Approximation does not help in quantum unitary time-reversal

Kean Chen* Nengkun Yu† Zhicheng Zhang‡

Abstract

Access to the time-reverse U^{-1} of an unknown quantum unitary process U is widely assumed in quantum learning, metrology, and many-body physics. The fundamental task of unitary time-reversal dictates implementing U^{-1} to within diamond-norm error ϵ using black-box queries to the d-dimensional unitary U. Although the query complexity of this task has been extensively studied, existing lower bounds either hold only for the exact case (i.e., $\epsilon = 0$) or are suboptimal in d. This raises a central question: does approximation help reduce the query complexity of unitary time-reversal? We settle this question in the negative by establishing a *robust* and *tight* lower bound $\Omega((1-\epsilon)d^2)$ with explicit dependence on the error ϵ. This implies that unitary time-reversal retains optimal exponential hardness (in the number of qubits) even when constant error is allowed. Our bound applies to adaptive and coherent algorithms with unbounded ancillas and holds even when ϵ is an average-case distance error.

1 Introduction

The time evolution of a closed quantum system is governed by a reversible unitary U. Time-reversal of this evolution can, in principle, be simulated by implementing the inverse unitary U^{-1}. However, this requires complete knowledge of the quantum system, which is typically inaccessible when the dynamics is dictated by nature. A straightforward solution is to perform unitary tomography to obtain an approximate description of the unknown d-dimensional unitary U, which enables approximate unitary time-reversal with error ϵ, using $O(d^2/\epsilon)$ queries [HKOT23] to U.

In this paper, we study the *unitary time-reversal* task: given black-box query access to an unknown d-dimensional unitary U, approximately implementing U^{-1} to within diamond-norm error ϵ. Beyond its root as a natural concept in physics, this task is fundamental to quantum information theory and has a deep connection to the power of out-of-time-order correlators (OTOCs) [CSM23, SNC+23, SS14, YGS+16, SBSSH16, VES+19, XS24] and advanced quantum learning algorithms with time-reversed unitary access [vACGN23, WZC+22, GP22, MW16, SHH25, Zha25b, TW25a]. In quantum cryptography, there is also a growing attention [MH25, Zha25a, SML+25] to security against an adversary that can query time-reversed unitary oracles.

Quantum algorithms for unitary time-reversal have been extensively studied in the literature [SCHL16, SBZ19, QDS+19a, QDS+19b, EHM+23, QE22, Nav18, GO24, TDN20, TDN23, SST+23, YH21, MZC+25, ZMCW24, MLW25, YSM23, CML+24, GYMO25, BYMQ25, ZCJ+25]. Surprisingly, Yoshida, Soeda, and Murao [YSM23] demonstrated that, when the dimension $d = 2$, the unitary time-reversal can be done *deterministically* and *exactly*[1] using only four queries to the

*University of Pennsylvania, Philadelphia, USA. Email: keanchen.gan@gmail.com
†Stony Brook University, NY, USA. Email: nengkunyu@gmail.com
‡University of Technology Sydney, Sydney, Australia. Email: iszczhang@gmail.com

[1]Exactness and determinism mean that the algorithm has no error and has success probability 1.

unknown unitary. Thereafter, Chen, Mo, Liu, Zhang, and Wang [CML+24] significantly generalized this result by providing a deterministic and exact time-reversal algorithm for unitaries of *any dimension* d, using $O(d^2)$ queries. In a complementary breakthrough, Odake, Yoshida, and Murao [OYM25] showed that any deterministic and exact (i.e., $\epsilon = 0$) unitary time-reversal algorithm must use at least $\Omega(d^2)$ queries, matching the upper bound in [CML+24].

However, the prior best lower bound $\Omega(d^2)$ [OYM25] holds only for the exact case. A remaining gap is whether allowing approximation with a nonzero error $\epsilon > 0$ could yield any improvement over the $O(d^2)$ query complexity. In other words, the following question is still open:

Question 1.1. *Does approximation help reduce the query complexity of unitary time-reversal?*

1.1 Our results

In this paper, we provide a negative answer to the above question by showing a *robust* and *tight* query lower bound $\Omega((1-\epsilon)d^2)$, demonstrating that unitary time-reversal retains optimal exponential hardness (in the number of qubits) even when constant error is allowed. Our lower bound, combined with the upper bound $O(d^2)$ by the exact algorithm given in [CML+24], fully settles the query complexity of the unitary time-reversal task with approximation. Formally, our main result is as follows:

Theorem 1.2 (Theorem 3.1 and Corollary 3.2 restated)**.** *Given query access to an unknown d-dimensional unitary U, any algorithm that approximates the time-reversed unitary U^{-1} to within diamond norm or average-case distance[2] error ϵ, must use at least $\Omega((1-\epsilon)d^2)$ queries.*

One interesting aspect of our result is the error scaling $(1-\epsilon)$, which fundamentally differs from the typical scaling $\mathrm{poly}(1/\epsilon)$ in quantum learning or metrology tasks. This error scaling is natural as the unitary time-reversal can be done exactly (i.e., $\epsilon = 0$). Thus, Theorem 1.2 provides a *robust and optimal* hardness guarantee: permitting approximation even with constant errors cannot lead to better efficiency. In comparison, most prior works focus on the lower bounds only for the *exact* case, which has recently been settled by Odake, Yoshida, and Murao [OYM25] through a differentiation-based SDP framework (see further comparison in Section 1.3.1 and Table 1). We note that their method depends on the exact differentiation of the transformation map, and therefore cannot provide hardness guarantee in the approximate (non-exact) regime: it cannot rule out the possibility that an approximate unitary time-reversal algorithm could achieve better efficiency.

We note that our lower bound is strong in various senses. First, it holds not only for diamond norm error ϵ (i.e., worst-case guarantee), but also for an average-case distance error ϵ (i.e., average-case guarantee over Haar random input states), a stronger robustness notion in many settings. Then, it applies to the general adaptive and coherent algorithms with unbounded ancillas. Using the techniques [Kit95, SMM09, TW25b] of unitary controlization, it further extends to unitary time-reversal algorithms that have access to controlled queries. Furthermore, we note that our bounds match the numerical results obtained in [GYMO25] (see Theorem 3.7 and the discussion thereafter), suggesting that our query lower bound is likely optimal even non-asymptotically.

Lower bound for generalized time-reversal. We can provide a lower bound for the *generalized unitary time-reversal* task: approximately implementing the time-reverse $U^{-t} := e^{-iHt}$ for a given reversing time $t > 0$. Here, H is allowed to be any Hamiltonian satisfying $e^{iH} = e^{i\theta}U$ for some real number θ and $\|H\| \leq \pi$, where $\|\cdot\|$ denotes the operator norm, the largest singular value. Using Theorem 1.2, we can provide a robust and tight query lower bound $\Omega(d^2)$ for this task with

[2]The average-case distance is given in Definition 2.2.

constant error and constant reversing time, which matches the upper bound $O(d^2)$ by unitary tomography [HKOT23].

Corollary 1.3 (Corollary 3.3 restated)**.** *Given query access to an unknown d-dimensional unitary U, any algorithm that approximates the generalized time-reversed unitary U^{-t} to within constant diamond norm error $\epsilon \leq 10^{-5}$ for constant reversing time $t \geq 0.1$, must use at least $\Omega(d^2)$ queries.*

Here, the ranges $t \geq 0.1$ and $\epsilon \leq 10^{-5}$ are rather loose and chosen for simplicity of proof.

Lower bound for low-depth unitaries. Our main result, when combined with the low-depth construction of pseudorandom unitaries [MH25, SHH25, SML+25], can give a super-polynomial lower bound for the time complexity of unitary time-reversal even when the target unitary is restricted to depth $\mathrm{poly}(\log\log\log d)$[3], where $\log d$ represents the number of qubits. This strengthens the prior lower bound by [SHH25] that requires larger depth $\mathrm{poly}(\log\log d)$. Formally,

Corollary 1.4. *Under cryptographic assumptions, no algorithm can solve the unitary time-reversal task in polynomial time (with respect to the number of qubits), even only for* $\mathrm{poly}(\log\log\log d)$*-depth unitaries and with constant average-case distance error.*

Proof. The idea follows from noting that the Haar expectation in Theorem 3.6 can be simulated using $\mathrm{poly}(\log\log\log d)$-depth pseudorandom unitaries [SHH25, Corollary 2]. Specifically, we assume, to the contrary, that an algorithm R performs unitary time-reversal for these low-depth pseudorandom unitaries in polynomial time. Then, the quantum channel shown in Figure 3 must be close to the identity channel. On the other hand, our Theorem 3.6 shows that, for Haar random unitaries, the quantum channel shown in Figure 3 must be close to the completely depolarizing channel. This is because Theorem 3.6 in fact shows that the maximal eigenvalue of the Choi representation of the channel in Figure 3 is no more than $(n+1)/d$ and n is polynomial in $\log d$ by assumption. This means the algorithm R can be used to distinguish the pseudorandom and Haar random unitaries, by distinguishing whether the channel in Figure 3 is close to identity or depolarizing channel to some constant precision, which only incurs a constant overhead (e.g., inputting $|0\rangle$ and then measuring $\{|0\rangle\langle 0|, I - |0\rangle\langle 0|\}$). This contradicts the indistinguishability between pseudorandom unitaries and Haar random unitaries under cryptographic assumptions [MH25]. □

Moreover, when combined with the low-depth construction of unitary t-designs [SML+25, Theorem 1], our result establishes a lower bound of $\Omega(t)$ for the unitary time-reversal of $\widetilde{O}(t)$[4]-depth unitaries (for any $t \leq O(d^2)$) under the average-case distance. Recalling that an arbitrary $\log d$-qubit unitary can be implemented with depth $O(d^2)$ [MVBS04], our lower bound $\Omega(t)$ is optimal up to polylogarithmic factors. This is because the range $t \leq O(d^2)$ of our lower bound saturates the upper bound of the depth of $\log d$-qubit unitaries, and general $O(d^2)$-depth unitaries can be learned with $\Theta(d^2)$ queries. This result can be viewed as a *graded* version of our main lower bound. Notably, these graded bounds are purely information-theoretic and do not rely on the cryptographic assumptions required by pseudorandom unitaries.

1.2 Overview of techniques

Here, we briefly outline the main idea for the proof of Theorem 1.2. Suppose R is a unitary time-reversal algorithm with error bounded by ϵ that uses n queries to the unknown unitary U. Then, R

[3] A t-depth unitary is a unitary that can be implemented by a t-depth circuit using single- and two-qubit gates

[4] $\widetilde{\Omega}(\cdot)$ hides polylogarithmic factors.

can be described by an $(n+1)$-comb [CDP08, CDP09] (see Section 2.3 for a brief introduction) on $(\mathcal{H}_0, \mathcal{H}_1, \ldots, \mathcal{H}_{2n+1})$ where each $\mathcal{H}_i$ is a copy of $\mathbb{C}^d$. We make an additional query to U at the end of R, so that the overall channel approximates the identity channel $\mathcal{I}_{\mathcal{H}_0 \to \mathcal{H}_{2n+2}}$ from $\mathcal{H}_0$ to $\mathcal{H}_{2n+2}$ (see Figure 3). Here, the $n+1$ queries to U in total can be also described by an $(n+1)$-comb:

$$C_U \coloneqq |U\rangle\!\rangle\langle\!\langle U|^{\otimes n+1},$$

where $|U\rangle\!\rangle$ denotes the vector obtained by flattening the matrix U. Consequently, the overall channel is represented by the 1-comb $R \star C_U \overset{\epsilon}{\approx} |I_{\mathcal{H}_0 \to \mathcal{H}_{2n+2}}\rangle\!\rangle\langle\!\langle I_{\mathcal{H}_0 \to \mathcal{H}_{2n+2}}|$, where $\star$ denotes the link product [CDP09] for quantum combs, $|I_{\mathcal{H}_0 \to \mathcal{H}_{2n+2}}\rangle\!\rangle\langle\!\langle I_{\mathcal{H}_0 \to \mathcal{H}_{2n+2}}|$ is the Choi representation of the identity channel $\mathcal{I}_{\mathcal{H}_0 \to \mathcal{H}_{2n+2}}$, and $\overset{\epsilon}{\approx}$ denotes approximation with error bounded by ϵ. Taking Haar expectation over $U \in \mathbb{U}_d$ and using the bilinearity of $\star$, we can see that:

$$R \star \mathop{\mathbf{E}}_U[C_U] \overset{\epsilon}{\approx} |I_{\mathcal{H}_0 \to \mathcal{H}_{2n+2}}\rangle\!\rangle\langle\!\langle I_{\mathcal{H}_0 \to \mathcal{H}_{2n+2}}|. \tag{1}$$

Here, $\mathbf{E}_U[C_U]$ is called the $(n+1)$-th moment of the Haar random unitary under Choi representation (also called performance operator [QE22] for the time-reversal task). Then, we use the *stair operators* $\{A_k\}_{k=1}^n$ (see Definition 3.8) as a tool to analyze the Haar moment $\mathbf{E}_U[C_U]$. The stair operators are defined in the Young-Yamanouchi basis [CSST10, Har05] under the Schur-Weyl duality of an asymmetric bipartite quantum system. We prove two lemmas (see Lemma 3.9 and Lemma 3.10) for the stair operators. The first lemma shows that they provide a good upper bound (w.r.t. the Löwner order) for the Haar moment:

$$\mathop{\mathbf{E}}_U[C_U] \sqsubseteq I_{\mathcal{H}_{2n+1}} \otimes A_n, \tag{2}$$

where $I_{\mathcal{H}_{2n+1}}$ is the identity operator on $\mathcal{H}_{2n+1}$. The second lemma indicates that, the stair operators are compatible with quantum combs through the link product $\star$, i.e., when linked with an arbitrary $(n+1)$-comb R, they can be sequentially contracted and bounded, until we obtain the identity operator (with a coefficient) on $\mathcal{H}_{2n+2} \otimes \mathcal{H}_0$, as shown in Corollary 3.11:

$$R \star (I_{\mathcal{H}_{2n+1}} \otimes A_n) \sqsubseteq \frac{n+1}{d} I_{\mathcal{H}_{2n+2}} \otimes I_{\mathcal{H}_0}. \tag{3}$$

Combining Equation (2) and Equation (3), we can show that

$$R \star \mathop{\mathbf{E}}_U[C_U] \sqsubseteq R \star (I_{\mathcal{H}_{2n+1}} \otimes A_n) \sqsubseteq \frac{n+1}{d} I_{\mathcal{H}_{2n+2}} \otimes I_{\mathcal{H}_0}, \tag{4}$$

where we also use the property that $\star$ preserves the Löwner order. Intuitively, this means if n is small and U is Haar random, the overall channel $R \star \mathbf{E}_U[C_U]$ is highly depolarizing, for any algorithm R. Therefore, combining Equation (4) with Equation (1), we can see that n must be $\Omega(d^2)$ even for a constant non-zero error ϵ, since the maximum eigenvalue of $|I_{\mathcal{H}_0 \to \mathcal{H}_{2n+2}}\rangle\!\rangle\langle\!\langle I_{\mathcal{H}_0 \to \mathcal{H}_{2n+2}}|$ is d. As our proof is based on the Choi representation, it is natural to expect that our lower bound applies even for the average-case distance error, according to Equation (7).

To prove these two lemmas, we perform a representation-theoretic analysis of the stair operators. Specifically, we exploit the raising and lowering techniques of Young diagrams based on the branching rule of symmetric group representations to show Equation (2) and Equation (3), respectively. Since the subsystem to be traced out is not the "last" subsystem (w.r.t. the ordering fixed by the action of symmetric groups) in the stair operators, the lowering techniques are not directly applicable here. To address this, we use the Young's orthogonal form [CSST10] for adjacent transpositions to swap the last two subsystems and then analyze the resulting expressions. By this, we can reduce the original problems to pure combinatorics on Young diagrams (see Section 4.2), which is then solved with the assistance of Kerov's interlacing sequences [Ker93, Ker00].

Algorithm Types		Lower Bounds	
Exact ($\epsilon = 0$)	Deterministic	$\Omega(d^2)$	[OYM25]
	Probabilistic†	$\Omega(d)$	[QDS+19a, QDS+19b]
Ancilla-Free		$\Omega(d^{1/4})$	[CSM23]
Clean‡		$\Omega(d)$	[GST24]
General		$\Omega(\sqrt{d})$	[FK18]
		$\Omega(d)$	[YKS+26]
		$\Omega(d^2)$	Our Theorem 1.2

Table 1: Comparison of query lower bounds for unitary time-reversal. In the approximate (non-exact) settings, the lower bounds hold for constant error $\epsilon > 0$. †Probabilistic exact algorithms refer to algorithms that are exact conditioned on a flag output bit being "success". ‡Clean means the ancilla output must be returned in $|0\rangle$ assuming the algorithm is only unitary with postselection.

1.3 Related work

In what follows, we review previous lower bound results for the unitary time-reversal task.

1.3.1 Lower bounds from semidefinite programming

Quintino, Dong, Shimbo, Soeda, and Murao [QDS+19a, QDS+19b] developed a systematic semidefinite programming (SDP) approach to find probabilistic exact algorithms (i.e., exact conditioned on a flag output bit being "success") for unitary transformation tasks such as transposition, complex conjugation and time-reversal. Query lower bounds are established for such algorithms; in particular, $\Omega(d)$ queries are required for probabilistic exact unitary time-reversal. This approach is extended by Yoshida, Koizumi, Studziński, Quintino, and Murao [YKS+26] to derive the same $\Omega(d)$ query lower bound but applicable to approximate unitary time-reversal. Odake, Yoshida, and Murao [OYM25] further developed a general framework for deriving lower bounds on deterministic exact unitary transformation tasks. Specifically, they consider transformations described by differentiable functions $f : \mathbb{U}_d \to \mathbb{U}_d$ and analyze the induced Lie algebra mappings obtained by differentiation at various points. Within the Lie algebra framework, they formulate an SDP such that the number of queries required to implement $f(\cdot)$ is lower bounded by the SDP solution. Consequently, they show an $\Omega(d^2)$ query lower bound for deterministic exact unitary time-reversal (i.e., $f(U) = U^{-1}$), matching the upper bound $O(d^2)$ by the algorithm in [CML+24].

However, the differentiation-based method in [OYM25] depends on the exact form of the transformation map and appears challenging to extend directly to the analysis of approximate (non-exact) algorithms, so that their lower bound $\Omega(d^2)$ applies only to the *exact* case. This leaves open the possibility that approximate unitary time-reversal algorithms may achieve better efficiency and motivates our central Question 1.1.

1.3.2 Lower bounds from topological obstructions

Gavorová, Seidel, and Touati [GST24] established lower bounds for unitary transformation tasks from a topological perspective. By proving any m-homogeneous function from $\mathbb{U}_d$ to the 1-sphere

$\mathcal{S}^1$ must satisfy $d \mid m$ (i.e., d divides m), they show that any *clean* algorithm (i.e., the ancilla output must be returned in $|0\rangle$ assuming the algorithm is only unitary with postselection) for unitary time-reversal task requires $\Omega(d)$ queries. Note that the cleanness requirement is a strong constraint, since uncomputation of ancillas is generally hard when the algorithm has access only to the time-forward U.

1.3.3 Lower bounds from quantum learning

Cotler, Schuster, and Mohseni [CSM23] designed a quantum learning task such that when ancillas are not allowed: $\Omega(d^{1/4})$ queries to U are required to solve it; and only $O(1)$ queries are sufficient given additional access to the time-reverse U^{-1}. This separation implies an $\Omega(d^{1/4})$ query lower bound for the unitary time-reversal task, but it applies only to algorithms without ancillas [SHH25]. Schuster, Haferkamp, and Huang [SHH25] later established a stronger separation in time complexity without this limitation. Their new learning task leverages the low depth construction of pseudo-random unitaries [SHH25, MH25] and the connectivity features of quantum dynamics [SNC+23]. Specifically, this task cannot be solved in polynomial time; but it can be solved in $O(1)$ time given additional access to the time-reverse U^{-1}. This implies a super-polynomial lower bound $\omega(\operatorname{poly}\log d)$ for the time complexity of unitary time-reversal.

1.3.4 Lower bounds from quantum adversary method

Fefferman and Kimmel [FK18] investigated the task of finding the pre-image of an unknown in-place permutation oracle. Specifically, an in-place permutation oracle O_π is a unitary that maps $|i\rangle$ to $|\pi(i)\rangle$, where $i \in \{1, \ldots, d\}$ and $\pi \in \mathfrak{S}_d$ is an unknown permutation on d elements. The task is to find the index i such that $\pi(i) = 1$. They proved that at least $\Omega(\sqrt{d})$ queries to the in-place permutation oracle O_π are required to solve this task, using the techniques from the quantum adversary method [Amb00]. On the other hand, we can directly find the index if we are able to apply $(O_\pi)^{-1}$ on $|1\rangle$. Thus, their result implies an $\Omega(\sqrt{d})$ lower bound for the unitary time-reversal.

1.3.5 Lower bounds from compressed oracle method.

Tang and Wright [TW25a] showed a query lower bound $\Omega(\min\{d, 1/\delta^2\})$ for amplitude estimation without access to U^{-1}, revealing the importance of the time-reverse unitary in the Grover-type quadratic speedup. Here, d is the dimension of the unknown unitary U and δ is the required precision. Specifically, they reduced the problem of distinguishing two diagonal unitary ensembles to the amplitude estimation problem. The hardness of the former problem was then proved using the compressed oracle method [Zha19], which purifies the randomness from the unitary ensembles and yields a finer lower bound given access only to U. Combined with the well-known upper bound $O(1/\delta)$, their result implies a lower bound of $\Omega(\min\{\sqrt{d}, 1/\sqrt{\epsilon}\})$ for unitary time-reversal to within diamond norm error ϵ.[5] We note that when ϵ is a constant, this lower bound becomes $\Omega(1)$.

