

Topological obstructions to quantum computation with unitary oracles

Zuzana Gavorová ^{1,2,*}, Matan Seidel ³ and Yonathan Touati ¹

¹*Rachel and Selim Benin School of Computer Science and Engineering, The Hebrew University of Jerusalem, 91904 Jerusalem, Israel*

²*Física Teòrica: Informació i Fenòmens Quàntics, Departament de Física, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain*

³*School of Mathematical Sciences, Tel Aviv University, 69978 Tel Aviv, Israel*



(Received 9 November 2023; revised 10 January 2024; accepted 29 February 2024; published 28 March 2024)

Algorithms with unitary oracles can be nested, which makes them extremely versatile. An example is the phase estimation algorithm used in many candidate algorithms for quantum speedup. The search for new quantum algorithms benefits from understanding their limitations: Some tasks are impossible in quantum circuits, although their classical versions are easy, for example, cloning. An example with a unitary oracle U is the if clause, the task to implement controlled U (up to the phase on U). In classical computation the conditional statement is easy and essential. In quantum circuits the if clause was shown impossible from one query to U . Is it possible from polynomially many queries? Here we unify algorithms with a unitary oracle and develop a topological method to prove their limitations: No number of queries to U and U^\dagger lets quantum circuits implement the if clause, even if admitting approximations, postselection, and relaxed causality. We also show limitations of process tomography, oracle neutralization, and $\sqrt[\dim U]{U}$, U^T , and U^\dagger algorithms. Our results strengthen an advantage of linear optics, challenge the experiments on relaxed causality, and motivate new algorithms with many-outcome measurements.

DOI: [10.1103/PhysRevA.109.032625](https://doi.org/10.1103/PhysRevA.109.032625)

I. INTRODUCTION

Quantum computation is believed to outperform classical computation. Many candidate algorithms for quantum speedup [1–6] use Kitaev’s phase estimation algorithm [7]. Phase estimation is so versatile because it is an *oracle* algorithm; it estimates the phase of any unitary oracle U . In oracle algorithms, the input is an oracle: an unknown subroutine, a black box whose inner workings are hidden. It is only possible to *query* an oracle: to set its input and receive its output. With no assumptions on the subroutines, versatility is the main advantage of oracle algorithms. They can be gradually composed into complicated computation in independent building blocks.

Another advantage of oracles is the provability of limitations of computation. Factoring has a low-complexity quantum algorithm due to Shor [1], while it is *believed* to have only high-complexity classical algorithms. Rigorously proving complexity lower bounds is hard. Unlike factoring, *oracle* problems do have some lower bound proof methods: the polynomial [8], hybrid [9], and adversary [10,11] methods. For example, the problem solved by Grover’s quantum algorithm [12] is an oracle problem. It has a *proven* complexity lower bound in classical [12] as well as quantum [8,9] circuits. In

Grover’s quantum algorithm, the oracle is classical; it is a function on bit strings (representing classical computation). In phase estimation [7] the oracle is quantum; representing any quantum circuit, it is a unitary matrix. Additional versatile algorithms could be singled out by a new, possibly purely quantum, lower-bound method.

Discovering limitations of quantum computation has further advantages. Consider the classically easy tasks, whose quantum analogs are impossible: cloning [13], deleting [14], universal-NOT gate [15], programming [16]. These impossibilities lead to new protocols and deepen our understanding of physics; for example, no cloning inspires quantum cryptography [17] and serves as a check for various axiomatizations of quantum mechanics [18–22].

Usually, after an initial impossibility result, stronger no-go theorems and precise bounds follow by studying approximation, postselection, and complexity generalizations. The first two generalizations are often considered separately: approximate cloning [23,24] independently from postselection (probabilistic) cloning [25], and similarly for other tasks [16,26–30] (see [31] for an exception). The third generalization allows repeated input. Repreparing a state or rerunning a procedure is often necessary, for example, for obtaining statistics from experiments. Hopefully, polynomially many repetitions suffice. One such complexity generalization is N -to- M cloning [31–35]: if $M = N + 1$, N is the *sample complexity* of creating one clone. When the input is an oracle, we are interested in the algorithm’s *query complexity*.

Here we develop a unified formalism for quantum computation with unitary oracles, which captures all of the above generalizations simultaneously. Thus, we are able to derive the strongest-yet no-go theorem for the *if-clause* task [36];

*zuzana.gavorova@uab.cat

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.

the task to implement controlled U given an oracle U , the quantum version of the conditional statement essential to classical computation. We show that an if clause of *any query complexity* is impossible, thus losing what could be a fundamental building block in quantum circuits. This, for example, limits the flexibility of phase estimation, which *assumes* the controlled- U building blocks. Our method exploits features unique to unitary oracles; its novel topological approach adds to the few previously known lower-bound methods. To further demonstrate its applicability, we prove limitations of other previously studied unitary-oracle tasks: the neutralization, fractional power, inverse and transpose tasks. The limitations hold for quantum circuits (even with postselection and relaxed causality), but not for linear optics. This motivates developing implementation-dependent algorithms for oracle problems.

A. Problem statement

To state the problem precisely, we first discuss some preceding works. In addition to phase estimation [7], several known algorithms use oracle access to U for implementing a function of U : the complex conjugate \bar{U} [37], the transpose U^T and inverse U^\dagger [38,39], and the fractional power U^q for some $q \in \mathbb{Q}$ [40,41]. The fractional power algorithms differ from the others in two ways: First, they require queries to U (or controlled U) and to its inverse. Second, they succeed not for *all*, but only for *most*, n -qubit unitary oracles $U \in U(2^n)$: they are (exponential¹) average-case algorithms in the sense of *heuristic schemes* [42, Definition 11]. Complexity theory mostly studies worst-case algorithms, algorithms succeeding even on the worst input. Worst-case algorithms satisfy input-independent guarantees, which is preferred for reliable computation.

In different computational models the same oracle problem might have a different *query complexity*: the minimal number of queries the model needs to solve the problem. In the quantum-circuit model the queries have predefined causal order, and in process matrices the causal order is relaxed [43,44]. While quantum circuits need three queries, process matrices provide a two-query solution to the quantum switch [43], the task to superpose orders of two oracles. Thus, theoretically, process matrices are more powerful. But are they physically relevant? The debate about the physicality of the process-matrix solution to the quantum switch is ongoing, for example, whether it was implemented [45] or only simulated [46] by certain linear optics experiments [47–49].

A discrepancy between quantum circuits and linear optics is already exhibited by a simpler unitary-oracle problem, the *if clause*. Its classical version is fundamental for classical computation: Apply a subroutine if a control bit is 1 and do nothing otherwise. In the quantum version, the control bit is a qubit and the subroutine is a d -dimensional unitary $U \in U(d)$,

¹The algorithms work for inputs U that satisfy a condition captured by the gap parameter g [40]. As a function of g , the algorithms require $O(1/g)$ queries to succeed on these “good” inputs. Haar measure over $U(2^n)$ induces the probability of a “bad” input $\delta = 1 - (1 - g)^{2^n} \leq 2^{-n}g$, thus the algorithms fit within a requested δ with $O(2^n/\delta)$ queries. Since $\delta \geq 1 - e^{-2^n g}$, the exponential query complexity is necessary.

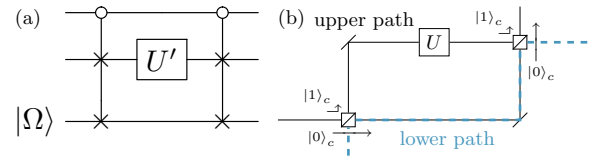


FIG. 1. (a) This algorithm of Kitaev [7] implements $c_0(U')$ if given $|\Omega\rangle$ such that $U'|\Omega\rangle = |\Omega\rangle$. (b) The interferometric if clause [36]: polarizing beam splitters transmit or reflect the photon depending on its polarization degree of freedom $|i\rangle_c$. U acts on other degrees of freedom.

acting on $n = \lceil \log_2 d \rceil$ qubits. Following [36], we identify this with the linear operator

$$c_\phi(U) := |0\rangle\langle 0| \otimes I + e^{i\phi(U)} |1\rangle\langle 1| \otimes U \quad (1)$$

for some real function ϕ . In the special case when ϕ is identically zero, the operator, $c_0(U) = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes U$, is controlled U . Access to $c_0(U)$ is assumed in phase estimation [7]. $c_0(U)$ is implementable given a +1 eigenstate [Fig. 1(a)] or a classical description [50,51] of U . Such information is not available in our oracle setting, which entertains more flexibility. When U is an oracle, the freedom in $\phi(U)$ is necessary: It grants the if clause an insensitivity to U 's global phase [36,52,53]. Araújo, Feix, Costa, and Brukner [36] proved that implementing (1) for all $U \in U(d)$ from only one query to U is impossible in quantum circuits, though it is possible in linear optics [Fig. 1(b)]. To explain the discrepancy, the authors argued that the gate U in Fig. 1(b) is not completely unknown: its position is known, revealing that on the lower path modes it acts as the identity. Any physical gate is restricted to a specific space (specific modes), so Ref. [36] suggested adding the direct sum composition $U \oplus I$ to the quantum-circuit model.

Dong, Nakayama, Soeda, and Muraō [54] found an algorithm for a related task, of implementing $c_\phi^d(U)$ for all $U \in U(d)$, where

$$c_\phi^m(U) := |0\rangle\langle 0| \otimes I + e^{i\phi(U)} |1\rangle\langle 1| \otimes U^m. \quad (2)$$

In their algorithm $m = d$, $\phi(U) = \det(U)^{-1}$, and U is queried d times. Alternatively, with d queries to the fractional power $U^{\frac{1}{d}}$ the algorithm implements $c_\phi^1(U) = c_\phi(U)$ ($m = 1$), the if clause. Thus, the algorithm of [54] relates two unitary-oracle problems: the if clause and the fractional power.

The fractional power dates back to the \sqrt{U} question in Aaronson's 2006 list of “The ten most annoying questions in quantum computing” [55]. For a given fraction, a given unitary has several possible fractional powers. Sheridan, Maslov, and Mosca [40] chose a particular fractional power, and, by finding a contradiction, proved that a worst-case algorithm implementing it for an unknown U is impossible [40, Lemma 1]. This oracle impossibility means that if a worst-case algorithm for some \sqrt{U} exists, it depends on the details of the implementation of U .

Another important oracle task is to remove the effect of an unknown and undesired evolution: the *neutralization* task. Often the evolution, the repeated application of some unitary U , cannot be avoided. However, initializing the state and/or interlacing U with fixed gates can cancel out its effect. For example, spin echo [56] is a neutralization algorithm for

restricted U 's: for qubit rotations around the z axis. Spin echo neutralizes two queries to U . The neutralization algorithm of [54] works for general d -dimensional unitaries. It neutralizes d queries to U .

The tasks mentioned above are all tasks with unitary oracles. Some algorithms and no-go theorems have been found, but maybe a unified approach is possible, with more emphasis on query complexity in order to strengthen the no-gos and to understand the optimality of the known algorithms. We have seen that the linear-optics complexity of the if clause is constant, 1, and that its quantum-circuit complexity is strictly bigger. To bridge the gap, Ref. [36] suggested modifying the quantum circuit formalism. But, is the gap large enough, e.g., beyond $\text{poly}(n)$,² to justify this? What is the quantum-circuit complexity of the if clause? It seems at most exponential because U 's classical description suffices to build an if clause [51], and process tomography [57–59] yields the oracle's classical description from $\exp(n)$ queries.

B. Results overview

In this work we develop a unified formalism for computation with unitary oracles U . We phrase such tasks and worst-case algorithms as functions on unitaries. The algorithm functions have two important properties: continuity and homogeneity. We add to the few previously known lower bound methods a different, topological method and derive results.

The first and the most surprising result shows that the quantum-circuit complexity of the if clause is infinite. We show that no matter how many times a postselection circuit queries U and U^\dagger , it cannot implement the if clause with a nonzero success probability for every $U \in U(d)$, not even approximately! Surprisingly, process tomography suggested above fails for the if clause, and only works for a relaxed variant of the task. We prove this limitation of process tomography directly. Second, we give a different proof that implementing the fractional power $U^{\frac{1}{d}}$ for all $U \in U(d)$ is impossible, from any number of queries to U and U^\dagger . Our contradiction is independent of the particular d th root function. Third, we show that quantum circuits and process matrices fail to neutralize some specific numbers of queries. Last, we prove related limitations regarding the transpose and inverse tasks.

The above results limit versatile quantum computation, and impact our understanding of tomography, measurements, linear optics, and causality. Using process tomography for the if clause has a caveat: Instead of a superoperator estimate of $\rho \mapsto U \rho U^\dagger$, the if clause requires a matrix estimate of U . We show the limitation of such matrix tomography. Defining a relaxed if clause circumvents the limitation, but the algorithm must use a measurement beyond the binary success or fail type. This splits measurements into two groups with different effects on the quantum-circuit query complexity. The quantum-circuit model itself is compared to other models: linear optics and process matrices. On the if clause the quantum-circuit model turns out to be infinitely less efficient than linear optics! One might put some hope into relaxing

causality; maybe linear optics is better matched by process matrices. While arguably true for the quantum switch task, this is wrong for the more fundamental if-clause task: its process-matrix complexity is infinite, too. The advantage of linear optics stems from restricting the oracle. The models with fully general unitary oracles have a property central to our impossibility proofs: homogeneity. Linear optics restrict oracles to the form $1 \oplus U$ which breaks homogeneity. Differently from [36], we attribute the direct sum to the linearity of linear optics (see Sec. V).

The rest of the paper is organized as follows. Section II defines oracle computation using functions on d -dimensional unitaries, $U \in U(d)$. Section III proves the if-clause impossibility and the process tomography limitation by exploiting the continuity of algorithms and the topology of the space $U(d)$ (Lemma 1). Section IV uses this topological approach to prove results regarding the neutralization, $1/d$ th power, transpose, and inverse. In Sec. V, we emphasize that our method applies to the worst-case models with the exception of linear optics. We discuss the cause and the significance of this exception. Then we discuss relaxed causality and measurements.

II. ALGORITHM AS A FUNCTION OF THE ORACLE

An algorithm should solve a problem. Each problem mentioned in the Introduction asks to implement some operator $t(U) \in L(\mathcal{H})$, where $L(\mathcal{H})$ is the set of linear operators from a finite-dimensional Hilbert space \mathcal{H} to itself. We call the function $t : U(d) \rightarrow L(\mathcal{H})$ *task function*. The allowed types of queries to the oracle, for example, U and U^\dagger queries, are also represented by functions on operators (*query functions*) in our example by $id : U \mapsto U$ and $inv : U \mapsto U^\dagger$. Another example is the c_0 query function in phase estimation. A set of allowed query functions is a *query alphabet* Σ . A task function and a query alphabet together form a pair (t, Σ) called *task*. Examples are in Table I, which phrases the above-mentioned preceding works in terms of *tasks*.

Next we represent algorithms by functions of U . In [8,60] a unitary quantum circuit with N queries to an oracle U corresponds to a sequence of unitary transformations

$$V_0, U, V_1, U, \dots, V_{N-1}, U, V_N$$

for some fixed unitaries V_i acting on a possibly larger Hilbert space than the oracle U . We generalize the above in two ways. First, we add the possibility to query U via query functions σ_i :

$$V_0, \sigma_1(U), V_1, \sigma_2(U), \dots, V_{N-1}, \sigma_N(U), V_N.$$

Regarding σ_i as a single character, the string $\mathbf{s} = \sigma_1 \sigma_2 \dots \sigma_N$, called the *query sequence*, is fixed for the algorithm. Second, we allow for a projective measurement at the end $\{\Pi_{\text{succ}}, \Pi_{\text{fail}}\}$, followed by the postselection on success Π_{succ} .

Definition 1 (Postselection oracle algorithm). A *postselection oracle algorithm* (A, Σ_A) is a function $A : U(d) \rightarrow L(\mathcal{H} \otimes \mathcal{K})$ of the form

$$A(U) = \Pi_{\text{succ}} V_N [\sigma_N(U) \otimes I_{\mathcal{K}_N}] \dots V_1 [\sigma_1(U) \otimes I_{\mathcal{K}_1}] V_0, \quad (3)$$

where $\sigma_i \in \Sigma_A$ for all $i \in [N]$. The algorithm implements the total linear operator $A(U)$ (Fig. 2) whenever its projective measurement $\{\Pi_{\text{succ}}, \Pi_{\text{fail}}\}$ yields success. The probability of success must be nonzero if the ancilla Hilbert space \mathcal{K} is

²In the absence of oracles, the quantum-circuit model can simulate optics with only a polynomial overhead [79–82].

TABLE I. Examples of tasks and their approximate solvability by oracle algorithms with postselection.

	Task	For all $U \in \mathcal{U}(d)$	For most $U \in \mathcal{U}(d)$
if clause	$(c_\phi, \{id, inv\})$	<i>Impossible</i> ^a	Unitary algorithm [40,54] ^b
if clause from fractions	$(c_\phi, \{U \mapsto U^{\frac{1}{d}}\})$	Exact unitary algorithm [54]	
Complex conjugate	$(U \mapsto \bar{U}, \{id\})$	Exact unitary algorithm [37]	
Transpose	$(U \mapsto U^T, \{id\})$	Exact algorithm [38,39]	
Inverse	$(inv, \{id\})$	Exact algorithm [37–39] ^b	
Phase estimation	$(phase_est, \{c_0\})$	Unitary algorithm [7]	
Fractional power	$(U \mapsto U^{\frac{1}{d}}, \{id, inv\})$	<i>Impossible</i> [40] ^{a, b}	Unitary algorithm [40]

^aThis result.

^bBy composing two tasks.

initialized to the all-zero state, i.e., for all $U \in \mathcal{U}(d)$ and for all $|\xi\rangle \in \mathcal{H}$,

$$\|A(U)(|\xi\rangle \otimes |\mathbf{0}\rangle_{\mathcal{K}})\|^2 > 0, \quad (4)$$

which is the usual postselection condition [61].

