Quantum advantage for noisy channel discrimination

Zane M. Rossi, Jeffery Yu, Isaac L. Chuang, and Sho Sugiura

Phys. Rev. A **105**, 032401 — Published 2 March 2022

DOI: 10.1103/PhysRevA.105.032401

Quantum advantage for noisy channel discrimination

Zane M. Rossi,[1] Jeffery Yu,[1] Isaac L. Chuang,[2] and Sho Sugiura[3]

[1] *Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[2] *Department of Physics, Department of Electrical Engineering and Computer Science, and Co-Design Center for Quantum Advantage, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[3] *Physics and Informatics Laboratory, NTT Research, Inc., 940 Stewart Dr., Sunnyvale, California, 94085, USA*

(Dated: January 28, 2022)

Many quantum mechanical experiments can be viewed as multi-round interactive protocols between known quantum circuits and an unknown quantum process. Fully quantum "coherent" access to the unknown process is known to provide an advantage in many discrimination tasks compared to when only incoherent access is permitted, but it is unclear if this advantage persists when the process is noisy. Here, we show that a quantum advantage can be maintained when distinguishing between two noisy single-qubit rotation channels. Numerical and analytical calculations reveal a distinct transition between the performance of fully coherent and fully incoherent protocols as a function of noise strength. Moreover, the size of the region of coherent quantum advantage shrinks inverse polynomially in the number of channel uses, and in an intermediate regime an improved strategy is a hybrid of fully-coherent and fully-incoherent subroutines. The fully coherent protocol is based on quantum signal processing, suggesting a generalizable algorithmic framework for the study of quantum advantage in the presence of realistic noise.

I. INTRODUCTION

Experimental progress over the past twenty years has increasingly enabled the coherent manipulation of complex quantum mechanical systems, bolstering the ongoing search for settings where quantum protocols permit advantage over their classical counterparts. Often it is assumed that multiple unitary operations can be coherently applied serially to an initial quantum state, and indeed numerous results support the intuition that quantum advantage often relies on the ability to perform deep quantum circuits.

Multiple recently developed frameworks have considered quantum advantage through the lens of 'quantum' versus 'classical' access models [1, 2]. In particular, for certain inference problems, it has been shown that quantum advantage is recoverable for models in which an oraclized quantum process can be applied coherently, as compared to models without such coherent access. This approach, first fixing a problem and then comparing algorithmic performance across *differing access models*, has permitted novel complexity-theoretic insights into the sources of quantum advantage.

A missing piece in the work on access model dependent quantum advantage is an *explicit and constructive* study of the effect of noise. In [1], the exponential query complexity advantage studied does not survive the introduction of noise, while relatedly the quantum advantage studied in [2] is proven for finite noise, but only information-theoretically. In the era of noisy quantum devices, it is vital to understand the gap between these two approaches: namely, constructive investigations of (1) the effect of noise for realistic inference tasks, and (2) problems which permit a difference in performance among access models to survive the introduction of nonzero noise.

This work aims to understand the gaps demarcated by [1] and [2] by investigating an instance of quantum advantage which (1) is amenable to constructive methods, (2) incorporates noise, and (3) introduces concrete parameters to enable visualization of where this advantage exists. Specifically, while [2] demonstrates an information-theoretic no-go theorem for quantum advantage in average case regression tasks for machine learning, we choose a classification problem, to which their results are not directly applicable. Moreover, while [1] considers comparisons between noiseless settings and unparameterized noise, we incorporate noise which is parameterized by a single, continuous real value. Considering this binary classification task and simply parameterized noise, we are able to identify, and compellingly depict, regions of quantum advantage with respect to simple parameters representing signal and noise.

The specific problem we consider is discrimination among two noisy single qubit rotation channels, where the noise is defined by classical distributions over the rotation angle. Given consistent access to one among two possible quantum channels, where sampling rotates the querent's qubit by said noisy angle, the querent is challenged to determine which distribution underlies their sampling power. The normally distributed noise we consider is parameterized solely by its mean and standard deviation. Consequently, plotting the the performance of coherent and incoherent access protocols against these two parameters reveals thresholds between regions in the defining parameter space for which each access model exhibits relative advantage.

This problem is perhaps the simplest instance of a more general class of quantum channel discrimination problems. This class contains instances which are known to be difficult and rely on sophisticated use of quantum resources, e.g., entanglement and auxiliary space.

We find that, up to a certain noise threshold, a coherent access protocol can always outperform its incoherent access counterpart for our hypothesis testing problem, but that the reverse is true above this noise threshold. Moreover, below the threshold, we show that there exist a family of even better performing hybrid protocols, which are alternately coherent and incoherent. For such protocols we find that one should compute coherently for a certain time, measure, and repeat, and we compute the optimal query complexity, or *coherence length*, for the coherent subroutines of these protocols.

The problem proposed in this work is one for which the recently developed algorithmic primitive of quantum signal processing (QSP) [3–6] is natural. Indeed, much of our analysis relies on the application of known properties and guarantees of QSP, though we here extend these methods to a new, noisy context. Even simple noise within QSP protocols has not been evaluated beyond questions of simple error propagation, and thus this work presents one avenue toward lifting QSP into more exotic noise settings.

This work is structured as follows: in Section II we discuss a noiseless instance of the hypothesis testing problem introduced in Definition I.1, for which optimal quantum protocols are known, and generalize these results to the case of noise in Section III, where various limits permit closed form analysis of the behavior of these discrimination protocols. Finally, we examine hybrid protocols in Section IV and discuss their significance.

A. Problem statement

Consider an instance of symmetric hypothesis testing among quantum channels. Let two distinct distributions be Θ_0, Θ_1, over the reals. These distributions have well-behaved probability density functions denoted $\Theta_b(\theta)$ for $b \in \{0, 1\}$. We involve these distributions in an "RDG" game involving quantum channels, as given in Definition I.1.

Definition I.1. *Rotation discrimination game (RDG). A party with a single qubit is afforded multiple-shot query access to a single-qubit quantum channel taking the following form: when queried, some θ unknown to the party is drawn from Θ_b (either 0 or 1, with b fixed for all queries), and the unitary channel*

$$\mathcal{E}_b \equiv \exp(i\theta\sigma_x) \tag{1}$$

is applied to the party's qubit, where σ_x is the Pauli X operator with determinant 1.

The distributions Θ_b are taken to be normal, defined by two parameters

$$\Theta_b(\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\theta_b - \mu_b)^2}{2\sigma^2}}, \tag{2}$$

where μ_b and σ are the mean and the standard deviation respectively. The challenge is to, in as few queries as possible, determine which classical distribution parameterizes the quantum channel; i.e., to determine b with high confidence and low query complexity. The total query complexity is denoted N. We use a shorthand for the separation of the means, $\delta = |\mu_0 - \mu_1|$.

Finally, the party is assumed to only apply serial, non-adaptive protocols*, i.e., they have access to only one qubit, and their strategy must be independent of intermediate measurement results. The resources otherwise afforded to an algorithm playing an RDG, specifically* coherent access *or* incoherent access*, are defined below.*

This problem is simple to describe, yet sufficiently rich to exhibit a distinct transition between optimal performance of coherent and incoherent access models.

We pose this problem because it will be natural to consider two limits in the single noise parameter σ and the mean separation δ. One limit is the near noiseless case, i.e., $\sigma \to 0$; in the absence of noise, the coherent model is strictly better than the incoherent model, which we discuss in Section II. The other limit is when the noise is much larger than the mean angular separation δ, or simply when $\delta \to 0$ for any fixed non-zero σ. In this limit coherence is almost immediately lost, and a fully incoherent protocol will show advantage.

