Simple Upper and Lower Bounds on the Ultimate Success Probability for Discriminating Arbitrary Finite-Dimensional Quantum Processes

Kenji Nakahira and Kentaro Kato
Quantum Information Science Research Center, Quantum ICT Research Institute, Tamagawa University, Machida, Tokyo 194-8610, Japan

(Received 4 January 2021; revised 19 April 2021; accepted 22 April 2021; published 18 May 2021)

We consider the problem of discriminating finite-dimensional quantum processes, also called quantum supermaps, that can consist of multiple time steps. Obtaining the ultimate performance for discriminating quantum processes is of fundamental importance, but is challenging mainly due to the necessity of considering all discrimination strategies allowed by quantum mechanics, including entanglement-assisted strategies and adaptive strategies. In the case in which the processes to be discriminated have internal memories, the ultimate performance would generally be more difficult to analyze. In this Letter, we present a simple upper bound on the ultimate success probability for discriminating arbitrary quantum processes. In the special case of multishot channel discrimination, it can be shown that the ultimate success probability increases by at most a constant factor determined by the given channels if the number of channel evaluations increases by one. We also present a lower bound based on Bayesian updating, which has a low computational cost. Our numerical experiments demonstrate that the proposed bounds are reasonably tight. The proposed bounds do not explicitly depend on any quantum phenomena, and can be readily extended to a general operational probabilistic theory.

DOI: 10.1103/PhysRevLett.126.200502

A quantum process, which is a mathematical object that models the probabilistic behavior of quantum devices, plays an essential role in quantum information science. Discriminating between quantum processes is a fundamental and challenging problem, which forms the basis of a large class of problems in quantum information theory such as quantum communication, quantum cryptography, and quantum metrology. The simplest instance of this problem is a quantum state discrimination problem, which has been widely studied since the end of the 1960s [1–3]. Since the maximum success probability is often quite difficult to obtain accurately, its upper and lower bounds have been developed [4–9]. Discrimination problems of quantum measurements [10–16] and quantum channels [17–22] are also particular instances. In quantum channel discrimination, entanglement with an ancillary system and an adaptive strategy may be required to achieve the ultimate performance, which makes this problem difficult in general. A quantum process describes the most general transformation that maps channels to channels [23,24]. A process can consist of several memory channels [25–29], whose output states can depend on the previous input states. As an example of process discrimination, we can consider the problem of retrieving the value of the bit that is encoded into the reflectivity of a certain memory cell, which is often referred to as quantum reading [30]. This problem can be seen as the discrimination of two processes, even when a finite number of uses of the memory cell are allowed and the reflectivity may change depending on the previous inputs to it. Although finding the ultimate performance for discriminating such processes is extremely difficult, it is of fundamental importance in various fields including quantum cryptography [31], quantum game theory [32], and quantum algorithms.

In this Letter, we derive a simple upper bound on the ultimate success probability for discriminating arbitrary finite-dimensional quantum processes. In the special case of multishot channel discrimination, our approach can ensure that the ultimate success probability increases by at most a constant factor, which is determined by the given channels, if the number of channel evaluations increases by one. Note that an upper bound for channel discrimination has been reported very recently [33], which is based on port-based teleportation [34,35]. We present numerical simulations that show that, at least in a certain multishot channel discrimination problem, our upper bound is significantly tighter than that of Ref. [33].

A tight lower bound is also required to accurately evaluate the ultimate performance. Since the success probability of any discrimination allowed by quantum mechanics yields a lower bound on the ultimate success probability, a natural approach to derive such a bound is to find good discrimination. As an illustration of this approach, certain nonadaptive discrimination has sometimes been discussed [33,36]. However, an adaptive strategy would outperform the best nonadaptive strategy except for some special cases [21,24,37–41]. For example, it is known that there exist two channels that can be perfectly distinguished by using an adaptive strategy with only two uses of the channel, while they cannot be perfectly distinguished by using any nonadaptive strategy with a finite number of uses [42]. We present a lower bound that is obtained by an adaptive discrimination strategy based on Bayesian updating. Our work is motivated by the fact that, for quantum state discrimination, a Bayesian updating approach has been shown to be effective [43–46] and to be optimal at least for discriminating two identical copies of a pure state [47–49]. Our numerical results demonstrate the tightness of the proposed lower bound. We should emphasize that the proposed upper and lower bounds do not explicitly depend on any quantum phenomena, such as entanglement and quantum teleportation, and can be readily extended to operational probabilistic theory (or generalized probabilistic theory) [50–54].