We will equivalently represent a postselection oracle algorithm by (\mathcal{A}, Σ_A) , where the calligraphic font denotes the corresponding function into *superoperators*, i.e.,

$$\mathcal{A}(U)(\rho) := \text{tr}_{\mathcal{K}}[A(U)(\rho \otimes |\mathbf{0}\rangle\langle\mathbf{0}|_{\mathcal{K}})A(U)^\dagger], \quad (5)$$

where we already include the correct initialization of the ancilla, and its tracing out at the end of the algorithm. Denote by $\mathcal{S}_+(\mathcal{H})$ the completely positive superoperators $L(\mathcal{H}) \rightarrow L(\mathcal{H})$ that are trace nonvanishing, i.e., they map density states $\mathcal{D}(\mathcal{H}) \subset L(\mathcal{H})$ to linear operators with strictly positive trace. Then, the postselection condition (4) translates to the requirement that $\mathcal{A} : \mathcal{U}(d) \rightarrow \mathcal{S}_+(\mathcal{H})$.

The following generalization of postselection oracle algorithms relaxes the causal order of the queries.

Definition 2 (Process-matrix algorithm). A *process-matrix algorithm* (\mathcal{A}, Σ_A) is a function $\mathcal{A} : \mathcal{U}(d) \rightarrow \mathcal{S}_+(\mathcal{H})$ such that

$$J_{\mathcal{A}(U)} = \text{tr}_{\text{edges}}^{\text{inner}}[W(J_{\tilde{\sigma}_1(U)} \otimes \cdots \otimes J_{\tilde{\sigma}_N(U)} \otimes I_{\text{open}}^{\text{edges}})] \quad (6)$$

is the contraction in Fig. 3, where W is a fixed matrix, $J_\Psi := \sum_{i,j} |i\rangle\langle j| \otimes \Psi(|i\rangle\langle j|)$ is the Choi isomorphism of Ψ , and $\tilde{\sigma}_i(U)$ are query superoperators, $\tilde{\sigma}_i(U)(\rho) := \sigma_i(U)\rho\sigma_i(U)^\dagger$ for the query functions $\sigma_i \in \Sigma_A$.

As opposed to quantum circuits, the wires in Fig. 3 do not fix the order of the queries. How the queries are connected is determined by W , which may correspond to a fixed causal order, but also to a superposition of orders [43], and beyond [44,62].

It remains to define what it means for an algorithm to achieve a task. Note that we consider only unitary $t(U)$.

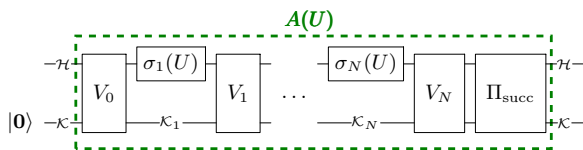


FIG. 2. A postselection oracle algorithm has fixed unitary gates V_i , a fixed projection Π_{succ} , and special slots for queries $\sigma_i(U)$, while identity is applied to the remaining Hilbert space \mathcal{K}_i .

Definition 3 (Exactly achieving a task). A postselection oracle algorithm (\mathcal{A}, Σ_A) *exactly achieves* the task (t, Σ) if $\Sigma_A \subseteq \Sigma$ and if it implements $t(U)$ to the Hilbert space \mathcal{H} , provided that the ancilla Hilbert space \mathcal{K} was initialized to the all-zero state $|\mathbf{0}\rangle$, i.e., the operator equation

$$A(U)(I \otimes |\mathbf{0}\rangle\langle\mathbf{0}|) = t(U) \otimes |g(U)\rangle\langle\mathbf{0}| \quad (7)$$

holds for some $|g(U)\rangle \in \mathcal{K}$.

The freedom in the unnormalized state $|g(U)\rangle$ allows for any global phase and any postselection probability. By the postselection condition (4) $|g(U)\rangle$ is never zero.

To define achieving a task for *both* postselection oracle algorithms and process-matrix algorithms, we use an equivalent equation in terms of (trace nonvanishing) superoperators $\mathcal{A}(U) \in \mathcal{S}_+(\mathcal{H})$:

$$\frac{\mathcal{A}(U)(\rho)}{\text{tr}[\mathcal{A}(U)(\rho)]} = t(U)\rho t(U)^\dagger. \quad (8)$$

One direction of the equivalence is immediate from (5), the other follows from Theorem 2.3 of [63]. The theorem also relates the *errors* in the operator and superoperator languages. Here we continue with superoperators.

Definition 4 (ϵ -approximately achieving a task). Algorithm (\mathcal{A}, Σ_A) ϵ -*approximately achieves* the task (t, Σ)

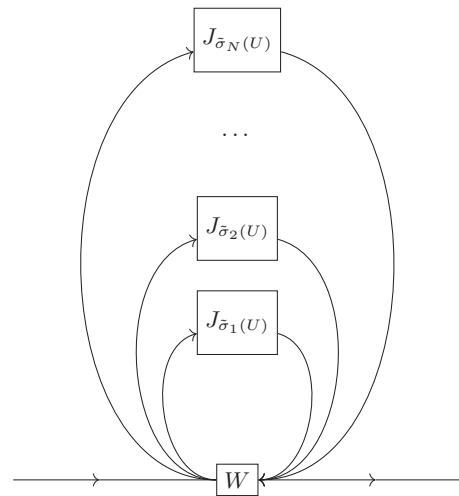


FIG. 3. A process-matrix algorithm is a contraction of queries with a fixed matrix W . W may or may not determine the queries' order.

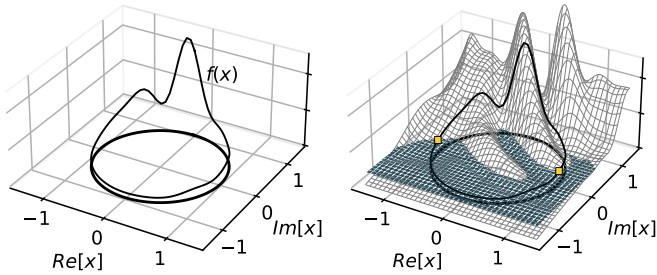


FIG. 4. Intuition for the $n = 1$ case of the Borsuk-Ulam theorem: At some moment the rising tide must hit points that are exactly opposite on the circle.

if $\Sigma_A \subseteq \Sigma$ and if the renormalized output of the algorithm is always close to the task output, i.e., for all $U \in U(d)$,

$$\sup_{\substack{\mathcal{H}' \\ \rho \in \mathcal{D}(\mathcal{H} \otimes \mathcal{H}')}} \left\| \frac{\mathcal{A}(U) \otimes \mathcal{I}(\rho)}{\text{tr}[\mathcal{A}(U) \otimes \mathcal{I}(\rho)]} - [t(U) \otimes I] \rho [t(U) \otimes I]^\dagger \right\|_{\text{tr}} \leq \epsilon, \quad (9)$$

where $\|\cdot\|_{\text{tr}}$ is the trace norm, and $I \in L(\mathcal{H}')$ and $\mathcal{I} \in \mathcal{S}_+(\mathcal{H}')$ are the identity operator and superoperator.

The left-hand side of inequality (9) is *postselection diamond distance* [63]. It simplifies to diamond distance if $\mathcal{A}(U)$ is trace preserving. Our use of the postselection condition and of diamond distance, a worst-case measure, means that ϵ -approximately achieving is worst case: even with the worst inputs ρ and U the algorithm succeeds with nonzero probability and then satisfies the error bound. Thus, our formalism excludes the two algorithms in column four of Table I, and includes all the algorithms in column three, where “exact” means $\epsilon = 0$, and “unitary” means $\Pi_{\text{succ}} = I$.

III. IF-CLAUSE IMPOSSIBILITY VIA TOPOLOGY

Algorithms as functions of U [Eqs. (5) and (6)] have useful properties. The postselection oracle algorithm, function (5), is continuous whenever the query functions are [see (3)]. So is the process-matrix algorithm, function (6), because Choi isomorphism and its inverse are continuous. Moreover, both functions are 0-homogeneous according to the definition below.

Definition 5 (Homogeneity). A function $f : X \rightarrow Y$ is m -homogeneous for some $m \in \mathbb{Z}$ iff $f(\lambda x) = \lambda^m f(x)$ for any $x \in X$ and any scalar λ such that $\lambda x \in X$.

For example, the scalars are restricted to the unit circle (one-sphere) $\lambda \in S^1$ if the domain of f is the space of unitaries $U(d)$. The space $SU(2)$ is homeomorphic to the three-sphere S^3 and the following prominent result in topology studies continuous functions on n -spheres S^n :

Borsuk-Ulam theorem [64]. For any $f : S^n \rightarrow \mathbb{R}^n$ continuous there exist antipodal points mapped to the same value $f(x) = f(-x)$.

While the Borsuk-Ulam theorem (Fig. 4) implies a special case (see Appendix A), for the fully general if-clause impossibility we prove a topological lemma that also exploits homogeneity.

Lemma 1. Let $m \in \mathbb{Z}$. If a continuous m -homogeneous function $f : U(d) \rightarrow S^1$ exists, then d divides m .

For example, the determinant, emerging in the c_ϕ^d algorithm reviewed in Sec. IV A, is a d -homogeneous function: $\det(\lambda U) = \lambda^d \det(U)$. By Lemma 1 this homogeneity degree and its multiples are the only ones possible for continuous functions $U(d) \rightarrow S^1$. Consequently, in our main result c_ϕ^m obeys a dichotomy that depends on whether d divides m ($d|m$):

Theorem 1. Let d be the oracle dimension, let $m \in \mathbb{Z}$ and $\epsilon \in [0, 1/2)$.

(i) If $d|m$ a postselection oracle algorithm exists ϵ -approximately achieving $(c_\phi^m, \{id, inv\})$ for some ϕ . It makes $|m|$ queries.

(ii) If $d \nmid m$ no such postselection oracle algorithm or process-matrix algorithm exists, for any number of queries.

The if-clause ($m = 1$) impossibility is immediate. In addition to the following full proof of Theorem 1, the Appendix contains two more proofs for only the exact, $\epsilon = 0$, impossibility. The “operational” proof in Appendix B reaches a contradiction by using the supposed c_ϕ^m algorithm as a building block in a larger circuit. The proof in Appendix C for only the $m = 1$ case³ gives additional intuition: The special unitary group $SU(d)$ is a d th cover of $PU(d)$, the projective unitary group. This prevents the existence of a continuous map from a unitary superoperator to a matching operator, $PU(d) \rightarrow U(d)$, preventing an exact if-clause algorithm. The full proof below holds for approximations and relies on Lemma 1 proven next.

Proof of Theorem 1. For $d|m$ direction we repeat $|m|/d$ times the algorithm of [54] (Sec. IV A), which exactly achieves $(c_\phi^d, \{id\})$ using d queries. We switch the id to inv queries if $m < 0$.

For the $d \nmid m$ direction, assume an algorithm $(\mathcal{A}, \{id, inv\})$ ϵ -approximately achieves $(c_\phi^m, \{id, inv\})$ for some $\epsilon < \frac{1}{2}$. As mentioned, \mathcal{A} is a continuous, 0-homogeneous function of $U \in U(d)$. Consider the input state $\rho \in \mathcal{D}(\mathcal{H})$, $\rho = |+\rangle\langle +|_c \otimes |0\rangle\langle 0|$, where the first register is the control qubit, and the second the target qudit. Using ρ define $f : U(d) \rightarrow \mathbb{C}$ as

$$f(U) := \text{tr} \left[\frac{\mathcal{A}(U)(\rho)}{\text{tr}[\mathcal{A}(U)(\rho)]} (|1\rangle\langle 0|_c \otimes U^m) \right],$$

which is continuous, m -homogeneous. Observing that

$$\text{tr} \left[c_\phi^m(U) \rho c_\phi^m(U)^\dagger (|1\rangle\langle 0|_c \otimes U^m) \right] = \frac{e^{-i\phi(U)}}{2}, \quad (10)$$

and that multiplication by $|1\rangle\langle 0|_c \otimes U^m$ and taking trace cannot increase trace norm, we get

$$\left| f(U) - \frac{e^{-i\phi(U)}}{2} \right| \leq \left\| \frac{\mathcal{A}(U)(\rho)}{\text{tr}[\mathcal{A}(U)(\rho)]} - c_\phi^m(U) \rho c_\phi^m(U)^\dagger \right\|_{\text{tr}},$$

which is less than ϵ by the assumption. Thus,

$$|f(U)| \geq \frac{1}{2} - \epsilon > 0$$

for all $U \in U(d)$ because $\epsilon < \frac{1}{2}$ by the assumption. This lets us define the continuous function $\widehat{f}(U) := f(U)/|f(U)|$. Re-

³We thank an anonymous referee from the QIP conference for observing that such a proof is possible.

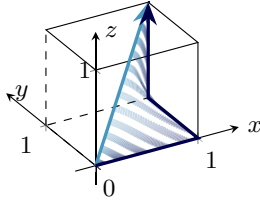


FIG. 5. Sequential and simultaneous increase of each coordinate are homotopic inside the cube $[0, 1]^3$.

call that homogeneity of functions on $U(d)$ is defined with respect to scalars of unit norm only, $\lambda \in S^1$. Thus, $|f(U)|$ is 0-homogeneous, and the continuous function $\hat{f} : U(d) \rightarrow S^1$ is m -homogeneous. By Lemma 1, $d|m$. ■

Proof of Lemma 1. This elementary proof (for a shorter proof see Appendix D) studies the structure of the space $U(d)$, finding paths on this space that are continuously deformable, or *homotopic*, to each other. Denoting a path by $U(t)$, $t \in [0, 1]$, we split $[0, 1]$ to d identical intervals labeled by $j \in \{0, 1, \dots, d-1\}$ and use the parameter $\Delta \in [0, 1]$ to move inside each interval. Consider the paths $U(t) = e^{i2\pi t} I$,

$$U'\left(\frac{j+\Delta}{d}\right) = \begin{matrix} & & 0 & \dots & j & \dots & d-1 \\ \begin{matrix} 0 \\ \vdots \\ j \\ \vdots \\ d-1 \end{matrix} & \begin{pmatrix} \ddots & & & & & & \\ & 1 & & & & & \\ & & e^{i2\pi\Delta} & & & & \\ & & & 1 & & & \\ & & & & \ddots & & \end{pmatrix} & & & & & \end{matrix}$$

and

$$U''\left(\frac{j+\Delta}{d}\right) = \begin{pmatrix} e^{i2\pi\Delta} & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix},$$

where all the off-diagonal matrix elements are zero and the diagonal elements change: $U(t)$ loops on all d of them simultaneously, $U'(t)$ loops on all d of them in sequence, and $U''(t)$ loops only on the zeroth one.

First, we show that U is homotopic to U' . We can write $U(t) = \exp(i2\pi \text{diag}[t, t, \dots, t])$ and $U'(\frac{j+\Delta}{d}) = \exp(i2\pi \text{diag}[1, \dots, 1, \Delta, 0, \dots, 0])$, where \exp is matrix exponentiation and $\text{diag } v$ is the square matrix with the vector v on the diagonal and zeros elsewhere. For any vector in the d -dimensional hypercube, $v \in [0, 1]^d$, the continuous function $v \mapsto \exp(i2\pi \text{diag } v)$ outputs a d -dimensional unitary. If fed a hypercube path $v(t)$ from $[0, \dots, 0]$ to $[1, \dots, 1]$, the function outputs a unitary path starting and ending at the identity: the unitary path $U(t)$, if the hypercube path is the straight line; the unitary path $U'(t)$, if the hypercube path travels along the edges of the hypercube (Fig. 5). Since the two hypercube paths are homotopic, so are U and U' .

To show that U' and U'' are homotopic, note that $U'(\frac{j}{d}) = U''(\frac{j}{d})$ for all j . Therefore, showing their homotopy on each of the d subintervals suffices. Indeed, there exists a continuous unitary path $R_j(s)$, $s \in [0, 1]$ such that conjugating by $R_j(s)$ continuously deforms U' into U'' on the j th subinterval; at $s = 0$, U' stays unchanged, at $s = 1$ it becomes U'' —the role

of the path $R_j(s)$ being analogous to the thick lines in Fig. 5:

$$R_j(s) = \begin{matrix} 0 \\ 1 \\ \vdots \\ j \\ \vdots \end{matrix} \begin{pmatrix} \cos\left(\frac{\pi}{2}s\right) & & & & & & & & i \sin\left(\frac{\pi}{2}s\right) \\ & 1 & & & & & & & \\ & & \ddots & & & & & & \\ & & & 1 & & & & & \\ & & & & \ddots & & & & \\ & i \sin\left(\frac{\pi}{2}s\right) & & & & \cos\left(\frac{\pi}{2}s\right) & & & \\ & & & & & & & & 1 \\ & & & & & & & & \ddots \end{pmatrix}$$

so, indeed, the map $U'(\frac{j+\Delta}{d}) \mapsto R_j(s)U'(\frac{j+\Delta}{d})R_j(s)^\dagger$ at $s = 0$ does nothing and at $s = 1$ swaps the 0th and the j th basis states. This fixes a homotopy between U' and U'' . Combining with the first homotopy, we conclude that U and U'' are homotopic.

Assuming a continuous m -homogeneous function $f : U(d) \rightarrow S^1$, next we show that d must divide m . Since f is m -homogeneous, we have $f(U(t)) = e^{im2\pi t} f(I)$, which corresponds to m loops on the circle S^1 starting and ending at $f(I)$. $U''(t)$ is a periodic function with d periods inside $[0, 1]$, and so is the function $f(U''(t))$. Therefore, if $f(U''(t))$ makes $k \in \mathbb{Z}$ loops on the interval $[0, \frac{1}{d}]$, then it makes kd loops on S^1 on the entire $[0, 1]$ interval. Since f is continuous, $f(U(t))$ must be homotopic to $f(U''(t))$. But on the circle S^1 two paths can be homotopic if and only if they make the same number of loops (this is captured by the fundamental group of S^1). We must have $m = kd$, which proves Lemma 1. ■

A. Process tomography for the if clause?

Theorem 1 seems to contradict process tomography. After we allow approximations and arbitrarily many queries, process tomography seems to offer a solution to the if clause. $\text{exp}(n)$ queries suffice to estimate any oracle's *superoperator* with exponentially small error [65]. To build the if clause, our process tomography should output a matrix estimate $X \in U(d)$, instead of the superoperator estimate $\rho \mapsto X\rho X^\dagger$, because controlled X ,

$$\rho \mapsto (|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes X)\rho(|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes X)^\dagger, \quad (11)$$

contains cross terms. Thus, in this section we study process tomography (of a unitary oracle) that outputs a *matrix* estimate X .

