Quantum Information and Computation, Vol. 22, No. 7&8 (2022) 0594–0626

A TIGHT LOWER BOUND FOR NON-COHERENT INDEX ERASURE

NATHAN LINDZEY
Department of Computer Science, University of Colorado at Boulder
430 UCB, 1111 Engineering Drive, Boulder, Colorado 80309, USA

ANSIS ROSMANIS
Graduate School of Mathematics, Nagoya University
Furocho, Chikusa Ward, Nagoya, Aichi, 464-8601, Japan

Received September 27, 2021
Revised March 24, 2022

The *index erasure problem* is a quantum state generation problem that asks a quantum computer to prepare a uniform superposition over the image of an injective function given by an oracle. We prove a tight $\Omega(\sqrt{n})$ lower bound on the quantum query complexity of the *non-coherent* case of the problem, where, in addition to preparing the required superposition, the algorithm is allowed to leave the ancillary memory in an arbitrary function-dependent state. This resolves an open question of Ambainis et al., who gave a tight bound for the coherent case, the case where the ancillary memory must return to its initial state.

To prove our main result, we first extend the *automorphism principle* of Høyer et al. to the *general adversary method* of Lee et al. for state generation problems, which allows one to exploit the symmetries of these problems to lower bound their quantum query complexity. Using this method, we establish a strong connection between the quantum query complexity of non-coherent symmetric state generation problems and the *Krein parameters* of an association scheme defined on injective functions. In particular, we use the spherical harmonics a finite symmetric Gelfand pair associated with the space of injective functions to obtain asymptotic bounds on certain Krein parameters, from which the main result follows.

Keywords: Quantum query complexity, quantum state generation, association schemes, representation theory

1 Introduction

For proving lower bounds in the *oracle query model*, one assumes access to an oracle O_f that evaluates a black-box function $f\colon [n] \to [m]$ on input queries, where $[n] := \{1, 2, \cdots, n\}$ and $[m] := \{1, 2, \cdots, m\}$, and the goal is to prove that any algorithm for solving the computational problem at hand must make a certain number of oracle queries. This principle for proving lower bounds applies to both classical and quantum computation, and in the latter we allow the oracle to be queried in a superposition.

Quantum query algorithms are known to surpass their classical counterparts for many

important classical tasks, such as unstructured search, game tree evaluation, random walks, and others (see [15, 2] for recent surveys). Classical tasks aside, one may also be interested in *quantum mechanical tasks*, such as *quantum state generation*. A quantum state generation problem simply asks for a certain quantum state $|\psi_f\rangle$ to be generated on the target register. In this paper, we consider a particular state generation problem known as Index Erasure.

Given an injective function $f\colon [n] \to [m]$ via a black-box oracle O_f, Index Erasure is the task of preparing the quantum state that is the uniform superposition over the image of f, namely,

$$|\psi_f\rangle := \frac{1}{\sqrt{n}} \sum_{x=1}^{n} |f(x)\rangle \,.$$

The name of the problem stems from the fact that a quantum computer can prepare the uniform superposition $\frac{1}{\sqrt{n}} \sum_{x=1}^{n} |x\rangle\,|f(x)\rangle$ using a single query to O_f, yet the task of ignoring or *erasing* the first register that records the *index* x is non-trivial. Indeed, if one could solve Index Erasure using a poly-logarithmic number of queries, one would obtain a time-efficient algorithm for Graph Isomorphism (see Section 1.3).

The question of the complexity of Index Erasure was first raised by Shi in [21], where he already observed that the problem can be solved in $\mathrm{O}(\sqrt{n})$ queries by an algorithm based on Grover's search. In the same paper, Shi also introduced the Set Equality problem, which asks to decide whether two injective functions f, f' given via black-box oracles $O_f, O_{f'}$ have the same image or have disjoint images, given a promise that either is the case. Set Equality can be easily reduced to Index Erasure via the swap test, increasing the number of oracle queries by at most a constant factor; therefore, when Midrijānis presented an $\Omega((n/\log n)^{1/5})$ lower bound on the quantum query complexity of Set Equality [14], the same lower bound automatically applied to Index Erasure, ruling out the existence of poly-logarithmic query algorithms for these two problems.

Quantum state generation comes in two forms: the *coherent state generation*, where all memory aside from the target state must return to its initial state, $|\mathbf{0}\rangle := |0\cdots 0\rangle$, and the *non-coherent state generation*, where there is no such a requirement, namely, where the ancillary memory can remain in some function-dependent state $|t_f\rangle$. Ambainis, Magnin, Roetteler, and Roland devised the *hybrid adversary method* [1], which they used to prove a tight $\Omega(\sqrt{n})$ lower bound for Index Erasure in the *coherent* regime, and left the non-coherent case as an open question. Later, the lower bound for Set Equality was improved to a tight $\Omega(n^{1/3})$ [22, 4], which in turn led to an improved query lower bound for the non-coherent Index Erasure.

In this paper, we close the gap for the non-coherent Index Erasure problem, by proving a tight lower bound on its quantum query complexity under the condition that the range of the black-box function f is sufficiently large. More formally, we show the following.

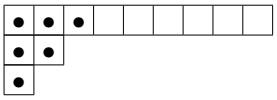
Theorem 1 (Main Result) *The bounded-error quantum query complexity of the* Index Erasure *problem is* $\Theta(\sqrt{n})$ *in the non-coherent state generation regime, provided that* $m \geq n^{3+\epsilon}$ *for some* $\epsilon > 0$.

We outline the proof of Theorem 1 below.

1.1 Outline of the Proof of Theorem 1

The symmetries of INDEX ERASURE are paramount in our proof (see Section 4 and Section 5 for a detailed discussion of these symmetries and any undefined terminology in this outline). The product $S_n \times S_m$ of two symmetric groups acts on a function $f\colon [n] \to [m]$ as $(\pi, \rho)\colon f \mapsto \rho * f * \pi^{-1}$, where $(\pi, \rho) \in S_n \times S_m$ and $*$ denotes the composition of functions. This group action on injective functions defines a representation of $S_n \times S_m$. This representation is multiplicity-free, meaning that it contains no more than one instance of any irrep (irreducible representation) of $S_n \times S_m$. Moreover, it consists of those and only those irreps $\lambda \otimes \lambda'$ where the Young diagram $\lambda \vdash n$ is contained in the Young diagram $\lambda' \vdash m$ and the skew shape λ'/λ has no more than one cell per column. Throughout the paper, we often abuse the terminology and we interchangeably use the terms partition λ of n, denoted $\lambda \vdash n$, the irreducible representation (irrep) corresponding to λ, and the n-cell Young diagram corresponding to λ.

The following class of irreps plays a distinguished role in our proof. Given $\lambda \vdash n$, we call the irrep $\lambda \otimes \bar{\lambda}$ where $\bar{\lambda} \vdash m$ is obtained from λ by adding $m - n$ cells to the first row of λ a *minimal* irrep. In other words, if $\theta \vdash k$ and $\lambda := (n - k, \theta) \vdash n$, then $(n - k, \theta) \otimes (m - k, \theta)$ is a minimal irrep. For example, if $m = 12$, $n = 6$, $k = 3$, and $\theta = (2, 1)$, then the minimal irrep with respect to θ is



(1)

where the •'s indicate an S_n-irrep and the □'s indicate an S_m-irrep. To lower bound the quantum query complexity of the non-coherent INDEX ERASURE, we use essentially the same adversary matrix Γ that [1] used for the coherent INDEX ERASURE, which is specified through minimal irreps (see Section 6 for a formal definition of this matrix).

An adversary matrix is a symmetric real matrix whose rows and columns are labeled by all the functions in the domain of the problem, and it is the central object of most adversary methods. In our case, the adversary matrix acts on the same $m^{\underline{n}}$-dimensional space as the representation matrices of $S_n \times S_m$ mentioned above, where $m^{\underline{n}} := m!/(m - n)!$ is the total number of functions. Similarly to [1], we choose

$$\Gamma := \sum_{k=0}^{\sqrt{n}-1} \left(\sqrt{n} - k\right) \sum_{\theta \vdash k} E_{(n-k,\theta)\otimes(m-k,\theta)},$$

where $E_{\lambda\otimes\lambda'}$ is the orthogonal projector on the irrep $\lambda \otimes \lambda'$ (note that we have only used projectors on certain minimal irreps to construct Γ). We also note that the Gram matrices $T_{\lambda\otimes\lambda'} = m^{\underline{n}} E_{\lambda\otimes\lambda'}/d_{\lambda\otimes\lambda'}$, where $d_{\lambda\otimes\lambda'} := \operatorname{tr}[E_{\lambda\otimes\lambda'}]$ is the dimension of $\lambda \otimes \lambda'$, play an important role in our proof.

In order to take advantage of the inherent symmetries of the INDEX ERASURE problem, we first extend the the *automorphism principle* of Høyer, Lee, and Špalek [11] to the *general adversary method* for state generation and conversion problems [13] (see Corollary 1 and Theorem 3). This extension leads us to consider the Gram matrix corresponding to the final state $|\psi_f, t_f\rangle$ of an algorithm run with oracle O_f (assuming no error). The Gram matrix

corresponding to $|\psi_f\rangle$ is

$$\frac{n}{m}T_{(n)\otimes(m)} + \left(1 - \frac{n}{m}\right) T_{(n)\otimes(m-1,1)} =: T^{\odot},$$

therefore the Gram matrix corresponding to $|\psi_f, t_f\rangle$ is $T^{\odot} \circ T$, where $T_{f,f'} := \langle t_f | t_{f'} \rangle$ and $\circ$ denotes the Schur (i.e., entrywise) matrix product. For the coherent regime lower bound, $\langle \mathbf{0} | \mathbf{0} \rangle = 1$ and $T = J = T_{(n)\otimes(m)}$ is the all-ones matrix. For the non-coherent regime, the Gram matrix T can be arbitrary, but one of the consequences of the generalization of the automorphism principle is that it suffices to consider T such that $T_{f,f'} = T_{\sigma(f),\sigma(f')}$ for all functions f, f' and all $\sigma \in S_n \times S_m$.

To prove the $\Omega(\sqrt{n})$ lower bound, we must show, for all such Gram matrices T, that

$$\operatorname{tr}\left[\Pi_\Gamma \frac{T^{\odot} \circ T}{m^{\underline{n}}}\right] = o(1), \tag{2}$$

where Π_Γ is the orthogonal projector on the image of Γ, and that $\|\Gamma \circ \Delta_x\| = \mathrm{O}(1)$ for all $x \in [n]$, where Δ_x is the binary matrix with $(\Delta_x)_{f,f'} := 1$ if and only if $f(x) \neq f'(x)$.[a] Here we only need to prove the former condition because we use essentially the same adversary matrix as [1], and the latter condition is shown in their work. On the other hand, showing condition (2) was a triviality in [1] because $T = J$ in the coherent regime and thus the trace evaluates to n/m. Showing that condition (2) holds is the main technical contribution of this work, and appears to require significantly more algebraic results on the space of injective functions than the main technical result of [1], which we develop in Sections 4 and 5.

We now present the three main simplifying steps used to narrow the scope of condition (2). First, we use linearity to show that it suffices to prove

$$\operatorname{tr}\left[\Pi_\Gamma \frac{T_{(n)\otimes(m-1,1)} \circ T_{\lambda\otimes\lambda'}}{m^{\underline{n}}}\right] = o(1)$$

for all irreps $\lambda \otimes \lambda'$. That is, we can restrict our attention from a continuum of choices for T to a finite set $\{T_{\lambda\otimes\lambda'}\}_{\lambda\otimes\lambda'}$ of choices, where we have also used that the term $T_{(n)\otimes(m-1,1)}$ "dominates" $T_{(n)\otimes(m)}$ in $T^{\odot}$ (since we assume that $m \gg n$).

Second, we use the connection between $T_{(n)\otimes(m-1,1)}$ and a specific primitive idempotent of the Johnson (association) scheme to obtain

$$\operatorname{tr}\left[E_{\lambda\otimes\bar{\lambda}} \frac{T_{(n)\otimes(m-1,1)} \circ T_{\lambda\otimes\lambda'}}{m^{\underline{n}}}\right] = o(1)$$

as a sufficient condition, where we have to consider only Young diagrams $\lambda \vdash n$ that have less than $\sqrt{n}$ cells below the first row.

Third, for such λ, we show that the dimension of $\lambda \otimes \bar{\lambda}$ is much smaller than the dimension of any other $\lambda \otimes \lambda'$ (thus the nomenclature "minimal irrep"); therefore, we show it suffices to prove for all $\lambda \vdash n$ that

$$\operatorname{tr}\left[E_{\lambda\otimes\bar{\lambda}} \frac{T_{(n)\otimes(m-1,1)} \circ T_{\lambda\otimes\bar{\lambda}}}{m^{\underline{n}}}\right] = o(1) \tag{3}$$

[a] The terms in condition (2) and similar expressions are written in such a way to emphasize that $\frac{T^{\odot} \circ T}{m^{\underline{n}}}$ is a density operator.

It is convenient to think of (2) and its simplifications in terms of the following association scheme (see Section 5 for more details). For a pair of functions (f, f'), consider the orbit $\mathcal{O}_\mu := \{(\sigma(f), \sigma(f')) \colon \sigma \in S_n \times S_m\}$, and let A_μ be the binary matrix with $(A_\mu)_{h,h'} = 1$ if and only if $(h, h') \in \mathcal{O}_\mu$. Here we use μ to label distinct orbits and let $\mathcal{C}_n$ be the set of all of them. The set of matrices $\{A_\mu \colon \mu \in \mathcal{C}_n\}$ forms a symmetric association scheme, denoted $\mathcal{A}_{n,m}$, which has been called the *injection scheme* [16]. Note that there is an obvious bijection between injective functions $f \colon [n] \to [m]$ and n-partial permutations of $[m]$ via $f \leftrightarrow (f(1), f(2), \cdots, f(n))$.[b]

In the terminology of association schemes, the projectors $E_{\lambda\otimes\lambda'}$ are called the *primitive idempotents*, and their entries corresponding to the orbit $\mathcal{O}_\mu$ multiplied by $m^{\underline{n}}$ are called *dual eigenvalues* of the association scheme, which we denote as $q_{\lambda\otimes\lambda'}(\mu)$. The *valency* v_μ is the size of $\mathcal{O}_\mu$ divided by $m^{\underline{n}}$, thus, in terms of dual eigenvalues, the left hand side of condition (3) can be written as

$$\frac{\sum_{\mu\in\mathcal{C}_n} v_\mu \cdot q_{(n)\otimes(m-1,1)}(\mu) \cdot q^2_{\lambda\otimes\bar{\lambda}}(\mu)}{m^{\underline{n}}\, d_{(n)\otimes(m-1,1)} d_{\lambda\otimes\bar{\lambda}}} = o(1). \tag{4}$$

Finally, to prove (4), we consider the spherical harmonics of a finite symmetric Gelfand pair associated with the space of injective functions along with some estimates of combinatorial coefficients related to the unsigned Stirling numbers of the first kind.

