

Oracle problems as communication tasks and optimization of quantum algorithms

Amit Te'eni^{1,*}, Zohar Schwartzman-Nowik^{1,2}, Marcin Nowakowski^{3,4,5}, Paweł Horodecki^{6,3}, and Eliahu Cohen^{1,7}¹*Faculty of Engineering and Institute of Nanotechnology and Advanced Materials, Bar Ilan University, Ramat Gan, Israel*²*Department of Computer Science and Engineering, Hebrew University of Jerusalem, Jerusalem, Israel*³*Faculty of Applied Physics and Mathematics, Gdansk University of Technology, 80-952 Gdansk, Poland*⁴*National Quantum Information Center, 80-309 Gdansk, Poland*⁵*ObserveQ, 80-890 Gdansk, Poland*⁶*International Centre for Theory of Quantum Technologies (ICTQT), University of Gdansk, 80-309 Gdansk, Poland*⁷*Institute for Quantum Studies, Chapman University, Orange, California 92866, USA*

(Received 16 October 2025; accepted 14 July 2026; published 1 September 2026)

Quantum query complexity studies the number of queries needed to learn some property of a black box. A closely related question is how well an algorithm can succeed with this learning task using only a fixed number of queries. In this work, we propose measuring an algorithm's performance using the mutual information between the output and the actual value. The task of optimizing this mutual information using a single query is similar to a basic task of quantum communication, where one attempts to maximize the mutual information of the sender and receiver. We make this analogy precise by splitting the algorithm between two agents, obtaining a communication protocol. The oracle's target property plays the role of a message that Alice encodes into a quantum state, which is subsequently sent over to Bob. The first part of the algorithm performs this encoding, and the second part measures the state and aims to deduce the message from the outcome. Moreover, we formally consider the oracle as a separate subsystem, whose state records the unknown oracle identity. Within this construction, Bob's optimal measurement basis minimizes the quantum correlations between the two subsystems. We also find a lower bound on the mutual information, which is related to quantum coherence. These results extend to multiple-query nonadaptive algorithms. As a result, we describe the optimal nonadaptive algorithm that uses at most a fixed number of queries, for any oracle classification problem. Crucially, this mutual-information perspective was shown elsewhere [A. Levi *et al.*, *Quantum Sci. Technol.* **11**, 015029 (2026)] to have practical utility for algorithmic design, providing the theoretical foundation to optimize iterative subroutines in hybrid quantum-classical schemes. Within the present work, we apply this framework to analyze the stage-by-stage information flow and track partial progress in several standard quantum algorithms, including the Deutsch-Jozsa, Bernstein-Vazirani, and Shor-Kitaev algorithms.

DOI: [10.1103/5qsb-5jy5](https://doi.org/10.1103/5qsb-5jy5)

I. INTRODUCTION

Quantum query complexity explores the performance of quantum algorithms in terms of the number of queries they make to an oracle [1]. There are at least two good reasons to study query complexity. First, many important quantum algorithms solve oracle problems, including some of the first examples of quantum advantage (such as the Deutsch-Jozsa [2], Bernstein-Vazirani [3], and Simon [4] algorithms). Second, query complexity is much simpler to study than time complexity (the number of gates used in an algorithm). Indeed, it is sometimes possible to prove lower bounds on query complexity [5–10] (lower bounds are usually difficult to obtain in both classical and quantum complexity theory). On the other hand, results about query complexity may allow us to

obtain insight regarding time complexity. Therefore, quantum query complexity remains widely studied, with many interesting problems still open [11,12].

One avenue of research seeks to study the origins of quantum speedup, i.e., the computational advantage of quantum computers with respect to classical ones. The role played by various resources in expediting quantum computation has been studied extensively [13–30] and remains an active area of research. One of the quantum resources that seems to be essential for quantum speedup is quantum discord [20,28,31–33]. However, research in this direction has focused on mixed-state quantum computation, since for pure states the notion of quantum discord coincides with entanglement entropy. Interestingly, quantum advantage in communication tasks seems to require similar resources. Indeed, quantum nonlocality is known to be a resource for quantum communication [34], and discord in particular [35]. Relations between quantum computation and communication were studied as well [36–47].

In this work, we develop a framework for studying single-query quantum algorithms solving oracle classification tasks.

*Contact author: amit.teeni@biu.ac.il

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

We fix a probability distribution over the set of oracles that may appear in the black box. The oracle class—the correct solution to the classification problem—thus becomes a random variable. We measure the algorithm's performance via the mutual information between this solution and the algorithm's output. This perspective eliminates the need to consider any classical postprocessing performed after measuring the quantum state. We derive an upper bound on the mutual information via an analogy between our setup and quantum communication tasks.

Using our analogy, we prove a theorem that characterizes the optimal single-query algorithm. To do so, we consider the oracle as a separate subsystem in a classical mixture of all its possible states. We also find a lower bound on the mutual information, derived from a measure of quantum coherence. Finally, we study how several physical and information-theoretic quantities of interest vary throughout known algorithms.

A. Main contributions

The main contribution of this paper is a theorem that describes the optimal quantum algorithm for an arbitrary oracle classification task. Here, optimality is defined with respect to the mutual information between the correct result and the algorithm's output. As finding the optimal algorithm is generally difficult, we focus in this work on nonadaptive algorithms with a fixed number of queries. The optimal algorithm is given by solving a double optimization problem, which involves information-theoretic quantities such as Holevo's quantity, mutual information, and quantum discord, thereby highlighting their fundamental significance for quantum computation. Our results also demonstrate the utility of measuring algorithms' performance via mutual information.

While the primary focus of this paper is establishing the theoretical framework, we emphasize that this mutual-information perspective carries immediate practical utility for algorithmic design. Specifically, this framework has already been operationalized to significantly improve hybrid quantum-classical algorithms. In recent work [48], we applied these exact principles to quantum likelihood estimation (QLE) for Hamiltonian learning. By modeling each QLE iteration as a single-query oracle task and utilizing the optimality theorem developed here to maximize mutual information, we were able to systematically minimize quantum discord and substantially accelerate the algorithm's convergence. This demonstrates that reformulating oracle problems as communication tasks is not merely a conceptual exercise, but a concrete tool for optimizing NISQ-era algorithms.

The present work does not aim to establish a quantum query-complexity lower bound, a success-probability separation, or an efficient general-purpose method for solving the resulting optimization problem. Rather, its contribution is an information-theoretic characterization of fixed-query nonadaptive oracle algorithms, showing how optimal extraction of mutual information is tied to Holevo information, discord minimization, and coherence. This viewpoint is complementary to query-complexity approaches that optimize the number of queries or worst-case success probability.

B. Related work

Although independently derived, one may notice slight similarity between our approach and the one proposed by Buhrman *et al.* [36], in the sense that both constructively demonstrate the existence of a protocol for a quantum communication task, given a quantum algorithm for an oracle problem. However, our analogy always yields a very simple communication task, where Alice wishes to communicate a single message to Bob. Our analogy is used to find an upper bound on the mutual information obtainable in a single query. This is in contrast to existing works, where the total query complexity of the computational problem is related to the bit complexity of the communication task. Therefore, the analogy presented here yields insights regarding information-theoretic quantities, including quantum coherence and discord.

We note that the problem of finding optimal quantum query algorithms has been addressed from a different perspective by Belovs [49], who provided tight upper and lower bounds for adaptive query complexity (which encompasses the nonadaptive setting) by formulating the task as an exponentially large semidefinite program (SDP). While both the SDP approach and our framework ultimately lead to optimization problems that are generally difficult to solve for arbitrary tasks, their fundamental objectives and insights differ. Belovs's SDP is designed to minimize query complexity for a target success probability. In contrast, our approach fixes the query count and optimizes the extracted mutual information. Crucially, rather than yielding a purely mathematical program, our formulation analytically links computational optimality to physical, information-theoretic resources—specifically, the Holevo quantity, quantum discord, and relative entropy of coherence. This allows us not only to characterize the optimal algorithm but also to track the physical flow of information throughout the stages of the quantum circuit. Consequently, the two approaches answer different questions: Belovs's SDP gives a tight query-complexity characterization, whereas the present framework gives a physical/information-theoretic account of what information is stored, inaccessible, and finally extracted at a fixed number of nonadaptive queries.

Finally, regarding our representation of the oracle as a physical subsystem, our classical mixture construction is intimately related to an older one by Ambainis [50] (see also Ref. [51]). In that work, the oracle subsystem is in a coherent superposition of its possible states; indeed, our classical mixture can be viewed as the result of applying a dephasing channel to that pure state (with the oracle identity traced out). However, working explicitly with this classical-quantum mixed state is conceptually advantageous for our purposes, as it naturally allows the (nonminimized) quantum discord between the computer and oracle subsystems to play a key role in our characterization theorem.

C. Paper organization

The remainder of the paper is organized as follows. In Sec. II, we define the formal setting, detailing the oracle classification problem and the scope of the quantum algorithms considered. Section III outlines our main results: We establish our general information-theoretic framework, derive the optimality conditions for single-query algorithms, prove upper

and lower bounds on algorithmic performance, and demonstrate the direct application of these principles to optimizing hybrid quantum-classical schemes such as quantum likelihood estimation. In Sec. IV, we apply this framework to analyze the stage-by-stage information flow in several standard quantum algorithms, including the Deutsch-Jozsa, Bernstein-Vazirani, and phase estimation algorithms, and sketch the analysis for the Shor-Kitaev algorithm, using Simon's algorithm as an example. Finally, in Sec. V, we summarize our findings and discuss possible directions for future work.

II. PRELIMINARIES

In this section, we establish the formal mathematical setting for our study by defining the oracle classification problem and delineating the structure of the quantum algorithms under consideration.

A. Single-query oracle classification

Here, we introduce the setting for the type of computational problems we study, as well as the notation used later. Our definitions are fairly similar to the ones found in several papers [52–54]. We do not claim that the setting presented here is the most general or important one possible; however, we do provide justification for some of our choices here.

An oracle classification problem is described by the following data:

- (1) A set $\mathcal{F}$ of all allowed oracle “identities.”
- (2) A partition: $\mathcal{F} = \bigsqcup_{j \in \mathcal{J}} A_j$, where $\bigsqcup$ denotes a disjoint union and the sets A_j are all nonempty.
- (3) A protocol for oracle queries, in the form of unitary gates $\{U_f \mid f \in \mathcal{F}\}$.

Usually, the problem is, given an unknown oracle $f \in \mathcal{F}$, to find its class with respect to the partition. In other words, it is required to find the unique index $j \in \mathcal{J}$ for which $f \in A_j$, using as few oracle queries as possible. However, we consider a strongly relevant variation where the problem contains a fourth component:

- (4) A probability distribution $p_f := \Pr(F = f)$ on $\mathcal{F}$, i.e., a random variable F that obtains its values in $\mathcal{F}$.

Introducing a prior distribution is a fundamentally important variation; by Yao's minimax principle [55], the worst-case success probability of a quantum algorithm over all oracles is equal to its average success probability against an adversarially chosen distribution (see also Sec. 2 of the Supplemental Material [56]). While such probabilistic settings typically aim to maximize this average success probability using a single oracle query, our objective here is to maximize the mutual information between the true oracle class and the algorithm's output. As we demonstrate, optimizing this information-theoretic objective is intimately related to, yet distinct from, maximizing the probability of success.

Our guiding example would be the Deutsch-Jozsa- k problem [hereon denoted as $DJ(k)$], where:

- (1) $\mathcal{F}$ is the set of all functions $\{0, 1\}^k \rightarrow \{0, 1\}$, which are either constant or balanced;
- (2) $\mathcal{F}$ is partitioned as a disjoint union of the subsets A_b, A_c of balanced and constant functions, respectively;

- (3) the protocol is given by the unitary gates U_f , defined on computational basis states as

$$U_f|x\rangle \otimes |a\rangle = |x\rangle \otimes |a \oplus f(x)\rangle, \quad x \in \{0, 1\}^k, a \in \{0, 1\}, \quad (1)$$

where $\oplus$ denotes addition modulo 2.