If such tomography is possible, its matrix output X approximately describes the oracle $U \in U(d)$ only up to the global phase because distinguishing the global phase is unphysical. Thus, to quantify the error, we compare $\kappa(X)$ to $\kappa(U)$, where the function $\kappa : U(d) \rightarrow U(d)$ fixes one canonical form for all unitaries equal up to the global phase. Formally, $\kappa(e^{i\alpha}U) = \kappa(U)$ for all $\alpha \in [0, 2\pi)$ (κ is 0-homogeneous) and $\kappa(U) = e^{i\phi(U)}U$ for some real function ϕ . An example is a function that makes the first nonzero matrix element of its input real positive.

While any practical process tomography procedure includes classical postprocessing of measurement outcomes, in principle it is convertible to a fully quantum algorithm with

one measurement at the end that yields the full estimate (in our case a full matrix X): all its matrix elements written to some N -dependent precision. Being a measurement outcome, this matrix estimate X is a random variable that takes values in the discrete space $\Omega_N \subset U(d)$.

Definition 6 (Process tomography yielding a unitary matrix). Process tomography of a unitary oracle is an algorithm making N queries to $U \in U(d)$ and then measuring, obtaining the outcome $X \in \Omega_N \subset U(d)$ with probability $p_{N,U}(X)$. The output distribution $p_{N,U}$ is such that for all $U \in U(d)$,

$$\Pr_{X \sim p_{N,U}(\cdot)}[\|\kappa(X) - \kappa(U)\|_{\text{op}} \leq \epsilon_N] \geq 1 - \delta_N,$$

where $\lim_{N \rightarrow \infty} \epsilon_N = 0$ and $\lim_{N \rightarrow \infty} \delta_N = 0$.

The operator norm $\|\cdot\|_{\text{op}}$ distinguishes Definition 6 from existing process tomography algorithms that ensure a vanishing distance of *superoperators* [65]. Can we relate the two distances? Vanishing operator norm implies vanishing diamond distance of the corresponding superoperators [66, Lemma 12.6]. Conversely, vanishing diamond distance $\|X(\cdot)X^\dagger - U(\cdot)U^\dagger\|_{\diamond}$ implies vanishing

$$\|X - \langle v|U^\dagger X|v\rangle U\|_{\text{op}} = \|\kappa(X) - \langle v|U^\dagger \kappa(X)|v\rangle U\|_{\text{op}} \quad (12)$$

for any unit vector $|v\rangle$ [63, Theorem 1.3]. However, Definition 6 requires the U term to be independent of X . This makes all the difference.

Theorem 2. Process tomography of Definition 6 is impossible.

Proof. The triangle inequality gives

$$\begin{aligned} \|\kappa(U) - \kappa(V)\|_{\text{op}} &\leq \sigma_{N,U} + \sum_{X \in \Omega_N} |p_{N,U}(X) - p_{N,V}(X)| \\ &\quad + \sigma_{N,V}, \end{aligned} \quad (13)$$

where $\sigma_{N,U} := \sum_{X \in \Omega_N} p_{N,U}(X) \|\kappa(U) - \kappa(X)\|_{\text{op}}$ is a sum which can be split to two sums, over “good” and “bad” X values, and then by Definition 6 upper bounded by $\epsilon_N + 2\delta_N$. The middle term of (13) is the total variational distance of two probability distributions. Equivalently, it is the trace distance of two classical density states $\rho_{N,U} = \text{diag}(p_{N,U})$ and $\rho_{N,V} = \text{diag}(p_{N,V})$, the outputs of the process tomography algorithm making N queries to U or V , respectively. By the subadditivity of diamond distance, the trace distance of the outputs is upper bounded $\|\rho_{N,U} - \rho_{N,V}\|_{\text{tr}} \leq N\|U \cdot U^\dagger - V \cdot V^\dagger\|_{\diamond}$. Using the inequality $\|U \cdot U^\dagger - V \cdot V^\dagger\|_{\diamond} \leq 2\|U - V\|_{\text{op}}$ [66, Lemma 12.6] we get

$$\|\kappa(U) - \kappa(V)\|_{\text{op}} \leq 2N\|U - V\|_{\text{op}} + 2\epsilon_N + 4\delta_N. \quad (14)$$

Thus, κ is continuous. But if κ is continuous (it is 0-homogeneous by definition), then defining $f(U) := \langle 0|\kappa(U)U^\dagger|0\rangle = e^{i\phi(U)}$ gives us $f : U(d) \rightarrow S^1$ continuous and -1 -homogeneous, contradicting Lemma 1. ■

Building on top of known process tomography algorithms cannot produce matrix estimates in the sense of Definition 6 and the proof reveals the reason: The arbitrary closeness of the matrices $\kappa(X)$, $\kappa(U)$ cannot be guaranteed because any κ must be discontinuous (see also Lemma 2 in Appendix C). Consider the κ example that makes the first nonzero matrix element real positive. This κ is discontinuous. At the discontinuity is, for example, the Pauli- X operator

TABLE II. If-clause achievability by algorithms that use a different type of measurement.

Measurement with	if clause	Random if clause
One success outcome	✗	✗ (Appendix F)
Many success outcomes	✗ (Appendix F)	✓ (Appendixes E,F)

$U = \kappa(U) = \sigma_x$, because its first matrix element is zero, $\langle 0|\sigma_x|0\rangle = 0$. We can choose an estimate X with small but negative $\langle 0|X|0\rangle$, so that $\kappa(X) \approx -\sigma_x$. This maximizes the operator norm difference $\|\kappa(X) - \sigma_x\|_{\text{op}} \approx 2$. Thus, near the discontinuity, $\|X(\cdot)X^\dagger - U(\cdot)U^\dagger\|_{\diamond}$ vanishes, but $\|\kappa(X) - \kappa(U)\|_{\text{op}}$ is large: the discontinuity in κ amplifies the error arbitrarily.

We suggest modifying Definition 6 to use multiple κ_j , each continuous on some subset of $U(d)$ and applied only if the sampled X lies deep inside this subset. For example, the following functions, indexed by $j \in \{0, 1, \dots, d-1\}$, split $U(d)$ to d such (overlapping) subsets:

$$\kappa_j(U) := \begin{cases} \frac{\langle j|U|0\rangle}{|\langle j|U|0\rangle|} U, & \text{if } \langle j|U|0\rangle \neq 0 \\ \kappa_{j+1 \bmod d}(U), & \text{otherwise.} \end{cases} \quad (15)$$

Having sampled X choose κ_r , where $r = r(X) = \min\{j : |\langle j|X|0\rangle| \geq 1/\sqrt{d}\}$. The chosen function is continuous around the sampled X , circumventing the impossibility. Thus, given some set $\{\kappa_r\}_r$ and a rule $r = r(X)$ we define the following.

Definition 7 (Process tomography yielding a unitary matrix, revised). Process tomography of a unitary oracle is an algorithm making N queries to $U \in U(d)$ and then measuring, obtaining the outcome $X \in \Omega_N \subset U(d)$ with probability $p_{N,U}(X)$. The output distribution $p_{N,U}$ is such that for all $U \in U(d)$,

$$\Pr_{X \sim p_{N,U}(\cdot)}[\|\kappa_r(X) - \kappa_r(U)\|_{\text{op}} \leq \epsilon_N] \geq 1 - \delta_N,$$

where $\lim_{N \rightarrow \infty} \epsilon_N = 0$, $\lim_{N \rightarrow \infty} \delta_N = 0$, and $r = r(X)$.

Appendix E argues that any standard process tomography [65] combined with (15) satisfies this definition. The dependence of κ_r on the sampled X yields randomness beyond the estimation error: There are oracles U for which the sampling yields either an approximation of $\kappa_0(U)$ or an approximation of $\kappa_1(U)$, both with a non-negligible probability. [Consider the Fourier transform oracle $\langle j|U|k\rangle = e^{i\frac{2\pi}{d}jk}/\sqrt{d}$ and κ_r , $r(X)$ of example (15).] Having obtained a classical estimate of $\kappa_r(U)$, we build a circuit close to $|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes \kappa_r(U) = c_{\phi_r}(U)$, but the value r changes the circuit’s operator (and the superoperator) beyond the estimation error. The index r is known and random; calculated from the measurement outcome obtained along implementing $c_{\phi_r}(U)$, instead of the if clause, process tomography yields the *random if clause* (see Appendix F). A many-outcome measurement is necessary for randomness. Thus, on the random if clause, quantum circuits using a many-outcome measurement are more powerful than those using a binary success or fail measurement (Table II).

We can make the tomography-based strategy fully quantum, deferring the r measurement to an additional ancilla register. Before the measurement, this implements a

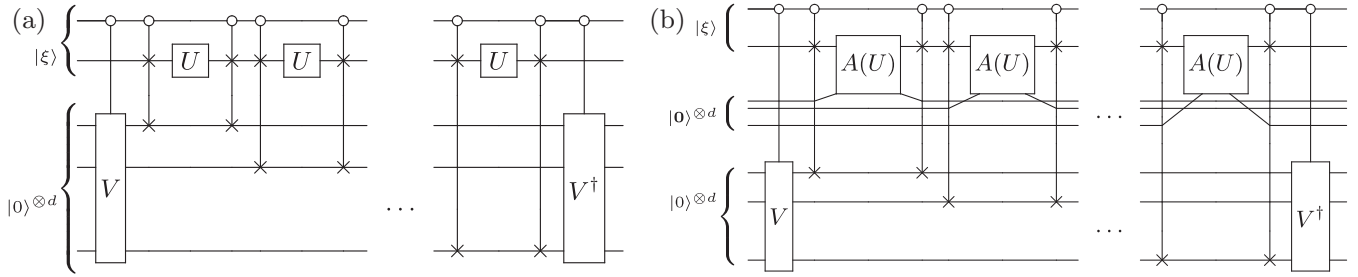


FIG. 6. (a) The unitary algorithm due to [54] makes d queries to U and implements $c_\phi^d(U)$ to its first two registers. (b) The queries to U are replaced by subroutines $A(U)$ that possibly act on additional ancillae. $A(U)$ is assumed to approximately apply $t(U) = U^{\frac{1}{d}}$ to its non-ancilla Hilbert space.

superoperator close to

$$\rho \mapsto \sum_r p_{rU} [c_{\phi_r}(U) \otimes |r\rangle] \rho [c_{\phi_r}(U)^\dagger \otimes \langle r|], \quad (16)$$

corresponding to the *entangled if clause* (see Appendix F).

After the deferring, a d -outcome measurement suffices to achieve the random if clause because r in $c_{\phi_r}(U) = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes e^{i\phi_r(U)}U$ takes d values. This measurement is less complex than the full process tomography.⁴ Since the $|\Omega_N|$ -outcome tomography measurement turns out to be unnecessary, one may ask whether the d -outcome measurement is optimal or can be simplified further. Even more pressing is the question of the query complexity of the random if clause. The brute-force tomography approach is likely far from optimal.

IV. APPLICATIONS TO OTHER TASKS

Here we use our method to study other tasks, namely, the neutralization, fractional power, transpose, and inverse task. For each task we first review some preceding work.

A. Review of the c_ϕ^d algorithm via the neutralization

In this section we review the unitary algorithm of Dong, Nakayama, Soeda, and Murao [54] that exactly implements $c_\phi^d(U)$. Suppose we try to implement $c_\phi^m(U)$ taking the following approach: (i) use the controlled-swap gates as in Fig. 1(a); if the control is $|1\rangle$ the oracle U is queried in sequence m times, if the control is $|0\rangle$ the m queries are “moved” to an ancilla register. (ii) If the control is $|0\rangle$ apply to the ancilla a subroutine that makes the “moved” queries have no effect other than a multiplication by a scalar. In the words of Ref. [54], the subroutine of step (ii) *neutralizes* the moved queries. Fitting the neutralization into our framework we define the following:

Definition 8 (ϵ -approximately neutralizing a query sequence). Given a query sequence \mathbf{s} , a postselection oracle algorithm $A_\emptyset : U(d) \rightarrow L(\mathcal{H})$ ϵ -approximately neutralizes \mathbf{s} ,

⁴With the full process tomography, one can also implement $c_{\phi_X}(U)$ with $X \in \Omega_N$ because the outcomes X yield $|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes X$ which is close to $c_{\phi_X}(U) = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes e^{i\phi_X(U)}U$ via expression (12) and by choosing $|v\rangle$ to be an eigenvector of $U^\dagger X$ so that $\langle v|U^\dagger X|v\rangle =: e^{i\phi_X(U)}$.

if its query sequence is \mathbf{s} and still it leaves an all-zero input (almost) untouched:

$$\left\| \frac{A_\emptyset(U)|\mathbf{0}\rangle\langle\mathbf{0}|A_\emptyset(U)^\dagger}{\text{tr}[A_\emptyset(U)|\mathbf{0}\rangle\langle\mathbf{0}|A_\emptyset(U)^\dagger]} - |\mathbf{0}\rangle\langle\mathbf{0}| \right\|_{\text{tr}} \leq \epsilon. \quad (17)$$

Definition 8 uses the algorithm’s operator instead of the superoperator in order to forbid the (partial) trace; applying the oracles to an ancilla that is later traced out is not considered a neutralization.

Observing that the totally antisymmetric state is an eigenstate of $U^{\otimes d}$, Ref. [54] presented an algorithm *exactly neutralizing* ($\epsilon = 0$) the sequence $\mathbf{s} = (id)^d$ of *parallel* queries. This postselection oracle algorithm $A_\emptyset^{\parallel d}(U)$ is unitary. Reference [54] used $A_\emptyset^{\parallel d}(U)$ to build the algorithm of Fig. 6(a), the (unitary) postselection oracle algorithm that exactly achieves $(c_\phi^d, \{id\})$:

Construction 1 (due to [54]). For the correctly chosen fixed unitary V , the unitary algorithm of Fig. 6(a) makes d queries to U and implements the operator $A(U)$ such that $A(U)(I_{\mathcal{H}} \otimes |\mathbf{0}\rangle\langle\mathbf{0}|_{\mathcal{K}})$ equals

$$[|\mathbf{0}\rangle\langle\mathbf{0}| \otimes I + \det(U)^{-1}|1\rangle\langle 1| \otimes U^d] \otimes \det(U)|\mathbf{0}\rangle\langle\mathbf{0}|_{\mathcal{K}}. \quad (18)$$

Proof. First note that we can write the determinant as

$$\begin{aligned} \det(U) &= \sum_{\pi \in S_d} \text{sgn}(\pi) \prod_{i=1}^d \langle i|U|\pi(i)\rangle \\ &= \frac{1}{d!} \sum_{\pi, \tau \in S_d} \text{sgn}(\tau) \text{sgn}(\pi) \prod_{i=1}^d \langle \tau(i)|U|\pi(i)\rangle \\ &= \langle \chi_d | U^{\otimes d} | \chi_d \rangle, \end{aligned} \quad (19)$$

where $|\chi_d\rangle = \frac{1}{\sqrt{d!}} \sum_{\pi \in S_d} \text{sgn}(\pi) |\pi(1)\rangle \otimes \cdots \otimes |\pi(d)\rangle \in \mathcal{K}$ is the totally antisymmetric state. For a proof of Eq. (19) see Appendix G. Note that we have indexed the basis vectors starting with $i = 1$. Switching to the usual labeling that starts with zero, observe that $|\chi_d\rangle$ is normalized and therefore there exists a unitary $V \in L(\mathcal{K})$ such that $V|\mathbf{0}\rangle = |\chi_d\rangle$. Define the unitary oracle algorithm $A_\emptyset^{\parallel d}(U) := V^\dagger U^{\otimes d} V$. Since $\langle \chi_d | U^{\otimes d} | \chi_d \rangle = 1$, the unitary $A_\emptyset^{\parallel d}(U)$ on input $|\mathbf{0}\rangle$ outputs $|\mathbf{0}\rangle$ with probability 1 for all $U \in U(d)$. Therefore, $A_\emptyset^{\parallel d}(U)$ exactly neutralizes its queries, $A_\emptyset^{\parallel d}(U)|\mathbf{0}\rangle\langle\mathbf{0}| = \det(U)|\mathbf{0}\rangle\langle\mathbf{0}|$. Observe that the algorithm of Fig. 6(a) implements $|0\rangle\langle 0| \otimes I \otimes A_\emptyset^{\parallel d}(U)_{\mathcal{K}} + |1\rangle\langle 1| \otimes U^d \otimes I_{\mathcal{K}}$, from which (18) follows. \blacksquare

B. Neutralizable query sequences

Complementary to the neutralizing algorithm $A_{\emptyset}^{\parallel d}(U)$ above, here we prove the impossibility of neutralizing the remaining numbers of parallel, as well as sequential, id queries. At first glance, it seems that one could always add a query and then make it have no effect in the algorithm. This is less obvious when tracing out subsystems is disallowed, as in the neutralization (Definition 8). Indeed, for most query numbers, the neutralization is impossible. Conceptually, this no-neutralization of oracle queries is analogous to the no-deleting of states [14].

Theorem 3. Let $d \in \mathbb{N}$ be the dimension of the oracle U , let $s = (id)^m$ and $\epsilon \in [0, 1)$. If there exists a postselection oracle algorithm ϵ -approximately neutralizing s , then $d|m$.

Proof. Suppose $(A_{\emptyset}, \{id\})$ is a postselection oracle algorithm ϵ -approximately neutralizing s . By Eq. (3), $A_{\emptyset} : U(d) \rightarrow L(\mathcal{H})$ is continuous and m -homogeneous and so is

$$f_{A_{\emptyset}}(U) := \frac{\langle \mathbf{0} | A_{\emptyset}(U) | \mathbf{0} \rangle}{\sqrt{\text{tr} [A_{\emptyset}(U) | \mathbf{0} \rangle \langle \mathbf{0} | A_{\emptyset}(U)^{\dagger}]}}$$

Since taking $X \mapsto \langle \mathbf{0} | X | \mathbf{0} \rangle$ can only contract trace norm, inequality (17) implies that

$$||f_{A_{\emptyset}}(U)|^2 - 1| \leq \epsilon < 1$$

so $|f_{A_{\emptyset}}(U)| \neq 0$ holds for all $U \in U(d)$ and we can define $\widehat{f}_{A_{\emptyset}}(U) := f_{A_{\emptyset}}(U)/|f_{A_{\emptyset}}(U)|$ which maps from $U(d)$ to the circle and is continuous m -homogeneous. Then by Lemma 1, m is a multiple of d . ■

Theorem 3 does not distinguish between parallel and sequential queries: it applies to both. More generally, all of the above applies also to process-matrix algorithms (a model already without the parallel or sequential distinction).