1.2 Krein Parameters of the Injection Scheme

In the context of quantum query complexity, the injection association scheme was already considered in [19], where a conjecture on its eigenvalues implied tight adversary bounds for the Collision and Set Equality problems. Along these lines, our work shows a connection between quantum query complexity and the *Krein parameters* $q_{i,j}(k)$ of association schemes (see Section 6 for a formal definition). Indeed, condition (3) is equivalent to the conditions

$$q_{\lambda\otimes\bar{\lambda},\,\lambda\otimes\bar{\lambda}}((n)\otimes(m-1,1)) = o(d_{\lambda\otimes\bar{\lambda}}) \quad \text{and} \quad q_{\lambda\otimes\bar{\lambda},\,(n)\otimes(m-1,1)}(\lambda\otimes\bar{\lambda}) = o(m)$$

on the Krein parameters of $\mathcal{A}_{n,m}$, and (4) gives an expression of these parameters in terms of dual eigenvalues.

The Krein parameters of an association scheme are important because they are the *dual structure constants* of its corresponding *Bose-Mesner algebra*. While the structure constants (i.e., intersection numbers) of Bose-Mesner algebras admit an obvious combinatorial meaning, its dual structure constants do not (e.g., they can be irrational) and are difficult to interpret. Indeed, the question of whether or not there exists a "good" interpretation of these constants has often been asked in algebraic combinatorics, so we find their connection to quantum query complexity to be interesting.

1.3 Connection to Graph Isomorphism through Set Equality

Given a graph G of the vertex set $[k] := \{1, 2, \cdots, k\}$ and a permutation $\pi \in S_k$, let $\pi \cdot G$ denote the graph obtained by the natural action of π on the vertices of G. Let us assume that G is rigid, so the "permuting" function $f \colon \pi \mapsto \pi \cdot G$ is injective, and let O_f be the oracle evaluating this function.

[b] For describing particular injections, we prefer the latter representation as it is a bit more succinct.

Now suppose we have two rigid graphs G_0, G_1 of the vertex set $[k]$, and let f_0, f_1 be the corresponding permuting functions. If the two graphs are isomorphic, then $\operatorname{im} f_0 = \operatorname{im} f_1$, while, if they are non-isomorphic, then $\operatorname{im} f_0 \cap \operatorname{im} f_1 = \emptyset$. As a result, we can employ a query-optimal algorithm for Set Equality [5] which performs $\mathrm{O}(\sqrt[3]{k!})$ queries to oracles O_{f_0}, O_{f_1} and tests isomorphism of G_0 and G_1 without having to look into internal structure of these graphs.

Because of the optimality of the query algorithm for Set Equality, one may want to say that any algorithm for Graph Isomorphism that does not employ the internal structure of graphs must perform $\Omega(\sqrt[3]{k!})$ queries to oracles O_{f_0}, O_{f_1}. However, to formally prove such a statement, one would have to formalize what is meant by "not employing the internal structure of a graph". A potential approach to do that would be to encrypt all graphs using a uniformly random injective function $\mathcal{E}$ from all k-vertex graphs to bit-strings of length $\mathsf{const} \cdot k$ and to provide an algorithm with encryptions $\mathcal{E}(G_0), \mathcal{E}(G_1)$ and an oracle-access to the function $F\colon (\pi, \mathcal{E}(G)) \mapsto \mathcal{E}(\pi \cdot G)$.[c]

We conjecture that, in this setting, testing isomorphism of G_0 and G_1 requires $\Theta(\sqrt[3]{k!})$ oracle evaluations of F. However, note that, compared to completely random injective functions on S_k, the function F has an additional structure. For example, $F(\tau, F(\tau, \mathcal{E}(G))) = \mathcal{E}(G)$ for any transposition τ. To prove the desired lower bound, one would have to show that no algorithm can take advantage of this additional structure, and such task is beyond the scope of the present work.

The Set Equality problem can be reduced to Index Erasure. Let us describe two natural reductions, one that requires Index Erasure to be coherent and one that permits it to be non-coherent. Here we assume that we are given oracle access to two injective functions $f_0, f_1\colon [n] \to [m]$.

In the first reduction, one prepares state

$$(|0\rangle\,|\psi_{f_0}\rangle\,|t_0\rangle + |1\rangle\,|\psi_{f_1}\rangle\,|t_1\rangle)/\sqrt{2} \qquad \text{such that} \qquad |\psi_f\rangle := \sum\nolimits_{x=1}^{n} |f(x)\rangle/\sqrt{n}$$

using an Index Erasure algorithm, and then measures the first qubit in its Fourier basis $\{|+\rangle, |-\rangle\}$, where $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$. For this reduction to work, the temporary registers must be left in a default state $|t_0\rangle = |t_1\rangle = |\mathbf{0}\rangle$, which requires index erasure to be coherent.

In the second reduction, one prepares the state $|S_n(G_0)\rangle\,|t_0\rangle\,|S_n(G_1)\rangle\,|t_1\rangle$ using two runs of an Index Erasure algorithm, and then performs the SWAP test on the registers containing $|S_n(G_0)\rangle$ and $|S_n(G_1)\rangle$. Unlike the reduction before, this reduction does not require Index Erasure to be coherent. That is, the states $|t_0\rangle, |t_1\rangle$ of the "garbage" qubits are inconsequential, and the algorithm does not have to clean them up.

Ambainis et al. [1] showed that the quantum query complexity of *coherent* Index Erasure is $\Theta(\sqrt{n})$, therefore strictly separating complexities of Set Equality and coherent Index Erasure. Before the present work, there was still a possibility that non-coherent Index Erasure might be as fast as Set Equality, but we prove that it is not the case, strictly separating complexities of Set Equality and non-coherent Index Erasure as well.

[c] The constant const must be at least 2, but one may wish to choose it larger so that a randomly guessed bit string is unlikely to be an encryption of any graph.

1.4 Organization of the paper

The paper is organized as follows. In Section 2, we present preliminaries on the quantum query model, with emphasis on state generation problems, including INDEX ERASURE, the general adversary method, and the automorphism principle. In Section 4, we present preliminaries on the representation theory, particularly focusing on the symmetric group and its action on injections. The automorphism principle of the general adversary method requires us to analyze highly symmetric matrices, which are elements of the Bose–Mesner algebra corresponding to the injection scheme. In Section 5, we formally define this association scheme, establishing the labeling of its various parameters and computing some of them, as well as addressing its connection to the Johnson scheme. With this formalism at our disposal, in Section 6, we show that the proof for the $\Omega(\sqrt{n})$ lower bound on the quantum query complexity of the non-coherent INDEX ERASURE can be reduced to showing upper bounds on certain Krein parameters of the injection scheme. Finally, we place the required bounds on these Krein parameters in Section 7.

2 Quantum state generation

In this paper, we address limitations of quantum query algorithms for solving the INDEX ERASURE problem. We assume that the reader is familiar with foundations of quantum computing (see [17] for an introductory reference), some of which we review here. The basic memory unit of a quantum computer is a qubit, which is a two-dimensional complex Euclidean space $\mathbb{C}[\{0,1\}]$ having *computational* orthonormal basis $\{|0\rangle, |1\rangle\}$. Similarly, a k-qubit system corresponds to Euclidean space $\mathbb{C}[\{0,1\}^k]$ with computational basis $\{|b\rangle : b \in \{0,1\}^k\}$. Unit vectors $|\Psi\rangle \in \mathbb{C}[\{0,1\}^k]$ are called (pure) *quantum states* and they represent superpositions over various computational basis states.

Quantum bits are often grouped together in *registers* for the ease of algorithm design and analysis. If $|\psi\rangle, |\phi\rangle$ are states of two registers, then the state of the joint system is $|\psi\rangle \otimes |\phi\rangle$. We often shorten the notation $|\psi\rangle \otimes |\phi\rangle$ to $|\psi\rangle |\phi\rangle$ or $|\psi, \phi\rangle$. Due to *entanglement*, it is not always the case that the state of the joint system can be written as a tensor product of states of the individual registers.

Quantum information is processed by unitary transformations, which correspond to square matrices U such that $UU^* = U^*U = I$, and they map quantum states to quantum states. This unitary processing of quantum information implies that any (noiseless) quantum computation is reversible.

2.1 Quantum query model

In the oracle model, we are given an access to a black-box oracle O_f that evaluates some unknown function $f\colon [n] \to [m]$. The goal of a query algorithm is to perform some computational task that depends on f, for example, to compute some function of f, such as $\textsc{Parity}(f) := f(1) \oplus f(2) \oplus \cdots \oplus f(n)$ when $m = 2$. In quantum computing, one can query the oracle in superposition. On the other hand, due to the requirement for reversibility, the oracle is typically designed so that it preserves the input query x. Namely, given $|x, y\rangle$ as an input, the oracle O_f outputs $|x, y \oplus f(x)\rangle$ (see Figure 1). Here and below we may assume $x, y, f(x)$ to be represented in binary. Even if f is injective—as it is for INDEX ERASURE—unless one knows how to compute the inverse of f, implementing $|x\rangle \mapsto |f(x)\rangle$ in practice

might be much harder than $|x, y\rangle \mapsto |x, y \oplus f(x)\rangle$.

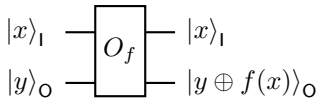


Fig. 1. A schematic of a quantum oracle O_f. We assume that y and $f(x)$ are encoded in binary, and thus O_f is its own inverse.

A quantum query algorithm with oracle O_f consists of

- four registers: input and output registers I and O for accessing the black-box function f, the target register T for storing the result of the computation, and a workspace register W;
- an indexed sequence of unitary transformations $U_0, U_1, \cdots, U_Q$ acting on those four registers.

The quantum query algorithm starts its computation in state $|\mathbf{0}\rangle := |00\cdots0\rangle$, and then performs $2Q+1$ unitary operations, alternating between U_i, which acts on all the registers, and O_f, which acts on registers IO. Thus the final state of the computation is

$$|\Psi_f\rangle := U_Q(O_f \otimes I_{\mathsf{TW}})U_{Q-1}(O_f \otimes I_{\mathsf{TW}}) \cdots U_1(O_f \otimes I_{\mathsf{TW}})U_0 \,|\mathbf{0}\rangle\,,$$

where I_{TW} is the identity operator on registers TW. Figure 2 gives a schematic of a quantum query algorithm. Note that Q is the number of oracle queries performed by the algorithm, and we also refer to it as the *query complexity of the algorithm.*

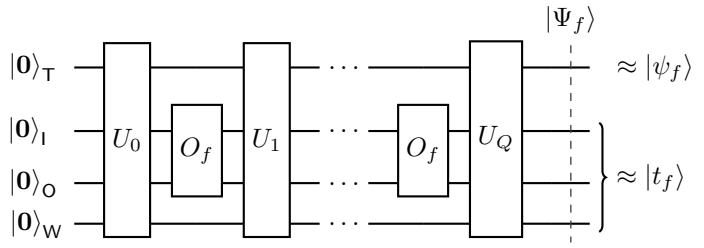


Fig. 2. A schematic of a quantum algorithm that uses an oracle O_f. The registers labeled $\mathsf{T}, \mathsf{I}, \mathsf{O}, \mathsf{W}$ are, respectively, the target, input, output, and workspace registers of the algorithm. The target register of the final state $|\Psi_f\rangle$ of the algorithm should be in a state close to the target state $|\psi_f\rangle$.

In this paper we are interested in quantum query algorithms whose goal is to generate a specific f-dependent state $|\psi_f\rangle$ by accessing f via O_f. We note that this generalizes classical function evaluation by a quantum algorithm, where each $|\psi_f\rangle$ is asked to be a computational basis vector. In the next section we describe two distinct regimes of quantum state generation, as well as why they are exactly the same for classical function evaluation.

2.2 Coherent vs. Non-coherent State Generation

When we talk about quantum state generation with oracle O_f, we implicitly assume the domain $[n]$ and the range $[m]$ of f to be fixed. A *quantum state generation* problem is thus

specified by a subset $\mathcal{F}$ of functions in form $f\colon [n] \to [m]$, which we call the *domain of the problem*, a complex Euclidean space called the *target space*, and, for every $f \in \mathcal{F}$, a quantum state $|\psi_f\rangle$ in the target space called the *target state*.

One may consider quantum state generation in two regimes: coherent and non-coherent. In the *coherent* state generation regime, all the computational memory other than the target register (i.e., registers IOW) must be returned to its initial state $|\mathbf{0}\rangle$. Therefore, if one was running an algorithm for a superposition of oracles, the final quantum state would be a superposition of the target states. In contrast, for *non-coherent* state generation, one does not place any requirements on the ancillary memory. More precisely, in the coherent case, for every input $f \in \mathcal{F}$ we require that the final state $|\Psi_f\rangle$ satisfies

$$\Re\langle\psi_f, \mathbf{0}|\Psi_f\rangle \geq \sqrt{1-\epsilon},$$

where $|\mathbf{0}\rangle$ is the initial state of the ancillary registers and a constant $\epsilon \geq 0$ is the desired precision [13]. We call the minimum among quantum query complexities among quantum query algorithms that achieve this task the *(ϵ-error) quantum query complexity of the coherent version of the problem*. On the other hand, in the non-coherent case, the final state $|\Psi_f\rangle$ has to satisfy

$$\|(\langle\psi_f| \otimes I)|\Psi_f\rangle\| = \max_{|t_f\rangle} \Re\langle\psi_f, t_f|\Psi_f\rangle \geq \sqrt{1-\epsilon},$$

where the maximum is over unit vectors $|t_f\rangle$ on the system of registers IOW [13], and we analogously define the quantum query complexity of the non-coherent version of the problem.

It is worth noting that evaluation of classical functions can be considered as a special case of quantum state generation, where one is asked to prepare the computational basis state $|\psi_f\rangle$. Since quantum mechanics permits cloning of orthogonal states (computational basis states, in this case), there is no difference between coherent and non-coherent function evaluation, if one is willing to tolerate a two-fold increase in query complexity: at the end of a non-coherent computation, one can copy the target register into an additional register, and then run the whole computation in reverse, restoring all but this additional register to their initial state.

Finally, note that an algorithm for a coherent case of a problem solves its non-coherent case as well. Conversely, a lower bound on the non-coherent version of the problem is a lower bound on the coherent version as well.

2.3 Index Erasure

Throughout this work, we let n and m be positive integers such that $n \leq m$. The domain of Index Erasure is the set of all injective functions $f\colon [n] \to [m]$. These functions are in one-to-one correspondence with n-partial permutations of $[m]$ and thus $|\mathcal{F}| = m^{\underline{n}} := m!/(m-n)!$. Index Erasure is the task of preparing the quantum state that is the uniform superposition

$$|\psi_f\rangle := \frac{1}{\sqrt{n}} \sum_{x=1}^{n} |f(x)\rangle$$

over the image of f. Note that the state

$$\frac{1}{\sqrt{n}} \sum_{x=1}^{n} |x\rangle|f(x)\rangle$$

can be prepared using a single query to O_f. This would give us the superposition that we seek if we could only ignore or *erase* the first register that records the *index* x, which gives the problem its namesake.

