When is the Chernoff Exponent for Quantum Operations finite?

Nengkun Yu
Centre for Quantum Software and Information,
Faculty of Engineering and Information Technology,
University of Technology Sydney, NSW 2007, Australia
Email: nengkunyu@gmail.com
Li Zhou
Max Planck Institute for Security and Privacy, Bochum, Germany
zhou31416@gmail.com

***Abstract*—We consider the problem of testing two hypotheses of quantum operations in a setting of many uses where an arbitrary prior probability distribution is given. The Chernoff exponent for quantum operations is investigated to track the minimal average error probability of discriminating two quantum operations asymptotically. We answer the question, "When is the Chernoff exponent for quantum operations finite?" We show that either two quantum operations can be perfectly distinguished with finite uses, or the minimal discrimination error decays exponentially with respect to the number of uses asymptotically. That is, the Chernoff exponent is finite if and only if the quantum operations can not be perfectly distinguished with finite uses. This rules out the possibility of super-exponential decay of error probability. Upper bounds of the Chernoff exponent for quantum operations are provided.**

I. Introduction

A fundamental problem in quantum information theory is to test a device that may be prepared for implementing one of many quantum operations. The testing treated in the framework of quantum mechanics is performed by inputting a quantum state and performing a quantum measurement. The general noncommutative feature and the complex structure of quantum operations make quantum statistics a much richer field than its classical counterpart.

In the degenerate case where the outputs of the quantum operations are fixed, the freedom of choosing input states becomes useless. The asymptotic behavior of the average error, in discriminating a set of quantum states $\{\rho_1^{\otimes n}, \ldots, \rho_r^{\otimes n}\}$ with prior probability distribution $\{\Pi_1, \ldots, \Pi_r\}$ is of great interest. In [26], Parthasarathy showed that the average error decays exponentially, asymptotically. Significant efforts have been made to identify the Chernoff exponent as the optimal error exponent. In two breakthrough papers, [3] and [25], the closed-form of the optimal error exponent was obtained, which can be regarded as the quantum generalization of the Chernoff bound in classical hypothesis testing [6]. Li proved that multiple Chernoff exponent equals the minimal mutual Chernoff exponent [22].

It is highly desirable to generalize the results of quantum states to quantum operations. Given that considerable experimental effort has been devoted to the field of quantum mechanics to prepare quantum systems and measure quantum states, it is of fundamental importance to develop a theory that can discriminate the different quantum operations. We note that for classical channel discrimination, the optimal exponential error rate problem has been well understood, where it is proven that adaptive choice does not improve the exponential error rate in these settings [13].

Where quantum operations are only allowed to be used once, the problem has been extensively studied, with fruitful results. By employing Holevo-Helstrom's celebrated theorem on the one-copy quantum state discrimination [14], [17], a completely bounded trace norm, known as the diamond norm, was introduced to characterize the difference between quantum channels by Kitaev [19]. This norm becomes a fundamental tool in almost all aspects of quantum information science [2], [31], [32], [27] since it is the most physically meaningful notion of distance between quantum operations.

The problem becomes much more complicated when quantum operations are used multiple times [15], [7], [28]. Much effort has been devoted to characterizing the conditions of perfect distinguishability, in the sense that two quantum operations can be distinguished without error by a finite number of uses [1], [9], [37], [21], [20].

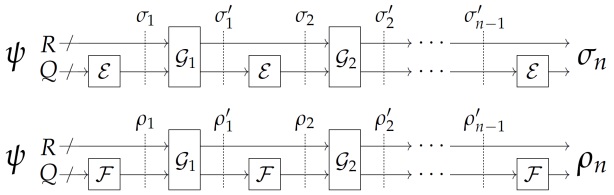

Fig. 1. Model of discriminating two quantum operations $\mathcal{E}$ and $\mathcal{F}$ on system Q with n uses by employing an arbitrary ancilla system R. $\mathcal{G}_i$ are arbitrary quantum operations applied on the joint system RQ and between two uses of the device. One can show that any arbitrary adaptive strategy can be translated into this model, by following the fact that any quantum measurement can be fully implemented by a quantum operation with outcomes stored as qubits.

Unlike classical channel discrimination, unitary operations exist that cannot be distinguished without error for single-use, while multiple uses can help achieve perfect discrimination. A complete solution to this problem is obtained in [10] with a feasible, necessary, and sufficient condition.

In this paper, we investigate the concept of Chernoff exponent for quantum operations to characterize the asymptotic behavior of the average error probability of distinguishing given quantum operations under any prior probability distribution. Suppose we have a quantum device that is secretly chosen from $\{\mathcal{E}, \mathcal{F}\}$, and a known set of two quantum operations according to a prior probability distribution $\{\Pi_0, \Pi_1\}$. Our goal is to identify whether the device is $\mathcal{E}$ or $\mathcal{F}$ by using this device many times. We explore the Chernoff exponent for two quantum operations $\mathcal{E}$ and $\mathcal{F}$ to track the optimal error probability by using the following definition:

$$\xi_{\mathcal{E},\mathcal{F}} = -\overline{\lim_{n\to\infty}} \frac{\log P_{err,min,n}}{n} \tag{1}$$

where $P_{err,min,n}$ denotes the infimum discrimination error over all possible output states ρ_n and σ_n as illustrated in Figure 1 where the quantum device is used n time.

Notice that all possible strategies can be described by, or translate to, the model showed in Figure 1 with a sufficiently large ancilla system. This model is the most general scheme, and quantum operations $\mathcal{G}_i$ are freely chosen. For instance, parallel uses of devices can always be simulated by sequential uses and employing swap operators.

We show that the Chernoff exponent for quantum operations is finite if and only if they cannot be distinguished perfectly with finite uses. More precisely, we show that the average error probability decays at most, according to exponential function for quantum operations, if the quantum operations can not be perfectly distinguishable. This indicates that the error probability can never decay super-exponentially, such as $\exp(-\alpha n^2)$. Computable upper bounds on the Chernoff exponent for quantum operations are provided. Finally, we generalize our results to deal with multiple quantum operations.

