Quantum state discrimination circuits inspired by Deutschian closed timelike curves

Christopher Vairogs [1,2] Vishal Katariya [2] and Mark M. Wilde [2]

[1]*Department of Mathematics and Department of Physics, University of Florida, Gainesville, Florida 32611, USA*

[2]*Hearne Institute for Theoretical Physics, Department of Physics and Astronomy, and Center for Computation and Technology, Louisiana State University, Baton Rouge, Louisiana 70803, USA*

It is known that a party with access to a Deutschian closed timelike curve (DCTC) can perfectly distinguish multiple nonorthogonal quantum states. In this paper we propose a practical method for discriminating multiple nonorthogonal states, by using a previously known quantum circuit designed to simulate DCTCs. This method relies on multiple copies of an input state, multiple iterations of the circuit, and a fixed set of unitary operations. We first characterize the performance of this circuit and study its asymptotic behavior. We also show how it can be equivalently recast as a local adaptive circuit that may be implemented simply in an experiment. Finally, we prove that our state discrimination strategy achieves the multiple Chernoff bound when discriminating an arbitrary set of pure qubit states.

I. INTRODUCTION

Closed timelike curves (CTCs) arise as solutions to the Einstein field equations in general relativity. While the existence of CTCs is unverified, they bring up the possibility of time travel and the paradoxes associated with it [1]. To better understand the properties of these objects, several quantum information theoretic models of CTCs have been proposed [2–5].

One such CTC model is that given by Deutsch [2], where paradoxes associated with time travel using CTCs are resolved by a self-consistency condition. This self-consistency condition introduces a nonlinearity in the evolution of a quantum state through a Deutschian CTC (DCTC). Standard quantum mechanics demands that the evolution of an arbitrary state is linear, which places restrictions on physical evolutions, such as the no-cloning theorem [6–8] and the Heisenberg uncertainty principle.

Thus, in contrast to standard quantum mechanics, this nonlinearity allows for many remarkable characteristics associated with DCTCs beyond what is allowed by standard quantum mechanics. Deutschian CTCs may be utilized to violate the no-cloning theorem [9], the Holevo bound [10], and the second law of thermodynamics [11] and enable quantum computers to solve problems in the computational complexity class PSPACE [12]. (However, note that these claims have been debated in the literature [13,14].)

The aspect of DCTCs that we are most interested in here is their use in perfectly distinguishing multiple nonorthogonal quantum states, violating Heisenberg's uncertainty principle [10]. We use ideas contained in the DCTC-assisted state discrimination method to create a practical iterative state discrimination circuit that works by approximating the behavior of a DCTC.

Even though DCTCs are inaccessible, we may simulate the evolution of the state of a system traveling along a DCTC. Such simulations are important to us not only because they allow us to gain a better understanding of the properties of DCTCs in an accessible setting, but also because they enable us to exploit their unique characteristics for applications. Simulating a DCTC is directly related to computing the fixed point of a quantum channel, which is a difficult task [12]. One DCTC simulation method uses polarization-encoded photons as qubits [15], which involves computing the self-consistent solution for the state of a system traveling along a DCTC. This computation is practical for simple quantum systems, but it becomes prohibitively expensive for larger systems. Circumventing this issue, the authors of [16] proposed a method for simulating CTCs that uses an iterative quantum circuit, with the circuit approaching the behavior of a DCTC with an increasing number of iterations. It is also "self-contained" in the sense that it does not involve the discarding of experimental data, unlike that in [15].

Strategies for discriminating nonorthogonal quantum states have been analyzed in many different contexts. In this paper we restrict our attention to the minimum-error discrimination of pure states. The case of discriminating two quantum states has been well studied [17–20]. Optimal minimum-error approaches for discriminating multiple quantum states have been characterized in a variety of specific cases [21–26]. A strategy using a theoretical apparatus for distinguishing multiple arbitrary quantum states was proposed in [27]. Our method is similar in that it applies to the general task of discriminating multiple (and possibly nonorthogonal) quantum states.

The major contribution of our paper is a practical state discrimination strategy for multiple nonorthogonal pure states that combines the DCTC simulation circuit of [16] and the CTC-assisted state discrimination strategy of [10]. We briefly state our strategy here. Assume that the set of states to be

discriminated is $\{|\psi_i\rangle\}_{i=0}^{N-1}$ and that we are given n copies of an unknown state randomly selected from this set. Assume that measurements are made in the basis $\{|i\rangle\}_{i=0}^{N-1}$. Suppose that we have access to a set of unitary operations $\{U_i\}_{i=0}^{N-1}$ such that $U_i|\psi_i\rangle = |i\rangle$. Suppose that the n copies of the unknown state are ordered. Perform a measurement on the first copy of the unknown state. If outcome j is obtained, then apply the unitary operation U_j to the second copy of the unknown state and perform a measurement on the resulting state. If outcome k is obtained, then apply the unitary operation U_k to the third copy of the unknown state. Repeat this procedure until a unitary operation has been applied to all n states. If outcome l is obtained after measuring the nth state, then claim that the unknown state is the state $|\psi_l\rangle$.

Our strategy proves advantageous because it relies only on a fixed set of local operations and may be implemented by performing measurements on each copy of a state, with each successive measurement being determined by the outcome of the previous one, i.e., a local adaptive strategy [18–20]. In this way, our state discrimination strategy is amenable to a practical experimental implementation. [See Fig. 3(c) for a schematic of the method.]

Furthermore, we calculate the asymptotic rate of decay of the average probability of error of our state discrimination circuit, which we use to show that our state discrimination scheme attains the fundamental limit, i.e., the multiple Chernoff bound [28,29], for the general task of discriminating an arbitrary set of pure qubit states. This is another desirable property that our state discrimination scheme possesses.

The rest of our paper is structured as follows. We provide some preliminaries and set up notation in Sec. II. Then, in Sec. III, we provide our state discrimination circuit and also show how it can be implemented as a local adaptive circuit. In Sec. IV we calculate its average probability of error in distinguishing states. We show that this probability of error converges to zero in the limit of infinitely many iterations of the circuit. Finally, in Sec. V, we consider two examples and show how our scheme achieves the multiple Chernoff bound when discriminating an arbitrary set of pure qubit states.

II. PRELIMINARIES

A. State discrimination

We first describe the problem of state discrimination considered in this paper. The goal is to distinguish the states in the set $\{\rho_i\}_{i=0}^{N-1}$, where the state ρ_i is chosen with probability p_i. The minimum error approach to state discrimination may be pictured as the following game between Alice and Bob. Alice and Bob agree on the set $\{\rho_i\}_{i=0}^{N-1}$ of quantum states and probability distribution $\{p_i\}_i$ that they will use. Alice prepares a state ρ_j from that set with probability p_j and sends it to Bob. Bob then, in an attempt to identify Alice's state, performs a measurement described by the set $\{M_i\}_i$ of measurement operators. He guesses that the state Alice prepared is ρ_k if he measures outcome k. Bob's goal is to find the measurement that minimizes his average probability of error, defined as follows:

$$p_e := \sum_{k=0}^{N-1} p_k \sum_{j\neq k} \text{Tr}\{M_j\rho_k\} = 1 - \sum_{k=0}^{N-1} p_k \,\text{Tr}\{M_k\rho_k\}. \quad (1)$$

In the case when $N = 2$, the Helstrom measurement is an optimal measurement [17].

An alternative approach to the state discrimination problem is to assume that Bob has n available copies of the state ρ_j that Alice selects. In this case, an optimal measurement is a collective measurement on all n states $\rho_j^{\otimes n}$. In the limit of large n, the optimal error probability p_e^{opt} decays exponentially as

$$p_e^{\text{opt}} \sim e^{-n\xi^{\text{opt}}}, \quad (2)$$

where the value ξ^{opt} is known as the multiple Chernoff bound [29] and is given by

$$\xi^{\text{opt}} = -\ln\left[\max_{i\neq j} \min_{0\leqslant s\leqslant 1} \text{Tr}\left\{\rho_i^s \rho_j^{1-s}\right\}\right]. \quad (3)$$

The multiple Chernoff bound places a fundamental limit on how fast the error probability decays for a multiple-copy state discrimination scheme [29]. See [28] for the special case when all of the states in the set $\{\rho_i\}_i$ are pure.

B. Deutschian closed timelike curves

Next we describe the model for CTCs that is applicable to our work, namely, the Deutschian (DCTC) model [2]. We note that there also exist other models for quantitatively describing the behavior of CTCs, namely, postselected quantum teleportation CTCs [4] and transition probability CTCs [5].

