Identification of a reversible quantum gate: assessing the resources

Giulio Chiribella

Center for Quantum Information, Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing 100084, China.

Giacomo Mauro D'Ariano

QUIT group, Dipartimento di Fisica "A. Volta", INFN Sezione di Pavia, via Bassi 6, 27100 Pavia, Italy.

Martin Roetteler

Microsoft Research, One Microsoft Way, Redmond, WA 98052, USA.

Abstract.   We assess the resources needed to identify a reversible quantum gate among a finite set of alternatives, including in our analysis both deterministic and probabilistic strategies. Among the probabilistic strategies, we consider unambiguous gate discrimination—where errors are not tolerated but inconclusive outcomes are allowed—and we prove that parallel strategies are sufficient to unambiguously identify the unknown gate with minimum number of queries. This result is used to provide upper and lower bounds on the query complexity and on the minimum ancilla dimension. In addition, we introduce the notion of generalized t-designs, which includes unitary t-designs and group representations as special cases. For gates forming a generalized t-design we give an explicit expression for the maximum probability of correct gate identification and we prove that there is no gap between the performances of deterministic strategies an those of probabilistic strategies. Hence, evaluating of the query complexity of perfect deterministic discrimination is reduced to the easier problem of evaluating the query complexity of unambiguous discrimination. Finally, we consider discrimination strategies where the use of ancillas is forbidden, providing upper bounds on the number of additional queries needed to make up for the lack of entanglement with the ancillas.

1. Introduction

Identifying an unknown unitary evolution, available as a black box, is a fundamental problem in quantum theory [1, 2, 3, 4, 5, 6, 7, 8, 9], with a wide range of applications in quantum information and computation. In quantum computation, the problem is known as oracle identification [10, 11, 12, 13, 14] and is the core of paradigmatic quantum algorithms such as Grover's [15] and Bernstein-Vazirani's [16]. In quantum information processing, the identification of an unknown unitary gate plays a key role in the stabilizer formalism of quantum error correction [17, 18] and in its generalization to unitary error bases [19, 20, 21, 22, 23], in the security analysis of quantum cryptographic protocols that encode secret data into unitary gates [24, 25, 26, 27, 28, 29], in the alignment of reference frames via quantum communication [30, 31, 32, 33, 34], in the design of quantum communication protocols that work in the absence of shared reference frames [36, 37, 38] and quantum machines that learn to execute a desired operation from a training set of examples [39]. For all these applications, the crucial step is to find efficient strategies that discriminate among a set of unknown gates with minimum expenditure of resources. Typical resources considered are: the number of black box queries needed to identify the unknown gate, the number of time steps and the size of the auxiliary systems (ancillas) employed in the discrimination strategy, and the total number of elementary gates needed to implement the discrimination strategy.

A striking feature of gate discrimination is that any two distinct unitaries can be perfectly distinguished from one another in a finite number of queries, either using entanglement [1, 2] or using a sequential strategy where different queries are called at different time steps [5]. This feature implies that an unknown gate in a finite set $\mathsf{U} := (U_x)_{x \in \mathsf{X}}$ can be perfectly identified in a finite number of queries, e.g. by running $|\mathsf{U}| - 1$ pairwise tests each of which eliminates one wrong alternative [1]. However, in terms of efficiency the method of pairwise elimination leaves large room for improvement: For example, when the unitaries are mutually orthogonal, one can identify the black box in a single query using an ancilla, following the lines of the dense coding protocol [40]. In general, finding the minimum number of queries needed for perfect discrimination is a hard problem: solving it would automatically give a general solution for the query complexity of oracle identification. One way to approach the problem is to consider the less demanding task of *unambiguous gate discrimination* [3, 4, 6, 7, 41], where the unknown gate is identified without errors but one allows for an inconclusive result. General conditions for unambiguous discrimination were given in Refs. [4, 7, 41] under the assumption that the available queries are used in parallel. However, the case of general strategies and the quantification of the resources required for unambiguous gate discrimination have remained largely unaddressed up to now.

In this paper we provide a systematic study of the resources needed to identify an unknown gate, focussing in particular on the following resources: number of queries, size of the ancillary systems, and number of time steps in the discrimination strategy. We start from the observation that parallel strategies are sufficient for unambiguous gate

discrimination: if unambiguous discrimination can be achieved in N queries, then it can be achieved by applying the N queries in parallel (in general, using ancillas). Based on the reductions to parallel strategies, we provide lower and upper bounds on the query complexity of unambiguous discrimination and on the size of the ancilla systems needed by the discrimination strategy. The bounds are general and can often be improved in specific cases. Nevertheless, they suffice to show that unambiguous discrimination of the gates U is always possible with no more than $|\mathsf{U}| - 1$ queries. Since $|\mathsf{U}| - 1$ is the minimum number of queries that would be needed by the method of pairwise elimination, our result shows that a joint discrimination strategy typically offers an advantage.

After having discussed the resources for unambiguous discrimination, we ask under which conditions one can achieve the stronger task of perfect discrimination, where inconclusive outcomes are not allowed. This is important because in practice the usefulness of unambiguous discrimination can be undermined by the fact that the inconclusive outcome occurs too frequently. To this purpose, we introduce the notion of *generalized t-designs*, which includes as special cases the unitary t-designs of Refs. [43, 44, 45, 46] and all the examples where the unknown gates form a group [47, 48, 49, 50]. Relative to gate identification, generalized t-designs have three important features:

 (i) there is no difference between the performances of deterministic and probabilistic strategies allowing for inconclusive outcomes

 (ii) there is no difference between the performances of strategies using the queries in parallel and general strategies using the queries in a sequence of time steps

(iii) there is a simple analytic formula for the maximum probability of correct gate identification with given number of queries.

The feature *i)* implies that, if unambiguous discrimination is possible in N queries, then also perfect deterministic discrimination must be possible in N queries. This result reduces the query complexity of perfect discrimination to the query complexity of unambiguous discrimination, which is much simpler to evaluate. The reduction to unambiguous discrimination has a fairly large range of applications, especially in the case when the set of gates forms a group. Particular examples are the group of all Boolean oracles [10], the groups of linear [16] and quadratic [42] Boolean functions, the group of permutations [30], and the group of all oracles corresponding to functions from a given finite set to another [7]. The feature *ii)* implies that the number of time steps needed to identify a gate picked from a generalized t-design is minimum: applying the queries in parallel one can reduce the discrimination strategy to three steps: the preparation of an entangled state, the parallel application of the unitary gates, and the execution of a suitable measurement. Note that the number of time steps in a discrimination strategy should not be confused with the number of elementary gates needed to implement the strategy: preparing the joint state and performing the joint measurement may require a large number of elementary gates. Nevertheless, the fact that in principle the number of time steps can be reduced to the minimum is an interesting and non-trivial property.

In general, such a property does not hold when the unitaries do not form a generalized t-design: for example, using the available queries in parallel would spoil the quadratic speed-up in Grover's algorithm [51, 52].

Finally, we address the quantification of resources for strategies where the use of ancillary systems is forbidden. This analysis is important for applications to reference frame alignment [30, 31, 32, 33, 34, 35] and quantum communication in the absence of shared reference frames [36, 37, 38]. Our contribution to these research topics is *a)* to show that every gate discrimination using ancillas can be converted into a strategy using a number of extra queries to the unknown gate and *b)* to provide bounds on the number of extra queries. When the dimension of the ancilla used in the original strategy is large, we show that the number of extra queries scales logarithmically with the ancilla dimension: a strategy using $N_A \gg 1$ ancillary qubits can be replaced by a strategy using $O(N_A)$ extra queries. More specific bounds ben be obtained when the unitaries form a generalized t-design or a group. In all these cases, we show that, again, there is no difference between the performances of deterministic strategies and those of probabilistic strategies allowing for inconclusive outcomes.

The paper is structured as follows. In section 2 we give a synopsis of the main results. The basic facts about general gate discrimination strategies, along with the ovservation that unambiguous discrimination can be parallelized, are provided in Section 3, and exploited in Section 4 to derive upper and lower bounds on the query complexity of unambiguous discrimination. In Section 5 we introduce the notion of generalized t-designs, giving an explicit formula for the maximum probability of correct gate identification and showing that parallel deterministic strategies achieve the same performances of arbitrary probabilistic strategies. Bounds on the size of the ancilla needed for unambiguous/perfect discrimination with minimum queries are provided in Section 6, while Section 7 considers discrimination schemes where the ancilla is not allowed, providing estimates for the query overhead. The conclusions are drawn in Section 8. The Appendix contains all the technical proofs of the results presented in the paper.

2. Main results

We provide here a synopsis of the main results of the paper. A more extended discussion, including the precise definition of the framework, additional results and applications will be the object of the following Sections.

Unambiguous gate discrimination: parallelizability and bounds on the query complexity. We start by showing that unambiguous gate discrimination can be parallelized: if the gates in a given set can be distinguished unambiguously with N queries, then they can be distinguished unambiguously by applying the queries in parallel, possibly using ancillas. This fact is extremely useful to provide bounds of the query complexity of unambiguous discrimination. Denoting by $N_{\min}$ the minimum number of queries needed

to unambiguously identify a unitary gate in the set U, we prove the bounds

$$|\mathsf{U}| \leq \binom{N_{\min} + d^2 - 1}{d^2 - 1} \tag{1}$$

and

$$N_{\min} \leq |\mathsf{U}| - \dim(\mathsf{U}) + 1, \tag{2}$$

where d is the dimension of the Hilbert space where the gates act and $\dim(\mathsf{U})$ is the number of linearly independent unitaries in U. Both bounds are tight, in the sense that for every size $|\mathsf{U}|$ one can find a set of gates achieving the equality. The upper bound of Eq. (2) proves that a joint discrimination strategy typically needs less queries than a strategy based on pairwise eliminations. The bound of Eq. (1) contains implicitly a lower bound on $N_{\min}$, which can be estimated as

$$N_{\min} > |\mathsf{U}|^{\frac{1}{d^2-1}} - 1 \tag{3}$$

When d is fixed and $|\mathsf{U}|$ is large, this estimate gives the actual scaling of the tight bound of Eq. (1).

In addition to the above bounds, we also provide a bound in term of the maximum fidelity between pairs of gates. The bound is obtained from a simple observation about unambiguous state discrimination of pure states, which to the best of our knowledge did not appear in the previous literature on the subject: the states in a generic set $\{|\psi_x\rangle\}_{x\in\mathsf{X}}$ can be unambiguously discriminated using N identical copies whenever N satisfies

$$N > \frac{\log(|\mathsf{X}| - 1)}{\log\left(F^{-\frac{1}{2}}\right)} \qquad F := \max_{x\neq y} |\langle\psi_x|\psi_y\rangle|^2. \tag{4}$$

In the case of gate discrimination, this result can be applied to the set of bipartite states $|\Psi_x\rangle := (U_x \otimes I)|\Psi\rangle$, where $|\Psi\rangle$ is a bipartite input state. Optimizing over $|\Psi\rangle$, we then get the *fidelity bound*

$$N_{\min} \leq \left\lfloor \frac{\log(|\mathsf{U}| - 1)}{\log\left(F_{\mathsf{U}}^{-1/2}\right)} \right\rfloor + 1, \tag{5}$$

where F_{U} is the *minimax fidelity* $F_{\mathsf{U}} := \min_{|\Psi\rangle\in\mathcal{H}\otimes\mathcal{H}, \|\Psi\|=1} \max_{x\neq y} \left|\langle\Psi|(U_x^\dagger U_y \otimes I)|\Psi\rangle\right|^2$. The fidelity bound is important because it connects a measure of pairwise distinguishability with the the performances of general joint strategies for unambiguous discrimination. Moreover, in several examples it gives a better estimate than the linear bound of Eq. (2).

Generalized t-designs: maximum error probability and optimality of parallel deterministic strategies. We introduce the notion of generalized t-designs, which includes as special cases the unitary t-designs of Refs. [43, 44, 45, 46] and all the examples where the unknown gates form a group [47, 48, 49, 50]. When the unitary gates form a generalized t-design, we consider the problem of gate discrimination with minimum error probability, or, equivalently, with maximum probability of correct gate identification.