1.4 Discussion and further implication

In this paper, we settle the complexity of the unitary time-reversal, a fundamental task in quantum information processing. We prove a robust and tight query lower bound $\Omega((1-\epsilon)d^2)$ for this task.

[5] This can be seen by the following argument. Suppose $\epsilon \le \delta^2$ and the time-reversal to within diamond norm error ϵ can be done using n queries. By the original amplitude estimation algorithm, we can use $O(n/\delta)$ queries to U to obtain an estimate of the amplitude with error $O(\delta + \epsilon/\delta) = O(\delta)$. Therefore, the lower bound due to [TW25a] implies $n = \Omega(\min\{d\delta, 1/\delta\})$. If $\epsilon \le 1/d$, we set $\delta = 1/\sqrt{d}$ and then $n = \Omega(\sqrt{d})$. Otherwise if $\epsilon > 1/d$, we set $\delta = \sqrt{\epsilon} \ge 1/\sqrt{d}$ and then $n = \Omega(1/\sqrt{\epsilon})$.

This answers the central Question 1.1 in the negative: approximation does not help in the unitary time-reversal. Here, we further discuss the implication of our results.

An interesting question is how to develop a general framework for the robust hardness of quantum query transformations, like the SDP-based framework established by [OYM25] for the exact cases. Beyond unitary time-reversal, other two canonical query transformations are the unitary conjugation $U \mapsto U^*$ and unitary transposition $U \mapsto U^{\mathrm{T}}$. For the unitary conjugation, it is shown that $O(d)$ queries suffice [MSM19, EHM+23] and $\Omega(d)$ queries are necessary even when allowing approximation [EHM+23]. For unitary transposition, our Theorem 1.2 implies the *state-of-the-art* result in the following Corollary 1.5, extending the exact-case lower bound in [OYM25].

Corollary 1.5. *Given query access to an unknown d-dimensional unitary U, any algorithm that approximates the transpose U^{T} to within diamond norm or average-case distance error ϵ, must use at least $\Omega((1-\epsilon)d)$ queries.*

This is by noting that any m-query unitary transposition algorithm combined with the $O(d)$-query exact unitary conjugation algorithm [MSM19, EHM+23] yields an $O(md)$-query unitary time-reversal algorithm with the same approximation error. However, there remains a gap to the best known upper bound $O(d^2)$ for unitary transposition by [CML+24].

1.5 Organization

In Section 2, we review the notation and preliminaries used in this paper. In Section 3, we present the details of our main results, where Theorem 1.2 is proved through Sections 3.1 to 3.4, and Corollary 1.3 is proved in Section 3.5. Technical lemmas and their proofs are deferred to Section 4.

2 Preliminaries

2.1 Notation

We use $\mathcal{L}(\mathcal{H})$ to denote the set of linear operators on the Hilbert space $\mathcal{H}$. Similarly, we use $\mathcal{L}(\mathcal{H}_0, \mathcal{H}_1)$ to denote 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 each 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 facts:

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

where $|\phi^*\rangle$ is the entry-wise complex conjugate of $|\phi\rangle$ w.r.t. to a given orthonormal basis, and Z^{T} is the transpose of the matrix Z. The inner product can be denoted by $\langle\!\langle X|Y\rangle\!\rangle = \mathrm{tr}(X^\dagger Y)$.

The *Choi-Jamiołkowski operator* of a quantum channel $\mathcal{E} : \mathcal{L}(\mathcal{H}_0) \to \mathcal{L}(\mathcal{H}_1)$ is defined by:

$$\mathfrak{C}(\mathcal{E}) = (\mathcal{E} \otimes \mathcal{I}_{\mathcal{H}_0})(|I\rangle\!\rangle\langle\!\langle I|),$$

where $\mathcal{I}_{\mathcal{H}_0} : \mathcal{L}(\mathcal{H}_0) \to \mathcal{L}(\mathcal{H}_0)$ is the identity channel and $|I\rangle\!\rangle = \sum_i |i\rangle|i\rangle$ is the (unnormalized) maximally entangled state. Then, the application of the channel $\mathcal{E}$ can be described by its Choi-Jamiołkowski operator:

$$\mathcal{E}(X) = \mathrm{tr}_{\mathcal{H}_0}\left(\mathfrak{C}(\mathcal{E})^{\mathrm{T}_{\mathcal{H}_0}} \cdot (I_{\mathcal{H}_1} \otimes X)\right),$$

where $\mathrm{tr}_{\mathcal{H}_0}$ and $\mathrm{T}_{\mathcal{H}_0}$ denote the partial trace and partial transpose on the subsystem $\mathcal{H}_0$, respectively.

For a sequence of Hilbert spaces $\mathcal{H}_0, \ldots, \mathcal{H}_n$ indexed by consecutive integers, we use $\mathcal{H}_{i:j}$ to denote the Hilbert space $\mathcal{H}_i \otimes \cdots \otimes \mathcal{H}_j$. For two linear operators X, Y, we use $X \sqsubseteq Y$ to denote that $Y - X$ is positive semidefinite.

2.2 Distance between quantum channels

Worst-case distance. First, we introduce the diamond norm, which servers as the worst-case distance between quantum channels.

Definition 2.1 (Diamond norm [AKN98])**.** *The diamond norm of two quantum channels $\mathcal{E}_1, \mathcal{E}_2$ is defined as*

$$\|\mathcal{E}_1 - \mathcal{E}_2\|_\diamond := \sup_\rho \|(\mathcal{E}_1 \otimes \mathcal{I})(\rho) - (\mathcal{E}_2 \otimes \mathcal{I})(\rho)\|_1,$$

where $\|\cdot\|_1$ denotes the Schatten 1-norm (trace norm).

There is a simple relation between the diamond norm of quantum channels and trace norm of the corresponding Choi-Jamiołkowski operators. Specifically, let $\mathcal{E}_1, \mathcal{E}_2 : \mathcal{L}(\mathbb{C}^d) \to \mathcal{L}(\mathbb{C}^d)$ be two d-dimensional quantum channels. Then, by the definition of diamond norm, we have

$$\begin{aligned}\|\mathcal{E}_1 - \mathcal{E}_2\|_\diamond &\geq \left\|(\mathcal{E}_1 \otimes \mathcal{I})\left(\frac{|I\rangle\!\rangle\langle\!\langle I|}{d}\right) - (\mathcal{E}_2 \otimes \mathcal{I})\left(\frac{|I\rangle\!\rangle\langle\!\langle I|}{d}\right)\right\|_1 \\ &= \frac{1}{d}\|\mathfrak{C}(\mathcal{E}_1) - \mathfrak{C}(\mathcal{E}_2)\|_1, \end{aligned} \tag{5}$$

where $|I\rangle\!\rangle\langle\!\langle I|/d$ is a maximally entangled state.

Average-case distance. Then, we introduce the average-case distance between quantum channels, which is widely used in quantum learning [HLB+24, VH25], benchmarking [MGE12, MGE11] and metrology [CWLY23, YRC20]. For this, we first need the notion of fidelity. Recall that the fidelity of two quantum states ρ, σ is defined as

$$F(\rho, \sigma) := \mathrm{tr}\left(\sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}}\right)^2.$$

Note that when $\sigma = |\psi\rangle\langle\psi|$ is a pure state, we have $F(\rho, \sigma) = \mathrm{tr}(\rho\sigma) = \langle\psi|\rho|\psi\rangle$. Then, the definition of average-case distance is as follows:

Definition 2.2 (Average-case distance [HLB+24, VH25])**.** *The average-case distance between two quantum channels $\mathcal{E}_1, \mathcal{E}_2$ is defined as*

$$\mathcal{D}_{\mathrm{avg}}(\mathcal{E}_1, \mathcal{E}_2) := \mathop{\mathbf{E}}_{|\psi\rangle}[1 - F(\mathcal{E}_1(|\psi\rangle\langle\psi|), \mathcal{E}_2(|\psi\rangle\langle\psi|))],$$

where $\mathbf{E}_{|\psi\rangle}$ denotes the expectation over the Haar random state $|\psi\rangle$.

A closely related measure is the *channel fidelity* [Rag01] (or *entanglement fidelity* [Nie02]). Specifically, for d-dimensional quantum channels $\mathcal{E}_1$ and $\mathcal{E}_2$, the channel fidelity F_{ch} of them is defined as the fidelity of their normalized Choi operators:

$$F_{\mathrm{ch}}(\mathcal{E}_1, \mathcal{E}_2) := F\left(\frac{\mathfrak{C}(\mathcal{E}_1)}{d}, \frac{\mathfrak{C}(\mathcal{E}_2)}{d}\right).$$

This measure is also commonly used in higher-order quantum transformation literature (see, e.g., [TMM+25]).

Now, we introduce some properties of these measures. First, it is easy to see that both the average-case distance and the channel fidelity are unitarily invariant, i.e.,

$$\begin{aligned}\mathcal{D}_{\mathrm{avg}}(\mathcal{U} \circ \mathcal{E}_1 \circ \mathcal{V}, \mathcal{U} \circ \mathcal{E}_2 \circ \mathcal{V}) &= \mathcal{D}_{\mathrm{avg}}(\mathcal{E}_1, \mathcal{E}_2), \\ F_{\mathrm{ch}}(\mathcal{U} \circ \mathcal{E}_1 \circ \mathcal{V}, \mathcal{U} \circ \mathcal{E}_2 \circ \mathcal{V}) &= F_{\mathrm{ch}}(\mathcal{E}_1, \mathcal{E}_2), \end{aligned} \tag{6}$$

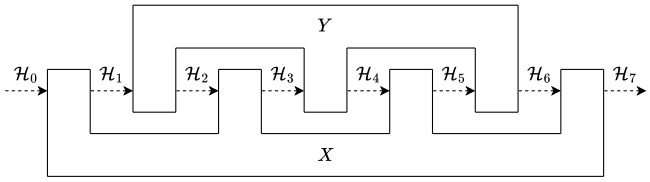


Figure 1: The combination of a 4-comb X with a 3-comb Y, resulting in a 1-comb $X \star Y$ on $(\mathcal{H}_0, \mathcal{H}_7)$.

for any unitary channels $\mathcal{U}, \mathcal{V}$. Second, the average-case distance is closely related to the channel fidelity when one of $\mathcal{E}_1$ and $\mathcal{E}_2$ is a unitary channel [Nie02]:

$$\mathcal{D}_{\text{avg}}(\mathcal{E},\mathcal{U}) = \frac{d}{d+1}\left(1 - F_{\text{ch}}(\mathcal{E},\mathcal{U})\right) = \frac{d}{d+1}\left(1 - \frac{1}{d^2}\operatorname{tr}(\mathfrak{C}(\mathcal{E}) \cdot |U\rangle\!\rangle\langle\!\langle U|)\right), \tag{7}$$

where d is the dimension of $\mathcal{U}$ and $\mathcal{E}$. Then, it is easy to see that

$$\mathcal{D}_{\text{avg}}(\mathcal{E},\mathcal{U}) \leq \frac{d}{d+1}\left\|\frac{\mathfrak{C}(\mathcal{E})}{d} - \frac{|U\rangle\!\rangle\langle\!\langle U|}{d}\right\|_1 \leq \frac{d}{d+1}\|\mathcal{E}-\mathcal{U}\|_\diamond, \tag{8}$$

where the first inequality is due to the well-known property [FVDG02]: $1 - \sqrt{F(\rho,\sigma)} \leq \frac{1}{2}\|\rho-\sigma\|_1$ and $F(\rho,\sigma) \leq 1$, the second inequality is by Equation (5).

We remark that the average-case distance can be much smaller than the diamond norm. For example, let $U = (|0\rangle\langle 1| + |1\rangle\langle 0|) \oplus I_{d-2}$ be a d-dimensional unitary, where I_{d-2} is the identity operator of dimension $d-2$. In such case, we have $\mathcal{D}_{\text{avg}}(\mathcal{U},\mathcal{I}) = O(1/d)$ but $\|\mathcal{U}-\mathcal{I}\|_\diamond = \Omega(1)$. This implies that hardness results with average-case distance are generally much stronger than those with diamond norm.

2.3 Quantum combs

In this section, we introduce the quantum comb [CDP08, CDP09], which is a powerful tool to describe (higher) transformations of quantum processes. Specifically, the Choi-Jamiołkowski representation of quantum channels (i.e., transformations of quantum states) can be generalized to a higher-level concept (i.e., transformations of quantum processes), which is called *quantum comb*.

Definition 2.3 (Quantum comb [CDP09])**.** *For an integer $n \geq 1$, a quantum 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 $X^{(n)}, X^{(n-1)}, \ldots, X^{(1)}, X^{(0)}$ such that*

$$\operatorname{tr}_{\mathcal{H}_{2j-1}}\left(X^{(j)}\right) = I_{\mathcal{H}_{2j-2}} \otimes X^{(j-1)}, \quad 1 \leq j \leq n, \tag{9}$$

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

Note that in Definition 2.3, the operators $X^{(j)}$ are again quantum j-combs (for convenience, we define any quantum 0-comb to be the scalar 1). It is also worth noting that a quantum 1-comb is simply the Choi-Jamiołkowski operator of a quantum channel.

The combination of any two quantum combs is defined by the link product "$\star$":

Definition 2.4 (Link product "$\star$" [CDP09])**.** *Suppose X is 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 Y is 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 similar for $\boldsymbol{j} = (j_1, \ldots, j_m)$. Let $\boldsymbol{a} = \boldsymbol{i} \cap \boldsymbol{j}$ be the set of indices in both $\boldsymbol{i}$ and $\boldsymbol{j}$ and $\boldsymbol{b} = \boldsymbol{i} \cup \boldsymbol{j}$ be the set of indices in either $\boldsymbol{i}$ or $\boldsymbol{j}$. Then, the combination of X and Y is defined by*

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

where $\mathcal{H}_{\boldsymbol{a}}$ means the tensor product of subsystems labeled by the indices in $\boldsymbol{a}$, $\mathrm{T}_{\mathcal{H}_{\boldsymbol{a}}}$ means 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 has a very good property [CDP09, Theorem 2]:

$$X, Y \sqsupseteq 0 \Longrightarrow X \star Y \sqsupseteq 0. \tag{10}$$

Moreover, it characterizes the channel concatenation under the Choi representation: 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 $\mathfrak{C}(\mathcal{E}_2 \circ \mathcal{E}_1) = \mathfrak{C}(\mathcal{E}_2) \star \mathfrak{C}(\mathcal{E}_1)$. The link product is also used to describe the combination of two higher-order combs such as that shown in Figure 1. More generally, 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

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

turns out to be a 1-comb on $(\mathcal{H}_0, \mathcal{H}_{2n-1})$. To see this, first note that X, Y are positive semidefinite, which means $X \star Y$ is also positive semidefinite. On the other hand, note that

$$\begin{aligned}
\mathrm{tr}_{\mathcal{H}_{2n-1}}(X \star Y) &= \mathrm{tr}_{\mathcal{H}_{1:2n-2}}\left(\mathrm{tr}_{\mathcal{H}_{2n-1}}(X)^{\mathrm{T}_{\mathcal{H}_{1:2n-2}}} \cdot (Y \otimes I_{\mathcal{H}_0})\right) \\
&= \mathrm{tr}_{\mathcal{H}_{1:2n-2}}\left((I_{\mathcal{H}_{2n-2}} \otimes X^{(n-1)})^{\mathrm{T}_{\mathcal{H}_{1:2n-2}}} \cdot (Y \otimes I_{\mathcal{H}_0})\right) && (11) \\
&= \mathrm{tr}_{\mathcal{H}_{1:2n-3}}\left((X^{(n-1)})^{\mathrm{T}_{\mathcal{H}_{1:2n-3}}} \cdot (\mathrm{tr}_{\mathcal{H}_{2n-2}}(Y) \otimes I_{\mathcal{H}_0})\right) \\
&= \mathrm{tr}_{\mathcal{H}_{1:2n-3}}\left((X^{(n-1)})^{\mathrm{T}_{\mathcal{H}_{1:2n-3}}} \cdot (I_{\mathcal{H}_{2n-3}} \otimes Y^{(n-2)} \otimes I_{\mathcal{H}_0})\right) && (12) \\
&= \mathrm{tr}_{\mathcal{H}_{1:2n-3}}\left(((X^{(n-1)})^{\mathrm{T}_{\mathcal{H}_{1:2n-4}}})^{\mathrm{T}_{\mathcal{H}_{2n-3}}} \cdot (I_{\mathcal{H}_{2n-3}} \otimes Y^{(n-2)} \otimes I_{\mathcal{H}_0})\right) \\
&= \mathrm{tr}_{\mathcal{H}_{1:2n-3}}\left((X^{(n-1)})^{\mathrm{T}_{\mathcal{H}_{1:2n-4}}} \cdot (I_{\mathcal{H}_{2n-3}} \otimes Y^{(n-2)} \otimes I_{\mathcal{H}_0})\right) && (13) \\
&= \mathrm{tr}_{\mathcal{H}_{2n-3}}\left(X^{(n-1)} \star Y^{(n-2)}\right),
\end{aligned}$$

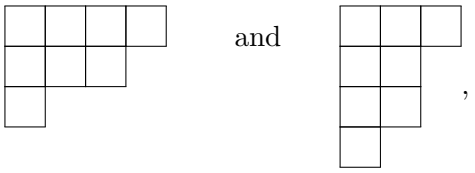
where in Equation (11), $X^{(n-1)}$ is an $(n-1)$-comb obtained from taking partial trace on X (see Definition 2.3) and similarly in Equation (12), $Y^{(n-2)}$ is an $(n-2)$-comb obtained from Y, Equation (13) is due to the property of partial trace $\mathrm{tr}_{\mathcal{H}}(X^{\mathrm{T}_{\mathcal{H}}}) = \mathrm{tr}_{\mathcal{H}}(X)$. Then, we can iterate this process and finally obtain

$$\mathrm{tr}_{\mathcal{H}_{2n-1}}(X \star Y) = \mathrm{tr}_{\mathcal{H}_1}\left(X^{(1)}\right) = I_{\mathcal{H}_0},$$

which means $X \star Y$ is a 1-comb on $(\mathcal{H}_0, \mathcal{H}_{2n-1})$.

2.4 Young diagrams

Young diagrams and standard Young tableaux. A *Young diagram* λ consisting of n boxes and ℓ rows is a partition $(\lambda_1, \ldots, \lambda_\ell)$ of n such that $\sum_{i=1}^{\ell} \lambda_i = n$ and $\lambda_1 \geq \cdots \geq \lambda_\ell > 0$. The i-th rows consists of λ_i boxes. By convention, the Young diagram is drawn with left-justified rows, arranged from top to bottom. The *conjugate of a Young diagram* is obtained by exchanging rows and columns. For example, the Young diagram $\lambda = (4, 3, 1)$ and its conjugate $\lambda' = (3, 2, 2, 1)$ are drawn as:



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

A *standard Young tableau* of shape $\lambda \vdash n$ is obtained by a filling of the n boxes of λ using the integers $1, \ldots, n$, each appearing exactly once, such that the entries increase from left to right in each row and from top to bottom in each column. An example of standard Young tableau of shape $(4, 3, 1)$ is

1	3	4	6
2	7	8	
5			

.

Without causing confusion, we refer to the box in T containing the integer i as the i-box. We use $\mathrm{Sh}(T)$ to denote the shape (Young diagram) of a standard Young tableau T. We use $\mathrm{Tab}(\lambda)$ to denote the set of all standard Young tableaux T such that $\mathrm{Sh}(T) = \lambda$ and use $\mathrm{Tab}(n)$ to denote $\bigcup_{\lambda \vdash n} \mathrm{Tab}(\lambda)$. Given a Young diagram $\lambda \vdash n$, the number of standard Young tableaux of shape λ (i.e., $|\mathrm{Tab}(\lambda)|$) is characterized by the *hook length formula* [Ful97]:

$$|\mathrm{Tab}(\lambda)| = \frac{n!}{\prod_{\square \in \lambda} h_\lambda(\square)}, \tag{14}$$

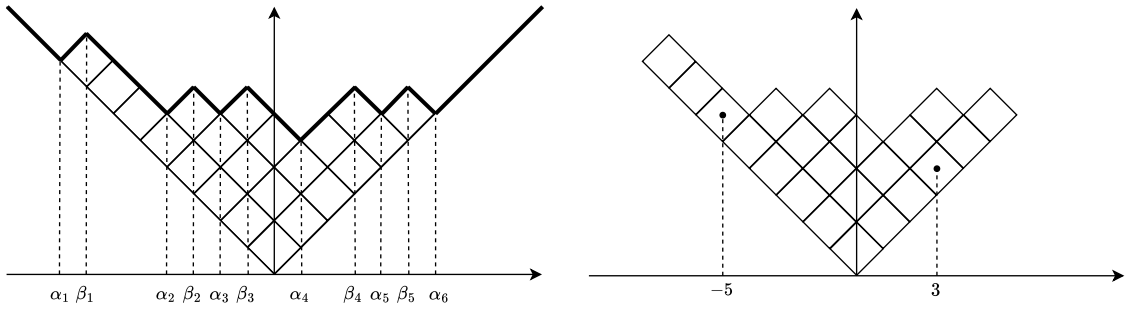
where $h_\lambda(\square)$ is the number of boxes that are directly to the right and below the box $\square$, including the box itself (i.e., the "hook length").

