\mathcal{Q}_\lambda^d$ .

Furthermore, we can see that the path in Equation (7) uniquely corresponds to a semistandard Young tableau  $S$  of shape  $\lambda = \lambda^{(d)}$ , where the integer  $i$  is assigned to the boxes in  $\lambda^{(i)} \setminus \lambda^{(i-1)}$ , in which  $\lambda^{(i)} \setminus \lambda^{(i-1)}$  denotes the shape by removing from the shape of  $\lambda^{(i)}$  all the boxes belonging to  $\lambda^{(i-1)}$ . For example, the chain in Equation (2) corresponds to the semistandard Young tableau in Equation (3). By this, we can identify a semistandard Young tableau  $S$  with alphabet  $[d]$  uniquely by a length- $d$  interlacing sequence of Young diagrams as in Equation (7). Thus each vector in the Gelfand-Tsetlin basis of  $\mathcal{Q}_\lambda^d$  is labeled by a semistandard Young tableau  $S$  of shape  $\lambda$  with alphabet  $[d]$ , and we will use  $|S\rangle$  to denote this vector.

### 2.3 Clebsch-Gordan coefficients

Let us consider the tensor product of the irreducible representations  $\mathcal{Q}_\lambda^d$  and  $\mathcal{Q}_\square^d$ , where  $\square$  denotes the Young diagram with only one box and thus  $\mathcal{Q}_\square^d \cong \mathcal{H}$  is just the defining representation of  $\mathbb{U}_d$ . This tensor product representation is generally reducible and decomposes as

$$\mathcal{Q}_\lambda^d \otimes \mathcal{H} \stackrel{\mathbb{U}_d}{\cong} \bigoplus_{\substack{\mu: \lambda \nearrow \mu \\ \ell(\mu) \leq d}} \mathcal{Q}_\mu^d,$$

due to the Pieri's rule (see, e.g., [GW09]). Under this isomorphism, we can write the tensor product of vectors  $|S\rangle \in \mathcal{Q}_\lambda^d$  and  $|i\rangle \in \mathcal{H}$ , where  $S \in \text{SSYT}(\lambda, d)$  and  $i \in [d]$ , as

$$|S\rangle \otimes |i\rangle = \sum_{\substack{\mu: \lambda \nearrow \mu \\ \ell(\mu) \leq d}} \sum_{T \in \text{SSYT}(\mu, d)} C_T^{S,i} |T\rangle,$$

where  $C_T^{S,i}$  is called the Clebsch-Gordan coefficients which can be further expressed as a product of reduced Wigner coefficients [Har05, BCH07, BCH06] or scalar factors [VK92]. When  $i = d$ , the decomposition has a simpler form:

$$|S\rangle \otimes |d\rangle = \sum_{k=1}^d C_k^S |S_k\rangle, \quad (8)$$

where  $S_k$  denotes the semistandard Young tableau obtained from  $S$  by adding a box filling with  $d$  to the  $k$ -th row.<sup>3</sup> Let  $\lambda^{(1)} \rightarrow \dots \rightarrow \lambda^{(d)} = \lambda$  be the interlacing sequence of Young diagrams corresponding to  $S$ , then  $C_k^S$  can be expressed as (see, e.g., Equation 3 in [VK92, Sec. 18.2.10], and also [PSW25, Sec. 10], [GBO24, Apdx. A]):

$$C_k^S = \left| \frac{\prod_{j=1}^{d-1} (\lambda_j^{(d-1)} - j - \lambda_k + k - 1)}{\prod_{j=1, j \neq k}^d (\lambda_j - j - \lambda_k + k)} \right|^{1/2}. \quad (9)$$

### 3 Estimation algorithm

In this section, we introduce our estimation algorithm. Let  $\mathcal{H} \cong \mathbb{C}^d$  be a  $d$ -dimensional Hilbert space and  $U$  be an unknown unitary on  $\mathcal{H}$ . We apply  $U^{\otimes n}$  on a state  $|p\rangle \in \mathcal{H}^{\otimes n}$ , called the probe state. Then we perform a POVM on  $U^{\otimes n}|p\rangle$ , obtaining a “partial” estimate  $\hat{U}$  of  $U$ . Finally, we set  $\hat{U}|d\rangle$  as our estimate of  $U|d\rangle$ .

#### 3.1 A coherent system on $\Gamma$ -shaped tableaux

Let  $n$  be the number of queries. Without loss of generality, we can assume  $n$  is a multiple of  $2d$ . Let

$$L := \frac{n}{2d} \quad \text{and} \quad N := (d+1)L = \frac{(d+1)n}{2d}.$$

Then, we define a family of semistandard Young tableaux that have “ $\Gamma$ ”-like shape.

**Definition 3.1** ( $\Gamma$ -shaped tableaux). *For any integer  $i$  such that  $0 \leq i \leq L$ , let  $\Gamma_i$  and  $\Gamma_i^+$  be the semistandard Young tableaux defined in Figure 1.*

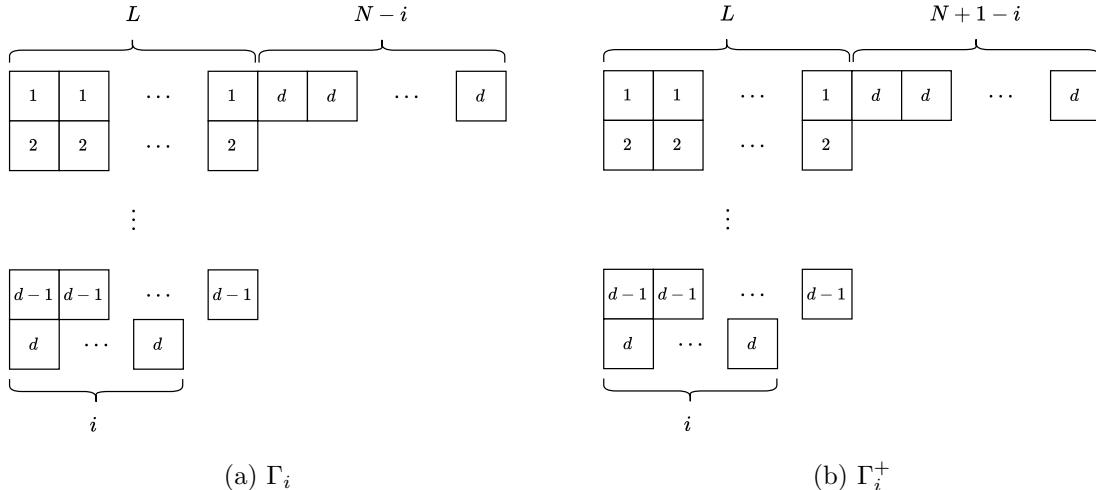


Figure 1: Definition of the semistandard Young tableaux  $\Gamma_i$  and  $\Gamma_i^+$ .