Optimizing over all possible discrimination strategies, we show that the that maximum probability of correct identification with $N \leq t$ queries (conditional to the occurrence of conclusive outcomes) is given by

$$p_N^{\mathrm{max}} = \frac{\dim \mathsf{U}_N}{|\mathsf{U}|} \qquad \mathsf{U}_N := \left(U_x^{\otimes N} \right)_{x \in \mathsf{X}}.$$ (6)

Moreover, we show that this optimum value can be achieved by a deterministic strategy that uses the N queries in parallel. As a consequence, this shows that *i)* there is no difference between the performances of deterministic and probabilistic discrimination strategies, and *ii)* there is no difference between the performances of parallel strategies and those of strategies using a sequence of multiple time steps. In particular, if a set of gates U is a generalized $|\mathsf{U}|$-design, then there is no difference between perfect and unambiguous discrimination: whenever unambiguous discrimination is possible, the probability of the inconclusive result can be reduced to zero. This result is important from the practical point of view, because unambiguous discrimination by itself may not be a useful primitive if the probability of the inconclusive result is too high. Moreover, thanks to the reduction to deterministic strategies, Eqs. (1), (2) and (5) become bounds on the query complexity of perfect deterministic gate discrimination.

Minimum ancilla dimension. Another important resource, in addition to the number of queries, is the dimension of the ancilla needed to achieve gate discrimination [48, 53]. The ancilla dimension quantifies the extra memory space used for the discrimination task. We show that, when N queries to the black boxes are used, the minimum ancilla dimension can be upper bounded as

$$d_{A,N}^{\mathrm{min}} \leq \begin{pmatrix} N + d - 1 \\ d - 1 \end{pmatrix}.$$ (7)

Since the binomial can be upper bounded as $(N+1)^{d-1}$, our result implies that the size of the ancilla scales at most polynomially in the number of queries. In other words, this means that the number of ancillary qubits needed for gate discrimination in N queries is at most logarithmic in N. The bound is independent of the gate set U. When more information on the gates is available, further estimates can be provided: For example, if the set U is contained in a representation of a finite group G, the dimension of the ancilla can be upper bounded as

$$d_{A,N}^{\mathrm{min}} \leq \sqrt{|\mathsf{G}|},$$ (8)

independently of the number of queries and of the dimension of the Hilbert space. This bound provides a fast estimate of the ancilla dimension, and, in several situations, the estimate is actually accurate. For example, in the case of the Pauli matrices $\{I, X, Y, Z\}$ the bound gives correctly $d_{A,N}^{\mathrm{min}} \leq 2$, meaning that unambiguous discrimination is possible using a single ancilla qubit. This is indeed what is achieved by the dense coding protocol [40]. An even stronger result holds if the unitaries in U commute: in this case, no ancilla at all is needed, a result that was already known for discrimination strategies using parallel queries [48, 7].

Ancilla-free gate discrimination. For applications in reference frame alignment [30, 31, 32, 33, 34, 35] and quantum communication in the absence of shared reference frames [36, 37, 38], it is useful to consider discrimination strategies that do not use ancillas, here referred to as *ancilla-free.* In this case, one can make up for the lack of ancillas using a number of additional queries to the black box, using the invariant encoding of Ref. [54]. For a strategy using a large number of ancilla qubits N_A, we show that the scaling of the minimum overhead $\Delta N_{\min}$ is upper bounded as

$$\Delta N_{\min} \leq O(N_A). \tag{9}$$

In other words, the N_A ancilla qubits can be replaced with (order of) N_A extra queries to the black box, showing that the use of extra queries is a more powerful resource than the use of ancillas. In addition, we provide the ancilla-free version of the upper bounds of Eqs. (2) and (5), showing that, even if the use of ancillas is prohibited, a joint discrimination strategy will still outperform the method of pairwise elimination.

The conditions for unambiguous discrimination are sufficient to guarantee perfect deterministic discrimination when the set U is a generalized t-design. Indeed, also in the ancilla-free case we show that for generalized t-designs there is no difference between probabilistic and deterministic discrimination strategies. Finally, when the gates in U form a representation of a group $\mathbf{G}$, one can prove more specific results [55]:

(i) A perfect discrimination strategy using d_A-dimensional ancilla can be replaced by a perfect ancilla-free discrimination strategy using

$$\Delta N_{\min} \leq \left\lceil \frac{\log d_A + \log \sqrt{|\mathbf{G}|}}{\log d} \right\rceil \tag{10}$$

extra queries. This result is consistent with the scaling with the number of qubits promised by Eq. (9): for a strategy using a large number of ancillary qubits $N_A \gg \log \sqrt{|\mathbf{G}|}$, the number of extra queries to the black box scales as $O(N_A)$.

(ii) The query complexity of ancilla-free discrimination can be upper bounded with an expression involving the maximum entanglement fidelity between pairs of different gates (cf. Subsection 7.2 for the actual expression). This is quite surprising because the operational interpretation of the entanglement fidelity is the fidelity between the output states obtained by applying the unitaries on one side of the maximally entangled state, a strategy that is forbidden in ancilla-free gate discrimination. When the entanglement fidelity is zero, we obtain that the minimum number of queries needed for ancilla-free discrimination is given by

$$N_{\min}^{AF} = \lceil \log_d |\mathbf{G}| \rceil. \tag{11}$$

In principle, this is the most favourable scaling possible: indeed, with less than $\lceil \log_d |\mathbf{G}| \rceil$ queries it would be impossible to pack $|\mathbf{G}|$ orthogonal vectors in the joint Hilbert space of the systems used by the discrimination strategy.

Eqs. (10) and (11) allow one to quantify the resources needed for protocols of quantum communication [36] and decoherence-free encoding [38] in the absence of shared

reference frames. Indeed, in these cases $N_{\min}$ is equal to the number of physical systems needed to construct the "token state" used in the communication protocol [36, 38]. In the case of the alignment protocols [30, 31, 32, 33, 34, 35], $N_{\min}^{AF}$ is the minimum number of quantum systems that have to be exchanged in order to establish a reference frame.

3. Gate discrimination: framework and basic definitions

We consider the problem of identifying an unknown unitary gate under the promise that the gate belongs to a finite set U. For simplicity, we assume that the gates act on a system with Hilbert space $\mathcal{H}$ of finite dimension $d < \infty$. Moreover, all throughout the paper we assume that every two unitaries $U_x, U_y \in \mathsf{U}$ are statistically distinguishable, that is, there exists at least one input state, with density matrix ρ, such such that

$$U_x \rho U_x^\dagger \neq U_y \rho U_y^\dagger. \tag{12}$$

If this were not the case, there would be no point in making an experiment to distinguish between U_x and U_y, because these two gates would give rise to the same outcome probabilities for every possible experiment, and, therefore, there would be no operational way to tell them apart.

3.1. Discrimination strategies

In order to identify the action of the unknown gate, one is allowed to make queries to the corresponding black box and to use them in an arbitrary quantum circuit. As long as there is no constraint on the use of ancillas, one can focus without loss of generality on circuits consisting of pure states and unitary gates, of the form

 (13)
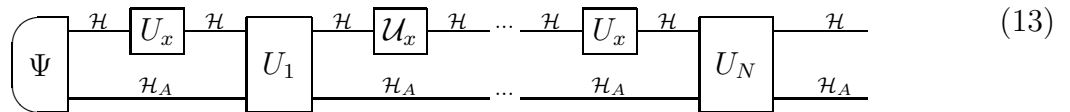
where

(i) $|\Psi\rangle \in \mathcal{H} \otimes \mathcal{H}_A$ is the joint state state of the input of U_x and of an ancilla with Hilbert space $\mathcal{H}_A$

(ii) U_t is a unitary gate representing a joint evolution of the system and the ancilla at the time step $t \in \{1, \ldots, N\}$ (the unitary U_N is added just for convenience of notation).

Once the input state $|\Psi\rangle$ and the unitaries $(U_t)_{t=1}^N$ have been chosen, identifying the unitary U_x is equivalent to identifying the output state

$$|\Psi_x\rangle := \left[\prod_{n=1}^N U_n (U \otimes I_A) \right] |\Psi\rangle. \tag{14}$$

To this purpose, one has to perform a suitable quantum measurement, described by a joint POVM $(P_y)_{y \in \mathsf{Y}}$. Here we allow for measurements with a set of outcomes $\mathsf{Y} = \mathsf{X} \cup \{?\}$, including an inconclusive outcome $y = ?$, which corresponds to the case when

the experimenter abstains from producing a guess [3]. Among all possible strategies, the deterministic ones are those for which the inconclusive outcome never occurs, namely $P_? = 0$.

3.2. Optimal, error-free, unambiguous, and perfect discrimination

As a figure of merit for gate discrimination, we choose the probability that that the unknown gate is identified correctly, provided that the measurement does not output the inconclusive outcome $y = ?$. This probability is given by

$$p_N := \frac{\sum_{x \in \mathsf{X}} p_N(x|x) \, p_x}{\sum_{x,y \in \mathsf{X}} p_N(y|x) \, p_x}. \tag{15}$$

where p_x is the prior probability of U_x and $p_N(y|x) = \langle \Psi_x | P_y | \Psi_x \rangle$ is the conditional probability of the measurement outcome y given that the gate is U_x and that N queries are used.

The optimal discrimination strategy is the one that maximizes the success probability p_N. We denote the corresponding probability by $p_N^{\max}$. Note that, in general, a deterministic strategy may not be able to reach the value $p_N^{\max}$: in order to achieve the optimal performances one may be forced to have an inconclusive outcome. For this reason, the probability $p_N^{\max}$ is an upper bound on the maximum probability of success over deterministic strategies, which is the quantity normally considered in minimum-error discrimination.

In this paper we will be particularly interested in discrimination strategies that are *error-free*, in the sense that they never misidentify the gate ($p_N = 1$). Note that the error-free condition $p_N = 1$ is much weaker than it may seem at first sight: this can be seen in the example of the qubit gates

$$U_0 = (I + Z)/\sqrt{2} \qquad U_k = \cos(2\pi/K)I + i \sin(2\pi/K)X \quad k = 1, \dots, K,$$

where K is an arbitrary integer number. In this case, one can achieve success probability $p_N = 1$ by applying the unknown unitary on one side of the maximally entangled state $|\Phi^+\rangle = (|0\rangle|0\rangle + |1\rangle|1\rangle)/\sqrt{2}$ and by measuring the output state with the POVM given by

$$P_0 = (Z \otimes I)|\Phi^+\rangle\langle\Phi^+|(Z^\dagger \otimes I) \qquad P_k = 0 \quad \forall k = 1, \dots, K \qquad P_? = I - P_0.$$

Clearly, *when the inconclusive result does not occur*, the unknown gate has been identified with certainty: indeed, the outcome 0 can only occur when the gate is U_0.

In some situations, having a discrimination strategy that detects only one gate and aborts otherwise may not be useful. Instead, one may require that every gate in the set U have a non-zero probability of being identified. We say that a discrimination strategy achieves *unambiguous discrimination* if it is error-free ($p_N = 1$) and, in addition, $p(x|x) > 0$ for every x. For example, the two unitaries $U_0 = I$ and $U_1 = \exp[i\theta Z]$ can

be distinguished unambiguously by preparing the input state $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ and by measuring the POVM defined by

$$P_0 = \frac{U_1|-\rangle\langle-|U_1^\dagger}{1 + \cos(\theta/2)} \qquad P_1 = \frac{U_0|-\rangle\langle-|U_0^\dagger}{1 + \cos(\theta/2)} \qquad P_? = I - P_0 - P_1,$$

with $|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$. This strategy is error-free $p_N = 1$ and both gates have the chance of being detected: in this particular case, one has $p(0|0) = p(1|1) = [\sin(\theta/2)]^2/[1 + \cos(\theta/2)]$.

Note, that the definition of unambiguous discrimination does not include any requirement on the probability of the inconclusive outcome, which in principle can be arbitrarily close to 1. In some situations, this feature can undermine the usefulness of the discrimination scheme. On the opposite end, one can restrict the attention to discrimination strategies such that the probability of the inconclusive outcome is equal to 0. We refer to these strategies as *perfect discrimination strategies*.