C. Fractional power impossibility

Here we are interested in the $(1/d)$ th power task $(U \mapsto U^{\frac{1}{d}}, \{id, inv\})$. It is listed in Table I. For this task, Sheridan, Maslov, and Mosca [40] found an average-case algorithm, an algorithm that works for most, but not all, $U \in U(d)$. Given any fraction $q \in \mathbb{Q}$, the algorithm makes id, inv queries (after the adjustment in Appendix A of [40] that removes the controlled- U queries) and implements an operator that for most $U \in U(d)$ is close to $U^q \otimes U^{-q} \otimes I_{\mathcal{K}_2}$. This can be arbitrarily close at the cost of increasing the dimension of the ancilla Hilbert space \mathcal{K}_2 . In [40] $U \mapsto U^q$ corresponds to a specific fractional-power function. Many exist because each unitary has many possible roots. For example, for any $j, k \in \{0, 1, \dots, m-1\}$ a possible m th root of

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \text{ is } \begin{pmatrix} e^{i\frac{2\pi j}{m}} & 0 \\ 0 & e^{i\frac{\pi+2\pi k}{m}} \end{pmatrix}.$$

Here we show that, unlike in the average case, in the worst case we cannot implement $U^{\frac{1}{d}}$ to within an arbitrarily small error ϵ . For the specific $(1/d)$ th power function chosen in [40], the impossibility was already proven in [40, Lemma 1]. However, the following corollary of Theorem 1 holds for any $(1/d)$ th power function.

Corollary 1. Let $t : U(d) \rightarrow U(d)$ be any function such that $t(U)^d = U$. No postselection oracle algorithm or

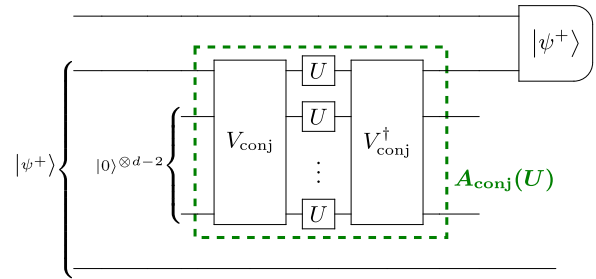


FIG. 7. After postselecting the first two registers on outcome $|\psi^+\rangle$, this algorithm of [38,39] implements U^\dagger . Its subroutine in the dashed box, the unitary algorithm of [37], implements the complex conjugate \bar{U} from $d - 1$ parallel queries to U .

process-matrix algorithm can ϵ -approximately achieve the $(1/d)$ th power task $(t, \{id, inv\})$ for $\epsilon < 1/2d$, no matter how many queries it makes.

The impossibility of errors smaller than $1/\exp(n)$ for n -qubit unitaries is discouraging because the “inverse exponential scaling” here is qualitatively different from how errors scale in useful algorithms. In a useful algorithm the error ϵ scales inverse exponentially or even inverse polynomially with respect to the complexity of the algorithm. In other words, the complexity scales as $\log_2(1/\epsilon)$ or $\text{poly}(1/\epsilon)$ [41,67–69]. Here, instead, we found ϵ values for which arbitrarily complex algorithms fail.

Another way to understand the strength of the result is to consider only two-dimensional unitary oracles, which intuitively should be the simplest to work with. According to the result, the square root of a two-dimensional unitary is impossible to implement within $\epsilon < \frac{1}{4}$ by quantum circuits of any complexity.

The intuition for a complexity-independent lower bound on ϵ is simple: we are trying to approximate a discontinuous function t by some continuous algorithm.

Proof. Suppose $(A, \{id, inv\})$ ϵ -approximately achieves the $(1/d)$ th power task $(t, \{id, inv\})$. Use $A(U)$ instead of each query to U in the $(c_\phi^d, \{id, inv\})$ algorithm of Ref. [54] as in Fig. 6. The resulting algorithm in the exact ($\epsilon = 0$) case $A(U)(I \otimes |\mathbf{0}\rangle\langle\mathbf{0}|) = t(U) \otimes |g(U)\rangle\langle\mathbf{0}|$ implements

$$\begin{aligned} &(|\mathbf{0}\rangle\langle\mathbf{0}| \otimes I + \det[t(U)]^{-1} |1\rangle\langle 1| \otimes U) \otimes [|g(U)\rangle\langle\mathbf{0}|]^{\otimes d} \\ &\otimes \det[t(U)] |\mathbf{0}\rangle\langle\mathbf{0}|, \end{aligned}$$

which corresponds to the if clause. For A with error ϵ , the composed algorithm in Fig. 6(b) has error $d\epsilon$ because it uses the $A(U)$ subroutine d times and because postselection diamond distance satisfies weak subadditivity (Theorem 2.1 in [63]). By Theorem 1 we must have $d\epsilon \geq \frac{1}{2}$ implying that $\epsilon \geq 1/2d$. ■

D. Review of the complex-conjugation algorithm

Complex conjugation is used as a subroutine in the inverse algorithm discussed in the subsequent sections. Here we review the complex-conjugation algorithm of Miyazaki, Soeda, and Murao [37], giving an alternative proof of its correctness.

Construction 2 (due to [37]). For a correctly chosen fixed unitary V_{conj} , the unitary algorithm in the dashed box in Fig. 7

implements $A_{\text{conj}}(U)$ such that

$$A_{\text{conj}}(U)(I \otimes |0\rangle\langle 0|^{\otimes d-2}) = \bar{U} \otimes \det(U)|0\rangle\langle 0|^{\otimes d-2} \quad (20)$$

while making $d - 1$ queries to U .

Therefore, the (unitary) postselection oracle algorithm exactly achieves the complex-conjugation task ($U \mapsto \bar{U}, \{id\}$).

Proof. Cramer's rule for matrix inversion reads as $U^{-1} = \det(U)^{-1}C(U)^T$, where $C(U)$ is the cofactor matrix of U . We define it presently, but first note that from Cramer's rule $C(U) = \det(U)\bar{U}$. It remains to show that the algorithm implements $C(U)$ to its first register corresponding to a d -dimensional Hilbert space \mathcal{H} . Cofactor matrix of U is the matrix of minors of U , $C(U)_{i,j} = (-1)^{i+j} \det(U_{i,j})$, or

$$C(U) = \sum_{i,j=1}^d (-1)^{i+j} \det(U_{i,j}) |i\rangle\langle j|, \quad (21)$$

where the labeling of the computational basis vectors starts with $i = 1$, and where $\det(U_{i,j})$ is the (i, j) th minor, defined as the determinant of the matrix U with the i th row and j th column deleted. Formally, we write

$$\det(U_{i,j}) = \frac{(-1)^{i+j}}{(d-1)!} \sum_{\substack{\tau, \pi \in S_d \\ \tau(1)=i \\ \pi(1)=j}} \text{sgn}(\tau) \text{sgn}(\pi) \prod_{k=2}^d \langle \tau(k) | U | \pi(k) \rangle. \quad (22)$$

See Appendix G for the proof. Substituting (22) into (21) we get $C(U) = E^\dagger U^{\otimes d-1} E$ with $E = \sum_{j=1}^d |v_j\rangle\langle j|$ and

$$|v_j\rangle = \frac{1}{\sqrt{(d-1)!}} \sum_{\substack{\pi \in S_d \\ \pi(1)=j}} \text{sgn}(\pi) |\pi(2)\rangle \otimes \cdots \otimes |\pi(d)\rangle.$$

$E : \mathcal{H} \rightarrow \mathcal{H}^{\otimes d-1}$ is an isometry because the vectors $\{|v_j\rangle\}_{j \in [d]}$ in $\mathcal{H}^{\otimes d-1}$ are orthonormal. We switch back to the basis labeling that starts with zero. Since $\{|v_j\rangle\}_{j \in \{0,1,\dots,d-1\}}$ are orthonormal, there exists a (nonunique) unitary V_{conj} such that $V_{\text{conj}}(|j\rangle \otimes |0\rangle^{\otimes d-2}) = |v_j\rangle = E|j\rangle$, which implies that

$$(I \otimes \langle 0|^{\otimes d-2}) V_{\text{conj}}^\dagger U^{\otimes d-1} V_{\text{conj}} (I \otimes |0\rangle^{\otimes d-2}) = C(U) \quad (23)$$

is a unitary. Therefore, the dashed box $A_{\text{conj}}(U) = V_{\text{conj}}^\dagger U^{\otimes d-1} V_{\text{conj}}$ in Fig. 7 on input $|\xi\rangle \otimes |0\rangle^{\otimes d-2}$ outputs the all-zero state on the ancilla with probability

$$\begin{aligned} p &= \|(I \otimes \langle 0|^{\otimes d-2}) A_{\text{conj}}(U) (|\xi\rangle \otimes |0\rangle^{\otimes d-2})\|^2 \\ &= \|C(U)|\xi\rangle\|^2 = 1. \end{aligned}$$

In other words, projecting the ancilla output onto the all-zero state has no effect, i.e., $A_{\text{conj}}(U)(I \otimes |0\rangle\langle 0|^{\otimes d-2})$ equals

$$(I \otimes |0\rangle\langle 0|^{\otimes d-2}) A_{\text{conj}}(U) (I \otimes |0\rangle\langle 0|^{\otimes d-2}),$$

which by Eq. (23) equals to $C(U) \otimes |0\rangle\langle 0|^{\otimes d-2}$. ■

E. Review of the simplest transpose and inverse algorithms

We listed the transpose and inverse task in Table I. Quintino, Dong, Shimbo, Soeda, and Murao [38,39] presented a family of algorithms for each task. Here we review only the simplest algorithm of each family: the one with the fewest queries and the lowest probability of success.

Construction 3 (due to [38,39]). Figure 7 with the dashed box replaced by U is a single-query algorithm implementing U^T with the probability of success $1/d^2$.

Prepending to Construction 3 a unitary that prepares the input $|\psi^+\rangle = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |ii\rangle$ (the generalized Bell state) gives a postselection oracle algorithm exactly achieving the task ($U \mapsto U^T, \{id\}$).

Proof. Note that if the dashed box is replaced by the identity operator I , Fig. 7 corresponds to the generalized teleportation, which succeeds with probability $\frac{1}{d^2}$. In other words, $(\langle \psi^+ | \otimes I)(I \otimes |\psi^+\rangle) = \frac{1}{d} I$. Next, for any complex $d \times d$ matrix M , $(M \otimes I)|\psi^+\rangle = I \otimes M^T |\psi^+\rangle$, which can be verified by writing M in terms of its matrix elements. Thus, the algorithm of Fig. 7 with the dashed box replaced by U implements

$$\begin{aligned} & (\langle \psi^+ | \otimes I)(I \otimes U \otimes I)(I \otimes |\psi^+\rangle) \\ &= U^T (\langle \psi^+ | \otimes I)(I \otimes |\psi^+\rangle) = \frac{1}{d} U^T. \end{aligned}$$

■

Composing the transpose algorithm with the unitary algorithm for the complex conjugate [37] gives a postselection oracle algorithm exactly achieving the task ($U \mapsto U^\dagger, \{id\}$).

Construction 4 (due to [38,39]). The algorithm of Fig. 7 makes $d - 1$ queries to U and implements U^\dagger with the probability of success $1/d^2$.

Both algorithms above have a useful feature: their effect on the ancillae can easily be uncomputed. Let V_{prep} be a unitary that prepares $|\psi^+\rangle$ from the all-zero state. Substituting $|\psi^+\rangle = V_{\text{prep}}|0\rangle^{\otimes 2}$ into the algorithm of Fig. 7 for both occurrences of $|\psi^+\rangle$ gives algorithms with ancillae initialized to zeros and with $\Pi_{\text{succ}} = |0\rangle\langle 0|^{\otimes 2}$. Adding after the dashed box a gate that swaps the first and the last register results in transpose and inverse algorithms A_T and A_\dagger such that

$$\begin{aligned} A_T(U)(I \otimes |0\rangle\langle 0|^{\otimes 2}) &= U^T \otimes \frac{1}{d} |0\rangle\langle 0|^{\otimes 2}, \\ A_\dagger(U)(I \otimes |0\rangle\langle 0|^{\otimes d}) &= U^\dagger \otimes \frac{1}{d} |0\rangle\langle 0|^{\otimes d}. \end{aligned}$$

We call such algorithms *clean* (Definition 9). Note the complex-conjugation algorithm, Eq. (20), is also clean. *Cleanness* can be achieved whenever the ancilla output is independent of the oracle U , by adding the appropriate fixed unitary gate as we did here.

F. Query complexity of the clean transpose and inverse

We reviewed the simplest transpose and inverse algorithms. References [38,39] also presented other algorithms, with higher probabilities of success, that require more queries. Are they all clean? We answer this negatively in a much more general way: In this section we show that *clean* transpose and inversion algorithms with certain numbers of queries are impossible. First we define clean algorithms: Apart from achieving their task [inequality (9)], they satisfy an additional condition regarding their effect on their ancillae: if the ancilla input is the all-zero state, so should be the ancilla output. In the definition we merge the task and the ancilla conditions into one inequality.

Definition 9. (Clean postselection oracle algorithm). A postselection oracle algorithm $A : U(d) \rightarrow L(\mathcal{H} \otimes \mathcal{K})$ that ϵ -approximately achieves the task (t, Σ) is *clean* if for all \mathcal{H}' and $\rho \in \mathcal{D}(\mathcal{H}' \otimes \mathcal{H})$, ϵ upper bounds also

$$\left\| \frac{\widetilde{A(U)}(\rho \otimes |\mathbf{0}\rangle\langle\mathbf{0}|)\widetilde{A(U)}^\dagger}{\text{tr}[\widetilde{A(U)}(\rho \otimes |\mathbf{0}\rangle\langle\mathbf{0}|)\widetilde{A(U)}^\dagger]} - t(\widetilde{U})\rho t(\widetilde{U})^\dagger \otimes |\mathbf{0}\rangle\langle\mathbf{0}| \right\|_{\text{tr}},$$

where the wide tilde denotes the extension $\widetilde{X} := I_{\mathcal{H}'} \otimes X$.

In other words, the ancilla output is independent of the unknown oracle. *Clean* algorithms enable ancillae recycling and preserve coherences if used inside an interferometer [Fig. 1(b)]. The following impossibility of certain clean algorithms is a corollary of Theorem 3.

Corollary 2. Let $\epsilon \in [0, 1)$. If an N -query clean postselection oracle algorithm ϵ -approximately achieves the task

- (1) $(\text{inv}, \{id\})$, then $N \equiv -1 \pmod{d}$.
- (2) $(U \mapsto U^T, \{id\})$, then $N \equiv 1 \pmod{d}$.

Proof of 1. Such an algorithm $(A, \{id\})$ satisfies

$$\left\| \frac{A(U)|\mathbf{0}\rangle\langle\mathbf{0}|A(U)^\dagger}{\text{tr}[A(U)|\mathbf{0}\rangle\langle\mathbf{0}|A(U)^\dagger]} - (U^\dagger \otimes I)|\mathbf{0}\rangle\langle\mathbf{0}|(U \otimes I) \right\|_{\text{tr}} \leq \epsilon \quad (24)$$

by Definition 9 with \mathcal{H}' trivial and $\rho = |\mathbf{0}\rangle\langle\mathbf{0}|$. If it makes N queries, the new algorithm $(U \mapsto (U \otimes I)A(U), \{id\})$ makes $N + 1$ queries. Represent by $\|X\|_{\text{tr}}$ the left-hand side of (24). Since $\|(U \otimes I)X(U^\dagger \otimes I)\|_{\text{tr}} = \|X\|_{\text{tr}}$, the algorithm ϵ -approximately neutralizes the query sequence $\mathbf{s} = (id)^{N+1}$. Theorem 3 then guarantees that d divides $N + 1$.

Proof of 2. Suppose some clean N -query algorithm ϵ -approximately achieves the transpose task. Replace each query to U by $A_{\text{conj}}(U)$ of Eq. (20), i.e., by the clean, unitary, $(d - 1)$ -query algorithm of [37] that exactly implements \overline{U} . The algorithm makes $N(d - 1)$ queries to U , ϵ -approximately implements U^\dagger , and is clean. Then by the preceding proof d divides $N(d - 1) + 1$, so $N \equiv 1 \pmod{d}$. ■

References [38,39] found for every $k \in \mathbb{N}$ a k -query transpose algorithm and a $k(d - 1)$ -query inverse algorithm. The success probability of these grows with k . We have shown that their algorithms can be clean only if $k \equiv 1 \pmod{d}$.

V. DISCUSSION AND OPEN PROBLEMS

In this work we introduce a unified framework for algorithms with unitary oracles $U \in U(d)$, and develop a topological method to prove their limitations. The method yields new limitations for algorithms implementing the if clause $c_\phi(U)$, neutralization, $(1/d)$ th power $U^{1/d}$, transpose U^T , and inverse U^\dagger . Our method adds to the very few previously known proof methods for limiting query complexity.

Similarly to the preceding methods [8–11], our method is worst case: it expects algorithms to work on *every* input. The distinction is important. While a quantum circuit implementing $U^{1/d}$ for most $U \in U(d)$ exists [40], one working for all $U \in U(d)$ is impossible. Similarly, the worst-case if clause is impossible (Theorem 1), and the (exponential) average-case if clause is implemented by Fig. 6(b) after substituting the $U^{1/d}$ heuristic scheme [40] for $A(U)$. Whether a more efficient average-case algorithm exists remains open. Between the finite average-case and the infinite worst-case complexity, other

complexity notions fit, and could be explored. One notion is that of Levin's average-case complexity [71], equivalent to the notion of *errorless* heuristic scheme [42]: the scheme gives a warning (e.g., the \perp symbol) if its output breaks the requested specifications (in our case, if the implemented operator exceeds the requested diamond distance from the task operator). Errorless heuristic schemes remain a possibility for the if clause and the fractional power. Another notion is that of smoothed complexity [72], which considers the worst input, but after perturbing it slightly. Although previously defined for errorless heuristic schemes [72,73], smoothed complexity can also be defined for heuristic schemes with errors (as Ref. [74] did specifically for the perceptron algorithm). Such smoothed analysis could resolve the transition between the two problems' average-case and worst-case complexities. Focusing here on worst-case query complexity, we leave all the above questions open.