It is easy to see that  $\Gamma_i$  contains  $(d-1)L + N = n$  boxes and  $\Gamma_i^+$  contains  $(d-1)L + N + 1 = n + 1$  boxes. Let  $\gamma_i := \text{Sh}(\Gamma_i)$  and  $\gamma_i^+ = \text{Sh}(\Gamma_i^+)$  be the shapes of  $\Gamma_i$  and  $\Gamma_i^+$ . Then,  $|\Gamma_i\rangle$  and  $|\Gamma_i^+\rangle$  are two vectors in the Gelfand-Tsetlin bases of  $\mathcal{Q}_{\gamma_i}^d$  and  $\mathcal{Q}_{\gamma_i^+}^d$ , respectively.

---

<sup>3</sup>Note that  $S_k$  may not be a valid semistandard Young tableau, but in such case  $C_k^S = 0$ .

We note that these tableaux are related if we consider the tensor product of  $|\Gamma_i\rangle$  and  $|d\rangle$ . By the Clebsch-Gordan decomposition (see Section 2.3),  $|\Gamma_i\rangle \otimes |d\rangle$  decomposes into a sum of vectors corresponding to the semistandard Young tableaux obtained from  $\Gamma_i$  by adding one box filling with  $d$ . It is easy to see that such box can only be added at the first and the  $d$ -th rows. More specifically, from Equation (8) and Equation (9), we have

$$\begin{aligned} |\Gamma_i\rangle \otimes |d\rangle &= \left| \frac{\prod_{1 \leq j \leq d-1} (N-i+j)}{(L+N+d-2i-1) \prod_{2 \leq j \leq d-1} (N-i+j-1)} \right|^{1/2} |\Gamma_i^+\rangle \\ &\quad + \left| \frac{\prod_{1 \leq j \leq d-1} (L+d-i-j-1)}{(L+N+d-2i-1) \prod_{2 \leq j \leq d-1} (L+d-i-j)} \right|^{1/2} |\Gamma_{i+1}^+\rangle \\ &= \sqrt{\frac{N+d-i-1}{L+N+d-2i-1}} |\Gamma_i^+\rangle + \sqrt{\frac{L-i}{L+N+d-2i-1}} |\Gamma_{i+1}^+\rangle. \end{aligned} \quad (10)$$

We define

$$\alpha_i := \frac{N+d-i-1}{L+N+d-2i-1}, \quad \beta_i := \frac{L-i}{L+N+d-2i-1}. \quad (11)$$

Then, Equation (10) can be written as

$$|\Gamma_i\rangle \otimes |d\rangle = \sqrt{\alpha_i} |\Gamma_i^+\rangle + \sqrt{\beta_i} |\Gamma_{i+1}^+\rangle, \quad (12)$$

which is also shown in Figure 2.

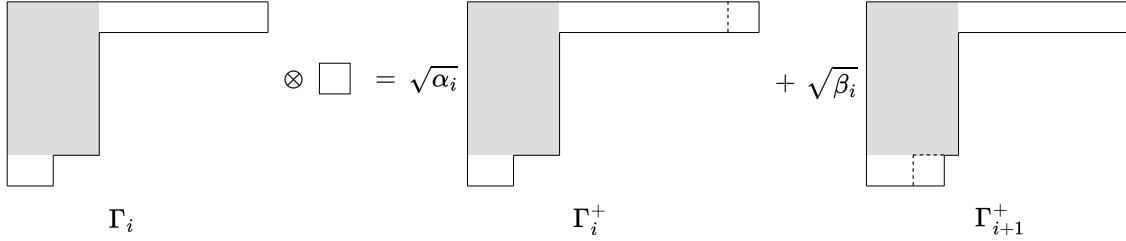


Figure 2: The decomposition of  $|\Gamma_i\rangle \otimes |d\rangle$

### 3.2 Probe state

Then, we use the  $\Gamma$ -shaped tableaux introduced in Section 3.1 to define the probe state  $|p\rangle$ . For each  $0 \leq i \leq L$ , fix an arbitrary unit vector  $|\perp_{\gamma_i}\rangle \in \mathcal{P}_{\gamma_i}$ , where  $\mathcal{P}_{\gamma_i}$  the irreducible representation of  $\mathfrak{S}_n$  labeled by  $\gamma_i$ . The state  $|\perp_{\gamma_i}\rangle$  will be used in our definition of the probe state and measurement, while the final results are independent of the choice of  $|\perp_{\gamma_i}\rangle$ .

**Definition 3.2** (The probe state).

$$|p\rangle := \sum_{i=0}^L f_i |\Gamma_i\rangle |\perp_{\gamma_i}\rangle \in \mathcal{H}^{\otimes n}, \quad (13)$$

where  $f_i$  are real numbers such that  $\sum_{i=0}^L f_i^2 = 1$ , and will be determined in Section 4.1.