3.3. Basic facts about error-free and unambiguous discrimination

Error-free and unambiguous discrimination can be nicely characterized in terms of linear independence:

Theorem 1 *The unitaries in* U *can be discriminated in* N *queries*

(i) *in an error-free way if and only if there exists a unitary* U_{x_0} *that is not a linear combination of the other unitaries in* U

(ii) *in an unambiguous way if and only if the unitaries* $(U_x^{\otimes N})_{x \in \mathsf{X}}$ *are linearly independent.*

The equivalence between unambiguous gate discrimination and linear independence of the unitaries was previously observed in Ref. [7] in the case of *parallel strategies*, i.e. strategies where the N queries are applied in parallel to a suitable multipartite state $|\Psi\rangle \in \mathcal{H}^{\otimes N} \otimes \mathcal{H}_A$, thus producing the output state $|\Psi_x\rangle = \left(U_x^{\otimes N} \otimes I_A\right)|\Psi\rangle$, as in figure

$$(16)$$
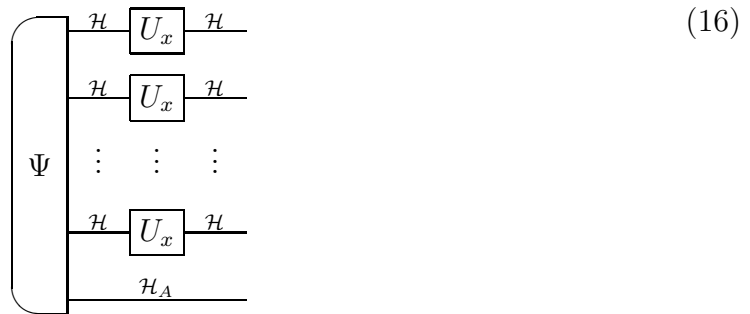
Parallel strategies are a special case of the strategies of Eq. (13), where one has the freedom to apply the N queries at different time steps in a quantum circuit. The parallel strategies of Eq. (17) can be recovered as a special case from the general strategies of

Eq. (13) by setting the gates $(U_t)_{t=1}^N$ to be suitable swap gates, as in the following example

$$\tag{17}$$
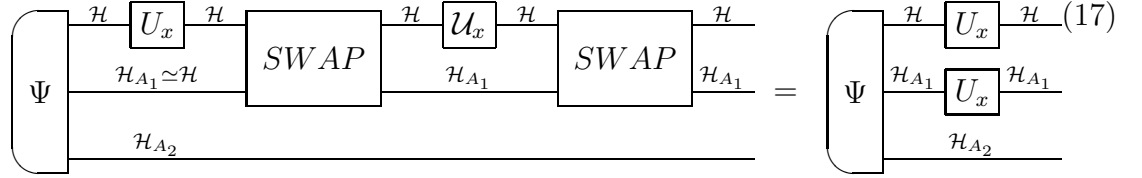
Theorem 1 has an important consequence: it implies that error-free discrimination and unambiguous discrimination discrimination can be parallelized:

Corollary 1 (Parallelization of error-free and unambiguous discrimination) *If the gates* U *can be distinguished unambiguously (respectively, in a error-free fashion) with N queries, then they can be distinguished unambiguously (respectively, in a error-free fashion) using the N queries in parallel.*

In other words, the identification of the gate can be achieved in the shortest possible number of time-steps: in the problem of error-free and unambiguous discrimination the time resource can be completely replaced by spatial resources. This fact is also useful as a technical tool: it implies that the query complexity of error-free/unambiguous discrimination—defined as minimum number N_{min} needed to unambiguously identify a gate in U—does not change if one restricts to parallel strategies. Note, however, that general sequential strategies can help in reducing the probability of the inconclusive result. This fact is well illustrated by the example of Grover's algorithm:

Example 1 (Discrimination of Grover's oracles) *Grover's algorithm is designed to identify a unitary gate in the set* U *containing the gates* $U_x = I - 2|x\rangle\langle x|$, $x \in \mathsf{X} = \{1,\ldots,d\}$. *Clearly, the gates* U *are linearly independent for every $d > 2$ and therefore they can be unambiguously discriminated in a single query, as originally observed by Chefles* et al *in Ref. [7]. However, the probability of unambiguous discrimination in a single query must be necessarily low. One way to see it is the following: as showed by Brassard, Hoyer, Boyer, and Tapp [51] and Zalka [52], Grover's algorithm cannot be efficiently parallelized: there is no deterministic parallel strategy that can achieve in N queries the same probability of correct gate identification as in Grover's algorithm. Now, if the probability of unambiguous discrimination were sufficiently large, one could run different rounds of unambiguous discrimination, and use this fact to construct an efficient parallel strategy. In the case of search in a large database (d ≫ 1), denoting by N_G is the number of queries needed by Grover's algorithm to achieve maximum probability of correct gate identification, one can show that the probability of unambiguous discrimination in one query must be upper bounded by $O(\log N_G/N_G)$.*

4. General bounds on the query complexity of unambiguous gate discrimination

The possibility of parallelizing unambiguous gate discrimination, established by theorem 1, leads immediately to general bounds on the query complexity. These bounds do not

assume any structure of the set of unitaries U and can typically be improved when more information about U is available.

4.1. Lower bound

Here we give a lower bound to the number of queries that are necessary to for unambiguous gate identification.

Proposition 1 (Dimensional bound) *If the the gates in U can be unambiguously discriminated using N queries, then*

$$|\mathsf{U}| \le \left(\begin{array}{c} N + d^2 - 1 \\ d^2 - 1 \end{array} \right). \tag{18}$$

If we do not impose any structure on the set of unitaries U, then the bound of Eq. (18) is the best we can hope for. Indeed, for any fixed Hilbert space dimension d and for every size $|\mathsf{U}|$ we can always find a set of unitaries U such that the minimum number of queries needed to unambiguously identify a gate in U is exactly the minimum N compatible with Eq. (18) (cf. the proof in the Appendix).

Eq. (18) can be used to provide an easy lower bound on the query complexity: Combining it with the inequality

$$\left(\begin{array}{c} N + k \\ k \end{array} \right) < (N + 1)^k, \tag{19}$$

we obtain the bound $N > |\mathsf{U}|^{\frac{1}{d^2-1}} - 1$, which is a necessary condition for unambiguous discrimination with N queries. The bound is not tight, but provides the right scaling with $|\mathsf{U}|$ in the regime when d is fixed and N is large compared to d^2. Indeed, in this case one has

$$\left(\begin{array}{c} N + d^2 - 1 \\ d^2 - 1 \end{array} \right) = \frac{N^{d^2-1}}{(d^2 - 1)!} + O(d^2/N),$$

which means that the scaling of the tight bound associated to Eq. (18) is actually $N = \Omega\left(|\mathsf{U}|^{\frac{1}{d^2-1}} \right)$.

4.2. Upper bounds

An upper bound on the query complexity can be obtained by observing that the number of linearly independent unitaries in U_N grows at least linearly with N, a fact that can be proved using an earlier result by Chefles [57]:

Proposition 2 (Linear bound) *The query complexity of unambiguous discrimination of the gates in U is upper bounded by*

$$N_{\min} \le |\mathsf{U}| + 1 - \dim(\mathsf{U}). \tag{20}$$

In general, the bound of Eq. (20) can be achieved: for every fixed Hilbert space dimension d and for every fixed cardinality $|\mathsf{U}|$ we can find a set of unitaries such that $N_{\min} = |\mathsf{U}| - \dim(\mathsf{U}) + 1$. This can be seen in the following

Example 2 (Unambiguous discrimination of discrete phase-shifts) *Consider the problem of identifying an unknown phase shift*

$$U_x := \omega^x \, |1\rangle\langle 1| + (I - |1\rangle\langle 1|) \qquad \omega := e^{\frac{2\pi i}{|\mathsf{X}|}},$$

with $x = 1, \ldots, |\mathsf{X}|$. In this case the number of linearly independent unitaries in $(U_x^{\otimes N})_{x \in \mathsf{X}}$ is exactly equal to $N+1$, as it can be seen from the fact that the unitaries $(U_x^{\otimes N})_{x \in \mathsf{X}}$ are in one-to-one correspondence with the vectors of their eigenvalues, given by $(v_x)_{x \in \mathsf{X}} \subset \mathbb{C}^{N+1}$ where $v_x := (1, \omega, \omega^2, \ldots, \omega^N)^T$. Since the number of linearly independent unitaries in $(U_x^{\otimes N})_{x \in \mathsf{X}}$ is $N+1$, the minimum number needed for unambiguous discrimination is exactly $N_{\min} = |\mathsf{X}| - 1 = |\mathsf{U}| - \dim(\mathsf{U}) + 1$.

Another example where the bound of Eq. (20) gives the exact value is the identification of a "shift-and-multiply" gate:

Example 3 (Unambiguous discrimination of shift-and-multiply gates) *Consider the problem of identifying a shift-and-multiply gate*

$$U_{pq} = S^p M^q \qquad (p, q) \in \mathbb{Z}_d \times \mathbb{Z}_d \ , \tag{21}$$

where $S = \sum_{k=1}^d |(k+1) \bmod d\rangle\langle k|$ and $M = \sum_{k=1}^d e^{(2\pi i k)/d} |k\rangle\langle k|$. In this case, the unitaries $(U_{pq})_{(p,q) \in \mathbb{Z}_d \times \mathbb{Z}_d}$ are linearly independent, and therefore the bound gives $N_{\min} = 1$. Note that, in fact, the unitaries are orthogonal in the Hilbert-Schmidt product, and, therefore, an unknown unitary U_{pq} can be identified perfectly and deterministically, as in the dense coding protocol [40].

Proposition 2 provides an estimate of $N_{\min}$ that is always better than the number of pairwise tests $|\mathsf{U}| - 1$ that would be needed to identify a gate in U with the method of pairwise eliminations outlined in [1, 2]. Note however that Eq. (20) only ensures *unambiguous* discrimination, while the pairwise elimination method ensures *perfect* discrimination. In the next Section we will see that the distinction between unambiguous and perfect discrimination disappears when the gates in U form a group representation, or, more generally, a generalized t-design.

Before adding more structure on the set U, we give here a second upper bound that often yields a better estimate than Proposition 2. To state the bound we introduce the *minimax fidelity* of the unitaries in U, defined as

$$F_{\mathsf{U}} := \min_{|\Psi\rangle \in \mathcal{H} \otimes \mathcal{H}, \|\Psi\|=1} \ \max_{x,y \in \mathsf{X}, x \neq y} \left| \langle \Psi | (U_x^\dagger U_y \otimes I) | \Psi \rangle \right|.$$

The minimax fidelity quantifies the pairwise distinguishability of the gates in U when single-shot ancilla-assisted strategies are used. Clearly, if $F_{\mathsf{U}} = 0$, the unitaries can be perfectly distinguished in one shot using a suitable input state. Note also that, under the standing assumption of this paper, F_{U} must be strictly smaller than 1: indeed, the

distinguishability condition of Eq. (12) implies that for every two distinct unitaries U_x and U_y there exists at least an input state $|\psi\rangle$ such that $\left|\langle\psi|U_x^\dagger U_y|\psi\rangle\right|^2 < 1$. In terms of the minimax fidelity, we have the following

Proposition 3 (Fidelity bound) *The query complexity of unambiguous discrimination of the gates in* U *is upper bounded as*

$$N_{\min} \leq \left\lfloor \frac{\log(|\mathsf{U}| - 1)}{\log\left(F_{\mathsf{U}}^{-\frac{1}{2}}\right)} \right\rfloor + 1. \tag{22}$$

The proof of the bound is based on a simple observation about unambiguous state discrimination, which is interesting *per se*:

Lemma 1 *Let* $(|\psi_x\rangle)_{x\in\mathsf{X}} \in \mathcal{H}$ *be a set of pure states and let* $F := \max_{x \neq y} |\langle\psi_x|\psi_y\rangle|^2$ *be the maximum fidelity between two distinct states in the set. If* $F^{N/2} < 1/(|\mathsf{X}| - 1)$*, then the states* $(|\psi_x\rangle^{\otimes N})_{x\in\mathsf{X}}$ *are linearly independent, and, therefore, unambiguously distinguishable.*

For qubits, this simple observation gives exactly the minimum number of copies needed for unambiguous discrimination of the states in a *symmetric informationally-complete (SIC) POVM* [58, 59]. Recall that a SIC-POVM in dimension d is a set of d^2 unit vectors with the property that the overlap between any two distinct vectors is the same:

$$|\langle\psi_x|\psi_y\rangle|^2 = \frac{1}{d+1} \qquad \forall x \neq y.$$

In general, for d^2 pure states one can easily see from dimensional arguments that unambiguous discrimination requires at least 3 copies (cf. the lower bound in Ref. [57]). On the other hand, Lemma 1 shows that for qubits $N = 3$ copies are sufficient, thus implying that $N = 3$ is actually the minimum number of copies needed for unambiguous discrimination of a SIC-POVM. For general d-dimensional systems, Lemma 1 gives the upper bound $N \leq 4$, almost matching the dimensional lower bound [60], and thus showing that the number of copies does not scale up with the dimension of the system. The exact value of the minimum number of copies is equal to $N_{\min} = 3$ for all the known examples of SIC-POVMs, except for the one example of SIC POVM in dimension $d = 3$, which actually requires $N_{\min} = 4$ copies [60].