Process discrimination problems.—Suppose that we want to discriminate between M quantum processes $\mathcal{E}_1, \ldots, \mathcal{E}_M$ as accurately as possible, where each $\mathcal{E}_m$ is a process consisting of T channels $\Lambda_m^{(1)}, \ldots, \Lambda_m^{(T)}$. The most general discrimination protocol can be expressed as the collection of a state σ_1, channels $\sigma_2, \ldots, \sigma_T$, and a measurement $\Pi \coloneqq \{\Pi_k\}_{k=1}^M$ (see Fig. 1). Channels $\Lambda_m^{(t)}$ and $\Lambda_m^{(t+1)}$ are connected by an ancillary system W'_t. A process $\mathcal{E}_m$, which is also called a quantum supermap or a quantum comb [24], is equivalent to a sequence of memory channels [55]. In the first step of process discrimination, a bipartite system $V_1 \otimes V'_1$ is prepared in an initial state σ_1. Its part V_1 is sent through the channel $\Lambda_m^{(1)}$, followed by a channel σ_2. Then, we send the system V_2 through the channel $\Lambda_m^{(2)}$ and so on. After T steps, a quantum measurement Π is performed on the system W_T. The problem of discriminating M channels $\Lambda_1, \ldots, \Lambda_M$ with T queries can be regarded as a special case of a processes discrimination problem with $\Lambda_m^{(t)} = \Lambda_m$ and $W'_t = \mathbb{C}$ $(\forall\, m, t)$. For simplicity, we focus on the case of equal prior probabilities. Let $P_{k|m}$ be the conditional probability that the measurement outcome is k given that the given process is $\mathcal{E}_m$, which is expressed by

$$P_{k|m} \coloneqq \Pi_k \circ \Lambda_m^{(T)} \circ \sigma_T \circ \cdots \circ \sigma_2 \circ \Lambda_m^{(1)} \circ \sigma_1.$$

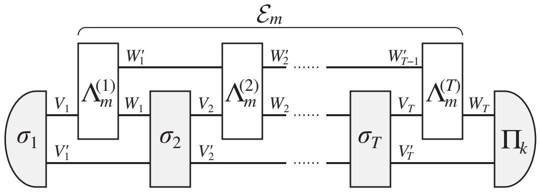


FIG. 1. General protocol of quantum process discrimination. $\mathcal{E}_m$ is a process consisting of T channels $\Lambda_m^{(1)}, \ldots, \Lambda_m^{(T)}$. Discrimination is characterized by the collection of a state σ_1, channels $\sigma_2, \ldots, \sigma_T$, and a measurement $\{\Pi_k\}_{k=1}^M$.

The success probability P is written as

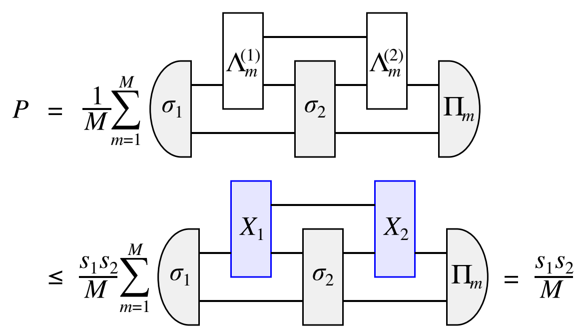
$$P \coloneqq \frac{1}{M} \sum_{m=1}^{M} P_{m|m}, \tag{1}$$

where $\circ$ denotes function composition. Our objective is to find discrimination $(\sigma_1, \ldots, \sigma_T, \Pi)$ that maximizes the success probability. It is known that this optimization problem is formulated as a semidefinite programming (SDP) problem of order $\tilde{N} \coloneqq \prod_{t=1}^T N_{V_t} N_{W_t}$ [56], where N_{V_t} and N_{W_t} are, respectively, the dimensions of the systems V_t and W_t. Solving this problem requires time polynomial in $\tilde{N}$, and thus is generally intractable for large T. Indeed, in the case of $N_{V_t} = N_{W_t} \eqqcolon N$ for each t, for example, $\tilde{N} = N^{2T}$ is exponentially increasing with T.