The query complexity bounds proven here reveal surprising differences between oracle computational models. First, we show how quantum circuits differ from linear optics. By Theorem 1, the if clause is an oracle problem on which quantum circuits perform much worse than linear optics. The query complexities of the if clause in these models are not polynomially, and not even exponentially equivalent. While the linear-optics complexity of the if clause was known to equal one, its quantum-circuit and process-matrix complexities are worked out here to be infinite! Linear optics avoid our impossibility because the direct sum $U \oplus I$ lacks homogeneity, a property required by our method. According to Ref. [36], the direct sum appears due to the gate's localization to the upper path modes in Fig. 1(b). However, instead of a direct sum, a second-quantization view yields a *tensor product* of modes. We suggest a conceptually different reason for the direct sum: the *linearity* of linear optics. Figure 1(b) implicitly assumes that U is a linear (passive) gate: the vacuum $|\Omega\rangle$ is its $+1$ eigenstate. If this hidden assumption is made explicit, the oracle is not completely unknown and should be represented by $U' = 1 \oplus U$ as in Fig. 1(a). Thus, while restricting to linear optics yields a weaker computational model, in the presence of oracles (Appendix H), it also imposes additional knowledge about the oracles. Surprisingly, overall this gives a remarkable advantage on the if clause.

The demonstrated advantage of linear optics motivates studying restricted computational models, especially if the restriction makes the model more implementable in the near term. For example, aside from linear optics, we could restrict optics to Gaussian optics or to various levels of non-Gaussianity [75], applying the same restriction to the oracle. This could capture varying amounts of device independence in experiments, adversarial power in cryptography, or distrust in communication protocols. Restricting the computational model, and with it the oracle, may change the query complexity of problems other than the if clause. Studying these changes under restrictions that are relevant in near-term computation is an interesting future direction.

Process matrices compare equally unfavorably to linear optics on the if clause. While the linear restriction improves efficiency, relaxed causality does not. This affects the debate of whether the linear optics experiments [47–49] implement the process-matrix solution to the quantum switch: Any

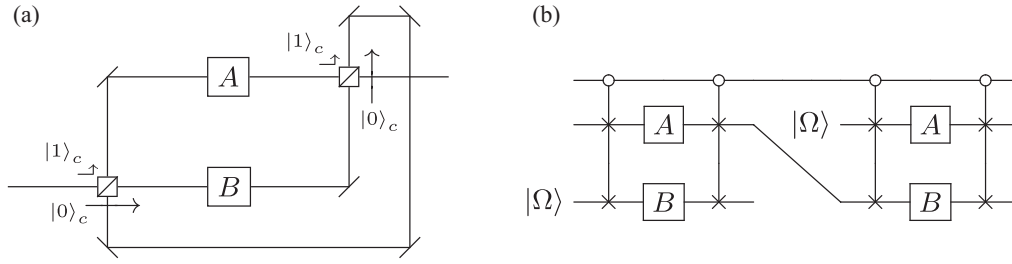


FIG. 8. (a) In this quantum-switch implementation [70], the photon passes through the interferometer twice. In the first pass, A acts on one photon if the polarization is $|1\rangle_c$ and on vacuum otherwise. In the second pass, A acts on one photon if the polarization is $|0\rangle_c$ and on vacuum otherwise. A can act nontrivially on vacuum. (b) Analogously to Fig. 1, the equivalent circuit captures this possibility.

claimed linear-optics efficiency could stem from the implicit restriction on the oracles, and not from relaxed causality. To discount the restriction effect, we suggest to check the query complexity in the experiments by imagining that the oracles are fully general and act nontrivially on vacuum. For example, for the specific linear-optics quantum switch in Fig. 8(a), this leads to an equivalent circuit in Fig. 8(b) whose query complexity does not match the process-matrix quantum switch. This approach to query complexity has a specific, causality-related purpose. It might not work for other purposes. Indeed, in computational models that allow changing or superposing particle numbers, calculating complexity is an interesting and difficult open problem.

The last query complexity gap is between measurements. Understanding why process tomography fails at the if clause [it fails at the $(1/d)$ th power for a similar reason] leads us to define modified tasks which are achievable: the *entangled* and *random if clause*. If simpler implementations exist, the entangled if clause could serve as a subroutine in larger algorithms. The random if clause reveals the complexity gap between measurements: its query complexity is infinite in the one-success-outcome model, but finite (at most exponential) in a model with $\exp(n)$ success outcomes (Table II). Are there algorithms requiring fewer queries and fewer

outcomes? An optimal algorithm will exactly quantify the advantage of many outcomes; finding one remains an open question.

ACKNOWLEDGMENTS

Z.G. and Y.T. thank D. Aharonov for the supervision and support. All authors thank A. Kamber for sketching the more elementary proof of Lemma 1, M. Araújo for enlightening discussions of the tomography issue, A. Dotan, I. Leigh, G. Gasbarri, and M. Skotiniotis for useful feedback, and an anonymous referee from the QIP conference for suggesting an alternative approach via $SU(d)$ covering spaces (Appendix C is our take on this approach). This work was supported by the Simons Foundation (Grant No. 385590), by the Israel Science Foundation (Grants No. 2137/19 and No. 1721/17), by the European Commission QUANTERA grant ExTRaQT (Spanish MICINN Project No. PCI2022-132965), by the Ministry for Digital Transformation and of Civil Service of the Spanish Government through the QUANTUM ENIA project call, Quantum Spain project, by the European Union through the Recovery, Transformation and Resilience Plan NextGenerationEU within the framework of the Digital Spain 2026 Agenda, by Spanish Agencia Estatal de Investigación (Project No. PID2022-141283NB-I00).

APPENDIX A: IF-CLAUSE IMPOSSIBILITY (SPECIAL CASE) VIA THE BORSUK-ULAM THEOREM

Here we prove a special case of Theorem 1 from the famous result of algebraic topology: the Borsuk-Ulam theorem [64] (see Fig. 4).

Borsuk-Ulam theorem. If $f : S^n \rightarrow \mathbb{R}^n$ is a continuous function, then there exists $x \in S^n$ such that $f(x) = f(-x)$.

Proof of Theorem 1 (d even, m odd). Let d be an even and m an odd integer. The first part is the same as in the general proof of Theorem 1. Assume towards contradiction that there exists a continuous 0-homogeneous function $\mathcal{A} : U(d) \rightarrow S_+(\mathbb{C}^{2d})$ such that

$$\left\| \frac{\mathcal{A}(U)(\rho)}{\text{tr}[\mathcal{A}(U)(\rho)]} - (|0\rangle\langle 0| \otimes I + e^{i\phi(U)}|1\rangle\langle 1| \otimes U^m) \rho (|0\rangle\langle 0| \otimes I + e^{i\phi(U)}|1\rangle\langle 1| \otimes U^m)^\dagger \right\| < \frac{1}{2} \quad (\text{A1})$$

holds for all $U \in U(d)$, all $\rho \in \mathcal{D}(\mathbb{C}^{2d})$, and for some real function ϕ . Let $\rho_+ = |+\rangle\langle +| \otimes |0\rangle\langle 0| \in \mathbb{C}^2 \otimes \mathbb{C}^d$, and define $f : U(d) \rightarrow \mathbb{C}$ by

$$f(U) := \text{tr} \left[\frac{\mathcal{A}(U)(\rho_+)}{\text{tr}[\mathcal{A}(U)(\rho_+)]} (|1\rangle\langle 0|_c \otimes U^m) \right]. \quad (\text{A2})$$

This is well defined because \mathcal{A} maps into $S_+(\mathbb{C}^{2d})$ so the superoperator $\mathcal{A}(U)$ never vanishes the trace of its input. Observe that by its definition f is a continuous odd function. Substituting the inequality (A1) into (A2), we get that $|f(U) - \frac{1}{2}e^{-i\phi(U)}| < \frac{1}{2}$ so f never maps to zero. We can define $\hat{f}(U) = f(U)/\sqrt{(f(U))f(U)}$, which is also continuous and odd, but maps into the circle S^1 .

Now, since d is even we can define the following continuous function $g : S^3 \rightarrow U(d)$ such that given a vector $\mathbf{x} = (x_1, x_2, x_3, x_4)^T$ on the 3-sphere $S^3 \subset \mathbb{R}^4$:

$$g(\mathbf{x}) = \begin{pmatrix} x_1 + ix_2 & & & & & & & -x_3 + ix_4 \\ & \ddots & & & & & & \ddots \\ & & x_1 + ix_2 & & -x_3 + ix_4 & & & \\ & & x_3 + ix_4 & & x_1 - ix_2 & & & \\ & & \ddots & & \ddots & & & \\ & & & & & & & \ddots \\ x_3 + ix_4 & & & & & & & x_1 - ix_2 \end{pmatrix},$$

where the dots denote that the elements are repeated across the diagonal and the anti-diagonal, and there are zeros everywhere else. Note that g is odd, $g(-\mathbf{x}) = -g(\mathbf{x})$. By the Borsuk-Ulam theorem, since $\hat{f} \circ g : S^3 \rightarrow \mathbb{C} \subset \mathbb{R}^3$ is continuous, there exists $\mathbf{x} \in S^3$ such that $\hat{f} \circ g(\mathbf{x}) = \hat{f} \circ g(-\mathbf{x}) = -\hat{f} \circ g(\mathbf{x})$, where the last equality holds because $\hat{f} \circ g$ is odd. So for this \mathbf{x} we have $\hat{f} \circ g(\mathbf{x}) = 0$. But this contradicts the fact that \hat{f} maps into S^1 . ■

APPENDIX B: EXACT IF-CLAUSE IMPOSSIBILITY: AN OPERATIONAL PROOF

Proof of Theorem 1 ($\epsilon = 0$). Assume a postselection oracle algorithm $(A, \{id, inv\})$ exactly achieves the task $(c_\phi^m, \{id, inv\})$. Since the query functions are continuous and ± 1 homogeneous, by Eq. (3) $A(U)$ is continuous and w -homogeneous for some $w \in \mathbb{Z}$. Equation 7 implies

$$A(U)(I \otimes |0\rangle\langle 0|) = (|0\rangle\langle 0| \otimes I + e^{i\phi(U)}|1\rangle\langle 1| \otimes U^m) \otimes |g(U)\rangle\langle 0|. \tag{B1}$$

Consider the circuit of Fig. 9. It outputs $B(U)|0\rangle = e^{-i\phi(U)} ||g(U)\rangle\langle 0||^2 |0\rangle$. Moreover, $B(U)$ is continuous and m -homogeneous. Then

$$f(U) := \frac{\langle 0 | B(U) | 0 \rangle}{\langle 0 | A(U)^\dagger A(U) | 0 \rangle}$$

is continuous, m -homogeneous, and by Eq. (4) well defined on all $U \in U(d)$. Observe that $f(U) = e^{-i\phi(U)}$. Lemma 1 applies to f , so d divides m . ■

The above can be extended to approximations, as long as the two sides of (B1) are close enough to guarantee $f(U) \neq 0$.

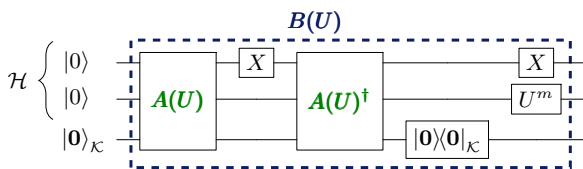


FIG. 9. Circuit $B : U(d) \rightarrow L(\mathcal{H} \otimes \mathcal{K})$ running A as a subroutine. The control qubit and the target qubit of \mathcal{H} are shown explicitly. X is a Pauli- X gate.

APPENDIX C: EXACT IF-CLAUSE IMPOSSIBILITY VIA COVERING SPACES

This proof uses the fact that $PU(d)$ is a d th cover of $SU(d)$.⁵ The covering space approach gives useful intuition, but, we did not find a way to extend it to the approximate case ($\epsilon > 0$); we discuss the difficulty at the end of this section.

Proof of Theorem 1 ($m = 1, \epsilon = 0$). Assume towards contradiction the existence of an algorithm (A, Σ_A) that ϵ -approximately achieves $(c_\phi, \{id, inv\})$ for $\epsilon = 0$. Observe from Eqs. (5) and (6) that we can define a function on unitary superoperators $\mathcal{A}_S(\mathcal{U}) := \mathcal{A}(U)$, where the unitary superoperator \mathcal{U} is defined as $\mathcal{U}(\rho) := U\rho U^\dagger$. Since the query alphabet $\Sigma_A \subseteq \{id, inv\}$ only contains continuous functions, $\mathcal{A}_S(\mathcal{U})$ is continuous. Next define

$$s(\mathcal{U}) := 2d \frac{\langle (|1\rangle \otimes I) \mathcal{A}_S(\mathcal{U}) (|+\rangle\langle +| \otimes \frac{I}{d}) (|0\rangle \otimes I) \rangle}{\text{tr}[\mathcal{A}_S(\mathcal{U}) (|+\rangle\langle +| \otimes \frac{I}{d})]} \tag{C1}$$

and observe that s is continuous and that applying Eq. (8) we get $s(\mathcal{U}) = e^{i\phi(U)}U$, which produces a contradiction with the lemma below. ■

We will show that there is no continuous function s (called *section*) that maps a unitary superoperator to some matching unitary.

Lemma 2. Let $\pi : U(d) \rightarrow \mathcal{S}_+(\mathbb{C}^d)$ map unitaries to the corresponding superoperators, i.e., $\pi : U \mapsto \mathcal{U}$, where $\mathcal{U}(\rho) = U\rho U^\dagger$. There is no continuous map $s : \text{Im}(\pi) \rightarrow U(d)$ such that $\pi \circ s = id_{\text{Im}(\pi)}$.

First we will characterize the space $\text{Im}(\pi)$. Defining the equivalence relation $U \sim e^{i\alpha}U$ for all $\alpha \in [0, 2\pi)$, we will show that $\text{Im}(\pi)$ is homeomorphic to the quotient space $U(d)/\sim = U(d)/Z[U(d)] =: PU(d)$ called the projective unitary space. In particular, π maps into the same element precisely those unitaries that differ by the global phase. To see this, consider the continuous map π and the quotient map $p : U(d) \rightarrow PU(d)$:

$$\begin{array}{ccc} U(d) & \xrightarrow{\pi} & \text{Im}(\pi) \\ p \downarrow & \nearrow h & \\ PU(d) & & \end{array}$$

and observe that whenever $p(U) = p(U')$ we have that $\pi(U) = \pi(U')$. By the universal property of quotient spaces

⁵Suggested by an anonymous referee of the QIP conference, whom we would like to thank in this way.

in topology there exists a continuous map $h : \text{PU}(d) \rightarrow \text{Im}(\pi)$, $h([U]) = \mathcal{U}$ such that $h \circ p = \pi$. Therefore, h is surjective (because π is). h is also injective: assume $h([U]) = \mathcal{U} = \mathcal{U}' = h([U'])$ and let $U \in [U]$ be some unitary from the equivalence class. Choose coordinates $k, l \in [d]$ such that the matrix element u_{lk} of U is nonzero. Observing that $\sum_j \mathcal{U}(|j\rangle\langle k|)|l\rangle\langle j| = \bar{u}_{lk}U$ and similarly for some $U' \in [U']$, we get $\bar{u}_{lk}U = \bar{u}'_{lk}U' \neq 0$ and since both U and U' are unitary we have $U = e^{i\alpha}U'$ for some $\alpha \in [0, 2\pi)$. Therefore, $[U] = [U']$ and h is also injective. A bijective continuous function from a compact space to a Hausdorff space is a homeomorphism, so $\text{Im}(\pi)$ and $\text{PU}(d)$ are homeomorphic. We can restate Lemma 2:

Lemma 3. Let $p : \text{U}(d) \rightarrow \text{PU}(d)$ be the quotient map. There is no continuous map $s : \text{PU}(d) \rightarrow \text{U}(d)$ such that $p \circ s = \text{id}$.

We can prove this via Lemma 1.⁶ But our aim here is to prove Lemma 3 directly, using the fact that $\text{SU}(d)$ is a finite cover of $\text{PSU}(d) := \text{SU}(d)/\mathbb{Z}[\text{SU}(d)] = \text{SU}(d)/\sim$.

Proof. First we show that $\text{PU}(d)$ and $\text{PSU}(d)$ are homeomorphic. Let $\iota : \text{SU}(d) \rightarrow \text{U}(d)$ be the embedding of the special unitaries into the unitaries and sp and p the quotient maps as depicted:

$$\begin{array}{ccc} \text{SU}(d) & \xrightarrow{\iota} & \text{U}(d) \\ sp \downarrow & & \downarrow p \\ \text{PSU}(d) & \xrightarrow{g} & \text{PU}(d) \end{array}$$

Observing that whenever $sp(U) = sp(U')$ we must have that $U = \lambda U'$ (with $\lambda^d = 1$) and therefore $(p \circ \iota)(U) = (p \circ \iota)(U')$, we get by the universal property of quotient spaces in topology that there exists a continuous $g : \text{PSU}(d) \rightarrow \text{PU}(d)$ that makes the diagram commute. Next define the function $r : \text{U}(d) \rightarrow \text{SU}(d)$, $r(U) = U/\det(U)$. This is a continuous function such that $r \circ \iota = \text{id}_{\text{SU}(d)}$ (a retraction):

$$\begin{array}{ccc} \text{SU}(d) & \xleftarrow{r} & \text{U}(d) \\ sp \downarrow & & \downarrow p \\ \text{PSU}(d) & \xleftarrow{g'} & \text{PU}(d) \end{array}$$

Whenever $p(U) = p(U')$ we have that $U = e^{i\alpha}U'$ for some $\alpha \in [0, 2\pi)$ and therefore $(sp \circ r)(U) = (sp \circ r)(U')$. By the universal property of quotient spaces in topology there exists a continuous $g' : \text{PU}(d) \rightarrow \text{PSU}(d)$ that makes the diagram commute. The following argument shows that g and g' are the inverse of each other: Take any $\bar{x} \in \text{PSU}(d)$ and some preimage of it $x \in \text{SU}(d)$, $sp(x) = \bar{x}$. Observe

that

$$\begin{aligned} g' \circ g(\bar{x}) &= g' \circ g \circ sp(x) = g' \circ p \circ \iota(x) \\ &= sp \circ r \circ \iota(x) = sp(x) = \bar{x}, \end{aligned}$$

where the second equality follows from the first commuting diagram, the third equality from the second commuting diagram, and the fourth equality from the fact that r is a retraction. Similarly, take any $[U] \in \text{PU}(d)$. For any preimage $U \in \text{U}(d)$ of $[U]$, $p(U) = [U]$ we have

$$\begin{aligned} g \circ g'([U]) &= g \circ g' \circ p(U) = g \circ sp \circ r(U) \\ &= p \circ \iota \circ r(U) = [U], \end{aligned}$$

where the second equality is from the second commuting diagram, the third equality is from the first commuting diagram, and the last equality holds because p maps $U/\det(U)$ and U to the same equivalence class. We got that the continuous maps g and g' are the inverse of each other, so $\text{PSU}(d)$ and $\text{PU}(d)$ are homeomorphic. In particular, we can find the fundamental group of $\text{PU}(d)$ by calculating it for $\text{PSU}(d)$.