The question of the complexity of Index Erasure was first raised by Shi [21]. As for the upper bound, there is a simple quantum query algorithm for coherent Index Erasure given access to O_f. Thinking of the injective function f as a database with entries in $[m]$, for any y in the image of f we may use Grover's algorithm with O_f to find the unique index x of f such that $f(x) = y$. In other words, there is a circuit that sends the superposition

$$\frac{1}{\sqrt{n}} \sum_{x=1}^{n} |f(x)\rangle \quad \text{to} \quad \frac{1}{\sqrt{n}} \sum_{x=1}^{n} |x\rangle|f(x)\rangle.$$

Inverting this circuit effectively "erases" the index register, which implies that the quantum query complexity of Index Erasure is $\mathrm{O}(\sqrt{n})$.

The first non-trivial lower bounds on the quantum query complexity of Index Erasure were obtained via the Set Equality problem, which asks to decide whether two injective functions f, f' given via black-box oracles $O_f, O_{f'}$ have the same image or have disjoint images, given a promise that either is the case. Set Equality can be easily reduced to non-coherent (and, thus, coherent too) Index Erasure via the swap test, increasing the number of oracle queries by at most a constant factor. Thus, when Midrijānis presented an $\Omega((n/\log n)^{1/5})$ lower bound on the quantum query complexity of Set Equality [14], the same lower bound automatically applied to Index Erasure. Ambainis, Magnin, Roetteler, and Roland devised the *hybrid adversary method* [1], which they used to prove a tight $\Omega(\sqrt{n})$ lower bound for Index Erasure in the *coherent* regime, and left the non-coherent case as an open question. Later, the lower bound for Set Equality was improved to a tight $\Omega(n^{1/3})$ [22, 4], which in turn led to an improved query lower bound for the non-coherent Index Erasure.

The focus of this work is to prove a tight lower bound on the quantum query complexity of Index Erasure in the *non-coherent* case. To show this, we use the so-called *general adversary method* [13] which we review in Section 3.

3 General Adversary Method

The general adversary method places optimal lower bounds on the quantum query complexity of any state conversion problem [13]. State conversion problems generalize state generation problems, yet in this paper it will suffice to introduce the adversary bound only for the latter.

The general adversary bound is stated via the γ_2 and filtered γ_2 norms, which are defined as follows. Let M be any matrix and let $\Delta = \{\Delta_x \colon x \in [n]\}$ be a family of matrices of the same dimensions as M. Define

$$\begin{aligned} \gamma_2(M) \quad &:= \max_{\Gamma'} \{\|M \circ \Gamma'\| \colon \|\Gamma'\| \leq 1\}, \\ \gamma_2(M|\Delta) &:= \max_{\Gamma} \big\{\|M \circ \Gamma\| \colon \max_{x \in [n]} \|\Delta_x \circ \Gamma\| \leq 1\big\}, \end{aligned}$$

where $\circ$ denotes the Schur (i.e., entrywise) product of two matrices and, thus, Γ and Γ' are required to have the same dimensions as M. One can show that $\gamma_2(\cdot)$ is a norm over the set of all matrices and $\gamma_2(\cdot|\Delta)$ is a norm over the set of matrices M that has $M_{f,f'} = 0$ whenever

$(\Delta_x)_{f,f'} = 0$ for all $x \in [n]$ (see [13] for details). The two norms are called the γ_2 *norm* and the *filtered* γ_2 *norm*, respectively.

The general adversary bound employs various real symmetric matrices whose rows and columns are labeled by black-box functions $f \in \mathcal{F}$ in the same order. The family of *difference matrices* Δ is defined as follows. For each $x \in [n]$, the Δ_x is a binary matrix such that $(\Delta_x)_{f,f'} := 1$ if and only if $f(x) \neq f'(x)$. A *state matrix* is any positive-semidefinite matrix T such that $T \circ I = I$. In other words, it is a Gram matrix corresponding to some family of unit vectors. Note that $\gamma_2(\cdot|\Delta)$ is a norm on the set of matrices whose diagonals are all-zeros, and a difference of any two state matrices belongs to this set.

Let $\mathcal{T}$ be the set of all state matrices. (In Section 6, we will narrow the definition of $\mathcal{T}$ to contain only state matrices possessing certain symmetries.) Note that $\mathcal{T}$ is a compact set and it is closed under the Schur product. Two particular state matrices of our interest are the all-ones matrix J, which corresponds to the family $\{|\mathbf{0}\rangle : f \in \mathcal{F}\}$, and the *target matrix* $T^{\odot}$ defined as $(T^{\odot})_{f,f'} := \langle\psi_f|\psi_{f'}\rangle$.

Theorem 2 is a special case of [13, Theorem 4.9].

Theorem 2 *The ϵ-error quantum query complexity of a non-coherent state generation problem with the target matrix $T^{\odot}$ and the family of difference matrices Δ is both*

$$\Omega\big(\mathsf{Adv}_{2\sqrt{2\epsilon}}\big) \quad \textit{and} \quad \mathrm{O}\big(\mathsf{Adv}_{\epsilon^4/16}\,\epsilon^{-2}\log\epsilon^{-1}\big),$$

where

$$\mathsf{Adv}_{\delta} := \min_{R,T\in\mathcal{T}}\{\gamma_2(J - R|\Delta)\colon \gamma_2(R - T^{\odot}\circ T) \leq \delta\}. \tag{5}$$

In the case of coherent state generation, one imposes $T = J$ in the expression for Adv_{δ}.

In the expression for Adv_{δ}, the state matrix T essentially corresponds to the ancillary states that are prepared in addition to the target states. Thus, assuming there were no error, $T^{\odot} \circ T$ would be the Gram matrix corresponding to the final states of the whole system. However, since one allows some error—determined by the parameter δ—it suffices that the state matrix R corresponding *exactly* to the final states of the algorithm is close to $T^{\odot} \circ T$.

When applying the adversary bound, it is convenient to actually apply it to the zero-error case therefore eliminating the matrix R from the consideration. In particular, this leads to the following corollary of Theorem 2.

A symmetric matrix Γ that satisfies $\|\Delta_x \circ \Gamma\| \leq 1$ for all x is called an *adversary matrix*. Let Π_{Γ} denote the orthogonal projector on the image of Γ.

Corollary 1 *Let Γ be an adversary matrix for a non-coherent state generation problem with the target matrix $T^{\odot}$ and the family of difference matrices Δ, let ω be a principal eigenvector of Γ of norm 1, and let*

$$\eta' := \max_{T\in\mathcal{T}} \omega^{\top}(T^{\odot} \circ T \circ \Gamma/\|\Gamma\|)\,\omega.$$

The ϵ-error quantum query complexity of the problem is

$$\Omega\big((1 - \eta' - 2\sqrt{2\epsilon})\,\|\Gamma\|\,\big).$$

If ω is a uniform superposition over $\mathcal{F}$, then $\eta' \leq \eta$ for

$$\eta := \max_{T \in \mathcal{T}} \operatorname{tr}\left[\Pi_\Gamma (T^{\odot} \circ T)/|\mathcal{F}|\right].$$

Proof. For the first part of the corollary, suppose $R, T \in \mathcal{T}$ satisfy $\gamma_2(R - T^{\odot} \circ T) \leq 2\sqrt{2\epsilon}$ and are thus a feasible solution to the minimization in $\mathsf{Adv}_{2\sqrt{2\epsilon}}$. We have

$$\begin{aligned}
\gamma_2(J - R|\Delta) \geq & \|(J - R) \circ \Gamma\| \\
\geq & \|(J - T^{\odot} \circ T) \circ \Gamma\| - \|\Gamma\| \left\|(R - T^{\odot} \circ T) \circ \Gamma / \|\Gamma\|\right\| \\
\geq & \, \omega^{\top} \Gamma \, \omega - \omega^{\top} (T^{\odot} \circ T \circ \Gamma) \, \omega - 2\sqrt{2\epsilon} \|\Gamma\| \\
\geq & (1 - \eta' - 2\sqrt{2\epsilon}) \|\Gamma\|.
\end{aligned}$$

For the second part, note that, if ω is a uniform superposition over $\mathcal{F}$, then, for any two symmetric $|\mathcal{F}| \times |\mathcal{F}|$ matrices M, M', we have $\omega^{\top}(M \circ M') \, \omega = \operatorname{tr}[MM']/|\mathcal{F}|$. The inequality $\eta' \leq \eta$ results from both $T^{\odot} \circ T$ and $\Pi_\Gamma - \Gamma/\|\Gamma\|$ being positive-semidefinite $\square$.

3.1 Automorphism Principle for State Generation

The *automorphism principle* of [11] addresses the adversary bound for function evaluation problems and states that, without loss of generality, the optimal adversary matrix can be required to respect symmetries of the problem. The main result of this section is Theorem 3, which is a generalization of the automorphism principle to state generation problems. It is not difficult to see that the proof of Theorem 3 can be generalized further to state conversion problems *mutatis mutandis*; however, since the current application to Index-Erasure is a state generation problem, we have elected not to prove Theorem 3 in this generality.

The wreath product $S_m \wr S_n$ of groups S_m and S_n is the group whose elements are $(\pi, \boldsymbol{\sigma}) \in S_n \times S_m^n$ and whose group operation is

$$\big(\pi', (\sigma'_1, \cdots, \sigma'_n)\big)\big(\pi, (\sigma_1, \cdots, \sigma_n)\big) = \big(\pi'\pi, (\sigma'_1 \sigma_{(\pi')^{-1}(1)}, \cdots, \sigma'_n \sigma_{(\pi')^{-1}(n)})\big)$$

(see [12, Ch. 4]). Similarly to (7) below, the action of $S_m \wr S_n$ on $f\colon [n] \to [m]$ is given by

$$\big((\pi, \boldsymbol{\sigma}) f\big)(x) = \sigma_x(f(\pi^{-1}(x))) \qquad \text{for all } x \in [n]. \tag{6}$$

The action of a subgroup $G \leq S_m \wr S_n$ on the set of black-box functions $\mathcal{F}$ is *closed* if $g(f) \in \mathcal{F}$ for all $f \in \mathcal{F}$ and $g \in G$.

Suppose M is a symmetric $|\mathcal{F}| \times |\mathcal{F}|$ matrix whose rows and columns are labeled by $f \in \mathcal{F}$ in the same order and suppose the action of a subgroup $G \leq S_m \wr S_n$ on $\mathcal{F}$ is closed. We say that M is *G-invariant* if $M_{g(f),g(f')} = M_{f,f'}$ for all $f, f' \in \mathcal{F}$ and $g \in G$. Similarly, a vector $\omega \in \mathbb{C}[\mathcal{F}]$ is G-invariant if $\omega_{g(f)} = \omega_f$ for all $f \in \mathcal{F}$ and $g \in G$. A subgroup G is an *automorphism group* for a state generation problem with a target matrix $T^{\odot}$ if G's action on $\mathcal{F}$ is closed and $T^{\odot}$ is G-invariant.[d]

Note that the free product of two automorphism groups is an automorphism group, so one can consider the maximum automorphism group of a problem. For example, the maximum

[d] The G-invariance of $T^{\odot}$ is equivalent to the existence of a unitary representation U_g of G acting on the target space such that $U_g|\psi_f\rangle = |\psi_{g(f)}\rangle$ for all $f \in \mathcal{F}$ and $g \in G$.

automorphism group of PARITY is the whole wreath product $S_2 \wr S_n$ while the maximum automorphism groups of OR and INDEX ERASURE are, respectively,

$$\{(\pi, (\varepsilon, \cdots, \varepsilon)) \colon \pi \in S_n\} \cong S_n,$$
$$\{(\pi, (\sigma, \cdots, \sigma)) \colon \pi \in S_n \text{ and } \sigma \in S_m\} \cong S_n \times S_m,$$

where ε is the identity permutation in S_2. Note that for PARITY, the target states $|\psi_f\rangle$ and $|\psi_{g(f)}\rangle$ may differ for g in the maximum automorphism group, and the same is true for INDEX ERASURE.

Theorem 3 *Let G be an automorphism group for a non-coherent state generation problem. The value of Adv_δ remains the same if one restricts the minimization in the expression defining Adv_δ and the maximization in the expressions defining the γ_2 and filtered γ_2 norms to R, T, Γ, Γ' that are all G-invariant.*

The proof of Theorem 3 splits into two parts according to the two types of symmetrizations of the matrices R, T, Γ, Γ', which depend on whether they are arguments in the aforementioned minimization or maximization.

Proof. (Theorem 3) Let M be a generic symmetric matrix whose rows and columns are labeled by black-box functions $f \in \mathcal{F}$ in the same order. Let $g(M)$ be obtained by permuting the rows and the columns of M according to the action of $g \in G$ of $\mathcal{F}$ (see (6)). Namely, entrywise we define $g(M)$ as

$$(g(M))_{f,f'} := M_{g^{-1}(f), g^{-1}(f')}.$$

Similarly, for a vector $\omega \in \mathbb{C}[\mathcal{F}]$, define $g(\omega)$ entrywise as $(g(\omega))_{f'} := \omega_{g^{-1}(f)}$. For the sake of conciseness, we also occasionally write M^g and ω^g instead of $g(M)$ and $g(\omega)$, respectively. Note that M is G-invariant if $M^g = M$ for all $g \in G$, and $T^{\odot}, I, J$ are G-invariant. Also note that $(M \circ M')^g = M^g \circ M'^g$.

Let $\Delta = \{\Delta_1, \cdots, \Delta_n\}$ be the family of difference matrices. This family is closed under the action of G is the following sense.

Claim 1 *We have $(\pi, \boldsymbol{\sigma})(\Delta_x) = \Delta_{\pi(x)}$ for all $(\pi, \boldsymbol{\sigma}) \in G$.*

Proof. Fix $(\pi, \boldsymbol{\sigma}) \in G$ and let $g := (\pi, \boldsymbol{\sigma})^{-1}$. Note that $g = (\pi^{-1}, \boldsymbol{\sigma}')$ for some $\boldsymbol{\sigma}' \in S_m^n$. From (6), we have $(g(f))(x) = (g(f'))(x)$ if and only if $f(\pi(x)) = f'(\pi(x))$. As a result, we have

$$((\pi, \boldsymbol{\sigma})(\Delta_x))_{f,f'} = (\Delta_x)_{g(f), g(f')} = 1$$

if and only if $f(\pi(x)) = f'(\pi(x))$ $\square$.

Note that M^g equals M with its rows and columns permuted. Permuting rows and columns does not affect the γ_2 norm, so we have $\gamma_2(M^g) = \gamma_2(M)$ for all $g \in G$. And, if the diagonal of M is all-zeros, then Claim 1 also implies that $\gamma_2(M^g|\Delta) = \gamma_2(M|\Delta)$ for all $g \in G$.

Claim 2 *Restricting $R, T \in \mathcal{T}$ to be G-invariant does not change the optimal value of the minimization problem defining* Adv_δ.