The DCTC model involves two subsystems: the chronology-respecting (CR) system S, which does not travel through the CTC, and the chronology-violating (CV) system C, which does travel through the CTC. The different CTC models differ in the manner in which they resolve or avoid causality paradoxes. In the DCTC model, this is accomplished by requiring the state of the chronology-violating system to be a fixed point of the evolution that results from the interaction between the CR and CV systems.

Let σ_C be the state of the CV system and let ρ_S be the state of the CR system. In Deutsch's model, systems S and C are assumed to be in a tensor-product state $\rho_S \otimes \sigma_C$ before they interact unitarily via the CTC. They then interact according to an interaction unitary V_{SC} before the CV system enters the future mouth of its wormhole, so that the state of the composite system after the evolution is $V_{SC}(\rho_S \otimes \sigma_C)V_{SC}^\dagger$. We refer to states of the CR system as the CV system emerges from the past mouth of its wormhole and as it enters the future mouth of its wormhole as "initial" CR states and "final" CR states, respectively. We define initial CV states and final CV states similarly.

The evolution of the CV system is represented by the quantum channel

$$\mathcal{N}_{V,\rho} : \sigma_C \mapsto \text{Tr}_S\{V_{SC}(\rho_S \otimes \sigma_C)V_{SC}^\dagger\}, \quad (4)$$

which maps each possible initial CV state to its corresponding final CV state when the initial CR state is ρ_S. Furthermore, as stated earlier, the DCTC model enforces a self-consistency condition so as to avoid "grandfatherlike" causality paradoxes. This requires that an arbitrary state σ_C of the CV system is a fixed point of the quantum channel $\mathcal{N}_{V,\rho}$, i.e.,

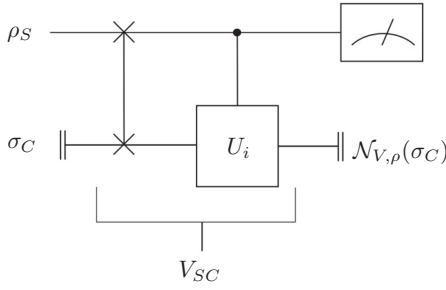$$\mathcal{N}_{V,\rho}(\sigma_C) = \sigma_C. \quad (5)$$

FIG. 1. The upper system is the CR system and the lower system is the CV system. The past and future mouths of the wormhole are represented by the double bars on the left and right. The CR and CV systems interact according to the unitary V_{SC}, defined in (7), before the CV system enters the future mouth of the wormhole.

A solution σ_C to (5) always exists, although it is not necessarily unique [2]. The evolution of the CR system is represented by

$$\mathcal{M}_V : \rho_S \mapsto \mathrm{Tr}_C\{V_{SC}(\rho_S \otimes \sigma_C)V_{SC}^\dagger\}, \qquad (6)$$

which maps every possible initial CR state to its corresponding final CR state whenever the CV state is σ_C. Note that $\mathcal{M}_V(\rho_S)$ depends not only on ρ_S, but also on σ_C, whose dependence on ρ_S is given by (5). Therefore, $\mathcal{M}_V$ is not a linear map. The nonlinear nature of this evolution leads to the host of interesting properties mentioned earlier in Sec. I.

III. CTC-INSPIRED STATE DISCRIMINATION CIRCUIT

Before we provide our CTC-inspired state discrimination circuit, we recall how to discriminate multiple nonorthogonal pure states using a DCTC [10]. The set of states to be discriminated is $\{|\psi_i\rangle\langle\psi_i|\}_{i=0}^{N-1}$, and at the end of the DCTC-assisted interaction, we may perform a measurement in the orthonormal basis $\{|i\rangle\}_{i=0}^{N-1}$. To do so, we require a set of unitaries $\{U_i\}_{i=0}^{N-1}$ such that $U_i|\psi_i\rangle = |i\rangle$ and $\langle j|U_i|\psi_j\rangle \neq 0$ for all $0 \leqslant i$, $j \leqslant N-1$, and $i \neq j$. Such a set of unitaries exists for every set $\{|\psi_i\rangle\langle\psi_i|\}_{i=0}^{N-1}$ of states [10]. Let the interaction unitary V_{SC} of the DCTC be

$$V_{SC} = \left(\sum_{i=0}^{N-1} |i\rangle\langle i|_S \otimes (U_i)_C\right) \circ \mathrm{SWAP}, \qquad (7)$$

where SWAP is the unitary operator that swaps systems S and C. The circuit representation of this unitary is shown in Fig. 1. The authors of [10] demonstrated that $\sigma_C = |i\rangle\langle i|$ is the unique self-consistent solution to (5) whenever $\rho_S = |\psi_i\rangle\langle\psi_i|$. That is, the DCTC-assisted circuit can be used to map nonorthogonal states to distinct orthogonal basis states, and hence one can perfectly discriminate the nonorthogonal states in question.

To outline the functioning of the circuit, we briefly explain why each $\sigma_C = |i\rangle\langle i|$, for $0 \leqslant i \leqslant N-1$, is a fixed point of $\mathcal{N}_{V,\rho}$, i.e., a solution to (5). Suppose that $\rho_S = |\psi_i\rangle\langle\psi_i|$ and $\sigma_C = |i\rangle\langle i|$. The SWAP gate acts first and transforms the CR system to the state $|i\rangle\langle i|$. Next the U_i unitary is triggered. The unitary U_i acts on $|\psi_i\rangle\langle\psi_i|$, which leads to the self-consistency condition $\mathcal{N}_{V,\rho}(\sigma_C) = |i\rangle\langle i|$. Now if one performs
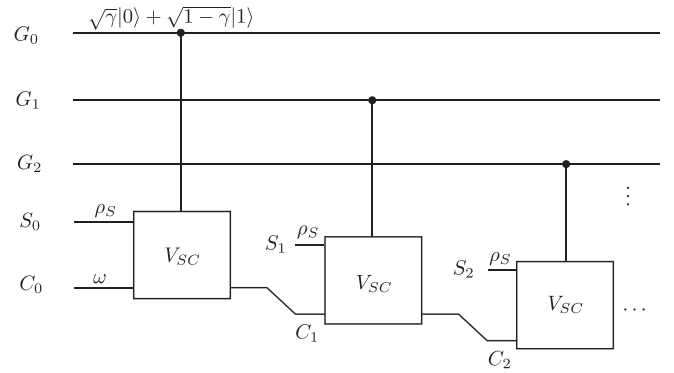

FIG. 2. This circuit simulates a DCTC with interaction unitary V_{SC}. This figure shows three rounds of interaction. As the number of rounds n increases, the simulation approaches the behavior of a DCTC.

a measurement in the basis $\{|i\rangle\}_i$ on the final CR state, one may determine ρ_S with certainty. Measurement outcome j corresponds to the initial state $\rho_S = |\psi_j\rangle\langle\psi_j|$. Thus, using the DCTC, in principle, we are able to discriminate perfectly the possibly nonorthogonal states $\{|\psi_i\rangle\langle\psi_i|\}_{i=0}^{N-1}$.

The construction outlined above works if we have access to a DCTC. In its absence, we can only construct iterative circuits that approximate its behavior. Our major contribution is one such circuit; i.e., we combine the CTC-assisted state discrimination scheme described above with a version of a quantum circuit given by [16] that simulates the behavior of a DCTC. The circuit consists of multiple copies of three registers whose nth copies we label G_n, S_n, and C_n (Fig. 2). We will initialize each of the G_n registers to the state vector $\sqrt{\gamma}|0\rangle + \sqrt{1-\gamma}|1\rangle$. We find, however, that for our purposes, it suffices to set $\gamma = 0$. Each of the S_n registers is initialized to the initial CR state ρ_S and the C_0 register is initialized to a state ω. At every step of the circuit, the G_n, S_n, and C_n systems interact via the controlled unitary

$$|0\rangle\langle 0|_{G_n} \otimes I_{S_n C_n} + |1\rangle\langle 1|_{G_n} \otimes V_{S_n C_n}, \qquad (8)$$

where $V_{S_n C_n}$ is the interaction unitary in (7) acting on the S_n and C_n registers. In other words, the procedure is the following.

(1) Apply the controlled unitary between the G_n, S_n, and C_n registers.

(2) Discard the S_n and G_n registers.

(3) Load the resulting state of the C_n system into the C_{n+1} register, increment n by one, and repeat.

By applying this procedure, for every ρ_S, the state of the C_n register converges to the fixed point of $\mathcal{N}_{V,\rho}$ as $n \to \infty$ [16]. Note that the rate of convergence is dependent on V_{SC} and the state ω to which C_0 is initialized. We will discuss how to optimize the rate of convergence with respect to them.