II. Notations and Preliminaries

We use the symbols $\mathcal{H}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}$ to denote finite-dimensional Hilbert spaces over complex numbers and $\mathrm{L}(\mathcal{H})$ to denote the set of linear operators mapping from $\mathcal{H}$ into itself. For Hermitian matrices A, B, we use $\langle A, B\rangle = \mathrm{Tr}(A^\dagger B) = \mathrm{Tr}(AB)$ to denote their inner product. Let $\mathrm{Pos}(\mathcal{H}) \subset \mathrm{L}(\mathcal{H})$ be the set of positive (semidefinite) matrices, and $\mathcal{D}(\mathcal{H}) \subset \mathrm{Pos}(\mathcal{H})$ is the set of positive matrices with trace one. A pure quantum state of $\mathcal{H}$ is just a normalized vector $|\psi\rangle \in \mathcal{H}$, while a general quantum state is characterized by a density operator $\rho \in \mathcal{D}(\mathcal{H})$. For simplicity, we use ψ to represent the density operator of a pure state $|\psi\rangle$ which is just the projector $\psi = |\psi\rangle\langle\psi|$. A density operator ρ can always be decomposed into a convex combination of pure states:

$$\rho = \sum_{k=1}^{n} p_k |\psi_k\rangle\langle\psi_k|,$$

where the coefficients p_k are strictly positive numbers and add up to one. The support of ρ is defined as $\mathrm{supp}(\rho) = \mathrm{span}\{|\psi_k\rangle : 1 \le k \le n\}$. We say two pure states $|\psi\rangle$ and $|\phi\rangle$ are orthogonal if and only if their inner product $\langle\psi, \phi\rangle$ is equal to zero, and the orthogonality of two density operators ρ and σ is defined by the orthogonality of their supports, namely, ρ and σ are orthogonal if and only if $\mathrm{supp}(\rho) \perp \mathrm{supp}(\sigma)$. Two density operators ρ and σ are said to be disjoint if $\mathrm{supp}(\rho) \cap \mathrm{supp}(\sigma) = \{0\}$ and joint if the intersection of their support contains some non-zero vectors.

There are two commonly used measures to characterize the difference between the quantum states: trace distance and fidelity. The trace distance D between two density operators ρ and σ is defined as

$$D(\rho, \sigma) \equiv \frac{1}{2}\mathrm{Tr}|\rho - \sigma|$$

where we define $|A| \equiv \sqrt{A^\dagger A}$ to be the positive square root of $A^\dagger A$.

The fidelity of states ρ and σ is defined to be

$$F(\rho, \sigma) \equiv \mathrm{Tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}.$$

For pure states $|\psi\rangle$ and $|\phi\rangle$, $F(\psi, \phi) = |\langle\psi|\phi\rangle|$.

The strong concavity property for the fidelity is quite useful, which can be formalized as

Fact 1 ([24]). *For quantum states ρ_i, σ_i and probability distributions $(p_0, p_1, \cdots, p_n)$ and $(q_0, q_1, \cdots, q_n)$*

$$F\left(\sum_i p_i\rho_i, \sum_i q_i\sigma_i\right) \geq \sum_{i=0}^{n} \sqrt{p_i q_i} F(\rho_i, \sigma_i).$$

If $\rho_i = \psi_i$ and $\sigma_i = \phi_i$ are all pure states, we obtain

$$\begin{aligned} F\left(\sum_i p_i\psi_i, \sum_i q_i\phi_i\right) &\geq \sum_{i=0}^{n} \sqrt{p_i q_i} F(\psi_i, \phi_i) \\ &= \sum_{i=0}^{n} |\langle \sqrt{p_i}\psi_i | \sqrt{q_i}\phi_i \rangle|. \end{aligned}$$

Definition 1. *We say that a pure state $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ is a purification of some state ρ if $\mathrm{tr}_A(\psi) = \rho$.*

Fact 2 (Uhlmann's theorem, [29]). *Given quantum states ρ, σ, and a purification $|\psi\rangle$ of ρ, it holds that $F(\rho, \sigma) = \max_{|\phi\rangle} |\langle\phi|\psi\rangle|$, where the maximum is ranging over all purifications of σ.*

The following fact connects the trace distance and the fidelity between two states.

Fact 3 (Fuchs-van de Graaf inequalities [11]). *For quantum states ρ and σ, it holds that*

$$1 - F(\rho, \sigma) \leq D(\rho, \sigma) \leq \sqrt{1 - F(\rho, \sigma)^2}.$$

For pure states $|\phi\rangle$ and $|\psi\rangle$, we have

$$D(\phi, \psi) = \sqrt{1 - F(\phi, \psi)^2} = \sqrt{1 - |\langle\phi|\psi\rangle|^2}.$$

The trace distance is a static measure quantifying how close two quantum states are and is closely related to the discrimination of quantum states. Let us consider the two hypotheses, H_0 and H_1. Hypothesis H_0 assumes that a given unknown quantum state is equal to ρ_0, and Hypothesis H_1 assumes that a given unknown quantum state is equal to ρ_1. We assume that the prior probability distribution of ρ_0 and ρ_1 are Π_0 and Π_1, respectively, which add up to one.

A physical strategy to discriminate between these two hypotheses is to perform a positive-operator valued measure (POVM) on the quantum state with two outcomes, 0 and 1. Such a POVM has two elements $\{E_0, E_1\}$ satisfying $E_0, E_1 \in \mathrm{Pos}(\mathcal{H})$ and $E_0 + E_1 = I$, where I is the identity matrix of $\mathcal{H}$. The aim of quantum state discrimination is to find the elements E_0 and E_1 that minimize the total error P_{err}, which is

$$P_{err} = \Pi_0 \mathrm{Tr}[E_1\rho_0] + \Pi_1 \mathrm{Tr}[E_0\rho_1].$$

This optimal error has been identified by Helstrom as expressed in the following equation

$$P_{err,min} = \frac{1}{2}\left(1 - \mathrm{Tr}|\Pi_1\rho_1 - \Pi_0\rho_0|\right).$$

A quantum operation $\mathcal{E}$ from $\mathrm{L}(\mathcal{H})$ to $\mathrm{L}(\mathcal{Z})$ is a completely positive and trace-preserving map used to describe the evolution of an open quantum system. A quantum operation $\mathcal{E}$ can always be represented using the Kraus representation as

$$\mathcal{E}(\rho) = \sum_{i=1}^{k} E_i \rho E_i^\dagger,$$

where $\{E_i\}_{i=1,\cdots,k}$ are the Kraus operators of $\mathcal{E}$ satisfying $\sum_{i=1}^{k} E_i^\dagger E_i = I$, the identity of $\mathcal{H}$.

The following fact states that the fidelity between two states is non-decreasing under quantum operations.

Fact 4 ([24]). *For states ρ, σ, and quantum operation $\mathcal{E}(\cdot)$, it holds that*

$$F(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \geq F(\rho, \sigma).$$

Quantum operations $\mathcal{E}$ and $\mathcal{F}$ are said to be perfectly distinguishable with finite uses if there exists a strategy illustrated as Figure 1 such that σ_n and ρ_n are orthogonal.