Let us now comment on the tightness of the fidelity bound for gate discrimination. The bound gives good estimates when F_{U} is close to zero (in particular, in the extreme case where $F_{\mathsf{U}} = 0$ it predicts correctly $N_{\min} = 1$). However, it tends to overestimate $N_{\min}$ when F_{U} approaches 1. To understand this fact, note that for $F \geq 1 - \epsilon$, the estimate of $N_{\min}$ becomes

$$N_{\min} \leq \frac{2\ln(|\mathsf{U}| - 1)}{\epsilon},$$

having chosen the logarithm in base e. Now, two unitaries can have fidelity arbitrarily close to 1 and still be linearly independent. For example, the unitaries $U_0 = I$ and

$U_1 = e^{i\theta Z}$ are linearly independent, and their fidelity is $F = \cos\theta$. This implies that for every set U containing these two unitaries one has $F_{\mathsf{U}} \geq \cos\theta$ and, therefore, the fidelity bound gives $N_{\min} \leq \frac{4\ln(|\mathsf{U}|-1)}{\theta^2}$, where the r.h.s. can be arbitrarily large when the angle θ is small. Hence, for $F_{\mathsf{U}} \approx 1$ the fidelity bound can be arbitrarily far from the correct value (think of the case when the unitaries U are linearly independent, and the correct value is $N_{\min} = 1$). Another example of the gap between the value of the upper bound and true value of $N_{\min}$ for $F_{\mathsf{U}} \approx 1$ is illustrated in the following

Example 4 (Permutation gates) *Consider identification of an unknown permutation gate*

$$U_\pi = \sum_{k=1}^{d} |\pi(k)\rangle\langle k|, \tag{23}$$

where π is an element of the permutation group S_d. In this case it is clear that the unitary U_π can be perfectly identified with d queries (applying U_π to all the d vectors in the computational basis we can surely identify the permutation $\pi \in S_d$). One the other hand, applying the unitary U_π on a maximally entangled state gives the bound $F_{\mathsf{U}} \geq \left(\frac{d-2}{d}\right)^2$, which inserted in the fidelity bound gives $N_{\min} \leq \log(d!)/\log[d/(d-2)] = O(d^2 \log d)$, which is off by a factor $d \log d$ from the actual value.

5. Discrimination of generalized t-designs

Here we impose additional structure on the set of gates U. Our analysis includes the case where the set X labelling the gates in U is a finite group and $x \mapsto U_x$ is a representation of X. Also, it includes the case where the unitaries U from a unitary t-design [43, 44, 45, 46]. In order to treat these two cases in a unified way, we introduce the notion of *generalized t-designs*. For the discrimination of generalized t-designs we will show the following properties

(i) among all possible discrimination strategies using $N \leq t$ queries, the deterministic strategies using all queries in parallel maximize the probability of correct gate identification

(ii) for strategies using $N \leq t$ queries, there is no difference between error-free, unambiguous, and perfect discrimination.

(iii) the maximum probability of correct gate identification with $N \leq t$ queries has a simple analytic formula.

5.1. Generalized t-designs: definition and characterization

The notion of unitary t-design plays an important role in quantum information, with applications essentially in all protocols that require to extract a random gate from the uniform distribution over all possible unitaries [43, 44, 45, 46]. Unitary t-designs are

defined as follows: Consider an ensemble $(U_x, p_x)_{x \in \mathsf{X}}$ where U_x is a unitary and p_x is a probability. The ensemble is a unitary t-design iff

$$\sum_{x \in \mathsf{X}} p_x U_x^{\otimes t} \otimes \overline{U}_x^{\otimes t} = \int \mathrm{U} \; U^{\otimes t} \otimes \overline{U}^{\otimes t},$$

where $\overline{U}$ denotes the complex conjugate of the matrix U and the integral in the l.h.s. runs over the normalized Haar measure on the unitary group $U(d)$.

We will now generalize the above definition considering, instead of the full unitary group, a smaller group of unitary gates:

Definition 1 (Generalized t-designs) *Let $U : \mathbf{G} \mapsto U_g$ be a representation of a group* $\mathbf{G}$ *and let* X *be a subset of* $\mathbf{G}$. *An ensemble* $(U_x, p_x)_{x \in \mathsf{X}}$ *is a* generalized t-design *iff*

$$\sum_{x \in \mathsf{X}} p_x \; U_x^{\otimes t} \otimes \overline{U}_x^{\otimes t} = \int g \; U_g^{\otimes t} \otimes \overline{U}_g^{\otimes t}, \tag{24}$$

where $\int g f(g)$ denotes the integral of f with respect to the normalized Haar measure.

Note that, by definition, every generalized t-design is also a generalized $(t-1)$-design. Of course, a special case of generalized t-design is obtained by taking $\mathbf{G}$ to be a finite group and $\mathsf{X} = \mathbf{G}$:

Example 5 *The ensemble consisting of unitary gates in a representation of a finite group* $\mathbf{G}$, *randomly sampled with uniform probability distribution $p_g = 1/|\mathbf{G}|$, is a generalized t-design for every t.*

The definition of t-designs is extremely convenient, because it allows to easily transfer properties of groups to finite sets of quantum gates. In the next sections we will use this trick to prove strong properties of gate discrimination in the case of generalized t-designs.

5.2. Optimal discrimination of generalized t-designs

We start from general result about optimal probabilistic gate discrimination. Precisely, we show that the maximum probability $p_N^{\max}$ of correct gate identification can be always achieved with a parallel strategy. In addition, we give an analytic expression for $p_N^{\max}$.

Theorem 2 (Optimal probabilistic gate discrimination) *For every choice of prior probabilities, the maximum success probability $p_N^{\max}$ [cf.Eq. (15)] is achieved by applying the N queries in parallel on an entangled state. The probability $p_N^{\max}$ is given by*

$$p_N^{\max} = \max_{x \in \mathsf{X}} \; p_x \langle\!\langle U_x |^{\otimes N} \; R_N^{-1} \; |U_x \rangle\!\rangle^{\otimes N}, \tag{25}$$

with $|U_x\rangle\!\rangle := (U_x \otimes I)|I\rangle\!\rangle$, $|I\rangle\!\rangle := \sum_{n=1}^d |n\rangle|n\rangle$, $R_N := \sum_{x \in \mathsf{X}} p_x \; (|U_x\rangle\!\rangle\langle\!\langle U_x|)^{\otimes N}$, and R_N^{-1} being the inverse of R_N on its support.

The explicit formula of Eq. (25) is useful even if one is interested in deterministic strategies, rather than probabilistic ones. Indeed, by definition $p_N^{\max}$ provides an upper bound to the probability of success of arbitrary deterministic strategies. In some cases, one may even be able to achieve the upper bound with a deterministic strategy. This is actually the case for generalized t-designs:

Theorem 3 (Optimal gate discrimination for generalized t-designs) *Let $(U_x, p_x)_{x \in \mathsf{X}}$ be a generalized t-design. Then, the maximum of the probability of correct discrimination over all probabilistic strategies consisting of $N \leq t$ queries is*

$$p_N^{\max} = \dim(\mathsf{U}_N) \ \max_{x \in \mathsf{X}} p_x \tag{26}$$

For uniform prior $p_x = 1/|\mathsf{U}|$, the maximum probability $p_N^{\max} = \dim(\mathsf{U}_N)/|\mathsf{U}|$ can be achieved by a deterministic strategy that uses the N queries in parallel.

The general result of theorem 3 is well illustrated by the case of discrete phase shifts:

Example 6 (Discrete phase shifts) *Consider problem of identifying a discrete phase-shift gate*

$$U_x = \sum_{y=0}^{d-1} \omega^{xy} \ |y\rangle\langle y| \qquad \omega = e^{\frac{2\pi i}{|\mathsf{X}|}}, \tag{27}$$

U_x chosen at random with uniform probability $p_x = 1/|\mathsf{U}|$. Here $U : x \mapsto U_x$ is a unitary representation of the Abelian group $\mathsf{X} = \mathbb{Z}^K$, and, therefore, it is a generalized t-design for every t. Now, the number of linearly independent unitaries in U is d. Hence, the probability of correct identification of a unitary with a single query is $p_1^{\max} = d/|\mathsf{U}|$. Similarly, the number of linearly independent unitaries in U_N is $\min\{N(d-1)+1, |\mathsf{U}|\}$, and therefore, Eq. (26) gives

$$p_N^{\max} = \frac{Nd - N + 1}{|\mathsf{U}|} \qquad N \leq \frac{|\mathsf{U}| - 1}{d - 1}. \tag{28}$$

5.3. Perfect discrimination of generalized t-designs

From now on, we restrict our attention to generalized t-designs with *uniform* probability distribution $p_x = 1/|\mathsf{X}|$. Since there is no ambiguity, we will just refer to them as "generalized t-designs". An immediate consequence of Theorem 3 is that for generalized t-designs there is no difference between error-free, unambiguous, and perfect gate discrimination:

Corollary 2 *If the unitaries $(U_x)_{x \in \mathsf{X}}$ form a generalized t-design, then the following are equivalent:*

(i) error-free discrimination is possible with $N \leq t$ queries

(ii) unambiguous discrimination is possible with $N \leq t$ queries

(iii) perfect discrimination is possible in $N \leq t$ queries.

In particular, for a generalized |U|-design there is no difference between error-free, unambiguous, and perfect discrimination.

For generalized *t*-designs the evaluation of the query complexity of perfect discrimination is reduced to the simpler problem of evaluating the query complexity of unambiguous discrimination. In particular, the bounds in Propositions 1, 2, and 3 become automatically bounds on the query complexity of perfect discrimination.

6. Bounding the dimension of the ancilla

In addition to the query complexity, it is useful to bound the size of the ancilla needed for gate discrimination [48, 53]. Here we show that the size of the optimal ancilla scales at most polynomially with the number of queries N. We prove this result as a particular case of a more powerful statement about discrimination of unitaries picked from a group representation.

Proposition 4 *Let $U : g \mapsto U_g$ be a representation of a group $\mathbf{G}$ and let U be a subset of $(U_g)_{g \in \mathbf{G}}$ Then, the minimum dimension of the ancilla needed for unambiguous discrimination of the gates U in N queries is upper bounded by*

$$d_{A,N}^{\min} \leq \max_{\mu \in \mathrm{Irr}(U^{\otimes N})} \left\lceil \frac{d_\mu}{m_\mu} \right\rceil, \tag{29}$$

where the maximum runs over the set $\mathrm{Irr}\left(U^{\otimes N}\right)$ of irreducible representations contained in the decomposition of $U^{\otimes N}$, d_μ and m_μ are the dimension and the multiplicity of the irreducible representation μ (see the Appendix for some background information on representation theory).

Proposition 4 has many useful consequences. The first is that ancillas are not needed for the unambiguous discrimination of commuting unitaries, a fact that was noted in Ref. [7] for parallel discrimination strategies.

Corollary 3 *If the unitaries in U commute, then no ancilla is needed for unambiguous discrimination.*

The proof is immediate: a commuting set of unitaries is a subset of the Abelian group of all unitaries diagonal in a given basis. Since in this case the irreducible subspaces are one-dimensional, Eq. (29) gives $d_{A,N}^{\min} \leq 1$, which means that no ancilla is required.

As anticipated, another consequence of Proposition 4 is the fact that, no matter which set of gates U we are considering, the dimension of the ancilla needed for discrimination in N queries cannot grow faster than a polynomial in N.

Corollary 4 *The minimal dimension of the ancilla needed for unambiguous discrimination with N queries is upper bounded by $d_{A,N}^{\min} \leq \begin{pmatrix} N + d - 1 \\ d - 1 \end{pmatrix}$.*

The proof is provided in the Appendix. Bounding the binomial with Eq. (19), we have that the dimension of the ancilla is upper bounded by $(N+1)^{d-1}$. In other words, the number of ancillary qubits needed for unambiguous discrimination scales at most as the logarithm of the number of queries.

Another application of Proposition 4 can be found when the group $\mathbf{G}$ is finite. In this case, one can provide an upper bound on the size of the ancilla that is independent of the number of queries:

Corollary 5 *The minimum dimension of the ancilla needed for unambiguous discrimination of the gates* $\mathsf{U} \subseteq (U_g)_{g \in \mathbf{G}}$ *is upper bounded by* $d_{A,N}^{\min} \leq \sqrt{|\mathbf{G}|}$, *independently of the number of queries.*

The bound follows from the fact that, for a finite group $\mathbf{G}$, the dimensions of the irreducible subspaces are upper bounded by $\sqrt{\mathbf{G}}$ (cf. the background on representation theory provided in the Appendix). Corollary 5 is useful to give a quick estimate of the size of the ancilla. Such an estimate is actually tight in the example of the shift-and-multiply gates U_{pq} (cf. Example 3), which form a representation of the group $\mathbf{G} = \mathbb{Z}_d \times \mathbb{Z}_d$. Since the size of the group is $|\mathbf{G}| = d^2$, the bound gives $d_{A,N}^{\min} \leq d$. In other words, this means that unambiguous discrimination can be achieved with an ancilla that is of the same size of the input system. This is exactly what is done by the dense coding protocol [40].