Proposed upper bound.—The basic idea is quite simple: for each t, we only have to replace $\Lambda_m^{(t)}$ of Eq. (1) by $s_t X_t$, where s_t and X_t are, respectively, a positive real number and a channel satisfying $s_t X_t \geq \Lambda_m^{(t)}$ $(\forall\, m)$. For two single-step processes Λ and Λ', the inequality $\Lambda \geq \Lambda'$ denotes that $\Lambda - \Lambda'$ is completely positive. Such a pair (s_t, X_t) obviously exists. From Eq. (1), we have

$$P \leq \frac{1}{M} \sum_{m=1}^{M} \Pi_m \circ s_T X_T \circ \sigma_T \circ \cdots \circ \sigma_2 \circ s_1 X_1 \circ \sigma_1 = \frac{1}{M} \prod_{t=1}^{T} s_t,$$

where the equality follows from $\sum_{m=1}^M \Pi_m \circ X_T \circ \sigma_T \circ \cdots \circ \sigma_2 \circ X_1 \circ \sigma_1 = 1$. This gives that the ultimate success probability is upper bounded by $M^{-1} \prod_{t=1}^T s_t$. For example, in the case of $T = 2$, it is diagrammatically depicted as



To make this bound as tight as possible, we need to minimize $s_1, \ldots, s_T$. This problem is written as

$$\begin{aligned} &\text{minimize} \quad s_t \\ &\text{subject to} \quad s_t X_t \geq \Lambda_m^{(t)} \ (\forall\, m) \end{aligned} \tag{2}$$

with a real number s_t and a channel X_t from $W'_{t-1} \otimes V_t$ to $W'_t \otimes W_t$. Let $s_t^\star$ be the optimal value of problem (2). The proposed upper bound $\bar{P}_1$ is given by

$$\bar{P}_1 \coloneqq \frac{1}{M}\prod_{t=1}^{T} s_t^\star, \tag{3}$$

where the subscript 1 indicates that $\bar{P}_1$ is obtained by optimization problems for finding single-step processes (the same for $\bar{P}_2$, which will be defined below).

The above argument can be readily extended to obtain a tighter bound at the expense of additional complexity. For instance, instead of finding a single-step process $s_t X_t$ that is larger than $\Lambda_m^{(t)}$ as in Eq. (2), we can consider finding a pair of single-step processes that is larger than the pair $[\Lambda_m^{(t-1)}, \Lambda_m^{(t)}]$. Specifically, we consider the following optimization problem

$$\begin{aligned} &\text{minimize} \quad s_{t,2} \\ &\text{subject to} \quad s_{t,2} X_t \circ \eta \circ X_{t-1} \geq \Lambda_m^{(t)} \circ \eta \circ \Lambda_m^{(t-1)} \\ &\qquad\qquad (\forall\, m, \eta) \end{aligned} \tag{4}$$

with a real number $s_{t,2}$ and channels X_{t-1} and X_t, which are the same type as $\Lambda_m^{(t-1)}$ and $\Lambda_m^{(t)}$, respectively. η is any channel that can be sequentially connected to channels $\Lambda_m^{(t)}$ and $\Lambda_m^{(t-1)}$ such as $\Lambda_m^{(t)} \circ \eta \circ \Lambda_m^{(t-1)}$. Its optimal solution, $s_{t,2}^\star$, can be used to obtain an upper bound instead of $s_{t-1}^\star s_t^\star$. Thus, we obtain the following upper bound

$$\overline{P_2} \coloneqq \frac{1}{M}\prod_{t=1}^{T/2} s_{2t,2}^\star \quad \text{or} \quad \overline{P_2} \coloneqq \frac{s_T^\star}{M}\prod_{t=1}^{(T-1)/2} s_{2t,2}^\star, \tag{5}$$

for even or odd T, respectively. We can easily see $P \leq \overline{P_2} \leq \overline{P_1}$ [57].

With the so-called Choi-Jamiołkowski representation of X_t [58,59], Eqs. (2) and (4) can be formulated as SDP problems. Thus, their numerical optimal solutions can be efficiently obtained by several well-known SDP solvers. Analytical optimal solutions to these problems can be obtained in some cases, such as the case in which processes $\mathcal{E}_1, \ldots, \mathcal{E}_M$ have some kind of symmetry [60]; another example is shown in Sec. III of the Supplemental Material (SM) [61].