Two conditions introduced by [10] characterize the perfect distinguishability between quantum operations.

Definition 2. *Two quantum operations, $\mathcal{E}$ and $\mathcal{F}$, acting on the same principal system, denoted by Q, are said to be disjoint if there is an auxiliary system R, and a pure state $|\psi^{RQ}\rangle$, such that $(\mathcal{I}^R \otimes \mathcal{E}^Q)(\psi^{RQ})$ and $(\mathcal{I}^R \otimes \mathcal{F}^Q)(\psi^{RQ})$ are disjoint, where $\mathcal{I}^R$ is the identity operation on R, and the superscripts only identify which systems the operations acted on. Otherwise, they are called joint.*

Intuitively, this disjointness guarantees that the outputs do not have a common part with the carefully chosen input. This disjointness is necessary to achieve perfect distinguishability. Otherwise, according to an inductive argument, there is always a non-zero common part between the outputs for an arbitrary strategy with finite uses.

Another relationship is to ensure that non-orthogonal states, ρ and σ, exist such that $(\mathcal{I}^R \otimes \mathcal{E}^Q)(\rho^{RQ})$ and $(\mathcal{I}^R \otimes \mathcal{F}^Q)(\sigma^{RQ})$ become orthogonal, then one can distinguish $(\mathcal{I}^R \otimes \mathcal{E}^Q)(\rho^{RQ})$ and $(\mathcal{I}^R \otimes \mathcal{F}^Q)(\sigma^{RQ})$ without error. This is the final step of any strategy to achieve perfect distinguishability between $\mathcal{E}$ and $\mathcal{F}$.

Interestingly, these two conditions are not only necessary but also sufficient for the perfect distinguishability between quantum operations [10] .

Proposition 1. *Two quantum operations $\mathcal{E}$ and $\mathcal{F}$ are perfectly distinguishable if and only if: 1). They are disjoint; 2). They can map some non-orthogonal states into orthogonal states.*

We remark here that ancillary systems are also allowed to achieve perfect discrimination.

In the following, we give an analytical characterization of the negation of the second condition, i.e., that $\mathcal{E}$ and $\mathcal{F}$ cannot map some non-orthogonal states into orthogonal states even with the help of ancillary system.

Remark 1. *For $\mathcal{E}(\cdot)=\sum_i E_i\cdot E_i^\dagger$ and $\mathcal{F}(\cdot)=\sum_j F_j\cdot F_j^\dagger$, Condition 2) of Proposition 1 is equivalent to $I\notin \mathrm{span}\{(E_i{}^\dagger F_j); 1\le i,j\le m\}$.*

We want to emphasize this proof of the characterization is precisely the same as given in [10]. We provide the following argument for the readers' convenience.

Without loss of generality, we assume that $\mathcal{E}$ and $\mathcal{F}$ have the same number of Kraus operators by adding zero Kraus operators, if necessary. That is, $\mathcal{E}(\cdot)=\sum_{i=1}^m E_i\cdot E_i^\dagger$ and $\mathcal{F}(\cdot)=\sum_{j=1}^m F_j\cdot F_j^\dagger$, cannot make non-orthogonal states orthogonal. That is, if ρ^{RQ} and σ^{RQ} are not orthogonal, then $(\mathcal{I}^R\otimes\mathcal{E}^Q)(\rho^{RQ})$ and $(\mathcal{I}^R\otimes\mathcal{F}^Q)(\sigma^{RQ})$ are not orthogonal. Equivalently, if pure states $\rho^{RQ}=|\psi^{RQ}\rangle\langle\psi^{RQ}|$ and $\sigma^{RQ}=|\phi^{RQ}\rangle\langle\phi^{RQ}|$ are not orthogonal, then $(\mathcal{I}^R\otimes\mathcal{E}^Q)(\rho^{RQ})$ and $(\mathcal{I}^R\otimes\mathcal{F}^Q)(\sigma^{RQ})$ are not orthogonal. In other words,

$$\mathrm{Tr}[(\mathcal{I}^R\otimes\mathcal{E}^Q)(\psi^{RQ})(\mathcal{I}^R\otimes\mathcal{F}^Q)(\phi^{RQ})]=0$$

implies

$$\langle\psi^{RQ}|\phi^{RQ}\rangle=0.$$

That is, if $\forall 1\le i,j\le m$,

$$(I^R\otimes E_i^Q)|\psi^{RQ}\rangle\langle\psi^{RQ}|(I^R\otimes E_i^Q)^\dagger$$

is orthgonal to

$$(I^R\otimes F_j^Q)|\phi^{RQ}\rangle\langle\phi^{RQ}|(I^R\otimes F_j^Q)^\dagger],$$

then

$$\langle\psi^{RQ}|\phi^{RQ}\rangle=0.$$

The above condition is equivalent to for $\left|\psi^{RQ}\right\rangle$ and $\left|\phi^{RQ}\right\rangle$, if

$$\langle\psi^{RQ}|(I^R\otimes E_i^Q)^\dagger(I^R\otimes F_j^Q)\left|\phi^{RQ}\right\rangle=0$$

for all $1\le i,j\le m$, then

$$\langle\psi^{RQ}|\phi^{RQ}\rangle=0.$$

That is, if $\forall 1\le i,j\le m$

$$\langle\psi^{RQ}|(I^R\otimes E_i{}^\dagger F_j)\left|\phi^{RQ}\right\rangle=0,$$

then

$$\langle\psi^{RQ}|\phi^{RQ}\rangle=0.$$

For any $M\in\mathrm{L}(\mathcal{H}_Q)$, one can find $\left|\phi^{RQ}\right\rangle$ and $\left|\psi^{RQ}\right\rangle$ such that $M=\mathrm{Tr}_R|\phi^{RQ}\rangle\langle\psi^{RQ}|$. We know that $\forall 1\le i,j\le m$

$$\mathrm{Tr}(ME_i{}^\dagger F_j)=\langle\psi^{RQ}|(I^R\otimes E_i{}^\dagger F_j)\left|\phi^{RQ}\right\rangle=0,$$

implies

$$\mathrm{Tr}\,M=0.$$

That is satisfied if and only if $I^{RQ}\in\mathrm{span}\{(I^R\otimes E_i{}^\dagger F_j); 1\le i,j\le m\}$, which in turn is equivalent to $I\in\mathrm{span}\{E_i^\dagger F_j; 1\le i,j\le m\}$.