As the concrete protocols we consider will encompass well-performing strategies for both limits discussed above, they will also be useful in discussing the more difficult to analyze intermediate regions in σ, δ.

Taking a step back, we are most interested in performance differences between quantum algorithms situated in coherent and incoherent access models for specific tasks. We want to make concrete the distinction between quantum strategies for RDGs in these two models, depicted in Figure 1. Before this, however, we give a short definition. In the single-qubit and binary outcome setting a *complete measurement*, in analogy to a term used in [1], will mean simply a projective measurement, defined by $\{|\psi\rangle\langle\psi|, I - |\psi\rangle\langle\psi|\}$ for some pure single qubit state $|\psi\rangle$. More generically, a complete measurement is one that prepares quantum states which are completely defined by (classical) measurement results. We now define our access models of interest.

- **Incoherent access:** The querying party is forced to perform a *complete measurement* (see above) between each channel application. Measurement outcomes are processed through some classical thresholding (or post-processing) procedure.

- **Coherent access:** The querying party *must* now defer complete measurement, and may instead perform intermediate unitary quantum gates on the qubit between successive applications of $\mathcal{E}_b$.

The protocols we consider will, as stated in Def. I.1, take place in a serial, non-adaptive setting regardless of access model. That is, the single-qubit channel $\mathcal{E}_b$ is not applied jointly to many qubits, and quantum operations are performed independent of intermediate measurement results. This is a restriction, but one which does not remove all interesting properties of the comparison.

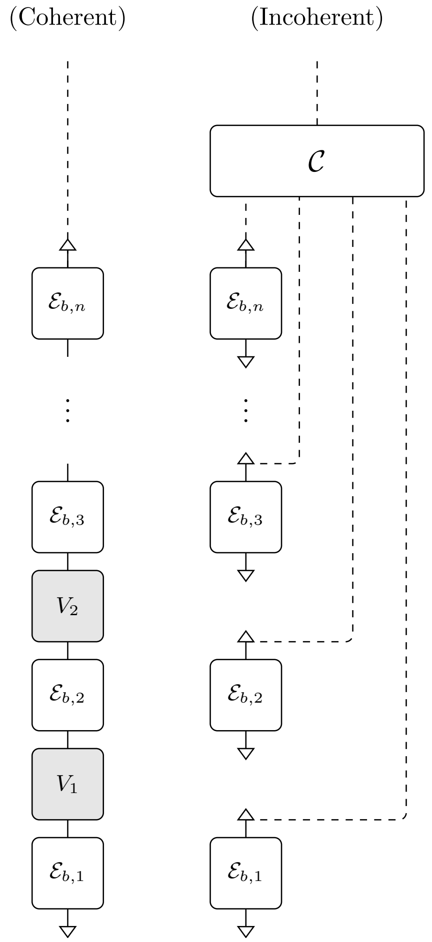

FIG. 1. Non-adaptive serial incoherent access (right) and coherent access (left) protocols for RDGs. Here $\mathcal{C}$ is some classical thresholding procedure performed on the results of measurements, dotted lines indicate transmission of classical information, and triangles are complete measurements when leaving a channel application and classically controlled quantum state preparations going into a channel application. The V_j are unitary quantum gates, while $\mathcal{E}_{b,j}$ is the j-th application of b-labelled hypothesis channel. This figure follows conventions established in [1], in which time progresses as one moves up the diagram.

B. Prior work

Much of the work discussing quantum advantage in relation to coherence does so for problems in inference [2, 7–9]. Statistical inference, a central and broad tool in experimental contexts, has been investigated with respect to quantum mechanical systems for over fifty years, and hypothesis testing among quantum channels specifically, the subject of this work, has been studied in a variety of contexts for the past twenty years [10–13]. While binary hypothesis testing in quantum settings has been characterized for both quantum states [7] and restricted sets of quantum channels [8, 10, 11], such initial treatments did not consider generally noisy settings. Moreover, these seminal results are still being elaborated currently to pertain to more diverse settings, most often the presence of noise and adaptiveness [14–16].

The specific approach of analyzing quantum algorithmic performance according to access model is more recent [1]. Along with the complexity-theoretic treatment of [1], a similar distinction has been investigated for machine learning problems [2], differentiating between the performance of coherent access and incoherent access protocols for worst-case approximation of specific observables. In the latter work, [2], as their inference task involves all n-qubit Pauli operators, which can be efficiently measured in the coherent access model by clever constructions, a performance separation is proven by information theoretic methods against any, even adaptive incoherent access protocol. Surprisingly, the work of [2] guarantees the persistence of this quantum advantage in the presence of noise of any strength. While their problem is one of regression, inferring expectation values averaged over noise, a new challenge appears when considering classification. Such problems, primarily investigated in [1], can lead to a vanishing of quantum advantage under finite noise. It is the fragility of quantum advantage for certain classification tasks, and its apparent robustness in regression problems, that presents an interesting gap. Extending such robustness to general noisy classification problems remains an intriguing obstacle to overcome.

Quantum mechanics is known to be a resource convex theory [17, 18], i.e., that there exist discrimination problems for which any additional resource can provide algorithmic advantage. However, investigating algorithmic advantage among access models, even for simple access models and tasks, is in general a difficult problem, sensitive to the algorithmic families considered, and is thus often approached by non-constructive methods. Therefore, providing constructive evidence for any gap in performance, even within relatively restrictive access models, can provide insight into the benefits of coherence. Moreover, if these restrictions are nonetheless physically motivated, conclusions drawn can serve as a good basis for future constructive generalizations, and their use in experiment.

Various bounds exist for the performance of algorithms for quantum hypothesis testing among noisy quantum channels, though these results rely on novel information-theoretic proofs in complicated resource models [12, 13]. This complexity is required by the broadness of the settings they consider. While it is surprising and intriguing that these results exist, most still do not consider (1) relative performance among coherent and incoherent access protocols specifically, and (2) classification problems for these protocols. For even simple classification games like RDGs, no existing work investigates the relative advantage between coherent and incoherent access models.

Moreover, although the protocols that are considered in this paper will be non-adaptive, a great deal of literature exists which (non-constructively) proves that adaptiveness is not required to achieve optimial discrimination in many quantum channel discrimination problems. Most notably these include classical feed-forward adaptive quantum channel discrimination problems [19] as well as a variety of channel sets whose structure is richer than those investigated in this work [9, 15, 20]. Consequently, within the resource models introduced in this work, separations in performance are robust to even the introduction of adaptiveness.

In a seemingly entirely separate subfield of quantum algorithms, the development of quantum signal processing (QSP) [3–5] and the quantum singular value transform (QSVT) [6] has encouraged a fresh look at the common mechanisms underlying notable quantum algorithms [21]. QSP and QSVT, which permit the efficient polynomial transformation of eigenvalues of unitary operators and singular values of embedded linear operators respectively, have proven flexible ansätze: able to reproduce optimal quantum algorithms for problems as diverse as Hamiltonian simulation, the quantum linear systems problem, and factoring [6, 21]. Since the polynomial transformations within QSP and QSVT can be efficiently performed even on unknown unitary processes, they are expected to be powerful methods for addressing problems in inference as well. That said, their application to noisy settings has been limited, and introduces caveats. Building a bridge between quantum algorithms and quantum inference in noisy settings (e.g. from QSP to RDGs) is consequently of great interest for understanding constructive methods to achieve quantum advantage for inference problems, and fleshing out further uses for QSP and QSVT.

II. NOISELESS DISCRIMINATION

Before discussing solutions to the major problem discussed in this work, i.e., the discrimination of noisy quantum processes in the form of RDGs (Definition I.1), it is worthwhile to discuss the noiseless case. This simpler problem is not only amenable to closed form results but also provides intuition that will help guide us through the introduction of noise.