### 3.3 Pretty good measurement

By applying  $U^{\otimes n}$  to the probe state  $|p\rangle$ , we obtain a state  $U^{\otimes n}|p\rangle$  that encodes some information of  $U$ . To extract the information from this state, we view it as a state discrimination problem. That is, suppose we are given an unknown state drawn from the ensemble  $\{(dV, V^{\otimes n}|p\rangle)\}_{V \in \mathbb{U}_d}$ , where  $dV$  is the Haar measure, and we want to determine which state we have received (i.e., to identify the corresponding label  $V$ ). Our idea is to perform the so-called ‘‘pretty good measurement’’ (PGM) [HW94, Bel75a, Bel75b, BK02] (recall that given an ensemble  $\{(p_1, \rho_1), \dots, (p_m, \rho_m)\}$ , the PGM defines the POVM  $E_i := E^{-1/2}p_i\rho_iE^{-1/2}$ , where  $E := \sum_i p_i\rho_i$ ).

The PGM with respect to the ensemble  $\{(dV, V^{\otimes n}|p\rangle)\}_{V \in \mathbb{U}_d}$  is defined as a continuous POVM

$$\{E_V\}_{V \in \mathbb{U}_d},$$

where

$$E_V := dV E^{-\frac{1}{2}} V^{\otimes n} |p\rangle \langle p| V^{\dagger \otimes n} E^{-\frac{1}{2}},$$

in which  $dV$  is the Haar measure,  $E^{-\frac{1}{2}}$  is the pseudo-inverse of  $E^{\frac{1}{2}}$  and  $E$  is the average of the ensemble:

$$\begin{aligned} E &:= \int dV V^{\otimes n} |p\rangle \langle p| V^{\dagger \otimes n} \\ &= \int dV \sum_{i,j=0}^L f_i f_j \mathbf{q}_{\gamma_i}(V) |\Gamma_i\rangle \langle \Gamma_j| \mathbf{q}_{\gamma_j}(V^\dagger) \otimes |\perp_{\gamma_i}\rangle \langle \perp_{\gamma_j}| \\ &= \sum_{i=0}^L \frac{f_i^2}{\dim(\mathcal{Q}_{\gamma_i}^d)} I_{\mathcal{Q}_{\gamma_i}^d} \otimes |\perp_{\gamma_i}\rangle \langle \perp_{\gamma_i}|, \end{aligned} \quad (14)$$

where Equation (14) is by the Schur’s lemma. Thus, we can see that

$$E_V = dV |\eta_V\rangle \langle \eta_V|,$$

where

$$|\eta_V\rangle := E^{-1/2} V^{\otimes n} |p\rangle = \sum_{i=0}^L \sqrt{\dim(\mathcal{Q}_{\gamma_i}^d)} \mathbf{q}_{\gamma_i}(V) |\Gamma_i\rangle |\perp_{\gamma_i}\rangle. \quad (15)$$

### 3.4 The algorithm

Then, the estimation algorithm is shown in Algorithm 1.

---

**Algorithm 1** Estimating  $U|d\rangle$ 


---

**Input:**  $n$  forward queries to  $U$ .

**Output:** An estimate of  $U|d\rangle$ .

- 1: Prepare the probe state  $|p\rangle$ . ▷ see Equation (13)
  - 2: Apply  $U^{\otimes n}$  on  $|p\rangle$ .
  - 3: Apply the POVM  $\{dV |\eta_V\rangle \langle \eta_V|\}_{V \in \mathbb{U}_d}$  on  $U^{\otimes n}|p\rangle$  and obtain a classical outcome  $\hat{U}$ . ▷ see Equation (15)
  - 4: **Return**  $\hat{U}|d\rangle$ .
-

## 4 Bounding the error

In this section, we bound the error of Algorithm 1.

For  $0 \leq i \leq L$ , we define

$$\begin{aligned} x_i &:= \sqrt{\frac{(N+d-i-1)(N-i+1)}{(L+N+d-2i-1)(L+N+d-2i)}}, \\ y_i &:= \sqrt{\frac{(L-i+1)(L+d-i-1)}{(L+N+d-2i+1)(L+N+d-2i)}}. \end{aligned} \quad (16)$$

Then, we have the following lemma.

**Lemma 4.1.** *Let  $\widehat{U}|d\rangle$  be the output of Algorithm 1. The expectation of fidelity between  $\widehat{U}|d\rangle$  and  $U|d\rangle$  is*

$$\mathbf{E}_{\widehat{U}}\left[\left|\langle d|\widehat{U}^\dagger U|d\rangle\right|^2\right] = f_0^2 x_0^2 + \sum_{i=1}^L (f_i x_i + f_{i-1} y_i)^2. \quad (17)$$

*Proof.* Consider the probability  $\Pr(\widehat{U}; U)$  of getting the measurement outcome  $\widehat{U}$  when performing the PGM on the state  $U^{\otimes n}|p\rangle$ :

$$\begin{aligned} \Pr(\widehat{U}; U) &= d\widehat{U} |\langle \eta_{\widehat{U}} | U^{\otimes n} | p \rangle|^2 \\ &= d\widehat{U} \left| \sum_{i=0}^L f_i \sqrt{\dim(\mathcal{Q}_{\gamma_i}^d)} \cdot \langle \Gamma_i | \mathbf{q}_{\gamma_i}(\widehat{U}^\dagger U) | \Gamma_i \rangle \right|^2. \end{aligned}$$

Upon getting measurement outcome  $\widehat{U}$ , we set  $\widehat{U}|d\rangle$  as the estimate of  $U|d\rangle$ . Thus the expectation of (squared) fidelity is

$$\begin{aligned} \mathbf{E}_{\widehat{U}}\left[\left|\langle d|\widehat{U}^\dagger U|d\rangle\right|^2\right] &= \int d\widehat{U} \left| \sum_{i=0}^L f_i \sqrt{\dim(\mathcal{Q}_{\gamma_i}^d)} \cdot \langle \Gamma_i | \mathbf{q}_{\gamma_i}(\widehat{U}^\dagger U) | \Gamma_i \rangle \right|^2 \cdot \left| \langle d|\widehat{U}^\dagger U|d\rangle \right|^2 \\ &= \int d\widehat{U} \left| \sum_{i=0}^L f_i \sqrt{\dim(\mathcal{Q}_{\gamma_i}^d)} \cdot \langle \Gamma_i | \langle d | \mathbf{q}_{\gamma_i}(\widehat{U}^\dagger U) \otimes (\widehat{U}^\dagger U) | \Gamma_i \rangle | d \rangle \right|^2 \\ &= \int d\widehat{U} \left| \sum_{i=0}^L f_i \sqrt{\dim(\mathcal{Q}_{\gamma_i}^d)} \cdot \left( \alpha_i \langle \Gamma_i^+ | \mathbf{q}_{\gamma_i^+}(\widehat{U}^\dagger U) | \Gamma_i^+ \rangle + \beta_i \langle \Gamma_{i+1}^+ | \mathbf{q}_{\gamma_{i+1}^+}(\widehat{U}^\dagger U) | \Gamma_{i+1}^+ \rangle \right) \right|^2 \quad (18) \end{aligned}$$