To use this circuit, which is also shown in Fig. 2, in order to discriminate the set $\{|\psi_i\rangle\langle\psi_i|\}_i$ of states, the interaction unitary V_{SC} is set to the one in (7) and the input state ρ_S is restricted to belong to the set $\{|\psi_i\rangle\langle\psi_i|\}_i$. Since the state of the C_n register converges to the fixed point of $\mathcal{N}_{V,\rho}$, the

C_n register converges to the state $|j\rangle\langle j|$ if $\rho_S = |\psi_j\rangle\langle\psi_j|$ [10]. Therefore, by implementing a standard basis measurement on the C_n register after a suitably chosen n, we have a method for approximately discriminating the states in the set $\{|\psi_i\rangle\langle\psi_i|\}_i$.

One of our major contributions is an equivalent recasting of our state discrimination circuit as a local adaptive circuit. The adaptive circuit, which we describe below, consists of repeated iterations of a fixed two-register protocol and lends itself directly to experimental implementation. Another benefit of the adaptive circuit is that it is not necessary to update a quantum memory throughout the length of the circuit. The only information necessary to store in a quantum memory is the n copies of the unknown state $|\psi_i\rangle$.

We now explain how the original circuit in Fig. 2 can be recast in the simple adaptive form of Fig. 3(c). To perform the state discrimination in an adaptive manner, we first assume that we have n copies of the unknown state $|\psi_i\rangle$. By expanding the V_{SC} unitaries defined in (7) and by applying the fact that the state $|1\rangle$ suffices for each of the G_n registers in Fig. 2, it can be simplified to the circuit in Fig. 3(a). We then note that since the various S registers are traced out at the end of the circuit, the coherent controls in Fig. 3(a) can be equivalently replaced with classical controls. That further means that the classical control is effectively implemented by first performing a measurement and then choosing the corresponding U_i gate using the measurement outcome. What we are applying here is the well-known principle of deferred measurement [30]. This enables a further simplification of the circuit to that in Fig. 3(b).

Finally, we note that the circuit in Fig. 3(b) consists of iterations of a single atomic unit, which we denote concisely in Fig. 3(c). This depiction of our state discrimination scheme is of particular import: To perform our state discrimination circuit, one only needs two quantum registers, the ability to perform a standard basis measurement, and classical control. This makes it directly amenable to experimental implementation.

IV. ERROR ANALYSIS OF THE STATE DISCRIMINATION CIRCUIT

Here we define the average probability of error in our state discrimination scheme. Next we demonstrate how it decays with n, the number of iterations of the state discrimination circuit in Fig. 2. Finally, we show how the error probability exponentially converges to zero in the asymptotic limit.

A. Calculating the error probability

As earlier, we assume that we are given a set $\{|\psi_i\rangle\}_{i=0}^{N-1}$ of states, a set $\{|i\rangle\}_{i=0}^{N-1}$ of basis states, and a set $\{U_i\}_{i=0}^{N-1}$ of unitaries such that $U_i|\psi_i\rangle = |i\rangle$ for $0 \leqslant i \leqslant N-1$. We assume that the states $\{|\psi_i\rangle\}_{i=0}^{N-1}$ are to be discriminated using the circuit shown in Fig. 2 with

$$V_{SC} := \left(\sum_{i=0}^{N-1} |i\rangle\langle i| \otimes U_i \right) \circ \text{SWAP}. \tag{9}$$
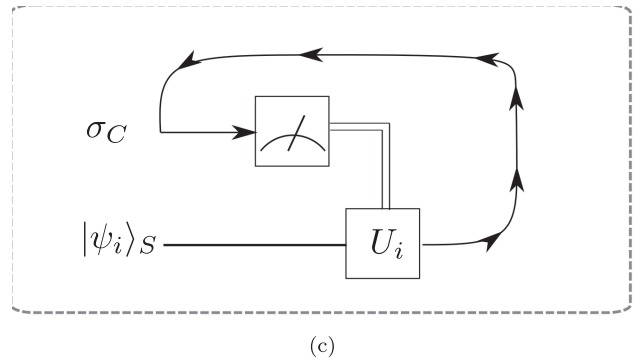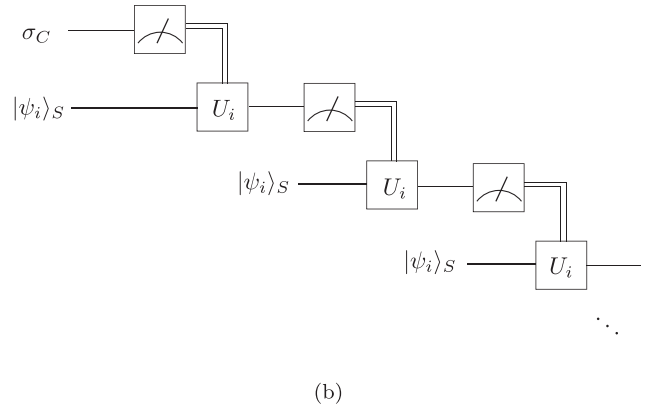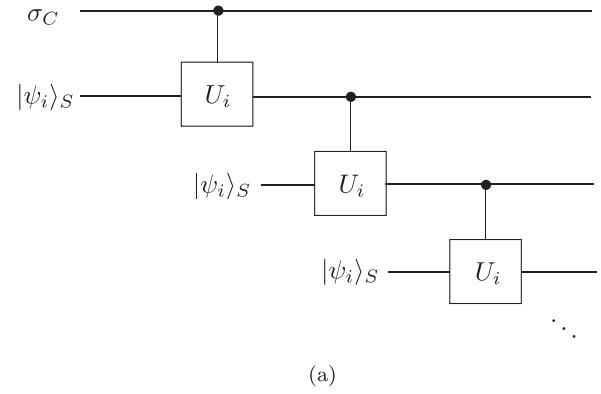

(a)

(b)

(c)

FIG. 3. Series of simplifications of the original state discrimination circuit in Fig. 2 that ultimately leads to the simple iterative circuit in Fig. 3(c). First, we explicitly expand V_{SC} and set the state of the G_n registers to $|1\rangle\langle 1|$, as this is optimal for performing our state discrimination scheme. This leads to the circuit in (a). We also note that the controlled-U_i gate is short for the gate $\sum_i |i\rangle\langle i| \otimes U_i$. Further, as the controlling registers are effectively classical, the circuit can be simplified to that in (b). Finally, the iterative circuit is rewritten in the compact form of (c). The arrows do not depict a time-travel loop, but instead depict that the output of the unitary U_i at one time interval is fed as input to the C system at the next time interval.

First, we provide an expression for the average probability of incorrectly discriminating the set of states after n iterations of the circuit.

Definition 1. For each $0 \leqslant k \leqslant N-1$, let $\sigma_{n,k}$ be the state of the C_n system when $\rho_S = |\psi_k\rangle\langle\psi_k|$. (That is, $\sigma_{n,k}$ is the state

obtained after n rounds of executing the circuit in Fig. 2 with ρ_S set equal to $|\psi_k\rangle\langle\psi_k|$.) We define the average probability of error $p_e^{(n)}$ and the average probability of success $p_s^{(n)}$ after n iterations as follows:

$$p_e^{(n)} := \sum_{k=0}^{N-1} \sum_{j \neq k} p_k \, \mathrm{Tr}\{|j\rangle\langle j|\sigma_{n,k}\}, \tag{10}$$

$$p_s^{(n)} := \sum_{k=0}^{N-1} p_k \, \mathrm{Tr}\{|k\rangle\langle k|\sigma_{n,k}\}. \tag{11}$$

Our goal now is to quantitatively describe the behavior of $p_e^{(n)}$ with the number n of iterations. We will find that the functioning of our circuit, particularly the error and success probabilities, possesses a Markov property. To obtain this property, for each $0 \leqslant k \leqslant N-1$, we define the $N \times N$ stochastic matrix P_k to be

$$P_k := \begin{pmatrix} |\langle 0|U_0|\psi_k\rangle|^2 & \cdots & |\langle 0|U_{N-1}|\psi_k\rangle|^2 \\ \vdots & \ddots & \vdots \\ |\langle N-1|U_0|\psi_k\rangle|^2 & \cdots & |\langle N-1|U_{N-1}|\psi_k\rangle|^2 \end{pmatrix}. \tag{12}$$

Also, we define, for $0 \leqslant k \leqslant N-1$, the $N \times 1$ column vector $u_k^{(n)}$ such that its ith element is

$$\mathrm{Tr}\{|i-1\rangle\langle i-1|\sigma_{n,k}\}. \tag{13}$$

Define $u^{(0)}$ to be $u_k^{(0)}$ for any $0 \leqslant k \leqslant N-1$. The vector $u^{(0)}$ is well defined since $\rho_{0,k} = \omega$ for all $0 \leqslant k \leqslant N-1$, where ω is defined just before (8).