It is the aim of this section to answer the following questions in the noiseless case, as each will become important and non-trivial in the general setting.

- Do there exist pairs of quantum channels for which coherent access (e.g., QSP-based) discrimination methods outperform incoherent access protocols, and what is the quantitative nature of this relative speedup?
- Given a pair of quantum channels (hypotheses), what is the simplest resource model in which an optimal discrimination protocol is possible? More fundamentally, what methods exist for defining optimality.

We consider these questions below, first analytically, and then with a concrete example.

A. Sufficiency of QSP protocols for optimal noiseless discrimination

To investigate these questions we state a simple channel discrimination problem concretely in the noiseless setting in Definition II.1.

Definition II.1. Noiseless RDGs. *We consider one concrete instantiation of an RDG (Definition I.1): distinguishing between two quantum channels $\mathcal{E}_0, \mathcal{E}_1$ with explicit form $\exp\{i\theta_0\sigma_x\}$ and $\exp\{i\theta_1\sigma_x\}$ respectively, where σ_x is a Pauli operator as before, and each of θ_0, θ_1 is fixed.*

A party is given the ability to apply a quantum channel $\mathcal{E}_b$ for some consistent $b \in \{0, 1\}$ without knowledge of b. The party is tasked with the following goal: determine, given repeated access to $\mathcal{E}_b$, the hidden bit b.

In comparison to general RDGs, in this game the distributions Θ_b over angles return unitary channels $\mathcal{E}_b$, and are thus Dirac distributions peaked about $\mu_b = \theta_b$. In other words, this is the problem of discriminating two fixed, known rotations. The problem of unitary channel discrimination, for which this problem is one example, has been studied for both coherent and incoherent access protocols [7, 8]. This section translates these results into the language of QSP protocols, and provides a new statement for the separation in performance of coherent access and incoherent access protocols, toward analysis of the more complicated case where each $\mathcal{E}_b$ is non-unitary (Section III).

In this setting we have a complete characterization of both coherent access and incoherent access protocols (in the non-adaptive case). To make this more clear, we present a new family of quantum protocols.

Definition II.2. *A* QSP protocol *for an RDG is defined by (1) a series of QSP phase angle lists, $\{\Phi_1, \Phi_2, \cdots, \Phi_m\}$, each of which is in $\mathbb{R}^{r_j}$ for $j \in \{1, 2, \cdots, m\}$ and r_j a positive integer, and (2) a series of classical descriptions of qubit preparations and binary projective qubit measurements $\{(\psi_1, \psi_1'), (\psi_2, \psi_2'), \cdots, (\psi_m, \psi_m')\}$.*

Here the probability to measure $|\psi_j\rangle$ is $|\langle\psi_j'|Q_{\Phi_j}|\psi_j\rangle|^2$, and Q_{Φ_j} has the form

$$Q_{\Phi_j} = e^{i\phi_0\sigma_z} \prod_{\ell=1}^{r_j} \left(\mathcal{E}_b\, e^{i\phi_\ell\sigma_z}\right). \tag{3}$$

The total number of z-rotations is equal to the number of channel applications: $\sum_j r_j = N$.

Evidently QSP protocols have non-trivial intersection with coherent access protocols, and moreover when $m = 1$ they are a proper subset of coherent access protocols. It is not difficult to see that in the case of $r_j = 1$ the QSP protocol given in Definition II.2 reduces to an incoherent access protocol. For the incoherent access protocol we will consider in comparison with general QSP protocols, the N classical measurement outcomes are processed through a simple thresholding procedure: majority vote.

When comparing coherent and incoherent access protocols, we will optimize over both QSP angles and choice of state preparations and projective measurements such that the overall error probability in discrimination is minimized. Performing this comparison, QSP protocols are shown to be a useful subset of coherent access quantum algorithms for RDGs.

It is worthwhile to cast our procedure in the framework of [1]. The problem they consider also concerns the inference of an unknown quantum channel from among two possibilities, though the advantage they prove is for a noiseless variant. The quantum circuit they propose for a coherent access protocol, like ours, comprises a set of possible gates, while in the incoherent case any two unknown channel applications must be interrupted by local operations and classical communication (LOCC). The majority vote we introduce can always be implemented with LOCC, and thus our work, while more general in the channels it considers, is narrower in the protocols proposed to distinguish these channels. We nevertheless show quantum advantage for the noiseless case in this section, and in Section III we go further than [1] by showing a quantum advantage which persists even in the presence of finite noise.

Given well-known results for the form of QSP-generated unitary operators, and the ease of analysis of the corresponding embedded eigenvalue transforms, QSP provides an excellent starting point for the analysis of quantum hypothesis testing. Using QSP, we first state the following theorem for noiseless RDGs.

Theorem II.1. *For noiseless RDGs (Definition II.1), coherent access protocols can always match or exceed the performance of incoherent access protocols. Moreover, there exists a finite positive integer N and a coherent access non-adaptive protocol using N queries such that this protocol perfectly decides the bit b naming the hidden channel, where $N = \mathcal{O}(\delta^{-1})$ is optimal.*

Proof. The existence of such an N follows from the major result of [8] by recognition that this is a unitary channel discrimination problem, and direct construction for rotations about a fixed axis can be found in [22] (Lemma IV.4). That coherent access protocols can always outperform incoherent access ones follows from the set describing the latter strictly containing the set of protocols comprising the former. This shows the sufficiency of QSP protocols for optimal noiseless discrimination. □

It is worthwhile to explain why the fully coherent QSP protocol given in Theorem II.1 performs obviously better than its incoherent access counterpart. We motivate a simple and old result from quantum information which will appear again and again in analyzing the basic behavior of incoherent access protocols.

The one-shot distinguishability of two unitary quantum channels $\mathcal{E}_0, \mathcal{E}_1$ is determined by the maximum over initial density operators σ and all POVMs, and gives the minimum error probability

$$p_{err} = \min_{\sigma} \frac{1}{2}\left(1 - \frac{\|\mathcal{E}_0(\sigma) - \mathcal{E}_1(\sigma)\|}{2}\right), \tag{4}$$

where $\|\cdot\|$ is the trace distance, the implicit optimal POVM the *Helstrom measurement*, and the overall statement the *Helstrom bound*. For noiseless RDGs, this bound takes the form

$$p_{err} = \frac{1}{2}(1 - \sin|\theta_0 - \theta_1|), \tag{5}$$

where it will be taken without loss of generality that $\delta = |\theta_0 - \theta_1| \leq \pi/2$. This simple bound completely defines the performance of incoherent protocols up to polynomial factors when paired with non-adaptive post-processing on measurement results, and will serve to constitute a performance bound in the comparison to coherent access protocols.

One possible post-processing method in the incoherent access setting is the majority vote: a simple thresholding procedure; this work will solely be considering this type of non-linear, non-adaptive post-processing, though many additional classical statistical methods and thresholding procedures are possible. We are able to forgo, in this instance, adaptive strategies due to previous work in Sec. I B showing their lack of benefit in our setting.

We define the majority vote below and discuss its performance in comparison to QSP protocols for noiseless RDGs. A majority vote (here denoted MAJ) unsurprisingly returns the majority result from a set of $(2M+1)$ i.i.d. Bernoulli samples defined by some underlying success probability $1/2 \leq p \leq 1$. The distribution which defines the output of a majority vote is itself a Bernoulli distribution with a modified p. I.e., it corresponds to a single sample returned from the Bernoulli distribution defined by

$$p' = \sum_{k=0}^{M} \binom{2M+1}{k} p^{2M+1-k}(1-p)^k, \tag{6}$$

which for the constraints given satisfies $p' \geq p$ when $M > 0$, an integer. An analogous reversed statement can be made for $0 \leq p < 1/2$.