Next we use the covering space argument to calculate $\pi_1[\text{PSU}(d)]$, the fundamental group of $\text{PSU}(d)$. The group of d th roots of unity (same group as $\mathbb{Z}/d\mathbb{Z}$) acts on $\text{SU}(d)$ by scalar multiplication: $\lambda U = \lambda U$. Since the group of d th roots of unity is finite, its action is a *covering space action*: for every $A \in \text{SU}(d)$ there is a neighborhood W of A such that if for some d th root of unity λ the neighborhood λW intersects W then $\lambda = 1$. Then the covering space theorem (Proposition 1.40 in [76]) states that $sp : \text{SU}(d) \rightarrow \text{PSU}(d)$ is a covering map and that the group of d th roots of unity $\mathbb{Z}/d\mathbb{Z}$ is isomorphic to $\pi_1[\text{PSU}(d)]/sp_*(\pi_1[\text{SU}(d)])$. We know that $sp_*(\pi_1[\text{SU}(d)]) = \{1\}$ because the induced homomorphism sp_* must map the trivial group to the trivial group and $\pi_1[\text{SU}(d)]$ is indeed trivial because $\text{SU}(d)$ is simply connected. We get that $\pi_1[\text{PSU}(d)]$ is isomorphic to $\mathbb{Z}/d\mathbb{Z}$ and so is $\pi_1[\text{PU}(d)]$.

Now assume towards contradiction that $s : \text{PU}(d) \rightarrow \text{U}(d)$ as in Lemma 3 exists. In other words we have the following commuting diagram:

$$\begin{array}{ccccc} \text{PU}(d) & \xrightarrow{s} & \text{U}(d) & \xrightarrow{p} & \text{PU}(d) \\ & & \searrow & \text{id} & \nearrow \end{array}$$

and knowing that $\pi_1[\text{PU}(d)] = \mathbb{Z}/d\mathbb{Z}$ and that $\pi_1[\text{U}(d)] = \mathbb{Z}$, we have the corresponding diagram for the induced homomorphisms between the fundamental groups:

$$\begin{array}{ccccc} \mathbb{Z}/d\mathbb{Z} & \xrightarrow{s_*} & \mathbb{Z} & \xrightarrow{p_*} & \mathbb{Z}/d\mathbb{Z} \\ & & \searrow & \text{id} & \nearrow \end{array}$$

Since s_* is a homomorphism, it must map the generator of $\mathbb{Z}/d\mathbb{Z}$ to some $x \in \mathbb{Z}$ such that $\overbrace{x + x + \dots + x}^{d \text{ times}} = 0$. Therefore s_* must map the generator, and consequently any element, of $\mathbb{Z}/d\mathbb{Z}$ to zero. s_* must be trivial and this contradicts $p_* \circ s_* = \text{id}$. ■

From the above proofs, extensions to approximations seem nontrivial. We reduced any if-clause algorithm to a map s of Lemma 3 that maps an equivalence class $[U]$ to some

⁶*Proof.* p is continuous and 0-homogeneous, i.e., $p(e^{i\alpha}U) = p(U)$ for all α real. Suppose that s as in Lemma 3 exists, then the opposite composition $s \circ p$ must be continuous and 0-homogeneous (as p loses the information about α). Also $(s \circ p)(U)$ differs from U only by the global phase (because $p \circ s = \text{id}$). Define $f : \text{U}(d) \rightarrow S^1$, $f(U) = \langle \mathbf{0} | U^{-1}(s \circ p)(U) | \mathbf{0} \rangle$. The function f is continuous and (-1) -homogeneous so its existence contradicts Lemma 1. ■

representative from that class $U_{\text{rep}} \in [U]$. But if the algorithm is only approximate, then so is s obtained from the reduction. Its output could be outside the input equivalence class $s([U]) \notin [U]$ or $p \circ s \neq id$. To reach contradiction we would need an extension of Lemma 3 to approximations $p \circ s \approx id$.

APPENDIX D: ALTERNATIVE PROOF OF LEMMA 1

Central to this proof is the fact that the determinant map $\det : U(d) \rightarrow S^1$ induces an isomorphism on the corresponding fundamental groups (Proposition 2.2.6 of [77]). The required algebraic-topology background, all of which is covered in Chap. 9 of [78], is just the basics of fundamental groups (paths, homotopies, time reparametrization, etc.) together with the description of the fundamental group of the circle S^1 as infinite cyclic and generated by the homotopy class of the loop $t \mapsto e^{2\pi it}$. Throughout the proof, we will use the notation $\pi_1(X, p)$ for the fundamental group of the topological space X based at p . For a path $\gamma : [0, 1] \rightarrow X$, we will denote by $[\gamma]$ its homotopy class, and for another path γ' such that $\gamma(1) = \gamma'(0)$ we will denote by $\gamma * \gamma'$ their composition.

Lemma 4. Let $\gamma : [0, 1] \rightarrow U(d)$ be the loop $\gamma(t) = e^{2\pi it}I$, where $I \in U(d)$ is the identity matrix. Then its homotopy class $[\gamma]$ is a d th power in the fundamental group $\pi_1(U(d), I)$, i.e., there exists some other loop $\gamma' : [0, 1] \rightarrow U(d)$ based at I such that $[\gamma']^d = [\gamma]$.

We proved Lemma 4 (less formally) in the main text. Here we give another, shorter proof.

Proof of Lemma 4. We use the fact (see Proposition 2.2.6 of [77]) that the determinant map $\det : U(d) \rightarrow S^1$ induces an isomorphism on the fundamental groups $\det_* : \pi_1(U(d), I) \rightarrow \pi_1(S^1, 1)$. The fundamental group $\pi_1(S^1, 1)$ is isomorphic to \mathbb{Z} and generated by the homotopy class of the loop $\nu : [0, 1] \rightarrow S^1$ defined by $\nu(t) = e^{2\pi it}$. Since $\det \circ \gamma : [0, 1] \rightarrow S^1$ is the loop $t \mapsto e^{2\pi idt}$, which winds d times around the circle, it is homotopic to the composition of ν with itself d times (as both loops are equal up to reparametrization of time). Thus,

$$\begin{aligned} \det_*([\gamma]) &= [\det \circ \gamma] = \left[\underbrace{\nu * \nu * \dots * \nu}_{d \text{ times}} \right] \\ &= \underbrace{[\nu][\nu] \dots [\nu]}_{d \text{ times}} = [\nu]^d. \end{aligned} \quad (\text{D1})$$

Since $\det_*([\gamma])$ is a d -th power in $\pi_1(S^1, 1)$ and \det_* is a group isomorphism, we deduce that $[\gamma]$ is also a d th power [to see this, apply the inverse of \det_* to both sides of Eq. (D1)]. ■

To complete the proof of Lemma 1 we give here the formal version of the arguments in the main text.

Proof of Lemma 1. Without the loss of generality we may assume that $f(I) = 1$ since otherwise we may replace f with its rotation $U \mapsto f(I)^{-1}f(U)$. Let $\gamma : [0, 1] \rightarrow U(d)$ be the loop $\gamma(t) = e^{2\pi it}I$. By Lemma 4, the homotopy class $[\gamma]$ is a d th power in $\pi_1(U(d), I)$, and so must map to a d th power under any group homomorphism. In particular, $h_*([\gamma]) = [f \circ \gamma]$ is a d th power in $\pi_1(S^1, 1)$, and so we can write $[f \circ \gamma] = [\gamma']^d$ for some loop $\gamma' : [0, 1] \rightarrow S^1$. Since the fundamental group of the circle $\pi_1(S^1, 1)$ is (infinite)

cyclic and generated by the homotopy class of the loop $\nu(t) = e^{2\pi it}$, we can write $[\gamma'] = [\nu]^k$ for some $k \in \mathbb{Z}$, and so we have $[f \circ \gamma] = [\nu]^{kd}$. On the other hand, the m -homogeneity of f allows us to derive an explicit formula for $f \circ \gamma$:

$$f \circ \gamma(t) = f(e^{2\pi it}I) = (e^{2\pi it})^m f(I) = e^{2\pi imt}.$$

The loop $t \mapsto e^{2\pi imt}$, which winds m times around the circle, is homotopic to the composition of ν with itself m times (as both loops are equal up to reparametrization of time) and so $[f \circ \gamma] = [\nu]^m$. Combining the two expressions we got for $[f \circ \gamma]$ in terms of $[\nu]$, we deduce that $[\nu]^{kd} = [\nu]^m$. Since the order of $[\nu]$ in $\pi_1(S^1, 1)$ is infinite, the two powers must be equal and $m = kd$. ■

APPENDIX E: PROCESS TOMOGRAPHY OF DEFINITION 7 EXISTS

In this Appendix we argue that algorithms of Definition 7 exist. For any U , they output a classical description of a unitary operator close to $\kappa_r(U)$, where r depends on a measurement outcome, so r is random. In example (E5) $r \in \{0, 1, \dots, d-1\}$. The classical description allows for building a new circuit that is close to implementing $|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes \kappa_r(U)$. This does not approximately achieve the if clause, and the only reason is the randomness of r . Intuitively, it achieves a random relaxation of the if clause, discussed rigorously in Appendix F.

Known process tomography algorithms achieve

$$\Pr_{X \sim p_{N,U}(\cdot)}[||X \cdot X^\dagger - U \cdot U^\dagger||_\diamond \leq \epsilon_{\diamond N}] \geq 1 - \delta_{\diamond N}, \quad (\text{E1})$$

where $\lim_{N \rightarrow \infty} \epsilon_{\diamond N} = 0$ and $\lim_{N \rightarrow \infty} \delta_{\diamond N} = 0$ [65].

Theorem 4. Any process tomography algorithm that satisfies (E1) also satisfies Definition 7.

Proof. Suppose that $||X \cdot X^\dagger - U \cdot U^\dagger||_\diamond \leq \epsilon_{\diamond N} < 1/d$. By Theorem 1.3 of [63] there exists $\alpha_{U,X} \in \mathbb{C}$,

$$(1 - \epsilon_{\diamond N})^2 \leq 1 - \epsilon_{\diamond N} \leq |\alpha_{U,X}|^2 \leq 1, \quad (\text{E2})$$

such that

$$||X - \alpha_{U,X}U||_{\text{op}} \leq 2\sqrt{\epsilon_{\diamond N}}. \quad (\text{E3})$$

This implies that the inequality

$$|\langle r|X|0\rangle - \alpha_{U,X}\langle r|U|0\rangle| \leq 2\sqrt{\epsilon_{\diamond N}} \quad (\text{E4})$$

holds for any $r \in \{0, 1, \dots, d-1\}$.

We aim to satisfy Definition 7 with the κ_j functions and the $r = r(X)$ rule described in Eq. (15) and below it:

$$\kappa_j(U) = \begin{cases} \frac{\langle j|U|0\rangle}{|\langle j|U|0\rangle|} U = \frac{|j\rangle\langle 0|}{\langle j|U|0\rangle} U & \text{if } |\langle j|U|0\rangle| \neq 0, \\ \kappa_{j+1 \bmod d}(U) & \text{otherwise,} \end{cases} \quad (\text{E5})$$

$$r = r(X) = \min\{j : |\langle j|X|0\rangle| \geq 1/\sqrt{d}\}, \quad (\text{E6})$$

which ensures that $|\langle r|X|0\rangle| \geq 1/\sqrt{d}$ and because of our supposition also that $\langle r|U|0\rangle \neq 0$, so we can safely divide by each. We will show that the left-hand side of (E3) is close to $||\kappa_r(X) - \kappa_r(U)||_{\text{op}}$, by using the following upper

bound:

$$\begin{aligned}
 \left| \alpha_{U,X} - \frac{|\langle r|U|0\rangle| |\langle r|X|0\rangle|}{|\langle r|X|0\rangle| |\langle r|U|0\rangle|} \right| &\leq \left| \alpha_{U,X} - \frac{\alpha_{U,X}}{|\alpha_{U,X}|} \right| + \left| \frac{\alpha_{U,X}}{|\alpha_{U,X}|} - \alpha_{U,X} \frac{|\langle r|U|0\rangle|}{|\langle r|X|0\rangle|} \right| + \left| \alpha_{U,X} \frac{|\langle r|U|0\rangle|}{|\langle r|X|0\rangle|} - \frac{|\langle r|U|0\rangle| |\langle r|X|0\rangle|}{|\langle r|X|0\rangle| |\langle r|U|0\rangle|} \right| \\
 &= \left| \alpha_{U,X} - \frac{|\alpha_{U,X}|}{|\alpha_{U,X}|} \right| + |\alpha_{U,X}| \left| \frac{1}{|\alpha_{U,X}|} - \frac{|\langle r|U|0\rangle|}{|\langle r|X|0\rangle|} \right| + \frac{|\langle r|U|0\rangle|}{|\langle r|X|0\rangle|} \left| \alpha_{U,X} - \frac{\langle r|X|0\rangle}{\langle r|U|0\rangle} \right| \\
 &\leq |\alpha_{U,X} - 1| + |\alpha_{U,X}| \frac{2\sqrt{\epsilon_{\diamond N}}}{|\alpha_{U,X}| |\langle r|X|0\rangle|} + \frac{|\langle r|U|0\rangle|}{|\langle r|X|0\rangle|} \frac{2\sqrt{\epsilon_{\diamond N}}}{|\langle r|U|0\rangle|}, \tag{E7}
 \end{aligned}$$

where the last inequality upper bounded: the first term using $|a| - |b| \leq |a - b|$, $a, b \in \mathbb{C}$, the second term using the same applied to Eq. (E4) with $\alpha_{U,X} \langle r|X|0\rangle$ taken out of the norm, the third term using Eq. (E4) with $\langle r|U|0\rangle$ taken out. Applying (E2) to the first term of (E7) and simplifying the rest, we get

$$\left| \alpha_{U,X} - \frac{|\langle r|U|0\rangle| |\langle r|X|0\rangle|}{|\langle r|X|0\rangle| |\langle r|U|0\rangle|} \right| \leq \epsilon_{\diamond N} + \frac{4\sqrt{\epsilon_{\diamond N}}}{|\langle r|X|0\rangle|}. \tag{E8}$$

We obtain

$$\begin{aligned}
 \|\kappa_r(X) - \kappa_r(U)\|_{\text{op}} &= \left\| \frac{|\langle r|X|0\rangle|}{|\langle r|X|0\rangle|} X - \frac{|\langle r|U|0\rangle|}{|\langle r|U|0\rangle|} U \right\|_{\text{op}} \\
 &= \left\| X - \frac{|\langle r|U|0\rangle| |\langle r|X|0\rangle|}{|\langle r|X|0\rangle| |\langle r|U|0\rangle|} U \right\|_{\text{op}} \\
 &\leq \|X - \alpha_{U,X} U\|_{\text{op}} \\
 &\quad + \left\| \alpha_{U,X} U - \frac{|\langle r|U|0\rangle| |\langle r|X|0\rangle|}{|\langle r|X|0\rangle| |\langle r|U|0\rangle|} U \right\|_{\text{op}} \\
 &\leq 2\sqrt{\epsilon_{\diamond N}} + \epsilon_{\diamond N} + 4 \frac{\sqrt{\epsilon_{\diamond N}}}{|\langle r|X|0\rangle|}. \tag{E9}
 \end{aligned}$$

The last inequality upper bounds the first term by (E3) and the second term by (E8).

Provided that condition (E1) holds, we set the parameters of Definition 7 in the following way. Let $\delta_N = \delta_{\diamond N}$. Let N_0 be some number such that $\epsilon_{\diamond N} \leq 1/d$ for all $N \geq N_0$. We set

$$\epsilon_N = \begin{cases} 2 & \text{if } N < N_0, \\ 2\sqrt{\epsilon_{\diamond N}} + \epsilon_{\diamond N} + 4\sqrt{d}\sqrt{\epsilon_{\diamond N}} & \text{if } N \geq N_0. \end{cases} \tag{E10}$$

The $r = r(X)$ rule guarantees that for $N \geq N_0$, ϵ_N upper bounds (E9). At the same time $\lim_{N \rightarrow \infty} \epsilon_N = 0$ and $\lim_{N \rightarrow \infty} \delta_N = 0$ hold, satisfying Definition 7. ■

APPENDIX F: RANDOM IF CLAUSE AND ENTANGLED IF CLAUSE

Here we show how algorithms of Definition 7 yield the random if clause and the entangled if clause. For a rigorous treatment we first need several definitions.

Definition 10 (*k-task*). A *k-task* is a pair (\mathbf{t}, Σ) , where

(1) \mathbf{t} is a vector $(t_0, t_1, \dots, t_{k-1})$ of task functions $t_i : U(d) \rightarrow L(\mathcal{H}_i)$.

(2) The query alphabet Σ is a set of query functions.