Proposition 1. The average probability of error and success after n iterations, denoted by $p_e^{(n)}$ and $p_s^{(n)}$, are respectively given by

$$p_e^{(n)} = \sum_{k=0}^{N-1} \sum_{\substack{j=0, \\ j \neq k}}^{N-1} p_k e_{j+1}^T P_k^n u^{(0)}, \tag{14}$$

$$p_s^{(n)} = \sum_{k=0}^{N-1} p_k e_{k+1}^T P_k^n u^{(0)}, \tag{15}$$

where e_j is the standard basis column vector with a one in the jth row and all other elements set to zero.

Proof. The probability of measuring outcome j on the $(n+1)$th iteration of the circuit, assuming the unitary V_{SC} always acts on the S_n and C_n registers, is

$$\mathrm{Tr}\{|j\rangle\langle j|\mathcal{N}_{V,\rho}(\sigma_{n,k})\}$$

$$= \mathrm{Tr}\{|j\rangle\langle j|[\mathrm{Tr}_S\{V_{SC}(|\psi_k\rangle\langle\psi_k| \otimes \sigma_{n,k})V_{SC}^\dagger\}]\}$$

$$= \mathrm{Tr}\left\{|j\rangle\langle j|\left(\sum_{l=0}^{N-1} \mathrm{Tr}\{|l\rangle\langle l|\sigma_{n,k}\}U_l|\psi_k\rangle\langle\psi_k|U_l^\dagger\right)\right\}$$

$$= \sum_{l=0}^{N-1} \mathrm{Tr}\{|l\rangle\langle l|\sigma_{n,k}\}|\langle j|U_l|\psi_k\rangle|^2, \tag{16}$$

where we used the definition of $\mathcal{N}_{V,\rho}$ given in (4) in the first equality and the definition of V_{SC} given in (7) (including the SWAP operation) in the second equality.

It follows from (16) that $u_k^{(n+1)} = P_k u_k^{(n)}$. From this we conclude that

$$u_k^{(n)} = P_k^n u^{(0)}. \tag{17}$$

Since $U_k|\psi_k\rangle = |k\rangle$, the $(k+1)$th column of P_k consists of zeros except for a one in the $(k+1)$th row.

Hence, we have

$$p_e^{(n)} = \sum_{k=0}^{N-1} \sum_{j=0: j \neq k}^{N-1} p_k \, \mathrm{Tr}\{|j\rangle\langle j|\sigma_{n,k}\} \tag{18}$$

$$= \sum_{k=0}^{N-1} \sum_{j=0: j \neq k}^{N-1} p_k e_{j+1}^T u_k^{(n)} \tag{19}$$

$$= \sum_{k=0}^{N-1} \sum_{j=0: j \neq k}^{N-1} p_k e_{j+1}^T P_k^n u^{(0)}. \tag{20}$$

In the above, the first equality arises due to the definition of $p_e^{(n)}$. The second equality is due to the definition of $u_k^{(n)}$ and the final equality is due to (17). We also have that

$$p_s^{(n)} = \sum_{k=0}^{N-1} p_k \, \mathrm{Tr}\{|k\rangle\langle k|\sigma_{n,k}\} \tag{21}$$

$$= \sum_{k=0}^{N-1} p_k e_{k+1}^T u_k^{(n)} \tag{22}$$

$$= \sum_{k=0}^{N-1} p_k e_{k+1}^T P_k^n u^{(0)}. \tag{23}$$

This concludes the proof. ∎

The functioning of our quantum circuit has a Markov property, which we explain here. Recall from earlier that the state discrimination circuit in Fig. 2 can be implemented as an adaptive quantum state discrimination scheme (Fig. 3). The adaptive scheme is as follows. One starts with n copies of ρ_S. If one obtains outcome k after measuring the nth copy of ρ_S, then one applies the unitary U_k to the $(n+1)$th copy of ρ_S and repeats the procedure. Suppose that we have $\rho_S = |\psi_k\rangle\langle\psi_k|$ and that we obtain outcome i after measuring the nth copy. The probability of obtaining the outcome j after measuring the $(n+1)$th copy is $|\langle j|U_i|\psi_k\rangle|^2$. That is, the probability of obtaining each successive measurement outcome given the previous measurement outcome depends only on the previous measurement outcome. Thus, we may view each successive measurement outcome as an element of a Markov chain with transition matrix P_k as defined in (12).

We now state and prove a simple result that allows us to maximize the probability of success after n iterations with respect to ω, the initial state of the C_0 register.

Proposition 2. For each $0 \leqslant i \leqslant N-1$, let $p_{s,i}^{(n)}$ be the average probability of success after n iterations given that $\omega = |i\rangle\langle i|$. Then the average probability of success (for arbitrary ω) is

$$p_s^{(n)} = \sum_{i=0}^{N-1} \mathrm{Tr}\{|i\rangle\langle i|\omega\}p_{s,i}^{(n)}. \tag{24}$$

It follows from Proposition 1 that

$$p_s^{(n)} = \sum_{k=0}^{N-1} p_k e_{k+1}^T P_k^n \left(\sum_{i=0}^{N-1} \text{Tr}\{|i\rangle\langle i|\omega\} e_{i+1} \right) \quad (25)$$

$$= \sum_{i=0}^{N-1} \text{Tr}\{|i\rangle\langle i|\omega\} \left(\sum_{k=0}^{N-1} p_k e_{k+1}^T P_k^n e_{i+1} \right) \quad (26)$$

$$= \sum_{i=0}^{N-1} \text{Tr}\{|i\rangle\langle i|\omega\} p_{s,i}^{(n)}. \quad (27)$$

The first equality arises due to Proposition 1 and the definition of $u^{(0)}$. The second equality arises due to algebraic manipulation and the final equality follows from identifying that $p_{s,i}^{(n)} = \sum_{k=0}^{N-1} p_k e_{k+1}^T P_k^n e_{i+1}$. ∎

It follows from Proposition 2 that the success probability after n iterations is maximized if ω is simply a basis state of the form $|j\rangle\langle j|$. The particular optimal basis state is dependent on the value of j that maximizes $p_{s,j}^{(n)}$.

B. Asymptotic analysis of error probability

We now consider how the average probability of error decays with n. First we recall that each P_k is the transition matrix of the Markov chain of successive measurement outcomes. We are interested in the equilibrium state, or steady state, of this Markov chain. It is a standard fact in Markov chain theory that the rate of convergence is determined by the second largest eigenvalue of the transition matrix. To utilize this fact, for $0 \leqslant k \leqslant N-1$, we construct the $(N-1) \times (N-1)$ matrix Q_k by deleting the $(k+1)$th row and the $(k+1)$th column of P_k. We will also construct, for $0 \leqslant k \leqslant N-1$, column vectors $v_k^{(n)}$ by deleting the $(k+1)$th entry in $u_k^{(n)}$. The vector $v^{(0)}$ is similarly constructed from $u^{(0)}$. This construction enables us to write the expression for $p_e^{(n)}$ in a more useful way.

Proposition 3. After n iterations of the state discrimination circuit, the average probability of error is given by

$$p_e^{(n)} = \sum_{k=0}^{N-1} \sum_{j=1}^{N-1} p_k e_j^T Q_k^n v^{(0)}. \quad (28)$$

Proof. We indicate here that this is a rewriting of Proposition 1, in which we make use of the definitions introduced directly above. For completeness, we provide details below.