For incoherent access protocols we consider that the querent is forced to completely measure after each use of the quantum channel, but is however free to choose this measurement as well as the preparation of the qubit state to which the channel is applied. In this restricted setting the Helstrom bound can be combined with majority vote

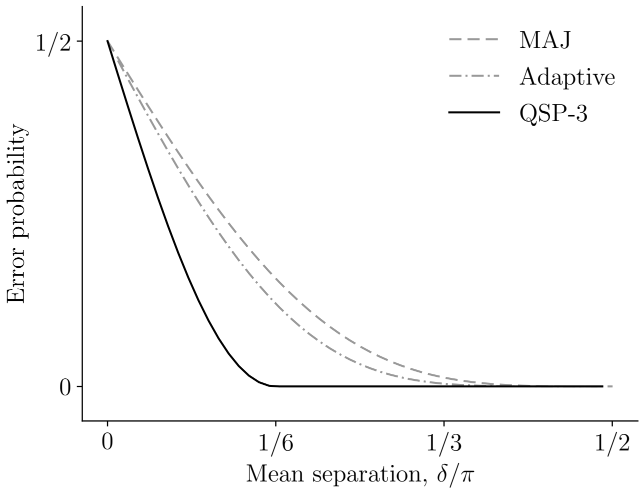

FIG. 2. Error probability plotted as a function of distance between signal means δ, with no noise, i.e., $\sigma = 0$. This figure compares protocols that use three total queries ($N = 3$), and which are either fully coherent (solid), fully incoherent and non-adaptive (dashed), and fully incoherent and adaptive (dotted dashed) as described in Appendix A.

to give a more complete picture of the performance of incoherent access protocols.

For coherent access protocols, in comparison, we consider instead that the querent may successively apply the quantum channel multiple times to a chosen input, possibly interspersing these applications with quantum operations of their own choosing before finally measuring with respect to a chosen complete measurement. While it is clear that the use of the hypothesis channel in a QSP sequence is a subset of such protocols, it is not clear that optimal discrimination protocols can be performed serially and without entanglement. However, this is demonstrated to be the case in settings like those of Theorem II.1, as well as for many noiseless channel discrimination protocols [11]. This paper will continue to consider discrimination strategies which do not use entanglement due to these strong indications that, at least in small noise settings, the benefit of entanglement vanishes [11, 20].

Having made the statement that coherent protocols can always outperform incoherent ones in the noiseless setting, we are now interested in an upper bound for the magnitude of the relative advantage between the two protocols with respect to a particular pair of quantum channels. One such bound is discussed in a concrete, simple example below.

B. A simple concrete example

If we consider a concrete hypothesis testing problem and employ constructive QSP protocols for its solution, corresponding performance differences become easier to visualize. In Figure 2 we compare the performance of two methods for symmetric binary hypothesis testing for rotations about a fixed axis in the $N = 3$ case. The same functional form derived here will extend easily to larger N.

Considering the performance of incoherent protocols first, the probability of error from successive Helstrom measurements followed by majority votes can be calculated explicitly, taking $N = 2M + 1$ an odd positive integer

$$p_{\mathrm{MAJ}} = 2^{-(2M+1)} \sum_{k=0}^{M} \binom{2M+1}{k} \times (1 - \sin\delta)^{2M+1-k} (1 + \sin\delta)^k \tag{7}$$

for $\delta \in [0, \pi/2]$, where p_{MAJ} is the error probability of the majority vote.

We can compare the the behavior of (7) with the behavior of an optimized QSP sequence of length 3 (QSP-3, precisely the length 3 form of Def. II.2), which for this noiseless setting is provably optimal in query complexity, and has a simple form. For any fixed N, the error probability of the optimized QSP sequence is given by

$$p_{\mathrm{QSP}} = \frac{1}{2}(1 - \sin N\delta) \tag{8}$$

for $\delta = |\theta_0 - \theta_1| \leq \pi/N$ and zero otherwise (in which case the results of [8, 22] enable perfect discrimination). The form of Eq. 8 is merely the Helstrom bound applied to the N-th power of the signal rotation. It is not difficult to numerically verify (see, e.g., Fig. 2) that this function is always strictly less than (7) for any positive choice of N and any non-zero separation δ within $[0, \pi/N]$. While this result can be shown empirically, its proof for some finitely sized region in δ follows from simple observation that the respectively derivatives of each continuous function around $\delta = 0$ (at which they both equal 1/2) are respectively

$$p_{\mathrm{QSP}} \propto 2M + 1, \quad p_{\mathrm{MAJ}} \propto 2^{-2M} \frac{(2M+1)!}{(M!)^2}, \tag{9}$$

the later of which is derived in Appendix C. Moreover, for some finite M the right hand side of Eq. 10 is greater than the left as they asymptotically approach

$$p_{\mathrm{QSP}} \to 2M, \quad p_{\mathrm{MAJ}} \to 2\sqrt{\frac{M}{\pi}}, \tag{10}$$

respectively, by Stirling's approximation. Finally, the magnitude of the ratio of the error probabilities is discussed in Figure 2, and can for certain choices of parameters be arbitrarily large. Consequently as the functions involved are bounded and differentiable, there exists some finite positive δ for which our coherent protocol beats its incoherent access counterpart in this simple discrimination problem.

As an addendum, it is known from literature [19, 20] that there is no asymptotic improvement added in the consideration of adaptive protocols when only classical feed-forward is allows, and so beyond constant factors, we need only compare non-adaptive incoherent access and fully coherent access models for our problem. However, although we will not discuss adaptive protocols in the rest of this work, we do depict the performance of one adaptive strategy for an incoherent access protocol in Figure 2 for illustrative purposes. The details of the method employed are given in Appendix A.

This adaptive incoherent protocol asymptotically converges to the measurement which minimizes the error probability, and is strictly better than plain majority vote [23], though not asymptotically better for large N. Moreover, such protocols provably *cannot* achieve perfect discrimination when the hypothesis channels are not orthogonal, i.e., when $\delta \neq \pi$, as expected [19].

III. NOISY DISCRIMINATION

The consideration of noise in channel discrimination is essential, given both that realistic quantum computers exhibit noise and more pressingly that the introduction of noise can impart important nuance on statements of advantage for specific quantum algorithms [1, 24].

This section explores if and when the observed gap in performance between coherent and incoherent access protocols for RDGs discussed in the previous section is robust to finite noise. We will show that even for simple, parameterized noise, this question leads to a transition boundary for quantum advantage among access models. We present these findings in two subsections below, first numerically, then analytically.

A. Quantum advantage for noisy discrimination

In QSP, we can tailor the sequence of phase angles defining the QSP protocol to optimize for discriminating between two channels parameterized by underlying classical distributions, Θ_0 and Θ_1. When the channels are noiseless, i.e., $\sigma = 0$, QSP protocols were shown in the previous section to have better success in hypothesis testing than incoherent access protocols.

In the limit of large noise, we might expect that a coherent protocol can do no better than its incoherent counterpart. Consequently an interesting regime occurs when, by some reasonable metric, the signal to noise ratio for the quantum channel is in an intermediate range. We investigate this region, and give evidence for a transition boundary for relative quantum advantage in the parameter space defining the underlying noise. Additionally, we analytically determine this boundary in two distinct limits where its computation is simplified.