Adjacency relation. For $\lambda \vdash n+1$, $\mu \vdash n$, we write $\mu \nearrow \lambda$ or $\lambda \searrow \mu$ if λ can be obtained from μ by adding one box. For a standard Young tableau T, we use $T^{\downarrow}$ to denote the standard Young tableau obtained from T by removing the box containing the largest integer in T. Suppose $\mu \nearrow \lambda$ and $T \in \mathrm{Tab}(\mu)$. We use $T^{\uparrow \lambda}$ to denote the standard Young tableau obtained from T by adding a box containing number $n+1$, resulting in shape λ. We define $\mathrm{Tab}(\lambda, \mu) \subseteq \mathrm{Tab}(\lambda)$ to be the set of all standard Young tableaux T such that $\mathrm{Sh}(T) = \lambda$ and $\mathrm{Sh}(T^{\downarrow}) = \mu$. It is easy to see that:

$$|\mathrm{Tab}(\lambda, \mu)| = |\mathrm{Tab}(\mu)|.$$

Let $\mu, \nu \vdash n$ and $\mu \neq \nu$. We call μ and ν are *adjacent* to each other if one of the following equivalent conditions is true:

- There exists a $\lambda \vdash n+1$ such that $\mu \nearrow \lambda$ and $\nu \nearrow \lambda$.
- There exists a $\tau \vdash n-1$ such that $\tau \nearrow \mu$ and $\tau \nearrow \nu$.



(a) The interlacing sequences of λ.

(b) Axial coordinates of two boxes in λ.

Figure 2: Rotated version of the Young diagram $\lambda = (6, 5, 3, 3, 2, 1, 1, 1)$.

Such μ and ν can be obtained from each other by moving a removable box. Furthermore, $\lambda = \mu \cup \nu$ and $\tau = \mu \cap \nu$ are unique, where $\mu \cup \nu$ denotes the Young diagram consisting of the union of boxes in μ and in ν, and $\mu \cap \nu$ denotes the Young diagram consisting of the intersection of boxes in μ and in ν. Moreover, we use $\lambda \setminus \mu$ to denote the set of boxes (not a Young diagram) in λ but not in μ.

Kerov's interlacing sequences. Kerov's interlacing sequences [Ker93, Ker00] serve as an important tool in studying the combinatorics of Young diagram and their analytic generalizations. Suppose $\lambda \vdash n$ is a Young diagram. The rotated version of λ is defined by rotating the Young diagram λ' (the conjugate of λ) 135° counterclockwise. We define the x-axis so that the center of each box lies at an integer x-axis coordinate. Consider the shape of the rotated Young diagram extended with the linear curves $y = x$ and $y = -x$ (e.g., the bold line in Figure 2a), which is a piecewise linear function. Then, Kerov's interlacing sequences are defined as the sequence x-coordinates of the local minima $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_L)$ and the sequence of x-coordinates of the local maxima $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_{L-1})$. The two sequences interlace

$$\alpha_1 < \beta_1 < \alpha_2 < \cdots < \alpha_{L-1} < \beta_{L-1} < \alpha_L.$$

Figure 2a shows an example of the Young diagram $(6, 5, 3, 3, 2, 1, 1, 1)$ and its interlacing sequences are $\boldsymbol{\alpha} = (-8, -4, -2, 1, 4, 6)$ and $\boldsymbol{\beta} = (-7, -3, -1, 3, 5)$.

We define the *axial coordinate*[6] $c(\square)$ for a box $\square$ in a Young diagram as the x-axis coordinate of the center of this box when drawing in the rotated version. Figure 2b shows an example of the Young diagram $(6, 5, 3, 3, 2, 1, 1, 1)$ and the axial coordinates of two boxes. The *axial distance* from box $\square_1$ to box $\square_2$ is defined as the difference $c(\square_1) - c(\square_2)$.

Adding or removing a box from a Young diagram can be described naturally using Kerov's interlacing sequences. A position where a box can be added corresponds precisely to a local minima α_i. Conversely, a removable box corresponds to a local maxima β_i. Specifically, the axial coordinate of the added (or removed) box $\square$ (i.e., $c(\square)$) coincides with the α_i (or β_i). Suppose $\mu \vdash n$ has interlacing sequences $\alpha_1, \ldots, \alpha_L$ and $\beta_1, \ldots, \beta_{L-1}$. Let $\lambda \vdash n + 1$ be obtained from μ by adding a box at α_k. Then, we have [Ker93, Eq. (3.2.3)]

$$\frac{|\mathrm{Tab}(\lambda)|}{|\mathrm{Tab}(\mu)|} = (n+1) \prod_{i=1}^{k-1} \frac{\alpha_k - \beta_i}{\alpha_k - \alpha_i} \prod_{i=k+1}^{L} \frac{\alpha_k - \beta_{i-1}}{\alpha_k - \alpha_i}. \tag{15}$$

[6]The axial coordinate is also called *content* in the literature [CSST10, OV96]

On the other hand, suppose $\lambda \vdash n$ has interlacing sequences $\alpha_1, \ldots, \alpha_L$ and $\beta_1, \ldots, \beta_{L-1}$. Let $\mu \vdash n-1$ be obtained from λ by removing a box at β_k. Then, we have [Ker00, Lemma 3.3]

$$\frac{|\mathrm{Tab}(\mu)|}{|\mathrm{Tab}(\lambda)|} = \frac{(\alpha_L - \beta_k)(\beta_k - \alpha_1)}{n} \prod_{i=1}^{k-1} \frac{\beta_k - \alpha_{i+1}}{\beta_k - \beta_i} \prod_{i=k+1}^{L-1} \frac{\beta_k - \alpha_i}{\beta_k - \beta_i}. \tag{16}$$

2.5 Schur-Weyl duality

Let $\mathcal{H}_1, \mathcal{H}_2, \ldots, \mathcal{H}_n$ be a sequence of Hilbert spaces such that $\mathcal{H}_i \cong \mathbb{C}^d$ for $1 \le i \le n$. Consider the Hilbert space $\bigotimes_{i=1}^n \mathcal{H}_i$. This space 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 $U \in \mathbb{U}_d$ and the symmetric group acts by permuting tensor factors:

$$P(\pi)|\psi_1\rangle \cdots |\psi_n\rangle = |\psi_{\pi^{-1}(1)}\rangle \cdots |\psi_{\pi^{-1}(n)}\rangle, \tag{17}$$

where $\pi \in \mathfrak{S}_n$.

Remark 2.5. *In this paper, when we say the symmetric group $\mathfrak{S}_n$ acts on $\mathcal{H}_{i_1} \otimes \mathcal{H}_{i_2} \otimes \cdots \otimes \mathcal{H}_{i_n}$, the order of tensor products matters instead of the indices i_j of the Hilbert spaces. For example, if $\mathfrak{S}_3$ acts on $\mathcal{H}_1 \otimes \mathcal{H}_3 \otimes \mathcal{H}_2$ and let $|\psi_1\rangle|\psi_3\rangle|\psi_2\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_3 \otimes \mathcal{H}_2$, then we have*

$$P(\pi_{(13)})|\psi_1\rangle|\psi_3\rangle|\psi_2\rangle = |\psi_2\rangle|\psi_3\rangle|\psi_1\rangle,$$

where $\pi_{(13)} \in \mathfrak{S}_3$ is the permutation that swaps integers 1 and 3.

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

$$\bigotimes_{i=1}^n \mathcal{H}_i \overset{\mathbb{U}_d \times \mathfrak{S}_n}{\cong} \bigoplus_{\lambda \vdash_d n} \mathcal{Q}_\lambda \otimes \mathcal{P}_\lambda, \tag{18}$$

where $\mathcal{Q}_\lambda$ and $\mathcal{P}_\lambda$ are irreducible representations of $\mathbb{U}_d$ and $\mathfrak{S}_n$ labeled by Young diagram λ, respectively.

Remark 2.6. *In general, the representation $\mathcal{Q}_\lambda$ depends on d and is sometimes denoted by $\mathcal{Q}_\lambda^d$. However, since the dimension d is fixed throughout this work, we can omit it.*

Remark 2.7. *We take the convention that $\mathcal{Q}_\lambda = 0$ (the trivial space) if $\ell(\lambda) > d$. Therefore, the summation in Equation (18) might be simplified to $\bigoplus_{\lambda \vdash n} \mathcal{Q}_\lambda \otimes \mathcal{P}_\lambda$.*

Now, suppose we have two sequences of Hilbert spaces $(\mathcal{H}_1^{\mathrm{A}}, \ldots, \mathcal{H}_n^{\mathrm{A}})$ and $(\mathcal{H}_1^{\mathrm{B}}, \ldots, \mathcal{H}_m^{\mathrm{B}})$, where $\mathcal{H}_i^{\mathrm{A}} \cong \mathcal{H}_j^{\mathrm{B}} \cong \mathbb{C}^d$. We define the action of group $\mathfrak{S}_n \times \mathfrak{S}_m$ on $\bigotimes_{i=1}^n \mathcal{H}_i^{\mathrm{A}} \otimes \bigotimes_{j=1}^m \mathcal{H}_j^{\mathrm{B}}$ as $P(\pi^{\mathrm{A}}) \otimes P(\pi^{\mathrm{B}})$ for $(\pi^{\mathrm{A}}, \pi^{\mathrm{B}}) \in \mathfrak{S}_n \times \mathfrak{S}_m$. Similarly, we define the action of $\mathbb{U}_d \times \mathbb{U}_d$ on $\bigotimes_{i=1}^n \mathcal{H}_i^{\mathrm{A}} \otimes \bigotimes_{j=1}^m \mathcal{H}_j^{\mathrm{B}}$ as $(U^{\mathrm{A}})^{\otimes n} \otimes (U^{\mathrm{B}})^{\otimes m}$ for $(U^{\mathrm{A}}, U^{\mathrm{B}}) \in \mathbb{U}_d \times \mathbb{U}_d$. Note that the action of $\mathfrak{S}_n \times \mathfrak{S}_m$ commutes with the action of $\mathbb{U}_d \times \mathbb{U}_d$. Therefore, we have a generalized Schur-Weyl duality on this bipartite system as

$$\bigotimes_{i=1}^n \mathcal{H}_i^{\mathrm{A}} \otimes \bigotimes_{j=1}^m \mathcal{H}_j^{\mathrm{B}} \overset{\mathbb{U}_d \times \mathbb{U}_d \times \mathfrak{S}_n \times \mathfrak{S}_m}{\cong} \bigoplus_{\substack{\lambda \vdash_d n \\ \mu \vdash_d m}} \mathcal{Q}_\lambda \otimes \mathcal{Q}_\mu \otimes \mathcal{P}_\lambda \otimes \mathcal{P}_\mu,$$

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

2.6 Young-Yamanouchi basis

Branching rule of $\mathfrak{S}_n$. We now look deeper into the structure of $\mathcal{P}_\lambda$. Consider the following chain of groups:

$$\mathfrak{S}_1 \subset \mathfrak{S}_2 \subset \cdots \subset \mathfrak{S}_n,$$

where the embedding $\mathfrak{S}_i \to \mathfrak{S}_{i+1}$ is by identifying each $\pi \in \mathfrak{S}_i$ as the permutation in $\mathfrak{S}_{i+1}$ that fixes the element $i+1$. Note that the irreducible representations of $\mathfrak{S}_i$ are in one-to-one correspondence to the Young diagrams consisting of i boxes. Let $\lambda \vdash i$ be a Young diagram of i boxes and $\mathcal{P}_\lambda$ be an irreducible representation of $\mathfrak{S}_i$. Then, $\mathcal{P}_\lambda$ is also a representation of $\mathfrak{S}_{i-1}$, which decomposes as a direct sum of irreducible representations of $\mathfrak{S}_{i-1}$. More specifically, we have the following branching rule [CSST10, Mél17]:

$$\mathrm{Res}^{\mathfrak{S}_i}_{\mathfrak{S}_{i-1}} \mathcal{P}_\lambda = \bigoplus_{\mu:\mu\nearrow\lambda} \mathcal{P}_\mu, \tag{19}$$

where $\mathrm{Res}^{\mathfrak{S}_i}_{\mathfrak{S}_{i-1}}$ denotes the restriction of representation from $\mathfrak{S}_i$ to $\mathfrak{S}_{i-1}$, $\mu \nearrow \lambda$ means that λ can be obtained from μ by adding one box. Equation (19) means that we can decompose the vector space $\mathcal{P}_\lambda$ into direct sum of smaller vector spaces. By iterating this process (i.e., taking further restrictions on $\mathcal{P}_\mu$), we finally obtain a direct sum of the trivial one-dimensional representations of $\mathfrak{S}_1$. Therefore, this process determines a basis (up to scalar factors) of $\mathcal{P}_\lambda$ and each basis vector can be parameterized by its branching path:

$$\lambda^{(1)} \to \lambda^{(2)} \to \cdots \to \lambda^{(n)}, \tag{20}$$

where $\lambda^{(n)} = \lambda$, $\lambda^{(i)} \nearrow \lambda^{(i+1)}$, and $\lambda^{(1)}$ is the one-box Young diagram. This basis is called *Young-Yamanouchi basis* or *Gelfand-Tsetlin basis* for $\mathfrak{S}_n$ [OV96]. For convenience, we will simply call it Young basis in this paper.

Furthermore, we can see that the path in Equation (20) uniquely corresponds to a standard Young tableau T of shape $\lambda = \lambda^{(n)}$, where the integer i is assigned to the box $\lambda^{(i)} \setminus \lambda^{(i-1)}$, in which $\lambda^{(i)} \setminus \lambda^{(i-1)}$ denotes the shape by removing from the shape of $\lambda^{(i)}$ all the boxes belonging to $\lambda^{(i-1)}$. By this mean, we may also denote the standard Young tableau as

$$T = \lambda^{(1)} \to \lambda^{(2)} \to \cdots \to \lambda^{(n)}.$$

By choosing an appropriate inner product making $\mathcal{P}_\lambda$ a unitary representation of $\mathfrak{S}_n$ and then normalizing the Young basis w.r.t. the inner product, we obtain an orthonormal basis $\{|T\rangle\}_{T\in\mathrm{Tab}(\lambda)}$ of $\mathcal{P}_\lambda$ labeled by standard Young tableaux $T \in \mathrm{Tab}(\lambda)$. This also implies that $\dim(\mathcal{P}_\lambda) = |\mathrm{Tab}(\lambda)|$.

Isotypic component projectors. Consider the group algebra $\mathbb{C}\mathfrak{S}_n$, which is the associative algebra containing formal linear combinations of the elements in $\mathfrak{S}_n$ with coefficients in $\mathbb{C}$. Note that any representation of $\mathfrak{S}_n$ is naturally a representation of $\mathbb{C}\mathfrak{S}_n$ by linearly extending the group action of $\mathfrak{S}_n$. For any $1 \le k \le n$ and $\lambda \vdash k$, we define the following element

$$e_\lambda \coloneqq \frac{\dim(\mathcal{P}_\lambda)}{k!} \sum_{\pi\in\mathfrak{S}_k} \chi^*_\lambda(\pi)\pi \in \mathbb{C}\mathfrak{S}_k \subseteq \mathbb{C}\mathfrak{S}_n, \tag{21}$$

where $\chi_\lambda(\cdot)$ is the character of $\mathcal{P}_\lambda$. By the orthogonality of characters (see e.g., [EGH+11, Sec. 4.5]), e_λ acts on $\mathcal{P}_\lambda$ as the identity and acts on $\mathcal{P}_\mu$ as the null map for $\mu \vdash k, \mu \neq \lambda$, thus is the projector onto the isotypic component of λ (in fact, for any unitary representation of $\mathfrak{S}_k$, e_λ acts

as an orthogonal projector). Therefore, for any standard Young tableau $T = \lambda^{(1)} \to \cdots \to \lambda^{(n)}$, we have

$$e_{\lambda^{(k)}}|T\rangle = |T\rangle, \tag{22}$$

for any $1 \le k \le n$. Moreover, $|T\rangle$ is the only vector (up to a scalar) satisfying Equation (22) for any $1 \le k \le n$. This means, under the algebra isomorphism $\mathbb{C}\mathfrak{S}_n \cong \bigoplus_{\mu \vdash n} \mathcal{L}(\mathcal{P}_\mu)$ (due to the semisimplicity of $\mathbb{C}\mathfrak{S}_n$), we can write

$$|T\rangle\langle T| = \prod_{k=1}^{n} e_{\lambda^{(k)}}, \tag{23}$$

where $|T\rangle\langle T|$ is considered as a linear operator in $\bigoplus_{\mu \vdash n} \mathcal{L}(\mathcal{P}_\mu)$.

Young's orthogonal form. For $1 \le i \le n-1$, let $s_i \in \mathfrak{S}_n$ be the i-th adjacent transposition (i.e., s_i swaps integers i and $i+1$). It is known that those s_i generate the group $\mathfrak{S}_n$ and are called Coxeter generators. The Young's orthogonal form [CSST10] gives an explicit description of the action of s_i on the Young basis:

$$s_i|T\rangle = \frac{1}{r(T)}|T\rangle + \sqrt{1 - \frac{1}{r(T)^2}}|s_i T\rangle, \tag{24}$$

where $r(T)$ is the axial distance from the $(i+1)$-box to the i-box in T, $s_i T$ is the Young tableau obtained from T by swapping the i-box and the $(i+1)$-box.[7]

3 Hardness of unitary time-reversal

In this section, we will prove our lower bounds for quantum unitary time-reversal. First, we provide a tight lower bound for time-reversal with the average-case distance error (see Definition 2.2).

Theorem 3.1. *Suppose there is an algorithm that approximately implements U^{-1} to within the average-case distance error ϵ, using n queries to the unknown d-dimensional unitary U. Then, it must satisfy $n \ge d(d+1)(1-\epsilon) - (d+1)$.*

The proof is deferred to Section 3.1. Note that this directly implies a tight lower bound for time-reversal with diamond norm error.

Corollary 3.2. *If we change the average-case distance in Theorem 3.1 to diamond norm, then we have $n \ge d^2(1-\epsilon) - 1$.*

Proof. By Equation (8), we know that any ϵ-approximate algorithm in diamond norm is an $(\frac{d}{d+1}\epsilon)$-approximate algorithm in the average-case distance. Therefore, the lower bound in Theorem 3.1 implies a lower bound $d^2(1-\epsilon) - 1$ for any diamond norm algorithm. □

From Corollary 3.2, we can further obtain a tight lower bound for the generalized unitary time-reversal task. Specifically, the task is to approximate the time-reverse unitary $U^{-t} \coloneqq e^{-iHt}$ given a reversing time $t > 0$, where H is allowed to be any specific Hamiltonian that satisfies $e^{iH} = e^{i\theta}U$ for some real number θ, and $\|H\| \le \pi$, in which $\|\cdot\|$ is the operator norm.

[7] Note that $s_i T$ may not be a standard Young tableau. Nevertheless, in such case $\sqrt{1 - 1/r(T)^2} = 0$.

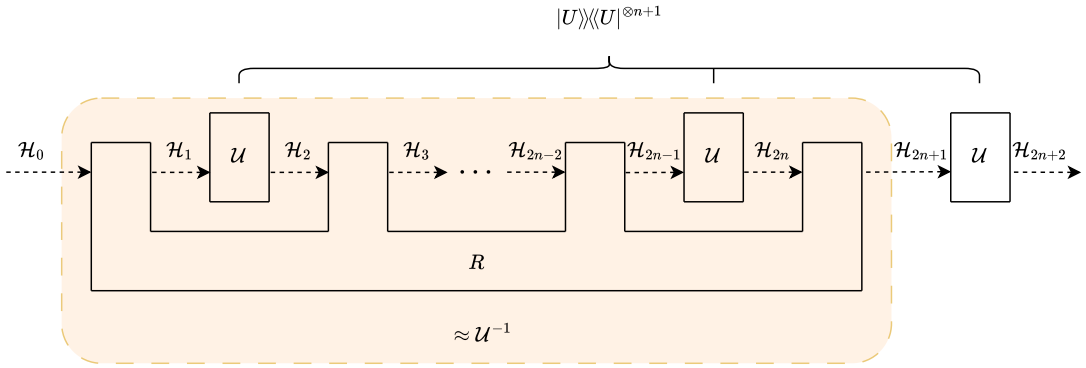


Figure 3: The quantum comb R is a unitary time-reversal algorithm with error bounded by ϵ using n queries to U and the overall channel from $\mathcal{H}_0$ to $\mathcal{H}_{2n+2}$ approximates the identity channel $\mathcal{I}$, i.e., $R \star C_U \overset{\epsilon}{\approx} |I\rangle\!\rangle\langle\!\langle I|$.

Corollary 3.3. *Suppose there is an algorithm that approximately implements U^{-t} to within constant diamond norm error $\epsilon \leq 10^{-5}$ for constant reversing time $t \geq 0.1$, using n queries to the unknown d-dimensional unitary U. Then, it must satisfy $n = \Omega(d^2)$.*

Here, the ranges $t \geq 0.1$ and $\epsilon \leq 10^{-5}$ are rather loose and chosen for simplicity of proof. The proof of Corollary 3.3 is deferred to Section 3.5.

3.1 Main framework

In this subsection, we introduce our main framework and settings for the proof of Theorem 3.1. Let $d \geq 2$ and $n \geq 1$ be arbitrary but fixed integers. We define a sequence of $2n+3$ Hilbert spaces $\mathcal{H}_0, \mathcal{H}_1, \ldots, \mathcal{H}_{2n+2}$ such that $\mathcal{H}_i \cong \mathbb{C}^d$ for $0 \leq i \leq 2n+2$.

Any algorithm, which takes n queries to an unknown d-dimensional unitary U and produces a d-dimensional quantum channel, can be described as an $(n+1)$-comb on $(\mathcal{H}_0, \mathcal{H}_1, \ldots, \mathcal{H}_{2n+1})$ where the i-th query is inserted on $(\mathcal{H}_{2i-1}, \mathcal{H}_{2i})$, and the produced quantum channel is from $\mathcal{H}_0$ to $\mathcal{H}_{2n+1}$ (see the orange region in Figure 3). Suppose R is a unitary time-reversal algorithm with average-case distance error bounded by ϵ. We make an additional query to U at the end of the algorithm R (see Figure 3). Then, the overall channel from $\mathcal{H}_0$ to $\mathcal{H}_{2n+2}$ should approximate the identity channel.