As a special case, we consider the T-shot discrimination of quantum channels. In this case, the optimal value $s_t^\star \eqqcolon s^\star$ of problem (2) is obviously independent of t. Let $P_T^\star$ be the ultimate success probability; then, since $P_T^\star$ increases as the number of evaluations T increases, $P_T^\star \geq P_{T-1}^\star \geq \cdots \geq P_1^\star$ holds. As an application of the above argument, we also obtain (see Sec. II B of the SM [61])

$$P_T^\star \leq s^\star P_{T-1}^\star \leq s^{\star 2} P_{T-2}^\star \leq \cdots \leq s^{\star T-1} P_1^\star = \frac{s^{\star T}}{M}, \tag{6}$$

where the equality follows from $P_1^\star = s^\star/M$ [56]. This equation implies that the ultimate success probability increases by at most $s^\star (\geq 1)$ times if the number of channel evaluations increases by one. Discrimination of quantum channels that are very close to each other is required in many application scenarios such as quantum illumination [63,64] and quantum reading [30]. In such a case, since $s^\star$ is very close to one, the inequality $P_{T-1}^\star \leq P_T^\star \leq s^\star P_{T-1}^\star$ provides a strong constraint. Equation (6) provides some useful properties. As an example, we can see that the given channels cannot be perfectly discriminated with T uses if $P_1^\star$ is smaller than $1/M^{1-1/T}$. As another example, in order for the ultimate success probability to be larger than a given threshold p, more than $\log_{s^\star} Mp$ evaluations are needed.

Proposed lower bound.—A natural approach for obtaining a lower bound is to restrict attention to certain types of discrimination strategies. A typical example is nonadaptive strategies. The success probability, $P^{(\mathrm{na})}$, of the best nonadaptive strategy would be more easily obtained than the ultimate success probability; for example, in the particular case of T-shot discrimination of two channels Λ_1 and Λ_2, it is well known that $P^{(\mathrm{na})}$ is given by $\frac{1}{2} + \frac{1}{4}\|\Lambda_1^{\otimes T} - \Lambda_2^{\otimes T}\|_\diamond$. However, adaptive strategies provide a clear advantage over nonadaptive ones in not a few cases.

We propose an adaptive strategy based on Bayesian updating to obtain a tight lower bound. In our method, channels $\sigma_2, \ldots, \sigma_T$ are restricted to measure-and-prepare (i.e., entanglement breaking) channels as illustrated in Fig. 2. The channel σ_t with $2 \leq t \leq T$ consists of a measurement, $\Pi^{(t-1)} \coloneqq \{\Pi_m^{(t-1)}\}_{m=1}^M$, followed by a state preparation, $\varrho^{(t)}$. The state preparation $\varrho^{(t)}$ and the measurement $\Pi^{(t)}$ can be connected by an ancillary system and may depend on the outcome of the previous measurement $\Pi^{(t-1)}$. Assume that they are independent of the outcome of measurements $\Pi^{(t-2)}, \Pi^{(t-3)}, \ldots$ to reduce the complexity. In such a scenario, we want to determine $\varrho^{(t)}$ and $\Pi^{(t)}$ such that the success probability is as high as possible. For practical computation, we need to optimize them sequentially for $t = 1, 2, \ldots$. Note that, since such discrimination

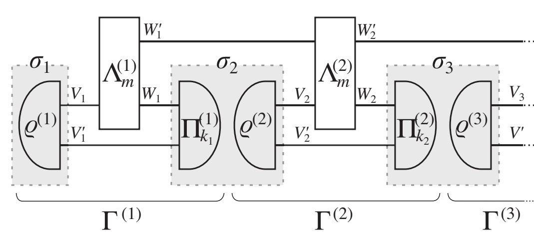


FIG. 2. Proposed protocol based on Bayesian updating. Each channel σ_t with $2 \leq t \leq T$ is restricted to a measure-and-prepare channel $\varrho^{(t)} \circ \Pi^{(t-1)}$. The state preparation $\varrho^{(t)}$ and the measurement $\Pi^{(t)}$ depend on the outcome of the previous measurement $\Pi^{(t-1)}$. To provide a tight lower bound on the ultimate success probability, $[\varrho^{(1)}, \Pi^{(1)}], [\varrho^{(2)}, \Pi^{(2)}], \ldots$ are sequentially optimized.

only requires state preparations and measurements, it has the advantage of being relatively easy to implement experimentally.

We here present a brief outline of the proposed method; we refer to Sec. IV of the SM [61] for details. Let $\varrho^{(1)} \coloneqq \sigma_1$ and $\Pi^{(T)} \coloneqq \Pi$; then, the sequence of processes shown in Fig. 2 is expressed by the sequential composition of

$$\Gamma^{(t)}_{k_t|m,k_{t-1}} \coloneqq \Pi^{(t)}_{k_t} \circ \Lambda^{(t)}_m \circ \varrho^{(t)},$$

where k_{t-1} is the outcome of $\Pi^{(t-1)}$. After some calculations, we find that the probability that the measurement $\Pi^{(t)}$ correctly distinguishes between the processes is given by

$$P^{(t)} \coloneqq \frac{1}{M} \sum_{m=1}^{M} q^{(t)}_m,$$

where $q^{(t)}_m$ is the conditional probability of the outcome of the measurement $\Pi^{(t)}$ being m given that the given process is $\mathcal{E}_m$, which is expressed by

$$q^{(t)}_m \coloneqq \mathrm{Tr} \sum_{k_{t-1}=1}^{M} \cdots \sum_{k_1=1}^{M} \Gamma^{(t)}_{m|m,k_{t-1}} \circ \Gamma^{(t-1)}_{k_{t-1}|m,k_{t-2}} \circ \cdots \circ \Gamma^{(1)}_{k_1|m}.$$