Where we cannot analytically approximate this transition boundary, for $RDG_{\delta,\sigma}$ with $\sigma > 0$, we use Monte Carlo methods to numerically optimize over QSP angles as described in Appendix B. The result for length three QSP protocols (QSP-3) is shown in Figure 3(b). Here, $\overline{p}$ denotes taking an average with respect to the probability distributions defining channel noise, i.e., Eq. (2). For a fair comparison we fix the number of channel applications (the query complexity) afforded to each quantum protocol, so we compare with MAJ for $M = 1$ (equivalently $N = 2M + 1 = 3$) in Figure 3(a). In a side-by-side comparison, it is evident that the QSP-3 protocol has a lower error probability in regions of low δ and low σ. On the other hand, however, both protocols perform poorly in regions of high noise, although this is expected, since in the limit of uniform Θ_0, Θ_1, these distributions become indistinguishable.

The application of QSP to this task is not merely the blind application of a useful ansatz; indeed in Fig. 5 we see that even in regions of non-zero noise, the quantum response function induced by the optimized angles is intuitively obvious, producing a large distance between the induced Bernoulli distributions.

To visualize the region of the quantum advantage more clearly, we show the log ratio of the success probabilities of QSP-3 versus MAJ-3, i.e., $\ln(\overline{p_{\mathrm{QSP}}/p_{\mathrm{MAJ}}})$, in terms of δ and σ in Fig 3(c). The darker blue region indicates the region of advantage, and later analytic results will show that this region extends to both finite δ and σ for any fixed N.

B. Transition boundaries in the low-separation limit

In the previous subsection the relative performance of QSP and incoherent access protocols was investigated numerically. To cross-check the validity of these results, as well as extend them to a region (i.e., low separation, or low δ) in which the simulations are numerically unstable, we can consider the limit of the multi-dimensional integrals which define these success probabilities.

For a noisy RDG, when the angular separation between the means of the possible rotations is sufficiently small, the optimal phase angles of a QSP protocol tend to zero (i.e., toward the *simple QSP protocol* which simply rotates along a fixed great circle on the Bloch sphere, and performs the Helstrom measurement). This limit follows from that such protocols can never achieve perfect discrimination, and thus their success probability is an monotonic function of the maximum possible mean separation between the overall rotation applied by the protocol unitary, in a natural metric on SU(2). This mean separation is maximized by rotating about a fixed axis (i.e., with all QSP angles zero), or else traveling along a locally straight path on this Riemannian manifold. We give a definition of such 'simple' protocols below.

Definition III.1. *Simple QSP protocol. A simple QSP protocol is one which uses the trivial choice of* $\Phi = \{0, 0, \cdots, 0\}$ *as well as a projective measurement which optimizes discrimination success for a particular pair*

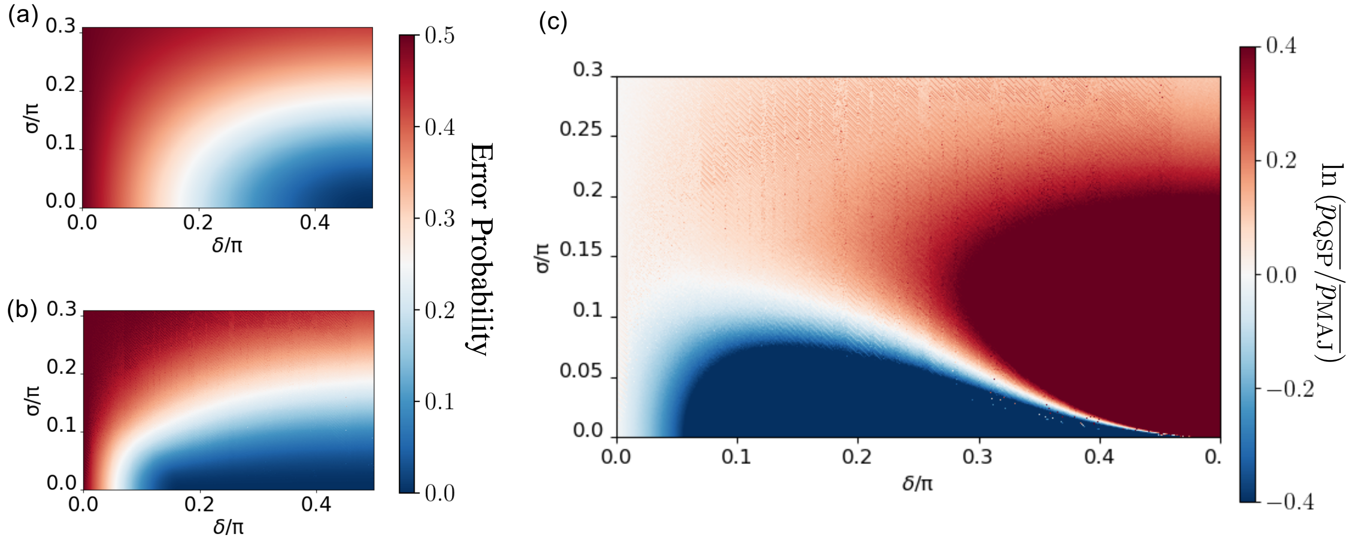

FIG. 3. (a, b) Error probabilities for MAJ-3 (a) and QSP-3 (b). In (c) is depicted the log of the ratio of the error probability of MAJ-3 and that of QSP-3. Red is used for the area where MAJ-3 performs better, white indicates roughly equal performance, and blue indicates where QSP-3 performs better. Quantum advantage for coherent access protocols is present when both the 'signal' (δ) and the 'noise' (σ) are not too large.

of hypothesis channels. Note that this physically corresponds to rotation about a fixed axis.

In this case, we can analytically calculate the expected success probability even in the noisy case and thus identify the transition boundary between coherent and incoherent access protocols. We first compare coherent and incoherent access protocol performance for a fixed odd query complexity $N = 2M + 1$. In this setting it is not too difficult to compute the expected error probability of a simple QSP protocol

$$\overline{p_{\mathrm{QSP}}} = \frac{1}{2}\left(1 - (2M+1)\delta e^{-2(2M+1)\sigma^2}\right), \tag{11}$$

where the angular separation is small enough that $\sin\delta \approx \delta$. Relatedly, we can take the $M = 0$ ($N = 1$) instance of (8), the Helstrom bound, and apply the majority vote (Eq. (6)), again keeping only leading order terms in δ,

$$\overline{p_{\mathrm{MAJ}}} = \frac{1}{2}\left(1 + \delta 2^{-2M}\frac{(2M+1)!}{(M!)^2}e^{-2\sigma^2}\right), \tag{12}$$

which is computed by summing the relevant binomial terms in the expansion of the majority vote function of the Helstrom success probability (see Appendix C). Equating these two in the limit of small angular separation $\delta \ll \pi/2$, one finds the condition for the transition boundary in terms of σ,

$$\sigma = \frac{1}{2}\sqrt{\frac{1}{M}\ln\left(\frac{2^{2M}(M!)^2}{(2M)!}\right)}, \tag{13}$$

which can be shown by approximation to scale as $\sigma^2 \approx \log(M)/M = \log(N)/N$ as N grows sufficiently large. For $M = 1$ (the length $N = 3$ sequences discussed at length in this paper's numerical results), this boundary occurs at $\sigma = \sqrt{\ln 2}/2$, which agrees well with numerical simulation. Qualitatively, we see that the region of advantage, for a fixed noise, grows smaller with increasing query complexity; one might naturally expect this as the Bloch sphere is compact, and thus a standard application of the law of large numbers cannot be applied to arbitrarily accurately threshold about a mean.

IV. HYBRID PROTOCOLS FOR LARGE N AND OPTIMAL COHERENCE LENGTH

We now pose a separate, related, and natural question regarding the protocols for RDGs we have been investigating. We will show that this question leads to yet another limit under which the transition boundary discussed above becomes amenable to analytic description.

Imagine that the querying party has some large budget for the total number of queries they can make to the quantum process defining some RDG. It is known from the previous sections that, if one knows that the underlying distributions for the two possible quantum channels are relatively narrow, then a coherent access protocol is best, while for sufficiently large noise an incoherent access protocol does well.