As mentioned before, many problems [such as $DJ(k)$] are not readily equipped with a probability distribution. However, there are at least two “natural” choices: the uniform distribution over $\mathcal{F}$ and the uniform distribution with respect to the partition. The latter is defined explicitly by

$$p_f = \frac{1}{|\mathcal{J}||A_j|}, \quad f \in A_j. \quad (2)$$

This distribution ensures that f belongs to each of the classes A_j with equal probabilities and that the distribution within each class is uniform. Let J denote the random variable that keeps track of the class A_j ; i.e., $J := \tilde{j}(F)$, where $\tilde{j}$ is the function that assigns to each $f \in \mathcal{F}$ its corresponding class, i.e., the unique $j \in \mathcal{J}$ such that $f \in A_j$.

As an aside, note that in $DJ(k)$ we have $|\mathcal{J}| = 2$. Classification problems with this property are also known as oracle decision problems; and for such problems, one often considers $\mathcal{F}$ as a subset of $\{0, 1\}^N$ for some N and identifies $\mathcal{J}$ with $\{0, 1\}$. Solving the oracle problem is then referred to as computing the partial Boolean function $\tilde{j} : \mathcal{F} \rightarrow \{0, 1\}$. Single-query quantum algorithms for partial Boolean functions were studied extensively [57–62].

The polar opposite of decision problems is the oracle identification problem, where the task is to identify the unknown oracle. Of course, this is a special case of oracle classification, given by the partition of $\mathcal{F}$ into singletons. Some instances of oracle identification are well known, including unstructured search (solved by Grover's algorithm) and the Bernstein-Vazirani problem. Moreover, the query complexity of the general oracle identification problem was studied extensively [63–66], with focus on generic algorithms that solve all oracle identification problems parametrized by N and the size of the set $\mathcal{F} \subseteq \{0, 1\}^N$.

Further remarks regarding the problem formulation appear in Sec. 1 of the Supplemental Material [56].

B. Single-query algorithms

A general quantum algorithm for an oracle that uses only a single query has the following structure:

- (0) Initialize the quantum computer to be in some initial n -qubit state $|\psi_0\rangle$, where $n \geq m$.
- (1) Apply some unitary gate V .
- (2) Perform a single oracle query $U_f \otimes I$, where U_f acts on the first m qubits and the identity matrix acts on the remaining $n - m$ qubits.
- (3) Apply an additional unitary gate W .
- (4) Measure all or some of the n qubits in the computational basis; obtain a bit-string y (hereon referred to as the measurement outcome).
- (5) Based on the bit string y , provide an output $\hat{j} \in \mathcal{J}$. This chosen output is given by some function $g : \{0, 1\}^n \rightarrow \mathcal{J}$, i.e. $\hat{j} = g(y)$,

where the algorithm's performance is either the probability of finding the right class, i.e., $p_s := \Pr(\hat{j} = j)$; or the mutual information $I(J; Y)$ (see remarks below). By $I(J; Y)$, we mean the (Shannon) mutual information of the (classical) random variables J and Y , where Y is the random variable that obtains the values y in step 4 (a more explicit definition is given at the beginning of Sec. III).

Remarks:

(1) The number m is determined by the problem setting via the dimensionality of the oracle unitaries U_f . Using $n > m$ qubits effectively allows providing a mixed state as the input to the oracle, as well as implementing a general positive operator-valued measure (POVM) instead of a projective measurement in step 4. Hereon we mostly assume some fixed value of n and use U_f to mean $U_f \otimes I$, in a slight abuse of notation.

(2) For any fixed choice of an initial state and gates V, W , one may compute the conditional probability distribution $\Pr(J = j | Y = y)$ (where J is a random variable as well, whose distribution depends on the distribution we have chosen for F). It is clear that any algorithm that does not output in the final step the value of $\hat{j}$ with highest conditional probability (given y from the fifth step) can be improved. Thus, once we have decided upon the first five steps, there is always an obvious choice for the final one.

(3) The algorithm succeeds with probability 1, i.e., $p_s = 1$, if and only if $I(J; Y) = H(J)$, where $H(J)$ is the Shannon entropy of the random variable J .

(4) More generally, Fano's inequality provides a relation between the probability of success and the mutual information. Assuming that J has the uniform distribution, the probability of success p_s is bounded from above according to

$$p_s \leq \frac{I(J; Y) + 1}{\log_2 |\mathcal{J}|}. \quad (3)$$

Thus, low mutual information implies low probability of success. Conversely, mutual information also provides a lower bound on the probability of success:

$$p_s \geq 2^{I(J; Y) - \log_2 |\mathcal{J}|}. \quad (4)$$

(5) In what follows, we focus on the mutual information $I(J; Y)$ as our chosen measure of performance. Since step 5 does not affect the mutual information, we shall mostly ignore it from now on. For further discussion and computations regarding the probability of success, including a proof of Eq. (4), see Sec. 2 of the Supplemental Material [56].

Let us emphasize an important point: The problem setting already determines step 2 (up to choosing the total number of qubits n); step 4 is always the same if we measure all qubits, which we can do without losing generality (we may simply ignore some bits of the measured bit string); and for any choice of steps 0, 1, and 3, there is an obvious choice for step 5. Moreover, we can always choose $|\psi_0\rangle = |0\rangle^{\otimes n}$ and “absorb” the change in the initial state into V . In other words, the various possible “reasonable” quantum algorithms for a given single-query oracle problem essentially differ in the value of n and their pre- and postquery gates, V and W . In the next section, we show how to choose the optimal postquery unitary W as well, given fixed choices for n and V .

As an example, let us describe a quantum algorithm for the $DJ(k)$ problem; this is the Deutsch-Jozsa algorithm, denoted by $QDJ(k)$.

The Deutsch-Jozsa algorithm $QDJ(k)$

(0) Initialize the $(k + 1)$ -qubit state $|\psi_0\rangle = |0\rangle^k \otimes |1\rangle$.

(1) Apply the unitary gate $V = H^{\otimes(k+1)}$.

(2) Perform a single oracle query U_f .

(3) Apply an additional unitary gate, $W = H^{\otimes k} \otimes I$.

(4) Measure the first k qubits in the computational basis (equivalently, measure all of them but ignore the outcome of the final qubit).

(5) If the measurement result was $0 \cdots 0$ (i.e., the all-zeros bit string), the output is $\hat{j} = c$ (constant); otherwise, $\hat{j} = b$ (balanced).

$QDJ(k)$ is a deterministic quantum algorithm for the $DJ(k)$ problem that always succeeds. It is known that this problem does not admit any single-query classical algorithms that always succeed, meaning it has a quantum advantage.

C. Nonadaptive algorithms

A multiple-query algorithm for an oracle problem can be either adaptive or nonadaptive. Here, “adaptive” means that outcomes of queries can be used to compute the inputs for later queries. If an algorithm makes several oracle queries in a nonadaptive manner, then it can be converted to an algorithm that makes all its queries at the same time [67]. Thus, we define a nonadaptive t -query algorithm as follows:

(0) Initialize the quantum computer to be in some initial n -qubit state $|\psi_0\rangle$, where $n \geq mt$.

(1) Apply some unitary gate V .

(2) Perform t simultaneous oracle queries $U_f^{\otimes t} \otimes I$, where $U_f^{\otimes t}$ acts on the first mt qubits and the identity matrix acts on the remaining $n - mt$ qubits.

(3) Apply an additional unitary gate W .

(4) Measure all or some of the n qubits in the computational basis; obtain a bit-string y (hereon referred to as the measurement outcome).

(5) Based on the bit string y , provide an output $\hat{j} \in \mathcal{J}$. This chosen output is given by some function $g: \{0, 1\}^n \rightarrow \mathcal{J}$, i.e., $\hat{j} = g(y)$.

Note that any t -query nonadaptive algorithm can be thought of as single-query algorithm with a different (“stronger”) oracle, with $m' = mt$ and $U'_f = U_f^{\otimes t}$. Thus, we may treat it as a single-query algorithm, and by virtue of this observation the entire discussion would also apply to multiple-query nonadaptive algorithms.

III. THEORETICAL FRAMEWORK AND OPTIMIZATION

As stated before, the measurement outcome of a quantum algorithm is generally a random variable; it is determined by the outcome of measuring all or some of the qubits in the computational basis. Thus, Y is defined to be the outcome of measuring the final state of the computation with respect to the observable $Z^{\otimes k} \otimes I_2^{\otimes(n-k)}$, where n is the total number of qubits and $k \leq n$ is the number of measured qubits. Equivalently, we may take Y to be the outcome of measuring all qubits of a suitable reduced subsystem; i.e., we first trace out the nonmeasured qubits and then measure all remaining qubits

in the computational basis. In what follows, we mostly use the latter point of view.

For the moment, let us disregard the oracle classes A_j and assume that our task is finding the exact oracle identity f . This task can be reformulated as follows: An agent (the oracle) prepares the pure state $|\psi_{\text{fin}}(f)\rangle$ with probability p_f , and the other party is tasked with deducing the index f from a measurement on the following mixed state:

$$\rho_Y := \mathbb{E}_{f \in \mathcal{F}}[|\psi_{\text{fin}}(f)\rangle\langle\psi_{\text{fin}}(f)|] = \sum_{f \in \mathcal{F}} p_f T_f |\psi_0\rangle\langle\psi_0| T_f^\dagger, \quad (5)$$

where $T_f = W U_f V$ is the total quantum circuit and $|\psi_{\text{fin}}(f)\rangle = T_f |\psi_0\rangle$ is the final state of the computation (before measurement). Thus, the state (5) defines precisely the type of mixture for which Holevo's bound applies. The subscript Y for ρ_Y is used to clarify that this state acts on a Hilbert space with a distinguished orthonormal basis (i.e., the computational basis), whose elements correspond to the values of the random variable Y . Thus, ρ_Y defines the probability distribution of the random variable Y , $\Pr(Y = y) := \langle y | \rho_Y | y \rangle$.

Recall that Holevo's bound deals with the following communication task: Alice randomly generates a letter x from some finite alphabet $\mathcal{X}$ and then prepares a quantum state ρ_x corresponding to that letter. She then sends the state to Bob, so he obtains the mixture $\rho = \sum_{x \in \mathcal{X}} p_x \rho_x$. Bob measures ρ in some basis and then uses the outcome to deduce the letter. Holevo's bound provides an upper bound on the mutual information between the random letter X and the measurement outcome Y (here, X, Y are the corresponding classical random variables):

$$I(X; Y) \leq S(\rho) - \sum_x p_x S(\rho_x). \quad (6)$$

The right-hand side is called the Holevo quantity, denoted by $\chi(\{p_x, \rho_x\}_{x \in \mathcal{X}})$.

In our setting, the unknown oracle identity f (and later the property j) is the random letter, and the quantum circuit plays the role of Alice, preparing a state based on f . The postquery unitary W chooses the measurement basis, and Y is the outcome of this final measurement. Thus, we obtain an upper bound on the mutual information, $I(F; Y) \leq S(\rho_Y)$. Note that if the first k qubits are measured then we can redefine ρ_Y as the appropriate reduced state of Eq. (5), and the inequality $I(F; Y) \leq S(\rho_Y)$ still holds.