First, we have as a direct consequence of (25) that

$$v_k^{(n)} = Q_k^n v^{(0)}. \quad (29)$$

Hence we have

$$p_e^{(n)} = \sum_{k=0}^{N-1} \sum_{\substack{j=0 \\ j \neq k}}^{N-1} p_k \text{Tr}\{|j\rangle\langle j|\sigma_{n,k}\} \quad (30)$$

$$= \sum_{k=0}^{N-1} \sum_{j=1}^{N-1} p_k e_j^T v_k^{(n)} \quad (31)$$

$$= \sum_{k=0}^{N-1} \sum_{j=1}^{N-1} p_k e_j^T Q_k^n v^{(0)}. \quad (32)$$

In the above, the first equality is due to the definition of $p_e^{(n)}$. The second equality is due to the definition of $v_k^{(n)}$. The final equality is due to the fact that $v_k^{(n)} = Q_k^n v^{(0)}$, which is a direct consequence of (17). ∎

Before we state our next result that quantifies the rate of decay of $p_e^{(n)}$, we establish some notation that we will use to prove it. First, for each $0 \leqslant k \leqslant N-1$, let s_k denote the number of distinct eigenvalues of Q_k and let $\lambda_{1,k}, \ldots, \lambda_{s_k,k}$ be the distinct eigenvalues of Q_k. That is, each Q_k has eigenvalues $\lambda_{1,k}, \lambda_{2,k}, \ldots, \lambda_{s_k,k}$, and let $m_{i,k}$ denote the algebraic multiplicity of eigenvalue $\lambda_{i,k}$. Also, let

$$\tau := \max_{i,k} |\lambda_{i,k}|. \quad (33)$$

That is, τ is the largest absolute value of the eigenvalues of all the Q_k matrices taken together for $0 \leqslant k \leqslant N-1$.

Proposition 4. The asymptotic error exponent of the state discrimination scheme outlined above is not smaller than the negative logarithm of τ, defined in (33). That is, the following inequality holds:

$$\xi := \lim_{n \to \infty} \frac{-\ln p_e^{(n)}}{n} \geqslant -\ln \tau. \quad (34)$$

Proof. We begin by denoting the $m_{i,k} \times m_{i,k}$ Jordan block of Q_k with eigenvalue $\lambda_{i,k}$ by $H_m(\lambda_{i,k})$, for each $1 \leqslant i \leqslant s_k$. For each $0 \leqslant k \leqslant N-1$, define the $(N-1) \times (N-1)$ block-diagonal matrix

$$J_k = \begin{pmatrix} H_{m_{1,k}}(\lambda_{1,k}) & 0 & \cdots & 0 \\ 0 & H_{m_{2,k}}(\lambda_{2,k}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & H_{m_{s_k,k}}(\lambda_{s_k,k}) \end{pmatrix} \quad (35)$$

such that $Q_k = S_k J_k S_k^{-1}$, where S_k is an invertible $(N-1) \times (N-1)$ matrix. Such a matrix exists since J_k is the Jordan form of Q_k. Recall the expression for $p_e^{(n)}$ in (30). We are interested in the matrix Q_k^n. Using the fact that $Q_k = S_k J_k S_k^{-1}$, we have that $Q_k^n = S_k J_k^n S_k^{-1}$. We then write [31], for $0 \leqslant k \leqslant N-1$ and $1 \leqslant i \leqslant s_k$,

$$H_{m_{i,k}}^n(\lambda_{i,k}) = \begin{pmatrix} \lambda_{i,k}^n & \binom{n}{1}\lambda_{i,k}^{n-1} & \binom{n}{2}\lambda_{i,k}^{n-2} & \cdots & \binom{n}{m_{i,k}-1}\lambda_{i,k}^{n-m_{i,k}+1} \\ & \lambda_{i,k}^n & \binom{n}{1}\lambda_{i,k}^{n-1} & \cdots & \binom{n}{m_{i,k}-2}\lambda_{i,k}^{n-m_{i,k}+2} \\ & & \ddots & \ddots & \vdots \\ & 0 & & \ddots & \ddots & \vdots \\ & & & & \lambda_{i,k}^n & \binom{n}{1}\lambda_{i,k}^{n-1} \\ & & & & & \lambda_{i,k}^n \end{pmatrix}, \quad (36)$$

where we have used the fact that, for $j > n$, $\binom{n}{j} = 0$.

Since for each integer n we have

$$J_k^n = \begin{pmatrix} H_{m_{1,k}}^n(\lambda_{1,k}) & 0 & \cdots & 0 \\ 0 & H_{m_{2,k}}^n(\lambda_{2,k}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & H_{m_{s_k,k}}^n(\lambda_{s_k,k}) \end{pmatrix}, \tag{37}$$

it follows from (14) and (36) that for sufficiently large n we have

$$p_e^{(n)} = \sum_{k=0}^{N-1} \sum_{\substack{i=1 \\ \lambda_{i,k} \neq 1}}^{s_k} \sum_{j=0}^{m_{i,k}-1} a_{i,j,k} \binom{n}{j} \lambda_{i,k}^{n-j} \tag{38}$$

for some set $\{a_{i,j,k}\}_{i,j,k}$ of constants, where $a_{i,j,k} \in \mathbb{C}$ for all i, j, k. To understand the above, we recall that

$$p_e^{(n)} = \sum_{k=0}^{N-1} \sum_{j=1}^{N-1} p_k e_j^T Q_k^n v^{(0)} \tag{39}$$

$$= \sum_{k=0}^{N-1} \sum_{j=1}^{N-1} p_k e_j^T S_k J_k^n S_k^{-1} v^{(0)}. \tag{40}$$

That is, $p_e^{(n)}$ is a linear combination of the elements of the matrices J_k^n for $0 \leqslant k \leqslant N-1$ and hence also a linear combination of powers of the eigenvalues $\lambda_{i,k}$. Further, by inspecting the elements of (36), the linear combination takes the form in (38).

We then have that

$$\xi := -\lim_{n \to \infty} \frac{\ln p_e^{(n)}}{n} \tag{41}$$

$$= -\lim_{n \to \infty} \frac{1}{n} \ln \left\{ \tau^n \left[\sum_{i,j,k} \frac{a_{i,j,k}}{\lambda_{i,k}^j} \binom{n}{j} \left(\frac{\lambda_{i,k}}{\tau} \right)^n \right] \right\} \tag{42}$$

$$= -\ln \tau - \lim_{n \to \infty} \frac{1}{n} \ln \left[\sum_{i,j,k} \frac{a_{i,j,k}}{\lambda_{i,k}^j} \binom{n}{j} \left(\frac{\lambda_{i,k}}{\tau} \right)^n \right] \tag{43}$$

$$\geqslant -\ln \tau. \tag{44}$$

In the above, the first equality is due to (38). The second equality is due to algebraic manipulation. To establish the final inequality, consider the following chain of reasoning:

$$\sum_{i,k} \sum_j \frac{a_{i,j,k}}{\lambda_{i,k}^j} \binom{n}{j} \left(\frac{\lambda_{i,k}}{\tau} \right)^n$$

$$= \left| \sum_{i,k} \sum_j \frac{a_{i,j,k}}{\lambda_{i,k}^j} \binom{n}{j} \left(\frac{\lambda_{i,k}}{\tau} \right)^n \right| \tag{45}$$

$$\leqslant \sum_{i,k} \sum_j \frac{|a_{i,j,k}|}{|\lambda_{i,k}|^j} \binom{n}{j} \left(\frac{|\lambda_{i,k}|}{\tau} \right)^n \tag{46}$$

$$= \sum_{(i,k) \in \mathcal{L}} \sum_j \frac{|a_{i,j,k}|}{|\lambda_{i,k}|^j} \binom{n}{j} \left(\frac{|\lambda_{i,k}|}{\tau} \right)^n$$

$$+ \sum_{(i,k) \notin \mathcal{L}} \sum_j \frac{|a_{i,j,k}|}{|\lambda_{i,k}|^j} \binom{n}{j} \left(\frac{|\lambda_{i,k}|}{\tau} \right)^n \tag{47}$$

$$= \sum_{(i,k) \in \mathcal{L}} \sum_j \frac{|a_{i,j,k}|}{|\lambda_{i,k}|^j} \binom{n}{j} + e^{-\Omega(n)} \tag{48}$$

$$= O(\text{poly}(n)) + e^{-\Omega(n)}. \tag{49}$$

Here $\mathcal{L}$ denotes the set of pairs (i, k) for which $|\lambda_{i,k}| = \tau$ [thus, if $(i, k) \notin \mathcal{L}$, then $|\lambda_{i,k}| < \tau$]. In the above, we have employed the triangle inequality and the fact that $\binom{n}{j} \leqslant \frac{n^j}{j!}$. By applying the negative logarithm and using its antimonotonicity, normalizing, and taking the limit $n \to \infty$, it follows that

$$- \lim_{n \to \infty} \frac{1}{n} \ln \left[\sum_{i,k} \sum_j \frac{a_{i,j,k}}{\lambda_{i,k}^j} \binom{n}{j} \left(\frac{\lambda_{i,k}}{\tau} \right)^n \right]$$

$$\geqslant - \lim_{n \to \infty} \frac{1}{n} \ln[O(\text{poly}(n)) + e^{-\Omega(n)}]$$

$$= 0. \tag{50}$$

This establishes the desired inequality

$$\xi \geqslant -\ln \tau \tag{51}$$

and concludes the proof. ∎

In Appendix A we state and prove simple lower bounds on ξ in terms of the unitaries $\{U_i\}_i$ and the states $\{|\psi_i\rangle\}_i$.