To formalize this, we first define the following $(n+1)$-comb C_U, which represents the $n+1$ queries to U.

Definition 3.4. *For any unitary $U \in \mathbb{U}_d$, we define the $(n+1)$-comb C_U on $(\mathcal{H}_1, \mathcal{H}_2, \ldots, \mathcal{H}_{2n+2})$ as*

$$C_U := \underbrace{|U\rangle\!\rangle\langle\!\langle U| \otimes \cdots \otimes |U\rangle\!\rangle\langle\!\langle U|}_{n+1},$$

where the i-th tensor factor $|U\rangle\!\rangle\langle\!\langle U|$ is the 1-comb on $(\mathcal{H}_{2i-1}, \mathcal{H}_{2i})$ corresponding to the unitary channel $\mathcal{U}$.

Therefore, we know that $R \star C_U \overset{\epsilon}{\approx} |I\rangle\!\rangle\langle\!\langle I|$. Furthermore, when the unknown unitary U is randomly sampled from the Haar measure, the overall channel should still approximate the identity channel i.e., $R \star \mathbf{E}_U[C_U] \overset{\epsilon}{\approx} |I\rangle\!\rangle\langle\!\langle I|$. Specifically, we have:

Proposition 3.5. *If an $(n+1)$-comb R on $(\mathcal{H}_0, \mathcal{H}_1, \ldots, \mathcal{H}_{2n+1})$ is a time-reversal algorithm with the average-case distance error bounded by ϵ, then we have*

$$\operatorname{tr}\left(|I\rangle\!\rangle\langle\!\langle I| \cdot \left(R \star \mathop{\mathbf{E}}_{U}[C_U]\right)\right) \geq d^2 - d(d+1)\epsilon,$$

where $|I\rangle\!\rangle\langle\!\langle I|$ is the 1-comb on $(\mathcal{H}_0, \mathcal{H}_{2n+2})$ corresponding to the identity channel, $\star$ is the link product (see Definition 2.4), $\mathbf{E}_U$ denotes the expectation over a Haar random unitary $U \in \mathbb{U}_d$.

Proof. For any $U \in \mathbb{U}_d$, R can produce a quantum channel that approximates $\mathcal{U}^{-1}$ to within the average-case distance error ϵ. Then, we apply the unitary channel $\mathcal{U}$ after the channel produced by R (see Figure 3). Due to the unitary invariance of the average-case distance (see Equation (6)), we obtain a channel that approximates the identity channel $\mathcal{I}$ to within the average-case distance error ϵ, i.e.,

$$\mathcal{D}_{\text{avg}}(\mathcal{E}_{R\star C_U}, \mathcal{I}) \leq \epsilon,$$

where $\mathcal{E}_{R\star C_U}$ denotes the quantum channel corresponding to the 1-comb (Choi state) $R \star C_U$. By Equation (7), this means

$$\frac{1}{d^2}\operatorname{tr}(|I\rangle\!\rangle\langle\!\langle I| \cdot (R \star C_U)) \geq 1 - \frac{d+1}{d}\epsilon.$$

Then, by taking the Haar expectation over $U \in \mathbb{U}_d$ and noting that $\star$ is bilinear, we have

$$\operatorname{tr}\left(|I\rangle\!\rangle\langle\!\langle I| \cdot \left(R \star \mathop{\mathbf{E}}_{U}[C_U]\right)\right) \geq d^2 - d(d+1)\epsilon.$$

□

On the other hand, we prove the following result, showing that when U is randomly sampled from the Haar measure and the number of queries n is not too large, the overall channel must be very depolarizing. The proof is deferred to Section 3.2.

Theorem 3.6. *For any $(n+1)$-comb X on $(\mathcal{H}_0, \mathcal{H}_1, \ldots, \mathcal{H}_{2n+1})$, we have*

$$X \star \mathop{\mathbf{E}}_{U}[C_U] \sqsubseteq \frac{n+1}{d} \cdot I_{\mathcal{H}_{2n+2}} \otimes I_{\mathcal{H}_0},$$

where $\mathbf{E}_U$ denotes the expectation over a Haar random unitary $U \in \mathbb{U}_d$.

Now, we are able to prove our main query lower bound (i.e., Theorem 3.1), assuming Theorem 3.6.

Proof of Theorem 3.1. Suppose R is an algorithm that approximately implements U^{-1} to within the average-case distance error ϵ, using n queries to U. Then, R can be described as an $(n+1)$-comb on $(\mathcal{H}_0, \mathcal{H}_1, \ldots, \mathcal{H}_{2n+1})$. Combining Proposition 3.5 and Theorem 3.6, we know that

$$n+1 \geq \operatorname{tr}\left(|I\rangle\!\rangle\langle\!\langle I| \cdot \left(R \star \mathop{\mathbf{E}}_{U}[C_U]\right)\right) \geq d^2 - d(d+1)\epsilon,$$

which means $n \geq d(d+1)(1-\epsilon) - (d+1)$. □

Equivalent to the query lower bound, we can also give an upper bound for the average fidelity achievable by an n-query unitary time-reversal algorithm for any fixed n, still using Theorem 3.6. Specifically, let $R(U)$ denote the channel produced by the algorithm R applied on n queries to U. The average performance of R can be evaluated using the channel fidelity between $R(U)$ and $\mathcal{U}^{-1}$ for Haar random U, i.e.,

$$F(R) \coloneqq \mathop{\mathbf{E}}_{U}\left[F_{\mathrm{ch}}(R(U),\mathcal{U}^{-1})\right].$$

Then, we can see the following result.

Theorem 3.7. *For any n-query algorithm R, we have $F(R) \leq \frac{n+1}{d^2}$.*

Proof. Note that

$$\begin{aligned}
\mathop{\mathbf{E}}_{U}\left[F_{\mathrm{ch}}(R(U),\mathcal{U}^{-1})\right] &= \mathop{\mathbf{E}}_{U}[F_{\mathrm{ch}}(\mathcal{U}\circ R(U),\mathcal{I})] && (25)\\
&= \frac{1}{d^2}\operatorname{tr}(R \star \mathop{\mathbf{E}}_{U}[C_U]\cdot |I\rangle\!\rangle\langle\!\langle I|)\\
&\leq \frac{n+1}{d^3}\operatorname{tr}(|I\rangle\!\rangle\langle\!\langle I|) && (26)\\
&= \frac{n+1}{d^2},
\end{aligned}$$

where Equation (25) is due to unitary invariance of the channel fidelity, and Equation (26) is due to Theorem 3.6 and the fact that R is an $(n+1)$-comb. □

We note that this upper bound matches the optimal fidelities numerically obtained by the semidefinite programming in [GYMO25, Table I]. This provides evidence that the fidelity upper bound in Theorem 3.7 (or equivalently, the query lower bound in Theorem 3.1) is likely optimal even in a non-asymptotic sense.

3.2 Proof of Theorem 3.6

Here, we define the following linear operators A_k for $1 \leq k \leq n$, which we refer to as *stair operators* and use as a tool in analyzing the Haar moment $\mathbf{E}_U[C_U]$.

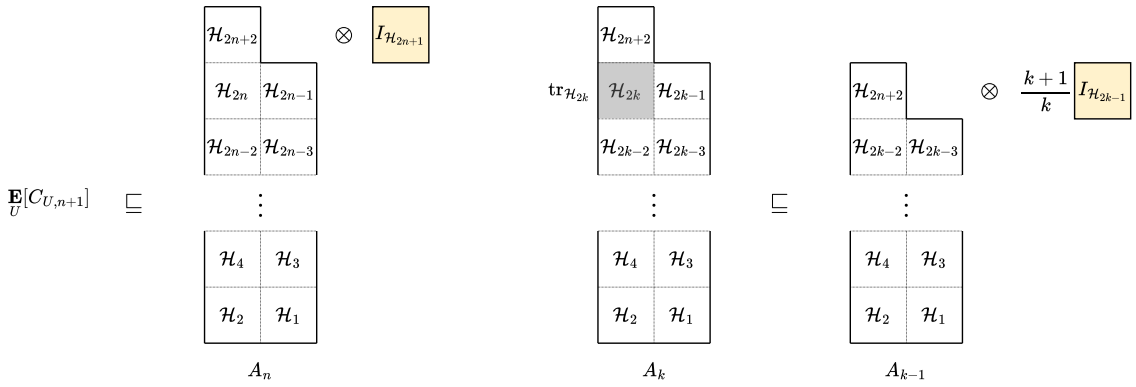
Definition 3.8 (Stair operators)**.** *For each $1 \leq k \leq n$, we define the linear operator A_k on*

$$\left(\left(\bigotimes_{i=1}^{k}\mathcal{H}_{2i}\right)\otimes\mathcal{H}_{2n+2}\right)\otimes\left(\bigotimes_{i=1}^{k}\mathcal{H}_{2i-1}\right) \overset{\mathbb{U}_d\times\mathbb{U}_d\times\mathfrak{S}_{k+1}\times\mathfrak{S}_k}{\cong} \bigoplus_{\substack{\lambda\vdash_d k+1\\ \mu\vdash_d k}} \mathcal{Q}_\lambda\otimes\mathcal{Q}_\mu\otimes\mathcal{P}_\lambda\otimes\mathcal{P}_\mu,$$

where $\mathfrak{S}_{k+1}$ acts on $\mathcal{H}_2\otimes\mathcal{H}_4\otimes\cdots\otimes\mathcal{H}_{2k}\otimes\mathcal{H}_{2n+2}$ and $\mathfrak{S}_k$ acts on $\mathcal{H}_1\otimes\mathcal{H}_3\otimes\cdots\otimes\mathcal{H}_{2k-1}$, as

$$A_k \coloneqq \bigoplus_{\lambda\vdash_d k+1}\bigoplus_{\mu:\mu\nearrow\lambda}\frac{\dim(\mathcal{P}_\lambda)}{\dim(\mathcal{P}_\mu)\dim(\mathcal{Q}_\lambda)}I_{\mathcal{Q}_\lambda}\otimes I_{\mathcal{Q}_\mu}\otimes\sum_{T,S\in\mathrm{Tab}(\lambda,\mu)}|T\rangle\langle S|\otimes|T^{\downarrow}\rangle\langle S^{\downarrow}|,$$

where $T^{\downarrow}$ denotes the standard Young tableau obtained from T by removing the box containing the largest integer, $\mathrm{Tab}(\lambda,\mu)$ denotes the set of standard Young tableaux $T\in\mathrm{Tab}(\lambda)$ such that $\mathrm{Sh}(T^{\downarrow})=\mu$.



(a) Bounding the Haar moment $\mathbf{E}_U[C_U]$. (b) Contraction of the stair operators.

Figure 4: Properties of the stair operators $\{A_k\}_{k=1}^n$ (see Lemma 3.9 and Lemma 3.10).

Note that each stair operator A_k is defined in an asymmetric bipartite system where one of the parties contains $k+1$ copies of $\mathbb{C}^d$ and the another contains k copies of $\mathbb{C}^d$. In addition, we will demonstrate a contraction process of A_k which leads to a smaller A_{k-1}. As shown in Figure 4b, each A_k resembles a stair, and the contraction process resembles walking down a stair. This is why we refer to $\{A_k\}_{k=1}^n$ as stair operators. A similar construction is adopted in [YKS+26]. However, their analysis applies only to the case of $n \leq d-1$, where a lower bound of $\Omega(d)$ is proved for the unitary time-reversal. In contrast, our construction and analysis apply to the general case without any constraint on n. Specifically, we provide the following lemmas to analyze the behavior of the Haar moment $\mathbf{E}_U[C_U]$ when combined with another quantum comb. The proofs are deferred to Section 3.3 and Section 3.4.

First, the stair operators provide a good upper bound (w.r.t. Löwner order) for $\mathbf{E}_U[C_U]$, see also Figure 4a.

Lemma 3.9. *Let C_U be that defined in Definition 3.4. Then*

$$\mathop{\mathbf{E}}_U[C_U] \sqsubseteq I_{\mathcal{H}_{2n+1}} \otimes A_n,$$

where $I_{\mathcal{H}_{2n+1}}$ is the identity operator on $\mathcal{H}_{2n+1}$.

Moreover, the stair operators share the similar tensor contraction property with quantum combs (see Equation (9)). Specifically, tracing out the "second-last" subsystem of a stair operator yields an operator bounded by the tensor product of the identity (with a coefficient) and a reduced stair operator, see also Figure 4b.

Lemma 3.10. *For any $2 \leq k \leq n$, we have*

$$\mathrm{tr}_{\mathcal{H}_{2k}}(A_k) \sqsubseteq \frac{k+1}{k} \cdot I_{\mathcal{H}_{2k-1}} \otimes A_{k-1}, \tag{27}$$

where $I_{\mathcal{H}_{2k-1}}$ is the identity operator on $\mathcal{H}_{2k-1}$. Moreover, we have

$$\mathrm{tr}_{\mathcal{H}_2}(A_1) = \frac{2}{d} \cdot I_{\mathcal{H}_1} \otimes I_{\mathcal{H}_{2n+2}}. \tag{28}$$

Lemma 3.10 implies that when a stair operator is combined with another quantum comb, they can be sequentially contracted. The following corollary is a direct application of Lemma 3.10, which characterizes the behavior of the stair operator A_k when combined with an arbitrary $(k+1)$-comb.

Corollary 3.11. *For any $1 \le k \le n$, and any $(k+1)$-comb X on $(\mathcal{H}_0, \mathcal{H}_1, \ldots, \mathcal{H}_{2k+1})$, we have*

$$X \star (I_{\mathcal{H}_{2k+1}} \otimes A_k) \sqsubseteq \frac{k+1}{d} \cdot I_{\mathcal{H}_{2n+2}} \otimes I_{\mathcal{H}_0}.$$

Proof. We use induction on k. For the case $k = 1$, we have

$$\begin{aligned}
X \star (I_{\mathcal{H}_3} \otimes A_1) &= \mathrm{tr}_{\mathcal{H}_{1:3}}\left(X^{\mathrm{T}_{\mathcal{H}_{1:3}}} \cdot (I_{\mathcal{H}_3} \otimes A_1)\right) \\
&= \mathrm{tr}_{\mathcal{H}_{1:2}}\left(\mathrm{tr}_{\mathcal{H}_3}(X)^{\mathrm{T}_{\mathcal{H}_{1:2}}} \cdot A_1\right) \\
&= \mathrm{tr}_{\mathcal{H}_{1:2}}\left(\left(I_{\mathcal{H}_2} \otimes X'\right)^{\mathrm{T}_{\mathcal{H}_{1:2}}} \cdot A_1\right) && (29) \\
&= \mathrm{tr}_{\mathcal{H}_1}\left(\left(X'\right)^{\mathrm{T}_{\mathcal{H}_1}} \cdot \mathrm{tr}_{\mathcal{H}_2}(A_1)\right) \\
&= \frac{2}{d} \cdot \mathrm{tr}_{\mathcal{H}_1}\left(\left(X'\right)^{\mathrm{T}_{\mathcal{H}_1}} \cdot I_{\mathcal{H}_1} \otimes I_{\mathcal{H}_{2n+2}}\right) && (30) \\
&= \frac{2}{d} \cdot I_{\mathcal{H}_0} \otimes I_{\mathcal{H}_{2n+2}}, && (31)
\end{aligned}$$

where Equation (29) is because X is a 2-comb on $(\mathcal{H}_0, \mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3)$, Equation (30) is due to Equation (28) in Lemma 3.10, and Equation (31) is because X' is a 1-comb on $(\mathcal{H}_0, \mathcal{H}_1)$.

Now, suppose the induction hypothesis holds for $k-1$, we prove it also hods for k.

$$\begin{aligned}
X \star (I_{\mathcal{H}_{2k+1}} \otimes A_k) &= \mathrm{tr}_{\mathcal{H}_{1:2k+1}}\left(X^{\mathrm{T}_{\mathcal{H}_{1:2k+1}}} \cdot (I_{\mathcal{H}_{2k+1}} \otimes A_k)\right) \\
&= \mathrm{tr}_{\mathcal{H}_{1:2k}}\left(\mathrm{tr}_{\mathcal{H}_{2k+1}}(X)^{\mathrm{T}_{\mathcal{H}_{1:2k}}} \cdot A_k\right) \\
&= \mathrm{tr}_{\mathcal{H}_{1:2k}}\left(\left(I_{\mathcal{H}_{2k}} \otimes X'\right)^{\mathrm{T}_{\mathcal{H}_{1:2k}}} \cdot A_k\right) && (32) \\
&= \mathrm{tr}_{\mathcal{H}_{1:2k-1}}\left(\left(X'\right)^{\mathrm{T}_{\mathcal{H}_{1:2k-1}}} \cdot \mathrm{tr}_{\mathcal{H}_{2k}}(A_k)\right) \\
&= \left(X'\right) \star \mathrm{tr}_{\mathcal{H}_{2k}}(A_k) \\
&\sqsubseteq \frac{k+1}{k} \cdot \left(X'\right) \star (I_{\mathcal{H}_{2k-1}} \otimes A_{k-1}) && (33) \\
&\sqsubseteq \frac{k+1}{d} \cdot I_{\mathcal{H}_{2n+2}} \otimes I_{\mathcal{H}_0}, && (34)
\end{aligned}$$

where Equation (32) is because X is an $(k+1)$-comb and thus X' is a k-comb, Equation (33) is due to Equation (27) in Lemma 3.10, and Equation (34) is due to the induction hypothesis for the case of $k-1$ □

Then, we are able to prove Theorem 3.6, assuming Lemma 3.9 and Lemma 3.10.

Proof of Theorem 3.6. Note that

$$\begin{aligned}
X \star \mathop{\mathbf{E}}_{U}[C_U] &\sqsubseteq X \star (I_{\mathcal{H}_{2n+1}} \otimes A_n) && (35) \\
&\sqsubseteq \frac{n+1}{d} \cdot I_{\mathcal{H}_{2n+2}} \otimes I_{\mathcal{H}_0}, && (36)
\end{aligned}$$

where Equation (35) is by Lemma 3.9 combined with Equation (10), and Equation (36) is by taking $k = n$ in Corollary 3.11. □

3.3 Bounding the Haar moment: proof of Lemma 3.9

In this subsection, we bound the Haar moment $\mathbf{E}_U[C_U]$ using the stair operators (see Lemma 3.9).

First, we need the following fact:

$$\mathop{\mathbf{E}}_{U}[C_U] = \bigoplus_{\lambda \vdash_d n+1} \frac{1}{\dim(\mathcal{Q}_\lambda)} I_{\mathcal{Q}_\lambda} \otimes I_{\mathcal{Q}_\lambda} \otimes |I_{\mathcal{P}_\lambda}\rangle\!\rangle\langle\!\langle I_{\mathcal{P}_\lambda}|,$$

which provides a representation-theoretic expression of the Haar moment $\mathbf{E}_U[C_U]$ (see Lemma 4.7). Then, we provide our proof of Lemma 3.9.