Now, we may derive a more general result, corresponding to our actual task of finding the oracle class A_j . We rewrite ρ_Y as follows:

$$\rho_Y = \sum_{j \in \mathcal{J}} p_j \sigma_j, \quad \sigma_j := \frac{1}{p_j} \sum_{f \in A_j} p_f T_f |\psi_0\rangle\langle\psi_0| T_f^\dagger, \quad (7)$$

where $p_j := \sum_{f \in A_j} p_f$. We still have two agents, where one wishes to communicate a letter of the alphabet $\mathcal{J}$ to the other. The first agent prepares the mixed state σ_j with probability p_j and sends the mixture ρ_Y to Bob. Bob then measures ρ_Y in the computational basis, resulting in the bit string Y . Holevo's bound then provides an upper bound on the mutual

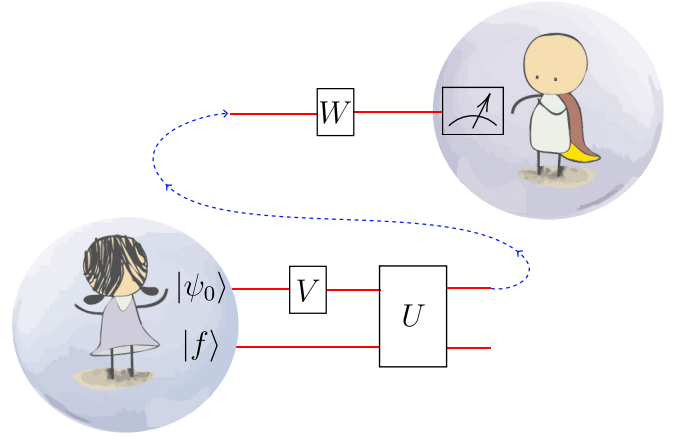


FIG. 1. The analogy with quantum communication: First, Alice initializes the quantum computer to be in the state $|\psi_0\rangle$ and chooses an oracle f . Then, she performs the first unitary gate V on the computer, followed by the controlled gate U . As a result, the state of the computer is $U_f V |\psi_0\rangle$, which she subsequently sends to Bob. Bob is then tasked with choosing a gate W to perform on the state, before measuring it in the computational basis, such that he would optimize the information he obtains on Alice's chosen oracle f or its class j .

information:

$$I(J; Y) \leq \chi(\{p_j, \sigma_j\}_{j \in \mathcal{J}}) = S(\rho_Y) - \sum_{j \in \mathcal{J}} p_j S(\sigma_j). \quad (8)$$

Hereon, we write χ for $\chi(\{p_j, \sigma_j\}_{j \in \mathcal{J}})$. Again, we can replace all states σ_j with their reduced states corresponding to the measured registers, and the inequality still holds.

Note that there is another—arguably more illuminating—way of delineating between the “Alice” and “Bob” roles in the algorithm. One could say that Alice's preparation ends immediately after the oracle query; then, Bob obtains an intermediate state, where T_f in Eq. (7) is replaced by $U_f V$ (this intermediate state is denoted ρ_Y^2 in the next section). Then Bob must choose the measurement basis that would allow him to extract as much information as possible about J ; effectively, he applies a corresponding unitary W and then measures in the computational basis (see Fig. 1). This point of view may help motivate the next subsection.

A. Quantum discord and optimal quantum algorithms

It turns out that a simple construction not only yields another way to derive Holevo's inequality, but also sheds light on the quantum informational structure of our setting. The idea is as follows: We wish to describe the oracle as a quantum subsystem, in addition to the quantum computer. Moreover, we want the oracle subsystem to be constructed in a manner that would allow us later to “forget” about the specific oracle identities $f \in \mathcal{F}$, and retain only information regarding the oracle classes A_j . Note that this construction—as well as the results obtained hereafter—holds a striking resemblance to a construction performed in the context of Quantum Darwinism [68]. Let $\mathcal{H} = \mathcal{H}_J \otimes \mathcal{H}_F \otimes \mathcal{H}_Y$ be the full Hilbert space,

where

$$\dim \mathcal{H}_J = |\mathcal{J}|, \quad \dim \mathcal{H}_F = |\mathcal{F}|, \quad \dim \mathcal{H}_Y = N = 2^n. \quad (9)$$

Let $\{|j\rangle\}_{j \in \mathcal{J}}$ be an orthonormal basis for $\mathcal{H}_J$ indexed by the set of oracle classes, and let $\{|f\rangle\}_{f \in \mathcal{F}}$ be an orthonormal basis for $\mathcal{H}_F$. This notation for the Hilbert spaces, as well as the corresponding notation for states on those Hilbert spaces (see below), follows the convention we have established earlier: A Hilbert space with subscript X has a distinguished orthonormal basis corresponding to the domain of the random variable X . Note this is a formal construction, where the states $|j\rangle, |f\rangle$ are not implied to represent an actual physical system. Now, define the classical-classical-classical state ρ_{JFY}^0 :

$$\rho_{JFY}^0 := \sum_{j \in \mathcal{J}} \sum_{f \in \mathcal{F}} p_f |j\rangle\langle j| \otimes |f\rangle\langle f| \otimes |\psi_0\rangle\langle\psi_0|. \quad (10)$$

We say that a state is “classical-classical-classical” if it can be written as a mixture of orthonormal basis elements on all three subsystems (generalizing the more common notion of classical-classical states). In our case, ρ_{JFY}^0 consists of three subsystems, which represent the oracle class J , the oracle identity F , and the quantum computer Y in its initial state, $|\psi_0\rangle$. This state evolves with every step of the quantum algorithm. After applying the unitary V , we still have a classical-classical-classical state:

$$\rho_{JFY}^1 = \sum_{j \in \mathcal{J}} \sum_{f \in \mathcal{F}} p_f |j\rangle\langle j| \otimes |f\rangle\langle f| \otimes V |\psi_0\rangle\langle\psi_0| V^\dagger. \quad (11)$$

However, the oracle query U_f generates correlations between the “oracle subsystem” and the quantum computer:

$$\rho_{JFY}^2 = \sum_{j \in \mathcal{J}} \sum_{f \in \mathcal{F}} p_f |j\rangle\langle j| \otimes |f\rangle\langle f| \otimes U_f V |\psi_0\rangle\langle\psi_0| V^\dagger U_f^\dagger. \quad (12)$$

Note that the above state is not necessarily classical-classical-classical, since the states $\{U_f V |\psi_0\rangle\}_{f \in \mathcal{F}}$ may not be orthogonal. Finally, after application of the second unitary W , we obtain the state of the system right before the measurement:

$$\rho_{JFY}^{\text{fin}} = \sum_{j \in \mathcal{J}} \sum_{f \in \mathcal{F}} p_f |j\rangle\langle j| \otimes |f\rangle\langle f| \otimes T_f |\psi_0\rangle\langle\psi_0| T_f^\dagger. \quad (13)$$

Since we do not care about the oracle identity f , we may trace out the subsystem F , to obtain

$$\begin{aligned} \rho_{JY} &= \sum_{j \in \mathcal{J}} \sum_{f \in \mathcal{F}} p_f |j\rangle\langle j| \otimes T_f |\psi_0\rangle\langle\psi_0| T_f^\dagger \\ &= \sum_{j \in \mathcal{J}} p_j |j\rangle\langle j| \otimes \sigma_j. \end{aligned} \quad (14)$$

Note that ρ_Y defined before is now simply given by taking the partial trace of ρ_{JY} with respect to the subsystem J . At this point we may also trace out the nonmeasured qubits.

Now, let us compute the nonoptimized quantum discord of the state ρ_{JY} with respect to measurement of subsystem Y in the computational basis. This is defined as follows:

$$D_Y(\rho_{JY}; Z^{\otimes n}) = S(\rho_Y) - S(\rho_{JY}) + S(\rho_J | Z^{\otimes n}), \quad (15)$$

where $\rho_J = \sum_{j \in \mathcal{J}} p_j |j\rangle\langle j|$ remains unchanged throughout the run of the algorithm. We have the following result (see Sec. 3 of the Supplemental Material [56] for the proof):

$$D_Y(\rho_{JY}; Z^{\otimes n}) = \chi - I(J; Y), \quad (16)$$

which is non-negative, as quantum discord always is. In fact, we have found here that the non-negativity of $D_Y(\rho_{JY}; Z^{\otimes n})$ is equivalent to Holevo’s bound. Note that this result was recently derived independently in the context of Bayesian metrology [69], although the nonoptimized discord appears there under the name ensemble relative entropy of coherence.

However, the more commonly used definition of quantum discord does not fix the measurement basis, but rather it optimizes over all possible measurements:

$$D_Y(\rho_{JY}) := \min_{\Pi} D_Y(\rho_{JY}; \Pi), \quad (17)$$

where $\Pi = \{\Pi_k\}_k$ is a projective measurement. A quantum algorithm can be optimal only if it applies unitary gates such that the computational basis is the optimal one, i.e., it yields the minimal oracle-outcome discord (and thus the mutual information obtained is as high as possible). Here, one should note that an f -independent unitary gate cannot change χ , but it can increase (or decrease) $I(J; Y)$. The source of information (about the oracle) in the quantum computer, χ , as one should expect, stems exclusively from oracle queries.

Let us reiterate the crucial point here. In what follows, $S(\rho_Y^i)$ denotes the von Neumann entropy of the subsystem $\rho_Y^i = \text{Tr}_{J,F,N}(\rho_{JFY}^i)$, where ρ_{JFY}^i is the entire system right after the i th step of the algorithm and N denotes the non-measured qubits; $I(J; Y_i)$ should be understood as the mutual information we would have obtained if we had measured in the computational basis right after the i th step of the algorithm; and $|\psi_1\rangle := V |\psi_0\rangle$. By changing our chosen V , we can set $|\psi_1\rangle$ to be any normalized pure state. Then, for such $|\psi_1\rangle$, we get the following state after the third step:

$$\rho_{JY}^2(|\psi_1\rangle) = \sum_{j \in \mathcal{J}} p_j |j\rangle\langle j| \otimes \left(\sum_{f \in \mathcal{F}} \frac{p_f}{p_j} U_f |\psi_1\rangle\langle\psi_1| U_f^\dagger \right), \quad (18)$$

where the expression in the parentheses is denoted by σ_j^2 . Hence, the oracle acts as a map that produces an ensemble $\{\sigma_j^2\}$ out of the pure state $|\psi_1\rangle$. In step (3), the best we can do is choose a unitary W such that the measurement in the computational basis would be the one with minimal discord. Thus, for any fixed $|\psi_1\rangle$, we have

$$I_{\max}(|\psi_1\rangle) = \chi(\{p_j, \sigma_j^2\}_{j \in \mathcal{J}}) - D_Y(\rho_{JY}^2), \quad (19)$$

where $I_{\max}(|\psi_1\rangle)$ denotes the maximally obtainable $I(J; Y)$ if we have the state $|\psi_1\rangle$ after the second step. Note that here $D_Y(\rho_{JY}^2)$ is the minimal discord.

Our analysis can be summarized in the following theorem:

Theorem. For any oracle problem of the form defined in Sec. II A and any fixed value of $n \geq m$, a single-query quantum algorithm of the form defined in Sec. II B that uses n qubits obtains the maximal value of $I(J; Y)$ out of all such algorithms (i.e., is optimal) if and only if the following

conditions are both satisfied:

- (1) the prequery unitary V has the property

$$I_{\max}(V|\psi_0\rangle) = \max_{|\psi_1\rangle} I_{\max}(|\psi_1\rangle),$$

where the maximization is over all n -qubit states $|\psi_1\rangle$;

- (2) the postquery unitary W has the property that $W^\dagger$ maps the computational basis to a basis which yields the minimal discord.

Note there is a single-query quantum algorithm that succeeds with probability 1 if and only if there exists a state $|\psi_1\rangle$ such that $I_{\max}(|\psi_1\rangle) = H(J)$. Since $\chi(\rho_{JY}^2) \leq H(J)$, this state must satisfy the following two conditions simultaneously:

- (1) $\chi(\rho_{JY}^2)$ obtains its maximal bound $H(J)$. This is an interesting condition, as there exist finer bounds that use properties of σ_j^2 's (see Refs. [70,71]). In particular, this occurs if and only if the states σ_j^2 have orthogonal support [72]. Since $\{\sigma_j^2\}$ are positive semidefinite, having orthogonal support is equivalent to orthogonality: $\forall j \neq j' \in \mathcal{J}, \text{Tr}(\sigma_j^2 \sigma_{j'}^2) = 0$.

- (2) $D_Y(\rho_{JY}^2) = 0$, i.e., the state ρ_{JY}^2 is not only classical-quantum, but also quantum-classical. This implies the state is classical-classical, and this scenario occurs if and only if the states σ_j^2 pairwise commute.

For any two positive semidefinite matrices ρ and σ , we have that $\text{Tr}(\rho\sigma) = 0$ implies $\rho\sigma = \sigma\rho = 0$ and, in particular, $[\rho, \sigma] = 0$. Thus, the first condition implies the second one (but not the other way around).