$$\begin{aligned} &= \int d\widehat{U} \left| f_0 \alpha_0 \sqrt{\dim(\mathcal{Q}_{\gamma_0}^d)} \langle \Gamma_0^+ | \mathbf{q}_{\gamma_0^+}(\widehat{U}^\dagger U) | \Gamma_0^+ \rangle \right. \\ &\quad \left. + \sum_{i=1}^L \left( f_i \alpha_i \sqrt{\dim(\mathcal{Q}_{\gamma_i}^d)} + f_{i-1} \beta_{i-1} \sqrt{\dim(\mathcal{Q}_{\gamma_{i-1}}^d)} \right) \langle \Gamma_i^+ | \mathbf{q}_{\gamma_i^+}(\widehat{U}^\dagger U) | \Gamma_i^+ \rangle \right|^2 \quad (19) \end{aligned}$$

$$= f_0^2 \alpha_0^2 \frac{\dim(\mathcal{Q}_{\gamma_0}^d)}{\dim(\mathcal{Q}_{\gamma_0^+}^d)} + \sum_{i=1}^L \left( f_i \alpha_i \sqrt{\frac{\dim(\mathcal{Q}_{\gamma_i}^d)}{\dim(\mathcal{Q}_{\gamma_i^+}^d)}} + f_{i-1} \beta_{i-1} \sqrt{\frac{\dim(\mathcal{Q}_{\gamma_{i-1}}^d)}{\dim(\mathcal{Q}_{\gamma_i^+}^d)}} \right)^2, \quad (20)$$

where in Equation (18) we use Equation (12), in Equation (19) we use the fact that  $\beta_L = 0$ , Equation (20) is by using the fact

$$\begin{aligned} \int d\widehat{U} q_{\gamma_i^+}(\widehat{U}^\dagger U)^\dagger |\Gamma_i^+\rangle\langle\Gamma_j^+| q_{\gamma_j^+}(\widehat{U}^\dagger U) &= \int d\widehat{U} q_{\gamma_i^+}(\widehat{U})^\dagger |\Gamma_i^+\rangle\langle\Gamma_j^+| q_{\gamma_j^+}(\widehat{U}) \\ &= \mathbb{1}_{i=j} \frac{1}{\dim(\mathcal{Q}_{\gamma_i^+}^d)} \cdot I_{\mathcal{Q}_{\gamma_i^+}^d}, \end{aligned}$$

due to the invariance of Haar measure and Schur's lemma. Then, from Fact 5.2 and Equation (20), we can see that

$$\mathbf{E}_{\widehat{U}} \left[ \left| \langle d | \widehat{U}^\dagger U | d \rangle \right|^2 \right] = f_0^2 x_0^2 + \sum_{i=1}^L (f_i x_i + f_{i-1} y_i)^2.$$

□

#### 4.1 Instantiating $f_i$

In this section, we will give an explicit form of  $f_i$  such that the expectation of fidelity (see Equation (17)) is as large as possible. For convenience, we discard the constraint  $\sum_{i=0}^L f_i^2 = 1$  and consider the homogeneous form of the RHS of Equation (17):

$$\frac{\sum_{i=0}^L (f_i x_i + f_{i-1} y_i)^2}{\sum_{i=0}^L f_i^2}, \quad (21)$$

where we set  $f_{-1} = 0$ . Therefore, we need to find a sequence of real numbers  $\{f_i\}_{i=-1}^L$  such that  $f_{-1} = 0$  and Equation (21) is as large as possible, or equivalently,  $1 - \text{Equation (21)}$  is as small as possible.

For  $-1 \leq i \leq L$ , we set

$$f_i = g_i \cdot \sqrt{(L + N + d - 2i - 1) \prod_{j=2}^{d-1} (N + j - i - 1) \prod_{j=2}^{d-1} (L + d - j - i)}, \quad (22)$$

where  $g_{-1} = 0$  and for  $0 \leq i \leq L$ ,

$$g_i = (i + 1)(L + N + d - i). \quad (23)$$

Then, we have the following results, where the proofs are based on repeatedly using Fact 5.3 and are deferred.

**Lemma 4.2.** *Using the instantiation of  $f_i$  in Equation (22), we can show that  $1 - \text{Equation (21)}$  equals*

$$\frac{\sum_{i=0}^L \frac{(g_i - g_{i-1})^2}{L+N+d-2i} \prod_{j=1}^{d-1} (N + j - i) \prod_{j=1}^{d-1} (L + j - i)}{\sum_{i=0}^L \frac{g_i^2 - g_{i-1}^2}{d-1} \prod_{j=1}^{d-1} (N + j - i) \prod_{j=1}^{d-1} (L + j - i)}. \quad (24)$$

**Lemma 4.3.** *Using the instantiation of  $g_i$  in Equation (23), we can show that Equation (24) equals*

$$\frac{d-1}{L + N + d + \frac{2}{d+1} NL}. \quad (25)$$

Then we can prove our Theorem 1.2.