The sets of $[\varrho^{(1)}, \Pi^{(1)}], [\varrho^{(2)}, \Pi^{(2)}], \ldots$ can be sequentially optimized. Specifically, for each t, we find $[\varrho^{(t)}, \Pi^{(t)}]$ that maximize $P^{(t)}$, which can be regarded as a single-shot channel discrimination problem and is formulated as an SDP problem. The success probability of our strategy is given by

$$\underline{P} \coloneqq P^{(T)}, \tag{7}$$

which is obviously a lower bound on the ultimate success probability. Since we need to optimize the T sets $[\varrho^{(1)}, \Pi^{(1)}], [\varrho^{(2)}, \Pi^{(2)}], \ldots,$ the computational complexity of obtaining $\underline{P}$ is roughly proportional to T.

We emphasize that since the proposed upper and lower bounds are based only on the concept of an operational probabilistic framework, it can be generalized to an arbitrary operational probabilistic theory. In such a theory, we need to solve some convex programming problems that are not SDP in general. However, these problems can be efficiently solved with existing techniques such as interior-point methods.

Numerical results.—First, we discuss a multishot channel discrimination problem. We here consider the problem of channel position finding [65] with two amplitude damping (AD) channels to compare our results with that in Ref. [33]. Let A_q be the AD channel with the damping parameter q, i.e., the qubit channel defined by

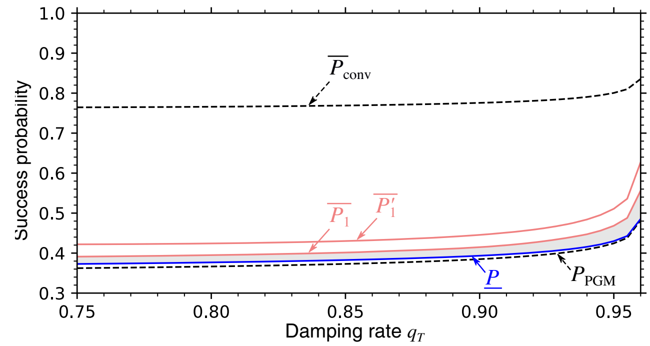


FIG. 3. Success probability in the problem of channel position finding with two AD channels A_{q_B} and A_{q_T}, where $T = 2$, $M = 3$, and $q_B = q_T + 0.04$. The ultimate success probability lies in the gray region, between our upper bound $\bar{P}_1$ of Eq. (3) and our lower bound $\underline{P}$ of Eq. (7). $\bar{P}'_1$ is the proposed upper bound described in the SM. $\bar{P}_{\rm conv}$ is the upper bound proposed in Ref. [33]. $P_{\rm PGM}$ is the success probability achieved by the maximally entangled pure state and the pretty good measurement [70,71], which is a lower bound on the ultimate success probability.

$$A_q(\rho) = E_0 \rho E_0^\dagger + E_1 \rho E_1^\dagger,$$
$$E_0 \coloneqq |0\rangle\langle 0| + \sqrt{1-q}|1\rangle\langle 1|, \qquad E_1 \coloneqq \sqrt{q}|0\rangle\langle 1| \tag{8}$$