Indeed, if the first condition holds, then a simple way to construct W would be to map the orthogonal support subspaces to the computational basis. Further details and proof of optimality appear in Sec. 4 of the Supplemental Material [56]. Moreover, even if the first condition does not hold but the second one does, the optimal W (for fixed V) is one that simultaneously diagonalizes all of the states σ_j^2 . This is evident from the observations above since $I(J; Y)$ obtains its upper bound χ , which cannot be changed by W . This fact remains true even if we are allowed to use arbitrary POVMs instead of projective measurements, since Holevo's bound is true for POVMs as well.

Remark. Suppose V is a monomial matrix, i.e., having exactly one nonzero entry in each row and column. Such a matrix cannot generate superpositions: It maps computational basis states to computational basis states multiplied by some phase. Usually, U_f has the stronger property of being a permutation matrix; this assumption is necessary if we think of computational basis states as being classical. In this scenario (i.e., if V is monomial and U_f is a permutation matrix), the matrices σ_j^2 are all diagonal in the computational basis, hence commuting. By the above discussion, $W = I$ is optimal, and since the phase is insignificant, such an algorithm can be executed by a classical computer. These facts can be interpreted as saying that if the oracle query is performed classically, then quantum postprocessing cannot be used to increase the amount of available classical information. Another consequence is that a quantum algorithm with monomial V can never outperform the optimal classical algorithm. Indeed, in all of the quantum algorithms analyzed as examples in Sec. IV, V is a generalized complex Hadamard matrix, i.e., a unitary matrix whose entries all have the same abso-

lute value. Such a matrix is as far from being monomial as possible.

To conclude this subsection, we note that the above theorem neatly characterizes optimal single-query quantum algorithms, and via the correspondence described in Sec. IIC it also applies to multiple-query nonadaptive algorithms with a fixed number of queries. However, it may be challenging to find the optimal $|\psi_1\rangle$ in practice for a general oracle problem (even for single-query algorithms).

Because of this practical difficulty, the theorem as stated is largely conceptual. It reveals that extracting maximal information is mathematically equivalent to minimizing quantum discord, thus translating an abstract computational condition into the language of fundamental physical resources. Thus, the theorem should be read as a structural characterization of optimal information extraction, not as an efficient algorithm for constructing the optimizer in arbitrary oracle problems.

While this conceptual lens does not directly yield standalone single-query algorithms, the underlying mathematical framework carries significant practical utility. As we discuss in Sec. IIIC, our mutual-information perspective provides the exact objective function needed to optimize iterative subroutines in hybrid quantum-classical algorithms, where accumulating partial information over multiple steps is more relevant than achieving a high single-shot success probability.

B. Coherence and a lower bound on the mutual information

Recall that $I(J; Y) = H(Y) - H(Y|J)$. A simple observation (see Sec. 5 of the Supplemental Material [56]) shows

$$H(Y) \geq S(\rho_Y), \quad (20)$$

where the left-hand side is the Shannon entropy of the bit string obtained by measuring the first k qubits, and the right-hand side is the von Neumann entropy of the reduced state obtained by tracing out the $n - k$ nonmeasured qubits. Moreover, the difference between the former and the latter is known to equal the relative entropy of coherence [73,74]:

$$C(\rho_Y) = H(Y) - S(\rho_Y) \geq 0. \quad (21)$$

Thus, just as the nonoptimized discord measures the “failure” of $I(J; Y)$ to obtain its upper bound, so does the relative entropy of coherence measure its failure to obtain the lower bound. Writing the lower and upper bounds together, we have

$$S(\rho_Y) - H(Y|J) \leq I(J; Y) \leq S(\rho_Y) - \sum_{j \in \mathcal{J}} p_j S(\sigma_j). \quad (22)$$

The sum $\mathfrak{S}_{Z^{\otimes n}}(\rho_{JY}) := C(\rho_Y) + D_Y(\rho_{JY}; Z^{\otimes n})$ is known as the irre realism [75–78] of measuring the Y subsystem in the computational basis. Note that $\mathfrak{S}_{Z^{\otimes n}}(\rho_{JY})$ vanishes if and only if the lower and upper bounds in Eq. (22) are both saturated.

C. Application to hybrid quantum-classical schemes

In separate work, some of us used the mutual-information objective developed here to optimize the information gain in QLE. We include this connection to illustrate how the present formalism can guide objective-function design in iterative settings, while the algorithmic results can be found in Ref. [48].

Consider the task of learning an unknown Hamiltonian H_{true} from a finite set of candidates $H_1, H_2, \dots, H_N$. The learner can access H_{true} only through its unitary evolution, $U(t) = e^{-itH_{\text{true}}/\hbar}$. A hybrid quantum-classical algorithm alternates between quantum and classical stages. Each quantum stage consists of a circuit involving a single application of $U(t_i)$ for some evolution time t_i , which may vary between iterations. Between successive quantum stages, classical processing uses the measurement outcomes to update the stored information about H_{true} and to determine the parameters of the next quantum stage. After a finite number of iterations, the algorithm outputs an estimate of H_{true} .

QLE provides a concrete example of such a hybrid scheme [79–84]. In its standard form, QLE maintains a probability distribution representing the information accumulated so far about H_{true} . However, only the evolution time t_i is adaptively chosen between iterations, while the pre- and postevolution unitaries are kept fixed throughout. In Ref. [48], it was shown that each quantum iteration can be modeled as a single-query quantum algorithm, and the results developed here were applied to optimize the information gain about H_{true} at each step. The resulting optimized version of QLE achieves substantially faster convergence, requiring far fewer iterations to learn H_{true} than the original algorithm. This shows that the proposed framework could have some practical ramifications in addition to its conceptual contributions.

IV. INFORMATION-THEORETIC ANALYSIS OF STANDARD QUANTUM ALGORITHMS

In this section, we demonstrate our approach by studying several examples of quantum algorithms. The information-theoretic figures we have discussed are computed and presented in detail, using a combination of analytic and numerical methods. Specifically, while we provide exact analytical derivations wherever possible, we utilize numerical simulations for cases where closed-form expressions for the subsystem entropies are mathematically intractable. These numerics are intended to empirically illustrate the dynamics of discord and coherence predicted by our framework, demonstrating that physical resources can be tracked even when exact analytical formulas evade us.

A. Simulation results for Deutsch-Jozsa

Starting with the Deutsch-Jozsa algorithm, we have run simulations for $k \leq 4$, where k denotes the number of input bits for the oracle functions (which is also the number of measured qubits). In particular, the simulation (a MATLAB script) goes over all possible oracle functions and emulates the algorithm's run for each, computing the state of the quantum computer (prequery, postquery, and final) and the probabilities for every possible measurement outcome. The states are mixed together to find ρ_Y^i , the probabilities are used to find the probability distribution of Y conditioned on J , and these are then used to calculate the required information-theoretic quantities. The results are displayed in Table I. Note that we have taken the “partition-uniform” probability distribution over the oracles.

TABLE I. Information-theoretic quantities for the Deutsch-Jozsa algorithm, found by simulating the algorithm's run over all possible inputs for $k \leq 4$. The three parts of the table depict quantities obtained by measuring the output qubits right before the oracle query (first row), immediately after the query (second row), or at the end of the algorithm (third row). Here, $D_Y^i := D_Y(\rho_{JY}^i; Z^{\otimes n})$ for $i = 1, 2$, and $D_Y := D_Y(\rho_{JY}; Z^{\otimes n})$.

k	$H(Y_1)$	$S(\rho_Y^1)$	$C(\rho_Y^1)$	$H(Y_1 J)$	χ_1	$I(J; Y_1)$	D_Y^1
1	1	0	1	1	0	0	0
2	2	0	2	2	0	0	0
3	3	0	3	3	0	0	0
4	4	0	4	4	0	0	0

k	$H(Y_2)$	$S(\rho_Y^2)$	$C(\rho_Y^2)$	$H(Y_2 J)$	χ_2	$I(J; Y_2)$	D_Y^2
1	1	1	0	1	1	0	1
2	2	1.7925	0.2075	2	1	0	1
3	3	2.4037	0.5963	3	1	0	1
4	4	2.9534	1.0466	4	1	0	1

k	$H(Y)$	$S(\rho_Y)$	$C(\rho_Y)$	$H(Y J)$	χ	$I(J; Y)$	D_Y
1	1	1	0	0	1	1	0
2	1.7925	1.7925	0	0.7925	1	1	0
3	2.4037	2.4037	0	1.4037	1	1	0
4	2.9534	2.9534	0	1.9534	1	1	0

As seen above, the simulations demonstrate that the discord $D_Y(\rho_{JY}; Z^{\otimes n})$ vanishes for Deutsch-Jozsa. For example, with $k = 4$ we obtain $S(\rho_Y) \approx 2.95$, while for the balanced and constant functions, respectively, we have $S(\rho_Y^b) \approx 3.91$ and $S(\rho_Y^c) = 0$. Since the Deutsch-Jozsa algorithm always succeeds, we have $I(J; Y) = 1$. Hence, the discord is $2.95 - \frac{1}{2} \cdot 3.91 - \frac{1}{2} \cdot 0 = 0$.

In fact, a closer look into the DJ algorithm reveals the following interesting structure. Upon initialization, the von Neumann entropy and mutual information are both zero, and this is still the case after applying V . After the oracle query, the von Neumann entropies of ρ_Y, σ_j already obtain their final values, i.e., $S(\rho_Y^2) = S(\rho_Y^{\text{fin}})$ and $S(\sigma_j^2) = S(\sigma_j^{\text{fin}})$. Thus, χ obtains its final value, which equals $H(J)$. $I(J; Y_2)$ obtains some value ϵ . After application of the second unitary W and before the measurement, the mutual information $I(J; Y)$ obtains its maximal possible value, i.e., $I(J; Y) = \chi = H(J)$, such that the discord takes its minimal value, which is zero.

B. Information-theoretic quantities for Bernstein-Vazirani

While the Bernstein-Vazirani algorithm itself is essentially identical to Deutsch-Jozsa, the promise on the function f is different, as well as the partition $\mathcal{J}$. In particular, we are promised that f is of the form $f_a(x) := a \cdot x$, where $a \in \{0, 1\}^n$. The objective is to determine the parameter a of f_a exactly, i.e., the classes are singletons; hence, $J = F$. These properties make this algorithm particularly well behaved, allowing us to compute the information-theoretic quantities analytically. The findings are summarized in Table II. Note that since σ_j are all pure states, we have $\chi = S(\rho_Y)$. As in the Deutsch-Jozsa algorithm, one can observe that the coherence $C(\rho_Y)$ and conditional entropy $H(Y|J)$ are initially large but

TABLE II. Information-theoretic quantities for the Bernstein-Vazirani algorithm, computed analytically for any value of n .

Quantity	Prequery	Postquery	Final
$H(Y)$	n	n	n
$S(\rho_Y)$	0	n	n
$C(\rho_Y)$	n	n	0
$H(Y F)$	n	n	0
$I(F; Y)$	0	0	n
$D_Y(\rho_{FY}; \mathbb{Z}^{\otimes n})$	0	n	0

are decreased by the second unitary W . In contrast, the von Neumann entropy and discord have initial values zero but are increased to n by the oracle query. The second unitary W can be said to “trade off” the discord in exchange for mutual information, which obtains the maximal value n in the final state. The detailed computation of these quantities can be found in Sec. 6 of the Supplemental Material [56].

C. The Shor-Kitaev algorithm

Suppose we are given a group G , a set X , and an unknown function $f : G \rightarrow X$ that hides a subgroup $H \subset G$. That is, we have the following commutative triangle:

$$\begin{array}{ccc} G & \xrightarrow{f} & X \\ & \searrow \pi & \nearrow \\ & G/H & \end{array} \quad (23)$$

In this case, f is the composition of the canonical projection $\pi : G \twoheadrightarrow G/H$ with any one-to-one map (of sets) $G/H \hookrightarrow X$ and we say that f hides the subgroup H . We are given oracle access to the function f , and the task is to find the subgroup H using as few queries as possible. This framework, known as the hidden subgroup problem (HSP), encompasses several well-known problems, including Simon’s problem, the discrete logarithm, and period and order finding. If G is finite and Abelian, the Shor-Kitaev algorithm solves the HSP efficiently using a quantum Fourier transform [72]. In fact, in the case of order finding, this algorithm forms the quantum subroutine in Shor’s prime factorization algorithm. Since this generic algorithm is nonadaptive, our framework applies. In fact, it is not even a proper multiple-query algorithm; rather it is the same single-query algorithm executed several times to learn the probability distribution of Y with greater accuracy.