Proof of Lemma 3.9. By Lemma 4.7, we can express $\mathbf{E}_U[C_U]$ in the Schur-Weyl basis

$$\bigoplus_{\substack{\lambda \vdash_d n+1 \\ \mu \vdash_d n+1}} \mathcal{Q}_\lambda \otimes \mathcal{Q}_\mu \otimes \mathcal{P}_\lambda \otimes \mathcal{P}_\mu \overset{\mathbb{U}_d \times \mathbb{U}_d \times \mathfrak{S}_{n+1} \times \mathfrak{S}_{n+1}}{\cong} \bigotimes_{i=1}^{n+1} \mathcal{H}_{2i} \otimes \bigotimes_{i=1}^{n+1} \mathcal{H}_{2i-1}.$$

On the other hand, by Definition 3.8, A_n is defined on the Schur-Weyl basis

$$\bigoplus_{\substack{\lambda \vdash_d n+1 \\ \mu \vdash_d n}} \mathcal{Q}_\lambda \otimes \mathcal{Q}_\mu \otimes \mathcal{P}_\lambda \otimes \mathcal{P}_\mu \overset{\mathbb{U}_d \times \mathbb{U}_d \times \mathfrak{S}_{n+1} \times \mathfrak{S}_{n}}{\cong} \bigotimes_{i=1}^{n+1} \mathcal{H}_{2i} \otimes \bigotimes_{i=1}^{n} \mathcal{H}_{2i-1}.$$

By comparing these two forms, we note that it is natural to study that for $\mu \vdash n$ and $T, S \in \mathrm{Tab}(\mu)$, how the linear operator $I_{\mathcal{H}_{2n+1}} \otimes I_{\mathcal{Q}_\mu} \otimes |T\rangle\langle S| \in \mathcal{L}(\mathcal{H}_{2n+1} \otimes \mathcal{Q}_\mu \otimes \mathcal{P}_\mu) \subseteq \mathcal{L}(\bigotimes_{i=1}^{n+1} \mathcal{H}_{2i-1})$ is expressed as an operator on $\bigoplus_{\mu \vdash_d n+1} \mathcal{Q}_\mu \otimes \mathcal{P}_\mu$. For this, we have the following result by using Lemma 4.1:

$$I_{\mathcal{H}_{2n+1}} \otimes I_{\mathcal{Q}_\mu} \otimes |T\rangle\langle S| = \bigoplus_{\nu : \mu \nearrow \nu} I_{\mathcal{Q}_\nu} \otimes |T^{\uparrow \nu}\rangle\langle S^{\uparrow \nu}|,$$

where $T^{\uparrow \nu}$ denotes the standard Young tableau obtained from T by adding a box filled with $n+1$, resulting in the shape ν. This means

$$\begin{aligned}
I_{\mathcal{H}_{2n+1}} \otimes A_n &= \bigoplus_{\lambda \vdash_d k+1} \bigoplus_{\mu : \mu \nearrow \lambda} \frac{\dim(\mathcal{P}_\lambda)}{\dim(\mathcal{P}_\mu)\dim(\mathcal{Q}_\lambda)} I_{\mathcal{Q}_\lambda} \otimes \bigoplus_{\nu : \mu \nearrow \nu} I_{\mathcal{Q}_\nu} \otimes \sum_{T,S \in \mathrm{Tab}(\lambda,\mu)} |T\rangle\langle S| \otimes |(T^{\downarrow})^{\uparrow \nu}\rangle\langle (S^{\downarrow})^{\uparrow \nu}| \\
&\sqsupseteq \bigoplus_{\lambda \vdash_d k+1} \bigoplus_{\mu : \mu \nearrow \lambda} \frac{\dim(\mathcal{P}_\lambda)}{\dim(\mathcal{P}_\mu)\dim(\mathcal{Q}_\lambda)} I_{\mathcal{Q}_\lambda} \otimes I_{\mathcal{Q}_\lambda} \otimes \sum_{T,S \in \mathrm{Tab}(\lambda,\mu)} |T\rangle\langle S| \otimes |T\rangle\langle S| \qquad (37) \\
&= \bigoplus_{\lambda \vdash_d k+1} \frac{1}{\dim(\mathcal{Q}_\lambda)} I_{\mathcal{Q}_\lambda} \otimes I_{\mathcal{Q}_\lambda} \otimes \sum_{\mu : \mu \nearrow \lambda} \frac{\dim(\mathcal{P}_\lambda)}{\dim(\mathcal{P}_\mu)} |\Phi_\mu^\lambda\rangle\langle \Phi_\mu^\lambda| \qquad (38) \\
&\sqsupseteq \bigoplus_{\lambda \vdash_d k+1} \frac{1}{\dim(\mathcal{Q}_\lambda)} I_{\mathcal{Q}_\lambda} \otimes I_{\mathcal{Q}_\lambda} \otimes \left(\sum_{\mu : \mu \nearrow \lambda} |\Phi_\mu^\lambda\rangle \right) \left(\sum_{\mu : \mu \nearrow \lambda} \langle \Phi_\mu^\lambda| \right) \qquad (39) \\
&= \bigoplus_{\lambda \vdash_d k+1} \frac{1}{\dim(\mathcal{Q}_\lambda)} I_{\mathcal{Q}_\lambda} \otimes I_{\mathcal{Q}_\lambda} \otimes |I_{\mathcal{P}_\lambda}\rangle\!\rangle\langle\!\langle I_{\mathcal{P}_\lambda}| \qquad (40) \\
&= \mathop{\mathbf{E}}_{U}[C_U],
\end{aligned}$$

where Equation (37) is by discarding those ν such that $\nu \neq \lambda$; in Equation (38), $|\Phi_\mu^\lambda\rangle \coloneqq \sum_{T\in \mathrm{Tab}(\lambda,\mu)} |T\rangle|T\rangle$; Equation (40) is because $|I_{\mathcal{P}_\lambda}\rangle\!\rangle = \sum_{T\in\mathrm{Tab}(\lambda)} |T\rangle|T\rangle = \sum_{\mu:\mu\nearrow\lambda} |\Phi_\mu^\lambda\rangle$; Equation (39) is by using Lemma 4.10 and the fact:

$$\begin{aligned}
&\left(\sum_{\mu:\mu\nearrow\lambda} \langle\Phi_\mu^\lambda|\right)\left(\sum_{\mu:\mu\nearrow\lambda} \frac{\dim(\mathcal{P}_\lambda)}{\dim(\mathcal{P}_\mu)} |\Phi_\mu^\lambda\rangle\langle\Phi_\mu^\lambda|\right)^{-1}\left(\sum_{\mu:\mu\nearrow\lambda} |\Phi_\mu^\lambda\rangle\right)\\
=&\left(\sum_{\mu:\mu\nearrow\lambda} \langle\Phi_\mu^\lambda|\right)\left(\sum_{\mu:\mu\nearrow\lambda} \frac{\dim(\mathcal{P}_\mu)}{\dim(\mathcal{P}_\lambda)}\cdot\frac{1}{\dim(\mathcal{P}_\mu)^2}\cdot |\Phi_\mu^\lambda\rangle\langle\Phi_\mu^\lambda|\right)\left(\sum_{\mu:\mu\nearrow\lambda} |\Phi_\mu^\lambda\rangle\right) \qquad (41)\\
=&\sum_{\mu:\mu\nearrow\lambda} \frac{\dim(\mathcal{P}_\mu)}{\dim(\mathcal{P}_\lambda)}\cdot\frac{1}{\dim(\mathcal{P}_\mu)^2}\cdot\dim(\mathcal{P}_\mu)^2\\
=&1, \qquad (42)
\end{aligned}$$

in which Equation (41) is because $|\Phi_\mu^\lambda\rangle$ are pairwise orthogonal and $|\mathrm{Tab}(\lambda,\mu)| = |\mathrm{Tab}(\mu)| = \dim(\mathcal{P}_\mu)$, Equation (42) is due to $\sum_{\mu:\mu\nearrow\lambda}\dim(\mathcal{P}_\mu) = \dim(\mathcal{P}_\lambda)$. □

3.4 Contraction of stair operators: proof of Lemma 3.10

In this subsection, we prove the contraction properties of the stair operators (see Lemma 3.10).

3.4.1 The case of $2 \le k \le n$

First, we prove for $2 \le k \le n$,

$$\mathrm{tr}_{\mathcal{H}_{2k}}(A_k) \sqsubseteq \frac{k+1}{k}\cdot I_{\mathcal{H}_{2k-1}} \otimes A_{k-1}.$$

Note that in the definition of A_k (see Definition 3.8), the symmetric group $\mathfrak{S}_{k+1}$ acts on $\mathcal{H}_2 \otimes \mathcal{H}_4 \otimes \cdots \otimes \mathcal{H}_{2k} \otimes \mathcal{H}_{2n+2}$. We can not directly use Lemma 4.2 since $\mathrm{tr}_{\mathcal{H}_{2k}}$ is tracing out the second-to-last subsystem $\mathcal{H}_{2k}$ instead of the last subsystem $\mathcal{H}_{2n+2}$. For this, we swap the last two tensor factors of A_k by using the k-th adjacent transposition s_k. Specifically, let $P(\pi)$ be the action of $\pi \in \mathfrak{S}_{k+1}$ on $\mathcal{H}_2 \otimes \mathcal{H}_4 \otimes \cdots \otimes \mathcal{H}_{2k} \otimes \mathcal{H}_{2n+2}$. Then, $P(s_k)A_kP(s_k)$ is the operator on $\mathcal{H}_2 \otimes \mathcal{H}_4 \otimes \cdots \otimes \mathcal{H}_{2k-2} \otimes \mathcal{H}_{2n+2} \otimes \mathcal{H}_{2k}$ obtained from A_k by swapping the positions of $\mathcal{H}_{2k}$ and $\mathcal{H}_{2n+2}$. By the Young's orthogonal form (see Equation (24)), we have

$$P(s_k)A_kP(s_k) = \bigoplus_{\lambda\vdash_d k+1}\bigoplus_{\mu:\mu\nearrow\lambda}\frac{\dim(\mathcal{P}_\lambda)}{\dim(\mathcal{P}_\mu)\dim(\mathcal{Q}_\lambda)} I_{\mathcal{Q}_\lambda}\otimes I_{\mathcal{Q}_\mu}\otimes \sum_{T,S\in\mathrm{Tab}(\lambda,\mu)} \perp_{T,S}\otimes |T^\downarrow\rangle\langle S^\downarrow|, \qquad (43)$$

where $\perp_{T,S}$ is defined as

$$\frac{1}{r(T)r(S)}\left(|T\rangle\langle S| + \sqrt{(r(T)^2-1)(r(S)^2-1)}|s_kT\rangle\langle s_kS| + \sqrt{r(T)^2-1}|s_kT\rangle\langle S| + \sqrt{r(S)^2-1}|T\rangle\langle s_kS|\right), \qquad (44)$$

where $r(T)$ is the axial distance from the $(k+1)$-box to the k-box in T. Then, we can use Lemma 4.2 to calculate the partial trace $\mathrm{tr}_{\mathcal{H}_{2k}}(P(s_k)A_kP(s_k))$.

To this end, let us consider each single summand in the RHS of Equation (43). Suppose $\lambda \vdash_d k+1$, $\mu \nearrow \lambda$, $T, S \in \mathrm{Tab}(\lambda,\mu)$. First note that it is not possible that $\mathrm{Sh}((s_kT)^\downarrow) = \mathrm{Sh}(S^\downarrow)$ since we know $\mathrm{Sh}(T^\downarrow) = \mathrm{Sh}(S^\downarrow)$. Thus by Lemma 4.2, we can ignore the terms $|s_kT\rangle\langle S|$ and $|T\rangle\langle s_kS|$ in Equation (44) since they do not contribute after the partial trace. Next, we consider the terms $|s_kT\rangle\langle s_kS|$ and $|T\rangle\langle S|$ separately.

Summands involving the term $|s_kT\rangle\langle s_kS|$. Consider the term $|s_kT\rangle\langle s_kS|$. We have

$$\begin{aligned}&\mathrm{tr}_{\mathcal{H}_{2k}}\left(\frac{\dim(\mathcal{P}_\lambda)}{\dim(\mathcal{P}_\mu)\dim(\mathcal{Q}_\lambda)}I_{\mathcal{Q}_\lambda}\otimes I_{\mathcal{Q}_\mu}\otimes\frac{\sqrt{(r(T)^2-1)(r(S)^2-1)}}{r(T)r(S)}|s_kT\rangle\langle s_kS|\otimes|T^{\downarrow}\rangle\langle S^{\downarrow}|\right)\\ =&\frac{\dim(\mathcal{P}_\lambda)}{\dim(\mathcal{P}_\mu)\dim(\mathcal{Q}_{\mathrm{Sh}((s_kT)^{\downarrow})})}I_{\mathcal{Q}_{\mathrm{Sh}((s_kT)^{\downarrow})}}\otimes I_{\mathcal{Q}_\mu}\\ &\otimes\mathbb{1}_{\mathrm{Sh}((s_kT)^{\downarrow})=\mathrm{Sh}((s_kS)^{\downarrow})}\cdot\frac{\sqrt{(r(T)^2-1)(r(S)^2-1)}}{r(T)r(S)}|(s_kT)^{\downarrow}\rangle\langle(s_kS)^{\downarrow}|\otimes|T^{\downarrow}\rangle\langle S^{\downarrow}|.\end{aligned}\tag{45}$$

Then, we consider their summation. For those summands of the form in Equation (45), by re-naming the variables $\mathrm{Sh}((s_kT)^{\downarrow})$ to ν, $(s_kT)^{\downarrow}$ to T and $(s_kS)^{\downarrow}$ to S, we can write the corresponding summation as

$$\bigoplus_{\nu\vdash_d k}\frac{1}{\dim(\mathcal{Q}_\nu)}I_{\mathcal{Q}_\nu}\otimes\bigoplus_{\substack{\lambda:\nu\nearrow\lambda\\ \ell(\lambda)\le d}}\bigoplus_{\substack{\mu:\mu\nearrow\lambda\\ \mu\neq\nu}}\frac{\dim(\mathcal{P}_\lambda)}{\dim(\mathcal{P}_\mu)}I_{\mathcal{Q}_\mu}\otimes\sum_{T,S\in\mathrm{Tab}(\nu,\mu\cap\nu)}\frac{\sqrt{(r(T^{\uparrow\lambda})^2-1)(r(S^{\uparrow\lambda})^2-1)}}{r(T^{\uparrow\lambda})r(S^{\uparrow\lambda})}|T\rangle\langle S|\otimes|T_\mu\rangle\langle S_\mu|,\tag{46}$$

where T_μ denotes the standard Young tableau obtained from T by moving the largest box of T to the position that results in shape μ (such moving always exists since μ is adjacent to ν). Since the summation is taken over all $\mu\vdash_d k$ that is adjacent to ν, by re-naming $\mu\cap\nu$ to τ, we can write Equation (46) in an equivalent form:

$$\bigoplus_{\nu\vdash_d k}\frac{1}{\dim(\mathcal{Q}_\nu)}I_{\mathcal{Q}_\nu}\otimes\bigoplus_{\tau:\tau\nearrow\nu}\bigoplus_{\substack{\mu:\tau\nearrow\mu\\ \mu\neq\nu\\ \ell(\mu)\le d}}\frac{\dim(\mathcal{P}_{\mu\cup\nu})}{\dim(\mathcal{P}_\mu)}I_{\mathcal{Q}_\mu}\otimes\sum_{T,S\in\mathrm{Tab}(\nu,\tau)}\frac{\sqrt{(r(T^{\uparrow\mu\cup\nu})^2-1)(r(S^{\uparrow\mu\cup\nu})^2-1)}}{r(T^{\uparrow\mu\cup\nu})r(S^{\uparrow\mu\cup\nu})}|T\rangle\langle S|\otimes|T_\mu\rangle\langle S_\mu|.\tag{47}$$

Note that for $T,S\in\mathrm{Tab}(\nu,\tau)$, $r(T^{\uparrow\mu\cup\nu})=r(S^{\uparrow\mu\cup\nu})=c(\mu\setminus\nu)-c(\nu\setminus\tau)$. Then, by Lemma 4.3, we can write Equation (47) as

$$\frac{k+1}{k}\bigoplus_{\nu\vdash_d k}\frac{1}{\dim(\mathcal{Q}_\nu)}I_{\mathcal{Q}_\nu}\otimes\bigoplus_{\tau:\tau\nearrow\nu}\bigoplus_{\substack{\mu:\tau\nearrow\mu\\ \mu\neq\nu}}\frac{\dim(\mathcal{P}_\nu)}{\dim(\mathcal{P}_\tau)}I_{\mathcal{Q}_\mu}\otimes\sum_{T,S\in\mathrm{Tab}(\nu,\tau)}|T\rangle\langle S|\otimes|T_\mu\rangle\langle S_\mu|,\tag{48}$$

where we ignore the constraint $\ell(\mu)\le d$ since $I_{\mathcal{Q}_\mu}=0$ when $\ell(\mu)>d$.

Summands involving the term $|T\rangle\langle S|$. Then, consider the term $|T\rangle\langle S|$. We have

$$\begin{aligned}&\mathrm{tr}_{\mathcal{H}_{2k}}\left(\frac{\dim(\mathcal{P}_\lambda)}{\dim(\mathcal{P}_\mu)\dim(\mathcal{Q}_\lambda)}I_{\mathcal{Q}_\lambda}\otimes I_{\mathcal{Q}_\mu}\otimes\frac{1}{r(T)r(S)}|T\rangle\langle S|\otimes|T^{\downarrow}\rangle\langle S^{\downarrow}|\right)\\ =&\frac{\dim(\mathcal{P}_\lambda)}{\dim(\mathcal{P}_\mu)\dim(\mathcal{Q}_\mu)}I_{\mathcal{Q}_\mu}\otimes I_{\mathcal{Q}_\mu}\otimes\frac{1}{r(T)r(S)}|T^{\downarrow}\rangle\langle S^{\downarrow}|\otimes|T^{\downarrow}\rangle\langle S^{\downarrow}|.\end{aligned}\tag{49}$$

For those summands of the form in Equation (49), by re-naming the variables μ to ν, $T^{\downarrow}$ to T and $S^{\downarrow}$ to S, we can write the corresponding summation as

$$\bigoplus_{\nu\vdash_d k}\frac{1}{\dim(\mathcal{P}_\nu)\dim(\mathcal{Q}_\nu)}I_{\mathcal{Q}_\nu}\otimes I_{\mathcal{Q}_\nu}\otimes\sum_{\substack{\lambda:\nu\nearrow\lambda\\ \ell(\lambda)\le d}}\sum_{T,S\in\mathrm{Tab}(\nu)}\frac{\dim(\mathcal{P}_\lambda)}{r(T^{\uparrow\lambda})r(S^{\uparrow\lambda})}|T\rangle\langle S|\otimes|T\rangle\langle S|.\tag{50}$$

Then, note that $\mathrm{Tab}(\nu) = \bigcup_{\tau:\tau\nearrow\nu} \mathrm{Tab}(\nu,\tau)$. We can write Equation (50) as

$$\bigoplus_{\nu\vdash_d k} \frac{1}{\dim(\mathcal{P}_\nu)\dim(\mathcal{Q}_\nu)} I_{\mathcal{Q}_\nu} \otimes I_{\mathcal{Q}_\nu} \otimes \sum_{\substack{\lambda:\nu\nearrow\lambda \\ \ell(\lambda)\le d}} \sum_{\substack{\tau:\tau\nearrow\nu \\ \kappa:\kappa\nearrow\nu}} \sum_{\substack{T\in\mathrm{Tab}(\nu,\tau) \\ S\in\mathrm{Tab}(\nu,\kappa)}} \frac{\dim(\mathcal{P}_\lambda)}{r(T^{\uparrow\lambda})r(S^{\uparrow\lambda})} |T\rangle\langle S| \otimes |T\rangle\langle S|. \tag{51}$$

Note that if we add

$$\bigoplus_{\nu\vdash_d k} \frac{1}{\dim(\mathcal{P}_\nu)\dim(\mathcal{Q}_\nu)} I_{\mathcal{Q}_\nu} \otimes I_{\mathcal{Q}_\nu} \otimes \sum_{\substack{\lambda:\nu\nearrow\lambda \\ \ell(\lambda)=d+1}} \sum_{\substack{\tau:\tau\nearrow\nu \\ \kappa:\kappa\nearrow\nu}} \sum_{\substack{T\in\mathrm{Tab}(\nu,\tau) \\ S\in\mathrm{Tab}(\nu,\kappa)}} \frac{\dim(\mathcal{P}_\lambda)}{r(T^{\uparrow\lambda})r(S^{\uparrow\lambda})} |T\rangle\langle S| \otimes |T\rangle\langle S| \tag{52}$$

to Equation (51), we can remove the constraint $\ell(\lambda) \le d$ and obtain:

$$\bigoplus_{\nu\vdash_d k} \frac{1}{\dim(\mathcal{P}_\nu)\dim(\mathcal{Q}_\nu)} I_{\mathcal{Q}_\nu} \otimes I_{\mathcal{Q}_\nu} \otimes \sum_{\lambda:\nu\nearrow\lambda} \sum_{\substack{\tau:\tau\nearrow\nu \\ \kappa:\kappa\nearrow\nu}} \sum_{\substack{T\in\mathrm{Tab}(\nu,\tau) \\ S\in\mathrm{Tab}(\nu,\kappa)}} \frac{\dim(\mathcal{P}_\lambda)}{r(T^{\uparrow\lambda})r(S^{\uparrow\lambda})} |T\rangle\langle S| \otimes |T\rangle\langle S|. \tag{53}$$

Note that Equation (52) is positive semidefinite. Therefore, Equation (51) $\sqsubseteq$ Equation (53). Note that $r(T^{\uparrow\lambda}) = c(\lambda\setminus\nu) - c(\nu\setminus\tau) > 0$ and $r(S^{\uparrow\lambda}) = c(\lambda\setminus\nu) - c(\nu\setminus\kappa) > 0$. Then, in Equation (53), by Lemma 4.4, we can discard those terms such that $\tau \ne \kappa$. Therefore, Equation (53) simplifies to:

$$\bigoplus_{\nu\vdash_d k} \frac{1}{\dim(\mathcal{P}_\nu)\dim(\mathcal{Q}_\nu)} I_{\mathcal{Q}_\nu} \otimes I_{\mathcal{Q}_\nu} \otimes \sum_{\lambda:\nu\nearrow\lambda} \sum_{\tau:\tau\nearrow\nu} \sum_{T,S\in\mathrm{Tab}(\nu,\tau)} \frac{\dim(\mathcal{P}_\lambda)}{r(T^{\uparrow\lambda})r(S^{\uparrow\lambda})} |T\rangle\langle S| \otimes |T\rangle\langle S|. \tag{54}$$

Then, by Lemma 4.5, we can further simplify Equation (54) to

$$\frac{k+1}{k} \cdot \bigoplus_{\nu\vdash_d k} \frac{1}{\dim(\mathcal{Q}_\nu)} I_{\mathcal{Q}_\nu} \otimes I_{\mathcal{Q}_\nu} \otimes \sum_{\tau:\tau\nearrow\nu} \frac{\dim(\mathcal{P}_\nu)}{\dim(\mathcal{P}_\tau)} \sum_{T,S\in\mathrm{Tab}(\nu,\tau)} |T\rangle\langle S| \otimes |T\rangle\langle S|. \tag{55}$$

In summary, we proved that the sum of the terms involving $|T\rangle\langle S|$ (i.e., Equation (50)) is less than (w.r.t. Löwner order) or equal to Equation (55).