If the querying party has knowledge of the magnitude of the channel noise then how should they choose to split the difference between coherent and incoherent access protocols? How long should they compute before measuring, where this length (in terms of query complexity) will be termed the *coherence length*. To make this question concrete we need one more definition.

Definition IV.1. *Hybrid protocols. A ξ-hybrid protocol for an RDG is one which, given a total budget of N queries to one among two quantum channels, performs ξ-length QSP protocols a total of N/ξ times, followed by a majority vote on the N/ξ measurement outcomes. While stating nothing of the methods for finding optimal ξ-hybrid protocols, it is easy to see that coherence lengths $\xi = N$ and $\xi = 1$ correspond to fully coherent and fully incoherent access protocols respectively.*

If we consider again (1) the limit of small separation between the means of the distributions defining the two possible quantum channels (i.e., where simple QSP protocols are optimal), and (2) the large N limit, then the work we did before can be re-purposed for differing ξ.

I.e., if we suitably scale (11) and (12) by ξ, we recover that in this limit, as $N \to \infty$, that the error probability for a ξ-hybrid protocol goes as

$$\frac{1}{2}\left(1 - \xi\delta\frac{[N/\xi]!}{[(N/\xi-1)/2)!]^2}2^{-(N/\xi-1)}e^{-2\xi\sigma^2}\right) \tag{14}$$

where N has been taken to replace $2M+1$ in (12) for simplicity, and is the total number of queries. This can easily be minimized over ξ (and in fact this minimum is unique), where $\xi_{\min}$ satisfies the relation

$$(2\sigma^2\xi - 1)\xi - n\log\left(1 - \frac{\xi}{n}\right) = 0, \tag{15}$$

where we have taken the small ξ (and large N) limits in order to replace digamma functions with familiar logarithms. Note that we have assumed that the limit of large N has permitted suitable analytic continuation of the discrete objects in (12). We now hope to further explain what solutions for $\xi_{\min}$ look like in various limits. We define these limits below:

1. First limit: consider large N, large σ (and consequently small $\xi \ll N$).

2. Second limit: consider large N, small σ, (and consequently $\xi \approx N$).

For large noise, and thus small coherence length, we can ignore the second, logarithmic term in Eq. 15 entirely (as it approaches one as ξ approaches one and N grows large,

$$\xi_{\min} \approx \frac{1}{2}\sigma^{-2}, \tag{16}$$

which tracks with intuition; in the limit of large noise (a classical limit) the optimal coherence length decreases according to the inverse variance. In the regime of small $\delta \ll 1/N$, underlying all of these limits, the simple QSP sequences used merely accrue the collected rotation about a single axis on the Bloch sphere.

In the small noise, large N limit (i.e., where the optimal ξ is expected to be large), we are not able to expand the expression given taking the derivative of Eq. 14 with respect to ξ as simply. We may, however numerically verify that up to an including this regime, the scaling of the optimal coherence length retains the same inverse polynomial behavior in σ seen in the low noise regime. While this observation is not exciting from the point of view of obtaining an unexpected quantum advantage, it is also not incomprehensible; this setting exhibits noise with little quantum structure, and thus we would not expect, even in highly coherent regimes, that an algorithm would perform polynomially better.

In Figure 4, numerical results for the optimal coherent length (solid line), and its closed form expression in the large σ limit (dotted dashed line) are shown. The dependence $\xi \propto \sigma^{-2}$ in the large σ region indicates that the optimal length is, as mentioned previously, determined classically, as the coherence length is such so that the accumulated error is kept constant. Even for small sigma (i.e., up to an until clipping in the optimal ξ occurs), the power-law dependence is roughly the same.

Note finally that (15) was obtained by analytic continuation of the factorial function, and thus, $\xi_{\min}$ is not, in principle, bounded between 1 and N. In Fig. 4, $\xi_{\min}$ becomes smaller than one around $\sigma/\pi = 0.2$, indicating that an incoherent access protocol gives a minimum error probability. Similarly, a fully coherent protocol becomes optimal around $\sigma/\pi = 0.02$. Evidently a hybrid protocol is optimal between these limits.

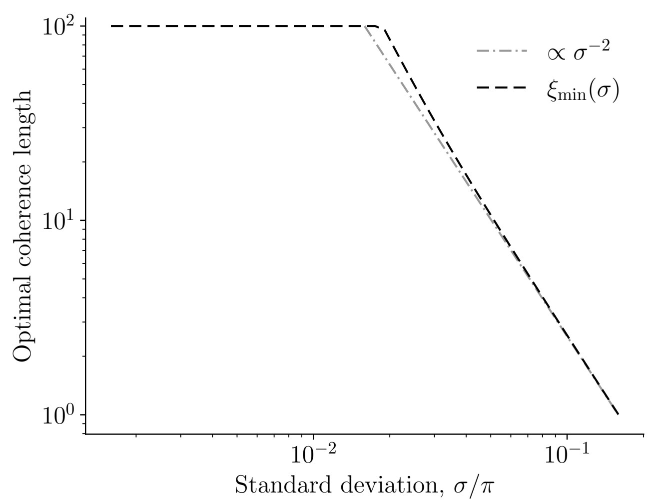

FIG. 4. Optimal coherence length for a hybrid simple QSP sequence for increasing noise, using a log-log scale. Note that not only in the limit of large noise does the coherence length fall as one would classically expect, $\xi \propto \sigma^{-2}$. For small noise the slope implies the same relation holds. This plot shows the case for $N = 100$ as an explicit example (and thus the optimal $\xi_{\min}$ is truncated between $N = 1$ and $N = 100$).

V. CONCLUSION AND DISCUSSION

In this work we show a quantitative difference in performance among quantum protocols for hypothesis testing against pairs of noisy quantum channels. The two families of quantum protocols considered for this task, coherent and incoherent access protocols, were constructively instantiated as QSP protocols (Definition II.2) and non-adaptive thresholding protocols infomed by the Helstrom bound. For these specific protocols, in the case of rotation discrimination games (RDGs; Definition I.1), we provided analytic arguments for their relative performance in multiple limits, as well as numerical results characterizing the conditions under which a crossover in relative advantage appears.

This work determines that for these specific quantum protocols the performance gap between incoherent and coherent access protocol performance is robust to small finite noise, although the region of relative advantage, plotted in space defined by σ (noise strength) and δ (signal strength), shrinks inverse-polynomially with increasing query complexity. We are able to depict this region in terms of these two defining parameters, and verify that numerical approximation of its boundary agrees with said boundary's analytic form in multiple limits.

Moreover, by analyzing hybrid protocols, we were able to compute the ideal method for distributing probes to the underlying quantum process between coherent and incoherent subroutines for optimal performance in RDGs. Here optimality is defined with respect to protocols of the hybrid type given, and such protocols outperform their fully incoherent and fully coherent counterparts. For small noise, we find that the optimal coherence length depends on the noise parameter according to the standard quantum limit, while for large noise the dependence on noise becomes more classical in character.

This work represents a new application of QSP methods to noisy settings, and a concrete series of methods for numerically and analytically investigating the robustness of such protocols to reasonable noise. Moreover, this work situates a physically realistic question of robustness within recent work regarding the complexity-theoretic characterization of quantum algorithms in differing access models [1]. We also provide a complement to the recent quantum machine learning work in [2], proposing a separate noise model and classification scheme which is not amenable to their analysis of adaptiveness, but which is amenable to generalization to large-dimension channels through the methods of QSVT. While the channels considered in this work are simple, they are minimal examples for which QSP-based methods can characterize the presence of quantum advantage.