Therefore, $\mathcal{E}$ and $\mathcal{F}$ can map some non-orthogonal states into orthogonal states, Condition 2) of Proposition 1, is equivalent to $I\notin\mathrm{span}\{(E_i{}^\dagger F_j); 1\le i,j\le m\}$.

III. Two useful Lemmas

The following lemma shows that if two quantum operations are joint, then for any input state, the output states always have a common semi-definite positive component whose size is positive and depends only on the operations.

Lemma 1. *If $\mathcal{E}$ and $\mathcal{F}$ are joint, there exists $\eta>0$, depending only on $\mathcal{E}$ and $\mathcal{F}$, such that for any quantum state ρ on a potentially larger Hilbert space RQ, there is a matrix A, such that $0\le A\le(\mathcal{I}^R\otimes\mathcal{E})(\rho^{RQ}),(\mathcal{I}^R\otimes\mathcal{F})(\rho^{RQ})$ and $\mathrm{Tr}(A)\ge\eta$.*

Proof. It is straightforward to verify that we only need to consider ρ to be a pure state. Thus, according to Schmidt decomposition, we can assume that ρ is a quantum state in $\mathcal{H}_{RQ}=\mathcal{H}'\otimes\mathcal{H}$ with the dimension of $\mathcal{H}'$ being equal to the dimension of $\mathcal{H}$, where $\mathcal{H}$ is the Hilbert space of Q.

Our goal is to show

$$\eta>0$$

where η is defined as

$$\begin{aligned}\eta&:=\inf_{\rho\in\mathcal{D}(\mathcal{H}_{RQ})}\ \sup_{\substack{0\le X\le(\mathcal{I}^R\otimes\mathcal{E})(\rho^{RQ}),\\ 0\le X\le(\mathcal{I}^R\otimes\mathcal{F})(\rho^{RQ})}}\mathrm{Tr}\,X\\&=\inf_{\rho\in\mathcal{D}(\mathcal{H}_{RQ})}\ \max_{\substack{0\le X\le(\mathcal{I}^R\otimes\mathcal{E})(\rho^{RQ}),\\ 0\le X\le(\mathcal{I}^R\otimes\mathcal{F})(\rho^{RQ})}}\mathrm{Tr}\,X.\end{aligned}$$

To prove this, we notice that for a fixed input ρ^{RQ}, the optimization problem

$$\max_{\substack{0\le X\le(\mathcal{I}^R\otimes\mathcal{E})(\rho^{RQ}),\\ 0\le X\le(\mathcal{I}^R\otimes\mathcal{F})(\rho^{RQ})}} \operatorname{Tr} X$$

can be formulated as the following semidefinite program [31]:

Primal problem

$$\begin{aligned} \text{maximize:}\quad & \langle I, X\rangle \\ \text{subject to:}\quad & \Phi(X)\le B, \\ & X\in \operatorname{Pos}(\mathcal{H}_{RQ}). \end{aligned}$$

Dual problem

$$\begin{aligned} \text{minimize:}\quad & \langle B, Y\rangle \\ \text{subject to:}\quad & \Phi^\dagger(Y)\ge I, \\ & Y\in \operatorname{Pos}(\mathcal{H}_{RQ}\oplus\mathcal{H}_{RQ}). \end{aligned}$$

In the above formula, Φ is the super-operator

$$\Phi : \mathrm{L}(\mathcal{H}_{RQ})\to \mathrm{L}(\mathcal{H}_{RQ}\oplus\mathcal{H}_{RQ})$$

as

$$\Phi(X)=\begin{pmatrix} X & 0\\ 0 & X\end{pmatrix},$$

where the adjoint super-operator

$$\Phi^\dagger : \mathrm{L}(\mathcal{H}_{RQ}\oplus\mathcal{H}_{RQ})\to \mathrm{L}(\mathcal{H}_{RQ})$$

is given by

$$\Phi^\dagger\begin{pmatrix} Z & \cdot\\ \cdot & W\end{pmatrix}=Z+W,$$

and

$$B=\begin{pmatrix} (\mathcal{I}^R\otimes\mathcal{E})(\rho^{RQ}) & 0\\ 0 & (\mathcal{I}^R\otimes\mathcal{F})(\rho^{RQ})\end{pmatrix}.$$

Choose $Y=I>0$, then $\Phi^\dagger(Y)=2I>0$. This dual program is strictly feasible. Thus, the primal value and dual value are the same [38].

Now we are going back to the original problem by considering the dual problem with ρ^{RQ} ranging over all possible states.

Let $\mathcal{B}$ denote the following set

$$\left\{\begin{pmatrix} (\mathcal{I}^R\otimes\mathcal{E})(\rho^{RQ}) & 0\\ 0 & (\mathcal{I}^R\otimes\mathcal{F})(\rho^{RQ})\end{pmatrix} : \rho\in \mathrm{D}(\mathcal{H}_{RQ})\right\}.$$

$\mathcal{B}$ is a compact set because it is a closed bounded set in a finite-dimensional space.

According to the compactness of $\mathcal{B}$, we have the following

$$\begin{aligned} &\inf_{\rho\in \mathrm{D}(\mathcal{H}_{RQ})}\ \max_{\substack{0\le X\le(\mathcal{I}^R\otimes\mathcal{E})(\rho),\\ 0\le X\le(\mathcal{I}^R\otimes\mathcal{F})(\rho)}} \operatorname{Tr} X \\ =&\inf_{B\in\mathcal{B}}\ \min_{\substack{\Phi^\dagger(Y)\ge I,\\ Y\in\operatorname{Pos}(\mathcal{H}_{RQ}\oplus\mathcal{H}_{RQ})}} \langle B, Y\rangle \\ =&\min_{\substack{\Phi^\dagger(Y)\ge I,\\ Y\in\operatorname{Pos}(\mathcal{H}_{RQ}\oplus\mathcal{H}_{RQ})}}\ \inf_{B\in\mathcal{B}} \langle B, Y\rangle \\ =&\min_{\substack{\Phi^\dagger(Y)\ge I,\\ Y\in\operatorname{Pos}(\mathcal{H}_{RQ}\oplus\mathcal{H}_{RQ})}}\ \min_{B\in\mathcal{B}} \langle B, Y\rangle \\ =&\langle B_0, Y_0\rangle, \end{aligned}$$

for some $B_0\in\mathcal{B}$ and $Y_0\in\operatorname{Pos}(\mathcal{H}_{RQ}\oplus\mathcal{H}_{RQ})$.