For example, the *k-if* clause is $((c_{\phi_0}, c_{\phi_1}, \dots, c_{\phi_{k-1}}), \{id, inv\})$ for any $\{\phi_i\}$. Another example is the *k-(q)th* power for some fraction $q \in \mathbb{Q}$, where each t_i is a specific *q*th power function on $U \in U(d)$ (recall that many such functions exist).

A random task or an entangled task is a *k-task* for any $k \geq 2$. The qualifiers ‘‘random’’ and ‘‘entangled’’ will invoke different notions of algorithms ‘‘achieving’’ the *k-task*. We defined postselection oracle algorithm (Definition 1), which includes only a binary success or fail measurement. Here we need a generalization:

Definition 11 (*m-postselection oracle algorithm*). An *m-postselection oracle algorithm* $((A_0, A_1, \dots, A_{m-1}), \Sigma_A)$ corresponds to an *m-tuple* of functions $A_r : U(d) \rightarrow L(\mathcal{H} \otimes \mathcal{K})$ of the form

$$A_r(U) = \Pi_r V_N[\sigma_N(U) \otimes I_{\mathcal{K}_N}] \dots V_1[\sigma_1(U) \otimes I_{\mathcal{K}_1}] V_0 \tag{F1}$$

that differ from each other only by the last multiplication by the projector Π_r . The operator $A_r(U)$ is implemented upon getting the success outcome $r \in \{0, 1, \dots, m-1\}$ in the measurement $\{\Pi_0, \Pi_1, \dots, \Pi_{m-1}, \Pi_{\text{fail}}\}$, where $\Pi_{\text{fail}} = I - \sum_{r=0}^{m-1} \Pi_r$. As before, $\sigma_i \in \Sigma_A$. Whenever the ancilla Hilbert space \mathcal{K} is initialized to the all-zero state, the total probability of success is nonzero, i.e., for all $U \in U(d)$ and for all $|\xi\rangle \in \mathcal{H}$:

$$\sum_{r=0}^{m-1} \|A_r(U)(|\xi\rangle \otimes |\mathbf{0}\rangle_{\mathcal{K}})\|^2 > 0. \tag{F2}$$

Observe that for $m = 1$ we are back to our original postselection oracle algorithm model. The superoperator implemented on the *r*th outcome is

$$\mathcal{A}_{rU}(\rho) = \text{tr}_{\mathcal{K}} [A_r(U)(\rho \otimes |\mathbf{0}\rangle\langle\mathbf{0}|_{\mathcal{K}})A_r(U)^\dagger], \tag{F3}$$

where we wrote U as a subscript to avoid too many parentheses. If $m \geq 2$, the probability $\text{tr}[\mathcal{A}_{rU}(\rho)]$ of an individual outcome $r \in \{0, 1, \dots, m-1\}$ can hit zero.

Next, we generalize approximately achieving (Definition 4) to random *k-tasks*. We would like to compare \mathcal{A}_{rU} to $t_r(U)$ as in Definition 4. The *r*th contribution to the error should be weighted by the conditional probability of the *r*th outcome, $p_{rU}(\rho) = \text{tr}[\widetilde{\mathcal{A}}_{rU}(\rho)]/p_{\text{succ}U}(\rho)$, where $\rho \in \mathcal{D}(\mathcal{H} \otimes \mathcal{H}')$, $\widetilde{\mathcal{A}}_{rU} = \mathcal{A}_{rU} \otimes \mathcal{I}_{\mathcal{H}'}$, and $p_{\text{succ}U}(\rho) = \sum_{r=0}^{m-1} \text{tr}[\widetilde{\mathcal{A}}_{rU}(\rho)]$ is the total probability of success. On inputs U, ρ that yield a nonzero probability of outcome *r*, the contribution is

$$\begin{aligned}
 p_{rU}(\rho) &\left\| \frac{\widetilde{\mathcal{A}}_{rU}(\rho)}{\text{tr}[\widetilde{\mathcal{A}}_{rU}(\rho)]} - t_r(U)\rho t_r(U)^\dagger \right\|_{\text{tr}} \\
 &= \frac{1}{p_{\text{succ}U}(\rho)} \left\| \widetilde{\mathcal{A}}_{rU}(\rho) - \text{tr}[\widetilde{\mathcal{A}}_{rU}(\rho)]t_r(U)\rho t_r(U)^\dagger \right\|_{\text{tr}}, \tag{F4}
 \end{aligned}$$

which is well defined on all $\rho \in \mathcal{D}(\mathcal{H})$ because of condition (F2). This justifies the following definition.

Definition 12 (m-postselection oracle algorithm ϵ -approximately achieving a random k-task). An m-postselection oracle algorithm $((A_0, A_1, \dots, A_{m-1}), \Sigma_A)$ ϵ -approximately achieves a random $((t_0, t_1, \dots, t_{k-1}), \Sigma)$ if $\Sigma_A \subseteq \Sigma$ and if for all $U \in \mathcal{U}(d)$

$$\begin{aligned} & \sup_{\substack{\mathcal{H}' \\ \rho \in \mathcal{D}(\mathcal{H} \otimes \mathcal{H}')}} \frac{1}{p_{\text{succ}} U(\rho)} \\ & \times \sum_{r=0}^{m-1} \|\widetilde{\mathcal{A}}_{rU}(\rho) - \text{tr}[\widetilde{\mathcal{A}}_{rU}(\rho)] t_r(\widetilde{U}) \rho t_r(\widetilde{U})^\dagger\|_{\text{tr}} \\ & \leq \epsilon, \end{aligned} \quad (\text{F5})$$

where the wide tilde denotes the \mathcal{H}' extension by the identity superoperator or operator, and for $r \geq k$ we set $t_r = t_{k-1}$.

If either $k = 1$ or $m = 1$, then (F5) simplifies to (9). Consequently, the if clause ($k = 1$) is impossible also in the $m > 1$ model, and the random if clause is impossible in the original $m = 1$ model. This justifies the remaining impossibilities in Table II.

Extending the algorithmic model is not necessary for the entangled if clause. Binary success or fail measurement suffices if the r value can remain in an additional, entangled register.

Definition 13 (Postselection oracle algorithm ϵ -approximately achieving an entangled k-task). A postselection oracle algorithm (\mathcal{A}, Σ_A) ϵ -approximately achieves an entangled $((t_0, t_1, \dots, t_{k-1}), \Sigma)$ if $\Sigma_A \subseteq \Sigma$ and if $\mathcal{A}_U := \mathcal{A}(U)$ maps $L(\mathcal{H}) \rightarrow L(\mathcal{H} \otimes \mathbb{C}^k)$ and

$$\begin{aligned} & \sup_{\substack{\mathcal{H}' \\ \rho \in \mathcal{D}(\mathcal{H} \otimes \mathcal{H}')}} \frac{1}{\text{tr}[\widetilde{\mathcal{A}}_U(\rho)]} \\ & \times \sum_{r=0}^{k-1} \|\langle r | \widetilde{\mathcal{A}}_U(\rho) | r \rangle - \text{tr}[\langle r | \widetilde{\mathcal{A}}_U(\rho) | r \rangle] t_r(\widetilde{U}) \rho t_r(\widetilde{U})^\dagger\|_{\text{tr}} \\ & \leq \epsilon, \end{aligned} \quad (\text{F6})$$

where $\{|r\rangle\}_{r \in [k]}$ is the computational basis of \mathbb{C}^k , and the wide tilde denotes the \mathcal{H}' extension by the identity superoperator or operator.

Writing $p_{rU}(\rho) = \text{tr}[\langle r | \widetilde{\mathcal{A}}_U(\rho) | r \rangle]$ the sum in (F6) is equal to

$$\begin{aligned} & \sum_{r=0}^{k-1} \|\langle r | \widetilde{\mathcal{A}}_U(\rho) | r \rangle \otimes |r\rangle\langle r| \\ & - p_{rU}(\rho) t_r(\widetilde{U}) \rho t_r(\widetilde{U})^\dagger \otimes |r\rangle\langle r|\|_{\text{tr}} \\ & = \|\sum_{r=0}^{k-1} |r\rangle\langle r | \widetilde{\mathcal{A}}_U(\rho) | r \rangle\langle r| \\ & - \sum_{r=0}^{k-1} p_{rU}(\rho) (t_r(\widetilde{U}) \otimes |r\rangle\langle r | \rho t_r(\widetilde{U})^\dagger \otimes |r\rangle\langle r|)\|_{\text{tr}}. \end{aligned} \quad (\text{F7})$$

Thus, expression (16) in the main text is indeed an example of an entangled k -task.

The last two definitions are closely related. Applying the principle of deferred measurement to an algorithm of Definition 12 will give an algorithm of Definition 13 for the

entangled version of the same k -task. In the opposite direction, adding the measurement with $\Pi_r = I \otimes |r\rangle\langle r|$ results in the corresponding k -postselection oracle algorithm.

Next we show that the random if clause is possible. By the above argument, the entangled if clause follows. We will show that example (E5) yields an algorithm satisfying Definition 12 for $k = m = d$, justifying the tick in Table II. Process tomography can obtain classical descriptions $\kappa_r(X) \in \mathcal{U}(d)$ that satisfy Definition 7. From this classical description, we can build a new circuit that implements an operator arbitrarily close to $c_r(X) := |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes \kappa_r(X)$. If we know the value of r , then this strategy implements a superoperator arbitrarily close to $\mathcal{A}_{rU,N}$, where

$$\mathcal{A}_{rU,N}(\rho) = \sum_{\substack{X \in \Omega_N, \\ r(X)=r}} p_{U,N}(X) c_r(X) \rho c_r(X)^\dagger. \quad (\text{F8})$$

Moreover, $p_{\text{succ}} U(\rho) = 1$. Then the expression on the left-hand side of (F5) is close to

$$\begin{aligned} & \sup_{\substack{\mathcal{H}' \\ \rho \in \mathcal{D}(\mathcal{H} \otimes \mathcal{H}')}} \sum_{r=0}^{d-1} \|\widetilde{\mathcal{A}}_{rU,N}(\rho) - \text{tr}[\widetilde{\mathcal{A}}_{rU,N}(\rho)] c_r(\widetilde{U}) \rho c_r(\widetilde{U})^\dagger\|_{\text{tr}} \\ & \leq \sup_{\substack{\mathcal{H}' \\ \rho \in \mathcal{D}(\mathcal{H} \otimes \mathcal{H}')}} \sum_{X \in \Omega_N} p_{U,N}(X) \|c_r(X) \rho c_r(X)^\dagger \\ & - \text{tr}[c_r(X) \rho c_r(X)^\dagger] c_r(\widetilde{U}) \rho c_r(\widetilde{U})^\dagger\|_{\text{tr}} \\ & \leq \sum_{X \in \Omega_N} p_{U,N}(X) \|c_r(X) \cdot c_r(X)^\dagger - c_r(\widetilde{U}) \cdot c_r(\widetilde{U})^\dagger\|_{\diamond} \\ & \leq \sum_{X \in \Omega_N} p_{U,N}(X) 2 \|\kappa_r(X) - \kappa_r(\widetilde{U})\|_{\text{op}} \\ & \leq 2\epsilon_N + 4\delta_N. \end{aligned} \quad (\text{F9})$$

In the first inequality we took the sum of (F8) and $p_{U,N}(X)$ out of the trace norm. The second inequality holds because $\text{tr}[c_r(X) \rho c_r(X)^\dagger] = 1$ and because moving the supremum into the sum can only increase the expression. The third inequality uses the upper bound on the diamond norm in terms of the operator norm [66, Lemma 12.6] and the equality $\|Y \otimes Z\|_{\text{op}} = \|Y\|_{\text{op}} \|Z\|_{\text{op}}$. The last inequality follows from Definition 7.

This shows that the described tomography strategy ϵ -approximately achieves the random if clause, for any $\epsilon > 0$. In particular, it achieves the random d -if clause, where $c_{\phi_r}(U) = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes \kappa_r(U)$ for κ_r of (E5).

APPENDIX G: SYMMETRIC EQUATIONS FOR THE DETERMINANT AND THE MINORS

In the following we first prove the symmetric formula for the determinant of Eq. (19), then we use it to prove Lemma 6, the symmetric formula for the minors of Eq. (22).

Lemma 5. The determinant of $M \in \text{Mat}_{n \times n}(\mathbb{C})$ is given by the formula:

$$\det(M) = \frac{1}{n!} \sum_{\pi, \tau \in \mathcal{S}_n} \text{sgn}(\tau) \text{sgn}(\pi) \prod_{i=1}^n M_{\tau(i), \pi(i)}.$$

Proof. By the usual Leibniz formula, we have $\det(M) = \sum_{\pi \in S_n} \text{sgn}(\pi) \prod_{i=1}^n M_{i,\pi(i)}$. For each $\tau \in S_n$, set M^τ to be the matrix obtained from M by permuting the rows of M according to τ^{-1} , i.e., $M_{i,j}^\tau = M_{\tau(i),j}$. On the one hand, by the properties of the determinant function we have $\det(M^\tau) = \text{sgn}(\tau) \det(M)$, and on the other hand, by the Leibniz formula

$$\det(M^\tau) = \sum_{\pi \in S_n} \text{sgn}(\pi) \prod_{i=1}^n M_{i,\pi(i)}^\tau = \sum_{\pi \in S_n} \text{sgn}(\pi) \prod_{i=1}^n M_{\tau(i),\pi(i)}.$$

Combining the two expressions for $\det(M^\tau)$ and multiplying by $\text{sgn}(\tau)$ gives

$$\det(M) = \sum_{\pi \in S_n} \text{sgn}(\pi) \text{sgn}(\tau) \prod_{i=1}^n M_{\tau(i),\pi(i)}.$$

Summing this equality over all $\tau \in S_n$ and dividing by $n!$ gives the desired result. ■

Lemma 6. The (i, j) minor of a matrix $M \in \text{Mat}_{n \times n}(\mathbb{C})$ is given by

$$\det(M_{i,j}) = \frac{(-1)^{i+j}}{(n-1)!} \cdot \sum_{\substack{\pi, \tau \in S_n \\ \pi(1)=j \\ \tau(1)=i}} \text{sgn}(\tau) \text{sgn}(\pi) \prod_{k=2}^n M_{\tau(k),\pi(k)}.$$

Proof. We think of S_{n-1} as the subset of S_n of permutations fixing n , i.e., $S_{n-1} = \{\alpha \in S_n \mid \alpha(n) = n\}$. For each $m \in \{i, j\}$ set $X_m = \{\alpha \in S_n \mid \alpha(1) = m\}$, and let $T_m : X_m \rightarrow S_{n-1}$ be the mapping $T_m(\alpha) = (n \ n-1 \ \dots \ m) \circ \alpha \circ (1 \ 2 \ 3 \ \dots \ n-1 \ n)$. Note that $T_m(\alpha)$ lies in S_{n-1} since it fixes n , that T_m is a bijection (since $T_m \alpha$ it is obtained from α by composing on both sides with fixed permutations), and finally that by the multiplicativity of sgn we have

$$\text{sgn}[T_m(\alpha)] = (-1)^{n-m} \text{sgn}(\alpha) (-1)^{n-1} = (-1)^{1-m} \text{sgn}(\alpha).$$

One can now verify that for all $k \in \{1, 2, \dots, n-1\}$ and for every $\tau \in X_i$ and $\pi \in X_j$ we have

$$(M_{i,j})_{T_i\tau(k), T_j\pi(k)} = M_{\tau(k+1), \pi(k+1)}.$$

Indeed, if $\tau(k+1) < i$ then deleting the i th row from M does not change the index of the $\tau(k+1)$ th row, and correspondingly $(n \ n-1 \ \dots \ i)$ fixes $\tau(k+1)$ and so $(T_i\tau)(k) = \tau(k+1)$. If $\tau(k+1) > i$ then then deleting the i th row from M decreases the index of the $\tau(k+1)$ th row by one, and correspondingly $(n \ n-1 \ \dots \ i)$ does the same thing to $\tau(k+1)$. A similar argument for π shows the validity of the last equality. Now, by Lemma 5,

$$\begin{aligned} \det(M_{i,j}) &= \frac{1}{(n-1)!} \sum_{\pi', \tau' \in S_{n-1}} \text{sgn}(\tau') \text{sgn}(\pi') \prod_{k=1}^{n-1} (M_{i,j})_{\tau'(k), \pi'(k)} \\ &= \frac{1}{(n-1)!} \sum_{\pi \in X_i, \tau \in X_j} \text{sgn}(T_i\tau) \text{sgn}(T_j\pi) \prod_{k=1}^{n-1} (M_{i,j})_{T_i\tau(k), T_j\pi(k)} \\ &= \frac{1}{(n-1)!} \sum_{\pi \in X_i, \tau \in X_j} (-1)^{1-i} \text{sgn}(\tau) (-1)^{1-j} \text{sgn}(\pi) \prod_{k=1}^{n-1} M_{\tau(k+1), \pi(k+1)} \\ &= \frac{(-1)^{i+j}}{(n-1)!} \sum_{\pi \in X_i, \tau \in X_j} \text{sgn}(\tau) \text{sgn}(\pi) \prod_{k=2}^n M_{\tau(k), \pi(k)}. \end{aligned}$$

■

APPENDIX H: LINEAR-OPTICS SOLUTION

First, we briefly review the relevant parts of the linear-optics model as described by Aaronson and Arkhipov [79]. A $d \times d$ unitary U specifies what happens to each of the d input modes, collected in the set S_U :

$$\forall k \in S_U \quad a_k^\dagger \mapsto \sum_{j \in S_U} U_{jk} a_j^\dagger,$$

where a_i^\dagger is the creation operator on the i th mode. The unitary acts trivially on other modes, which is captured by the following extension of the matrix:

$$U_{jk} := \delta_{jk} \quad \text{if } j \notin S_U \text{ or } k \notin S_U. \quad (\text{H1})$$

Now for any set of modes S we can describe the action of U by

$$\forall k \in S \quad a_k^\dagger \mapsto \sum_{j \in S \cup S_U} U_{jk} a_j^\dagger.$$

A general quantum state of n particles distributed among m modes is

$$|\psi\rangle_{m,n} = \sum_{\vec{n}=(n_1,n_2,\dots,n_m)\in\Phi_{m,n}} \alpha_{\vec{n}} \frac{(a_1^\dagger)^{n_1} (a_2^\dagger)^{n_2} \dots (a_m^\dagger)^{n_m}}{\sqrt{n_1! n_2! \dots n_m!}} |\Omega\rangle,$$

where $\Phi_{m,n}$ is the set of possible partitions of n particles into m slots. In words, a state is a result of taking the vacuum state $|\Omega\rangle = |\Omega\rangle_{m,0}$ and applying to it a degree- n polynomial of m variables, creation operators. A unitary U acts on the state as a linear transformation of the polynomial's *variables*. The output is

$$U[|\psi\rangle_{|S|,n}] = \sum_{\vec{n}=(n_1,n_2,\dots,n_{|S|})\in\Phi_{|S|,n}} \alpha_{\vec{n}} \frac{\left(\sum_{j\in S\cup S_U} U_{j1} a_j^\dagger\right)^{n_1} \left(\sum_{j\in S\cup S_U} U_{j2} a_j^\dagger\right)^{n_2} \dots \left(\sum_{j\in S\cup S_U} U_{j|S|} a_j^\dagger\right)^{n_{|S|}}}{\sqrt{n_1! n_2! \dots n_{|S|}!}} |\Omega\rangle, \quad (\text{H2})$$

where S are the input state's modes. In particular, for $k \in S \setminus S_U$, U leaves the k th mode of the input state untouched [since then Eq. (H1) applies]. Equation (H2) implies that it also leaves the vacuum state ($n = 0$) untouched. These are the two crucial reasons why building an if clause in the linear-optics model is easy.