with the standard basis $\{|0\rangle, |1\rangle\}$. Specifically, we consider T-shot discrimination of three channels, in which case the three channels are expressed in the form $A_{q_T} \otimes A_{q_B} \otimes A_{q_B}$, $A_{q_B} \otimes A_{q_T} \otimes A_{q_B}$, and $A_{q_B} \otimes A_{q_B} \otimes A_{q_T}$ with two damping parameters q_T and q_B. In Fig. 3, we show our numerical results. We computed our bounds $\bar{P}_1$ and $\underline{P}$ from Eqs. (3) and (7), respectively, where we solved the corresponding single-shot channel discrimination problems by the SDP solver CSDP (a C library for SDP) [66]. We also computed another proposed upper bound $\bar{P}'_1$ (detailed in Sec. II C of the SM [61]), which can be obtained at low computational cost. We can see that $\bar{P}_{\rm conv}$ is far from being optimal when the given channels are very close to each other. Indeed, $\bar{P}_{\rm conv} \geq (M+1)/2M$ always holds for any discrimination problem of M channels with equal prior probabilities [67], while the ultimate success probability $P^\star$ is close to $1/M$ when the given channels are nearly identical to each other. We can say that $\bar{P}_1$ is tighter than $\bar{P}_{\rm conv}$ in such a situation if T is not large enough [68]. Note that the proposed bound $\bar{P}_1$ becomes looser as T increases; in our preliminary numerical experiments, we observed that $\bar{P}_1$ is worse than $\bar{P}_{\rm conv}$ for large T (e.g., $T \geq 13$). As for the computational cost, computing $\bar{P}_1$ requires poly(4^M) time, whereas $\bar{P}_{\rm conv}$ requires $O(1)$ time (see Sec. VI of the SM [61]). The proposed method can be easily extended to obtain a slightly looser bound $\bar{P}'_1$ requiring $O(1)$ time. Computing $P_{\rm PGM}$ and $\underline{P}$ takes poly(4^{MT}) and $O(MT)$poly(4^M) times, respectively [69].

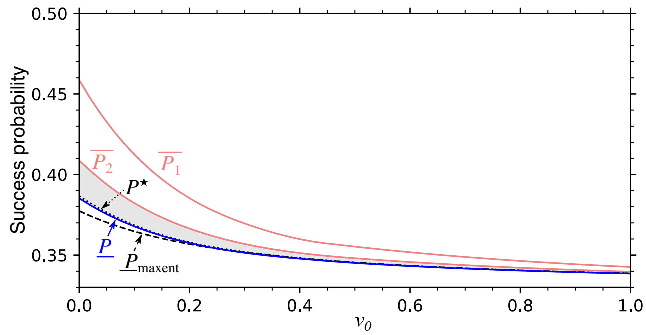


FIG. 4. Success probability in the discrimination between processes consisting of two generalized AD channels with correlated noise, where $T = M = 3$. The parameters that are associated with the zero-temperature dissipation rate are set to ν_0 and $\nu_0 + 0.04$, respectively. All the other parameters of these channels are the same. The two proposed upper bounds $\bar{P}_1$ of Eq. (3) and $\bar{P}_2$ of Eq. (5) and the proposed lower bound $\underline{P}$ of Eq. (7) are depicted. In this case, we can numerically compute the ultimate success probability $P^\star$. $\underline{P}_{\mathrm{maxent}}$ is the success probability achieved by the maximally entangled pure state and the optimal measurement, which gives a lower bound on $P^\star$.

Next, we discuss the problem of discriminating M processes where each process $\mathcal{E}_m$ consists of M memory channels each of which is the same channel, G_0, except the mth step, which is G_1. These processes are analogous to pulse-position modulated signals; We are here concerned with the case in which G_0 and G_1 are memory channels each of which is associated with two consecutive uses of generalized AD channel with correlated noise. Additional details including the exact definition of generalized AD channels are given in Sec. V of the SM [61]. Figure 4 shows the two upper bounds $\bar{P}_1$ of Eq. (3) and $\bar{P}_2$ of Eq. (5) and the lower bound $\underline{P}$. In this simulation, we set $M = 3$ to compute the exact value of the ultimate success probability $P^\star$ (note that $T = M$ holds in this problem). Since the cost of computing $P^\star$ increases exponentially with M, $P^\star$ is practically computable only for fairly small M (typically, $M \le 3$). We observe that $\underline{P}$ is very close to $P^\star$; the difference between them is less than 0.0015. $\bar{P}_1$, $\bar{P}_2$, and $\underline{P}$ have affordable computational costs; they require $O(1)$, $O(1)$, and $O(M^2)$ times, respectively.

Conclusions.—We presented upper and lower bounds on the ultimate success probability for discriminating arbitrary finite-dimensional quantum processes. In a special case of multishot channel discrimination, the ultimate success probability satisfies the relationship of Eq. (6). Our approach can be used to estimate the ultimate performances in various quantum information tasks, such as quantum sensing, quantum imaging, and quantum tomography.

We thank for O. Hirota and T. S. Usuda for comments and discussions. This work was supported by JSPS KAKENHI Grant No. JP19K03658.