We now describe the single-query routine. Let $\mathcal{F}$ be the set of allowed functions f , and $\mathcal{J}$ is the partition of $\mathcal{F}$ into classes such that functions that hide the same subgroup H are in the same class. Thus, the indices $j \in \mathcal{J}$ correspond to subgroups $H_j \subset G$. Assuming the standard “bit oracle” for the function f , the quantum computer should consist of two registers that correspond to elements of G and X , respectively. The joint state of the oracle and computer at each step i is of the form

$$\rho_{j_{FY}}^i = \sum_{j \in \mathcal{J}} \sum_{f \in A_j} p_f |j\rangle\langle j| \otimes |f\rangle\langle f| \otimes |\psi_i(f)\rangle\langle \psi_i(f)|, \quad (24)$$

where A_j is the set of all functions that hide the subgroup H_j of G and $|\psi_i(f)\rangle$ is the state of the computer at step i . In particular, right before the query, we have

$$|\psi_1\rangle = \frac{1}{|G|^{1/2}} \sum_{g \in G} |g\rangle|0\rangle. \quad (25)$$

After the query,

$$|\psi_2(f)\rangle = \frac{1}{|G|^{1/2}} \sum_{g \in G} |g\rangle|f(g)\rangle. \quad (26)$$

Choosing a representative $s \in gH_j$ for each coset $gH_j \in G/H_j$, we can write

$$|\psi_2(f)\rangle = \frac{1}{|G|^{1/2}} \sum_{s \in G/H_j} \left(\sum_{h \in H_j} |s+h\rangle \right) \otimes |f(s)\rangle \quad (27)$$

using the fact that $f(g)$ is constant on cosets. The usual description of the algorithm states that here the second register should be measured and discarded (and the measurement outcome is not used). By the principle of deferred measurement, we can move this measurement to the end of the algorithm and obtain the exact same action of the circuit.

Finally, we perform a quantum Fourier transform on the first register. This has the following effect:

$$\frac{1}{|H_j|^{1/2}} \sum_{h \in H_j} |s+h\rangle \mapsto \sqrt{\frac{|H_j|}{|G|}} \sum_{\chi_g \in H_j^\perp} \chi_g(s) |g\rangle, \quad (28)$$

where $\chi_g \in \hat{G} = \text{Hom}(G, \text{U}(1))$ is a character of $G = \prod_{i=1}^M (\mathbb{Z}/p_i\mathbb{Z})$ defined by

$$\chi_g(l_1, \dots, l_M) := \prod_i e^{2\pi i l_i g_i / p_i} \quad \text{for } g = (g_1, \dots, g_M), \quad (29)$$

and $H_j^\perp$ is the subgroup of $\hat{G}$ comprising the characters χ such that $\text{Res}_{H_j}^G(\chi) = \chi_e$, i.e., the characters that restrict to the trivial character on H_j . So, we have

$$|\psi_{\text{fin}}(f)\rangle = \frac{\sqrt{|H_j|}}{|G|} \sum_{s \in G/H_j} \left(\sum_{\chi_g \in H_j^\perp} \chi_g(s) |g\rangle \right) \otimes |f(s)\rangle. \quad (30)$$

Now, we wish to analyze the information-theoretic quantities. The random variable Y refers to the outcome of measuring only the first register in the computational basis (i.e., the one corresponding to elements of G). The pertinent matrices ρ_Y^i are

$$\rho_Y^i = \sum_{j \in \mathcal{J}} \sum_{f \in A_j} p_f \text{Tr}_O(|\psi_i(f)\rangle\langle \psi_i(f)|), \quad (31)$$

where Tr_O denotes partial trace over the output register. Let us perform the computation for each of the three steps.

1. Prequery

After discarding the output register, the state of the computer right before the query is a “democratic superposition”, hence

$$C(\rho_Y^1) = H(Y_1|J) = H(Y_1) = \log_2 |G| \quad (32)$$

and of course, we have

$$S(\rho_Y^1) = \chi_1 = I(J; Y) = D_Y(\rho_{JY}; G) = 0, \quad (33)$$

where $D_Y(\rho_{JY}; G)$ is the discord of ρ_{JY} with respect to measuring the first register of the computer in the computational basis, which here corresponds to the group G .

2. Postquery

The query does not change the diagonal entries of σ_j^2 , so we have $H(Y_2|J) = H(Y_2) = \log_2 |G|$ and $I(J; Y_2) = 0$. It turns out that for all $j, j' \in \mathcal{J}$, we have $[\sigma_j^2, \sigma_{j'}^2] = 0$, while $\text{Tr}(\sigma_j^2 \sigma_{j'}^2) \neq 0$; i.e., out of the two conditions for optimality described in Sec. III A, the second condition is satisfied while the first one is not. This means that the minimized discord $D_Y(\rho_{JY}^2)$ vanishes. It also means that the matrices σ_j^2 and ρ_Y^2 are all mutually diagonalizable; in fact, their common eigenbasis is $\{|v_\chi\rangle\}_{\chi \in \hat{G}}$, which are the column vectors of the inverse quantum Fourier transform over G (equivalently—the complex conjugates of the rows of the character table of G). This is also the basis in which the minimal discord (i.e., zero) is obtained, and indeed the final unitary W is the quantum Fourier transform, as we would have expected from a “good” algorithm.

These insights allow us to proceed with relative ease and generality. We find that $S(\sigma_j^2) = \log_2 |G| - \log_2 |H_j|$. All other quantities depend on finding $S(\rho_Y^2)$: We have $C(\rho_Y^2) = \log_2 |G| - S(\rho_Y^2)$ and $D_Y(\rho_{JY}^2; G) = \chi_2(\{\sigma_j^2; p_j\}) = S(\rho_Y^2) - \log_2 |G| + \sum_{j \in \mathcal{J}} p_j \log_2 |H_j|$. Unfortunately, there seems to be no concise expression for $S(\rho_Y^2)$; the best we can do here is provide expressions for the eigenvalues of ρ_Y^2 :

$$\lambda_\chi = \frac{1}{|G|} \sum_{j \in \mathcal{J} : \chi \in H_j^\perp} p_j |H_j|. \quad (34)$$

If J is uniformly random, we can also write them down as

$$\lambda_\chi = \frac{1}{\sqrt{|G||\mathcal{J}|}} \langle c | v_\chi \rangle, \quad (35)$$

where $|c\rangle = \sum_{g \in G} c_g |g\rangle$ and c_g is defined to be the number of subgroups H_j of which g is an element. This way, the eigenvalues can be computed directly from the character table by multiplying it from the left with a row vector. For proofs and further details, see Sec. 7 of the Supplemental Material [56].

3. Final state (premeasurement)

As mentioned before, after the quantum Fourier transform the matrices σ_j are all diagonal, hence so is ρ_Y . Thus, we have $C(\rho_Y) = 0$ and $H(Y) = S(\rho_Y) = S(\rho_Y^2)$, and the latter can be computed using the eigenvalues (34). We may easily compute

$$H(Y|J) = \sum_{j \in \mathcal{J}} p_j S(\sigma_j) = \log_2 |G| - \sum_{j \in \mathcal{J}} p_j \log_2 |H_j|. \quad (36)$$

Moreover, $S(\sigma_j) = S(\sigma_j^2) = \log_2 |G| - \log_2 |H_j|$ and $\chi(\{\sigma_j; p_j\}) = \chi(\{\sigma_j^2; p_j\})$. Since $D_Y(\rho_{JY}; G) = D_Y(\rho_{JY}^2) = 0$, we also have $I(J; Y) = \chi(\{\sigma_j; p_j\})$. Note that $I(J; Y)$ obtains both its lower and upper bounds. However, $I(J; Y) < H(J)$ since the σ_j do not have orthogonal support.

To conclude this subsection, we discuss the information-theoretic quantities obtained when performing t queries. We only do this for the final stage of the algorithm. We thus consider $(\sigma_j)^{\otimes t}$ instead of σ_j and define $\rho_Y^{(t)} := \sum_{j \in \mathcal{J}} p_j (\sigma_j)^{\otimes t}$. These matrices are still diagonal in the tensor product basis, so $C(\rho_Y^{(t)}) = 0$. Let Y_t denote the outcome of measuring all t input registers at the end of the algorithm. As it turns out, the conditional entropy is simply multiplied by t , i.e., $H(Y_t|J) = t \cdot H(Y|J)$. Describing $H(Y_t) = S(\rho_Y^{(t)})$ is slightly more complicated than before. The tensor product basis is indexed by tuples of t irreducible representations of G , and we have

$$\lambda(\chi_1, \dots, \chi_t) = \sum_{j \in \mathcal{J} : \chi_1, \dots, \chi_t \in H_j^\perp} p_j \left(\frac{|H_j|}{|G|} \right)^t. \quad (37)$$

As in the single-query case, we have vanishing discord, implying

$$I(J; Y_t) = S(\rho_Y^{(t)}) - t \cdot H(Y|J) = \chi(\{(\sigma_j)^{\otimes t}; p_j\}). \quad (38)$$

While deriving a single, closed-form analytical expression for $S(\rho_Y^{(t)})$ that holds universally across all arbitrary groups G is complicated, the exact values for any specific HSP instance can be computed algorithmically using Eq. (37) (as we demonstrate for Simon's algorithm in Sec. IV D). Furthermore, even without explicit computation, one can observe general asymptotic behaviors: We expect $S(\rho_Y^{(t)})$ to grow at least linearly with t , since the mutual information $I(J; Y_t)$ must be monotonically nondecreasing in t . On the other hand, we have $I(J; Y_t) \leq H(J)$, so for large values of t the entropy $S(\rho_Y^{(t)})$ cannot grow faster than a linear function. For any instance of the HSP, analyzing these quantities can be used to find the number of queries needed to obtain any desired amount of mutual information, i.e., t such that $I(J; Y_t) \geq H(J) - \varepsilon$.

It is worth noting that while traditional methods like the pretty good measurement [85] are highly effective for bounding the success probability and estimating asymptotic query requirements for state discrimination, analyzing $S(\rho_Y^{(t)})$ provides a complementary perspective. Specifically, rather than bounding discrimination error, our approach tracks the exact accumulation of partial mutual information per query. This is particularly useful in regimes where query counts are strictly limited and perfect discrimination is impossible.

4. Order finding

There is a subtlety in the description of order finding (hence also factorization) within this general framework. Let us briefly describe this problem: For a, N positive integers with $\gcd(a, N) = 1$, we define the order of a modulo N as the smallest positive integer r such that $a^r = 1 \pmod{N}$. The task is to find r for input a and N , given an oracle that computes $g(x) = ax \pmod{N}$. Since a and N are given to us directly as inputs, this is not even a proper oracle problem; indeed, even a classical algorithm can solve it without performing any queries. Nevertheless, this is still an interesting problem to study in terms of time complexity (i.e., the total number of computation steps, rather than counting only oracle queries). As it turns out, the path toward an efficient quantum algorithm

TABLE III. Information-theoretic quantities for a single query of Simon's algorithm, computed analytically for any value of n , up to $O(2^{-n})$.