Let ρ_0^{RQ} be such that

$$\begin{aligned} Y_0&=\begin{pmatrix} M & \cdot\\ \cdot & N\end{pmatrix} \\ B_0&=\begin{pmatrix} (\mathcal{I}^R\otimes\mathcal{E})(\rho_0^{RQ}) & 0\\ 0 & (\mathcal{I}^R\otimes\mathcal{E})(\rho_0^{RQ})\end{pmatrix}. \end{aligned}$$

For ρ_0^{RQ}, the intersection of the supports of $(\mathcal{I}^R\otimes\mathcal{E})(\rho_0^{RQ})$ and $(\mathcal{I}^R\otimes\mathcal{F})(\rho_0^{RQ})$ has a non-zero element. It indicates that there exists a non-zero $0\le G\le(\mathcal{I}^R\otimes\mathcal{E})(\rho_0^{RQ}), (\mathcal{I}^R\otimes\mathcal{F})(\rho_0^{RQ})$.

According to $\Phi^\dagger(Y_0)\ge I$, we have $M+N\ge I$. Let

$$\begin{aligned} \eta=&\langle B_0, Y_0\rangle \\ =&\langle(\mathcal{I}^R\otimes\mathcal{E})(\rho_0^{RQ}), M\rangle+\langle(\mathcal{I}^R\otimes\mathcal{F})(\rho_0^{RQ}), N\rangle \\ \ge&\langle G, M\rangle+\langle G, N\rangle \\ =&\langle G, M+N\rangle \\ \ge&\langle G, I\rangle \\ =&\operatorname{Tr} G>0. \end{aligned}$$

We can conclude that this η satisfies the wanted property. □

The following observation shows that if $I\in\operatorname{span}\{E_i^\dagger F_j\}$, then $\mathcal{E}$ and $\mathcal{F}$ cannot change the fidelity of two quantum states significantly.

Lemma 2. *If $I\in\operatorname{span}\{E_i^\dagger F_j\}$, then, there exists $\zeta>0$, depending only on $\mathcal{E}$ and $\mathcal{F}$, such that for all ρ,σ on a potentially larger Hilbert space RQ,*

$$F((\mathcal{I}^R\otimes\mathcal{E})(\rho^{RQ}), (\mathcal{I}^R\otimes\mathcal{F})(\sigma^{RQ}))\ge\zeta F(\rho,\sigma).$$

Proof. The condition $I\in\operatorname{span}\{E_i^\dagger F_j\}$ leads us to the existence of $\chi_{i,j}\in\mathbb{C}$ such that

$$I=\sum_{i,j=1}^m \chi_{i,j}E_i^\dagger F_j.$$

Using polar decomposition of the coefficient matrix $\chi_{i,j}$, we can always assume that

$$I = \sum_{i=1}^{m} \chi_i E_i^\dagger F_i,$$

and $\chi_i \geq 0$. We use $\chi = \max_i \chi_i$ to denote the largest χ_i.

For any ρ^{RQ} and σ^{RQ}, by Uhlmann's Theorem 2, there exist $\left|\psi_\rho^{RQT}\right\rangle$ and $\left|\psi_\sigma^{RQT}\right\rangle$ being ρ and σ's purifications respectively, and

$$F(\rho^{RQ}, \sigma^{RQ}) = F(\psi_\rho^{RQT}, \psi_\sigma^{RQT}) = |\langle\psi_\rho^{RQT}|\psi_\sigma^{RQT}\rangle|.$$

Now we can have the following

$$\begin{aligned}
&F((\mathcal{I}^R \otimes \mathcal{E})(\rho^{RQ}), (\mathcal{I}^R \otimes \mathcal{F})(\sigma^{RQ})) \\
\geq &F((\mathcal{I}^{RT} \otimes \mathcal{E})(\psi_\rho^{RQT}), (\mathcal{I}^{RT} \otimes \mathcal{F})(\psi_\sigma^{RQT})) \\
= &F[\sum_{i=1}^{m} (I^{RT} \otimes E_i)\psi_\rho^{RQT}(I^{RT} \otimes E_i)^\dagger, \\
&\sum_{i=1}^{m} (I^{RT} \otimes F_i)\psi_\sigma^{RQT}(I^{RT} \otimes F_i)^\dagger] \\
\geq &\sum_{i=1}^{m} F[(I^{RT} \otimes E_i)\psi_\rho^{RQT}(I^{RT} \otimes E_i)^\dagger, \\
&(I^{RT} \otimes F_i)\psi_\sigma^{RQT}(I^{RT} \otimes F_i)^\dagger] \\
= &\sum_{i=1}^{m} |\langle\psi_\rho^{RQT}|(I^{RT} \otimes E_i)^\dagger(I^{RT} \otimes F_i)|\psi_\sigma^{RQT}\rangle| \\
\geq &\frac{1}{\chi}\sum_{i=1}^{m} |\langle\psi_\rho^{RQT}|\chi_i(I^{RT} \otimes E_i^\dagger F_i)|\psi_\sigma^{RQT}\rangle| \\
\geq &\frac{1}{\chi}|\langle\psi_\rho^{RQT}|(I^{RT} \otimes (\sum_{i=1}^{m} \chi_i E_i^\dagger F_i)|\psi_\sigma^{RQT}\rangle| \\
= &\frac{1}{\chi}|\langle\psi_\rho^{RQT}|I^{RQT}|\psi_\sigma^{RQT}\rangle| \\
= &\frac{1}{\chi}|\langle\psi_\rho^{RQT}|\psi_\sigma^{RQT}\rangle| \\
= &\frac{1}{\chi}F(\psi_\rho^{RQT}, \psi_\sigma^{RQT}) \\
= &\frac{1}{\chi}F(\rho^{RQ}, \sigma^{RQ}).
\end{aligned}$$

The first inequality is due to Fact 4, the monotonicity of the fidelity under partial trace. The second inequality is due to Fact 1, the strong concavity of the fidelity, and positive homogeneity.

Therefore, we choose $\zeta = \frac{1}{\chi}$. □

IV. MAIN RESULTS

Our main result is as follows

Theorem 1. *The Chernoff exponent for quantum operations, Eq. 1, is finite if and only if they cannot be distinguished perfectly.*

For two distinct quantum operations, $\mathcal{E}$ and $\mathcal{F}$, it is straightforward to verify that $P_{err} \leq \exp(-n\xi')$ for some $\xi' > 0$ by observing the following process. First, one can always find an input state ρ such that $\mathcal{E}(\rho)$ and $\mathcal{F}(\rho)$ are distinct. Then we feed ρ as input through the device for n times. After that, the problem becomes to distinguish $\mathcal{E}(\rho)^{\otimes n}$ and $\mathcal{F}(\rho)^{\otimes n}$. Invoking the celebrated result on the Chernoff exponent for quantum states, we know that the error probability of distinguishing two different quantum states with identical copies decays according to an exponential function. Notice that this protocol only provides an upper bound on the minimal error probability of distinguishing $\mathcal{E}$ and $\mathcal{F}$, so one can conclude that $P_{err} \leq \exp(-n\xi')$ for some $\xi' > 0$.