[1] C. W. Helstrom, J. Stat. Phys. **1**, 231 (1969).
[2] A. S. Holevo, J. Multivariate Anal. **3**, 337 (1973).
[3] H. P. Yuen, K. S. Kennedy, and M. Lax, IEEE Trans. Inf. Theory **21**, 125 (1975).
[4] V. P. Belavkin, Stochastics **1**, 315 (1975).
[5] M. Hayashi, A. Kawachi, and H. Kobayashi, Quantum Inf. Comput. **8**, 345 (2008).
[6] A. Montanaro, in *2008 IEEE Information Theory Workshop* (IEEE, Porto, 2008), pp. 378–380, https://dx.doi.org/10.1109/ITW.2008.4578690.
[7] D. Qiu, Phys. Rev. A **77**, 012328 (2008).
[8] J. Tyson, Phys. Rev. A **79**, 032343 (2009).
[9] D. Qiu and L. Li, Phys. Rev. A **81**, 042329 (2010).
[10] Z. Ji, Y. Feng, R. Duan, and M. Ying, Phys. Rev. Lett. **96**, 200401 (2006).
[11] M. Ziman and T. Heinosaari, Phys. Rev. A **77**, 042321 (2008).
[12] M. Ziman, T. Heinosaari, and M. Sedlák, Phys. Rev. A **80**, 052102 (2009).
[13] M. Sedlák and M. Ziman, Phys. Rev. A **90**, 052312 (2014).
[14] Z. Puchała, Ł. Pawela, A. Krawiec, and R. Kukulski, Phys. Rev. A **98**, 042103 (2018).
[15] A. Krawiec, Ł. Pawela, and Z. Puchała, Quantum Inf. Process. **19**, 428 (2020).
[16] C. Datta, T. Biswas, D. Saha, and R. Augusiak, New J. Phys. **23**, 043021 (2021).
[17] A. Acin, Phys. Rev. Lett. **87**, 177901 (2001).
[18] M. F. Sacchi, Phys. Rev. A **71**, 062340 (2005).
[19] M. F. Sacchi, Phys. Rev. A **72**, 014305 (2005).
[20] L. Li and D. Qiu, J. Phys. A **41**, 335302 (2008).
[21] S. Pirandola and C. Lupo, Phys. Rev. Lett. **118**, 100502 (2017).
[22] S. Pirandola, R. Laurenza, C. Lupo, and J. L. Pereira, npj Quantum Inf. **5**, 50 (2019).
[23] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Europhys. Lett. **83**, 30004 (2008).
[24] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Phys. Rev. Lett. **101**, 060401 (2008).
[25] C. Macchiavello and G. M. Palma, Phys. Rev. A **65**, 050301(R) (2002).
[26] Y. Yeo and A. Skeen, Phys. Rev. A **67**, 064301 (2003).
[27] G. Bowen and S. Mancini, Phys. Rev. A **69**, 012306 (2004).
[28] D. Kretschmann and R. F. Werner, Phys. Rev. A **72**, 062323 (2005).
[29] M. B. Plenio and S. Virmani, Phys. Rev. Lett. **99**, 120504 (2007).
[30] S. Pirandola, Phys. Rev. Lett. **106**, 090504 (2011).
[31] G. M. D'Ariano, D. Kretschmann, D. Schlingemann, and R. F. Werner, Phys. Rev. A **76**, 032328 (2007).
[32] G. Gutoski and J. Watrous, in *Proceedings of the 39th Annual ACM Symposium on Theory of Computing* (Association for Computing Machinery, New York, 2007), pp. 565–574, https://dx.doi.org/10.1145/1250790.1250873.
[33] Q. Zhuang and S. Pirandola, Phys. Rev. Lett. **125**, 080505 (2020).
[34] S. Ishizaka and T. Hiroshima, Phys. Rev. Lett. **101**, 240501 (2008).
[35] S. Ishizaka and T. Hiroshima, Phys. Rev. A **79**, 042306 (2009).