7. Ancilla-free gate discrimination

We conclude the paper by briefly discussing parallel discrimination strategies that do not use ancillas. These strategies involve the preparation of a multipartite input state $|\Psi\rangle \in \mathcal{H}^{\otimes N}$ and on the application of the unknown gate on each of the N systems, thus obtaining the output state $|\Psi_x\rangle = (U_x^{\otimes N})|\Psi\rangle$, as in figure

$$\tag{30}$$

We refer to these strategies as *ancilla-free*. Note that the only difference between an ancilla-free strategy and a general parallel strategy, as in Eq. (17), is the presence of a non-trivial ancilla system. We will now show that the ancilla system can be always traded for an additional number of queries:

Theorem 4 *Every parallel discrimination strategy using a d_A-dimensional ancilla can be replaced by an ancilla-free strategy using a finite number of extra queries. For large d_A, the minimum number of extra queries scales as $O(\log_d(d_A))$.*

This result guarantees that a discrimination strategy using a large number N_A of ancillary qubits can be replaced by an ancilla-free strategy that uses $O(N_A)$ extra queries to the black box. Essentially, this means that the black box queries are a stronger resource than the use ancillary qubits, as the latter can be efficiently simulated using the former.

7.1. Upper bounds on the query complexity of ancilla-free unambiguous discrimination

Let us denote by $N_{\min}^{AF}$ the query complexity of ancilla-free unambiguous discrimination, i.e. the minimum number of queries needed to distinguish the gates U unambiguously in an ancilla-free way. It is immediate to see that ancilla-free unambiguous discrimination is possible in N queries if and only if there exist a state $|\Psi\rangle \in \mathcal{H}^{\otimes N}$ such that the output states $|\Psi_x\rangle = U_x^{\otimes N}|\Psi\rangle$ are linearly independent. Using this fact, the upper bounds of Propositions 2 and 3 can be easily adapted, by replacing the dimension $\dim(\mathsf{U})$ and the minimax fidelity F_{U} with the corresponding ancilla-free quantities:

Proposition 5 *The query complexity of ancilla-free unambiguous discrimination of the gates U is upper bounded as*

$$N_{\min}^{AF} \leq |\mathsf{U}| - \dim_{loc}(\mathsf{U}) + 1, \tag{31}$$

where $\dim_{loc}(\mathsf{U})$ is the maximum over all possible input states $|\psi\rangle \in \mathcal{H}$ of the dimension of the subspace spanned by the vectors $(U_x|\psi\rangle)_{x\in\mathsf{X}}$. Another upper bound is given by

$$N_{\min}^{AF} \leq \left\lfloor \frac{\log(|\mathsf{U}| - 1)}{\log\left(F_{loc,\mathsf{U}}^{-\frac{1}{2}}\right)} \right\rfloor + 1, \tag{32}$$

where $F_{loc,\mathsf{U}}$ is the local minimax fidelity $F_{loc,\mathsf{U}} := \min_{|\psi\rangle\in\mathcal{H}} \max_{x\neq y} |\langle\psi|U_x^\dagger U_y|\psi\rangle|^2$.

The proof is the obvious adaptation of proof for general parallel strategies, provided in the Appendix. Note that, due to the prohibition to use ancillas, one has $\dim_{loc}(\mathsf{U}) \leq \dim(\mathsf{U})$ and $F_{\mathsf{U},loc} \geq F_{\mathsf{U}}$. Therefore, the values of the upper bounds in Proposition 5 are larger than the values of the upper bounds in Propositions 2 and 3. Nevertheless, even without the use of ancillas, the minimum number of queries needed for unambiguous discrimination is always less than $|\mathsf{U}| - 1$, the minimum number of tests that would be needed to identify a gate in U via pairwise eliminations. The local fidelity bound of Eq. (32) provides an even better estimate when the unitaries in U generate a SIC-POVM [58, 59], in which case one has $F_{loc,\mathsf{U}} \leq (d+1)^{-1}$, implying that $N = 4$ queries are sufficient for ancilla-free discrimination. This estimate is much better than the estimate coming from the linear bound of Eq. (31), which in the case under consideration gives a quadratic scaling with the dimension $N_{\min}^{AF} \leq d^2 - d + 1$.

7.2. Perfect ancilla-free discrimination of generalized t-designs

For generalized t-designs, one can prove that unambiguous discrimination coincides with perfect discrimination, *even in the case of ancilla-free strategies*:

Proposition 6 *Let* U *be a generalized t-design. Then, the following are equivalent*

(i) *there exists an ancilla-free strategy for unambiguous discrimination using $N \leq t$ queries*

(ii) *there exists an ancilla-free strategy for perfect deterministic discrimination of the gates using $N \leq t$ queries.*

A case of special interest is when unitaries form a group representation, namely U $\equiv (U_g)_{g \in \mathbf{G}}$. Two group-theoretic lower bounds on the query complexity were originally derived in [55]. We include them here for completeness, concluding our general investigation of gate identification with multiple queries:

Proposition 7 ([55]) *For every perfect discrimination strategy using N parallel queries and a d_A-dimensional ancilla there exists a perfect ancilla-free strategy using $N + \Delta N_{\min}$ queries, with*

$$\Delta N_{\min} \leq \left\lceil \frac{\log d_A + \log \sqrt{|\mathbf{G}|}}{\log d} \right\rceil . \tag{33}$$

Note that for a strategy using a large number of ancillary qubits $N_A \gg \log \sqrt{|\mathbf{G}|}$, the number of extra queries to the black box scales as $O(N_A)$, as anticipated by theorem 4. Note also that the bound is independent of the number of queries of the initial ancilla-assisted strategy: in order to apply the bound, we only need to know that the original strategy allowed for perfect discrimination.

The last bound is expressed in terms of the maximum entanglement fidelity between pairs of unitaries. The entanglement fidelity between two unitaries U_g and U_h, defined as

$$F_{ent}(U_g, U_h) := \frac{\left| \text{Tr}[U_g^\dagger U_h] \right|^2}{d^2} ,$$

is the fidelity between the states $|\Phi_g\rangle = (U_g \otimes I)|\Phi\rangle$ and $|\Phi_h\rangle = (U_h \otimes I)|\Phi\rangle$, obtained by applying the two unitaries on one side of the maximally entangled state $|\Phi\rangle = (\sum_{n=1}^d |n\rangle|n\rangle)/\sqrt{d}$. Thanks to the group structure, the maximum entanglement fidelity over all pairs of unitaries in U is given by $F_{ent,\mathsf{U}} = \max_{g \in \mathbf{G}} |\text{Tr}[U_g]|^2/d^2$.

The other quantity appearing in the bound is the number of unitaries U_g that can be confused with the identity, given by

$$C = |\{U_g \neq I \mid F_{ent}(U_g, I) \neq 0\}| .$$

With the above definitions we have the following

Proposition 8 (Entanglement fidelity bound [55]) *If*

$$d^N \left(1 - F_{ent,\mathsf{U}}^{N/2} C \right) \geq |\mathbf{G}|, \tag{34}$$

then the unitaries in U *can be perfectly distinguished with an ancilla-free strategy using N queries.*

The condition of Eq. (34) contains implicitly an upper bound on the query complexity of ancilla-free discrimination. The fact that an *upper bound* on $N_{\min}^{AF}$ can be expressed in terms of the entanglement fidelity is quite surprising, because applying the unitaries on one side of a maximally entangled state is not an allowed strategy in ancilla-free gate discrimination.

Let us discuss the consequences of Eq. (34). The most immediate consequence is that when the entanglement fidelity is zero, perfect ancilla-free discrimination is possible with $N = \lceil \log_d |\mathbf{G}| \rceil$ queries. This value is actually the optimal one, because $\lceil \log_d |\mathbf{G}| \rceil$ is the minimum number of copies that is needed to pack $|\mathbf{G}|$ orthogonal vectors in the N-fold tensor product of a d-dimensional Hilbert space. An example of this situation is the discrimination of shift-and-multiply gates, already discussed in Example 3:

Example 7 (Ancilla-free discrimination of shift-and-multiply gates) *For a d-dimensional quantum system, the shift-and-multiply gates $U_{pq} = S^p M^q$ are d^2 mutually orthogonal unitaries. Eq. (34) then predicts that perfect ancilla-free discrimination is possible with $N_{\min}^{AF} = 2$ queries. A concrete discrimination strategy consists in preparing two probe systems in the state $|0\rangle|f_0\rangle$, where $|f_0\rangle = \left(\sum_{n=1}^d |n\rangle \right) / \sqrt{d}$ is the first vector of the Fourier basis. Applying two queries of the unknown gate U_{pq} we then obtain the output state $|p\rangle|q\rangle$, from which p and q can be read out in a perfect deterministic way.*

Before concluding, we note that, more generally, the optimal scaling $N_{\min}^{AF} = O(\log_d |\mathbf{G}|)$ can be achieved whenever the condition

$$\frac{\log \left(\frac{C}{1-\alpha} \right)}{\log \left(F_{ent,\mathsf{U}}^{-\frac{1}{2}} \right)} \leq \log_d |\mathbf{G}| \tag{35}$$

is satisfied for some constant $\alpha > 0$. Indeed, under this condition the query complexity of ancilla-free gate discrimination can be bounded between $\lceil \log_d |\mathbf{G}| \rceil$ and $\lceil \log_d |\mathbf{G}| \rceil + \log_d \alpha^{-1}$, thus implying

$$N_{\min}^{AF} = O(\log_d |\mathbf{G}|).$$

The argument is simple and is provided in the Appendix.

8. Conclusion

In this paper we investigated the problem of identifying an unknown unitary gate in a finite set of alternatives, using both deterministic and probabilistic discrimination strategies, and allowing the unknown gate to be queried multiple times and to be be used in parallel or in series in arbitrary quantum circuits. In this scenario, we provided upper and lower bounds on the amount of resources needed to achieve unambiguous and perfect gate identification. Specifically, we gave bounds on the query complexity and the minimum size of the ancillas.

Most of our results stem from two key observations. The first observation is that unambiguous gate discrimination can be parallelized: if unambiguous discrimination is

possible with N queries, then unambiguous gate discrimination must also be possible by applying the N queries in parallel on a suitable entangled state. The second key observation is based on the definition of generalized t-designs, a definition that includes unitary t-designs and group representations as special cases. The remarkable feature of generalized t-designs is that for strategies using $N \leq t$ queries there is no difference between unambiguous and perfect discrimination. Using this fact, one can reduce the analysis of perfect gate discrimination to the simpler analysis of unambiguous gate discrimination. Finally, motivated by the application to quantum communication in the lack of shared reference frames, we considered discrimination strategies where ancillas are not allowed, providing upper bounds on the number of extra queries that are needed to make up for this limitation.

Our results suggest several directions of further research. First of all, up to now we considered the question whether or not unambiguous discrimination is possible with a certain number of queries. However, as the example of Grover's algorithm clearly shows, sometimes the probability of unambiguous discrimination could be too small to have a useful application. Hence, for future developments it is important to have bounds on the probability of the inconclusive outcome as a function of the number of queries. Our work addressed the question in the simplest case, namely the case of generalized t-designs, where the probability of the inconclusive outcome is always zero. In other cases, like the discrimination of Grover's unitaries, an estimate of the probability of the inconclusive outcome as a function of the number of queries can be obtained with the techniques developed by Eldar [61]. Our work provides a useful first step in this direction, because knowing that the unitaries are unambiguously distinguishable, and therefore, that they produce linearly independent output states is a necessary condition for the application of the techniques of Ref. [61]. Besides the evaluation of the probability of the inconclusive outcome, it is also worth relaxing the requirement of perfect gate identification, allowing for a small probability of error $p_e \leq \epsilon$. In this case, the interesting quantity would be the minimum number of queries $N_{\min,\epsilon}$ needed to achieve gate identification with error probability smaller than ϵ. For generalized t-designs, our expression for the maximum probability of correct discrimination, given by $p_N^{\max} = \dim(\mathsf{U}_N)/|\mathsf{U}|$ (cf. theorem 3), gives a starting point for the evaluation of $N_{\min,\epsilon}$.