Quantity	Prequery	Postquery	Final
$H(Y)$	n	n	n
$S(\rho_Y)$	0	n	n
$C(\rho_Y)$	n	0	0
$H(Y J)$	n	n	$n - 1$
χ	0	1	1
$I(J; Y)$	0	0	1
$D_Y(\rho_{JY}; G)$	0	1	0

for this problem goes through defining a composite oracle out of the basic one that merely computes multiplication. What this composite oracle does is compute modular exponentiation:

$$U_a|x\rangle|y\rangle = |x\rangle|a^x y \pmod{N}\rangle, \quad (39)$$

which is an oracle for the function $f(x) = a^x \pmod{N}$. Since f hides the subgroup $H = r\mathbb{Z}$ of $G = \mathbb{Z}$, we get an instance of the hidden subgroup problem, and it can be solved using a quantum Fourier transform. Still, this problem is uninteresting in terms of query complexity; even if we are not given a as an input, a single classical query computes it: $f(1) = a$. Thus, after this query, the mutual information $I(F; Y)$ would already obtain its maximal value. Indeed, the advantage of Shor's algorithm does not lie in its ability to extract as much information about the function f using few queries; rather, it is about extracting the piece of information that we care about (the order r) without requiring many additional steps of computation. Therefore, one should keep in mind that thinking about the problems of order finding and integer factorization purely in terms of query complexity may not be a useful perspective.

That said, the quantum algorithms for these problems contain a key subroutine which can be naturally considered as solving an oracle problem, i.e., phase estimation. We study this important example in Sec. IV E.

D. Information-theoretic quantities in Simon's algorithm

Recall the description of Simon's problem: We are given oracle access to a function $f : \{0, 1\}^n \rightarrow \{0, 1\}^n$, with the promise that for some unknown $s \in \{0, 1\}^n$, we have $f(x) = f(y)$ iff $x \oplus y \in \{0^n, s\}$. The goal is to identify s . This is an instance of the hidden subgroup problem with $G = (\mathbb{Z}/2\mathbb{Z})^n$, where the subgroup is promised to be of the form $H_s = \{0, s\}$. Note that s may be zero, in which case H_s has only one element. We consider the uniform distribution and apply the results of the previous section.

The information-theoretic quantities after a single query are summarized in Table III. Details of the computation appear in Sec. 8 of the Supplemental Material [56]. Note that the evolution of the information-theoretic quantities throughout the algorithm somewhat differs from those of the Bernstein-Vazirani algorithm (see Table II). In Simon's algorithm, the intermediate coherence vanishes, whereas in Bernstein-Vazirani it is maximal. Moreover, the amount of information

TABLE IV. Information-theoretic quantities at the final step of Simon's algorithm, for $2 \leq n \leq 4$ and $t \leq 24/n$. Since simulating the run of the algorithm over all possible inputs is intractable, we have utilized the analytical findings of Sec. IV C 3. Numerical computations were used to find the eigenvalues of $\rho_Y^{(t)}$ by evaluating equation (37), to compute the entropy $H(Y_t) = S(\rho_Y^{(t)})$ using these eigenvalues, and then use it to find $I(J; Y_t) = \chi$ via Eq. (38). Note that we have assumed the identities $C(\rho_Y^{(t)}) = 0$ and $D_Y = 0$. It is noteworthy that for fixed n , $H(Y_t)$ grows slightly faster than $H(Y_t|J)$ (which grows linearly with t), thus increasing the mutual information.

n	t	$H(Y_t)$	$C(\rho_Y^{(t)})$	$H(Y_t J)$	χ	$I(J; Y_t)$	D_Y
2	1	1.8802	0	1.25	0.6302	0.6302	0
2	2	3.6157	0	2.5	1.1157	1.1157	0
2	3	5.2062	0	3.75	1.4562	1.4562	0
2	4	6.6777	0	5	1.6777	1.6777	0
2	5	8.0641	0	6.25	1.8141	1.8141	0
2	6	9.3949	0	7.5	1.8949	1.8949	0
2	7	10.6914	0	8.75	1.9414	1.9414	0
2	8	11.9678	0	10	1.9678	1.9678	0
2	9	13.2324	0	11.25	1.9824	1.9824	0
2	10	14.4905	0	12.5	1.9905	1.9905	0
2	11	15.7449	0	13.75	1.9949	1.9949	0
2	12	16.9972	0	15	1.9972	1.9972	0
3	1	2.9349	0	2.125	0.8099	0.8099	0
3	2	5.7994	0	4.25	1.5494	1.5494	0
3	3	8.4822	0	6.375	2.1072	2.1072	0
3	4	10.9777	0	8.5	2.4777	2.4777	0
3	5	13.3294	0	10.625	2.7044	2.7044	0
3	6	15.5863	0	12.75	2.8363	2.8363	0
3	7	17.7857	0	14.875	2.9107	2.9107	0
3	8	19.9517	0	17	2.9517	2.9517	0
4	1	3.9663	0	3.0625	0.9038	0.9038	0
4	2	7.8975	0	6.125	1.7725	1.7725	0
4	3	11.7534	0	9.1875	2.5659	2.5659	0
4	4	15.3994	0	12.25	3.1494	3.1494	0
4	5	18.8334	0	15.3125	3.5209	3.5209	0
4	6	22.1133	0	18.375	3.7383	3.7383	0

extracted in one query of Simon's algorithm is only of the order of a single bit, while in Bernstein-Vazirani it is n . This is reflected directly in comparing the last three rows: indeed, replacing 1 by n in the values for $H(Y|F)$, the mutual information and the discord in Simon's algorithm would provide the appropriate values for Bernstein-Vazirani.

These single-query quantities highlight a key advantage of measuring performance via mutual information. After a single query in Simon's algorithm, the probability of successfully identifying the hidden string s is exponentially small (roughly 2^{-n+1}). A traditional success-probability metric would therefore suggest that virtually no computational progress has been made. However, our framework reveals that 1 bit of mutual information about the oracle class has been concretely extracted.

Combining analytic and numerical methods, we can also find the information-theoretic quantities for t queries, at the final step of the algorithm, for sufficiently small values of n, t . The figures thus obtained are depicted in Table IV.

E. Information-theoretic quantities for the phase estimation algorithm

As mentioned in Sec. IV C 4, the procedure of phase estimation forms a key part of the order finding and factoring algorithms. Here, we wish to study phase estimation as an oracle problem.

The standard formulation of the problem is as follows: Suppose a unitary operator U has an eigenvector $|u\rangle$ with eigenvalue $e^{2\pi i\varphi}$ for $\varphi \in [0, 1)$. We have available an oracle that prepares the state $|u\rangle$, as well as oracles that perform controlled- U^{2^k} for any value of k , and we are tasked with finding φ using as few queries as possible. The standard quantum algorithm for this problem uses a controlled- U^{2^k} gate, where the control qubit is set to $|+\rangle$ and the target register is set to $|u\rangle$. This results in a phase of $2^k\varphi$ being “kicked back” to the control qubit. This procedure is performed for $k = 0, \dots, t-1$ using t separate control qubits, where t depends on our desired precision. The final state of these qubits is (up to normalization) $\sum_{l=0}^{2^t-1} e^{2\pi i\varphi l} |l\rangle$, and an inverse quantum Fourier transform is utilized to extract the phase.

It is not straightforward to cast this problem in our terms, since our formalism dictates that there should be only one family of oracles, and that the oracle identity is the only hidden information. In what follows, we suggest an alternative formulation of phase estimation that solves both of these issues, while retaining the essential features of the original problem and its quantum algorithm.

The idea is as follows: Deviating from our usual assumption that $\mathcal{F}$ is a finite set, let it be the interval of possible phases, i.e., $\mathcal{F} = [0, 1)$, and suppose the random variable F has the uniform distribution on this interval. For a fixed positive integer n , we want to define that the problem of finding the first n binary digits of the unknown phase f . That is, $f = 0.f_1f_2f_3\dots$ and we wish to find $f_1, \dots, f_n$. This is equivalent to finding the class of f with respect to the partition $\mathcal{F} = \bigsqcup_{j \in \mathcal{J}} A_j$, where $\mathcal{J} = \{0, 1, 2, \dots, 2^n - 1\}$ and $A_j := \{f \in \mathcal{F} \mid \frac{j}{2^n} \leq f < \frac{j+1}{2^n}\}$. Finally, we should describe the oracles U_f . These act on a register comprising t qubits and are diagonal in the computational basis with the following entries:

$$U_f|k\rangle = e^{2\pi i f k} |k\rangle, \quad k \in \{0, 1, 2, \dots, 2^t - 1\}, \quad (40)$$

where $t \geq n$ depends on the amount of information we wish to extract. Note that here we deviate from another one of our usual assumptions, i.e., that the oracle gates U_f should only have entries 0 or 1 in the computational basis. We usually assume this to ensure that the oracle can be accessed by a classical algorithm; indeed, the problem of finding an unknown phase of a unitary gate does not directly apply to classical algorithms.

While this problem does not quite fall into the framework of the hidden subgroup problem, it can be solved with high probability using a single-query algorithm. This algorithm is given by $V = H^{\otimes t}$ and W , the inverse quantum Fourier transform with respect to the group $\mathbb{Z}/2^t\mathbb{Z}$, i.e.,

$$\langle k|W|l\rangle = \frac{1}{2^{t/2}} e^{-\frac{2\pi i k l}{2^t}}, \quad k, l \in \{0, 1, 2, \dots, 2^t - 1\}. \quad (41)$$

Let us briefly present the information-theoretic quantities throughout this algorithm, which we find using a combination of analytic and numerical methods. The details appear in Sec. 9 of the Supplemental Material [56].

For the prequery state, we find

$$C(\rho_Y^1) = H(Y_1|J) = H(Y_1) = t \quad (42)$$

and

$$S(\rho_Y^1) = \chi_1 = I(J; Y_1) = D_Y(\rho_{JY}; Z^{\otimes t}) = 0. \quad (43)$$

After the query, we have $H(Y_2|J) = t$, $C(\rho_Y^2) = 0$, and $H(Y_2) = S(\rho_Y^2) = t$, implying $I(J; Y_2) = 0$. The value of χ_2 , which equals $D_Y^2(\rho_{JY}; Z^{\otimes n})$, was found numerically. For sufficiently large difference $t - n$, χ_2 seems to approach n (see χ in Table V), i.e., it is fairly close to the upper bound $H(J) = n$.

Right before the final measurement, we have $C(\rho_Y) = 0$, $H(Y) = S(\rho_Y) = t$, and $\chi = \chi_2$. Finding $H(Y|J)$ analytically is challenging, but numerical simulations show that this value approaches $t - n$ for large values of $t - n$. Hence, increasing the difference $t - n$ contributes to increasing the mutual information in two ways: It increases the Holevo quantity χ while decreasing the discord $D_Y(\rho_{JY}; Z^{\otimes n})$. Table V depicts the information-theoretic quantities in the final step, for several values of n and t .

V. DISCUSSION

In this work, we analyzed an insightful connection between oracle problems and quantum communication. We found a relation between optimality of single-query quantum algorithms and the optimization part in the definition of quantum discord. The measurement basis that minimizes quantum discord is also optimal in terms of the algorithm's performance (when measured by mutual information). This link is revealed only when formally considering the oracle as a separate physical subsystem with its own Hilbert space. Thus, Holevo's communication task bridges between quantum nonlocality and quantum computation. Curiously, we have also found a lower bound on the mutual information, equivalent to the non-negativity of the relative entropy of coherence. Thus, quantum informational quantities arise naturally in the study of quantum algorithms for black-box problems, when choosing mutual information as the figure of merit.

Our results apply not only for single-query algorithms but also for nonadaptive multiple-query algorithms. This is immediate from the insight detailed in Sec. II C. Indeed, to study a t -query nonadaptive algorithm, one needs only to substitute the oracle unitary U_f by its t th tensor power everywhere.

This paper yields an intuitive, physical heuristic for an optimal postquery unitary in any nonadaptive algorithm for an oracle problem. We also supply a function (19) of the prequery state, whose maximization is a necessary condition for optimality. However, the maximization here seems to be difficult to perform in practice. Nonetheless, as demonstrated by the successful application of our framework to quantum likelihood estimation (discussed in Sec. III C), this mutual-information perspective provides a highly effective, exact objective function for optimizing iterative, hybrid quantum-classical subroutines, where continuous accumulation of partial information is paramount.