*Proof of Theorem 1.2.* By Lemma 4.3 and noting that  $N = (d + 1)L$ ,  $L = \frac{n}{2d}$ , we can see that the expectation of infidelity is

$$\text{Equation (25)} = O\left(\frac{d}{L(L+d)}\right) = O\left(\frac{d^3}{n(n+d^2)}\right).$$

□

## 5 Deferred lemmas and proofs

### 5.1 Some facts

**Fact 5.1.** Suppose  $|\psi\rangle$ ,  $|\varphi\rangle$  are two pure states on a  $d$ -dimensional Hilbert space  $\mathcal{H}$ , and  $\Pi$  is an orthogonal projector onto a subspace of  $\mathcal{H}$ . Then we have

$$\left| \sqrt{\langle\psi|\Pi|\psi\rangle} - \sqrt{\langle\varphi|\Pi|\varphi\rangle} \right| \leq \frac{1}{\sqrt{2}} \|\psi\rangle\langle\psi| - |\varphi\rangle\langle\varphi|\|_1,$$

where  $\|\cdot\|_1$  is the trace norm.

*Proof.* Let  $re^{i\theta} = \langle\varphi|\psi\rangle$ . Then, we have

$$\begin{aligned} \left| \sqrt{\langle\psi|\Pi|\psi\rangle} - \sqrt{\langle\varphi|\Pi|\varphi\rangle} \right| &= \left| \|\Pi|\psi\rangle\| - \|e^{i\theta}\Pi|\varphi\rangle\| \right| \\ &\leq \left\| \Pi(|\psi\rangle - e^{i\theta}|\varphi\rangle) \right\| \\ &\leq \|\psi\rangle - e^{i\theta}|\varphi\rangle\| \\ &= \sqrt{2 - 2\Re(e^{-i\theta}\langle\varphi|\psi\rangle)} \\ &= \sqrt{2 - 2|\langle\varphi|\psi\rangle|} \\ &\leq \sqrt{2 - 2|\langle\varphi|\psi\rangle|^2} \\ &= \frac{1}{\sqrt{2}} \|\psi\rangle\langle\psi| - |\varphi\rangle\langle\varphi|\|_1, \end{aligned}$$

where  $\|\psi\rangle\| = \sqrt{\langle\psi|\psi\rangle}$  is the vector 2-norm. □

**Fact 5.2.** For  $0 \leq i \leq L$ , we have

$$\alpha_i \sqrt{\frac{\dim(\mathcal{Q}_{\gamma_i}^d)}{\dim(\mathcal{Q}_{\gamma_i^+}^d)}} = \sqrt{\frac{(N+d-i-1)(N-i+1)}{(L+N+d-2i-1)(L+N+d-2i)}},$$

and for  $1 \leq i \leq L$ , we have

$$\beta_{i-1} \sqrt{\frac{\dim(\mathcal{Q}_{\gamma_{i-1}}^d)}{\dim(\mathcal{Q}_{\gamma_i^+}^d)}} = \sqrt{\frac{(L-i+1)(L+d-i-1)}{(L+N+d-2i+1)(L+N+d-2i)}}.$$

*Proof.* By Equation (5), we have

$$\frac{\dim(\mathcal{Q}_{\gamma_i}^d)}{\dim(\mathcal{Q}_{\gamma_i^+}^d)} = \frac{(L+N+d-2i-1) \cdot \prod_{2 \leq j \leq d-1} (N+j-i-1)}{(L+N+d-2i) \cdot \prod_{2 \leq j \leq d-1} (N+j-i)}$$

$$= \frac{(L+N+d-2i-1)(N-i+1)}{(L+N+d-2i)(N+d-i-1)}.$$

$$\begin{aligned} \frac{\dim(\mathcal{Q}_{\gamma_{i-1}}^d)}{\dim(\mathcal{Q}_{\gamma_i^+}^d)} &= \frac{(L+N+d-2i+1) \prod_{2 \leq j \leq d-1} (L+d-i-j+1)}{(L+N+d-2i) \prod_{2 \leq j \leq d-1} (L+d-i-j)} \\ &= \frac{(L+N+d-2i+1)(L+d-i-1)}{(L+N+d-2i)(L-i+1)}. \end{aligned}$$

Then, combined with Equation (11), we can see

$$\alpha_i \sqrt{\frac{\dim(\mathcal{Q}_{\gamma_i}^d)}{\dim(\mathcal{Q}_{\gamma_i^+}^d)}} = \sqrt{\frac{(N+d-i-1)(N-i+1)}{(L+N+d-2i-1)(L+N+d-2i)}},$$

and

$$\beta_{i-1} \sqrt{\frac{\dim(\mathcal{Q}_{\gamma_{i-1}}^d)}{\dim(\mathcal{Q}_{\gamma_i^+}^d)}} = \sqrt{\frac{(L-i+1)(L+d-i-1)}{(L+N+d-2i+1)(L+N+d-2i)}}.$$

□

**Fact 5.3.** For any integer  $k \geq 0$  and number  $a, b$ , we have

$$(a+b+k+1) \prod_{j=1}^k (a+j)(b+j) = \frac{1}{k+1} \left( \prod_{j=1}^{k+1} (a+j)(b+j) - \prod_{j=1}^{k+1} (a-1+j)(b-1+j) \right).$$

*Proof.*

$$\begin{aligned} \prod_{j=1}^{k+1} (a+j)(b+j) - \prod_{j=1}^{k+1} (a-1+j)(b-1+j) &= \prod_{j=1}^{k+1} (a+j)(b+j) - \prod_{j=0}^k (a+j)(b+j) \\ &= ((a+k+1)(b+k+1) - ab) \prod_{j=1}^k (a+j)(b+j) \\ &= (k+1)(a+b+k+1) \prod_{j=1}^k (a+j)(b+j) \end{aligned}$$

□

## 5.2 Proof of Lemma 4.2

*Proof of Lemma 4.2.* From Equation (16) and Equation (22), and by direct calculation,

$$f_i x_i = g_i \cdot (N-i+1) \sqrt{\frac{\prod_{j=2}^{d-1} (N+j-i) \prod_{j=2}^{d-1} (L+d-j-i)}{L+N+d-2i}},$$

$$f_{i-1} y_i = g_{i-1} \cdot (L+d-i-1) \sqrt{\frac{\prod_{j=2}^{d-1} (N+j-i) \prod_{j=2}^{d-1} (L+d-j-i)}{L+N+d-2i}}.$$