Another interesting development suggested by our work is the experimental demonstration of quantum advantages in gate discrimination with multiple queries. While the viability of unambiguous state discrimination has been demonstrated in many experiments (see e.g. [62, 63, 64, 65]), the experimental realization of gate discrimination strategies is a rather unexplored territory. A recent experiment demonstrating unambiguous discrimination of two non-orthogonal gates was reported in Ref. [66] in the single-query scenario. As the level of control in the experiments increases, it would be highly desirable to have proof of principle demonstrations of the advantage of joint multi-query discrimination strategies over strategies based on pairwise elimination, both in the case of perfect discrimination and of unambiguous gate discrimination. Moreover, our results on gate discrimination without ancillas show suggest one may have quantum

advantages even with relatively small amounts of entanglement.

Appendix

Proof of theorem 1.

We first prove necessity. The condition for error-free discrimination is equivalent to the existence of at least one $x_0 \in \mathsf{X}$ such that $p_N(x_0|x) = 0 \quad \forall x \neq x_0$, which in turn is equivalent to the condition that the output state $|\Psi_{x_0}\rangle$ in Eq. (14) is linearly independent from the states $(|\Psi_x\rangle)_{x \neq x_0}$. Since the function $U_x^{\otimes N} \mapsto |\Psi_x\rangle$ is linear, the condition $U_{x_0}^{\otimes N} \notin \mathsf{Span}(U_x^{\otimes N})_{x \neq x_0}$ is necessary for error-free discrimination. Similarly, the condition for unambiguous discrimination is equivalent to requirement that $p_N(x_0|x) = 0 \quad \forall x \neq x_0$, which in turn is equivalent to the requirement that the output states $(|\Psi_x\rangle)_{x \in \mathsf{X}}$ are linearly independent. Independence of the states $(|\Psi_x\rangle)_{x \in \mathsf{X}}$ implies independence of the unitaries $(U_x^{\otimes N})_{x \in \mathsf{X}}$. Both conditions are also sufficient, because the linear function $U_x^{\otimes N} \mapsto |\Phi_x\rangle^{\otimes N}$ defined by $|\Phi_x\rangle := (U_x \otimes I)|\Phi\rangle$, $|\Phi\rangle := \sum_{n=1}^{d} |n\rangle|n\rangle/\sqrt{d}$ is invertible, and therefore preserves linear independence. Note that the states $|\Phi_x\rangle$ can be obtained from a parallel strategy where N pairs of systems are prepared in the state $|\Phi\rangle^{\otimes N}$ and the unitary U_x is applied on the first system of each pair. ∎

Proof of proposition 1

By theorem 1, unambiguous discrimination is possible only if $\dim(\mathsf{U}_N) = |\mathsf{U}|$. On the other hand, thinking of each unitary $U^{\otimes N}$ as a vector in the symmetric subspace of $\left(\mathbb{C}^{d^2}\right)^{\otimes N}$ we have $\dim \mathsf{U}_N \leq \binom{N + d^2 - 1}{d^2 - 1}$. The bound is tight, because one can always choose the unitaries U_N to be a spanning set for the symmetric subspace of $\left(\mathbb{C}^{d^2}\right)^{\otimes N}$. ∎

Proof of proposition 2

Let $\mathsf{S} = (v_x)_{x \in \mathsf{X}}$ be a finite set of vectors in a vector space V, with the property that every two distinct vectors in S are linearly independent. Under this hypothesis, Chefles proved that $\dim \mathsf{Span}(v_x^{\otimes N+1}) \geq \dim \mathsf{Span}(v_x^{\otimes N}) + 1$ [57]. Applying the result to the set $\mathsf{U}_N := (U_x^{\otimes N})_{x \in \mathsf{X}}$ gives $\dim(\mathsf{U}_N) \geq \dim(\mathsf{U}) + N - 1$. Hence, for the unitaries in U_N are linearly independent for $N = |\mathsf{U}| - \dim(\mathsf{U}) + 1$. ∎

Proof of lemma 2

Suppose that $\sum_{y \in \mathsf{X}} c_y |\psi_y\rangle^{\otimes N} = 0$. Multiplying by $\langle \psi_x |^{\otimes N}$, taking the modulus, and summing over x we obtain

$$\sum_{x \in \mathsf{X}} |c_x| = \sum_{x \in \mathsf{X}} \left| \sum_{y \in \mathsf{X}, y \neq x} c_y \langle \psi_x | \psi_y \rangle^N \right|$$
$$\leq \sum_{x \in \mathsf{X}} \sum_{y \in \mathsf{X}, y \neq x} |c_y| F^{N/2}$$
$$= (|\mathsf{X}| - 1) F^{N/2} \left(\sum_{x \in \mathsf{X}} |c_x| \right).$$

Clearly, if $(|\mathsf{X}| - 1) F_{\mathsf{S}}^{N/2} < 1$, the only possible solution is $c_x = 0 \; \forall x \in \mathsf{X}$. Hence, the states $(|\psi_x\rangle^{\otimes N})_{x \in \mathsf{X}}$ are linearly independent. ∎

An alternative proof of lemma 2 can be obtained from the application of the Welch bound [67].

Proof of proposition 3

Choose the input state $|\Psi\rangle \in \mathcal{H} \otimes \mathcal{H}$ so that $\max_{x,y \in \mathsf{X}, x \neq y} |\langle \Psi | (U_x^\dagger U_y \otimes I) | \Psi \rangle|^2 = F_{\mathsf{U}}$. For $F_{\mathsf{U}}^{N/2} \leq 1/(|\mathsf{U}| - 1)$ the states $(|\Psi_x\rangle^{\otimes N})_{x \in \mathsf{X}}, |\Psi_x\rangle := (U_x \otimes I)|\Psi\rangle$ are linearly independent. Therefore, also the unitaries $(U_x^{\otimes N})_{x \in \mathsf{X}}$ are linearly independent, i.e. unambiguously distinguishable. ∎

Proof of theorem 2

Using the formalism of quantum combs [56, 68], we express the probability $p_N(y|x)$ as $p_N(y|x) = \langle\!\langle U_x |^{\otimes N} \, T_y \, |U_x\rangle\!\rangle^{\otimes N}$ where $(T_y)_{y \in \mathsf{Y}}$ is a collection of positive operators satisfying suitable normalization conditions [69, 56] (the actual form of the conditions is irrelevant here). The probability of correct identification can be bounded as

$$p_N = \frac{\sum_{x \in \mathsf{X}} p_x \, \langle\!\langle U_x |^{\otimes N} R_N^{-\frac{1}{2}} \left(R_N^{\frac{1}{2}} T_x R_N^{\frac{1}{2}} \right) R_N^{-\frac{1}{2}} |U_x\rangle\!\rangle^{\otimes N}}{\sum_{y \in \mathsf{X}} \mathrm{Tr}[T_y R_N]}$$
$$\leq \sum_{x \in \mathsf{X}} p_x \, \mathrm{Tr}[\rho_x \, R_N^{-\frac{1}{2}} (|U_x\rangle\!\rangle \langle\!\langle U_x|)^{\otimes N} R_N^{-\frac{1}{2}}] \qquad \rho_x := \frac{R_N^{\frac{1}{2}} T_x R_N^{\frac{1}{2}}}{\sum_{y \in \mathsf{X}} \mathrm{Tr}[R_N^{\frac{1}{2}} T_y R_N^{\frac{1}{2}}]}$$

$$\leq \sum_{x \in \mathsf{X}} p_x \; \mathrm{Tr}[\rho_x] \; \| R_N^{-\frac{1}{2}}(|U_x\rangle\!\rangle\langle\!\langle U_x|)^{\otimes N} R_N^{-\frac{1}{2}}\|_\infty$$

$$\leq \max_{x \in \mathsf{X}} p_x \; \langle\!\langle U_x|^{\otimes N} \; R_N^{-1} \; |U_x\rangle\!\rangle^{\otimes N},$$

the last inequality coming from the condition $\sum_{x \in \mathsf{X}} \mathrm{Tr}[\rho_x] = 1$. Defining

$$x_{\max} := \mathrm{argmax} \; p_x \; \langle\!\langle U_x|^{\otimes N} \; R^{-1} \; |U_x\rangle\!\rangle^{\otimes N},$$

the bound can be saturated by applying the N queries of U_x in parallel on the maximally entangled state $|\Phi\rangle^{\otimes N}, |\Phi\rangle := |I\rangle\!\rangle/\sqrt{d}$, and by performing the POVM $(P_y)_{y \in \mathsf{Y}}$ defined by $P_{x_{\max}} = R^{-1}(|U_{x_{\max}}\rangle\!\rangle\langle\!\langle U_{x_{\max}}|)^{\otimes N} R^{-1}/\langle\!\langle U_{x_{\max}}|)^{\otimes N} R^{-2}|U_{x_{\max}}\rangle\!\rangle$, $P_? = I - P_{x_{\max}}$, $P_y = 0$ for every $y \neq x_{\max}$. ∎

Basic representation-theoretic facts

Since generalized t-designs have an underlying group-theoretic structure, before proceeding to the next proofs it is useful to recall some basic facts about group representations.

Let us denote by $\mathsf{Lin}(\mathcal{H}$ the set of linear operators acting on $\mathcal{H}$ and consider a representation $U : \mathbf{G} \to \mathsf{Lin}(\mathcal{H})$, $g \mapsto U_g$ of some (compact) group $\mathbf{G}$. Here we allow U to be a a *unitary projective representation (UPR)*, with multiplier $\omega : \mathbf{G} \times \mathbf{G} \to \mathbb{C}$. In short, this means that $U_g U_h = \omega(g, h)U_{gh}$, $\forall g, h \in \mathbf{G}$. The most familiar case is the case of the unitary representations (UR), for which $\omega(g, h) = 1 \; \forall g, h \in \mathbf{G}$.

With a suitable choice of basis, the Hilbert space can be decomposed as a direct sum of tensor product pairs

$$\mathcal{H} = \bigoplus_{\mu \in \mathsf{Irr}(U)} (\mathcal{R}_\mu \otimes \mathcal{M}_\mu), \tag{36}$$

where the sum runs over the set $\mathsf{Irr}(U)$ of all inequivalent irreducible representations (*irreps*) contained in the decomposition of U, $\mathcal{R}_\mu$ is a *representation space* of dimension d_μ, carrying the irrep U^μ, and $\mathcal{M}_\mu$ is a *multiplicity space* of dimension m_μ, m_μ being the multiplicity of the irrep U^μ in the decomposition of U. Eq. (36) implies that the representation U can be written in the block diagonal form

$$U = \bigoplus_{\mu \in \mathsf{Irr}(U)} (U^\mu \otimes I_{\mathcal{M}_\mu}), \tag{37}$$

where $I_{\mathcal{M}_\mu}$ denotes the identity matrix on $\mathcal{M}_\mu$. Note that all the irreps $U^\mu \in \mathsf{Irr}(U)$ must have the same multiplier ω.

Using Eq. (37) and the orthogonality of matrix elements, one can prove that the set of unitaries $\mathsf{U} := (U_g)_{g \in \mathbf{G}}$ satisfies

$$\dim(\mathsf{U}) = \sum_{\mu \in \mathsf{Irr}(U)} d_\mu^2. \tag{38}$$

Due to the importance of linear independence in the gate discrimination problem, this equation will become very useful in the following section.