V. EXAMPLES

We now discuss how to optimize the performance of our state discrimination circuit in specific cases. To optimize the performance of our state discrimination circuit, it is necessary to find a set of unitaries that minimizes the probability of error. We may find expressions for these unitaries in simple cases, but this becomes difficult in the general case. An alternative route is to maximize the error exponent $\xi := -\lim_{n \to \infty} \frac{\ln p_e^{(n)}}{n}$. In the following, we discuss explicit state discrimination schemes for sets of qubit states.

A. Two-qubit states

Our first example is the simplest possible, where we consider that we are to discriminate between two pure qubit states $|\psi_0\rangle$ and $|\psi_1\rangle$.

To perform the state discrimination, we require unitaries U_0 and U_1 such that $U_0|\psi_0\rangle = |0\rangle$ and $U_1|\psi_1\rangle = |1\rangle$. We may write the two unitaries U_0 and U_1 in the form

$$U_0 = e^{i\phi_0}|0\rangle\langle\psi_0| + e^{i\phi_1}|1\rangle\langle\psi_0^\perp|,$$

$$U_1 = e^{i\phi_2}|1\rangle\langle\psi_1| + e^{i\phi_3}|0\rangle\langle\psi_1^\perp|, \tag{52}$$

where $|\psi_0^\perp\rangle$ and $|\psi_1^\perp\rangle$ are pure states orthogonal to $|\psi_0\rangle$ and $|\psi_1\rangle$, respectively. We then have

$$Q_0 = |\langle 1|U_1|\psi_0\rangle|^2 = |\langle\psi_0|\psi_1\rangle|^2,$$
$$Q_1 = |\langle 0|U_0|\psi_1\rangle|^2 = |\langle\psi_0|\psi_1\rangle|^2. \tag{53}$$

From Proposition 4 and the Chernoff bound [28,29], we have $\xi = -\ln |\langle\psi_0|\psi_1\rangle|^2$. We see that the average error probability $p_e^{(n)}$ is independent of the choice of the unitaries, as expected from Proposition 1. Further, the error probability $p_e^{(n)}$ scales according to the Chernoff bound in (3).

B. Arbitrary set of qubit states

We now show how to construct a set of unitaries for discriminating an arbitrary set of more than two pure qubit states, and we find that our construction ensures that the probability of error $p_e^{(n)}$ scales according to the multiple Chernoff bound, generalizing what we showed above for two qubit states. This means that the probability of error decays at the optimal rate, so our state discrimination scheme will perform better than

or as well as any other scheme designed to discriminate qubit states in the asymptotic case.

Let $\{|\psi_i\rangle\}_{i=0}^{N-1}$ be a set of $N > 2$ qubit states, each of which is in a two-dimensional Hilbert space $\mathcal{H}$. Let $\{|i\rangle\}_{i=0}^{N-1}$ be a basis for an N-dimensional Hilbert space $\mathcal{H}'$. We will use the following isometries to perform our state discrimination protocol:

$$V_i = |i\rangle\langle\psi_i| + |i \oplus 1\rangle\langle\psi_i^\perp|. \tag{54}$$

Here $|\psi_i^\perp\rangle$ is a pure state orthogonal to $|\psi_i\rangle$. Note that each V_i is an isometry mapping $\mathcal{H}$ to $\mathcal{H}'$ and satisfies $V_i|\psi_i\rangle = |i\rangle$. Given each isometry V_i, let U_i be its unitary extension, satisfying

$$V_i|\psi\rangle = U_i|\tilde{\psi}\rangle \tag{55}$$

for every $|\psi\rangle \in \mathcal{H}$ and where $|\tilde{\psi}\rangle \in \mathcal{H}'$ denotes an embedding of $|\psi\rangle$ in $\mathcal{H}'$. Note that if $N = 2^\ell$ for some integer ℓ, then we can set $|\tilde{\psi}\rangle = |\psi\rangle |0\rangle^{\otimes(\ell-1)}$.

We then get that for $0 \leqslant k \leqslant N - 1$, the matrix P_k takes on the form

$$P_k = \begin{pmatrix} |\langle\psi_0|\psi_k\rangle|^2 & 0 & \cdots & 1 - |\langle\psi_{N-1}|\psi_k\rangle|^2 \\ 1 - \langle\psi_0|\psi_k\rangle|^2 & |\langle\psi_1|\psi_k\rangle|^2 & & 0 \\ 0 & 1 - |\langle\psi_1|\psi_k\rangle|^2 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & |\langle\psi_{N-1}|\psi_k\rangle|^2 \end{pmatrix}. \tag{56}$$

To construct each matrix Q_k, we delete the $(k+1)$th rows and columns from P_k. Recall that the $(k+1)$th column of each P_k contains all zeros except for a one in the $(k+1)$th row. This means that Q_0 and Q_{N-1} will be lower-triangular matrices. The other Q_k matrices will consist of a diagonal, a subdiagonal containing a zero element, and a non-necessarily-zero element $(Q_k)_{1,(N-1)}$, with all other elements set to zero. For such matrices, the eigenvalues are given by their diagonal entries. This can be seen by writing out the characteristic polynomial of the matrix and taking care to expand the determinant along the row or column containing the zero element of the subdiagonal. Therefore, the largest eigenvalue of each matrix Q_k is equal to $\max_{(i,j)} |\langle\psi_i|\psi_j\rangle|^2$. By Proposition 4 and the multiple Chernoff bound [28,29], it follows that $\xi = -\ln\max_{i\neq j}\{|\langle\psi_i|\psi_j\rangle|^2\}$.

Therefore, by appending sufficiently many ancillary qubits, our state discrimination circuit can be used to discriminate an arbitrary set of qubit states with the optimal scaling of the probability of error given by the multiple Chernoff bound in (3). This result only concerns the scaling of the probability of error in the asymptotic case. Investigating how close to optimal the probability of error is when only a finite number of copies of the input state are available may prove to be a fruitful direction for future work.

Note that one may not extend in a straightforward way the above procedure for constructing a set of unitaries $\{U_i\}$ that produce optimal scaling of the probability of error to states in a Hilbert space of dimension greater than 2. Observe that the above procedure hinges on the matrices Q_k having eigenvalues

given by their diagonal. By the definition of the Q_k matrix, specifying a set of isometries in a way similar to that of (54) for a higher-dimensional system would require that other elements besides $(Q_k)_{1,N-1}$ and the elements on the diagonal and subdiagonal of Q_k be nonzero. Such matrices do not in general have eigenvalues equal to their diagonal elements.

When simulating a DCTC with unitary interaction given by $V_{SC} = (\sum_{k=0}^{N-1} |k\rangle\langle k| \otimes U_k) \circ \text{SWAP}$, it follows from our choice of the set $\{U_k\}_k$ that $|j\rangle\langle j|$ is the unique solution for σ_C in (5) whenever $\rho_S = |\psi_j\rangle\langle\psi_j|$. In Appendix B we provide a proof inspired by the argument given in [10] that this is indeed true for the example considered above, consisting of discriminating an arbitrary set of qubit states.

C. BB84 states

Here we study how our state discrimination circuit can be used to discriminate a specific set of nonorthogonal states, i.e., the BB84 states $|0\rangle$, $|1\rangle$, $|+\rangle$, and $|-\rangle$. In general, identifying or performing the set of unitaries $\{U_i\}$ described above may be difficult. The authors of [10] identified a set of unitaries that can discriminate the BB84 states, which we restate here. Building off of these authors' work, we show that the probability of error using this construction scales according to the multiple Chernoff bound in (3). This is of significance because these operators provide an example of unitaries that both produce a probability of error that saturates the multiple Chernoff bound and are constructed from well-studied standard quantum logic gates.