Proof. Let R, T be an optimal solution of the minimization in (5). We define their respective G-symmetrizations as

$$\overline{R} := \frac{1}{|G|}\sum_{g\in G} g(R) \qquad \text{and} \qquad \overline{T} := \frac{1}{|G|}\sum_{g\in G} g(T),$$

which are both clearly in $\mathcal{T}$. Since $g(T^\odot) = T^\odot$ for all $g \in G$, the triangle inequality yields

$$\begin{aligned}\gamma_2(\overline{R} - T^\odot \circ \overline{T}) = \gamma_2\bigg(\frac{1}{|G|}\sum_{g\in G} g(R - T^\odot \circ T)\bigg) &\le \frac{1}{|G|}\sum_{g\in G} \gamma_2\big(g(R - T^\odot \circ T)\big) \\ &= \frac{1}{|G|}\sum_{g\in G} \gamma_2(R - T^\odot \circ T) = \gamma_2(R - T^\odot \circ T) \le \delta.\end{aligned}$$

Hence we have show that the pair $\overline{R}, \overline{T}$ is a feasible solution to the minimization in (5), and it remains to show that it is also optimal. And, again by the triangle inequality,

$$\begin{aligned}\gamma_2(J - \overline{R}|\Delta) = \gamma_2\bigg(\frac{1}{|G|}\sum_{g\in G} g(J - R)\bigg|\Delta\bigg) &\le \frac{1}{|G|}\sum_{g\in G} \gamma_2\big(g(J - R)|\Delta\big) \\ &= \frac{1}{|G|}\sum_{g\in G} \gamma_2(J - R|\Delta) = \gamma_2(J - R|\Delta) = \mathsf{Adv}_\delta\end{aligned}$$

□.

Now, fix G-invariant $R, T \in \mathcal{T}$ and let $M := J - R$, which is also G-invariant. Let us now show that the maximization in

$$\gamma_2(M|\Delta) = \max_\Gamma \big\{\|M \circ \Gamma\| \colon \forall x\ \|\Delta_x \circ \Gamma\| \le 1\big\}$$

can be restricted to G-invariant Γ.

The proof now proceeds along the lines of the automorphism principle in [13]. Fix an optimal solution Γ, and without loss of generality assume that the largest eigenvalue of $M \circ \Gamma$ is positive and let it correspond to an eigenvector $\omega \in \mathbb{C}[\mathcal{F}]$ of norm 1. Namely,

$$\|M \circ \Gamma\| = \omega^\top (M \circ \Gamma)\omega.$$

Define the G-symmetrization $\overline{\omega}$ of ω entrywise as

$$\overline{\omega}_f := \sqrt{\frac{1}{|G|}\sum_{g\in G} |(\omega^g)_f|^2},$$

and note that $\overline{\omega}$ also has norm 1. Without loss of generality, all the entries of $\overline{\omega}$ are strictly positive (the rows and columns corresponding to f such that $\overline{\omega}_f = 0$ can be removed from the consideration), and thus we can entrywise define a vector μ as $\mu_f := 1/\overline{\omega}_f$. Let us define

$$\overline{\Gamma} := \mu\mu^\top \circ \frac{1}{|G|}\sum_{g\in G} \Gamma^g \circ \omega^g \omega^{g\top},$$

which is clearly G-invariant.

Let us start by showing that $\|\overline{\Gamma} \circ \Delta_x\| \leq 1$ for all x. Note that $\|\Gamma^g \circ \Delta_x\| \leq 1$ for all x and all $g \in G$ due to Claim 1, $\|\Gamma \circ \Delta_x\| \leq 1$ if and only if $I \pm \Gamma \circ \Delta_x$ is positive-semidefinite, and

$$I \circ \mu\mu^\top \circ \frac{1}{|G|} \sum_{g \in G} \omega^g \omega^{g\top} = I.$$

We thus have that

$$I \pm \overline{\Gamma} \circ \Delta_x = \mu\mu^\top \circ \frac{1}{|G|} \Big(\sum_{g \in G} \omega^g \omega^{g\top} \circ (I \pm \Gamma^g \circ \Delta_x) \Big)$$

is positive-semidefinite as the sum and the entrywise product of positive-semidefinite matrices are positive-semidefinite. Thus, indeed, $\|\overline{\Gamma} \circ \Delta_x\| \leq 1$ for all x.

Now let us use the fact that ω is a principal eigenvector of $M \circ \Gamma$, and, therefore, ω^g is a principal eigenvector of $M \circ \Gamma^g$ for all $g \in G$ (recall that M is G-invariant). We have

$$\begin{aligned} \|M \circ \overline{\Gamma}\| \geq \overline{\omega}\, (M \circ \bar{\Gamma})\, \overline{\omega}^\top &= \sum_{f,f' \in \mathcal{F}} \Big(\frac{1}{|G|} \sum_{g \in G} (M \circ \Gamma^g \circ \omega^g \omega^{g\top}) \Big)_{f,f'} \\ &= \frac{1}{|G|} \sum_{g \in G} \omega^{g\top} (M \circ \Gamma^g) \omega^g = \|M \circ \Gamma\|. \end{aligned}$$

Thus $\overline{\Gamma}$ is also an optimal solution of the maximization above. Also note that $\overline{\omega}$ is the principal eigenvector of $M \circ \overline{\Gamma}$.

A similar argument shows that, for G-invariant $M' := R - T^{\odot} \circ T$, one can restrict the maximization in

$$\gamma_2(M') = \max_{\Gamma'} \big\{ \|M' \circ \Gamma'\| \colon \|\Gamma'\| \leq 1 \big\}$$

to G-invariant Γ'. This completes the proof of Theorem 3 $\square$.

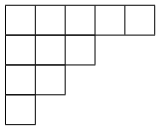
Note that the ability to restrict T and Γ to be G-invariant carries over from Theorem 2 to Corollary 1. The ability to restrict T will be paramount in our proof (see Section 6). On the other hand, the ability to restrict Γ is optional. Namely, Corollary 1 provides an adversary bound regardless of what restrictions one imposes on Γ, yet for too strict restrictions this bound would not be optimal.

As observed in [1], the set of $|\mathcal{F}| \times |\mathcal{F}|$ matrices indexed by $\mathcal{F}$ that are $(S_n \times S_m)$-invariant under the aforementioned action (6) afford a commutative matrix algebra. In particular, it is the Bose–Mesner algebra of a symmetric association scheme defined over injections, which we formally define in Section 5. Before we define this association scheme, some results from the representation theory of the symmetric group are needed, which we overview in the next section.

4 Representation Theory Preliminaries

We refer the reader to [7] for an introduction to group representation theory, [20] for more details on the representation theory of the symmetric group, and [6] for a more involved discussion on finite Gelfand pairs and their spherical functions.

Let $\mathrm{Sym}(X)$ denote the *symmetric group* on the symbol set X. If $X = [m] := \{1, 2, \cdots, m\}$, then we define $S_m := \mathrm{Sym}(X)$. It is well-known that the conjugacy classes of S_m and irreducible representations (irreps) of S_m are given by the cycle-types of permutations of S_m, which in turn are in one-to-one correspondence with *integer partitions* $\lambda \vdash m$, i.e., $\lambda := (\lambda_1, \lambda_2, \cdots, \lambda_k) \vdash m$ such that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k \geq 0$ and $\sum_{i=1}^{k} \lambda_i = m$. We may visualize λ as a *Young diagram*, a left-justified table of cells that contains λ_i cells in the ith row. When referencing a Young diagram, we alias λ as the *shape*. A *standard Young tableau* of shape $\lambda \vdash n$ is a Young diagram with unique entries from $[n]$ that are strictly increasing along rows and strictly increasing along columns. For example, the left Young diagram below has shape $(5, 3, 2, 1) \vdash 11$ and the tableau on the right is a standard Young tableau of the same shape



1	2	5	8	9
3	6	7		
4	10			
11				

.

Let $\mathcal{V}_\lambda$ denote the S_m-irrep corresponding to $\lambda \vdash m$. Let d_λ be the number of standard Young tableau of shape λ. It is well-known that d_λ is also the dimension of the S_m-irrep corresponding to $\lambda \vdash m$. The number of standard Young tableau can be counted elegantly via the *hook rule* (see [20] for a proof).

Theorem 4 (Hook rule) *Let $\lambda \vdash m$, and for any cell $c \in \lambda$ of the Young diagram of λ define the* hook-length $\mathsf{h}_\lambda(c)$ *to be the total number of cells below c in the same column and to the right of c in the same row, plus 1. Then we have $d_\lambda = m!/\mathsf{H}(\lambda)$ where $\mathsf{H}(\lambda) := \prod_{c \in \lambda} \mathsf{h}_\lambda(c)$.*

Another well-known result is the *branching rule*, which describes how an S_m-irrep decomposes into (S_{m-1})-irreps (see [20] for a proof). We say that a cell of a Young diagram is an *inner corner* if it has no cells to its right and no cells below it.

Theorem 5 (The Branching Rule) *If $\mathcal{V}_\lambda$ is an S_m-irrep, then $\mathcal{V}_\lambda \cong \bigoplus_{\lambda^-} \mathcal{V}_{\lambda^-}$ where λ^- ranges over all shapes obtainable by removing an inner corner from λ and $\mathcal{V}_{\lambda^-}$ is an (S_{m-1})-irrep corresponding to λ^-.*

The hook rule and the branching rule can be used to prove the following results. For any $\lambda \vdash n$, recall that $\bar{\lambda} \vdash m$ is obtained from λ by adding $m - n$ cells to the first row of λ.

Proposition 1 [7] *Let $\lambda \vdash n$ and $\ell = n - \lambda_1$. Then we have $d_\lambda d_{\bar{\lambda}} \leq \binom{n}{\ell}\binom{m}{\ell}\ell! \leq m^\ell n^\ell$.*

Theorem 6 *Let $\theta \vdash k$ and $\theta^+ \vdash (k+1)$ be any shape obtained by adding an inner corner to θ. For all $m \geq 2(k+1)$, we have*

$$\frac{d_{(m-k-1,\theta^+)}}{d_{(m-k,\theta)}} \geq \frac{m}{k} \cdot \left(1 - \frac{2k+1}{m}\right).$$

Proof. Recall that, by the hook rule, $\mathsf{H}(\theta)d_\theta = |\theta|!$. First, [18, Claim 6.3] states that

$$\frac{d_{(m-|\theta^+|,\theta)}}{d_{(m-|\theta|,\theta)}} \geq 1 - \frac{2k}{m}.$$

We reprove this claim here for completeness. Note that when we add a cell to the end of the top row of $(m-|\theta^+|,\theta)$ to obtain $(m-|\theta|,\theta)$, this increases the hook-lengths of the cells in the top row by 1, and the rest of the hook-lengths are unchanged. If we just consider the "overhang" and ignore everything else in the first row, then the product of the hook-lengths with respect to $(m-|\theta|,\theta)$ is $(m-2k)!$ whereas it is $(m-2k-1)!$ with respect to $(m-|\theta^+|,\theta)$. This gives us

$$\frac{d_{(m-|\theta^+|,\theta)}}{d_{(m-|\theta|,\theta)}} \geq \frac{m-2k}{m} = 1 - \frac{2k}{m},$$

which proves the claim.

Since $(m-|\theta^+|,\theta)$ is a partition of $(m-1)$, it corresponds to an (S_{m-1})-irrep. When we added one cell to $(m-|\theta^+|,\theta)$ to obtain $(m-|\theta^+|,\theta^+)$, only one hook-length in the first row increased, and before the increment it was at least $m-2k-1$. Thus, in the following derivation, all the other hook-lengths of the first rows of $(m-|\theta^+|,\theta)$ and $(m-|\theta^+|,\theta^+)$ have cancelled out.

$$\begin{aligned}\frac{d_{(m-|\theta^+|,\theta^+)}}{d_{(m-|\theta^+|,\theta)}} &\geq \frac{m!}{(m-1)!}\cdot\frac{m-2k-1}{m-2k}\cdot\frac{\mathsf{H}(\theta)}{\mathsf{H}(\theta^+)} \\ &= m\left(1-\frac{1}{m-2k}\right)\frac{k!d_{\theta^+}}{(k+1)!d_\theta} \geq \frac{m}{k+1}\left(1-\frac{1}{m-2k}\right),\end{aligned}$$

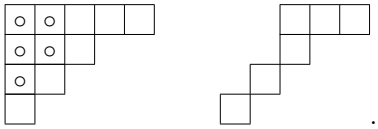
where the middle equality uses the hook rule once more, and the last inequality follows from the branching rule (namely, that $d_{\theta^+} \geq d_\theta$). Combining these two inequalities gives the result $\square$.

Note that the above fraction is greater than 1 when $m > 3k+1$.

4.1 The Representation Theory of Injections

Henceforth, let $S_{n,m}$ denote the collection of injective maps $f : [n] \to [m]$, equivalently, n-tuples $f := (f(1), f(2), \cdots, f(n))$ with no repeated elements such that $f(x) \in [m]$ for all $x \in [n]$. The latter representation is a bit more succinct, so we shall prefer it for describing particular injections. When $m = n$ we recover the symmetric group S_n on n symbols. To understand the representation theory of $S_{n,m}$ we must first broaden our Young tableau vocabulary.

For any $\lambda \vdash m$, let $l(\lambda)$ denote the *length* of λ, that is, the number of parts in the partition. For any integer partition μ, we let $|\mu|$ denote the size of μ, i.e., number of cells in its Young diagram. We say that a shape λ *covers* a shape μ if $\mu_i \leq \lambda_i$ for each i. If λ and μ are two shapes such that λ covers μ, then we obtain the *skew shape* λ/μ by removing the cells corresponding to μ from λ. For instance, the shape $(5,3,2,1)$ covers $(2,2,1)$, so we may consider the skew shape $(5,3,2,1)/(2,2,1)$:



A skew shape is a *horizontal strip* if each column has no more than one cell. For example, the skew shape $(5,3,2,1)/(3,3,1)$ is a horizontal strip, but the skew shape above is not.

We now observe that Theorem 6 has the following corollary.

Corollary 2 *Let $\lambda \vdash n$ be a shape such that $\lambda_1 \geq n - \sqrt{n}$ and let $\lambda' \vdash m$ be any shape that covers λ such that λ'/λ is a non-empty horizontal strip. Then $d_{\lambda'}/d_{\bar{\lambda}} \in \Omega(m/\sqrt{n})$.*

Henceforth, we let $S_n \times S_m$ act on $S_{n,m}$ as follows:

$$(\tau, \sigma) \cdot (f_1, \cdots, f_n) = (\sigma(f_{\tau^{-1}(1)}), \cdots, \sigma(f_{\tau^{-1}(n)})) \text{ for all } (\tau, \sigma) \in S_n \times S_m. \tag{7}$$

The stabilizer of the *identity injection* $f_{\mathrm{id}} := (1, 2, \cdots, n) \in S_{n,m}$ in $S_n \times S_m$ is isomorphic to the group

$$\mathrm{diag}(S_n \times S_n) \times S_{m-n} = \{(\tau, \tau, \pi) : \tau \in \mathrm{Sym}([n]), \pi \in \mathrm{Sym}(\{n+1, \cdots, m\}).$$

One can show (see [6]) that the permutation representation of $(S_n \times S_m)$ acting on $S_{n,m} \cong (S_n \times S_m)/\ (\mathrm{diag}(S_n \times S_n) \times S_{m-n})$ is *multiplicity-free*, that is, its decomposition has at most one copy of any $(S_n \times S_m)$-irrep, as shown in Theorem 7.