A representation that plays a key role in gate discrimination is the *regular representation*, which for finite groups is a representation of $\mathbf{G}$ on the Hilbert space $\mathcal{H} = \mathbb{C}^{|\mathbf{G}|}$, equipped with the orthonormal basis $\{|g\rangle \mid g \in \mathbf{G}\}$. Precisely, the regular representation with multiplier ω is the projective representation $U^{reg,\omega} : \mathbf{G} \to \mathsf{Lin}(\mathbb{C}^{|\mathbf{G}|})$ defined by

$$U_g^{reg,\omega}|h\rangle = \omega(g, h) |gh\rangle , \qquad \forall g, h \in \mathbf{G} \tag{39}$$

The regular decomposition is reducible and its decomposition is

$$U_g^{reg,\omega} = \bigoplus_{\mu \in \mathrm{Irr}(\mathbf{G}, \omega)} \left(U_g^\mu \otimes I_{\mathcal{M}_\mu} \right) \qquad \mathcal{M}_\mu \simeq \mathbb{C}^{d_\mu} \tag{40}$$

where $\mathrm{Irr}(\mathbf{G}, \omega)$ denotes the set of all the irreps of $\mathbf{G}$ with multiplier ω [in particular, $\mathrm{Irr}(\mathbf{G}, 1)$ is the set of all *unitary* irreps of $\mathbf{G}$]. Note that every irrep appears with multiplicity $m_\mu = d_\mu$. Choosing $g = e$ (the identity element in the group) and taking the trace on both sides of Eq. (40) one obtains

$$|\mathbf{G}| = \sum_{\mu \in \mathrm{Irr}(\mathbf{G}, \omega)} d_\mu^2 , \tag{41}$$

which holds for every possible multiplier ω. Finally, combining Eqs. (38) and (41), one gets the following statement:

Proposition 9 *Let $\mathbf{G}$ be a finite group and let $U : \mathbf{G} \to \mathsf{Lin}(\mathcal{H})$ be a UPR with multiplier ω. Then, the unitaries $(U_g)_{g \in \mathbf{G}}$ are linearly independent if and only if the decomposition of U contains all the irreps in $\mathsf{Irr}(\mathbf{G}, \omega)$.*

Proof of theorem 3

Let $\mathbf{G}$ the compact group such that $\sum_{x \in \mathsf{X}} (U_x \otimes \overline{U}_x)^{\otimes N} = \int \mathrm{g} \, (U_g \otimes \overline{U}_g)^{\otimes N}$, or equivalently, $\sum_{x \in \mathsf{X}} U_x^{\otimes N} A U_x^{\dagger \otimes N} = \int \mathrm{g} \, U_g^{\otimes N} A U_g^{\dagger \otimes N}$ for every operator $A \in \mathsf{Lin}(\mathcal{H}^{\otimes N})$. Exploiting the decomposition of $U^{\otimes N}$, one can write $U_x^{\otimes N} = \bigoplus_{\mu \in \mathsf{Irr}(U)} \left(U_x^\mu \otimes I_{\mathcal{M}_\mu} \right)$ and, therefore, $|U_x\rangle\!\rangle^{\otimes N} = \bigoplus_{\mu \in \mathsf{Irr}(U^{\otimes N})} |U_x^\mu\rangle\!\rangle |I_{\mathcal{M}_\mu}\rangle\!\rangle$. The operator R_N in theorem 2 can be directly computed as

$$\begin{aligned} R_N &= \sum_{x \in \mathsf{X}} p_x \left(|U_x\rangle\!\rangle \langle\!\langle U_x| \right)^{\otimes N} \\ &= \int \mathrm{g} \, \left(|U_g\rangle\!\rangle \langle\!\langle U_g| \right)^{\otimes N} \\ &= \bigoplus_{\mu \in \mathsf{Irr}(U^{\otimes N})} \frac{m_\mu}{d_\mu} \left(I_{\mathcal{R}_\mu} \otimes I_{\mathcal{R}_\mu} \otimes \frac{|I_{\mathcal{M}_\mu}\rangle\!\rangle \langle\!\langle I_{\mathcal{M}_\mu}|}{m_\mu} \right) , \end{aligned}$$

so that, computing the inverse, one has $\langle\!\langle U_x|^{\otimes N} R_N^{-1} |U_x\rangle\!\rangle^{\otimes N} = \sum_{\mu \in \mathsf{Irr}(U^{\otimes N})} d_\mu^2 = \dim(\mathsf{U}^{\otimes N})$ [cf. Eq. (38)]. Inserting this value in Eq. (25) proves Eq. (26). We now prove that for the uniform prior the maximum success probability can be obtained with

a deterministic strategy that uses the N queries in parallel. To this purpose, consider the maximum likelihood input state [47, 49]: this is the state in $\mathcal{H}^{\otimes N} \otimes \mathcal{H}_A$ given by

$$|\Phi_{ML}\rangle := \bigoplus_{\mu \in \mathsf{Irr}(U^{\otimes N})} \sqrt{\frac{d_\mu}{\dim(U^{\otimes N})}} \, |I_{\mathcal{R}_\mu}\rangle\!\rangle,$$

where $|I_{\mathcal{R}_\mu}\rangle\!\rangle = \sum_{n=1}^{d_\mu} |\alpha_n^\mu\rangle|\beta_n^\mu\rangle$, $(|\alpha_n^\mu\rangle)_{n=1}^{d_\mu}$ being an orthonormal basis for $\mathcal{R}_\mu$ and $(|\beta_n^\mu\rangle)_{n=1}^{d_\mu}$ being an orthonormal set of vectors in $\mathcal{M}_\mu \otimes \mathcal{H}_A$ [here the dimension of $\mathcal{H}_A$ is chosen in order to satisfy the relation $d_\mu \leq m_\mu d_A, \forall \mu \in \mathsf{Irr}(U^{\otimes N})$]. Applying the N queries in parallel one obtains the output states $|\Phi_{ML,x}\rangle := (U_x^{\otimes N} \otimes I_A)|\Phi_{ML}\rangle$. Optimal discrimination can be achieved deterministically using the square root measurement [70], which in this case has POVM elements $P_x := \frac{\dim(\mathsf{U}_N)}{|\mathsf{U}|} \, |\Phi_{ML,x}\rangle\langle\Phi_{ML,x}|$. ∎

Proof of proposition 4

The proof is an immediate generalization of the proof of Lemma 1 in Ref. [48]. We provide it here just for the sake of completeness. Consider the irreducible decomposition $\mathcal{H}^{\otimes N} = \bigoplus_{\mu \in \mathrm{Irr}(U^{\otimes N})} \mathcal{R}_\mu \otimes \mathcal{M}_\mu$ associated to the group representation $U : g \mapsto U_g$. Take an ancillary Hilbert space $\mathcal{H}_A$ of dimension $d_A = \max_{\mu \in \mathrm{Irr}(U^{\otimes N})} \lceil d_\mu/m_\mu \rceil$ and define $|\Phi\rangle \in \mathcal{H}^{\otimes N} \otimes \mathcal{H}_A$ to be the unit vector $|\Phi\rangle = \bigoplus_{\mu \in \mathrm{Irr}(U^{\otimes N})} c_\mu|\Phi_\mu\rangle$, where c_μ are non-zero coefficients and $|\Phi_\mu\rangle$ is a maximally entangled state in $\mathcal{R}_\mu \otimes \mathcal{M}'_\mu$, $\mathcal{M}'_\mu := \mathcal{M}_\mu \otimes \mathcal{H}_A$. Then, the number of linearly independent states of the form $|\Phi_x\rangle := (U_x^{\otimes N} \otimes I_A)|\Phi\rangle, x \in \mathsf{X}$ is equal to the number of linearly independent unitaries in the set $(U_x^{\otimes N})_{x \in \mathsf{X}}$. In particular, the unitaries $(U_x^{\otimes N})_{x \in \mathsf{X}}$ are linearly independent (i.e. unambiguously distinguishable) iff the states $(|\Phi_x\rangle)_{x \in \mathsf{X}}$ are linearly independent (i.e. unambiguously distinguishable), that is, iff unambiguous discrimination is possible using a d_A-dimensional ancilla satisfying $d_A = \max_{\mu \in \mathrm{Irr}(U^{\otimes N})} \lceil d_\mu/m_\mu \rceil$. ∎

Proof of corollary 4

Every unitary set U is contained in the group $U(d)$. The irreps of the tensor representation $U \mapsto U^{\otimes N}$ are labelled by Young diagrams of N boxes arranged in at most d rows and their dimensions (multiplicities) are given by [71]

$$d_\mu = \prod_{(i,j)\in\mu} \frac{d+j-i}{|h_{ij}|} \qquad \left(m_\mu = \frac{N!}{\prod_{(i,j)\in\mu} |h_{ij}|} \right), \tag{42}$$

where the products runs over the boxes in the Young diagram μ, each box being identified by its row-column coordinates (i, j). Here, $|h_{ij}|$ is the length of the hook centred on the box (i, j). Taking the ratio, one obtains

$$\frac{d_\mu}{m_\mu} \leq \frac{\prod_{(i,j)\in\mu} d+j-i}{N!} \leq \binom{N+d-1}{d-1} \qquad \forall \mu \in \mathrm{Irr}(U^{\otimes N}).$$

∎

Proof of theorem 4

The idea is to encode the ancilla space into one of the multiplicity spaces $\mathcal{M}_{\mu_0}$ contained in the decomposition of the M-fold tensor representation of $U(d)$. This can be done using the invariant encoding of Ref. [54]. Precisely, choosing $\mu_0 \in \mathsf{Irr}(U^{\otimes M})$ such that $m_{\mu_0} \geq d_A$, one can encode the input state $|\Psi\rangle \in \mathcal{H}^{\otimes N} \otimes \mathcal{H}_A$ through the isometric embedding $V_{\mu_0} : \mathcal{H}^{\otimes N} \otimes \mathcal{H}_A \to \mathcal{H}^{\otimes N} \otimes \mathcal{R}_{\mu_0} \otimes \mathcal{M}_{\mu_0} \subset \mathcal{H}^{\otimes(N+M)}$ defined by $V_{\mu_0}|\alpha\rangle|\beta\rangle := |\alpha\rangle|\psi_0\rangle|\beta\rangle$, $|\psi_0\rangle \in \mathcal{R}_\mu$ being a fixed unit vector. One way to satisfy the condition $m_{\mu_0} \geq d_A$ is to choose $M = d \cdot l$ and to set μ_0 to be the Young diagram with d rows of length l. By Eq. (42) one has

$$m_{\mu_0} = \frac{M!}{(l!)^d \prod_{k=0}^{d-1} \binom{l+k}{k}} \geq \frac{M!}{(l!)^d \prod_{k=0}^{d-1}(l+1)^k} = \frac{M!}{(l!)^d (l+1)^{d(d-1)/2}}.$$

Hence, for large M the Stirling approximation yields $\log_d(m_{\mu_0}) = M \log_d(M/e) - M \log_d(l/e) - O(\log_d l) = M - O(\log M)$. For large d_A, the condition $m_{\mu_0} \geq d_A$ is then satisfied whenever $M \geq (1 + \epsilon) \log_d(d_A)$, $\epsilon > 0$. $\blacksquare$

Proof of proposition 6

Suppose that the states $|\Psi_x\rangle := U_x^{\otimes N}|\Psi\rangle$ are linearly independent and define $\rho := 1/|\mathsf{U}| \sum_{x\in\mathsf{X}} |\Psi_x\rangle\langle\Psi_x|$. Then, the states $|\Phi_x\rangle := \sqrt{1/|\mathsf{U}|}\rho^{-1/2}|\Psi_x\rangle$ are mutually orthogonal. Since the unitaries form an t-design, we have $[\rho, U_x^{\otimes N}] = 0 \quad \forall x \in \mathsf{X}$. Hence, the orthogonal states $|\Phi_x\rangle$ are generated by applying $U_x^{\otimes N}$ to the state $|\Phi\rangle := \sqrt{1/|\mathsf{U}|}\rho^{-1/2}|\Psi\rangle$. This yields the desired strategies for perfect discrimination. $\blacksquare$

Proof of proposition 7

The proof invokes the following bound on the sum of the multiplicities arising in a tensor representation:

Lemma 2 *Let m_μ be the multiplicity of the irrep μ in the decomposition of $U^{\otimes M}$. Then,*

$$\sum_{\mu\in\mathsf{Irr}(U^{\otimes M})} m_\mu \geq \frac{d^M}{\sqrt{|\mathbf{G}|}} \tag{43}$$

Proof. Setting $g = e$ in the decomposition $U_g^{\otimes M} = \bigoplus_{\mu\in\mathsf{Irr}(U^{\otimes M})} (U_g^\mu \otimes I_{\mathcal{M}_\mu})$, and taking the trace of $U_g^{\otimes M}$ one gets $d^M = \sum_{\mu\in\mathrm{Irr}(U^{\otimes M})} d_\mu m_\mu$. On the other hand, $d_\mu \leq \sqrt{|\mathbf{G}|}$ for any μ, whence Eq. (43). $\blacksquare$

Proof of proposition 7. Since N queries are sufficient for perfect discrimination using an ancilla, the unitaries $(U_g^{\otimes N})_{g\in\mathbf{G}}$ are linearly independent (lemma 1). Now, suppose that M additional queries are available and consider the decomposition $U^{\otimes M} = \bigoplus_{\mu\in\mathsf{Irr}(U^{\otimes M})} (U_\nu \otimes I_{\mathcal{M}_\nu})$. Since the tensor product cannot decrease the number of linearly independent vectors, the unitaries $(U_g^{\otimes N} \otimes U_g^\mu)_{g\in\mathbf{G}}$ are linearly independent for every

$\mu \in \text{Irr}(U^{\otimes M})$. Equivalently, this means that the representation $U^{\otimes N} \otimes U^{\mu}$ contains all the irreps with multiplier ω^{N+M}, namely $\text{Irr}(U^{\otimes N} \otimes U^{\mu}) \equiv \text{Irr}(\mathbf{G}, \omega^{N+M}) \equiv \text{Irr}(U^{\otimes(N+M)})$. Denoting by $m_{\nu,N+M}$ the multiplicity of the representation $\nu \in \text{Irr}(U^{\otimes N+M})$ and by $\mathcal{M}_{\nu,N+M}$ the corresponding multiplicity space, we have

$$m_{\nu,N+M} \geq \sum_{\mu \in \text{Irr}(U^{\otimes M})} m_{\mu} \geq \frac{d^M}{\sqrt{|\mathbf{G}|}},$$

the last inequality due to lemma 2. Now, the condition $d^M \geq d_A \sqrt{|\mathbf{G}|}$ guarantees $m_{\nu,N+M} \geq d_A$ for every possible irrep $\nu \in \text{Irr}(\mathbf{G}, \omega^{M+N})$. Hence, by proposition 4 the optimal discrimination can be achieved without ancilla. Since $\dim(\mathsf{U}_N) = |\mathbf{G}| = \dim(\mathsf{U}_{N+M})$, the discrimination is perfect (theorem 3). ∎

Proof of proposition 8

The proof takes advantage of the connection between perfect distinguishability and the regular representation:

Lemma 3 *[55, 38] Let $U : \mathbf{G} \to \text{Lin}(\mathcal{H})$ be a UPR with multiplier ω. Then the following are equivalent*

(i) *the gates $(U_g^{\otimes N})_{g \in \mathbf{G}}$ are perfectly/unambiguously distinguishable without ancilla*

(ii) *$U^{\otimes N}$ contains as a sub-representation the regular representation with multiplier ω^N, defined by*

$$U_g^{reg}|h\rangle = \omega^N(g,h)\,|gh\rangle\,, \qquad \forall g, h \in \mathbf{G},$$

where $(|g\rangle)_{g \in \mathbf{G}}$ are orthonormal vectors.