As we described in the example earlier, we encode these four states into a four-dimensional Hilbert space. Let $|\psi_0\rangle \equiv |00\rangle$, $|\psi_1\rangle \equiv |10\rangle$, $|\psi_2\rangle \equiv |+0\rangle$, and $|\psi_3\rangle \equiv |-0\rangle$. That is, we obtain the set of states $\{|\psi_i\rangle\}_{i=0}^3$ by appending an ancillary qubit in the $|0\rangle$ state. Let $|0\rangle \equiv |00\rangle$, $|1\rangle \equiv |01\rangle$, $|2\rangle \equiv |10\rangle$, and $|3\rangle \equiv |11\rangle$. Now let

$$
\begin{aligned}
U_0 &= \text{SWAP}, \\
U_1 &= X \otimes X, \\
U_2 &= (X \otimes I) \circ (H \otimes I), \\
U_3 &= (X \otimes H) \circ \text{SWAP}.
\end{aligned}
\tag{57}
$$

Then $U_i|\psi_i\rangle = |i\rangle$ for $i \in \{0, 1, 2, 3\}$. It may be seen that the probability of error $p_e^{(n)}$ scales according to the Chernoff bound by constructing the Q_k matrices for $0 \leqslant k \leqslant 3$ and checking that their largest eigenvalues are each $\max_{i \neq j} |\langle \psi_i|\psi_j\rangle|^2 = \frac{1}{2}$.

Since an arbitrary set of four geometrically uniform qubit states is simply a rotation of the BB84 states on the Bloch sphere, one may saturate the Chernoff bound when using our state discrimination circuit to distinguish any set of four geometrically uniform qubit states using only compositions of the standard qubit gates X, H, SWAP, and rotations.

VI. CONCLUDING REMARKS

In this paper we have proposed a method for discriminating multiple nonorthogonal states, which is inspired by a construction considered in the context of closed timelike curves [10]. Our state discrimination method can be equivalently recast as a local iterative circuit whose simplicity lends itself to experimental implementation. Furthermore, we studied the average probability of error for our scheme and showed that in the general case of discriminating an arbitrary set of pure qubit states, it achieves the multiple Chernoff bound.

We would like to point out three aspects of our work that require further investigation. It has been shown that a two-state local adaptive state discrimination scheme may be optimal for any number of copies [18]. It remains open whether there exists a way to configure some aspect of our circuit differently so that it is possible for our scheme to be optimal for any number of copies. Also, while sets of unitaries do exist that optimize the performance of our circuit in the asymptotic limit for an arbitrary set of qubit states, it is unknown whether there exists a set of optimal unitaries for any set of qubit states. Furthermore, it is worth investigating this aspect of our work to see whether there exists a set of *product* unitaries to optimize, or even make the performance of our circuit suboptimally efficient, in the asymptotic limit. This would be beneficial because product operators are more convenient to implement experimentally.

Finally, any attempt at practical state discrimination will be subject to noisy conditions. Noise may have the ability to enhance or worsen the performance of a state discrimination scheme. A recent result due to [20] has shown that the optimal measurement in the discrimination of two pure qubit states is no longer optimal when these qubit states are subject to perturbations. Studying the behavior of our state discrimination circuit in the presence of noise remains a topic for future work.

APPENDIX A: UPPER AND LOWER BOUNDS ON THE MULTIPLE CHERNOFF EXPONENT ξ

We now state a result that allows us to bound the scaling of the average probability of error in terms of the set $\{U_i\}_i$ of unitaries and the set $\{|\psi_i\rangle\}_i$ of states.

Proposition 5. The inequalities

$$
\xi \geqslant -\ln\left(1 - \min_{j,k:j\neq k} |\langle k|U_j|\psi_k\rangle|^2\right),
\tag{A1}
$$

$$
\xi \geqslant -\ln\left(\max_{j,k:j\neq k} \sum_{i\neq k} |\langle j|U_i|\psi_k\rangle|^2\right)
\tag{A2}
$$

hold, with ξ defined in (41).

Proof. Let $0 \leqslant k \leqslant N - 1$. The maximum column sum of Q_k is

$$
\max_{j\neq k} \sum_{i\neq k} |\langle i|U_j|\psi_k\rangle|^2 = 1 - \min_{j\neq k} |\langle k|U_j|\psi_k\rangle|^2.
\tag{A3}
$$

It follows from the Gershgorin circle theorem [32] that the largest eigenvalue of Q_k is bounded from above by the maximum column sum of Q_k. We recall that $\tau \coloneqq \max_{i,k} |\lambda_{i,k}|$, i.e., τ is the largest of the absolute values of the eigenvalues of all of the Q_k matrices. It follows that

$$
\tau \leqslant 1 - \min_{j,k:j\neq k} |\langle k|U_j|\psi_k\rangle|^2.
\tag{A4}
$$

Then using Proposition 4, we have

$$
\xi \geqslant -\ln\tau \geqslant -\ln\left(1 - \max_{j,k:j\neq k} |\langle k|U_j|\psi_k\rangle|^2\right).
\tag{A5}
$$

It also follows from the Gershgorin circle theorem that for $0 \leqslant k \leqslant N - 1$, the largest eigenvalue of Q_k is bounded from above by the maximum row sum of Q_k. Hence, we have

$$
\tau \leqslant \max_{\substack{j,k \\ j\neq k}} \sum_{i\neq k} |\langle j|U_i|\psi_k\rangle|^2,
\tag{A6}
$$

so that

$$
\xi \geqslant -\ln\tau \geqslant -\ln\left(\max_{j,k:j\neq k} \sum_{i\neq k} |\langle j|U_i|\psi_k\rangle|^2\right).
\tag{A7}
$$

This concludes the proof. ∎

APPENDIX B: UNIQUENESS OF THE FIXED POINT OF THE CHANNEL $\mathcal{N}_{V,\rho}$

Here we show the uniqueness of the fixed point of the channel $\mathcal{N}_{V,\rho}$ when discriminating an arbitrary set of qubit states, which we studied in Sec. V B.

Proposition 6. Let $N \geqslant 3$. For each $0 \leqslant i \leqslant N - 1$, let U_i be a unitary extension of the isometry V_i defined in (54). Let $V_{SC} = (\sum_k |k\rangle\langle k| \otimes U_k) \circ \text{SWAP}$. Then for $0 \leqslant a \leqslant N - 1$, $|a\rangle\langle a|$ is the unique solution for σ_C in (5) whenever $\rho_S = |\psi_a\rangle\langle\psi_a|$.

Proof. We argued that $|a\rangle\langle a|$ is a solution for σ_C whenever $\rho_S = |\psi_a\rangle\langle\psi_a|$ in Sec. III. It remains to show uniqueness. Suppose $\rho_S = |\psi_a\rangle\langle\psi_a|$ and σ is a solution for σ_C in (5). Then we have

$$
\begin{aligned}
\sigma &= \text{Tr}_S\{V_{SC}(|\psi_a\rangle\langle\psi_a| \otimes \sigma)V_{SC}^\dagger\} \\
&= \text{Tr}_S\left\{\left(\sum_k |k\rangle\langle k| \otimes U_k\right)(\sigma \otimes |\psi_a\rangle\langle\psi_a|) \right. \\
&\quad \left. \times \left(\sum_l |l\rangle\langle l| \otimes U_l^\dagger\right)\right\} \\
&= \sum_{k,l} \text{Tr}_S\{|k\rangle\langle k|\sigma|l\rangle\langle l| \otimes U_k|\psi_a\rangle\langle\psi_a|U_l^\dagger\} \\
&= \sum_k \langle k|\sigma|k\rangle U_k|\psi_a\rangle\langle\psi_a|U_k^\dagger. \quad\quad (\text{B1})
\end{aligned}
$$

In the above, the first line is due to the self-consistency condition (5). The following equalities come from explicitly writing out V_{SC} and algebraic manipulation. Hence, the matrix elements of σ are given by

$$
\langle m|\sigma|n\rangle = \sum_k \langle k|\sigma|k\rangle\langle m|U_k|\psi_a\rangle\langle\psi_a|U_k^\dagger|n\rangle. \quad (\text{B2})
$$

We now show that all diagonal elements of σ other than $\langle a|\sigma|a\rangle$ are zero. We will proceed by induction to show that $\langle a \ominus s|\sigma|a \ominus s\rangle = 0$ for all $1 \leqslant s \leqslant N - 1$. (Here $\ominus$ denotes subtraction modulo N.) We first show that $\langle a \ominus 1|\sigma|a \ominus 1\rangle = 0$. It follows from (B2) that

$$
\langle a|\sigma|a\rangle = \langle a|\sigma|a\rangle + \sum_{k \neq a} \langle k|\sigma|k\rangle|\langle a|U_k|\psi_a\rangle|^2, \quad (\text{B3})
$$

which implies that $\sum_{k \neq a} \langle k|\sigma|k\rangle|\langle a|U_k|\psi_a\rangle|^2 = 0$. It follows that

$$
\begin{aligned}
&\langle a \ominus 1|\sigma|a \ominus 1\rangle(1 - |\langle\psi_{a\ominus 1}|\psi_a\rangle|^2) \\
&\quad = \langle a \ominus 1|\sigma|a \ominus 1\rangle|\langle\psi_{a\ominus 1}^\perp|\psi_a\rangle|^2 \\
&\quad = \langle a \ominus 1|\sigma|a \ominus 1\rangle|\langle a|U_{a\ominus 1}|\psi_a\rangle|^2 \\
&\quad = \sum_{k \neq a} \langle k|\sigma|k\rangle|\langle a|U_k|\psi_a\rangle|^2 \\
&\quad = 0. \quad\quad (\text{B4})
\end{aligned}
$$