Theorem 7 [6] *The complex-valued functions over $S_{n,m}$ denoted as $\mathbb{C}[S_{n,m}]$ admits the following decomposition into $(S_n \times S_m)$-irreps:*

$$\mathbb{C}[S_{n,m}] \cong \bigoplus_{\mu,\lambda} \mathcal{V}_\mu \otimes \mathcal{V}_\lambda$$

where μ, λ ranges over all pairs $\mu \vdash n, \lambda \vdash m$ such that λ/μ is a horizontal strip.

Let $\mathrm{Irr}(S_{n,m})$ denote the set of $(S_n \times S_m)$-irreps that appear in Theorem 7. Every multiplicity-free permutation representation gives rise to a commutative association scheme (see [3]), so a consequence of Theorem 7 is the existence of a symmetric association scheme $\mathcal{A}_{n,m}$ over $S_{n,m}$ that we call *the injection association scheme* [16]. In Section 5, we discuss this association scheme in more detail.

For any finite groups $K \leq G$, we say that (G, K) is a *finite Gelfand pair* if its double coset algebra $\mathbb{C}[K\backslash G/K]$ is commutative, or equivalently, if the permutation representation of G acting on G/K is multiplicity-free. For any Gelfand pair (G, K), let ω^i be the *spherical function* corresponding to irrep indexed by i, i.e., the projection of the irreducible character indexed by i onto the space of (left) K-invariant functions of $\mathbb{C}[G/K]$ (see [6]). The spherical functions are constant on double cosets $K\backslash G/K$, so we may define ω^i_j to be the evaluation of ω^i on the double coset indexed by j. For more details on the connection between Gelfand pairs and association schemes, see [3].

We recall two well-known and basic facts about the spherical functions of finite Gelfand pairs.

Proposition 2 [6] *For any spherical function ω^i and double coset j, we have $|\omega^i_j| \leq 1$.*

A finite Gelfand pair (G, K) is *symmetric* if $g^{-1} \in KgK$ for all $g \in G$. Define $\delta_{i,j}$ so that $\delta_{i,j} = 1$ if $i = j$; otherwise, $\delta_{i,j} = 0$.

Proposition 3 [6] *Let $X = G/K$ such that (G, K) is a finite symmetric Gelfand pair. Let ω^i denote the i-th spherical function corresponding to irrep i of dimension d_i. Then*

$$\sum_{x \in X} \omega^i(x)\overline{\omega^j(x)} = \sum_{x \in X} \omega^i(x)\omega^j(x) = \delta_{i,j}\frac{|X|}{d_i}.$$

It is known that $(S_n \times S_m, \mathrm{diag}(S_n \times S_n) \times S_{m-n})$ is a finite symmetric Gelfand pair (see [6]), and we will use these basic results in our proof of the main result.

Finally, as stated in the Section 1.1, the minimal irreps of $\mathrm{Irr}(S_{n,m})$ will be of particular importance in our proof of the main result, which we formally define below.

Definition 1 (Minimal Irreps) *For any $\lambda \vdash n$, the* minimal irrep *with respect to λ is $\lambda \otimes \bar{\lambda}$.*

See (1) in Section 1.1 for a graphical example. Note that if $\lambda_1 \geq n - \sqrt{n}$, then Theorem 6 implies that the minimal irreps indeed have the least dimension over all irreps of the form $\lambda \otimes \mu \in \mathrm{Irr}(S_{n,m})$ for sufficiently large m.

5 The Injection Association Scheme

The theory of association schemes will be a convenient language for describing the algebraic and combinatorial components of our work. We refer the reader to Bannai and Ito's reference [3] and Chris Godsil's notes [9] for a more thorough treatment.

Definition 2 (Association Schemes) *A* symmetric association scheme *is a collection of $d+1$ binary $|X| \times |X|$ matrices $\mathcal{A} = \{A_0, A_1, \cdots, A_d\}$ over a set X that satisfy the following axioms:*

1. *A_i is symmetric for all $0 \leq i \leq d$,*
2. *$A_0 = I$ where I is the identity matrix,*
3. *$\sum_{i=0}^{d} A_i = J$ where J is the all-ones matrix, and*
4. *$A_iA_j = A_jA_i \in Span\{A_0, A_1, \cdots, A_d\} =: \mathfrak{A}$ for all $0 \leq i, j \leq d$.*

The matrices $A_1, A_2, \cdots, A_d$ are called the associates*, and the algebra $\mathfrak{A}$ is called the* Bose–Mesner *algebra of the association scheme. Moreover, $\mathfrak{A}$ admits a unique dual basis of* primitive idempotents *$E_0, E_1, \cdots, E_d$, i.e., $E_i^2 = E_i$ for all $0 \leq i \leq d$ and $\sum_{i=0}^{d} E_i = I$.*

Since the permutation representation of $S_n \times S_m$ acting on $S_{n,m}$ is multiplicity-free (see Theorem 7), the orbits $A_0, A_1, \cdots, A_d$ (so-called *orbitals*) of the action of $S_n \times S_m$ on ordered

pairs $S_{n,m} \times S_{n,m}$ forms a *symmetric association scheme* (see [3] for a proof). We abuse the notation, and also use A_i to denote the binary matrix with entries 1 corresponding to exactly those pairs that are in the orbit A_i. Let $\mathcal{A}_{n,m} := \{I, A_1, \cdots, A_d\}$ denote the *n, m-injection association scheme*.

Although it is well-known that permutation representation of $S_n \times S_m$ acting on $S_{n,m}$ is multiplicity-free (see [1, 6, 16, 10] for example), the parameters of its corresponding association scheme $\mathcal{A}_{n,m}$ have not yet been fully worked out. We now give a more in-depth treatment of the injection association scheme.

5.1 The Associates

The following is a more combinatorial definition of the associates of $\mathcal{A}_{n,m}$ that gives a combinatorial bijection between the associates of $\mathcal{A}_{n,m}$ and $\mathrm{Irr}(S_{n,m})$, which are the eigenspaces of the association scheme. The bijection is readily observed by thinking of each element of $S_{n,m}$ graphically as a maximum matching of the complete bipartite graph $K_{n,m}$ (see Figure 3).

Recall that $f_{\mathrm{id}} = (1, 2, \cdots, n)$ is the identity injection, which we can view as the maximum matching of $K_{n,m}$ that pairs 1 with 1, 2 with 2, and so on (e.g., the red matching in Figure 3). For any two maximum matchings f, f' of $K_{n,m}$, let $G(f, f')$ be the multigraph whose edge multiset is the multiset union $f \cup f'$. Clearly $G(f, f') = G(f', f)$ and this graph is composed of disjoint even cycles and disjoint even paths. Let c denote the number of disjoint cycles and let $2\lambda_i$ denote the length of an even cycle. Let p denote the number of disjoint paths and let $2\rho_i$ denote the length of an even path. If we order the cycles and paths respectively from longest to shortest and divide each of their lengths by two, assuming $m \geq 2n$, we see that the graphs $G(f, f')$ are in bijection (up to graph isomorphism) with pairs $(\lambda|\rho)$ of integer partitions $\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_c), \rho = (\rho_1, \rho_2, \cdots, \rho_p)$ such that $(\lambda_1, \cdots, \lambda_c, \rho_1, \cdots, \rho_p) \vdash n$. Let $\mathrm{d}(f, f') := (\lambda|\rho)$ denote this bijection, which we refer to as the *cycle-path type of f' with respect to f*. Note that $\mathrm{d}(\sigma(f), \sigma(f')) = \mathrm{d}(f, f')$ for all injections f, f' and all $\sigma \in S_n \times S_m$. If one of the arguments is the identity matching, then we say $\mathrm{d}(f) := \mathrm{d}(f_{\mathrm{id}}, f)$ is *the cycle-path type of f*. Illustrations of the graphs $G_{(\varnothing|n)}$ and $G_{(n-1|1)}$, and $G_{(\varnothing|1^n)}$ are provided in Figure 3 where $n = 3$ and $m = 6$.

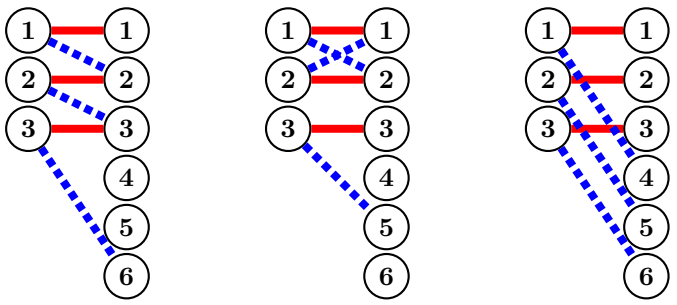
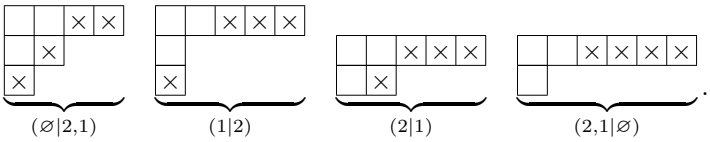


Fig. 3. $(2, 3, 6)$ on the left has type $(\varnothing|3)$, $(2, 1, 5)$ has type $(2|1)$, and $(4, 5, 6)$ has type $(\varnothing|1^3)$.

Recall that $|\lambda|$ denotes the size of the integer partition λ, i.e., the number of cells in its Young diagram, and $l(\lambda)$ denotes the number of parts of λ, i.e., the number of rows in its Young diagram. Let $\mathcal{C}_n := \{(\lambda|\rho) \colon |\lambda| + |\rho| = n\}$ where λ and ρ are partitions. When $m \geq 2n$, $\mathcal{C}_n$ is the set of all cycle-path types. Note that $(\varnothing|1^n)$ is not a cycle-path type when $m < 2n$,

and for $m = n$, all cycle-path types are of form $(\lambda|\varnothing)$, where $\lambda \vdash n$. We can decompose $\mathcal{C}_n$ as a disjoint union $\mathcal{C}_n = \bigcup_{k=0}^{n} \mathcal{C}_{n,k}$, where $\mathcal{C}_{n,k}$ consists of all $(\lambda|\rho) \in \mathcal{C}_n$ such that $l(\rho) = n - k$. Note that, for any two $f, f' \in S_{n,m}$, having $\mathrm{d}(f, f') \in \mathcal{C}_{n,k}$ implies $|\mathrm{im}\ f \cap \mathrm{im}\ f'| = k$.

Recall that any irrep in $\mathrm{Irr}(S_{n,m})$ is of the form $\lambda \otimes \lambda'$ where λ'/λ is a horizontal strip of size $m - n$. To see that cycle-path types $(\tau|\rho)$ have a natural correspondence with these irreducibles, consider a Young diagram of λ' such that the cells of λ'/λ are marked. Every columns of λ in λ' with a marked cell below it corresponds to a part in ρ whereas an unmarked column correspond to a part in τ. For instance, taking $\lambda = (2, 1)$ and $m = 7$, we have



Note that the marked singleton columns correspond to paths of length zero (i.e., isolated nodes). For each cycle-path type $(\tau|\rho)$, the $(\tau|\rho)$-*associate* of $\mathcal{A}_{n,m}$ is the following $m^{\underline{n}} \times m^{\underline{n}}$ binary matrix:

$$(A_{(\tau|\rho)})_{i,j} = \begin{cases} 1, & \text{if } \mathrm{d}(i,j) = (\tau|\rho) \\ 0, & \text{otherwise} \end{cases} \qquad \text{for all } i, j \in S_{m,n}.$$

5.2 The Valencies and Multiplicities

For each $0 \le i \le d$, let $d_i := \mathrm{tr} E_i$ denote the *multiplicity* of the ith eigenspace of an association scheme, that is, the dimension of its ith eigenspace. For each $0 \le i \le d$, define the *valency* v_i to be the row sum of an arbitrary row of A_i (equivalently, the largest eigenvalue of A_i). We now give formulas for the valencies $v_{(\lambda|\rho)}$ and multiplicities $d_{(\lambda|\rho)}$ of $\mathcal{A}_{n,m}$.

For each $(\lambda|\rho)$, define the $(\lambda|\rho)$-*sphere* to be the following set:

$$\Omega_{(\lambda|\rho)} := \{f \in S_{n,m} : \mathrm{d}(f) = (\lambda|\rho)\}.$$

The spheres partition $S_{n,m}$ and it useful to think of them as conjugacy classes. Indeed, when $n = m$, these spheres are the conjugacy classes of S_m. Note that $v_{(\lambda|\rho)} = |\Omega_{(\lambda|\rho)}|$, and basic combinatorial reasoning reveals the following.

Proposition 4 *For any cycle-path type $(\lambda|\rho)$, the size of the $(\lambda|\rho)$-sphere is*

$$v_{(\lambda|\rho)} = |\Omega_{(\lambda|\rho)}| = \frac{n!}{\prod_{i=1}^{n} i^{\ell_i} \ell_i! r_i!} (m-n)^{\underline{l(\rho)}}$$

where $\lambda = (n^{\ell_n}, \cdots, 1^{\ell_1})$, $\rho = (n^{r_n}, \cdots, 1^{r_1})$, and $l(\rho) = r_1 + \cdots + r_n$.

The multiplicities $d_{(\tau|\rho)}$ are easy to deduce due to the fact that each eigenspace of the scheme is isomorphic to an irrep $\mu \otimes \lambda$ of $S_n \times S_m$, and that $\dim \mu \otimes \lambda = \dim \mu \cdot \dim \lambda$. As we have seen, these dimensions are counted by the hook rule. In particular, for a cycle-path type $(\tau|\rho)$, let $\tau \cup \rho$ be the union of the set of parts of the two partitions. Then we have $d_{(\tau|\rho)} = d_{\lambda \otimes \lambda'}$ such that $\lambda = (\tau \cup \rho)^{\top} \vdash n$, $\lambda' = (\tau \cup (m - n, \rho^{\top})^{\top})^{\top}$, and '$\top$' denotes the transpose partition.

5.3 Stirling Numbers and Path Covers

The proof of the main result will rely on some combinatorial estimates of sizes of certain unions of spheres. These sizes are closely related to the (unsigned) Stirling numbers of the first kind. Recall that for any positive integers n, k, the *(unsigned) Stirling numbers of the first kind* are defined by the following recurrence:

$$\begin{bmatrix}0\\0\end{bmatrix} = 1; \quad \begin{bmatrix}n\\0\end{bmatrix} = \begin{bmatrix}0\\n\end{bmatrix} = 0; \quad \begin{bmatrix}n+1\\k\end{bmatrix} = n\begin{bmatrix}n\\k\end{bmatrix} + \begin{bmatrix}n\\k-1\end{bmatrix}.$$

A *directed cycle cover* of a directed graph is a union of directed cycles that partition the vertices of the graph. It is well-known that $\left[\begin{smallmatrix}n\\k\end{smallmatrix}\right]$ counts the number of directed cycle covers of the complete directed graph $\vec{K}_n := ([n], [n] \times [n])$ on n vertices and n^2 arcs that have precisely k cycles. The following upper bound is also well-known:

$$\begin{bmatrix}n\\n-k\end{bmatrix} \leq \frac{n^{2k}}{2^k k!}. \tag{8}$$

A *directed path cover* of a directed graph is a union of directed paths that partition the vertices of the graph. Let $\left[\begin{smallmatrix}n\\k\end{smallmatrix}\right]'$ denote the number of directed path covers of $\vec{K}_n$ that have precisely k paths, where an isolated vertex is considered a trivial path. These numbers are given by the following recurrence:

$$\begin{bmatrix}0\\0\end{bmatrix}' = 1; \quad \begin{bmatrix}n\\0\end{bmatrix}' = \begin{bmatrix}0\\n\end{bmatrix}' = 0; \quad \begin{bmatrix}n+1\\k\end{bmatrix}' = (n+k)\begin{bmatrix}n\\k\end{bmatrix}' + \begin{bmatrix}n\\k-1\end{bmatrix}'.$$

Indeed, there are $(n+k)$ ways of extending any given directed path cover of $\vec{K}_n$ with k paths to a directed path cover of $\vec{K}_{n+1}$ with k paths. It is not hard to show an upper bound akin to (8).