Therefore, Equation (21) can be written as

$$\frac{\sum_{i=0}^L \frac{((N-i+1)g_i + (L+d-i-1)g_{i-1})^2}{L+N+d-2i} \prod_{j=2}^{d-1} (N+j-i) \prod_{j=2}^{d-1} (L+j-i)}{\sum_{i=0}^L g_i^2 (L+N+d-2i-1) \prod_{j=2}^{d-1} (N+j-i-1) \prod_{j=2}^{d-1} (L+j-i)}. \quad (26)$$

Note that the denominator of Equation (26) equals

$$\begin{aligned} & \sum_{i=0}^L g_i^2 (L+N+d-2i-1) \prod_{j=1}^{d-2} (N+j-i) \prod_{j=1}^{d-2} (L+j-i) \\ &= \sum_{i=0}^L \frac{g_i^2}{d-1} \left( \prod_{j=1}^{d-1} (N+j-i)(L+j-i) - \prod_{j=1}^{d-1} (N+j-i-1)(L+j-i-1) \right) \end{aligned} \quad (27)$$

$$\begin{aligned} &= \sum_{i=0}^L \frac{g_i^2 - g_{i-1}^2}{d-1} \prod_{j=1}^{d-1} (N+j-i)(L+j-i) - \frac{g_L^2}{d-1} \prod_{j=1}^{d-1} (N-L+j-1)(j-1) \\ &= \sum_{i=0}^L \frac{g_i^2 - g_{i-1}^2}{d-1} \prod_{j=1}^{d-1} (N+j-i)(L+j-i), \end{aligned} \quad (28)$$

where Equation (27) is by using Fact 5.3. Also note that the numerator of Equation (26) equals

$$\begin{aligned} & \sum_{i=0}^L \frac{((N-i+1)g_i + (L+d-i-1)g_{i-1})^2}{L+N+d-2i} \prod_{j=1}^{d-2} (N+j-i+1) \prod_{j=1}^{d-2} (L+j-i) \\ &= \sum_{i=0}^L \frac{((L+N+d-2i)g_i - (L+d-i-1)(g_i - g_{i-1}))^2}{L+N+d-2i} \prod_{j=1}^{d-2} (N+j-i+1) \prod_{j=1}^{d-2} (L+j-i) \\ &= \sum_{i=0}^L \left( (L+N+d-2i)g_i^2 - 2(L+d-i-1)g_i(g_i - g_{i-1}) + \frac{(L+d-i-1)^2(g_i - g_{i-1})^2}{L+N+d-2i} \right) \\ & \quad \cdot \prod_{j=1}^{d-2} (N+j-i+1) \prod_{j=1}^{d-2} (L+j-i) \end{aligned} \quad (29a)$$

$$\begin{aligned} &= \sum_{i=0}^L (L+N+d-2i)g_i^2 \prod_{j=1}^{d-2} (N+j-i+1) \prod_{j=1}^{d-2} (L+j-i) \\ &+ \sum_{i=0}^L \left( -2(L+d-i-1)g_i(g_i - g_{i-1}) + \frac{(L+d-i-1)^2(g_i - g_{i-1})^2}{L+N+d-2i} \right) \prod_{j=1}^{d-2} (N+j-i+1) \prod_{j=1}^{d-2} (L+j-i). \end{aligned} \quad (29b)$$

The part in Equation (29a) equals

$$\begin{aligned} & \sum_{i=0}^L (L+N+d-2i)g_i^2 \prod_{j=1}^{d-2} (N+j-i+1) \prod_{j=1}^{d-2} (L+j-i) \\ &= \sum_{i=0}^L \frac{g_i^2}{d-1} \left( \prod_{j=1}^{d-1} (N+j-i+1)(L+j-i) - \prod_{j=1}^{d-1} (N+j-i)(L+j-i-1) \right) \end{aligned} \quad (30)$$

$$\begin{aligned}
&= \sum_{i=0}^L \frac{g_i^2 - g_{i-1}^2}{d-1} \prod_{j=1}^{d-1} (N+j-i+1)(L+j-i) - \frac{g_L^2}{d-1} \prod_{j=1}^{d-1} (N-L+j)(j-1) \\
&= \sum_{i=0}^L \frac{g_i^2 - g_{i-1}^2}{d-1} \prod_{j=1}^{d-1} (N+j-i+1)(L+j-i),
\end{aligned} \tag{31}$$

where Equation (30) is by using Fact 5.3. Then, the denominator of Equation (26) minus the numerator of Equation (26) is therefore (28) – (31) – (29b), which equals

$$\begin{aligned}
&\left( \sum_{i=0}^L \frac{g_i^2 - g_{i-1}^2}{d-1} [(N+1-i) - (N+d-i)] \prod_{j=2}^{d-1} (N+j-i) \prod_{j=1}^{d-1} (L+j-i) \right) - (29b) \\
&= \left( \sum_{i=0}^L (g_{i-1}^2 - g_i^2) \prod_{j=2}^{d-1} (N+j-i) \prod_{j=1}^{d-1} (L+j-i) \right) - (29b) \\
&= \sum_{i=0}^L \left( (g_{i-1}^2 - g_i^2)(L+d-i-1) + 2(L+d-i-1)g_i(g_i - g_{i-1}) - \frac{(L+d-i-1)^2(g_i - g_{i-1})^2}{L+N+d-2i} \right) \\
&\quad \cdot \prod_{j=2}^{d-1} (N+j-i) \prod_{j=1}^{d-2} (L+j-i) \\
&= \sum_{i=0}^L \frac{(L+d-i-1)(N-i+1)(g_i - g_{i-1})^2}{L+N+d-2i} \prod_{j=2}^{d-1} (N+j-i) \prod_{j=1}^{d-2} (L+j-i) \\
&= \sum_{i=0}^L \frac{(g_i - g_{i-1})^2}{L+N+d-2i} \prod_{j=1}^{d-1} (N+j-i) \prod_{j=1}^{d-1} (L+j-i).
\end{aligned} \tag{32}$$