The above arguments show that the error decays at least exponentially. In other words, $\xi_{\mathcal{E},\mathcal{F}}$ is greater than 0. However, this scheme can be far from optimal. Perfect discrimination between unknown processes chosen from a finite set is shown to be possible. For two quantum operations that can be distinguished perfectly, $\xi_{\mathcal{E},\mathcal{F}} = \infty$, we prove that this is the only case where $\xi_{\mathcal{E},\mathcal{F}} = \infty$. Moreover, we provide an easy computable upper bound of $\xi_{\mathcal{E},\mathcal{F}}$ for quantum operations that can not be distinguished perfectly, i.e., $P_{err} \geq \exp(-n\xi)$, where the parameter ξ is a positive constant that depends on the two operations only.

Proof. The only if part of Theorem 1 is trivial. The if part follows Proposition 1 in Section II, and we prove it for prior probability distribution $\Pi_0 = \Pi_1 = 1/2$. Also, we prove the Chernoff exponent is independent of a prior distribution in Proposition 2.

To prove the only if part of Theorem 1 under distribution $\Pi_0 = \Pi_1 = 1/2$, we only need to show that when either condition in Proposition 1 is violated, the error probability is at least an exponential function of the number of channel uses.

First, we suppose $\mathcal{E}$ and $\mathcal{F}$ are joint, in the sense that the produced quantum states have non-zero overlapping supports for any common input state, we show that there exists $\eta > 0$ such that $P_{err,n} \geq \eta^n/2$ in the following:

Refer to Figure 1 for our notations. By employing Lemma 1, we observe that there exists $0 \leq A_1 \leq \rho_1, \sigma_1$ such that $\operatorname{Tr} A_1 \geq \eta$. Then, $0 \leq A_1' = \mathcal{G}_1(A_1) \leq \rho_1', \sigma_1'$ such that $\operatorname{Tr} A_1' = \operatorname{Tr} A_1 \geq \eta$. Then, there exists $0 \leq A_2 \leq \rho_2, \sigma_2$ such that $\operatorname{Tr} A_2 \geq \eta^2$. Thus, $0 \leq A_2' =$

$\mathcal{G}_2(A_2) \leq \rho_1', \sigma_1'$ such that $\operatorname{Tr} A_2' = \operatorname{Tr} A_2 \geq \eta^2 \cdots$ There exists $0 \leq A_n \leq \rho_n, \sigma_n$ such that $\operatorname{Tr} A_n \geq \eta^n$.

By Helstrom's celebrated result on state discrimination [14], we know that the discrimination error satisfies the following

$$\begin{aligned} P_{err,min,n} &= \inf_{\rho_n,\sigma_n} \frac{1}{2}(1 - \frac{\operatorname{Tr}|\rho_n - \sigma_n|}{2}) \\ &= \inf_{\rho_n,\sigma_n} \frac{1}{2}(1 - \frac{\operatorname{Tr}|\rho_n - A_n - \sigma_n + A_n|}{2}) \\ &\geq \inf_{\rho_n,\sigma_n} \frac{1}{2}(1 - \frac{\operatorname{Tr}(\rho_n - A_n) + \operatorname{Tr}(\sigma_n - A_n)}{2}) \\ &= \frac{\operatorname{Tr} A_n}{2} \\ &\geq \frac{\eta^n}{2}. \end{aligned}$$

The first inequality is according to the triangle inequality and $0 \leq A_n \leq \rho_n, \sigma_n$.

Second, suppose two quantum operations $\mathcal{E}$ and $\mathcal{F}$ can not transform non-orthogonal states into orthogonal states. This is equivalent to $I \in \operatorname{span}\{E_i^* F_j\}$, as illustrated in Remark 1 at the end of Section II. We show in the following that there exists $\mu > 0$ such that $P_{err,n} \geq \mu^n/4$. The proof of this part is according to the observation that if two quantum operations cannot make nonorthogonal states orthogonal, they cannot change their fidelity significantly.

Refer to Figure 1 for our notations. By employing Lemma 2, we observe that there exists $\zeta > 0$ such that after n uses of the unknown quantum operation, the possible outcome states ρ_n and σ_n satisfy the following:

$$\begin{aligned} F(\rho_n, \sigma_n) &\geq \zeta F(\rho_{n-1}', \sigma_{n-1}') \\ &\geq \zeta F(\rho_{n-1}, \sigma_{n-1}) \\ &\geq \zeta^2 F(\rho_{n-2}', \sigma_{n-2}') \\ &\cdots \\ &\geq \zeta^{n-1} F(\rho_1, \sigma_1) \\ &\geq \zeta^n F(\psi, \psi) \\ &= \zeta^n, \end{aligned}$$

where $F(\cdot,\cdot)$ denotes the fidelity of quantum states.

The first inequality is due to Lemma 2. The second inequality is due to the monotonicity of fidelity under any quantum operation.

According to the relation between fidelity and trace distance, we have

$$\begin{aligned} P_{err,min,n} &= \inf_{\rho_n,\sigma_n} \frac{1}{2}(1 - \operatorname{Tr}|\rho_n - \sigma_n|/2) \\ &\geq \frac{1}{2}(1 - \sqrt{1 - \zeta^{2n}}) \\ &\geq \frac{\zeta^{2n}}{4}, \end{aligned}$$

where the minimization ranges across all possible output ρ_n and σ_n.

Therefore, we can choose $\mu = \zeta^2$.

Putting these two conditions together, we obtain that for indistinguishable quantum operations $\mathcal{E}, \mathcal{F}$ under uniform distribution,

$$\xi_{\mathcal{E},\mathcal{F}} \leq \min\{-\log \eta, -\log \mu\}. \tag{2}$$

□

If we have more than two quantum operations, suppose we have a quantum device that is secretly chosen from $\{\mathcal{E}_1, \cdots \mathcal{E}_r\}$, a known set of quantum operations according to prior probability distribution $\{\Pi_1, \cdots, \Pi_r\}$. Our goal is to see which quantum operation the quantum device implements by using the device many times. The definition of the Chernoff exponent for two quantum operations Eq.(1) can be easily generalized into multiple quantum operations, where $P_{err,min,n}$ now is defined as the infimum error probability for distinguishing multiple quantum operations with n uses. One can prove that the Chernoff exponent for multiple quantum operations shares the same properties as the Chernoff exponent for two quantum operations. This exponent does not depend on the prior probability distribution, and it is infinite if these quantum operations are mutually perfectly distinguishable. Moreover, it is, at most, the minimal mutual Chernoff exponent for quantum operations.