We define *oracle linear-optics* algorithm to be a sequence of gates $V_0, \sigma_1(U), V_1, \sigma_2(U), \dots, V_{N-1}, \sigma_N(U), V_N$; V_i representing fixed gates and $\sigma_i(U)$ query gates as in Fig. 2 in the main text, but the gates now act on a vector state according to Eq. (H2).

Next, we can describe the Mach-Zehnder-interferometer solution [Fig. 1(b)] of the if clause ($c_\phi, \{id, inv\}$) as a single-query *oracle linear-optics* algorithm. In the quantum-circuit model, a general pure-state input to $c_\phi(U)$ decomposes as $|\xi\rangle = \alpha_0|0\rangle \otimes |t_0\rangle + \alpha_1|1\rangle \otimes |t_1\rangle$, where the first Hilbert space is the control qubit and the second the target qudit. Suppose we encode the control in the photon's polarization. Collect in S_h the modes corresponding to horizontal polarization $|S_h| = d$ and in S_v the modes corresponding to the vertical $|S_v| = d$. We write this as the $2d$ -mode single-photon state

$$|\psi_{|\xi}\rangle_{2d,1} = \alpha_0 |\psi_{|t_0}\rangle_{|S_h|,1} |\Omega\rangle_{|S_v|,0} + \alpha_1 |\Omega\rangle_{|S_h|,0} |\psi_{|t_1}\rangle_{|S_v|,1},$$

i.e., in the first summand there is one photon on the d horizontal modes and zero photons on the d vertical modes and in the second summand vice versa. In Fig. 1(b) each polarizing beam splitter corresponds to the unitary V which effectively relabels polarization modes to path modes $h \mapsto$ lower and $v \mapsto$ upper. Thus, the resulting unitary $V^\dagger U V$ acts nontrivially on the S_v modes only. Equation (H2) implies that

$$\begin{aligned} (V^\dagger U V)[|\psi_{|\xi}\rangle_{2d,1}] &= \alpha_0 |\psi_{|t_0}\rangle_{|S_h|,1} U[|\Omega\rangle_{|S_v|,0}] + \alpha_1 |\Omega\rangle_{|S_h|,0} U[|\psi_{|t_1}\rangle_{|S_v|,1}] \\ &= \alpha_0 |\psi_{|t_0}\rangle_{|S_h|,1} |\Omega\rangle_{|S_v|,0} + \alpha_1 |\Omega\rangle_{|S_h|,0} |\psi_{U|t_1}\rangle_{|S_v|,1} \end{aligned} \quad (\text{H3})$$

and the second equality holds because U acts trivially on vacuum, and because $|\psi_{|t_1}\rangle_{|S_v|,1}$ has one photon on exactly those modes that are in S_U . The output state (H3) is the linear-optics version of

$$\alpha_0 |0\rangle \otimes |t_0\rangle + \alpha_1 |1\rangle \otimes U|t_1\rangle, \quad (\text{H4})$$

which equals $c_\phi(U)|\xi\rangle$ with ϕ identically zero.

-
- [1] P. W. Shor, Algorithms for quantum computation: discrete logarithms and factoring, in *Proceedings 35th Annual Symposium on Foundations of Computer Science (IEEE, Piscataway, NJ, 1994)*, pp. 124–134.
- [2] A. W. Harrow, A. Hassidim, and S. Lloyd, Quantum algorithm for linear systems of equations, *Phys. Rev. Lett.* **103**, 150502 (2009).
- [3] F. G. S. L. Brandao and K. M. Svore, Quantum speed-ups for solving semidefinite programs, in *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS) (IEEE, Piscataway, NJ, 2017)*, pp. 415–426.
- [4] J. van Apeldoorn, A. Gilyén, S. Gribling, and R. de Wolf, Quantum SDP-solvers: Better upper and lower bounds, *Quantum* **4**, 230 (2020).
- [5] K. Temme, T. J. Osborne, K. G. Vollbrecht, D. Poulin, and F. Verstraete, Quantum metropolis sampling, *Nature (London)* **471**, 87 (2011).
- [6] J. D. Whitfield, J. Biamonte, and A. Aspuru-Guzik, Simulation of electronic structure Hamiltonians using quantum computers, *Mol. Phys.* **109**, 735 (2011).
- [7] A. Yu. Kitaev, Quantum measurements and the abelian stabilizer problem, [arXiv:quant-ph/9511026](https://arxiv.org/abs/quant-ph/9511026).
- [8] R. Beals, H. Buhrman, R. Cleve, M. Mosca, and R. De Wolf, Quantum lower bounds by polynomials, *J. ACM* **48**, 778 (2001).
- [9] C. H. Bennett, E. Bernstein, G. Brassard, and U. Vazirani, Strengths and weaknesses of quantum computing, *SIAM J. Comput.* **26**, 1510 (1997).
- [10] A. Ambainis, Quantum lower bounds by quantum arguments, in *Proceedings of the Thirty-second Annual ACM Symposium on Theory of Computing (ACM, New York, 2000)*, pp. 636–643.
- [11] P. Hoyer, T. Lee, and R. Spalek, Negative weights make adversaries stronger, in *Proceedings of the Thirty-ninth Annual ACM*

- Symposium on Theory of Computing* (ACM, New York, 2007), pp. 526–535.
- [12] L. K. Grover, A fast quantum mechanical algorithm for database search, in *Proceedings of the Twenty-eighth Annual ACM Symposium on Theory of Computing* (ACM, New York, 1996), pp. 212–219.
- [13] W. K. Wootters and W. H. Zurek, A single quantum cannot be cloned, *Nature (London)* **299**, 802 (1982).
- [14] A. K. Pati and S. L. Braunstein, Impossibility of deleting an unknown quantum state, *Nature (London)* **404**, 164 (2000).
- [15] V. Bužek, M. Hillery, and R. F. Werner, Optimal manipulations with qubits: Universal-not gate, *Phys. Rev. A* **60**, R2626 (1999).
- [16] M. A. Nielsen and I. L. Chuang, Programmable quantum gate arrays, *Phys. Rev. Lett.* **79**, 321 (1997).
- [17] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Quantum cryptography, *Rev. Mod. Phys.* **74**, 145 (2002).
- [18] G. Chiribella, G. Mauro D’Ariano, and P. Perinotti, Probabilistic theories with purification, *Phys. Rev. A* **81**, 062348 (2010).
- [19] B. Coecke and R. Duncan, Interacting quantum observables: Categorical algebra and diagrammatics, *New J. Phys.* **13**, 043016 (2011).
- [20] A. Kent, Quantum tasks in Minkowski space, *Class. Quantum Gravity* **29**, 224013 (2012).
- [21] W. H. Zurek, Quantum darwinism, *Nat. Phys.* **5**, 181 (2009).
- [22] P. M. B. Vitányi, Quantum Kolmogorov complexity based on classical descriptions, *IEEE Trans. Inf. Theory* **47**, 2464 (2001).
- [23] V. Bužek and M. Hillery, Quantum copying: Beyond the no-cloning theorem, *Phys. Rev. A* **54**, 1844 (1996).
- [24] V. Bužek, S. L. Braunstein, M. Hillery, and D. Bruß, Quantum copying: A network, *Phys. Rev. A* **56**, 3446 (1997).
- [25] L.-M. Duan and G.-C. Guo, Probabilistic cloning and identification of linearly independent quantum states, *Phys. Rev. Lett.* **80**, 4999 (1998).
- [26] M. Hillery, M. Ziman, and V. Bužek, Approximate programmable quantum processors, *Phys. Rev. A* **73**, 022345 (2006).
- [27] Y. Yang, R. Renner, and G. Chiribella, Optimal universal programming of unitary gates, *Phys. Rev. Lett.* **125**, 210501 (2020).
- [28] G. Vidal, L. Masanes, and J. I. Cirac, Storing quantum dynamics in quantum states: A stochastic programmable gate, *Phys. Rev. Lett.* **88**, 047905 (2002).
- [29] M. Sedláč, A. Bisio, and M. Ziman, Optimal probabilistic storage and retrieval of unitary channels, *Phys. Rev. Lett.* **122**, 170502 (2019).
- [30] S. Ishizaka and T. Hiroshima, Quantum teleportation scheme by selecting one of multiple output ports, *Phys. Rev. A* **79**, 042306 (2009).
- [31] A. Chefles and S. M. Barnett, Strategies and networks for state-dependent quantum cloning, *Phys. Rev. A* **60**, 136 (1999).
- [32] N. Gisin and S. Massar, Optimal quantum cloning machines, *Phys. Rev. Lett.* **79**, 2153 (1997).
- [33] D. Bruss, A. Ekert, and C. Macchiavello, Optimal universal quantum cloning and state estimation, *Phys. Rev. Lett.* **81**, 2598 (1998).
- [34] M. Keyl and R. F. Werner, Optimal cloning of pure states, testing single clones, *J. Math. Phys.* **40**, 3283 (1999).
- [35] R. F. Werner, Optimal cloning of pure states, *Phys. Rev. A* **58**, 1827 (1998).
- [36] M. Araújo, A. Feix, F. Costa, and Č. Brukner, Quantum circuits cannot control unknown operations, *New J. Phys.* **16**, 093026 (2014).
- [37] J. Miyazaki, A. Soeda, and M. Murao, Complex conjugation supermap of unitary quantum maps and its universal implementation protocol, *Phys. Rev. Res.* **1**, 013007 (2019).
- [38] M. T. Quintino, Q. Dong, A. Shimbo, A. Soeda, and M. Murao, Reversing unknown quantum transformations: Universal quantum circuit for inverting general unitary operations, *Phys. Rev. Lett.* **123**, 210502 (2019).
- [39] M. T. Quintino, Q. Dong, A. Shimbo, A. Soeda, and M. Murao, Probabilistic exact universal quantum circuits for transforming unitary operations, *Phys. Rev. A* **100**, 062339 (2019).
- [40] L. Sheridan, D. Maslov, and M. Mosca, Approximating fractional time quantum evolution, *J. Phys. A: Math. Theor.* **42**, 185302 (2009).
- [41] A. Gilyén, Y. Su, G. H. Low, and N. Wiebe, Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics, in *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing* (ACM, New York, 2019), pp. 193–204.
- [42] A. Bogdanov and L. Trevisan, Average-case complexity, *Found. Trends Theor. Comput. Sci.* **2**, 1 (2006).
- [43] G. Chiribella, G. M. D’Ariano, P. Perinotti, and B. Valiron, Quantum computations without definite causal structure, *Phys. Rev. A* **88**, 022318 (2013).
- [44] O. Oreshkov, F. Costa, and Č. Brukner, Quantum correlations with no causal order, *Nat. Commun.* **3**, 1092 (2012).
- [45] O. Oreshkov, Time-delocalized quantum subsystems and operations: on the existence of processes with indefinite causal structure in quantum mechanics, *Quantum* **3**, 206 (2019).
- [46] N. Paunković and M. Vojinović, Causal orders, quantum circuits and spacetime: distinguishing between definite and superposed causal orders, *Quantum* **4**, 275 (2020).
- [47] L. M. Procopio, A. Moqanaki, M. Araújo, F. Costa, I. A. Calafell, E. G. Dowd, D. R. Hamel, L. A. Rozema, Č. Brukner, and P. Walther, Experimental superposition of orders of quantum gates, *Nat. Commun.* **6**, 7913 (2015).
- [48] K. Goswami, C. Giarmatzi, M. Kewming, F. Costa, C. Branciard, J. Romero, and A. G. White, Indefinite causal order in a quantum switch, *Phys. Rev. Lett.* **121**, 090503 (2018).
- [49] G. Rubino, L. A. Rozema, A. Feix, M. Araújo, J. M. Zeuner, L. M. Procopio, Č. Brukner, and P. Walther, Experimental verification of an indefinite causal order, *Sci. Adv.* **3**, e1602589 (2017).
- [50] A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. A. Smolin, and H. Weinfurter, Elementary gates for quantum computation, *Phys. Rev. A* **52**, 3457 (1995).
- [51] M. A. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2002).
- [52] A. Soeda, Limitations on quantum subroutine designing due to the linear structure of quantum operators, in *The 3rd International Conference on Quantum Information and Technology (ICQIT2013)* (NII, Tokyo, 2013).
- [53] J. Thompson, K. Modi, V. Vedral, and M. Gu, Quantum plug n’play: Modular computation in the quantum regime, *New J. Phys.* **20**, 013004 (2018).

- [54] Q. Dong, S. Nakayama, A. Soeda, and M. Murao, Controlled quantum operations and combs, and their applications to universal controllization of divisible unitary operations, [arXiv:1911.01645](https://arxiv.org/abs/1911.01645).
- [55] S. Aaronson, The ten most annoying questions in quantum computing, 2006, retrieved from <https://scottaaronson.blog/?p=112>.
- [56] E. L. Hahn, Spin echoes, *Phys. Rev.* **80**, 580 (1950).
- [57] I. L. Chuang and M. A. Nielsen, Prescription for experimental determination of the dynamics of a quantum black box, *J. Mod. Opt.* **44**, 2455 (1997).
- [58] J. F. Poyatos, J. I. Cirac, and P. Zoller, Complete characterization of a quantum process: The two-bit quantum gate, *Phys. Rev. Lett.* **78**, 390 (1997).
- [59] D. W. C. Leung, Towards robust quantum computation, Ph.D. thesis, Stanford University, 2000.
- [60] H. Buhrman and R. De Wolf, Complexity measures and decision tree complexity: A survey, *Theor. Comput. Sci.* **288**, 21 (2002).
- [61] S. Aaronson, Quantum computing, postselection, and probabilistic polynomial-time, *Proc. R. Soc. A* **461**, 3473 (2005).
- [62] Å. Baumeler and S. Wolf, The space of logically consistent classical processes without causal order, *New J. Phys.* **18**, 013036 (2016).
- [63] Z. Gavorová, Notes on distinguishability of postselected computations, [arXiv:2011.08487](https://arxiv.org/abs/2011.08487).
- [64] K. Borsuk, Drei sätze über die n-dimensionale euklidische sphäre, *Fundamenta Mathematicae* **20**, 177 (1933).
- [65] T. Surawy-Stepney, J. Kahn, R. Kueng, and M. Guta, Projected least-squares quantum process tomography, *Quantum* **6**, 844 (2022).
- [66] D. Aharonov, A. Kitaev, and N. Nisan, Quantum circuits with mixed states, in *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing* (ACM, New York, 1998), pp. 20–30.
- [67] S. Lloyd, Universal quantum simulators, *Science* **273**, 1073 (1996).
- [68] D. Bacon, I. L. Chuang, and A. W. Harrow, Efficient quantum circuits for schur and Clebsch-Gordan transforms, *Phys. Rev. Lett.* **97**, 170502 (2006).
- [69] W. V. Dam and G. Seroussi, Efficient quantum algorithms for estimating Gauss sums, [arXiv:quant-ph/0207131](https://arxiv.org/abs/quant-ph/0207131).
- [70] M. Araújo, F. Costa, and Č. Brukner, Computational advantage from quantum-controlled ordering of gates, *Phys. Rev. Lett.* **113**, 250402 (2014).
- [71] L. A. Levin, Average case complete problems, *SIAM J. Comput.* **15**, 285 (1986).
- [72] D. A. Spielman and S.-H. Teng, Smoothed analysis: An attempt to explain the behavior of algorithms in practice, *Commun. ACM* **52**, 76 (2009).
- [73] M. Bläser and B. Manthey, Smoothed complexity theory, *ACM Trans. Comput. Theory* **7**, 1 (2015).
- [74] A. Blum and J. Dunagan, Smoothed analysis of the perceptron algorithm for linear programming, in *Proceedings of the Thirteenth Annual ACM-SIAM Symposium on Discrete Algorithms, San Francisco California* (ACM, New York, 2002), pp. 905–914.
- [75] U. Chabaud and S. Mehraban, Holomorphic representation of quantum computations, *Quantum* **6**, 831 (2022).
- [76] A. Hatcher, *Algebraic Topology* (Cambridge University Press, Cambridge, 2000).
- [77] D. McDuff and D. Salamon, *Introduction to Symplectic Topology* (Oxford University Press, Oxford, 2017).
- [78] J. R. Munkres, *Topology*, 2nd ed. (Prentice Hall, Upper Saddle River, NJ, 2000).
- [79] S. Aaronson and A. Arkhipov, The computational complexity of linear optics, in *Proceedings of the Forty-third Annual ACM Symposium on Theory of Computing* (ACM, New York, 2011), pp. 333–342.
- [80] R. P. Feynman, Simulating physics with computers, *Int. J. Theor. Phys.* **21**, 467 (1982).
- [81] D. S. Abrams and S. Lloyd, Simulation of many-body Fermi systems on a universal quantum computer, *Phys. Rev. Lett.* **79**, 2586 (1997).
- [82] S. B. Bravyi and A. Y. Kitaev, Fermionic quantum computation, *Ann. Phys.* **298**, 210 (2002).