Proposition 5 *For all $k \leq n$ we have*

$$\begin{bmatrix}n\\n-k\end{bmatrix}' \leq \frac{(2n-k)^{2k}}{2^k k!} \leq 2^k \left(\frac{n^{2k}}{k!}\right).$$

Proof. By induction, we have

$$\begin{aligned}
\begin{bmatrix} n \\ n-k \end{bmatrix}' &= (2n-k-1)\begin{bmatrix} n-1 \\ n-k \end{bmatrix}' + \begin{bmatrix} n-1 \\ n-k-1 \end{bmatrix}' \\
&= (2n-k-1)\begin{bmatrix} n-1 \\ (n-1)-(k-1) \end{bmatrix}' + \begin{bmatrix} n-1 \\ (n-1)-k \end{bmatrix}' \\
&\leq (2n-k-1)\frac{(2n-k-1)^{2(k-1)}}{2^{k-1}(k-1)!} + \frac{(2n-k-2)^{2k}}{2^k k!} \\
&\leq (2n-k-1)\frac{(2n-k-1)^{2(k-1)}}{2^{k-1}(k-1)!} + \frac{(2n-k-1)^{2k}}{2^k k!} \\
&= \frac{1}{2^k k!}(2k+(2n-k-1))(2n-k-1)^{2k-1} \\
&\overset{\text{AM-GM}}{\leq} \frac{1}{2^k k!}\left(\frac{(2k+(2n-k-1))+(2k-1)(2n-k-1)}{2k}\right)^{2k} \\
&= \frac{(2n-k)^{2k}}{2^k k!} \\
&\leq 2^k\left(\frac{n^{2k}}{k!}\right)
\end{aligned}$$

□.

For the remainder of this section, assume that $m \geq n^{2+\alpha}$ for some $\alpha > 0$. For any injection $f \in S_{n,m}$, define the *path support* of f to be the subset of vertices of $[n]$ that belong to a path of $f \cup f_{\mathsf{id}}$. Let $S_{n,m,k} \subseteq S_{n,m}$ be the set of injections $f \in S_{n,m}$ such that $\mathrm{d}(f) \in \mathcal{C}_{n,k}$, i.e., $f \cup f_{\mathsf{id}}$ has exactly $n-k$ non-trivial paths. Let $S_{n,m,k,j} \subseteq S_{n,m,k}$ be the set of injections $f \in S_{n,m,k}$ with a path support of size $n-j$. Define the probabilities

$$p_k := p_{n,m,k} = |S_{n,m,k}|/m^{\underline{n}} \quad \text{and} \quad p_{k,j} := p_{n,m,k,j} = |S_{n,m,k,j}|/m^{\underline{n}},$$

so that $\sum_{j=0}^{k} p_{k,j} = p_k$ and $\sum_{k=0}^{n} p_k = 1$. In what follows, we have $j \leq k \leq n$, and we shall think of these probabilities p_k and $p_{k,j}$ as being functions of n. Basic combinatorial reasoning shows that

$$|S_{n,m,k,j}| = \binom{n}{j} j! \begin{bmatrix} n-j \\ n-k \end{bmatrix}' (m-n)^{\underline{n-k}}, \quad \text{thus} \quad p_{k,j} = \frac{\begin{bmatrix} n-j \\ n-k \end{bmatrix}' n^{\underline{j}}\ (m-n)^{\underline{n-k}}}{m^{\underline{k}} \cdot (m-k)^{\underline{n-k}}} \leq \begin{bmatrix} n-j \\ n-k \end{bmatrix}' \frac{n^{\underline{j}}}{m^{\underline{k}}}.$$

By Proposition 5, we have

$$p_{k,j} \leq \begin{bmatrix} n-j \\ n-k \end{bmatrix}' \frac{n^{\underline{j}}}{m^{\underline{k}}} = \begin{bmatrix} (n-j) \\ (n-j)-(k-j) \end{bmatrix}' \frac{n^{\underline{j}}}{m^{\underline{k}}} \leq 2^{k-j}\frac{(n-j)^{2(k-j)}}{(k-j)!}\frac{n^{\underline{j}}}{m^{\underline{k}}} \leq 4\left(\frac{(n^2)^{k-j/2}}{(n^{2+\alpha}-n)^k}\right).$$

For sufficiently large n, it is immediate for any positive integer k that $p_{k,j}$ is maximized when $j = 0$. In particular, we have the following proposition.

Proposition 6 *Let t be a positive integer. For all $k \geq t$ and $0 \leq j \leq k$, we have $p_{k,j} = \mathrm{O}\left(1/n^{t\alpha}\right)$.*

5.4 The Johnson Ordering of $\mathcal{A}_{n,m}$

The *Johnson scheme* $\mathcal{J}(m,n)$ is a symmetric association scheme defined over the n-subsets of $[m]$. The ith associate $A_i \in \mathcal{J}(m,n)$ of the Johnson scheme is defined such that $(A_i)_{X,Y} = 1$ if $n - |X \cap Y| = i$, and is 0 otherwise for any two n-subsets X, Y. It is well-known that the ith eigenspace of $\mathcal{J}(m,n)$ is isomorphic to the S_m-irrep associated to the partition $(m-i,i) \vdash m$. For proofs of these facts and more, see [8]. Henceforth, let E_i be the primitive idempotent of the Johnson scheme that projects onto $\mathcal{V}_{(m-i,i)}$.

There exists a natural ordering of the $S_{n,m}$ that we call *the Johnson ordering* that shows the Johnson scheme is a quotient of $\mathcal{A}_{n,m}$. First, we order $S_{n,m}$ by the corresponding n-subsets (the particular order does not matter). Next, we lexicographically order all $n!$ injections that map to the same n-subset (i.e., share the same image), e.g., for $n = 3$, we have:

$$(1,2,3),(1,3,2),(2,1,3),(2,3,1),(3,1,2),(3,2,1),(1,2,4),(1,4,2),(2,1,4),(2,4,1),\cdots .$$

Both S_n and S_m act on the domain and range of an injection respectively, and so each of their actions correspond to some collection of $m^{\underline{n}} \times m^{\underline{n}}$ permutation matrices (i.e., their corresponding permutation representations). The action of S_m on $S_{m,n}$ is transitive, but S_n's action has $\binom{m}{n}$ orbits, one for each n-subset. Note that on all $n!$ permutations of S_n corresponding to any given n-subset, the action of S_n corresponds to the regular representation of S_n.

Given $\lambda \vdash n$ and $\lambda' \vdash m$, let E_λ and $E_{\lambda'}$ be the orthogonal projectors on the λ-isotypic and λ'-isotypic subspaces, respectively. Since the actions of S_n and S_m on $S_{n,m}$ commute, E_λ and $E_{\lambda'}$ also commute, and we have $E_{\lambda\otimes\lambda'} = E_\lambda E_{\lambda'}$. From the specific way we ordered $S_{n,m}$ in the previous paragraph, for any $\lambda \vdash n$ we have

$$E_\lambda = I_{\binom{m}{n}} \otimes F_\lambda = \underbrace{F_\lambda \oplus F_\lambda \oplus \cdots \oplus F_\lambda}_{\binom{m}{n} \text{ times}},$$

where F_λ is the $n! \times n!$ orthogonal projector on the λ-isotypic subspace of the regular representation of S_n. Hence, we can write $E_{\lambda\otimes\lambda'}$ as a product of two block matrices:

$$E_{\lambda\otimes\lambda'} = \begin{pmatrix} B^{\lambda'}_{1,1} & B^{\lambda'}_{1,2} & \cdots \\ B^{\lambda'}_{2,1} & B^{\lambda'}_{2,2} & \\ \vdots & & \ddots \end{pmatrix} \begin{pmatrix} F_\lambda & 0 & \cdots \\ 0 & F_\lambda & \\ \vdots & & \ddots \end{pmatrix} = \begin{pmatrix} B^{\lambda'}_{1,1}F_\lambda & B^{\lambda'}_{1,2}F_\lambda & \cdots \\ B^{\lambda'}_{2,1}F_\lambda & B^{\lambda'}_{2,2}F_\lambda & \\ \vdots & & \ddots \end{pmatrix},$$

where the first matrix is $E_{\lambda'}$, in which each block $B^{\lambda'}_{i,j}$ is some $n! \times n!$ matrix. For every $\lambda \vdash n$, we have $E_{\lambda\otimes\lambda'} = E_\lambda E_{\lambda'} E_\lambda$, which means that, if the rank of F_λ is 1, then $E_{\lambda\otimes\lambda'}$ can be expressed as a tensor product with one of the factors being F_λ. When $\lambda = (n)$, we have $F_\lambda = F_{(n)} = J/n!$, where J is the $n! \times n!$ all-ones matrix. Thus, from the expression above, we have

$$E_{(n)\otimes(m-1,1)} = J/n! \otimes E_1 = \begin{pmatrix} b_{1,1}J & b_{1,2}J & \cdots \\ b_{2,1}J & b_{2,2}J & \\ \vdots & & \ddots \end{pmatrix}, \tag{9}$$

where $b_{i,j}$ are scalars, $1/m^{\underline{n}}$ times the dual eigenvalues $q_1(\cdot)$ (see Lemma 1 for explicit expressions). Thus we have

$$E_{(n)\otimes(m-1,1)} \circ E_{\lambda\otimes\lambda'} = \begin{pmatrix} b_{1,1}B^{\lambda'}_{1,1}F_\lambda & b_{1,2}B^{\lambda'}_{1,2}F_\lambda & \cdots \\ b_{2,1}B^{\lambda'}_{2,1}F_\lambda & b_{2,2}B^{\lambda'}_{2,2}F_\lambda & \\ \vdots & & \ddots \end{pmatrix}, \tag{10}$$

which is orthogonal to E_μ (and thus $E_{\mu\otimes\mu'}$) for all $\mu \vdash n$ such that $\mu \neq \lambda$ because of $F_\lambda F_\mu = 0$.

5.5 The Dual Eigenvalues of $\mathcal{A}_{n,m}$

It is well-known that the primitive idempotents E_i of an association scheme can be written as a unique linear combination of associates A_i of the scheme (see [9, Ch. 2.1]):

$$E_i = \frac{1}{|X|}\sum_{j=0}^{d} q_i(j)A_j.$$

The $q_i(j)$'s are called the *dual eigenvalues* of the scheme. In our case, this specializes to

$$E_{\lambda\otimes\lambda'} = \frac{1}{m^{\underline{n}}}\sum_{(\mu|\rho)\in\mathcal{C}_n} q_{\lambda\otimes\lambda'}(\mu|\rho)A_{(\mu|\rho)},$$

and these coefficients $q_{\lambda\otimes\lambda'}(\mu|\rho)$ are the *dual eigenvalues* of $\mathcal{A}_{m,n}$.

Proposition 7 [6] *For any finite symmetric Gelfand pair (G, K), the dual eigenvalues $q_i(j)$ of the symmetric association scheme on $X = G/K$ can be written as*

$$q_i(j) = d_i\omega^i_j$$

where d_i is the dimension of irreducible i corresponding to the spherical function ω^i.

Let us consider matrices in the Bose–Mesner algebra $\mathfrak{A}_{n,m}$ of the injection association scheme. By symmetry, every such matrix can be specified by a row or column corresponding to a single injection. Note that

$$(E_{\lambda\otimes\lambda'})_{f,h} = \frac{q_{\lambda\otimes\lambda'}(\mathrm{d}(f,h))}{m^{\underline{n}}}$$

and that

$$\sum_{f\in S_{n,m}} q_{\lambda\otimes\lambda'}(\mathrm{d}(f,h))\ 1_f = m^{\underline{n}}\,(E_{\lambda\otimes\lambda'})_h \in \lambda\otimes\lambda'$$

for all $\lambda\otimes\lambda'$ and $f, h \in S_{n,m}$, where $1_f \in \mathbb{C}[S_{n,m}]$ denotes the binary unit vector with the unique 1 in position f. It is well-known that the projector $E_{\lambda'}$ onto the λ'-isotypic component can be written as

$$(E_{\lambda'})_{f,h} = \frac{d_{\lambda'}}{m!}\sum_{\sigma\in S_m}\chi_{\lambda'}(\sigma^{-1})(V_\sigma)_{f,h}$$

where $V_\sigma : 1_f \mapsto 1_{\sigma * f}$ for all $f, h \in S_{n,m}$ and $\chi_{\lambda'}$ is the character corresponding to λ'. The foregoing, and the fact that $E_{\lambda'}E_{\lambda\otimes\lambda'} = E_{\lambda\otimes\lambda'}$ implies the following proposition.

Proposition 8 *For any $f, h \in S_{n,m}$, $\lambda \vdash n$, and $\lambda' \vdash m$, we have*

$$q_{\lambda\otimes\lambda'}\big(\mathrm{d}(f,h)\big) = \frac{d_{\lambda'}}{m!}\sum_{\sigma\in S_m}\chi_{\lambda'}(\sigma)q_{\lambda\otimes\lambda'}\big(\mathrm{d}(\sigma^{-1}*f,h)\big).$$

Proof. We have $\chi_{\lambda'}(\sigma^{-1}) = \chi_{\lambda'}(\sigma)$. By applying $E_{\lambda'}$ to $m^{\underline{n}}\left(E_{\lambda\otimes\lambda'}\right)_h$, we get

$$\begin{aligned} m^{\underline{n}}\left(E_{\lambda'}E_{\lambda\otimes\lambda'}\right)_h &= \sum_{f\in S_{n,m}}\frac{d_{\lambda'}}{m!}\sum_{\sigma\in S_m}\chi_{\lambda'}(\sigma)q_{\lambda\otimes\lambda'}(\mathrm{d}(f,h))\ 1_{\sigma*f} \\ &= \sum_{f\in S_{n,m}}\frac{d_{\lambda'}}{m!}\sum_{\sigma\in S_m}\chi_{\lambda'}(\sigma)q_{\lambda\otimes\lambda'}(\mathrm{d}(\sigma^{-1}*f,h))\ 1_f. \end{aligned}$$

The proposition follows by equating the coefficients of 1_f in the expression for $m^{\underline{n}}\left(E_{\lambda\otimes\lambda'}\right)_h$ and the expression for $m^{\underline{n}}\left(E_{\lambda'}E_{\lambda\otimes\lambda'}\right)_h$ $\square$.