Putting it all together. Then, combining the above results, we can see that the sum of Equation (48) and Equation (55) is:

$$\frac{k+1}{k} \cdot \bigoplus_{\nu\vdash_d k} \frac{1}{\dim(\mathcal{Q}_\nu)} I_{\mathcal{Q}_\nu} \otimes \bigoplus_{\tau:\tau\nearrow\nu} \bigoplus_{\mu:\tau\nearrow\mu} \frac{\dim(\mathcal{P}_\nu)}{\dim(\mathcal{P}_\tau)} I_{\mathcal{Q}_\mu} \otimes \sum_{T,S\in\mathrm{Tab}(\nu,\tau)} |T\rangle\langle S| \otimes |T_\mu\rangle\langle S_\mu|. \tag{56}$$

Note that in Equation (56), $\sum_{\mu:\tau\nearrow\mu} I_{\mathcal{Q}_\mu} \otimes |T_\mu\rangle\langle S_\mu|$ is an operator on $\bigotimes_{i=1}^k \mathcal{H}_{2i-1}$. Thus Lemma 4.1 implies that

$$\bigoplus_{\mu:\tau\nearrow\mu} I_{\mathcal{Q}_\mu} \otimes |T_\mu\rangle\langle S_\mu| = I_{\mathcal{H}_{2k-1}} \otimes I_{\mathcal{Q}_\tau} \otimes |T^{\downarrow}\rangle\langle S^{\downarrow}|.$$

This means Equation (56) is

$$\begin{aligned}
&\frac{k+1}{k} I_{\mathcal{H}_{2k-1}} \otimes \bigoplus_{\nu\vdash_d k} \bigoplus_{\tau:\tau\nearrow\nu} \frac{\dim(\mathcal{P}_\nu)}{\dim(\mathcal{P}_\tau)\dim(\mathcal{Q}_\nu)} I_{\mathcal{Q}_\nu} \otimes I_{\mathcal{Q}_\tau} \otimes \sum_{T,S\in\mathrm{Tab}(\nu,\tau)} |T\rangle\langle S| \otimes |T^{\downarrow}\rangle\langle S^{\downarrow}| \\
=&\frac{k+1}{k} \cdot I_{\mathcal{H}_{2k-1}} \otimes A_{k-1}.
\end{aligned}$$

3.4.2 The case of $k = 1$

Then, we prove

$$\mathrm{tr}_{\mathcal{H}_2}(A_1) = \frac{2}{d} \cdot I_{\mathcal{H}_1} \otimes I_{\mathcal{H}_{2n+2}}.$$

by explicit calculation. Specifically, let $\lambda = \square\square$ and $\mu = \begin{smallmatrix}\square\\\square\end{smallmatrix}$ be the two Young diagrams of 2 boxes and $\nu = \square$ be the Young diagram of 1 box. Let $T = \begin{smallmatrix}1&2\end{smallmatrix}$ and $S = \begin{smallmatrix}1\\2\end{smallmatrix}$. Thus, $\nu \nearrow \lambda$, $\nu \nearrow \mu$, $\dim(\mathcal{P}_\lambda) = \dim(\mathcal{P}_\mu) = \dim(\mathcal{P}_\nu) = 1$, $\dim(\mathcal{Q}_\lambda) = \frac{d(d+1)}{2}$, $\dim(\mathcal{Q}_\mu) = \frac{d(d-1)}{2}$ and $\dim(\mathcal{Q}_\nu) = d$. Then, we have

$$A_1 = \frac{2}{d(d+1)} I_{\mathcal{Q}_\lambda} \otimes I_{\mathcal{Q}_\nu} \otimes |T\rangle\langle T| \otimes |T^{\downarrow}\rangle\langle T^{\downarrow}| + \frac{2}{d(d-1)} I_{\mathcal{Q}_\mu} \otimes I_{\mathcal{Q}_\nu} \otimes |S\rangle\langle S| \otimes |S^{\downarrow}\rangle\langle S^{\downarrow}|.$$

Next, we swap the subsystems $\mathcal{H}_2$ and $\mathcal{H}_{2n+2}$. In fact, note that the swap s_1 acts trivially on $|T\rangle$ and $|S\rangle$. Thus, we can directly calculate the partial trace on A_1 by Lemma 4.2:

$$\begin{aligned}\mathrm{tr}_{\mathcal{H}_2}(A_1) &= \frac{1}{d} I_{\mathcal{Q}_\nu} \otimes I_{\mathcal{Q}_\nu} \otimes |T^{\downarrow}\rangle\langle T^{\downarrow}| \otimes |T^{\downarrow}\rangle\langle T^{\downarrow}| + \frac{1}{d} I_{\mathcal{Q}_\nu} \otimes I_{\mathcal{Q}_\nu} \otimes |S^{\downarrow}\rangle\langle S^{\downarrow}| \otimes |S^{\downarrow}\rangle\langle S^{\downarrow}| \\ &= \frac{2}{d} I_{\mathcal{H}_1} \otimes I_{\mathcal{H}_{2n+2}}, \end{aligned} \tag{57}$$

where Equation (57) is because $T^{\downarrow} = S^{\downarrow} = \begin{smallmatrix}1\end{smallmatrix}$ and thus $|T^{\downarrow}\rangle\langle T^{\downarrow}| = |S^{\downarrow}\rangle\langle S^{\downarrow}| = I_{\mathcal{P}_\nu}$.

3.5 Generalized time-reversal

In this section, let us prove Corollary 3.3, which gives a lower bound for generalized time-reversal of an unknown unitary. Here, we need to use the following Dirichlet's approximation theorem.

Lemma 3.12 (Dirichlet's approximation theorem)**.** *For any real number $t \in \mathbb{R}$ and integer $N \geq 1$, there exist integers a, b such that $1 \leq b \leq N$ and $|bt - a| < 1/N$.*

Then, we are able to give the proof of Corollary 3.3.

Proof of Corollary 3.3. Suppose $\mathcal{A}$ is an algorithm that implements the unitary $U^{-t} = e^{-iHt}$ to within constant diamond norm error $\epsilon \leq 10^{-5}$ for some constant $t \geq 0.1$, using n queries to U, where H is a Hamiltonian satisfying $e^{iH} = e^{i\theta}U$ for some real number θ and $\|H\| \leq \pi$. By Lemma 3.12, there exist integers a' and $1 \leq b' \leq 10^3$ such that $|b't - a'| < 10^{-3}$. Let $b = 10b', a = 10a'$. Then, we have $|bt - a| \leq 0.01$, where $10 \leq b \leq 10^4$. We can also see $a \geq 1$ since $t \geq 0.1$.

Now let us construct an algorithm to approximately implement U^{-1}, based on algorithm $\mathcal{A}$. First, we repeat algorithm $\mathcal{A}$ b times, and then we obtain an algorithm $\mathcal{B}$ to implement the unitary e^{-iHbt} to within diamond norm error $b\epsilon \leq 0.1$. Second, we show the unitary e^{-iHbt} is close to e^{-iHa} in diamond distance, as follows. Let $\mathcal{V}_1(\cdot) = e^{-iHbt}(\cdot)e^{iHbt}$ and $\mathcal{V}_2(\cdot) = e^{-iHa}(\cdot)e^{iHa}$ be two quantum channels for e^{-iHbt} and e^{-iHa}, respectively. Using the inequality between the diamond norm and the operator norm (see, e.g., [HKOT23, Proposition 1.6]) yields

$$\|\mathcal{V}_1 - \mathcal{V}_2\|_\diamond \leq 2\left\|e^{-iHbt} - e^{-iHa}\right\|, \tag{58}$$

where $\|\cdot\|$ is the operator norm. To further bound the operator norm on the RHS of Equation (58), we observe

$$\left\|e^{-iHbt} - e^{-iHa}\right\| = \left\|e^{-iHbt}\left(I - e^{-iH(a-bt)}\right)\right\| = \left\|I - e^{-iH(a-bt)}\right\|,$$

which can be upper bounded by $|a-bt|\cdot\|H\| \leq 0.01\cdot\pi$, through looking at the eigenvalues and using $\left|1-e^{ix}\right| \leq |x|$. Consequently, $\|\mathcal{V}_1-\mathcal{V}_2\|_\diamond \leq 0.02\cdot\pi < 0.1$. That is, $\mathcal{B}$ is also an algorithm to implement the unitary e^{-iHa} to within diamond norm error $b\epsilon + 0.02\cdot\pi < 0.2$.

Finally, if we use $a-1\geq 0$ more queries to U followed by algorithm $\mathcal{B}$, we obtain an algorithm $\mathcal{C}$ to implement the time-reverse U^{-1} (up to a global phase) to within diamond norm error 0.2. Note that algorithm $\mathcal{C}$ only uses $bn+a-1$ queries to U, and both a and b are constants. Therefore, by Corollary 3.2, we have $n=\Omega(d^2)$.

□

4 Deferred lemmas

4.1 Raising and lowering of Young diagrams

In this subsection, we present two representation-theoretic results concerning the raising and lowering of Young diagrams. These results follow from standard techniques involving the Clebsch-Gordan transform and Schur transform [BCH07, BCH06, Har05] (see also [YSM23, SMKH22]). However, for the reader's convenience, we provide simple and self-contained proofs here.

Lemma 4.1. *Let $n\geq 2$ be an integer and $\mathcal{H}_1,\mathcal{H}_2,\ldots,\mathcal{H}_n$ be a sequence of Hilbert spaces such that $\mathcal{H}_i\cong\mathbb{C}^d$ for $1\leq i\leq n$. Let $\mathfrak{S}_n$ act on $\bigotimes_{i=1}^n\mathcal{H}_i$ and consider the corresponding decomposition $\bigotimes_{i=1}^n\mathcal{H}_i \overset{\mathbb{U}_d\times\mathfrak{S}_n}{\cong} \bigoplus_{\lambda\vdash_d n}\mathcal{Q}_\lambda\otimes\mathcal{P}_\lambda$. Then, for $\mu\vdash_d n-1$, $T,S\in\mathrm{Tab}(\mu)$, we have*

$$I_{\mathcal{H}_n}\otimes I_{\mathcal{Q}_\mu}\otimes|T\rangle\langle S| = \bigoplus_{\lambda:\mu\nearrow\lambda} I_{\mathcal{Q}_\lambda}\otimes|T^{\uparrow\lambda}\rangle\langle S^{\uparrow\lambda}|,$$

where $T^{\uparrow\lambda}$ denotes the standard Young tableau obtained from T by adding a box filled with n, resulting in the shape λ, and we take the convention that $\mathcal{Q}_\lambda=0$ when $\ell(\lambda)>d$.

Proof. We consider the linear operator $I_{\mathcal{Q}_\mu}\otimes|T\rangle\in\mathcal{L}(\mathcal{Q}_\mu,\mathcal{Q}_\mu\otimes\mathcal{P}_\mu)\subseteq\mathcal{L}(\mathcal{Q}_\mu,\bigotimes_{i=1}^{n-1}\mathcal{H}_i)$. By Pieri's rule:

$$\mathbb{C}^d\otimes\mathcal{Q}_\mu \overset{\mathbb{U}_d}{\cong} \bigoplus_{\lambda:\mu\nearrow\lambda}\mathcal{Q}_\lambda,$$

we can see that

$$I_{\mathcal{H}_n}\otimes I_{\mathcal{Q}_\mu}\otimes|T\rangle = \bigoplus_{\lambda:\mu\nearrow\lambda} I_{\mathcal{Q}_\lambda}\otimes|T_\lambda\rangle, \tag{59}$$

where $|T_\lambda\rangle$ is a unit vector and by Schur-Weyl duality on $\bigotimes_{i=1}^n\mathcal{H}_i$, we know that $|T_\lambda\rangle\in\mathcal{P}_\lambda$. Suppose $T=\mu^{(1)}\to\cdots\to\mu^{(n-1)}$ where $\mu^{(n-1)}=\mu$. For an integer $1\leq k\leq n-1$, we apply $P(e_{\mu^{(k)}})$ (i.e., the action of $e_{\mu^{(k)}}$ on $\bigotimes_{i=1}^n\mathcal{H}_i$, where $e_{\mu^{(k)}}$ is defined in Equation (21)) to both sides of Equation (59). For the LHS, since $e_{\mu^{(k)}}\in\mathbb{C}\mathfrak{S}_k\subseteq\mathbb{C}\mathfrak{S}_{n-1}$, $P(e_{\mu^{(k)}})$ acts non-trivially only on $\bigotimes_{i=1}^{n-1}\mathcal{H}_i$, thus we have

$$P(e_{\mu^{(k)}})(I_{\mathcal{H}_n}\otimes I_{\mathcal{Q}_\mu}\otimes|T\rangle) = I_{\mathcal{H}_n}\otimes I_{\mathcal{Q}_\mu}\otimes e_{\mu^{(k)}}|T\rangle = I_{\mathcal{H}_n}\otimes I_{\mathcal{Q}_\mu}\otimes|T\rangle.$$

For the RHS, we have

$$P(e_{\mu^{(k)}})\left(\bigoplus_{\lambda:\mu\nearrow\lambda} I_{\mathcal{Q}_\lambda}\otimes|T_\lambda\rangle\right) = \bigoplus_{\lambda:\mu\nearrow\lambda} I_{\mathcal{Q}_\lambda}\otimes e_{\mu^{(k)}}|T_\lambda\rangle.$$

Therefore, we can see that

$$e_{\mu^{(k)}}|T_\lambda\rangle = |T_\lambda\rangle,$$

for any $1 \le k \le n-1$. This means $|T_\lambda\rangle \in \mathcal{P}_\lambda$, when restricted as a vector in a representation of $\mathfrak{S}_k$, is a vector in $\mathcal{P}_{\mu^{(k)}}$. Therefore, $|T_\lambda\rangle = |\mu^{(1)} \to \cdots \to \mu^{(n-1)} \to \lambda\rangle = |T^{\uparrow\lambda}\rangle$.

We can analogously prove $I_{\mathcal{H}_n} \otimes I_{\mathcal{Q}_\mu} \otimes \langle S| = \bigoplus_{\lambda:\mu\nearrow\lambda} I_{\mathcal{Q}_\lambda} \otimes \langle S^{\uparrow\lambda}|$. This completes the proof. □

The following result can be viewed, in some sense, as a "dual" of Lemma 4.1.

Lemma 4.2. *Let $n \ge 2$ be an integer and $\mathcal{H}_1, \mathcal{H}_2, \ldots, \mathcal{H}_n$ be a sequence of Hilbert spaces such that $\mathcal{H}_i \cong \mathbb{C}^d$ for $1 \le i \le n$. Let $\mathfrak{S}_n$ act on $\bigotimes_{i=1}^n \mathcal{H}_i$ and consider the corresponding decomposition $\bigotimes_{i=1}^n \mathcal{H}_i \overset{\mathbb{U}_d\times\mathfrak{S}_n}{\cong} \bigoplus_{\lambda\vdash_d n} \mathcal{Q}_\lambda \otimes \mathcal{P}_\lambda$. Then, for $\lambda \vdash_d n$, $T, S \in \mathrm{Tab}(\lambda)$, we have*

$$\mathrm{tr}_{\mathcal{H}_n}(I_{\mathcal{Q}_\lambda} \otimes |T\rangle\langle S|) = \mathbb{1}_{\mathrm{Sh}(T^\downarrow)=\mathrm{Sh}(S^\downarrow)} \cdot \frac{\dim(\mathcal{Q}_\lambda)}{\dim(\mathcal{Q}_{\mathrm{Sh}(T^\downarrow)})} \cdot I_{\mathcal{Q}_{\mathrm{Sh}(T^\downarrow)}} \otimes |T^\downarrow\rangle\langle S^\downarrow|,$$

where $T^\downarrow$ denotes the standard Young tableau obtained from T by removing the box containing the largest integer.

Proof. Let

$$W \coloneqq \mathrm{tr}_{\mathcal{H}_n}(I_{\mathcal{Q}_\lambda} \otimes |T\rangle\langle S|). \tag{60}$$

Note that $W \in \mathcal{L}(\bigotimes_{i=1}^{n-1} \mathcal{H}_i)$. For any $U \in \mathbb{U}_d$, we have

$$\begin{aligned} U^{\otimes n-1} W U^{\dagger\otimes n-1} &= \mathrm{tr}_{\mathcal{H}_n}\Big(U^{\otimes n}(I_{\mathcal{Q}_\lambda} \otimes |T\rangle\langle S|)U^{\dagger\otimes n}\Big) \\ &= \mathrm{tr}_{\mathcal{H}_n}(I_{\mathcal{Q}_\lambda} \otimes |T\rangle\langle S|) \\ &= \mathrm{tr}_{\mathcal{H}_n}(I_{\mathcal{Q}_\lambda} \otimes |T\rangle\langle S|). \end{aligned}$$

By Schur-Weyl duality on $\bigotimes_{i=1}^{n-1} \mathcal{H}_i$ and Schur's lemma, this means

$$W = \bigoplus_{\mu\vdash_d n-1} I_{\mathcal{Q}_\mu} \otimes X_\mu, \tag{61}$$

for some $X_\mu \in \mathcal{L}(\mathcal{P}_\mu)$. On the other hand, suppose $T = \lambda_T^{(1)} \to \lambda_T^{(2)} \to \cdots \to \lambda_T^{(n)}$ and $S = \lambda_S^{(1)} \to \lambda_S^{(2)} \to \cdots \to \lambda_S^{(n)}$, where $\lambda_T^{(n)} = \lambda_S^{(n)} = \lambda$. For $k \le n-1$ and $\nu, \tau \vdash k$, we left-apply $P(e_\nu)$ (i.e., the action of e_ν on $\bigotimes_{i=1}^{n-1} \mathcal{H}_i$, where e_ν is defined in Equation (21)) and right-apply $P(e_\tau)$ to W. First, by the definition of W (c.f. Equation (60)), we have

$$\begin{aligned} P(e_\nu) W P(e_\tau) &= \mathrm{tr}_{\mathcal{H}_n}(P(e_\nu)(I_{\mathcal{Q}_\lambda} \otimes |T\rangle\langle S|)P(e_\tau)) && (62) \\ &= \mathrm{tr}_{\mathcal{H}_n}(I_{\mathcal{Q}_\lambda} \otimes e_\nu |T\rangle\langle S| e_\tau) \\ &= \mathbb{1}_{\nu=\lambda_T^{(k)}} \cdot \mathbb{1}_{\tau=\lambda_S^{(k)}} \cdot \mathrm{tr}_{\mathcal{H}_n}(I_{\mathcal{Q}_\lambda} \otimes |T\rangle\langle S|) \\ &= \mathbb{1}_{\nu=\lambda_T^{(k)}} \cdot \mathbb{1}_{\tau=\lambda_S^{(k)}} \cdot W, && (63) \end{aligned}$$

where Equation (62) is because $e_\nu, e_\tau \in \mathbb{C}\mathfrak{S}_k \subseteq \mathbb{C}\mathfrak{S}_{n-1}$ acts non-trivially only on $\bigotimes_{i=1}^{n-1} \mathcal{H}_i$. Then, by recursively using Equation (63), we know that

$$W = P\Big(e_{\lambda_T^{(1)}}\Big) \cdots P\Big(e_{\lambda_T^{(n-1)}}\Big) \cdot W \cdot P\Big(e_{\lambda_S^{(n-1)}}\Big) \cdots P\Big(e_{\lambda_S^{(1)}}\Big).$$

On the other hand, by using Equation (61), we have

$$\begin{aligned} W &= P\Big(e_{\lambda_T^{(1)}}\Big)\cdots P\Big(e_{\lambda_T^{(n-1)}}\Big)\cdot W\cdot P\Big(e_{\lambda_S^{(n-1)}}\Big)\cdots P\Big(e_{\lambda_S^{(1)}}\Big) \\ &= \bigoplus_{\mu\vdash_d n-1} I_{\mathcal{Q}_\mu}\otimes\Big(e_{\lambda_T^{(1)}}\cdots e_{\lambda_T^{(n-1)}}\cdot X_\mu\cdot e_{\lambda_S^{(n-1)}}\cdots e_{\lambda_S^{(1)}}\Big) \\ &= \bigoplus_{\mu\vdash_d n-1} I_{\mathcal{Q}_\mu}\otimes\Big(|T^{\downarrow}\rangle\langle T^{\downarrow}|\cdot X_\mu\cdot|S^{\downarrow}\rangle\langle S^{\downarrow}|\Big), \end{aligned} \tag{64}$$

where Equation (64) is by using Equation (23). Note that $|T^{\downarrow}\rangle\langle T^{\downarrow}|$ is a linear projector on $\mathcal{P}_{\mathrm{Sh}(T^{\downarrow})}$ and $|S^{\downarrow}\rangle\langle S^{\downarrow}|$ is a linear projector on $\mathcal{P}_{\mathrm{Sh}(S^{\downarrow})}$. This means $|T^{\downarrow}\rangle\langle T^{\downarrow}|\cdot X_\mu\cdot|S^{\downarrow}\rangle\langle S^{\downarrow}|$ is non-zero only when $\mu=\mathrm{Sh}(T^{\downarrow})=\mathrm{Sh}(S^{\downarrow})$. Therefore, we have

$$W=\mathbb{1}_{\mathrm{Sh}(T^{\downarrow})=\mathrm{Sh}(S^{\downarrow})}\cdot c\cdot I_{\mathcal{Q}_{\mathrm{Sh}(T^{\downarrow})}}\otimes|T^{\downarrow}\rangle\langle S^{\downarrow}|, \tag{65}$$

for some number c.