Since $1 - |\langle\psi_{a\ominus 1}|\psi_a\rangle|^2 \neq 0$, we then have that $\langle a \ominus 1|\sigma|a \ominus 1\rangle = 0$. Now suppose $1 \leqslant s < N - 1$ is such that $\langle a \ominus s|\sigma|a \ominus s\rangle = 0$. Using $(B2)$ and the definition of the unitaries $\{U_i\}$, we have

$$
\begin{aligned}
\langle a \ominus s|\sigma|a \ominus s\rangle &= \sum_k \langle k|\sigma|k\rangle|\langle a \ominus s|U_k|\psi_a\rangle|^2 \\
&= \langle a \ominus s|\sigma|a \ominus s\rangle|\langle a \ominus s|U_{a\ominus s}|\psi_a\rangle|^2 \\
&\quad + \langle a \ominus (s+1)|\sigma|a \ominus (s+1)\rangle \\
&\quad \times |\langle a \ominus s|U_{a\ominus(s+1)}|\psi_a\rangle|^2. \quad (\text{B5})
\end{aligned}
$$

Since $\langle a \ominus s|\sigma|a \ominus s\rangle = 0$, it follows that

$$
\begin{aligned}
0 &= \langle a \ominus (s+1)|\sigma|a \ominus (s+1)\rangle|\langle a \ominus s|U_{a\ominus(s+1)}|\psi_a\rangle|^2 \\
&= \langle a \ominus (s+1)|\sigma|a \ominus (s+1)\rangle|\langle\psi_{a\ominus(s+1)}^\perp|\psi_a\rangle|^2 \\
&= \langle a \ominus (s+1)|\sigma|a \ominus (s+1)\rangle(1 - |\langle\psi_{a\ominus(s+1)}|\psi_a\rangle|^2).
\end{aligned}
$$
$$(\text{B6})$$

Since $1 \leqslant s < N - 1$, we must have $a \ominus (s+1) \neq a$, so $1 - |\langle\psi_{a\ominus(s+1)}|\psi_a\rangle|^2 \neq 0$. It then follows from (B6) that $\langle a \ominus (s+1)|\sigma|a \ominus (s+1)\rangle = 0$. By induction, we have $\langle a \ominus s|\sigma|a \ominus s\rangle = 0$ for $1 \leqslant s \leqslant N - 1$. Hence, all diagonal elements of σ other than $\langle a|\sigma|a\rangle$ are zero.

Consequently, we must have $\langle k|\sigma|k\rangle = \delta_{ka}$ since $\text{Tr}[\sigma] = 1$. Any density operator of this form has off-diagonal elements that are all zero. Therefore, we have $\sigma = |a\rangle\langle a|$. Hence, $|a\rangle\langle a|$ is the unique solution for σ_C in (5). $\blacksquare$

[1] K. Gödel, An example of a new type of cosmological solutions of Einstein's field equations of gravitation, Rev. Mod. Phys. **21**, 447 (1949).

[2] D. Deutsch, Quantum mechanics near closed timelike lines, Phys. Rev. D **44**, 3197 (1991).

[3] G. Svetlichny, Time travel: Deutsch vs. teleportation, Int. J. Theor. Phys. **50**, 3903 (2011).

[4] S. Lloyd, L. Maccone, R. Garcia-Patron, V. Giovannetti, and Y. Shikano, Quantum mechanics of time travel through post-selected teleportation, Phys. Rev. D **84**, 025007 (2011).

[5] J.-M. A. Allen, Treating time travel quantum mechanically, Phys. Rev. A **90**, 042107 (2014).

[6] J. L. Park, The concept of transition in quantum mechanics, Found. Phys. **1**, 23 (1970).

[7] W. K. Wootters and W. H. Zurek, A single quantum cannot be cloned, Nature (London) **299**, 802 (1982).

[8] D. Dieks, Communication by EPR devices, Phys. Lett. A **92**, 271 (1982).

[9] T. A. Brun, M. M. Wilde, and A. Winter, Quantum State Cloning Using Deutschian Closed Timelike Curves, Phys. Rev. Lett. **111**, 190401 (2013).

[10] T. A. Brun, J. Harrington, and M. M. Wilde, Localized Closed Timelike Curves Can Perfectly Distinguish Quantum States, Phys. Rev. Lett. **102**, 210402 (2009).

[11] M. Bartkiewicz, A. Grudka, R. Horodecki, J. Łodyga, and J. Wychowaniec, Closed timelike curves and the second law of thermodynamics, Phys. Rev. A **99**, 022304 (2019).

[12] S. Aaronson and J. Watrous, Closed timelike curves make quantum and classical computing equivalent, Proc. R. Soc. A **465**, 631 (2009).

[13] C. H. Bennett, D. Leung, G. Smith, and J. A. Smolin, Can Closed Timelike Curves or Nonlinear Quantum Mechanics Improve Quantum State Discrimination or Help Solve Hard Problems?, Phys. Rev. Lett. **103**, 170502 (2009).

[14] E. G. Cavalcanti and N. C. Menicucci, Verifiable nonlinear quantum evolution implies failure of density matrices to represent proper mixtures, arXiv:1004.1219.

[15] M. Ringbauer, M. A. Broome, C. R. Myers, A. G. White, and T. C. Ralph, Experimental simulation of closed timelike curves, Nat. Commun. **5**, 4145 (2014).

[16] T. A. Brun and M. M. Wilde, Simulations of closed timelike curves, Found. Phys. **47**, 375 (2017).

[17] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Elsevier, Amsterdam, 1976).

[18] A. Acín, E. Bagan, M. Baig, Ll. Masanes, and R. M. Tapia, Multiple-copy two-state discrimination with individual measurements, Phys. Rev. A **71**, 032338 (2005).

[19] B. L. Higgins, A. C. Doherty, S. D. Bartlett, G. J. Pryde, and H. M. Wiseman, Multiple-copy state discrimination: Thinking globally, acting locally, Phys. Rev. A **83**, 052314 (2011).

[20] K. Flatt, S. M. Barnett, and S. Croke, Multiple-copy state discrimination of noisy qubits, Phys. Rev. A **100**, 032122 (2019).

[21] M. Ban, K. Kurokawa, R. Momose, and O. Hirota, Optimum measurements for discrimination among symmetric quantum states and parameter estimation, Int. J. Theor. Phys. **36**, 1269 (1997).

[22] E. Andersson, S. M. Barnett, C. R. Gilson, and K. Hunter, Minimum-error discrimination between three mirror-symmetric states, Phys. Rev. A **65**, 052308 (2002).

[23] J. Bae and W.-Y. Hwang, Minimum-error discrimination of qubit states: Methods, solutions, and properties, Phys. Rev. A **87**, 012334 (2013).

[24] D. Ha and Y. Kwon, Complete analysis for three-qubit mixed-state discrimination, Phys. Rev. A **87**, 062302 (2013).

[25] S. Slussarenko, M. M. Weston, J.-G. Li, N. Campbell, H. M. Wiseman, and G. J. Pryde, Quantum State Discrimination Using the Minimum Average Number of Copies, Phys. Rev. Lett. **118**, 030502 (2017).

[26] G. Weir, S. M. Barnett, and S. Croke, Optimal discrimination of single-qubit mixed states, Phys. Rev. A **96**, 022312 (2017).

[27] R. Blume-Kohout, S. Croke, and M. Zwolak, Quantum data gathering, Sci. Rep. **3**, 1800 (2013).

[28] M. Nussbaum and A. Szkoła, in *Theory of Quantum Computation, Communication, and Cryptography*, edited by W. van Dam, V. M. Kendon, and S. Severini, Lecture Notes in Computer Science Vol. 6519 (Springer, Berlin, 2011), pp. 1–8.

[29] K. Li, Discriminating quantum states: The multiple Chernoff distance, Ann. Stat. **44**, 1661 (2016).

[30] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

[31] C. D. Meyer, in *Matrix Analysis and Applied Linear Algebra* (Society for Industrial and Applied Mathematics, Philadelphia, 2000), p. 618.

[32] R. Bhatia, *Matrix Analysis* (Springer, Berlin, 1997).