(iii) *the decomposition of $U^{\otimes N}$ contains every irrep $\mu \in \text{Irr}(\mathbf{G}, \omega^N)$ with multiplicity $m_{\mu} \geq d_{\mu}$.*

Proof of proposition 8 . Let $U^{\otimes N} = \bigoplus_{\mu \in \text{Irr}(U^{\otimes N})} U^{\mu} \otimes I_{\mathcal{M}_{\mu}}$ be the decomposition of $U^{\otimes N}$. By the orthogonality of the characters, the multiplicity m_{μ} is given by

$$m_{\mu} = \frac{1}{|\mathbf{G}|} \sum_{g \in \mathbf{G}} \overline{\text{Tr}[U_g^{\mu}]}\,\text{Tr}[U_g]^N.$$

Hence, defining the normalized characters $\nu(g) := \text{Tr}[U_g]/d$ and $\nu_{\mu}(g) := \text{Tr}[U_g^{\mu}]/d_{\mu}$, one has

$$m_{\mu} = \frac{d_{\mu} d^N}{|\mathbf{G}|} \sum_{g \in \mathbf{G}} \overline{\nu_{\mu}(g)}\,\nu^N(g)$$

$$\geq \frac{d_{\mu} d^N}{|\mathbf{G}|}\left(1 - \sum_{g \in \text{Supp}(\nu) \setminus \{e\}} |\nu_{\mu}(g)\,\nu^N(g)|\right)$$

$$\geq \frac{d_{\mu} d^N}{|\mathbf{G}|}\left(1 - F_{\mathsf{U},ent}^{N/2}\,C\right).$$

If N is such that $d^N(1 - F_{\mathsf{U},ent}^{N/2}\,C) \geq |\mathbf{G}|$, then we have $m_{\mu} \geq d_{\mu}$ for every $\mu \in \text{Irr}(\mathbf{G}, \omega^N)$. This means that $U^{\otimes N}$ contains the regular representation U^{reg,ω^N} as a subrepresentation, and, therefore, perfect discrimination is possible by proposition 3. ∎

8.1. Scaling of the ancilla-free query complexity under the condition of Eq. (35)

Eq. (35) can be rewritten as

$$\left(\frac{1}{F_{ent,\mathsf{U}}^{1/2}}\right)^{\log_d |\mathbf{G}|} \geq \frac{C}{1-\alpha},$$

or, equivalently, $F_{ent,\mathsf{U}}^{\log_d |\mathbf{G}|/2} C \leq 1 - \alpha$. By monotonicity of the exponential, this implies that, for every $N \geq \log_d |\mathbf{G}|$, one has $F_{loc,\mathsf{U}}^{N/2} C \leq 1 - \alpha$, which in turn implies

$$d^N \left(1 - F_{loc,\mathsf{U}}^{N/2} C\right) \geq \alpha \, d^N.$$

Hence, choosing $N = \lceil \log_d |\mathbf{G}| \rceil + \log_d \alpha^{-1}$ the condition of Eq. (34) is satisfied. In conclusion, the query complexity has been bounded has $N_{\min}^{AF} \leq \lceil \log_d |\mathbf{G}| \rceil + \log_d \alpha^{-1}$.

References

[1] Acín A 2001 *Phys. Rev. Lett.* **87**, 177901
[2] D'Ariano G M, Lo Presti P, and Paris M G A 2001 *Phys. Rev. Lett.* **87**, 270404
[3] Bergou J A, Herzog U, and Hillery M 2003 *Phys. Rev. Lett.* **90** 257901
[4] Chefles A and Sasaki M 2003 *Phys. Rev. A* **67** 032112
[5] Duan R, Feng Y, and Ying M 2007 *Phys. Rev. Lett.* **98** 100503
[6] Bergou J A and Hillery M 2005 *Phys. Rev. A* **72** 012302
[7] Chefles A, Kitagawa A, Takeoka M, Sasaki M, and Twamley J 2007 *J. Phys. A: Math. Theor.* **40** 10183
[8] Chiribella G, D'Ariano G M, and Perinotti P 2008 *Phys. Rev. Lett.* **101** 180501
[9] Duan R, Feng Y, and Ying M 2009 *Phys. Rev. Lett.* **103** 210501
[10] Ambainis A 2002 *Journal of Computer and System Science* **64** 750
[11] van Dam W 1998 *Proceedings of the 39th Annual IEEE Symposium on Foundations of Computer Science (FOCS)* (ACM Press, New York) 362
[12] Fahri E, Goldstone J, Gutman S, and Sipser M 1999 *Phys. Rev. A* **60** 4331
[13] Ambainis A, Iwama K, Kawachi A, Masuda H, Putra R H, and Yamashita S 2004 *Proceedings of STACS 2004, Lecture Notes in Computer Science* **2996** (Springer, Berlin) 105
[14] Ambainis A, Iwama K, Kawachi A, Raymond R, and Yamashita S 2006 *Lecture Notes in Computer Science* **4059** (Springer, Berlin) 105
[15] Grover L 1997 *Phys. Rev. Lett.* **79** 325
[16] Bernstein E and Vazirani U 1993 *Proceedings of the 25th Annual ACM Symposium on the Theory of Computing* (ACM Press, New York) 11
[17] Calderbank A R, Rains E M, Shor P W, and Sloane N J A 1998 *IEEE Transactions on Information Theory* **44** 1369
[18] Gottesman D 1996 *Phys. Rev. A* **54** 1862
[19] Knill E 1996 arxiv.org:quant-ph/9608048
[20] Knill E 1996 arXiv:quant-ph/9608049
[21] Ashikhmin A and Knill E 2001 *IEEE Transactions on Information Theory* **47** 3065
[22] Rains E M 1999 *IEEE Transactions on Information Theory* **45** 1827
[23] Klappenecker A and Rötteler M 2002 *IEEE Trans. Inform. Theory* **48** 2392
[24] Boström K and Felbinger T 2002 *Phys. Rev. Lett.* **89** 187902
[25] Deng F-G and Long G-L 2004 *Phys. Rev. A* **69** 052319
[26] Deng F-G and Long G-L 2004 *Phys. Rev. A* **70** 012311
[27] Lucamarini M and Mancini S 2005 *Phys. Rev. Lett.* **94** 140501

[28] Chiribella G, D'Ariano G M, and Perinotti P 2008 *Phys. Rev. Lett.* **101** 180504
[29] Pirandola S, Mancini S, Lloyd S, and Braunstein S 2008 *Nature Physics* **4** 726.
[30] von Korff J and Kempe J 2004 *Phys. Rev. Lett.* **93** 260502
[31] Chiribella G, D'Ariano G M, Perinotti P, and Sacchi M F 2004 *Phys. Rev. Lett.* **93**, 180503
[32] Bagan E, Baig M, and Muñoz-Tapia R 2004 *Phys. Rev. A* **70** 030301
[33] Collins D, Diosi L, Gisin G, Massar S, and Popescu S 2005 *Phys. Rev. A* **72** 022304
[34] Hayashi A, Hashimoto T, and Horibe M 2005 Phys. Rev. A **71** 012326
[35] Chiribella G, D'Ariano G M, and Perinotti P 2008 *AIP Conf. Proc.* **1110** 47
[36] Bartlett S D, Rudolph T, Spekkens R W, Turner P S 2009 *New J. Phys.* **11** 063013
[37] Chiribella G, Giovannetti V, Maccone L, and Perinotti P 2012 *Phys. Rev. A* **86** 010304(R)
[38] Skotiniotis M, Kraus B, and Dür W 2013 *Quantum Information and Computation* **13** 0290
[39] Bisio A, Chiribella G, D'Ariano G M, Facchini S, and Perinotti P 2010 *Phys. Rev. A* **81** 032324
[40] Bennett C H and Wiesner S J 1992 Phys. Rev. Lett. **69** 2881
[41] Wang G and Ying M 2006 *Phys. Rev. A* **73** 042301
[42] Rötteler M 2009 Letc. Notes Comp. Science (Springer, Berlin) 493
[43] Dankert C, Cleve R, Emerson J, and Livine E 2009 *Physical Review A* **80** 012304
[44] Gross D, Audenaert K, and Eisert J 2007 *J. Math. Phys.* **48** 052104
[45] Scott A J 2008 *J. Phys. A* **41** 055308
[46] Roy A and Scott A J 2009 Des. Codes Cryptogr. **53** 13
[47] Chiribella G, D'Ariano G M, Perinotti P, and Sacchi M F 2004 *Phys. Rev. A* **70** 062105
[48] Chiribella G, D'Ariano G M, and Sacchi M F 2005 *Phys. Rev. A* **72** 042338
[49] Chiribella G, D'Ariano G M, Perinotti P, and Sacchi M F 2006 *Int. J. Quantum Inf.* **4** 453
[50] Hayashi M 2012 arXiv:1209.3463
[51] Brassard G, Hoyer P, Boyer M, and Tapp A 1998 *Fortsch. Phys.* **46** 493
[52] Zalka C 1999 *Phys.Rev. A* **60** 2746
[53] Chen J and Ying M 2010 *Quantum Information and Computation* **10** 160
[54] Bartlett S D, Rudolph T, and Spekkens R W 2003 *Phys. Rev. Lett.* **91** 027901
[55] Chiribella G 2006 *Optimal estimation of quantum signals in the presence of symmetry*, PhD Thesis, University of Pavia, available at http://www.qubit.it/educational/thesis/ThesisRevised.pdf
[56] Chiribella G, D'Ariano G M and Perinotti P 2008 *Phys. Rev. Lett.* **101** 060401
[57] Chefles A 2002 *Phys. Rev. A* **65** 052314
[58] Renes J M, Blume-Kohout R, Scott A J, and Caves C M 2004 *J. Math. Phys.* **45** 2171
[59] Caves C M, Fuchs C A, and Schack R 2002 *J. Math. Phys.* **43** 4537
[60] Zhu H 2013 private communication
[61] Eldar Y C 2003 *IEEE Transactions on information theory* **49** 446
[62] Clarke R B M, Chefles A, Barnett S M, and Riis E 2001 *Phys. Rev. A* **63** 040305(R)
[63] Mohseni M, Steinberg A M, and Bergou J A 2004 *Phys. Rev. Lett.* **93**, 200403
[64] Bartušková L, Cernoch A, Soubusta J, and Dušek, M 2008 *Phys. Rev. A* **77** 034306
[65] Becerra F E, Fan J, and Migdall A 2013 *Nature Communications* **4**
[66] Laing A, Rudolph T, and O'Brien J L 2009 *Phys. Rev. Lett.* **102** 160502
[67] Welch L R 1974 *IEEE Trans. Inform. Theory* **20 (3)** 397
[68] Chiribella G, D'Ariano G M and Perinotti P 2009 *Phys. Rev. A* **80** 022339
[69] Gutoski G and Watrous J 2007 Proceedings of the 39th ACM Symposium on the Theory of Computation (STOC) **39** (ACM Press) 565
[70] Hausladen P and Wootters W K 1994 *J. Mod. Opt.* **41** 2385
[71] Fulton W and Harris J 1991 *Representation theory: a first course* (Springer-Verlag, New York)