Now suppose $\mathrm{Sh}(T^{\downarrow})=\mathrm{Sh}(S^{\downarrow})$. Thus both $|T^{\downarrow}\rangle$ and $|S^{\downarrow}\rangle$ are in the same irreducible representation $\mathcal{P}_{\mathrm{Sh}(T^{\downarrow})}$. Let $K\in\mathbb{C}\mathfrak{S}_{n-1}$ such that $K|T^{\downarrow}\rangle=|S^{\downarrow}\rangle$ (such K always exists by the Jacobson density theorem [EGH+11]). Then, K also satisfies $K|T\rangle=|S\rangle$. To see this, we view $|T\rangle$ and $|S\rangle$ as vectors in the restricted representation $\mathrm{Res}^{\mathfrak{S}_n}_{\mathfrak{S}_{n-1}}\mathcal{P}_\lambda$. By the definition of Young basis, they both are in the same irreducible representation $\mathcal{P}_{\mathrm{Sh}(T^{\downarrow})}$ and correspond to $|T^{\downarrow}\rangle$ and $|S^{\downarrow}\rangle$, respectively. Therefore, by the definition of K, K maps $|T\rangle$ to $|S\rangle$. Then, consider the trace $\mathrm{tr}(P(K)\cdot W)$, where $P(K)$ is the action of K on $\bigotimes_{i=1}^{n-1}\mathcal{H}_i$. On the one hand, by Equation (65), we have

$$\mathrm{tr}(P(K)\cdot W)=\mathrm{tr}\Big(c\cdot I_{\mathcal{Q}_{\mathrm{Sh}(T^{\downarrow})}}\otimes K|T^{\downarrow}\rangle\langle S^{\downarrow}|\Big)=\mathrm{tr}\Big(c\cdot I_{\mathcal{Q}_{\mathrm{Sh}(T^{\downarrow})}}\otimes|S^{\downarrow}\rangle\langle S^{\downarrow}|\Big)=c\cdot\dim(\mathcal{Q}_{\mathrm{Sh}(T^{\downarrow})}).$$

On the other hand, by Equation (60), we have

$$\mathrm{tr}(P(K)\cdot W)=\mathrm{tr}(\mathrm{tr}_{\mathcal{H}_n}(I_{\mathcal{Q}_\lambda}\otimes K|T\rangle\langle S|))=\mathrm{tr}(\mathrm{tr}_{\mathcal{H}_n}(I_{\mathcal{Q}_\lambda}\otimes|S\rangle\langle S|))=\dim(\mathcal{Q}_\lambda).$$

Therefore, $c=\dim(\mathcal{Q}_\lambda)/\dim(\mathcal{Q}_{\mathrm{Sh}(T^{\downarrow})})$, completing the proof upon substitution into Equation (65). □

4.2 Combinatorics on Young diagrams

Here, we introduce some combinatorial results on Young diagrams. The first result is by a direct calculation using the hook length formula.

Lemma 4.3. *Suppose $\mu,\nu\vdash n$, $\mu\neq\nu$ and μ,ν are adjacent. Then*

$$\frac{\dim(\mathcal{P}_\mu)\dim(\mathcal{P}_\nu)}{\dim(\mathcal{P}_{\mu\cup\nu})\dim(\mathcal{P}_{\mu\cap\nu})}=\frac{n}{n+1}\left(1-\frac{1}{(c(\mu\setminus\nu)-c(\nu\setminus\mu))^2}\right),$$

where $\mu\setminus\nu$ contains only one box and $c(\mu\setminus\nu)$ is the axial coordinate of this box (and similarly for $\nu\setminus\mu$).

Proof. Let $\square_\mu$ be the box in $\mu\setminus\nu$ and $\square_\nu$ be the box in $\nu\setminus\mu$. Define $\tau:=\mu\cap\nu$. Denote by $H_\lambda(\square)$ the set of boxes in λ that are directly to the left and above the box $\square$, including $\square$ itself (i.e., a hook in the reverse direction). By the hook length formula (see Equation (14)), we can easily see that

$$\frac{\dim(\mathcal{P}_\mu)}{\dim(\mathcal{P}_\tau)}=n\cdot\prod_{\square\in H_\tau(\square_\mu)}\frac{h_\tau(\square)}{h_\tau(\square)+1}.$$

Similarly

$$\frac{\dim(\mathcal{P}_\nu)}{\dim(\mathcal{P}_\tau)} = n \cdot \prod_{\square \in H_\tau(\square_\nu)} \frac{h_\tau(\square)}{h_\tau(\square)+1}.$$

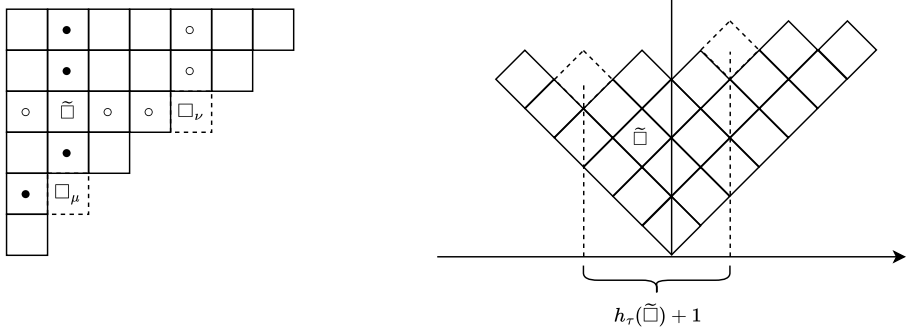
Then, note that $\square_\mu, \square_\nu$ must be at different rows and columns. Thus, there is exactly one box in $H_\tau(\square_\mu) \cap H_\tau(\square_\nu)$, which we denote by $\widetilde{\square}$. Then, we have

$$\frac{\dim(\mathcal{P}_{\mu\cup\nu})}{\dim(\mathcal{P}_\tau)} = n(n+1) \cdot \left(\prod_{\substack{\square \in H_\tau(\square_\mu) \\ \square \neq \widetilde{\square}}} \frac{h_\tau(\square)}{h_\tau(\square)+1} \right) \cdot \left(\prod_{\substack{\square \in H_\tau(\square_\nu) \\ \square \neq \widetilde{\square}}} \frac{h_\tau(\square)}{h_\tau(\square)+1} \right) \cdot \left(\frac{h_\tau(\widetilde{\square})}{h_\tau(\widetilde{\square})+2} \right).$$

Then, we have

$$\begin{aligned}
\frac{\dim(\mathcal{P}_\mu)\dim(\mathcal{P}_\nu)}{\dim(\mathcal{P}_{\mu\cup\nu})\dim(\mathcal{P}_\tau)} &= \frac{n}{n+1} \cdot \left(\frac{h_\tau(\widetilde{\square})}{h_\tau(\widetilde{\square})+1} \right)^2 \cdot \frac{h_\tau(\widetilde{\square})+2}{h_\tau(\widetilde{\square})} \\
&= \frac{n}{n+1} \cdot \left(1 - \frac{1}{(h_\tau(\widetilde{\square})+1)^2} \right) \\
&= \frac{n}{n+1} \left(1 - \frac{1}{(c(\square_\mu) - c(\square_\nu))^2} \right), \qquad (66)
\end{aligned}$$

where Equation (66) is because $h_\tau(\widetilde{\square}) + 1 = |c(\square_\mu) - c(\square_\nu)|$ (see, e.g., Figure 5b). □



(a) The sets $H_\tau(\square_\mu)$ and $H_\tau(\square_\nu)$. (b) The axial distance between $\square_\mu$ and $\square_\nu$.

Figure 5: An example illustrating the idea in the proof of Lemma 4.3.

Then, we introduce the following results. While Lemma 4.4 and Lemma 4.5 are known in the literature (see, e.g., [Kos03]), we provide different elementary proofs here for completeness. Our proofs are based on Kerov's interlacing sequences [Ker93].

Lemma 4.4. *Suppose $\nu \vdash n$, $\tau \nearrow \nu$, $\kappa \nearrow \nu$, $\tau \neq \kappa$. Then,*

$$\sum_{\lambda : \nu \nearrow \lambda} \frac{\dim(\mathcal{P}_\lambda)}{\dim(\mathcal{P}_\nu)} \frac{1}{c(\lambda \setminus \nu) - c(\nu \setminus \tau)} \frac{1}{c(\lambda \setminus \nu) - c(\nu \setminus \kappa)} = 0,$$

where $\lambda \setminus \nu$ contains only one box and $c(\lambda \setminus \nu)$ is the axial coordinate of this box (and similarly for $\nu \setminus \tau$ and $\nu \setminus \kappa$).

Proof. Note that

$$\sum_{\lambda:\nu\nearrow\lambda} \frac{\dim(\mathcal{P}_\lambda)}{\dim(\mathcal{P}_\nu)} \frac{1}{c(\lambda\setminus\nu)-c(\nu\setminus\tau)} \frac{1}{c(\lambda\setminus\nu)-c(\nu\setminus\kappa)}$$
$$=\frac{1}{c(\nu\setminus\tau)-c(\nu\setminus\kappa)} \sum_{\lambda:\nu\nearrow\lambda} \frac{\dim(\mathcal{P}_\lambda)}{\dim(\mathcal{P}_\nu)} \left(\frac{1}{c(\lambda\setminus\nu)-c(\nu\setminus\tau)} - \frac{1}{c(\lambda\setminus\nu)-c(\nu\setminus\kappa)}\right)$$

Therefore, it suffices to prove that for any $\tau \nearrow \nu$, we have

$$\sum_{\lambda:\nu\nearrow\lambda} \frac{\dim(\mathcal{P}_\lambda)}{\dim(\mathcal{P}_\nu)} \frac{1}{c(\lambda\setminus\nu)-c(\nu\setminus\tau)} = 0.$$

Let $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_L)$ and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_{L-1})$ be the interlacing sequences of ν. The box in $\nu\setminus\tau$ is a removable box of ν. Thus this box is at β_m for some m such that $\beta_m = c(\nu\setminus\tau)$. Any λ such that $\nu \nearrow \lambda$ is obtained by adding a box at an addable position of ν, which corresponds to an α_i for $1 \le i \le L$. Furthermore, for the λ obtained by adding a box at α_i, we know $c(\lambda\setminus\nu) = \alpha_i$. Then, by Equation (15), we have

$$\sum_{\lambda:\nu\nearrow\lambda} \frac{\dim(\mathcal{P}_\lambda)}{\dim(\mathcal{P}_\nu)} \frac{1}{c(\lambda\setminus\nu)-c(\nu\setminus\tau)} = (n+1)\cdot \sum_{i=1}^{L} \frac{1}{\alpha_i-\beta_m} \prod_{j=1}^{i-1} \frac{\alpha_i-\beta_j}{\alpha_i-\alpha_j} \prod_{j=i+1}^{L} \frac{\alpha_i-\beta_{j-1}}{\alpha_i-\alpha_j},$$

which is 0 by Lemma 4.6. □

Lemma 4.5. *Suppose $\nu \vdash n$, $\tau \nearrow \nu$. Then*

$$\sum_{\lambda:\nu\nearrow\lambda} \frac{\dim(\mathcal{P}_\lambda)}{\dim(\mathcal{P}_\nu)} \frac{1}{(c(\lambda\setminus\nu)-c(\nu\setminus\tau))^2} = \frac{n+1}{n}\cdot\frac{\dim(\mathcal{P}_\nu)}{\dim(\mathcal{P}_\tau)}, \tag{67}$$

where $\lambda\setminus\nu$ contains only one box and $c(\lambda\setminus\nu)$ is the axial coordinate of this box (and similarly for $\nu\setminus\tau$).

Proof. Let $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_L)$ and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_{L-1})$ be the interlacing sequences of ν. The box in $\nu\setminus\tau$ is a removable box of ν. Thus this box is at β_m for some m such that $\beta_m = c(\nu\setminus\tau)$. Any λ such that $\nu \nearrow \lambda$ is obtained by adding a box at an addable position of ν, which corresponds to an α_i for $1 \le i \le L$. Furthermore, for the λ obtained by adding a box at α_i, we know $c(\lambda\setminus\nu) = \alpha_i$. Then, by Equation (15) (applied to $\frac{\dim(\mathcal{P}_\lambda)}{\dim(\mathcal{P}_\nu)}$) and Equation (16) (applied to $\frac{\dim(\mathcal{P}_\tau)}{\dim(\mathcal{P}_\nu)}$), Equation (67) is equivalent to

$$\sum_{i=1}^{L} \frac{1}{(\alpha_i-\beta_m)^2} \prod_{j=1}^{i-1} \frac{\alpha_i-\beta_j}{\alpha_i-\alpha_j} \prod_{j=i+1}^{L} \frac{\alpha_i-\beta_{j-1}}{\alpha_i-\alpha_j} = \frac{1}{(\alpha_L-\beta_m)(\beta_m-\alpha_1)} \prod_{j=1}^{m-1} \frac{\beta_m-\beta_j}{\beta_m-\alpha_{j+1}} \prod_{j=m+1}^{L-1} \frac{\beta_m-\beta_j}{\beta_m-\alpha_j}. \tag{68}$$

Then, we prove Equation (68) by induction on L. For the case $L = 2$, m must be 1, and the holding of Equation (68) can be checked by direct calculation. Now, suppose Equation (68) holds for L. We want to prove it also holds for $L+1$, i.e., the following holds:

$$\sum_{i=1}^{L+1} \frac{1}{(\alpha_i-\beta_m)^2} \prod_{j=1}^{i-1} \frac{\alpha_i-\beta_j}{\alpha_i-\alpha_j} \prod_{j=i+1}^{L+1} \frac{\alpha_i-\beta_{j-1}}{\alpha_i-\alpha_j} = \frac{1}{(\alpha_{L+1}-\beta_m)(\beta_m-\alpha_1)} \prod_{j=1}^{m-1} \frac{\beta_m-\beta_j}{\beta_m-\alpha_{j+1}} \prod_{j=m+1}^{L} \frac{\beta_m-\beta_j}{\beta_m-\alpha_j}. \tag{69}$$

For the case of $L+1$, we can assume without loss of generality that $m \le L-1$. This is because if $m = L$, then we consider the reversed interlacing sequences $\boldsymbol{\alpha}' = (\alpha'_1, \dots, \alpha'_{L+1})$, $\boldsymbol{\beta}' = (\beta'_1, \dots, \beta'_L)$ such that $\alpha'_i = -\alpha_{L+2-i}$ and $\beta'_i = -\beta_{L+1-i}$. Then, it is easy to see that the holding of Equation (69) on $\boldsymbol{\alpha}, \boldsymbol{\beta}$ with $m = L$ is equivalent to that on $\boldsymbol{\alpha}'$, $\boldsymbol{\beta}'$ with $m = 1$.

It is easy to see that

$$\text{RHS of Equation (69)} - \left(\frac{\beta_m - \beta_L}{\beta_m - \alpha_{L+1}} \times \text{RHS of Equation (68)} \right) = 0.$$

By the induction hypothesis, it suffices to prove

$$\text{LHS of Equation (69)} - \left(\frac{\beta_m - \beta_L}{\beta_m - \alpha_{L+1}} \times \text{LHS of Equation (68)} \right) = 0. \tag{70}$$

Then, note that the LHS of Equation (69) can be written as

$$\sum_{i=1}^{L} \frac{\alpha_i - \beta_L}{(\alpha_i - \beta_m)^2 \cdot (\alpha_i - \alpha_{L+1})} \prod_{j=1}^{i-1} \frac{\alpha_i - \beta_j}{\alpha_i - \alpha_j} \prod_{j=i+1}^{L} \frac{\alpha_i - \beta_{j-1}}{\alpha_i - \alpha_j} + \frac{1}{(\alpha_{L+1} - \beta_m)^2} \prod_{j=1}^{L} \frac{\alpha_{L+1} - \beta_j}{\alpha_{L+1} - \alpha_j}.$$

Therefore, the LHS of Equation (70) is:

$$\begin{aligned}
&\sum_{i=1}^{L} \frac{\beta_L - \alpha_{L+1}}{(\alpha_i - \beta_m)(\beta_m - \alpha_{L+1})(\alpha_i - \alpha_{L+1})} \prod_{j=1}^{i-1} \frac{\alpha_i - \beta_j}{\alpha_i - \alpha_j} \prod_{j=i+1}^{L} \frac{\alpha_i - \beta_{j-1}}{\alpha_i - \alpha_j} + \frac{1}{(\alpha_{L+1} - \beta_m)^2} \prod_{j=1}^{L} \frac{\alpha_{L+1} - \beta_j}{\alpha_{L+1} - \alpha_j} \\
=&\frac{\beta_L - \alpha_{L+1}}{(\beta_m - \alpha_{L+1})^2} \left(\sum_{i=1}^{L} \left(\frac{1}{\alpha_i - \beta_m} - \frac{1}{\alpha_i - \alpha_{L+1}} \right) \prod_{j=1}^{i-1} \frac{\alpha_i - \beta_j}{\alpha_i - \alpha_j} \prod_{j=i+1}^{L} \frac{\alpha_i - \beta_{j-1}}{\alpha_i - \alpha_j} - \frac{1}{\alpha_{L+1} - \alpha_L} \prod_{j=1}^{L-1} \frac{\alpha_{L+1} - \beta_j}{\alpha_{L+1} - \alpha_j} \right) \\
=&- \frac{\beta_L - \alpha_{L+1}}{(\beta_m - \alpha_{L+1})^2} \left(\sum_{i=1}^{L} \frac{1}{\alpha_i - \alpha_{L+1}} \prod_{j=1}^{i-1} \frac{\alpha_i - \beta_j}{\alpha_i - \alpha_j} \prod_{j=i+1}^{L} \frac{\alpha_i - \beta_{j-1}}{\alpha_i - \alpha_j} + \frac{1}{\alpha_{L+1} - \alpha_L} \prod_{j=1}^{L-1} \frac{\alpha_{L+1} - \beta_j}{\alpha_{L+1} - \alpha_j} \right) \qquad (71) \\
=&- \frac{\beta_L - \alpha_{L+1}}{(\beta_m - \alpha_{L+1})^2} \left(\sum_{i=1}^{L+1} \frac{1}{\alpha_i - \beta_L} \prod_{j=1}^{i-1} \frac{\alpha_i - \beta_j}{\alpha_i - \alpha_j} \prod_{j=i+1}^{L+1} \frac{\alpha_i - \beta_{j-1}}{\alpha_i - \alpha_j} \right) \\
=&\, 0, \qquad (72)
\end{aligned}$$

where Equation (71) is by Lemma 4.6 and Equation (72) is also by Lemma 4.6. □

4.3 Auxiliary lemmas

The following Lemma 4.6 is repeatedly used in the proofs of Lemma 4.4 and Lemma 4.5.

Lemma 4.6. *Suppose $\alpha_1, \dots, \alpha_L, \beta_1, \dots, \beta_{L-1}$ are pairwise distinct. Let $1 \le m \le L-1$. Then we have*

$$\sum_{i=1}^{L} \frac{1}{\alpha_i - \beta_m} \prod_{j=1}^{i-1} \frac{\alpha_i - \beta_j}{\alpha_i - \alpha_j} \prod_{j=i+1}^{L} \frac{\alpha_i - \beta_{j-1}}{\alpha_i - \alpha_j} = 0. \tag{73}$$

Proof. Note that the LHS of Equation (73) can be written as

$$\sum_{i=1}^{L} \prod_{\substack{j=1 \\ j \ne m}}^{L-1} (\alpha_i - \beta_j) \prod_{\substack{j=1 \\ j \ne i}}^{L} \frac{1}{\alpha_i - \alpha_j}.$$

Then we multiply $\prod_{j=2}^{L}(\alpha_1-\alpha_j)$ on it and obtain

$$\prod_{\substack{j=1\\j\neq m}}^{L-1}(\alpha_1-\beta_j)-\sum_{i=2}^{L}\prod_{\substack{j=2\\j\neq i}}^{L}(\alpha_1-\alpha_j)\cdot\prod_{\substack{j=1\\j\neq m}}^{L-1}(\alpha_i-\beta_j)\prod_{\substack{j=2\\j\neq i}}^{L}\frac{1}{\alpha_i-\alpha_j}. \tag{74}$$

Then, we consider it as a polynomial in variable α_1 with parameters $\alpha_2,\ldots,\alpha_L,\beta_1,\ldots,\beta_{L-1}$. This polynomial is of degree $L-2$. Then, if we set $\alpha_1=\alpha_k$ for a $2\leq k\leq L$, we can see that Equation (74) becomes

$$\begin{aligned}
&\prod_{\substack{j=1\\j\neq m}}^{L-1}(\alpha_k-\beta_j)-\prod_{\substack{j=2\\j\neq k}}^{L}(\alpha_k-\alpha_j)\cdot\prod_{\substack{j=1\\j\neq m}}^{L-1}(\alpha_k-\beta_j)\prod_{\substack{j=2\\j\neq k}}^{L}\frac{1}{\alpha_k-\alpha_j}\\
=&\prod_{\substack{j=1\\j\neq m}}^{L-1}(\alpha_k-\beta_j)-\prod_{\substack{j=1\\j\neq m}}^{L-1}(\alpha_k-\beta_j)\\
=&0.
\end{aligned}$$

Therefore, this polynomial has $L-1$ roots $\alpha_2,\ldots,\alpha_L$, which means it must be the 0 polynomial. □

The following Lemma 4.7 provides a representation-theoretic expression of the Haar moment. This result was also used in the study of entanglement and bipartite quantum state (see, e.g., [MH07, CWZ24]).

Lemma 4.7. *Let C_U be that defined in Definition 3.4. We have*

$$\mathop{\mathbf{E}}_{U}[C_U]=\bigoplus_{\lambda\vdash_d n+1}\frac{1}{\dim(\mathcal{Q}_\lambda)}I_{\mathcal{Q}_\lambda}\otimes I_{\mathcal{Q}_\lambda}\otimes|I_{\mathcal{P}_\lambda}\rangle\!\rangle\langle\!\langle I_{\mathcal{P}_\lambda}|.$$

Proof. Note that $C_U=|U\rangle\!\rangle\langle\!\langle U|^{\otimes n+1}$ is an $(n+1)$-comb on $(\mathcal{H}_1,\mathcal{H}_2,\ldots,\mathcal{H}_{2n+2})$. We can easily check the following facts.