TABLE V. Information-theoretic quantities for the final step of the phase estimation algorithm, found by simulating the algorithm's run over all possible inputs for $2 \leq n \leq 7$ and $n \leq t \leq 10$. Notably, as t increases χ approaches n while the discord decreases [D_Y is shorthand notation for $D_Y(\rho_{JY}; \mathbb{Z}^{\otimes n})$], resulting in the mutual information approaching n . Note that $D_Y^2 = \chi_2 = \chi$, and all the other quantities for the first and second steps were found analytically, with the exact values given within the main body.

n	t	$H(Y)$	$S(\rho_Y)$	$C(\rho_Y)$	$H(Y J)$	χ	$I(J;Y)$	D_Y
2	2	2	2	0	1.3864	1.2090	0.6136	0.5954
2	3	3	3	0	1.9096	1.5250	1.0904	0.4346
2	4	4	4	0	2.5802	1.7220	1.4198	0.3022
2	5	5	5	0	3.3615	1.8405	1.6385	0.2019
2	6	6	6	0	4.2208	1.9099	1.7792	0.1306
2	7	7	7	0	5.1326	1.9498	1.8674	0.0823
2	8	8	8	0	6.0785	1.9723	1.9215	0.0507
2	9	9	9	0	7.0459	1.9848	1.9541	0.0307
2	10	10	10	0	8.0265	1.9918	1.9735	0.0183
3	3	3	3	0	1.5718	2.1865	1.4282	0.7583
3	4	4	4	0	2.0077	2.5146	1.9923	0.5223
3	5	5	5	0	2.6317	2.7170	2.3683	0.3487
3	6	6	6	0	3.3883	2.8380	2.6117	0.2263
3	7	7	7	0	4.2347	2.9087	2.7653	0.1434
3	8	8	8	0	5.1398	2.9492	2.8602	0.0890
3	9	9	9	0	6.0822	2.9720	2.9178	0.0542
3	10	10	10	0	7.0478	2.9847	2.9522	0.0325
4	4	4	4	0	1.6631	3.1810	2.3369	0.8441
4	5	5	5	0	2.0548	3.5120	2.9452	0.5668
4	6	6	6	0	2.6558	3.7157	3.3442	0.3716
4	7	7	7	0	3.4007	3.8374	3.5993	0.2381
4	8	8	8	0	4.2410	3.9084	3.7590	0.1494
4	9	9	9	0	5.1430	3.9490	3.8570	0.0920
4	10	10	10	0	6.0838	3.9719	3.9162	0.0558
5	5	5	5	0	1.7085	4.1797	3.2915	0.8881
5	6	6	6	0	2.0778	4.5114	3.9222	0.5892
5	7	7	7	0	2.6675	4.7154	4.3325	0.3829
5	8	8	8	0	3.4066	4.8373	4.5934	0.2438
5	9	9	9	0	4.2440	4.9083	4.7560	0.1523
5	10	10	10	0	5.1445	4.9490	4.8555	0.0935
6	6	6	6	0	1.7311	5.1793	4.2689	0.9104
6	7	7	7	0	2.0892	5.5112	4.9108	0.6004
6	8	8	8	0	2.6732	5.7154	5.3268	0.3886
6	9	9	9	0	3.4094	5.8372	5.5906	0.2467
6	10	10	10	0	4.2454	5.9083	5.7546	0.1537
7	7	7	7	0	1.7424	6.1793	5.2576	0.9216
7	8	8	8	0	2.0948	6.5112	5.9052	0.6060
7	9	9	9	0	2.6760	6.7153	6.3240	0.3914
7	10	10	10	0	3.4109	6.8372	6.5891	0.2481

To illustrate our results, we depicted the information-theoretic quantities for several known quantum algorithms. This includes a thorough examination of the hidden subgroup problem for finite Abelian groups. These examples may be used to develop intuition regarding the structure of successful algorithms. Let us elaborate on what these examples all have in common. First, the mutual information always vanishes in the postquery phase. The measurement outcome entropy $H(Y)$ remains high throughout all three stages. The condi-

tional entropy $H(Y|J)$ seems to follow a pattern of being initially high, remaining high in the intermediate step, and then reducing by the final value of the mutual information, $I(J;Y)$. The Holevo quantity χ follows a somewhat contrasting pattern: It starts out as zero, obtaining the value $I(J;Y)$ in the postquery phase, and then retaining this value in the final phase. The discord, however, is nonmonotonous: It is initially zero, rises to a high value after the query, and then reduces to a low value in the final step. This postquery gap, $\chi - I(J;Y) = D_Y$, provides a clear, stage-by-stage mechanical interpretation of the algorithm: The oracle query “stores” information in the quantum state (reflected by a high Holevo quantity χ and high discord, while accessible information remains low). The final unitary then serves to “unlock” this stored information, minimizing the discord and allowing the extracted mutual information to reach its upper bound. The coherence $C(\rho_Y)$ always starts out with high value and ends up being zero in the final step. However, its postquery may be either low or high; in fact, it seems that high postquery coherence is the exception (that is the case in the Bernstein-Vazirani algorithm).

Furthermore, our comparative analysis of these algorithms reveals deeper structural connections regarding deterministic success. For instance, the Bernstein-Vazirani problem, as well as the Deutsch-Jozsa problem for small input sizes ($k = 1, 2$), can be viewed as specific instances of the hidden subgroup problem where the hidden subgroup H is promised to have either index 1 or 2, and the oracle function's codomain X has cardinality 2. In these specific cases, the existing algorithms utilize these algebraic promises to solve the problem with probability one using a single query. Motivated by these observations, a subsequent study [86] introduces the general index- q hidden subgroup problem. In that follow-up work, we derived exact one-query algorithms that decide between index 1 and q for arbitrary Abelian codomain structures, and exactly identify the hidden subgroup under explicit cyclicity and compatibility conditions, further illustrating the profound role of algebraic promises in one-query solvability.

Future work could seek to tackle broader questions. It would be interesting to identify conditions for an oracle problem to admit a quantum algorithm that succeeds with probability 1. Alternatively, one may seek conditions for a problem to admit a quantum algorithm that outperforms any classical algorithm (i.e., a problem with quantum advantage).

While our framework successfully characterizes non-adaptive algorithms, fully adaptive algorithms—specifically “quantumly adaptive” protocols like Grover's algorithm, where multiple queries are applied sequentially with intermediate unitaries and no intermediate measurements—are currently outside its scope. Tracking the exact information-theoretic evolution of the state across multiple coherent, sequential channels presents a significant mathematical challenge.

However, our framework provides a conceptual roadmap for how this might be achieved in principle. An adaptive t -query algorithm can be modeled as an alternating sequence of operations, $T_f = W_t U_f W_{t-1} U_f \cdots W_1 U_f V$. The classical-quantum state ρ_{JFY} would continue to track the information flow across these stages, but the optimization landscape expands from a single pre- and postquery pair to a global, multistage parameter space. Within this lens, the intermediate

unitaries $\{W_1, \dots, W_{t-1}\}$ must serve a dual, coherent purpose: They must steer the ensemble to maximize the Holevo information (χ) injected by subsequent queries, while preserving a geometric structure that allows the final unitary W_t to successfully minimize the quantum discord (D_Y). Formulating this sequential interaction as an optimal control or dynamic programming problem remains a highly interesting direction for future work, potentially offering a systematic method for discovering adaptive algorithms.

Finally, let us note that optimizing over the mutual information is not equivalent to optimizing over the probability of success (which is arguably a more natural performance measure). It is conceivable that studying the latter may lead to analogous insights and results. It may also be interesting to try adapting our framework and using it to study alternative models of quantum computation, e.g., adiabatic, one-way, or topological quantum computers.

ACKNOWLEDGMENTS

We wish to thank Dorit Aharonov, Ernesto F. Galvão, Greg Kuperberg, Ashley Montanaro, and Thomas Vidick for help-

ful discussions. This work was supported by the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101178170, by the Israel Science Foundation under Grant Agreement No. 2208/24, by the Pazy Foundation, by the Israeli Ministry of Science and Technology, by the Israel Innovation Authority, and by the Israeli Council for Higher Education through "QERNEL" and the "Interdisciplinary Center for the Theory of Quantum Computing". A.T. was supported by the Prof. Rahamimoff Travel Grant of the U.S.-Israel Binational Science Foundation. P.H. acknowledges partial support of the IRA Programme, Project No. FENG.02.01-IP.05-0006/23, financed by the FENG program 2021–2027, Priority FENG.02, Measure FENG.02.01, with the support of the FNP.

The funders played no role in study design, data collection, analysis and interpretation of data, or the writing of this manuscript.

DATA AVAILABILITY

The source code for the numerical simulations that generated the data in the tables is publicly available in the GitHub repository [87].

-
- [1] H. Buhrman and R. de Wolf, Complexity measures and decision tree complexity: A survey, *Theor. Comput. Sci.* **288**, 21 (2002).
 - [2] D. Deutsch and R. Jozsa, Rapid solution of problems by quantum computation, *Proc. R. Soc. A* **439**, 553 (1992).
 - [3] E. Bernstein and U. Vazirani, Quantum complexity theory, *SIAM J. Comput.* **26**, 1411 (1997).
 - [4] D. R. Simon, On the power of quantum computation, *SIAM J. Comput.* **26**, 1474 (1997).
 - [5] C. H. Bennett, E. Bernstein, G. Brassard, and U. Vazirani, Strengths and weaknesses of quantum computing, *SIAM J. Comput.* **26**, 1510 (1997).
 - [6] R. Beals, H. Buhrman, R. Cleve, M. Mosca, and R. de Wolf, Quantum lower bounds by polynomials, *J. ACM* **48**, 778 (2001).
 - [7] S. Aaronson and Y. Shi, Quantum lower bounds for the collision and the element distinctness problems, *J. ACM* **51**, 595 (2004).
 - [8] B. W. Reichardt, Reflections for quantum query algorithms, in *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '11 (Society for Industrial and Applied Mathematics, Philadelphia, PA, 2011), pp. 560–569.
 - [9] C. Zhang and T. Li, Quantum lower bounds for finding stationary points of nonconvex functions, in *International Conference on Machine Learning* (PMLR, Honolulu, Hawaii, 2023), pp. 41268–41299.
 - [10] A. Lombardi, F. Ma, and J. Wright, A one-query lower bound for unitary synthesis and breaking quantum cryptography, in *Proceedings of the 56th Annual ACM Symposium on Theory of Computing* (ACM Press, New York, NY, 2024), pp. 979–990.
 - [11] S. Aaronson, Open problems related to quantum query complexity, *ACM Trans. Quantum Comput.* **2**, 1 (2021).
 - [12] S. Aaronson, How much structure is needed for huge quantum speedups? [arXiv:2209.06930](https://arxiv.org/abs/2209.06930).
 - [13] J. M. Madera, D. Egloff, N. Killoran, and M. B. Plenio, Coherent control of quantum systems as a resource theory, *Quantum Sci. Technol.* **1**, 01LT01 (2016).
 - [14] M. Howard, J. Wallman, V. Veitch, and J. Emerson, Contextuality supplies the “magic” for quantum computation, *Nature (London)* **510**, 351 (2014).
 - [15] G. Vidal, Efficient classical simulation of slightly entangled quantum computations, *Phys. Rev. Lett.* **91**, 147902 (2003).
 - [16] M. Hillery, Coherence as a resource in decision problems: The Deutsch-Jozsa algorithm and a variation, *Phys. Rev. A* **93**, 012111 (2016).
 - [17] D. Stahlke, Quantum interference as a resource for quantum speedup, *Phys. Rev. A* **90**, 022302 (2014).
 - [18] J. Ma, B. Yadin, D. Girolami, V. Vedral, and M. Gu, Converting coherence to quantum correlations, *Phys. Rev. Lett.* **116**, 160407 (2016).
 - [19] Y. Cai, H. N. Le, and V. Scarani, State complexity and quantum computation, *Ann. Phys.* **527**, 684 (2015).
 - [20] A. Datta, A. Shaji, and C. M. Caves, Quantum discord and the power of one qubit, *Phys. Rev. Lett.* **100**, 050502 (2008).
 - [21] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Quantum entanglement, *Rev. Mod. Phys.* **81**, 865 (2009).
 - [22] M. Van den Nest, Universal quantum computation with little entanglement, *Phys. Rev. Lett.* **110**, 060504 (2013).
 - [23] E. Biham, G. Brassard, D. Kenigsberg, and T. Mor, Quantum computing without entanglement, *Theor. Comput. Sci.* **320**, 15 (2004).
 - [24] N. Linden and S. Popescu, Good dynamics versus bad kinematics: Is entanglement needed for quantum computation? *Phys. Rev. Lett.* **87**, 047901 (2001).