Lemma 1 *Let $q_i(j)$ be a dual eigenvalue of the Johnson scheme $\mathcal{J}(m,n)$. Then we have*

$$q_1(j) = \frac{\binom{m}{n}}{\binom{m-2}{n-1}}\left(n-j-\frac{n^2}{m}\right).$$

Moreover, if $m \geq n^2$, then $q_1(j) \geq 0$ for all $j \neq n$.

Proof. Let $p_i(j)$ denote the j-th eigenvalue of the i-th associate of the Johnson scheme $\mathcal{J}(m,n)$. It is well-known (see [8] for example) that

$$p_i(j) = \sum_{r=i}^{n}(-1)^{(r-i+j)}\binom{r}{i}\binom{m-2r}{n-r}\binom{m-r-j}{r-j}.$$

Using basic relations between primal and dual eigenvalues (see [9]), we can write the first primitive idempotent E_1 of the Johnson scheme as follows:

$$\begin{aligned} E_1 &= \frac{1}{\binom{m}{n}}\sum_{j=0}^{n}q_1(j)A_j \\ &= \frac{m-1}{\binom{m}{n}}\sum_{j=0}^{n}\frac{p_j(1)}{\binom{n}{j}\binom{m-n}{j}}A_j \\ &= \frac{m-1}{\binom{m}{n}}\sum_{j=0}^{n}\frac{1}{\binom{n}{j}\binom{m-n}{j}}\left(\binom{n-1}{j}\binom{m-n-1}{j}-\binom{n-1}{j-1}\binom{m-n-1}{j-1}\right)A_j \\ &= \frac{m-1}{\binom{m}{n}}\sum_{j=0}^{n}\left(1-\frac{mj}{n(m-n)}\right)A_j \\ &= \frac{1}{\binom{m-2}{n-1}}\sum_{j=0}^{n}\left(n-j-\frac{n^2}{m}\right)A_j. \end{aligned}$$

Equating coefficients of A_j gives the result $\square$.

Lemma 2 *For all $\mu \in \mathcal{C}_{n,k}$, we have*

$$q_{(n)\otimes(m-1,1)}(\mu) = \frac{(km-n^2)(m-1)}{n(m-n)}.$$

Proof. Recall that the ith eigenspace of the Johnson scheme $\mathcal{J}(m,n)$ is isomorphic to the S_m-irrep associated to the partition $(m-i,i) \vdash m$, and that E_i denotes the primitive idempotent of the Johnson scheme that projects onto its $(m-i,i)$ eigenspace. In the proof of the previous lemma we saw that

$$E_1 = \frac{1}{\binom{m}{n}} \sum_{j=0}^{n} \left[\frac{\binom{m}{n}}{\binom{m-2}{n-1}} \left(n-j-\frac{n^2}{m} \right) \right] A_j = \frac{1}{\binom{m-2}{n-1}} \sum_{j=0}^{n} \left(n-j-\frac{n^2}{m} \right) A_j.$$

Since $(A_j)_{X,Y} = 1$ only if $|X \cap Y| = n-j$, we have the dual eigenvalue

$$\left(m^{\underline{n}} E_{(n)\otimes(m-1,1)}\right)_{f,h} = m^{\underline{n}} \frac{|\mathrm{im}\ f \cap \mathrm{im}\ h| - n^2/m}{n!\binom{m-2}{n-1}} = \frac{(m-1)(|\mathrm{im}\ f \cap \mathrm{im}\ h|m - n^2)}{n(m-n)}.$$

This value is clearly the same for all pairs of f and h that have the same $|\mathrm{im}\ f \cap \mathrm{im}\ h|$, in other words, for all pairs of f and h for which the cycle-path type $\mathrm{d}(f,h)$ is in the same $\mathcal{C}_{n,k}$. Therefore, it follows that

$$q_{(n)\otimes(m-1,1)}(\mu) = \frac{(km-n^2)(m-1)}{n(m-n)}$$

for all $\mu \in \mathcal{C}_{n,k}$, which completes the proof $\square$.

6 A sufficient condition on Krein parameters

In this section, we reduce the proof of Theorem 1 to an upper bound on certain Krein parameters of $\mathcal{A}_{n,m}$. Recall that $\circ$ denotes the Schur (entrywise) product of two matrices.

Definition 3 (Krein Parameters) *Let $\mathcal{A}$ be an association scheme on v vertices with d associates. For any $0 \le i,j \le d$, there exist constants $q_{i,j}(k)$ such that*

$$E_i \circ E_j = \frac{1}{v} \sum_{k=0}^{d} q_{i,j}(k) E_k,$$

which are called the Krein parameters *of $\mathcal{A}$. More explicitly, we have*

$$q_{i,j}(k) = v \frac{\mathrm{tr}\left[E_k(E_i \circ E_j)\right]}{d_k}.$$

The Krein parameters can alternatively be written as

$$q_{i,j}(k) = \frac{1}{vd_k} \sum_{\ell=0}^{d} \frac{q_i(\ell) q_j(\ell) \overline{q_k(\ell)}}{v_\ell} = \frac{d_i d_j}{v} \sum_{\ell=0}^{d} \frac{\overline{p_i(\ell) p_j(\ell)} p_k(\ell)}{v_\ell^2}, \tag{11}$$

where $p_i(j)$ denotes the j-th eigenvalue of A_i and $q_i(j)$ denotes the jth dual eigenvalue of E_i (see [9, Chap. 2.4] for a proof).

To prove the lower bound on non-coherent INDEX ERASURE, we use the same adversary matrix Γ as [1] used for the coherent case.[e] For simplifying the equations, without loss of generality let us assume that n is a square. As in [1], we choose

$$\Gamma := \sum_{k=0}^{\sqrt{n}-1} (\sqrt{n}-k) \sum_{\lambda \vdash k} E_{(n-k,\lambda)\otimes(m-k,\lambda)},$$

and thus the orthogonal projection onto its image is

$$\Pi_\Gamma := \sum_{\lambda\colon |\lambda|<\sqrt{n}} E_{(n-|\lambda|,\lambda)\otimes(m-|\lambda|,\lambda)}.$$

Note that the sole principal eigenvector ω of Γ is the uniform superposition over $\mathcal{F}$ (i.e., $\omega_f = 1/\sqrt{m^{\underline{n}}}$ for all $f \in \mathcal{F}$). Thus, as per Corollary 1, we are interested in the quantity

$$\eta = \max_{T\in\mathcal{T}} \operatorname{tr}\left[\Pi_\Gamma \frac{(T \circ T^{\odot})}{m^{\underline{n}}}\right].$$

As described by the automorphism principle (Theorem 3), here it suffices to consider T that are $(S_n \times S_m)$-invariant, that is, T that belong the Bose–Mesner algebra $\mathfrak{A}_{n,m}$. Because of that, for simplicity, let us redefine to $\mathcal{T}$ be the set of all state matrices in $\mathfrak{A}_{n,m}$.

For any primitive idempotent $E_{\lambda\otimes\lambda'}$, let

$$T_{\lambda\otimes\lambda'} := \left(\frac{m^{\underline{n}}}{\operatorname{tr} E_{\lambda\otimes\lambda'}}\right) E_{\lambda\otimes\lambda'} = \left(\frac{m^{\underline{n}}}{d_{\lambda\otimes\lambda'}}\right) E_{\lambda\otimes\lambda'}$$

be its associated state matrix. In [1] it is shown that the target matrix can be written as

$$T^{\odot} = \frac{n}{m} T_{(n)\otimes(m)} + \left(1-\frac{n}{m}\right) T_{(n)\otimes(m-1,1)}.$$

In the coherent case, recall that $T = J$, and therefore

$$\eta = \operatorname{tr}\left[\Pi_\Gamma \frac{T^{\odot}}{m^{\underline{n}}}\right] = \frac{n}{m} \operatorname{tr}\left[\frac{T_{(n)\otimes(m)}}{m^{\underline{n}}}\right] = \frac{n}{m}.$$

The most technically involved part of the proof of the lower bound by [1] is proving that $\|\Delta_x \circ \Gamma\| = \mathrm{O}(1)$. Since we are using the same adversary matrix Γ, we already have the above bound on $\|\Delta_x \circ \Gamma\|$. Our goal is to show that $\operatorname{tr}[\Pi_\Gamma (T \circ T^{\odot})/m^{\underline{n}}]$ is small for *all* state matrices T.

By dividing the elements in the set $\mathcal{T}$ by $m^{\underline{n}}$ we obtain the set of all density matrices (positive-semidefinite Hermitian matrices with trace 1) of the Bose–Mesner algebra $\mathfrak{A}_{n,m}$. Observe that $(T \circ T')/m^{\underline{n}}$ is a density matrix for all $T, T' \in \mathcal{T}$. For any $T \in \mathcal{T}$, we have

$$\operatorname{tr}\left[\Pi_\Gamma \frac{\left(T \circ T_{(n)\otimes(m)}\right)}{m^{\underline{n}}}\right] = \operatorname{tr}\left[\Pi_\Gamma \frac{T}{m^{\underline{n}}}\right] \le 1,$$

[e] Technically, the adversary matrix used here is $\sqrt{n}$ times that of [1] as the adversary method they use places slightly different conditions on the adversary matrix.

therefore

$$\operatorname{tr}\left[\Pi_\Gamma \frac{(T \circ T^{\odot})}{m^{\underline{n}}}\right] \leq \frac{n}{m} + \left(1 - \frac{n}{m}\right) \operatorname{tr}\left[\Pi_\Gamma \frac{\left(T \circ T_{(n)\otimes(m-1,1)}\right)}{m^{\underline{n}}}\right].$$

Our goal is to bound the latter term:

$$\left(1 - \frac{n}{m}\right) \operatorname{tr}\left[\Pi_\Gamma \frac{\left(T \circ T_{(n)\otimes(m-1,1)}\right)}{m^{\underline{n}}}\right]. \tag{12}$$

Note that $\mathcal{T} = \left\{\sum_\chi c_\lambda T_\chi \colon \sum_\chi c_\chi = 1\ ,\ c_\chi \geq 0\right\}$, where the sums range over $\chi \in \operatorname{Irr}(S_{n,m})$. Hence,

$$\begin{aligned}\max_{T\in\mathcal{T}} \operatorname{tr}\left[\Pi_\Gamma \frac{\left(T \circ T_{(n)\otimes(m-1,1)}\right)}{m^{\underline{n}}}\right] &= \max_{\substack{\{c_\chi \geq 0\}_\chi \\ \sum_\chi c_\chi = 1}} c_\chi \operatorname{tr}\left[\Pi_\Gamma \frac{\left(T_\chi \circ T_{(n)\otimes(m-1,1)}\right)}{m^{\underline{n}}}\right] \\ &= \max_\chi \operatorname{tr}\left[\Pi_\Gamma \frac{\left(T_\chi \circ T_{(n)\otimes(m-1,1)}\right)}{m^{\underline{n}}}\right].\end{aligned}$$

The following proposition simplifies (12).

Proposition 9 *For any $\lambda = (n - |\nu|, \nu)$ and $\bar{\lambda} = (m - |\nu|, \nu)$, we have*

$$\operatorname{tr}\left[\Pi_\Gamma \frac{\left(T_{\lambda\otimes\lambda'} \circ T_{(n)\otimes(m-1,1)}\right)}{m^{\underline{n}}}\right] = \operatorname{tr}\left[E_{\lambda\otimes\bar{\lambda}} \frac{\left(T_{\lambda\otimes\lambda'} \circ T_{(n)\otimes(m-1,1)}\right)}{m^{\underline{n}}}\right].$$

Moreover, if $|\nu| \geq \sqrt{n}$, then $\operatorname{tr}\left[\Pi_\Gamma \frac{\left(T_{\lambda\otimes\lambda'} \circ T_{(n)\otimes(m-1,1)}\right)}{m^{\underline{n}}}\right] = 0.$

Proof. By Equation (10), if $|\nu| < \sqrt{n}$, then we have

$$\begin{aligned}\Pi_\Gamma \frac{\left(T_{\lambda\otimes\lambda'} \circ T_{(n)\otimes(m-1,1)}\right)}{m^{\underline{n}}} &= \sum_{\mu\colon |\mu|<\sqrt{n}} E_{(n-|\mu|,\mu)\otimes(m-|\mu|,\mu)} \frac{\left(T_{\lambda\otimes\lambda'} \circ T_{(n)\otimes(m-1,1)}\right)}{m^{\underline{n}}} \\ &= E_{(n-|\nu|,\nu)\otimes(m-|\nu|,\nu)} \frac{\left(T_{\lambda\otimes\lambda'} \circ T_{(n)\otimes(m-1,1)}\right)}{m^{\underline{n}}} \\ &= E_{\lambda\otimes\bar{\lambda}} \frac{\left(T_{\lambda\otimes\lambda'} \circ T_{(n)\otimes(m-1,1)}\right)}{m^{\underline{n}}}.\end{aligned}$$

The second part of the proof is immediate from the definition of Γ $\square$.
The proposition above now allows us to bound (12) for all $\lambda \vdash n$ such that $|\lambda| - \lambda_1 \leq \sqrt{n}$.

Corollary 3 *Suppose $\lambda \vdash n$ has no more than $\sqrt{n}$ cells below the first row. Then for all $\lambda' \neq \bar{\lambda}$ we have*

$$\operatorname{tr}\left[E_{\lambda\otimes\bar{\lambda}} \frac{\left(T_{(n)\otimes(m-1,1)} \circ T_{\lambda\otimes\lambda'}\right)}{m^{\underline{n}}}\right] \in \mathrm{O}(\sqrt{n}/m).$$

Proof. Let sum$[\cdot]$ denote the sum of the entries of the matrix. We have $T_{\lambda\otimes\lambda''}/m^{\underline{n}} = E_{\lambda\otimes\lambda''}/d_{\lambda\otimes\lambda''}$ for all $\lambda'' \vdash m$. Therefore

$$\begin{aligned}\operatorname{tr}\left[E_{\lambda\otimes\overline{\lambda}}\frac{\left(T_{(n)\otimes(m-1,1)}\circ T_{\lambda\otimes\lambda'}\right)}{m^{\underline{n}}}\right] &= \frac{1}{d_{\lambda\otimes\lambda'}}\operatorname{sum}\left[E_{\lambda\otimes\overline{\lambda}}\circ T_{(n)\otimes(m-1,1)}\circ E_{\lambda\otimes\lambda'}\right]\\ &= \frac{d_{\lambda\otimes\overline{\lambda}}}{d_{\lambda\otimes\lambda'}}\operatorname{tr}\left[E_{\lambda\otimes\lambda'}\frac{\left(T_{(n)\otimes(m-1,1)}\circ T_{\lambda\otimes\overline{\lambda}}\right)}{m^{\underline{n}}}\right].\end{aligned}$$

Since $\left(T_{(n)\otimes(m-1,1)}\circ T_{\lambda\otimes\overline{\lambda}}\right)/m^{\underline{n}}$ is a density matrix, this trace is at most 1, thus

$$\leq \frac{d_{\overline{\lambda}}}{d_{\lambda'}} \in \mathrm{O}(\sqrt{n}/m),$$

where the asymptotic bound follows from Corollary 2, completing the proof $\square$.
We therefore have

$$\eta = \mathrm{O}\left(\frac{n}{m} + \operatorname{tr}\left[E_{\lambda\otimes\bar{\lambda}}\frac{\left(T_{\lambda\otimes\bar{\lambda}}\circ T_{(n)\otimes(m-1,1)}\right)}{m^{\underline{n}}}\right]\right),$$

and it remains to bound the value

$$\begin{aligned}\operatorname{tr}\left[E_{\lambda\otimes\bar{\lambda}}\frac{\left(T_{\lambda\otimes\bar{\lambda}}\circ T_{(n)\otimes(m-1,1)}\right)}{m^{\underline{n}}}\right] &= \frac{m^{\underline{n}}}{(m-1)d_{\lambda\otimes\bar{\lambda}}}\operatorname{sum}\left[E_{(n)\otimes(m-1,1)}\circ E_{\lambda\otimes\bar{\lambda}}\circ E_{\lambda\otimes\bar{\lambda}}\right]\\ &= \frac{\sum_{\mu\in\mathcal{C}_n} v_\mu \cdot q_{(n)\otimes(m-1,1)}(\mu)\cdot q^2_{\lambda\otimes\bar{\lambda}}(\mu)}{m^{\underline{n}}(m-1)d_{\lambda\otimes\bar{\lambda}}} \qquad (13)\end{aligned}$$

for all $\lambda \vdash n$ with no more than $\sqrt{n}$ cells below the first row.