Fact 4.8. $\mathbf{E}_U[C_U]$ *commutes with the action of the group $\mathbb{U}_d\times\mathbb{U}_d$, i.e., for $(V,W)\in\mathbb{U}_d\times\mathbb{U}_d$:*

$$(V\otimes W)^{\otimes n+1}\mathop{\mathbf{E}}_{U}[C_U]=\mathop{\mathbf{E}}_{U}[C_U](V\otimes W)^{\otimes n+1}.$$

Proof of Fact 4.8.

$$\begin{aligned}
(V\otimes W)^{\otimes n+1}\mathop{\mathbf{E}}_{U}[C_U]&=\mathop{\mathbf{E}}_{U}\left[(V\otimes W|U\rangle\!\rangle\langle\!\langle U|)^{\otimes n+1}\right]\\
&=\mathop{\mathbf{E}}_{U}\left[|VUW^{\mathrm{T}}\rangle\!\rangle\langle\!\langle U|^{\otimes n+1}\right]\\
&=\mathop{\mathbf{E}}_{U'}\left[|U'\rangle\!\rangle\langle\!\langle V^\dagger U'W^*|^{\otimes n+1}\right]\qquad(75)\\
&=\mathop{\mathbf{E}}_{U'}\left[(|U'\rangle\!\rangle\langle\!\langle U'|V\otimes W)^{\otimes n+1}\right]\\
&=\mathop{\mathbf{E}}_{U}[C_U](V\otimes W)^{\otimes n+1},
\end{aligned}$$

where Equation (75) is because U and U' are Haar random unitary. □

Fact 4.9. *$\mathbf{E}_U[C_U]$ is invariant under simultaneous permutation, i.e., for any $\pi \in \mathfrak{S}_{n+1}$, let $P^{\mathrm{odd}}(\pi)$ be the action of π on $\mathcal{H}_1 \otimes \mathcal{H}_3 \otimes \cdots \mathcal{H}_{2n+1} \cong (\mathbb{C}^d)^{\otimes n+1}$ and $P^{\mathrm{even}}(\pi)$ be that on $\mathcal{H}_2 \otimes \mathcal{H}_4 \otimes \cdots \mathcal{H}_{2n+2} \cong (\mathbb{C}^d)^{\otimes n+1}$, then*

$$P^{\mathrm{even}}(\pi) \otimes P^{\mathrm{odd}}(\pi) \mathop{\mathbf{E}}_U[C_U] = \mathop{\mathbf{E}}_U[C_U] P^{\mathrm{even}}(\pi) \otimes P^{\mathrm{odd}}(\pi) = \mathop{\mathbf{E}}_U[C_U].$$

Proof of Fact 4.9.

$$P^{\mathrm{even}}(\pi) \otimes P^{\mathrm{odd}}(\pi) \mathop{\mathbf{E}}_U[C_U] = \mathop{\mathbf{E}}_U\Big[P^{\mathrm{even}}(\pi) \otimes P^{\mathrm{odd}}(\pi) |U\rangle\!\rangle\langle\!\langle U|^{\otimes n+1}\Big] = \mathop{\mathbf{E}}_U\big[|U\rangle\!\rangle\langle\!\langle U|^{\otimes n+1}\big] = \mathop{\mathbf{E}}_U[C_U].$$

The other direction is similar. □

Note that we have the following decomposition of the space

$$(\mathcal{H}_1 \otimes \mathcal{H}_3 \otimes \cdots \otimes \mathcal{H}_{2n+1}) \otimes (\mathcal{H}_2 \otimes \mathcal{H}_4 \otimes \cdots \otimes \mathcal{H}_{2n+2}) \cong (\mathbb{C}^d)^{\otimes n+1} \otimes (\mathbb{C}^d)^{\otimes n+1}$$
$$\overset{\mathbb{U}_d \times \mathbb{U}_d \times \mathfrak{S}_{n+1} \times \mathfrak{S}_{n+1}}{\cong} \bigoplus_{\lambda,\mu \vdash_d n+1} \mathcal{Q}_\lambda \otimes \mathcal{Q}_\mu \otimes \mathcal{P}_\lambda \otimes \mathcal{P}_\mu.$$

Therefore, by Fact 4.8 and Schur's lemma, we have

$$\mathop{\mathbf{E}}_U[C_U] = \bigoplus_{\lambda,\mu \vdash_d n+1} I_{\mathcal{Q}_\lambda} \otimes I_{\mathcal{Q}_\mu} \otimes M_{\lambda,\mu},$$

for some $M_{\lambda,\mu} \in \mathcal{L}(\mathcal{P}_\lambda \otimes \mathcal{P}_\mu)$.

On the other hand, letting $P_{\mathrm{avg}} = \frac{1}{(n+1)!} \sum_{\pi \in \mathfrak{S}_{n+1}} P^{\mathrm{even}}(\pi) \otimes P^{\mathrm{odd}}(\pi)$, we have

$$P_{\mathrm{avg}} = \bigoplus_{\lambda,\mu \vdash_d n+1} I_{\mathcal{Q}_\lambda} \otimes I_{\mathcal{Q}_\mu} \otimes \frac{1}{(n+1)!} \sum_{\pi \in \mathfrak{S}_{n+1}} P_\lambda(\pi) \otimes P_\mu(\pi) \tag{76}$$

$$= \bigoplus_{\lambda,\mu \vdash_d n+1} I_{\mathcal{Q}_\lambda} \otimes I_{\mathcal{Q}_\mu} \otimes \frac{1}{(n+1)!} \sum_{\pi \in \mathfrak{S}_{n+1}} P_\lambda(\pi) \otimes P^*_\mu(\pi) \tag{77}$$

$$= \bigoplus_{\lambda \vdash_d n+1} I_{\mathcal{Q}_\lambda} \otimes I_{\mathcal{Q}_\lambda} \otimes \frac{1}{\dim(\mathcal{P}_\lambda)} |I_{\mathcal{P}_\lambda}\rangle\!\rangle\langle\!\langle I_{\mathcal{P}_\lambda}|, \tag{78}$$

where $P_\lambda(\pi)$ is the action of π on $\mathcal{P}_\lambda$, Equation (77) is because $P_\mu(\pi)$ is a real-valued matrix on the Young basis, Equation (78) is because the subspace in $\mathcal{P}_\lambda \otimes \mathcal{P}_\mu$ that is invariant under $P_\lambda(\pi) \otimes P^*_\mu(\pi)$ is $\mathrm{span}(|I_{\mathcal{P}_\lambda}\rangle\!\rangle)$ when $\lambda = \mu$ and the trivial subspace $\{0\}$ when $\lambda \neq \mu$.

Then, Fact 4.9 implies that $P_{\mathrm{avg}} \mathbf{E}_U[C_U] P_{\mathrm{avg}} = \mathbf{E}_U[C_U]$, which means

$$\mathop{\mathbf{E}}_U[C_U] = \bigoplus_{\lambda \vdash_d n+1} z_\lambda \cdot I_{\mathcal{Q}_\lambda} \otimes I_{\mathcal{Q}_\lambda} \otimes |I_{\mathcal{P}_\lambda}\rangle\!\rangle\langle\!\langle I_{\mathcal{P}_\lambda}|,$$

for some $z_\lambda \in \mathbb{C}$.

Moreover, since $\mathrm{tr}_{2,4,\ldots,2n+2}(C_U) = I^{\otimes n+1}$, we have

$$I^{\otimes n+1} = \bigoplus_{\lambda \vdash_d n+1} z_\lambda \cdot \mathrm{tr}_{\mathcal{Q}_\lambda}(I_{\mathcal{Q}_\lambda}) \cdot I_{\mathcal{Q}_\lambda} \otimes \mathrm{tr}_{\mathcal{P}_\lambda}(|I_{\mathcal{P}_\lambda}\rangle\!\rangle\langle\!\langle I_{\mathcal{P}_\lambda}|) \tag{79}$$

$$= \bigoplus_{\lambda \vdash_d n+1} z_\lambda \cdot \dim(\mathcal{Q}_\lambda) \cdot I_{\mathcal{Q}_\lambda} \otimes I_{\mathcal{P}_\lambda}, \tag{80}$$

which means $z_\lambda = \frac{1}{\dim(\mathcal{Q}_\lambda)}$. Therefore,

$$\mathop{\mathbf{E}}_{U}[C_U] = \bigoplus_{\lambda \vdash_d n+1} \frac{1}{\dim(\mathcal{Q}_\lambda)} \cdot I_{\mathcal{Q}_\lambda} \otimes I_{\mathcal{Q}_\lambda} \otimes |I_{\mathcal{P}_\lambda}\rangle\!\rangle\langle\!\langle I_{\mathcal{P}_\lambda}|.$$

□

Lemma 4.10. *For any positive semidefinite matrix M and vector $|\psi\rangle$ such that $|\psi\rangle \in \mathrm{supp}(M)$, we have*

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

where M^{-1} is the pseudo-inverse of M.

Proof.

$$M \sqsupseteq |\psi\rangle\langle\psi| \iff I_{\mathrm{supp}(M)} \sqsupseteq M^{-1/2}|\psi\rangle\langle\psi|M^{-1/2},$$

where $M^{-1/2}$ is the pseudo-inverse of $M^{1/2}$. Then

$$I_{\mathrm{supp}(M)} \sqsupseteq M^{-1/2}|\psi\rangle\langle\psi|M^{-1/2} \iff 1 \geq \mathrm{tr}(M^{-1/2}|\psi\rangle\langle\psi|M^{-1/2}) = \langle\psi|M^{-1}|\psi\rangle.$$

□

Acknowledgment

We thank John Wright and Ewin Tang for pointing out an error in our previous result on the hardness of unitary controlization and for directing us to the relevant reference [SMM09]. We also thank Qisheng Wang for helpful discussions. Finally, we thank the anonymous reviewers for their valuable comments and suggestions.

The work of Zhicheng Zhang was supported in part by the Australian Research Council under Grant DP250102952.

References

[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.

[Amb00] Andris Ambainis. Quantum lower bounds by quantum arguments. In *Proceedings of the thirty-second annual ACM symposium on Theory of computing*, pages 636–643, 2000.

[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.

[BYMQ25] Vanessa Brzić, Satoshi Yoshida, Mio Murao, and Marco Túlio Quintino. Higher-order quantum computing with known input states. *arXiv preprint arXiv:2510.20530*, 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.

[CML+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.

[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.

[EGH+11] Pavel I Etingof, Oleg Golberg, Sebastian Hensel, Tiankai Liu, Alex Schwendner, Dmitry Vaintrob, and Elena Yudovina. *Introduction to representation theory*, volume 59. American Mathematical Soc., 2011.

[EHM+23] Daniel Ebler, Michał Horodecki, Marcin Marciniak, Tomasz Młynik, Marco Túlio Quintino, and Michał Studziński. Optimal universal quantum circuits for unitary complex conjugation. *IEEE Transactions on Information Theory*, 69(8):5069–5082, 2023.

[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.

[Ful97] William Fulton. *Young tableaux: with applications to representation theory and geometry*. Number 35 in London Mathematical Society Student Texts. Cambridge University Press, 1997.

[FVDG02] Christopher A Fuchs and Jeroen Van De Graaf. Cryptographic distinguishability measures for quantum-mechanical states. *IEEE transactions on information theory*, 45(4):1216–1227, 2002.

[GO24] Dmitry Grinko and Maris Ozols. Linear programming with unitary-equivariant constraints. *Communications in Mathematical Physics*, 405(12):278, 2024.

[GP22] András Gilyén and Alexander Poremba. Improved quantum algorithms for fidelity estimation. *arXiv preprint arXiv:2203.15993*, 2022.

[GST24] Zuzana Gavorová, Matan Seidel, and Yonathan Touati. Topological obstructions to quantum computation with unitary oracles. *Physical Review A*, 109(3):032625, 2024.

[GYMO25] Dmitry Grinko, Satoshi Yoshida, Mio Murao, and Maris Ozols. Sequential quantum processes with group symmetries. *arXiv preprint arXiv:2510.07100*, 2025.

[Har05] Aram W Harrow. Applications of coherent classical communication and the schur transform to quantum information theory. *arXiv preprint quant-ph/0512255*, 2005.

[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.

[HLB+24] Hsin-Yuan Huang, Yunchao Liu, Michael Broughton, Isaac Kim, Anurag Anshu, Zeph Landau, and Jarrod R McClean. Learning shallow quantum circuits. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pages 1343–1351, 2024.

[Ker93] Sergei V Kerov. Transition probabilities for continual young diagrams and the markov moment problem. *Functional Analysis and its Applications*, 27(2):104–117, 1993.

[Ker00] Sergei Vasil'evich Kerov. Anisotropic young diagrams and jack symmetric functions. *Functional Analysis and Its Applications*, 34:41–51, 2000.

[Kit95] A Yu Kitaev. Quantum measurements and the abelian stabilizer problem. *arXiv preprint quant-ph/9511026*, 1995.

[Kos03] Masashi Kosuda. A new proof for some relations among axial distances and hook-lengths. *Tokyo Journal of Mathematics*, 26(1):199–228, 2003.

[Mél17] Pierre-Loïc Méliot. *Representation theory of symmetric groups*. Chapman and Hall/CRC, 2017.

[MGE11] Easwar Magesan, Jay M Gambetta, and Joseph Emerson. Scalable and robust randomized benchmarking of quantum processes. *Physical review letters*, 106(18):180504, 2011.

[MGE12] Easwar Magesan, Jay M Gambetta, and Joseph Emerson. Characterizing quantum gates via randomized benchmarking. *Physical Review A—Atomic, Molecular, and Optical Physics*, 85(4):042311, 2012.

[MH07] Keiji Matsumoto and Masahito Hayashi. Universal distortion-free entanglement concentration. *Physical Review A—Atomic, Molecular, and Optical Physics*, 75(6):062338, 2007.

[MH25] Fermi Ma and Hsin-Yuan Huang. How to construct random unitaries. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, pages 806–809, 2025.

[MLW25] Yin Mo, Tengxiang Lin, and Xin Wang. Efficient inversion of unknown unitary operations with structured hamiltonians. *arXiv preprint arXiv:2506.20570*, 2025.

[MSM19] Jisho Miyazaki, Akihito Soeda, and Mio Murao. Complex conjugation supermap of unitary quantum maps and its universal implementation protocol. *Physical Review Research*, 1(1):013007, 2019.

[MVBS04] Mikko Möttönen, Juha J. Vartiainen, Ville Bergholm, and Martti M. Salomaa. Quantum circuits for general multiqubit gates. *Physical Review Letters*, 93(13), September 2004.

[MW16] Ashley Montanaro and Ronald de Wolf. *A Survey of Quantum Property Testing*. Number 7 in Graduate Surveys. Theory of Computing Library, 2016.

[MZC+25] Yin Mo, Lei Zhang, Yu-Ao Chen, Yingjian Liu, Tengxiang Lin, and Xin Wang. Parameterized quantum comb and simpler circuits for reversing unknown qubit-unitary operations. *npj Quantum Information*, 11(1):32, 2025.

[Nav18] Miguel Navascués. Resetting uncontrolled quantum systems. *Physical Review X*, 8(3):031008, 2018.

[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.

[OV96] Andrei Okounkov and Anatoly Vershik. A new approach to representation theory of symmetric groups. *Selecta Mathematica*, 2(4):581, 1996.

[OYM25] Tatsuki Odake, Satoshi Yoshida, and Mio Murao. Analytical lower bound on query complexity for transformations of unknown unitary operations. *Physical Review Letters*, 135(23):230603, 2025.

[QDS+19a] Marco Túlio Quintino, Qingxiuxiong Dong, Atsushi Shimbo, Akihito Soeda, and Mio Murao. Probabilistic exact universal quantum circuits for transforming unitary operations. *Physical Review A*, 100(6):062339, 2019.

[QDS+19b] Marco Túlio Quintino, Qingxiuxiong Dong, Atsushi Shimbo, Akihito Soeda, and Mio Murao. Reversing unknown quantum transformations: Universal quantum circuit for inverting general unitary operations. *Physical Review Letters*, 123(21):210502, 2019.

[QE22] Marco Túlio Quintino and Daniel Ebler. Deterministic transformations between unitary operations: Exponential advantage with adaptive quantum circuits and the power of indefinite causality. *Quantum*, 6:679, 2022.

[Rag01] Maxim Raginsky. A fidelity measure for quantum channels. *Physics Letters A*, 290(1-2):11–18, 2001.

[SBSSH16] Brian Swingle, Gregory Bentsen, Monika Schleier-Smith, and Patrick Hayden. Measuring the scrambling of quantum information. *Physical Review A*, 94(4):040302, 2016.

[SBZ19] Michal Sedlák, Alessandro Bisio, and Mário Ziman. Optimal probabilistic storage and retrieval of unitary channels. *Physical review letters*, 122(17):170502, 2019.

[SCHL16] Imdad SB Sardharwalla, Toby S Cubitt, Aram W Harrow, and Noah Linden. Universal refocusing of systematic quantum noise. *arXiv preprint arXiv:1602.07963*, 2016.

[SHH25] Thomas Schuster, Jonas Haferkamp, and Hsin-Yuan Huang. Random unitaries in extremely low depth. *Science*, 389(6755):92–96, 2025.

[SMKH22] Michał Studziński, Marek Mozrzymas, Piotr Kopszak, and Michał Horodecki. Efficient multi port-based teleportation schemes. *IEEE Transactions on Information Theory*, 68(12):7892–7912, 2022.

[SML+25] Thomas Schuster, Fermi Ma, Alex Lombardi, Fernando Brandão, and Hsin-Yuan Huang. Strong random unitaries and fast scrambling. *arXiv preprint arXiv:2509.26310*, 2025.

[SMM09] Lana Sheridan, Dmitri Maslov, and Michele Mosca. Approximating fractional time quantum evolution. *Journal of Physics A: Mathematical and Theoretical*, 42(18):185302, 2009.

[SNC+23] Thomas Schuster, Murphy Niu, Jordan Cotler, Thomas O'Brien, Jarrod R McClean, and Masoud Mohseni. Learning quantum systems via out-of-time-order correlators. *Physical Review Research*, 5(4):043284, 2023.

[SS14] Stephen H Shenker and Douglas Stanford. Black holes and the butterfly effect. *Journal of High Energy Physics*, 2014(3):1–25, 2014.

[SST+23] Peter Schiansky, Teodor Strömberg, David Trillo, Valeria Saggio, Ben Dive, Miguel Navascués, and Philip Walther. Demonstration of universal time-reversal for qubit processes. *Optica*, 10(2):200–205, 2023.

[TDN20] David Trillo, Benjamin Dive, and Miguel Navascués. Translating uncontrolled systems in time. *Quantum*, 4:374, 2020.

[TDN23] David Trillo, Benjamin Dive, and Miguel Navascués. Universal quantum rewinding protocol with an arbitrarily high probability of success. *Physical Review Letters*, 130(11):110201, 2023.

[TMM+25] Philip Taranto, Simon Milz, Mio Murao, Marco Túlio Quintino, and Kavan Modi. Higher-order quantum operations. *arXiv preprint arXiv:2503.09693*, 2025.

[TW25a] Ewin Tang and John Wright. Amplitude amplification and estimation require inverses. *arXiv preprint arXiv:2507.23787*, 2025.

[TW25b] Ewin Tang and John Wright. Are controlled unitaries helpful? *arXiv preprint arXiv:2508.00055*, 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.

[VES+19] Benoît Vermersch, Andreas Elben, Lukas M Sieberer, Norman Y Yao, and Peter Zoller. Probing scrambling using statistical correlations between randomized measurements. *Physical Review X*, 9(2):021061, 2019.

[VH25] Francisca Vasconcelos and Hsin-Yuan Huang. Learning shallow quantum circuits with many-qubit gates. In *Conference on Learning Theory*, 2025.

[WZC+22] Qisheng Wang, Zhicheng Zhang, Kean Chen, Ji Guan, Wang Fang, Junyi Liu, and Mingsheng Ying. Quantum algorithm for fidelity estimation. *IEEE Transactions on Information Theory*, 69(1):273–282, 2022.

[XS24] Shenglong Xu and Brian Swingle. Scrambling dynamics and out-of-time-ordered correlators in quantum many-body systems. *PRX quantum*, 5(1):010201, 2024.

[YGS+16] Norman Y Yao, Fabian Grusdt, Brian Swingle, Mikhail D Lukin, Dan M Stamper-Kurn, Joel E Moore, and Eugene A Demler. Interferometric approach to probing fast scrambling. *arXiv preprint arXiv:1607.01801*, 2016.

[YH21] Yuxiang Yang and Masahito Hayashi. Representation matching for remote quantum computing. *PRX Quantum*, 2(2):020327, 2021.

[YKS+26] Satoshi Yoshida, Yuki Koizumi, Michał Studziński, Marco Túlio Quintino, and Mio Murao. One-to-one correspondence between deterministic port-based teleportation and unitary estimation. *IEEE Transactions on Information Theory*, 2026.

[YRC20] Yuxiang Yang, Renato Renner, and Giulio Chiribella. Optimal universal programming of unitary gates. *Physical review letters*, 125(21):210501, 2020.

[YSM23] Satoshi Yoshida, Akihito Soeda, and Mio Murao. Reversing unknown qubit-unitary operation, deterministically and exactly. *Physical Review Letters*, 131(12):120602, 2023.

[ZCJ+25] Guocheng Zhen, Yu-Ao Chen, Mingrui Jing, Jingu Xie, Ranyiliu Chen, and Xin Wang. Structure, optimality, and symmetry in shadow unitary inversion. *arXiv preprint arXiv:2510.24880*, 2025.

[Zha19] Mark Zhandry. How to record quantum queries, and applications to quantum indifferentiability. In *Annual International Cryptology Conference*, pages 239–268, 2019.

[Zha25a] Mark Zhandry. How to model unitary oracles. In *Annual International Cryptology Conference*, pages 237–268, 2025.

[Zha25b] Andrew Zhao. Learning the structure of any hamiltonian from minimal assumptions. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, pages 1201–1211, 2025.

[ZMCW24] Chengkai Zhu, Yin Mo, Yu-Ao Chen, and Xin Wang. Reversing unknown quantum processes via virtual combs for channels with limited information. *Physical Review Letters*, 133(3):030801, 2024.