Proposition 2. *The Chernoff exponent for multiple quantum operations, Eq. 1, does not depend on the prior distribution. Moreover, the multi-channel Chernoff exponent is upper bounded by the smallest pairwise Chernoff exponent.*

Proof. For prior $(\Pi_1, \Pi_2, \cdots, \Pi_r)$, $\Pi_1 \geq \Pi_2 \geq \cdots \Pi_r > 0$ and fixed n, we use $P_{err,min,n,\Pi}$ to denote the infimum error probability. $P_{err,min,n}$ denotes the infimum error probability for uniform prior $(1/r, 1/r, \cdots, 1/r)$. For any discrimination scheme, we

let $1-p_{n,i}$ be the probability of correctly identifying the i-th channel. Then we have

$$\Pi_1 \sum_{i=1}^{r} p_{n,i} \geq \sum_i \Pi_i p_{n,i} \geq \Pi_r \sum_{i=1}^{r} p_{n,i} \geq \frac{\Pi_r}{2}(p_{n,k}+p_{n,l})$$

for any $1 \leq k, l \leq r$.

Since this inequality holds for the error probabilities in any strategy, one can just take the infimum over all strategies in each term of the inequality and have

$$\begin{aligned} r\Pi_1 P_{err,min,n} \geq & P_{err,min,n,\Pi} \\ \geq & r\Pi_r P_{err,min,n} \\ \geq & \Pi_r P_{err,min,n,\{k,l\}} \end{aligned}$$

Therefore,

$$\begin{aligned} -\varlimsup_{n\to\infty} \frac{P_{err,min,n}}{n} = & -\varlimsup_{n\to\infty} \frac{P_{err,min,n,\Pi}}{n} \\ \leq & -\varlimsup_{n\to\infty} \frac{P_{err,min,n,\{k,l\}}}{n}. \end{aligned}$$

That is, the Chernoff exponent for multiple quantum operations does not depend on prior. □

According to the proof, we can also conclude that

Corollary 1. *The Chernoff exponent for multiple quantum operations, Eq. 1, is infinite if and only if the quantum operations are mutually perfectly distinguishable.*

Proof. Suppose we are given a quantum operation $\mathcal{E}$ being one of the quantum operations $\mathcal{E}_1, \mathcal{E}_2, \cdots, \mathcal{E}_r$, and any two quantum operations can be distinguished perfectly. Let a protocol produce orthogonal quantum states ρ_1 and ρ_2 for quantum operations $\mathcal{E}_1$ and $\mathcal{E}_2$, respectively.

Now we run the protocol on $\mathcal{E}$ and measure the output. We employ the measurement which can distinguish ρ_1 and ρ_2 perfectly. If the measurement outcome corresponds to ρ_1, then we know $\mathcal{E}$ can not be $\mathcal{E}_2$; otherwise, it can not be $\mathcal{E}_1$.

Therefore, via finite uses of $\mathcal{E}$, we can eliminate one candidate. By repeating this procedure, we can conclude that the Chernoff exponent is ∞.

Otherwise, if two quantum operations, say $\mathcal{E}_1$ and $\mathcal{E}_2$, can not be distinguished perfectly. According to the proof of Proposition 2, the multi-channel Chernoff exponent is upper bounded by the smallest pairwise exponent which is no more than the Chernoff exponent of $\mathcal{E}_1$ and $\mathcal{E}_2$, a finite number by Theorem 1. □

V. Conclusion and Open Problems

In this paper, we introduce the Chernoff exponent for quantum operations. We show the Chernoff exponent is finite if and only if the operations are not perfectly distinguishable. More precisely, we provide computable upper bounds of the Chernoff exponent by proving lower bounds on the error probability of distinguishing quantum operations with n uses. Our result is an asymptotic generalization of the diamond norm.

There are several open questions. One relates to the local operations and classical communication (LOCC)-Chernoff distance. Motivated by the quantum Chernoff theorem [3], [25], the LOCC-Chernoff exponent studies the distinguishability of two bipartite mixed states under the constraint of LOCC, in the limit of many copies [8], [23]. There is a significant difference between the LOCC Chernoff exponent and the standard Chernoff exponent. Orthogonality does not indicate perfect LOCC distinguishability. More precisely, there exist quantum states which cannot be locally distinguished but multicopy makes them perfectly distinguishable [35], [36]. This behavior is similar to the discrimination of quantum operations. A fundamental question regarding the LOCC Chernoff exponent is still not answered: For two quantum states that are not LOCC perfectly distinguishable, even in the limit of many copies, does the LOCC discrimination error always decay exponentially? The first difficulty is we do not have a characterization of LOCC distinguishability of quantum states, even though this problem has been studied for more than 20 years [34], [33], [12], [5], [30], [16], [4].

We thank the editor and the anonymous reviewers whose comments have greatly improved this manuscript. This work is supported by ARC Discovery Early Career Researcher Award DE180100156 and ARC Discovery Program DP210102449.

References

[1] A. Acin, "Statistical distinguishability between unitary operations." *Physical Review Letters*, **87**(17): 177901, 2001.

[2] D. Aharonov, A. Kitaev, and N. Nisan, "Quantum circuits with mixed states." *Proceeding of the Thirtieth Annual ACM Symposium on Theory of Computation*, pp. 20-30, 1997.

[3] K. M. R. Audenaert, J. Casamiglia, R. Munoz-Tapia, E. Bagan, Ll. Masanes, A. Acin, and F. Verstraete, "Discriminating states: the quantum Chernoff bound." *Physical Review Letters*, **98**(16): 160501, 2007.

[4] S. Bandyopadhyay, A. Cosentino, N. Johnston, V. Russo, J. Watrous, and N. Yu, "Limitations on separable measurements by convex optimization." *IEEE Transactions on Information Theory*, **61**(6): 3593, 2015.

[5] C. H. Bennett, D. P. DiVincenzo, T. Mor, P. W. Shor, J. A. Smolin, and B. M. Terhal, "Unextendible product bases and bound entanglement." *Physical Review Letters*, **82**(26): 5385, 1999.