Then,  $1 - \text{Equation (26)}$  is therefore (32)  $\div$  (28), which equals

$$\frac{\sum_{i=0}^L \frac{(g_i - g_{i-1})^2}{L+N+d-2i} \prod_{j=1}^{d-1} (N+j-i) \prod_{j=1}^{d-1} (L+j-i)}{\sum_{i=0}^L \frac{g_i^2 - g_{i-1}^2}{d-1} \prod_{j=1}^{d-1} (N+j-i) \prod_{j=1}^{d-1} (L+j-i)}.$$

□

### 5.3 Proof of Lemma 4.3

*Proof of Lemma 4.3.* It is easy to see that for  $0 \leq i \leq L$

$$g_i - g_{i-1} = L + N + d - 2i.$$

Thus, Equation (24) equals

$$\frac{(d-1) \sum_{i=0}^L (L+N+d-2i) \prod_{j=1}^{d-1} (N+j-i) \prod_{j=1}^{d-1} (L+j-i)}{\sum_{i=0}^L (g_i + g_{i-1})(L+N+d-2i) \prod_{j=1}^{d-1} (N+j-i) \prod_{j=1}^{d-1} (L+j-i)}. \tag{33}$$

The numerator of Equation (33) is

$$(d-1) \sum_{i=0}^L (L+N+d-2i) \prod_{j=1}^{d-1} (N+j-i) \prod_{j=1}^{d-1} (L+j-i)$$

$$= (d-1) \sum_{i=0}^L \frac{1}{d} \left( \prod_{j=1}^d (N+j-i)(L+j-i) - \prod_{j=1}^d (N+j-i-1)(L+j-i-1) \right) \quad (34)$$

$$\begin{aligned} &= \frac{(d-1)}{d} \left( \prod_{j=1}^d (N+j)(L+j) - \prod_{j=1}^d (N-L+j-1)(j-1) \right) \\ &= \frac{(d-1)}{d} \prod_{j=1}^d (N+j)(L+j), \end{aligned} \quad (35)$$

where Equation (34) is by using Fact 5.3. The denominator of Equation (33) is

$$\begin{aligned} &\sum_{i=0}^L (g_i + g_{i-1})(L+N+d-2i) \prod_{j=1}^{d-1} (N+j-i) \prod_{j=1}^{d-1} (L+j-i) \\ &= \sum_{i=0}^L \frac{g_i + g_{i-1}}{d} \left( \prod_{j=1}^d (N+j-i)(L+j-i) - \prod_{j=1}^d (N+j-i-1)(L+j-i-1) \right) \\ &= \frac{g_0}{d} \prod_{j=1}^d (N+j)(L+j) - \frac{g_L + g_{L-1}}{d} \prod_{j=1}^d (N-L+j-1)(j-1) + \sum_{i=1}^L \frac{g_i - g_{i-2}}{d} \prod_{j=1}^d (N+j-i)(L+j-i) \\ &= \frac{g_0}{d} \prod_{j=1}^d (N+j)(L+j) + \sum_{i=1}^L \frac{g_i - g_{i-2}}{d} \prod_{j=1}^d (N+j-i)(L+j-i) \\ &= \frac{g_0}{d} \prod_{j=1}^d (N+j)(L+j) + \frac{2}{d} \sum_{i=1}^L (L+N+d-2i+1) \prod_{j=1}^d (N+j-i)(L+j-i) \\ &= \frac{g_0}{d} \prod_{j=1}^d (N+j)(L+j) + \frac{2}{d(d+1)} \sum_{i=1}^L \left( \prod_{j=1}^{d+1} (N+j-i)(L+j-i) - \prod_{j=1}^{d+1} (N+j-i-1)(L+j-i-1) \right) \\ &= \frac{g_0}{d} \prod_{j=1}^d (N+j)(L+j) + \frac{2}{d(d+1)} \left( \prod_{j=1}^{d+1} (N+j-1)(L+j-1) - \prod_{j=1}^{d+1} (N-L+j-1)(j-1) \right) \\ &= \frac{g_0}{d} \prod_{j=1}^d (N+j)(L+j) + \frac{2}{d(d+1)} \prod_{j=1}^{d+1} (N+j-1)(L+j-1), \end{aligned} \quad (37)$$

where Equation (36) and Equation (37) are by using Fact 5.3. Therefore, Equation (33) equals (35)  $\div$  (38), which is

$$\frac{d-1}{g_0 + \frac{2}{d+1} NL} = \frac{d-1}{L+N+d+\frac{2}{d+1} NL}.$$

□

## Acknowledgments

We thank Ewin Tang, John Wright, and Masahito Hayashi for valuable discussions.