A major caveat of this work, and various works considering the relative performance of specific quantum algorithmic models, is the non-exhaustive characterization of incoherent and coherent access protocols. We do not consider adaptive hybrid protocols (though we show that in our simple setting, such consideration is not required), or entangled protocols (again unecessary by statements in [11]), which are known in general to offer (sometimes only modest) performance improvements for certain problems [17, 18]. However, for the limits we do consider, small noise and small mean separation, the limited protocols we consider perform well, and are physically reasonable to implement.

Finally while there are results indicating sufficient conditions under which the performance of adaptive protocols in similar discrimination problems tracks exactly with the performance of their non-adaptive counterparts (asymptotically) [19, 20], these bounds are in general difficult to compute outside highly symmetric noise-models (e.g., Haar-random channel access, as in [1]), in the presence of any non-Markovian channel noise, or for channel discrimination which permits coherent feed-forward subroutines. Such settings might permit more extensive applications of QSP methods, as well as hybrid QSP-adaptive protocols, but are beyond the scope of this work.

The broad question of the robustness of the advantage of coherent access protocols over incoherent access ones for *general noise models* remains open, and is not wholly addressed by this work. Indeed, the simple nature of the noise in this work leaves only the possibility of modest quantum advantage in the presence of quite small noise. That said, under additional assumptions or for more structured problems, such as the case of correlated noise or asymmetric hypothesis testing, our finding of classical character for the scaling of the coherence length in Section IV need not hold, leaving the door open to more dramatic statements of quantum advantage.

ACKNOWLEDGEMENTS

The authors thank Yoshihisa Yamamoto for useful discussions and valuable comments. This material is based upon work supported by the U.S. Department of Energy, Office of Science, National Quantum Information Science Research Centers, Co-design Center for Quantum Advantage (C2QA) under contract number DE-SC0012704. We also acknowledge NTT Research for their financial and technical support. ZMR was supported in part by the NSF EPiQC program.

[1] D. Aharonov, J. Cotler, and X.-L. Qi, Quantum algorithmic measurement (2021), arXiv:2101.04634 [quant-ph].

[2] H.-Y. Huang, R. Kueng, and J. Preskill, Information-theoretic bounds on quantum advantage in machine

learning, arXiv , 2101.02464 (2021).
[3] G. H. Low, T. J. Yoder, and I. L. Chuang, Methodology of resonant equiangular composite quantum gates, Phys. Rev. X **6**, 041067 (2016).
[4] G. H. Low and I. L. Chuang, Hamiltonian simulation by qubitization, Quantum **3**, 163 (2019).
[5] G. H. Low and I. L. Chuang, Optimal hamiltonian simulation by quantum signal processing, Phys. Rev. Lett. **118**, 010501 (2017).
[6] A. Gilyén, Y. Su, G. H. Low, and N. Wiebe, Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics, in *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2019 (Association for Computing Machinery, 2019) p. 193204.
[7] C. W. Helstrom, *Quantum Detection and Estimation Theory*, Mathematics in Science and Engineering: A Series of Monographs and Textbooks (Academic Press, Inc., 1976).
[8] A. Acín, Statistical distinguishability between unitary operations, Phys. Rev. Lett. **87**, 177901 (2001).
[9] S. Zhou and L. Jiang, Asymptotic theory of quantum channel estimation, PRX Quantum **2**, 010343 (2021).
[10] R. Duan, Y. Feng, and M. Ying, Perfect distinguishability of quantum operations, Phys. Rev. Lett. **103**, 210501 (2009).
[11] R. Duan, Y. Feng, and M. Ying, Entanglement is not necessary for perfect discrimination between unitary operations, Phys. Rev. Lett. **98**, 100503 (2007).
[12] S. Pirandola, R. Laurenza, C. Lupo, and J. L. Pereira, Fundamental limits to quantum channel discrimination, npj Quantum Inf. **5**, 50 (2019).
[13] Q. Zhuang and S. Pirandola, Ultimate limits for multiple quantum channel discrimination, Phys. Rev. Lett. **125**, 080505 (2020).
[14] R. Demkowicz-Dobrzański, J. Czajkowski, and P. Sekatski, Adaptive quantum metrology under general markovian noise, Phys. Rev. X **7**, 041009 (2017).
[15] S. Zhou, M. Zhang, J. Preskill, and L. Jiang, Achieving the heisenberg limit in quantum metrology using quantum error correction, Nature Communications **9**, 10.1038/s41467-017-02510-3 (2018).
[16] V. Giovannetti, S. Lloyd, and L. Maccone, Advances in quantum metrology, Nature Photonics **5**, 222229 (2011).
[17] R. Takagi and B. Regula, General resource theories in quantum mechanics and beyond: Operational characterization via discrimination tasks, Phys. Rev. X **9**, 031053 (2019).
[18] R. Takagi, B. Regula, K. Bu, Z. Liu, and G. Adesso, Operational advantage of quantum resources in subchannel discrimination, Phys. Rev. Lett. **122**, 140402 (2019).
[19] F. Salek, M. Hayashi, and A. Winter, When are adaptive strategies in asymptotic quantum channel discrimination useful? (2021), arXiv:2011.06569 [quant-ph].
[20] X. Wang and M. M. Wilde, Resource theory of asymmetric distinguishability for quantum channels, Physical Review Research **1**, 10.1103/physrevresearch.1.033169 (2019).
[21] J. M. Martyn, Z. M. Rossi, A. K. Tan, and I. L. Chuang, A grand unification of quantum algorithms (2021), arXiv:2105.02859 [quant-ph].
[22] Z. M. Rossi and I. L. Chuang, Quantum hypothesis testing with group structure (2021), arXiv:2102.02194 [quant-ph].
[23] T. Cooney, M. Mosonyi, and M. M. Wilde, Strong converse exponents for a quantum channel discrimination problem and quantum-feedback-assisted communication, Communications in Mathematical Physics **344**, 797 (2016).
[24] A. W. Cross, G. Smith, and J. A. Smolin, Quantum learning robust against noise, Phys. Rev. A **92**, 012327 (2015).
[25] `https://github.com/mitquanta/noisy-qsp-rdg`.

Appendix A: Adaptive incoherent strategy

We consider adaptive incoherent access protocols based on Bayesian inference, as discussed in the main text. We consider these protocols not because adaptiveness yields asymptotic improvement in discrimination for our setting, but because it is worthwhile to illustrate how such protocols can produce constant-factor improvements. Let the prior probability that Θ_b ($b = 0, 1$) is the correct hypothesis underlying an RDG be $P(\Theta_b)$, i.e., that $P(\Theta_0) + P(\Theta_1) = 1$. Moreover, before the first measurement, take $P(\Theta_b) = 0.5$.

We can update this prior following each measurement using Bayes' theorem. To be more concrete, denote the basis of the adaptive projective measurement by ψ_0 and ψ_1, corresponding to Θ_0 and Θ_1, respectively. We denote conditional probabilities $P(A \mid B)$ in the usual way. The conditional probability we're interested in is:

$$P(\Theta_b|\psi_{b'}) = \frac{P(\psi_{b'}|\Theta_b)P(\Theta_b)}{P(\psi_b)}, \tag{A1}$$

where

$$P(\psi_b) = \sum_{b'=0}^{1} P(\psi_b|\Theta_{b'})P(\Theta_{b'}). \tag{A2}$$

From the prior probability $P(\Theta_b)$ and the conditional probability $P(\psi_{b'}|\Theta_b)$, which we can analytically calculate, we update the probability $P(\Theta_b|\psi_{b'})$ after each measurement. After all measurements have been completed, we threshold to determine the most likely Θ_b.