[6] H. Chernoff, "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations." *The Annals of Mathematical Statistics*, **23**(4): 493, 1952.

[7] Tom Cooney, Milan Mosonyi, and Mark M. Wilde, "Strong converse exponents for a quantum channel discrimination problem and quantum-feedback-assisted communication." *Communications in Mathematical Physics*, **344**(3): 797-829, 2016.

[8] J. Calsamiglia, J. I. de Vicente, R. Muñoz-Tapia, and E. Bagan, "Local discrimination of mixed states." *Physical Review Letters*, **105**(8): 080504, 2010.

[9] R. Duan, Y. Feng, and M. Ying, "Entanglement is not necessary for perfect discrimination between unitary operations." *Physical Review Letters*, **98**(10): 100503, 2007.

[10] R. Duan, Y. Feng, and M. Ying, "Perfect distinguishability of quantum operations." *Physical Review Letters*, **103**(21): 210501, 2009.

[11] C. A. Fuchs, J. Van De Graaf, "Cryptographic distinguishability measures for quantum-mechanical states." *IEEE Transactions on Information Theory*, **45**(4): 1216, 1999. 1999.

[12] S. Ghosh, G. Kar, A. Roy, A. Sen(De) and U. Sen, "Distinguishability of Bell states." *Physical Review Letters*, **87**(27): 277902, 2001.

[13] M. Hayashi, "Discrimination of two channels by adaptive methods and its application to quantum system." *IEEE Transactions on Information Theory*, **55**(8): 3807, 2009.

[14] C. W. Helstrom, "Detection theory and quantum mechanics." *Information and Control*, **10**(3): 254, 1967.

[15] A. W. Harrow, A. Hassidim, D. W. Leung, and J. Watrous "Adaptive versus non-adaptive strategies for quantum channel discrimination." *Physical Review A*, **81**(3): 032339, 2010.

[16] M. Hayashi, D. Markham, M. Murao, M. Owari and S. Virmani, "Bounds on multipartite entangled orthogonal state discrimination using local operations and classical communication." *Physical Review Letters*, **96**(4): 040501, 2006.

[17] A. S. Holevo, "An analog of the theory of statistical decisions in noncommutative probability theory." *Transactions of the Moscow Mathematical Society*, **26**: 133, 1972.

[18] Z. Ji, Y. Feng, R. Duan, and M. Ying, "Identification and distance measures of measurement apparatus." *Physical Review Letters*, **96**(20): 200401, 2006.

[19] A. Kitaev, "Quantum computations: Algorithms and error correction." *Russian Mathematical Surveys*, **52**(6): 1191, 1997.

[20] A. Laing, T. Rudolph, and J. L. O'Brien, "Experimental quantum process discrimination." *Physical Review Letters*, **102**(16): 160502, 2009.

[21] L. Li and D. Qiu, "Local entanglement is not necessary for perfect discrimination between unitary operations acting on two qudits by local operations and classical communication." *Physical Review A*, **77**(03): 032337 , 2008.

[22] K. Li, "Discriminating quantum states: the multiple Chernoff distance." *Annals of Statistics*, 44 (4): 1661-1679 2016.

[23] W. Matthews, A. Winter, "On the Chernoff distance for asymptotic LOCC discrimination of bipartite quantum states." *Communications in Mathematical Physics*, **285**: 161, 2009.

[24] M. A. Nielsen, I. Chuang, "Quantum computation and quantum information." *Cambridge University Press*, Cambridge, UK, 2000.

[25] M. Nussbaum, A. Szkoła, "The Chernoff lower bound for symmetric quantum hypothesis testing." *The Annals of Statistics*, **37**(2): 1040, 2009.

[26] K. R. Parthasarathy, "On consistency of the maximum likelihood method in testing multiple quantum hypotheses." *Stochastics in Finite and Infinite Dimensions*, Birkhäuser Boston, 361, 2001.

[27] D. Puzzuoli, J. Watrous, "Ancilla dimension in quantum operation discrimination." *Annales Henri Poincaré*, Springer International Publishing, 2016.

[28] Masahiro Takeoka, and Mark M. Wilde, "Optimal estimation and discrimination of excess noise in thermal and amplifier channel." *arXiv:1611.09165*, 2016.

[29] A. Uhlmann, "The "transition probability" in the state space of a *-algebra", *Reports on Mathematical Physics*, **9**(2): 273, 1976.

[30] J. Watrous, "Bipartite subspaces having no bases distinguishable by local operations and classical communication." *Physical Review Letters*, **95**(8): 080505, 2005.

[31] J. Watrous, "Semidefinite programs for completely bounded norms." *Theory of Computing*, **5**(11): 217, 2009.

[32] J. Watrous, "Theory of quantum information." *University of Waterloo Fall*, 128, 2011.

[33] J. Walgate, A. J. Short, L. Hardy and V. Vedral, "Local distinguishability of multipartite orthogonal quantum states." *Physical Review Letters*, **85**(23): 4972, 2000.

[34] N. Yu, R. Duan and M. Ying, "Any $2 \otimes n$ subspace is locally distinguishable." *Physical Review A*, **84**(1): 012304, 2011.

[35] N. Yu, R. Duan and M. Ying, "Four locally indistinguishable ququad-ququad orthogonal maximally entangled states." *Physical Review Letters*, **109**(2): 020506, 2012.

[36] N. Yu, R. Duan, and M. Ying, "Distinguishability of quantum states by positive operator-valued measures with positive partial transpose." *IEEE Transactions on Information Theory*, **60**(4): 2069, 2014.

[37] X. F. Zhou, Y. S. Zhang, and G. C. Guo "Unitary Transformations Can Be Distinguished Locally." *Physical Review Letters*, **99**(17): 170401, 2007.

[38] M. Slater, "Lagrange Multipliers Revisited." Cowles Commission Discussion Paper No. 403 (Report).

Nengkun Yu is a Senior Lecturer in the Centre for Quantum Software and Information, University of Technology Sydney. He received the B.S. and Ph.D. degrees from the Department of Computer Science and Technology, Tsinghua University, Beijing, China, in July of 2008 and 2013. From January 2014 to July 2016, Nengkun was a postdoc at the Institute for Quantum Computing at the University of Waterloo, Canada. His research focuses on quantum computing.

Li Zhou is a postdoc at Max Planck Institute for Security and Privacy. He received my Ph.D. in Computer Science and Technology from Tsinghua University in 2019. His research focuses on quantum programs and protocols, including static analysis and verification, and runtime debugging and testing.