Thus to prove Theorem 1, it suffices to show for sufficiently large m, say $m \geq n^{3+\epsilon}$, that the value of (13), and thus η, is $\mathrm{O}(1/\sqrt{n})$ for all $\lambda \vdash n$ with no more than $\sqrt{n}$ cells below the first row. According to the expression (11) for Krein parameters, (13) equals

$$\frac{q_{\lambda\otimes\bar{\lambda},\,\lambda\otimes\bar{\lambda}}((n)\otimes(m-1,1))}{d_{\lambda\otimes\bar{\lambda}}} = \frac{q_{\lambda\otimes\bar{\lambda},\,(n)\otimes(m-1,1)}(\lambda\otimes\bar{\lambda})}{m-1},$$

and therefore the task of bounding (13) is equivalent to bounding the Krein parameters

$$q_{\lambda\otimes\bar{\lambda},\,\lambda\otimes\bar{\lambda}}((n)\otimes(m-1,1)) \quad \text{and} \quad q_{\lambda\otimes\bar{\lambda},\,(n)\otimes(m-1,1)}(\lambda\otimes\bar{\lambda}).$$

7 Proof of Theorem 1

We are now in a position to finish the proof of the main result. In Section 6 we saw that it suffices to show (13) is $\mathrm{O}(1/\sqrt{n})$ for sufficiently large m, which we state as the following claim.

Claim 3 *For all $\lambda \vdash n$ with $\ell \leq \sqrt{n}$ cells below the first row, there is a constant $\alpha > 1$ such that*

$$\frac{1}{m^{\underline{n}}} \sum_{(\mu|\rho)\in\mathcal{C}_n} \frac{v_{(\mu|\rho)} \cdot q_{(n)\otimes(m-1,1)}(\mu|\rho) \cdot q^2_{\lambda\otimes\bar{\lambda}}(\mu|\rho)}{d_{(n)\otimes(m-1,1)} \cdot d_{\lambda\otimes\bar{\lambda}}} = \mathrm{O}(1/\sqrt{n})$$

holds, provided $m \geq n^{2+\alpha}$.

We now prove this claim. Throughout the proof, at no loss of generality we shall assume that m and $\sqrt{n}$ are integers to avoid cumbersome notation.

Proof. Let $(\mu|\rho) \in \mathcal{C}_{n,k}$. Since $m \geq n^{2+\alpha}$, Lemma 2 implies that $q_{(n)\otimes(m-1,1)}(\mu|\rho) < 0$ if $k = 0$, and for $k > 0$ that

$$0 \leq q_{(n)\otimes(m-1,1)}(\mu|\rho) = \frac{(km - n^2)(m-1)}{n(m-n)} \leq \frac{k(m-n)(m-1)}{n(m-n)} = \frac{k(m-1)}{n}.$$

Note that $d_{(n)\otimes(m-1,1)} = (m-1)$. The $\mathcal{C}_{n,k}$'s partition $\mathcal{C}_n$, so we may rewrite the sum as

$$\frac{1}{m^{\underline{n}}} \sum_{(\mu|\rho)\in\mathcal{C}_n} \frac{v_{(\mu|\rho)} \cdot q_{(n)\otimes(m-1,1)}(\mu|\rho) \cdot q^2_{\lambda\otimes\bar{\lambda}}(\mu|\rho)}{d_{(n)\otimes(m-1,1)} \cdot d_{\lambda\otimes\bar{\lambda}}} \leq \frac{1}{m^{\underline{n}}} \sum_{k=1}^{n} \frac{k}{n} \sum_{(\mu|\rho)\in\mathcal{C}_{n,k}} \frac{v_{(\mu|\rho)} q^2_{\lambda\otimes\bar{\lambda}}(\mu|\rho)}{d_\lambda d_{\bar{\lambda}}}.$$

By Proposition 7, we can write the dual eigenvalue as

$$\begin{aligned} &= \frac{1}{m^{\underline{n}}} \sum_{k=1}^{n} \frac{k}{n} \sum_{(\mu|\rho)\in\mathcal{C}_{n,k}} \frac{v_{(\mu|\rho)} d^2_{\lambda\otimes\bar{\lambda}} \left[\omega^{\lambda\otimes\bar{\lambda}}_{(\mu|\rho)}\right]^2}{d_\lambda d_{\bar{\lambda}}} \\ &= \frac{d_\lambda d_{\bar{\lambda}}}{m^{\underline{n}}} \sum_{k=1}^{n} \frac{k}{n} \sum_{(\mu|\rho)\in\mathcal{C}_{n,k}} v_{(\mu|\rho)} \left[\omega^{\lambda\otimes\bar{\lambda}}_{(\mu|\rho)}\right]^2. \end{aligned}$$

For any $f \in S_{n,m}$, define k_f so that $n - k_f$ equals the number of non-trivial paths in the cycle-path type of f. From the foregoing, it suffices to show

$$\frac{d_\lambda d_{\bar{\lambda}}}{m^{\underline{n}}} \sum_{k=1}^{n} \frac{k}{n} \sum_{(\mu|\rho)\in\mathcal{C}_{n,k}} v_{(\mu|\rho)} \left[\omega^{\lambda\otimes\bar{\lambda}}_{(\mu|\rho)}\right]^2 = \frac{d_\lambda d_{\bar{\lambda}}}{m^{\underline{n}}} \sum_{f\in S_{n,m}} \frac{k_f}{n} \left[\omega^{\lambda\otimes\bar{\lambda}}(f)\omega^{\lambda\otimes\bar{\lambda}}(f)\right] = \mathrm{O}(1/\sqrt{n}). \tag{14}$$

By Proposition 3, we have

$$\frac{d_\lambda d_{\bar{\lambda}}}{m^{\underline{n}}} \sum_{f\in S_{n,m}} \mathrm{O}(1/\sqrt{n}) \left[\omega^{\lambda\otimes\bar{\lambda}}(f)\omega^{\lambda\otimes\bar{\lambda}}(f)\right] = \mathrm{O}(1/\sqrt{n});$$

therefore, it suffices to show there exists a constant $c > 0$ such that

$$\frac{d_\lambda d_{\bar{\lambda}}}{m^{\underline{n}}} \sum_{\substack{f\in S_{n,m} \\ k_f \geq c\sqrt{n}}} \omega^{\lambda\otimes\bar{\lambda}}(f)\omega^{\lambda\otimes\bar{\lambda}}(f) = \mathrm{O}(1/\sqrt{n}).$$

For all $0 \leq j \leq k$, define $\mathcal{C}_{n,k,j} := \{(\mu|\rho) \in \mathcal{C}_{n,k} : |\mu| = j\}$, so that $\mathcal{C}_{n,k} = \bigsqcup_{j=0}^{k} \mathcal{C}_{n,k,j}$. Recall that $p_{k,j}$ is the probability that the cycle-path type of $f \in S_{n,m}$ lies in $\mathcal{C}_{n,k,j}$. We have

$$\begin{aligned}
\frac{d_\lambda d_{\bar{\lambda}}}{m^{\underline{n}}} \sum_{\substack{f \in S_{n,m} \\ k_f \geq c\sqrt{n}}} \omega^{\lambda \otimes \bar{\lambda}}(f)\omega^{\lambda \otimes \bar{\lambda}}(f) &= \frac{d_\lambda d_{\bar{\lambda}}}{m^{\underline{n}}} \sum_{k=c\sqrt{n}}^{n} \sum_{j=0}^{k} \sum_{(\mu|\rho) \in \mathcal{C}_{n,k,j}} v_{(\mu|\rho)} |\omega^{\lambda \otimes \bar{\lambda}}_{(\mu|\rho)}|^2 \\
&\leq d_\lambda d_{\bar{\lambda}} \sum_{k=c\sqrt{n}}^{n} \sum_{j=0}^{k} \sum_{(\mu|\rho) \in \mathcal{C}_{n,k,j}} p_{(\mu|\rho)} \\
&= d_\lambda d_{\bar{\lambda}} \sum_{k=c\sqrt{n}}^{n} \sum_{j=0}^{k} p_{k,j},
\end{aligned}$$

where the inequality holds by Proposition 2. Let $c = 4$. Proposition 6 with $t = c\sqrt{n}$ implies

$$\leq d_\lambda d_{\bar{\lambda}} \cdot n^2 \cdot \mathrm{O}(1/n^{\alpha c\sqrt{n}}).$$

By Proposition 1, we have

$$\begin{aligned}
&\leq \mathrm{O}\left(n^2 \cdot \frac{(n^{2+\alpha})^\ell n^\ell}{n^{\alpha c\sqrt{n}}}\right) \\
&\leq \mathrm{O}\left(n^2 \cdot \frac{n^{3\ell}(n^\alpha)^\ell}{(n^\alpha)^{3\sqrt{n}}(n^\alpha)^{\sqrt{n}}}\right).
\end{aligned}$$

Since $\ell \leq \sqrt{n}$, setting $\alpha = 1 + \varepsilon$ gives us

$$\leq \mathrm{O}\left(n^2/n^{3\varepsilon\sqrt{n}}\right) = \mathrm{O}(1/\sqrt{n}).$$

This proves (14), and thus the claim, which completes the proof of the main result $\square$.

We have made no attempt to improve the dependency $m \geq n^{3+\varepsilon}$, and in fact, we believe that the claim above should be true for all $m \geq n(1+\varepsilon)$ such that $\varepsilon > 0$.

Acknowledgements

The authors would like to thank Aleksandrs Belovs, Chris Godsil, and Jérémie Roland for insightful discussions. The authors would also like to thank an anonymous referee for comments that substantially improved the readability of this paper. A.R. is supported by JSPS KAKENHI Grant Number JP20H05966 and MEXT Quantum Leap Flagship Program (MEXT Q-LEAP) Grant Number JPMXS0120319794. Part of this work was done while he was a JSPS International Research Fellow supported by the JSPS KAKENHI Grant Number JP19F19079, and when he was at the Centre for Quantum Technologies at the National University of Singapore supported by the Singapore Ministry of Education and the National Research Foundation under grant R-710-000-012-135.

References

1. A. Ambainis, L. Magnin, M. Roetteler, and J. Roland. Symmetry-assisted adversaries for quantum state generation. In *2011 IEEE 26th Annual Conference on Computational Complexity*, pages 167–177, June 2011.
2. A. Ambainis. Understanding quantum algorithms via query complexity. In *Proceedings of the 2018 International Congress of Mathematicians*, volume 3, pages 3249–3270, 2018.
3. E. Bannai and T. Ito. *Algebraic Combinatorics I: Association Schemes.* Mathematics lecture note series. Benjamin/Cummings Pub. Co., 1984.
4. A. Belovs and A. Rosmanis. Adversary lower bounds for the collision and the set equality problems. *Quantum Information & Computation*, 18:200–224, 2018.
5. G. Brassard, P. Høyer, and A. Tapp. Quantum cryptanalysis of hash and claw-free functions. In Cláudio L. Lucchesi and Arnaldo V. Moura, editors, *LATIN'98: Theoretical Informatics*, pages 163–169, Berlin, Heidelberg, 1998. Springer Berlin Heidelberg.
6. T. Ceccherini-Silberstein, F. Scarabotti, and F. Tolli. *Harmonic Analysis on Finite Groups: Representation Theory, Gelfand Pairs and Markov Chains.* Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2008.
7. P. Diaconis. *Group Representations in Probability and Statistics.* Institute of Mathematical Statistics Lecture Notes—Monograph Series, 11. Institute of Mathematical Statistics, Hayward, CA, 1988.
8. C. Godsil and K. Meagher. *Erdos-Ko-Rado Theorems: Algebraic Approaches.* Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2015.
9. C. Godsil. Notes on association schemes, June 2010.
10. A.S. Greenhalgh. Random walks on groups with subgroup invariance properties. Technical report, Stanford University, Department of Statistics, 04 1989.
11. P. Høyer, T. Lee, and R. Špalek. Negative weights make adversaries stronger. In *Proceedings of the Thirty-ninth Annual ACM Symposium on Theory of Computing*, STOC '07, pages 526–535, New York, NY, USA, 2007. ACM.
12. G.D. James and A. Kerber. *The Representation Theory of the Symmetric Group.* Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1984.
13. T. Lee, R. Mittal, B.W. Reichardt, R. Špalek, and M. Szegedy. Quantum query complexity of state conversion. In *Proceedings of the 2011 IEEE 52nd Annual Symposium on Foundations of Computer Science*, FOCS '11, pages 344–353, Washington, DC, USA, 2011. IEEE Computer Society.
14. G. Midrijānis. A polynomial quantum query lower bound for the set equality problem. In *Automata, Languages and Programming*, pages 996–1005, Berlin, Heidelberg, 2004. Springer Berlin Heidelberg.
15. A. Montanaro. Quantum algorithms: An overview. *npj Quantum Information*, 2, 11 2015.
16. A. Munemasa. The injection scheme of permutations (unpublished). 2001.
17. M.A. Nielsen and I.L. Chuang. *Quantum Computation and Quantum Information.* Cambridge Series on Information and the Natural Sciences. Cambridge University Press, 2000.
18. A. Rosmanis. Quantum adversary lower bound for element distinctness with small range. *Chicago Journal of Theoretical Computer Science*, 2014(4), July 2014.
19. A. Rosmanis and A. Belovs. On adversary lower bounds for the collision and the set equality problems, 2013. Available at arXiv:1310.5185v1 [quant-ph].
20. B. Sagan. *The Symmetric Group: Representations, Combinatorial Algorithms, and Symmetric Functions.* Graduate Texts in Mathematics. Springer New York, 2001.
21. Y. Shi. Quantum lower bounds for the collision and the element distinctness problems. In *Proceedings of the 43rd Symposium on Foundations of Computer Science*, FOCS '02, pages 513–519, Washington, DC, USA, 2002. IEEE Computer Society.
22. M. Zhandry. A note on the quantum collision and set equality problems. *Quantum Info. Comput.*, 15(7-8):557–567, May 2015.