[36] A. Jenčová and M. Plávala, J. Math. Phys. (N.Y.) **57**, 122203 (2016).
[37] M. Hayashi, IEEE Trans. Inf. Theory **55**, 3807 (2009).
[38] R. Duan, C. Guo, C.-K. Li, and Y. Li, in *Proceedings of the IEEE International Symposium on Information Theory (ISIT)* (IEEE, Barcelona, 2016), pp. 2259–2263, https://dx.doi.org/10.1109/ISIT.2016.7541701.
[39] S. Pirandola, B. R. Bardhan, T. Gehring, C. Weedbrook, and S. Lloyd, Nat. Photonics **12**, 724 (2018).
[40] Z. Puchała, Ł. Pawela, A. Krawiec, R. Kukulski, and M. Oszmaniec, Quantum **5**, 425 (2021).
[41] V. Katariya and M. M. Wilde, arXiv:2001.05376.
[42] A. W. Harrow, A. Hassidim, D. W. Leung, and J. Watrous, Phys. Rev. A **81**, 032339 (2010).
[43] R. S. Bondurant, Opt. Lett. **18**, 1896 (1993).
[44] A. Assalini, N. Dalla Pozza, and G. Pierobon, Phys. Rev. A **84**, 022342 (2011).
[45] F. Becerra, J. Fan, G. Baumgartner, J. Goldhar, J. Kosloski, and A. Migdall, Nat. Photonics **7**, 147 (2013).
[46] K. Flatt, S. M. Barnett, and S. Croke, Phys. Rev. A **100**, 032122 (2019).
[47] S. J. Dolinar, Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, 1976, Vol. 111.
[48] D. Brody and B. Meister, Phys. Rev. Lett. **76**, 1 (1996).
[49] A. Acin, E. Bagan, M. Baig, L. Masanes, and R. Muñoz-Tapia, Phys. Rev. A **71**, 032338 (2005).
[50] G. Ludwig, *An Axiomatic Basis of Quantum Mechanics* (Springer, New York, 1985 and 1987), Vols. I and II.
[51] *Foundations of Quantum Mechanics and Ordered Linear Spaces*, edited by A. Hartkämper and H. Neumann (Springer, New York, 1974), https://dx.doi.org/10.1007/3-540-06725-6.
[52] J. Barrett, Phys. Rev. A **75**, 032304 (2007).
[53] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Phys. Rev. A **81**, 062348 (2010).
[54] P. Janotta and R. Lal, Phys. Rev. A **87**, 052131 (2013).
[55] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Phys. Rev. Lett. **101**, 180501 (2008).
[56] G. Chiribella, New J. Phys. **14**, 125008 (2012).
[57] Let $(s_t^\star, X_t^\star)$ and $(s_{t-1}^\star, X_{t-1}^\star)$ be, respectively, the optimal solutions to Eq. (2) and that with t replaced by $t-1$; then, we can easily verify that $(s_{t-1}^\star s_t^\star, X_{t-1}^\star, X_t^\star)$ is a feasible solution to Eq. (4), which yields $s_{t-1}^\star s_t^\star \geq s_{t,2}^\star$. Thus, $\bar{P}_2 \leq \bar{P}_1$ holds. By the same discussion as $P \leq \bar{P}_1$, we obtain $P \leq \bar{P}_2$.
[58] M.-D. Choi, Linear Algebra Appl. **10**, 285 (1975).
[59] A. Jamiołkowski, Rep. Math. Phys. **3**, 275 (1972).
[60] K. Nakahira and K. Kato, arXiv:2104.09759.
[61] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.126.200502, which includes Ref. [62], for additional information about the proposed upper and lower bounds, numerical simulations, and computational complexity.
[62] Y. Jeong and H. Shin, Sci. Rep. **9**, 4035 (2019).
[63] S. Lloyd, Science **321**, 1463 (2008).
[64] S.-H. Tan, B. I. Erkmen, V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, S. Pirandola, and J. H. Shapiro, Phys. Rev. Lett. **101**, 253601 (2008).
[65] Q. Zhuang and S. Pirandola, Commun. Phys. **3**, 103 (2020).
[66] B. Borchers, Optim. Methods Software **11**, 613 (1999).
[67] From Eq. (7) of Ref. [33], we have $\bar{P}_{\rm conv} \geq 1-(M-1)/2M = (M+1)/2M$.
[68] Let $P_1^\star$ be the ultimate success probability in the case of $T=1$; then, $\bar{P}_1 = M^{T-1} P_1^{\star T}$ holds from $\bar{P}_1 = s^{\star T}/M$ and $P_1^\star = s^\star/M$, where $s^\star$ is the optimal value of problem (2). It follows from $\bar{P}_{\rm conv} \geq (M+1)/2M$ that $\bar{P}_1$ is tighter than $\bar{P}_{\rm conv}$ whenever $P_1^\star < [(M+1)/2]^{1/T}/M$.
[69] We did this numerical experiment on a PC with 16 GB memory, in which case neither $\bar{P}_1$ for $M \geq 4$ nor $P_{\rm PGM}$ for $TM \geq 9$ can be computed due to memory limitations.
[70] A. S. Holevo, Teor. Veroyatnost. i Primenen. **23**, 429 (1978) [Theory Probab. Appl. **23**, 411 (1978)].
[71] P. Hausladen and W. K. Wootters, J. Mod. Opt. **41**, 2385 (1994).