## References

- [BCH06] Dave Bacon, Isaac L Chuang, and Aram W Harrow. Efficient quantum circuits for schur and clebsch-gordan transforms. *Physical review letters*, 97(17):170502, 2006.
- [BCH07] Dave Bacon, Isaac L Chuang, and Aram W Harrow. The quantum schur and clebsch-gordan transforms: I. efficient qudit circuits. In *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 1235–1244, 2007.
- [Bel75a] Viacheslav P Belavkin. Optimal multiple quantum statistical hypothesis testing. *Stochastics: an international journal of probability and stochastic processes*, 1(1-4):315–345, 1975.
- [Bel75b] VP Belavkin. Optimum distinction of non-orthogonal quantum signals. *Radiotekhnika i Elektronika*, 20:1177–1185, 1975.
- [BHMT02] Gilles Brassard, Peter Hoyer, Michele Mosca, and Alain Tapp. Quantum amplitude amplification and estimation. *Contemporary Mathematics*, 2002.
- [BK02] Howard Barnum and Emanuel Knill. Reversing quantum dynamics with near-optimal quantum and classical fidelity. *Journal of Mathematical Physics*, 43(5):2097–2106, 2002.
- [BM99] Dagmar Bruß and Chiara Macchiavello. Optimal state estimation for d-dimensional quantum systems. *Physics Letters A*, 253(5-6):249–251, 1999.
- [CGJ19] Shantanav Chakraborty, András Gilyén, and Stacey Jeffery. The Power of Block-Encoded Matrix Powers: Improved Regression Techniques via Faster Hamiltonian Simulation. In *46th International Colloquium on Automata, Languages, and Programming (ICALP 2019)*, volume 132, pages 33:1–33:14, 2019.
- [CML<sup>+</sup>24] Yu-Ao Chen, Yin Mo, Yingjian Liu, Lei Zhang, and Xin Wang. Quantum algorithm for reversing unknown unitary evolutions. *arXiv preprint arXiv:2403.04704*, 2024.
- [CSM23] Jordan Cotler, Thomas Schuster, and Masoud Mohseni. Information-theoretic hardness of out-of-time-order correlators. *Physical Review A*, 108(6):062608, 2023.
- [CYZ25] Kean Chen, Nengkun Yu, and Zhicheng Zhang. Tight bound for quantum unitary time-reversal. *arXiv preprint arXiv:2507.05736*, 2025.
- [EGH<sup>+</sup>11] Pavel I Etingof, Oleg Golberg, Sebastian Hensel, Tiansai Liu, Alex Schwendner, Dmitry Vaintrob, and Elena Yudovina. *Introduction to representation theory*, volume 59. American Mathematical Soc., 2011.
- [FH13] William Fulton and Joe Harris. *Representation Theory: A First Course*, volume 129 of *Graduate Texts in Mathematics*. Springer, 2013.
- [FK18] Bill Fefferman and Shelby Kimmel. Quantum vs. Classical Proofs and Subset Verification. In *43rd International Symposium on Mathematical Foundations of Computer Science (MFCS 2018)*, volume 117, pages 22:1–22:23, 2018.
- [GBO24] Dmitry Grinko, Adam Burchardt, and Maris Ozols. Gelfand-tsetlin basis for partially transposed permutations, with applications to quantum information. In *27th Conference on Quantum Information Processing*, 2024.

- [GKKT20] Madalin Guță, 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. Query-optimal estimation of unitary channels via pauli dimensionality. *arXiv preprint arXiv:2510.00168*, 2025.
- [GSLW19] András Gilyén, Yuan Su, Guang Hao Low, and Nathan Wiebe. Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics. In *Proceedings of the 51st annual ACM SIGACT symposium on theory of computing*, pages 193–204, 2019.
- [GW09] Roe Goodman and Nolan R Wallach. *Symmetry, representations, and invariants*, volume 255. Springer, 2009.
- [Har05] Aram W Harrow. Applications of coherent classical communication and the schur transform to quantum information theory. *arXiv preprint quant-ph/0512255*, 2005.
- [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. *Physics Letters A*, 354(3):183–189, 2006.
- [Hay11] Masahito Hayashi. Comparison between the cramer-rao and the mini-max approaches in quantum channel estimation. *Communications in mathematical physics*, 304(3):689–709, 2011.
- [Hay16] Masahito Hayashi. Fourier analytic approach to quantum estimation of group action. *Communications in Mathematical Physics*, 347(1):3–82, 2016.
- [HHJ<sup>+</sup>16] Jeongwan Haah, Aram W Harrow, Zhengfeng Ji, Xiaodi Wu, and Nengkun Yu. Sample-optimal tomography of quantum states. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pages 913–925, 2016.
- [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.
- [HW94] Paul Hausladen and William K Wootters. A ‘pretty good’ measurement for distinguishing quantum states. *Journal of Modern Optics*, 41(12):2385–2390, 1994.
- [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.
- [Key06] Michael Keyl. Quantum state estimation and large deviations. *Reviews in Mathematical Physics*, 18(01):19–60, 2006.
- [KW99] Michael Keyl and Reinhard F Werner. Optimal cloning of pure states, testing single clones. *Journal of Mathematical Physics*, 40(7):3283–3299, 1999.
- [LC19] Guang Hao Low and Isaac L Chuang. Hamiltonian simulation by qubitization. *Quantum*, 3:163, 2019.

- [Ngu24] Quynh T Nguyen. The mixed schur transform: efficient quantum circuit and applications. In *27th Conference on Quantum Information Processing*, 2024.
- [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.
- [OYM25] Tatsuki Odake, Satoshi Yoshida, and Mio Murao. Analytical lower bound on query complexity for transformations of unknown unitary operations. *Physical Review Letters*, 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.
- [SHH25] Thomas Schuster, Jonas Haferkamp, and Hsin-Yuan Huang. Random unitaries in extremely low depth. *Science*, 389(6755):92–96, 2025.
- [SSW25] Thilo Scharnhorst, Jack Spilecki, and John Wright. Optimal lower bounds for quantum state tomography. *arXiv preprint arXiv:2510.07699*, 2025.
- [TW25] Ewin Tang and John Wright. Amplitude amplification and estimation require inverses. *arXiv preprint arXiv:2507.23787*, 2025.
- [vACGN23] Joran van Apeldoorn, Arjan Cornelissen, András Gilyén, and Giacomo Nannicini. Quantum tomography using state-preparation unitaries. In *Proceedings of the 2023 annual ACM-SIAM symposium on discrete algorithms (SODA)*, pages 1265–1318. SIAM, 2023.
- [vAG19] Joran van Apeldoorn and András Gilyén. Improvements in Quantum SDP-Solving with Applications. In *46th International Colloquium on Automata, Languages, and Programming (ICALP 2019)*, volume 132, pages 99:1–99:15, 2019.
- [VK92] N. Ja. Vilenkin and A. U. Klimyk. Representations in the gel’fand-tsetlin basis and special functions. In *Representation of Lie Groups and Special Functions: Volume 3: Classical and Quantum Groups and Special Functions*. Springer Netherlands, 1992.
- [Wer98] Reinhard F Werner. Optimal cloning of pure states. *Physical Review A*, 58(3):1827, 1998.
- [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.
- [YYM25] Satoshi Yoshida, Hironobu Yoshida, and Mio Murao. Asymptotically optimal unitary estimation in  $SU(3)$  by the analysis of graph laplacian. *arXiv preprint arXiv:2509.20608*, 2025.