In this procedure, there is a degree of freedom for choosing the measurement basis, and a good measurement basis is one for which states represented by choice of projector ψ_b are strongly correlated to the hypothesis distributions Θ_b for $b \in \{0, 1\}$. This correlation is made concrete in the maximization of the mutual information:

$$I(\Theta;\psi) = \sum_{b,b'} P(\Theta,\psi) \ln\left(\frac{P(\Theta,\psi)}{P(\Theta)P(\psi)}\right), \tag{A3}$$

where $P(\Theta, \psi)$ is the joint probability distribution $P(\Theta, \psi) = P(\psi_{b'}|\Theta_b)P(\Theta_b)$. In Figure 2, we numerically perform this maximization and plot the resulting discrimination performance, showing a slight improvement. Various results [19, 20] show that this improvement will asymptotically vanish in the large N limit.

Appendix B: Numerical methods

The error probabilities which constitute Fig. 2 and Fig. 3 are generated by numerical optimization on an 0.001×0.001 grid in δ, σ space. The relevant code is available in the `noisy-qsp-rdg` repository on Github [25].

The classical probabilities can be computed via analytical integration. Given two distributions with PDFs Θ_0 and Θ_1, the success probability we're interested in is given by

$$\int_{-\infty}^{\infty} \frac{1}{2}\left(\Theta_0(\theta)\cos^2(\alpha+\theta)+\Theta_1(\theta)\sin^2(\alpha+\theta)\right)d\theta, \tag{B1}$$

where α is chosen to maximize the success probability. Here $\cos^2(\alpha+\theta)$ is the probability of successfully measuring some known $|\psi_0\rangle$ if Θ_0 is chosen and $\sin^2(\alpha+\theta)$ is the probability of successfully measuring some known $|\psi_1\rangle$ perpendicular to $|\psi_0\rangle$ if Θ_1 is chosen. For normal distributions $\Theta_0 = \mathcal{N}(0,\sigma^2)$ and $\Theta_1 = \mathcal{N}(\delta,\sigma^2)$, the integral evaluates to

$$\frac{1}{4}\left(\left[1+e^{-2\sigma^2}\cos(2\alpha)\right]+\left[1-e^{-2\sigma^2}\cos(2(\alpha+\theta))\right]\right), \tag{B2}$$

with the optimal $\alpha = \frac{\pi}{4}-\frac{\delta}{2}$ given by the Helstrom bound [7]. The error probability is equal to the complement of this value, and majority votes are applied to this base probability in the calculation shown in Eq. 6.

For a given list of QSP-N phase angles Φ with prepared initial and final (i.e., measurement basis) states ψ and ψ', the success probability of the QSP protocol Q_Φ is simply

$$p_{\Phi,\psi,\psi'} = \int \Theta(\theta_0)\Theta(\theta_1)\Theta(\theta_2)|\langle\psi'|Q_\Phi(\theta)|\psi\rangle|^2\, d\theta_0\, d\theta_1\, d\theta_2. \tag{B3}$$

However, analytically integrating this is difficult, so we numerically calculate this by Monte Carlo methods. Specifically, we evaluate the success probability by random sampling the introduced θ (i.e., N samples per sequence):

$$p^{\text{approx}}_{\Phi,\psi,\psi'} = \frac{1}{N_r}\sum_r |\langle\psi'\,|\,Q_\Phi(\theta_r)\,|\,\psi\rangle|^2 \tag{B4}$$

where θ_r $(r = 1, 2, \cdots, N_r)$ are prepared according to the PDF of Θ. Note that p^{approx} converges to the exact probability p when $N_r \to \infty$.

Another merit of using (B4) instead of (B3) is that we can readily optimize Φ on a computer, e.g. with `scipy.optimize`. We optimize both the set of QSP angles Φ and the prepared states ψ and ψ' to minimize the error probability. Since there are many local minima given $N+4$ degrees of freedom, a single run may not generate the global optimum. Hence we run the optimization several times and take the best set of angles.

For a concrete view into whether these optimizations are fundamentally sensible, one can look at Fig. 5, which depicts the induced quantum response functions for fixed non-zero σ and varying δ. Evidently one observes that the optimization does indeed tend toward a unitary and pair of preparation and measurement which produces a large difference in probability defining the resulting Bernoulli distributions for each hypothesis distribution.

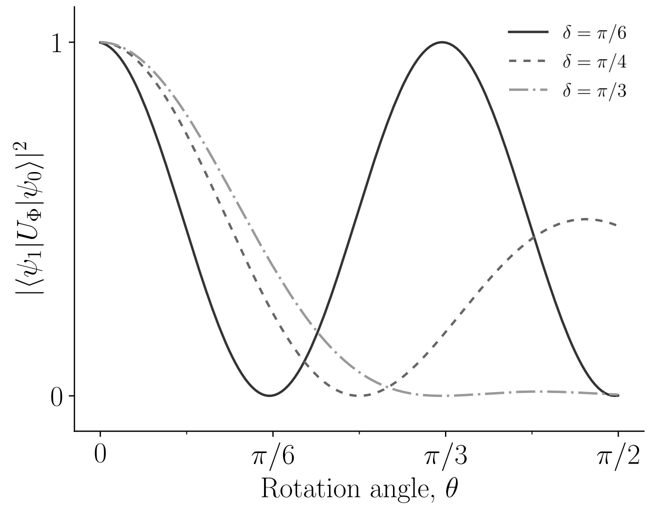

FIG. 5. Transition probability induced by a numerically optimized QSP-3 unitary from an optimized initial state $|\psi_0\rangle$ to an optimized final state $|\psi_1\rangle$ for fixed (small) $\sigma = 0.05$ and differing $\delta \in \{\pi/6, \pi/4, \pi/3\}$. Note that for the first hypothesis (implicitly $\delta = 0$), the transition probability is near 1, while for the chosen alternative δ it is near zero, indicating near perfect discrimination in the case of small noise and reasonably large δ.

The fact that the functions depicted in Fig. 5 are well-defined trigonometric polynomials further bolsters the intuition that smaller δ necessitates higher degree polynomials and thus longer QSP sequences for near-optimal discrimination.

Appendix C: Majority vote for Helstrom bound

We compute the success probability of a fully incoherent non adaptive protocol for an RDG. Given access to a binary process whose success probability follows, in the limit of small δ, the form

$$p = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{(\theta-\delta)^2/2\sigma^2}\frac{1}{2}(1+\sin\theta), \tag{C1}$$

which has the solution by simple integration methods

$$p = \frac{1}{2}(1+\delta e^{-2\sigma^2}), \tag{C2}$$

then the expected success probability of the classical statistical process which takes $2M+1$ samples from this Benoulli distribution and performs majority vote is

$$p' = \frac{1}{2}\left(1+\delta 2^{-2M}\frac{(2M+1)!}{(M!)^2}e^{-2\sigma^2}\right). \tag{C3}$$

Proof of this statement follows from applying the majority vote function to p

$$\sum_{j=0}^{M+1} \binom{2M+1}{j} p^{2M+1-j}(1-p)^j, \tag{C4}$$

and keeping only the first-order terms in δ (equivalently assuming δ small). This results in two terms when the sums are collected

$$2^{-(2M+1)}\left[\sum_{j=0}^{M+1} \binom{2M+1}{j}\right] + \tag{C5}$$

$$2^{-(2M+1)}\left[\sum_{j=0}^{M} \binom{2M+1}{j}(2M+1-2j)\right], \tag{C6}$$

the first of which is simply 1/2 by the known symmetry and total sum of the binomial coefficients, and the later of which is a known identity in combinatorics in terms of hypergeometric functions, namely providing the partial sum of j against binomial coefficients depending on j. Collecting the sum results in the previously given expression, both in Eq. C3 and Eq. 12. It is an interesting generalization to this method to compute the higher order δ terms, which have similar character, though that is not shown here.