- [25] A. Ekert and R. Jozsa, Quantum algorithms: Entanglement-enhanced information processing, *Philos. Trans. R. Soc. A* **356**, 1769 (1998).
- [26] R. Jozsa and N. Linden, On the role of entanglement in quantum-computational speed-up, *Proc. R. Soc. A* **459**, 2011 (2003).
- [27] N. Anand and A. K. Pati, Coherence and entanglement monogamy in the discrete analogue of analog Grover search, [arXiv:1611.04542](https://arxiv.org/abs/1611.04542).
- [28] V. Vedral, The elusive source of quantum speedup, *Found. Phys.* **40**, 1141 (2010).
- [29] S. Lloyd, Quantum search without entanglement, *Phys. Rev. A* **61**, 010301 (1999).
- [30] R. Jozsa, Entanglement and quantum computation, [arXiv:quant-ph/9707034](https://arxiv.org/abs/quant-ph/9707034).
- [31] A. Brodutch, Discord and quantum computational resources, *Phys. Rev. A* **88**, 022307 (2013).
- [32] H. Cable and D. E. Browne, Exact and efficient simulation of concordant computation, *New J. Phys.* **17**, 113049 (2015).
- [33] K. Modi, A. Brodutch, H. Cable, T. Paterek, and V. Vedral, The classical-quantum boundary for correlations: Discord and related measures, *Rev. Mod. Phys.* **84**, 1655 (2012).
- [34] H. Buhrman, Ł. Czekaj, A. Grudka, M. Horodecki, P. Horodecki, M. Markiewicz, F. Speelman, and S. Strelchuk, Quantum communication complexity advantage implies violation of a Bell inequality, *Proc. Natl. Acad. Sci. USA* **113**, 3191 (2016).
- [35] V. Madhok and A. Datta, Quantum discord as a resource in quantum communication, *Int. J. Mod. Phys. B* **27**, 1345041 (2013).
- [36] H. Buhrman, R. Cleve, and A. Wigderson, Quantum vs. classical communication and computation, in *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing* (ACM Press, New York, NY, 1998), pp. 63–68.
- [37] R. De Wolf, *Quantum Computing and Communication Complexity*, Ph.D. thesis, University of Amsterdam, 2001.
- [38] R. de Wolf, Quantum communication and complexity, *Theor. Comput. Sci.* **287**, 337 (2002).
- [39] G. Brassard, Quantum communication complexity, *Found. Phys.* **33**, 1593 (2003).
- [40] T. P. Spiller, K. Nemoto, S. L. Braunstein, W. J. Munro, P. van Loock, and G. J. Milburn, Quantum computation by communication, *New J. Phys.* **8**, 30 (2006).
- [41] H. Buhrman, R. Cleve, S. Massar, and R. de Wolf, Nonlocality and communication complexity, *Rev. Mod. Phys.* **82**, 665 (2010).
- [42] W. M. Hoza, Quantum communication-query tradeoffs, [arXiv:1703.07768](https://arxiv.org/abs/1703.07768).
- [43] S. Chakraborty, A. Chattopadhyay, N. S. Mande, and M. Paraashar, Quantum query-to-communication simulation needs a logarithmic overhead, in *Proceedings of the 35th Computational Complexity Conference, CCC '20* (Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, DEU, 2020).
- [44] S. Chakraborty, A. Chattopadhyay, P. Høyer, N. S. Mande, M. Paraashar, and R. de Wolf, Symmetry and Quantum Query-To-Communication Simulation, in *Proceedings of the 39th International Symposium on Theoretical Aspects of Computer Science (STACS 2022)*, edited by P. Berenbrink and B. Monmege, Leibniz International Proceedings in Informatics (LIPIcs) (Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 2022), Volume 219, pp. 20:1–20:23.
- [45] J. Blikstad, J. Van Den Brand, Y. Efron, S. Mukhopadhyay, and D. Nanongkai, Nearly optimal communication and query complexity of bipartite matching, in *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)* (IEEE, Piscataway, NJ, 2022), pp. 1174–1185.
- [46] A. Chakraborty, R. K. Patra, K. Agarwal, S. Sen, P. Ghosal, S. G. Naik, and M. Banik, Scalable and noise-robust communication advantage of multipartite quantum entanglement, *Phys. Rev. A* **111**, 032617 (2025).
- [47] S. Rout, N. Sakharwade, S. S. Bhattacharya, R. Ramanathan, and P. Horodecki, Unbounded quantum advantage in communication with minimal input scaling, *Phys. Rev. Res.* **7**, 023104 (2025).
- [48] A. Levi, Z. Ossi, E. Cohen, and A. Te'eni, Optimal quantum likelihood estimation, *Quantum Sci. Technol.* **11**, 015029 (2026).
- [49] A. Belovs, Variations on quantum adversary, [arXiv:1504.06943](https://arxiv.org/abs/1504.06943).
- [50] A. Ambainis, Quantum lower bounds by quantum arguments, *J. Comput. Syst. Sci.* **64**, 750 (2002).
- [51] 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 Press, New York, NY, 2007), pp. 526–535.
- [52] A. Montanaro, Structure, randomness and complexity in quantum computation, Ph.D. thesis, University of Bristol, 2007.
- [53] A. Montanaro, Quantum algorithms: An overview, *npj Quantum Inf.* **2**, 15023 (2016).
- [54] A. Chailloux, A note on the quantum query complexity of permutation symmetric functions, in *Proceedings of the 10th Innovations in Theoretical Computer Science Conference (ITCS 2019)*, edited by A. Blum, Leibniz International Proceedings in Informatics (LIPIcs), Vol. 124 (Schloss Dagstuhl–Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 2019), pp. 19:1–19:7.
- [55] A. C.-C. Yao, Probabilistic computations: Toward a unified measure of complexity, in *Proceedings of the 18th Annual Symposium on Foundations of Computer Science (sfcs 1977)* (IEEE Computer Society, USA, 1977), pp. 222–227.
- [56] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/5qsb-5jy5> for formal remarks on the problem setup, a discussion of the probability of success, a proof linking Holevo's bound to quantum discord, algebraic conditions for deterministic algorithms, and step-by-step information-theoretic analyses of the Bernstein-Vazirani, Shor-Kitaev, Simon, and phase estimation algorithms.
- [57] R. Jozsa, Characterizing classes of functions computable by quantum parallelism, *Proc. R. Soc. A* **435**, 563 (1991).
- [58] S. Aaronson, A. Ambainis, J. Iraids, M. Kokainis, and J. Smotrovs, Polynomials, quantum query complexity, and Grothendieck's inequality, in *Proceedings of the 31st Conference on Computational Complexity, CCC '16* (Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, DEU, 2016).
- [59] W. Chen, Z. Ye, and L. Li, Characterization of exact one-query quantum algorithms, *Phys. Rev. A* **101**, 022325 (2020).
- [60] G. Xu and D. Qiu, Partial Boolean functions with exact quantum query complexity one, *Entropy* **23**, 189 (2021).

- [61] D. Qiu and G. Xu, Exact quantum 1-query algorithms and complexity, *SPIN* **11**, 2140001 (2021).
- [62] Z.-K. Ye and L.-Z. Li, Characterization of exact one-query quantum algorithms for partial Boolean functions, *J. Comput. Sci. Technol.* **38**, 1423 (2023).
- [63] A. Ambainis, K. Iwama, A. Kawachi, H. Masuda, R. H. Putra, and S. Yamashita, Quantum identification of Boolean oracles, in *Annual Symposium on Theoretical Aspects of Computer Science* (Springer, Berlin, Heidelberg, 2004), pp. 105–116.
- [64] A. Ambainis, K. Iwama, A. Kawachi, R. Raymond, and S. Yamashita, Improved algorithms for quantum identification of Boolean oracles, *Theor. Comput. Sci.* **378**, 41 (2007).
- [65] R. Kothari, *An Optimal Quantum Algorithm for the Oracle Identification Problem* (Schloss Dagstuhl–Leibniz-Zentrum für Informatik, Germany, 2014).
- [66] L. Taghavi, Simplified quantum algorithm for the oracle identification problem, *Quantum Mach. Intell.* **4**, 19 (2022).
- [67] A. Montanaro, Nonadaptive quantum query complexity, *Inf. Process. Lett.* **110**, 1110 (2010).
- [68] M. Zolowak and W. H. Zurek, Complementarity of quantum discord and classically accessible information, *Sci. Rep.* **3**, 1729 (2013).
- [69] R. Lecamwasam, S. Assad, J. J. Hope, P. K. Lam, J. Thompson, and M. Gu, Relative entropy of coherence quantifies performance in Bayesian metrology, *PRX Quantum* **5**, 030303 (2024).
- [70] K. M. Audenaert, Quantum skew divergence, *J. Math. Phys.* **55**, 112202 (2014).
- [71] M. E. Shirokov, Upper bounds for the Holevo information quantity and their use, *Probl. Inf. Transm.* **55**, 201 (2019).
- [72] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2010).
- [73] T. Baumgratz, M. Cramer, and M. B. Plenio, Quantifying coherence, *Phys. Rev. Lett.* **113**, 140401 (2014).
- [74] Z. Xi, Y. Li, and H. Fan, Quantum coherence and correlations in quantum system, *Sci. Rep.* **5**, 10922 (2015).
- [75] A. Bilobran and R. Angelo, A measure of physical reality, *Europhys. Lett.* **112**, 40005 (2015).
- [76] P. R. Diegueze and R. M. Angelo, Information-reality complementarity: The role of measurements and quantum reference frames, *Phys. Rev. A* **97**, 022107 (2018).
- [77] A. C. Costa and R. M. Angelo, Information-based approach towards a unified resource theory, *Quantum Inf. Process.* **19**, 325 (2020).
- [78] I. L. Paiva, P. R. Diegueze, R. M. Angelo, and E. Cohen, Coherence and realism in the Aharonov-Bohm effect, *Phys. Rev. A* **107**, 032213 (2023).
- [79] C. E. Granade, C. Ferrie, N. Wiebe, and D. G. Cory, Robust online Hamiltonian learning, *New J. Phys.* **14**, 103013 (2012).
- [80] N. Wiebe, C. Granade, C. Ferrie, and D. Cory, Hamiltonian learning using imperfect quantum resources, *Phys. Rev. A* **89**, 042314 (2014).
- [81] N. Wiebe, C. Granade, C. Ferrie, and D. G. Cory, Hamiltonian learning and certification using quantum resources, *Phys. Rev. Lett.* **112**, 190501 (2014).
- [82] J. Wang, S. Paesani, R. Santagati, S. Knauer, A. A. Gentile, N. Wiebe, M. Petruzzella, J. L. O'brien, J. G. Rarity, A. Laing, *et al.*, Experimental quantum Hamiltonian learning, *Nat. Phys.* **13**, 551 (2017).
- [83] X.-L. Qi and D. C. Ranard, Determining a local Hamiltonian from a single eigenstate, *Quantum* **3**, 159 (2019).
- [84] A. Dutt, E. Pednault, C. W. Wu, S. Sheldon, J. Smolin, L. Bishop, and I. L. Chuang, Active learning of quantum system Hamiltonians yields query advantage, *Phys. Rev. Res.* **5**, 033060 (2023).
- [85] H. Barnum and E. Knill, Reversing quantum dynamics with near-optimal quantum and classical fidelity, *J. Math. Phys.* **43**, 2097 (2002).
- [86] A. Te'eni, Y. Oz, and E. Cohen, One-query quantum algorithms for the index- q hidden subgroup problem, [arXiv:2510.10538](https://arxiv.org/abs/2510.10538).
- [87] A. Te'eni, https://github.com/smitke6/Oracle_Problems_as_Communication_Tasks (2025).