Efficient estimation of the quantum Chernoff bound

Weijing Li
College of Statistics and Mathematics, Hebei University of Economics and Business, Shijiazhuang, Hebei 050000, China

Wei Xie*
School of Computer Science and Technology and CAS Key Laboratory of Wireless-Optical Communications, University of Science and Technology of China, Hefei, Anhui 230000, China

(Received 9 February 2024; accepted 17 July 2024; published 9 August 2024)

The quantum Chernoff bound is a famous result about discriminating two different states in a setting where large numbers of copies are available, which gives the analytic asymptotic rate at which the minimal error probability decays to zero exponentially. It is however a challenging task to calculate the quantum Chernoff bound exactly in practical scenarios. In this paper, from the viewpoint of differential geometry, we demonstrate a remarkable link between the quantum Chernoff bound and Wigner-Yanase skew information. As a result, the quantum Chernoff bound can be estimated efficiently by virtue of the skew information. We present several examples to illustrate the efficiency of estimation about the quantum Chernoff bound via Wigner-Yanase skew information.

DOI: 10.1103/PhysRevA.110.022415

I. INTRODUCTION

Quantum state discrimination is one of the most fundamental information processing tasks in quantum technologies [1] and it often form the basis for an analysis of other types of quantum information processes [2–8]. It is well known that nonorthogonal states cannot be discriminated perfectly [9], and various strategies for optimum discrimination with respect to some appropriately chosen criteria have been developed [10–12]. In the setting of ambiguous discrimination with minimal error, the state can be determined by performing a quantum measurement. Each measurement outcome identifies one of the possible states and the overall error probability is to be minimized; the target is to find out the optimal quantum measurement as well as the minimal error probability. The task of finding this optimal measurement is so fundamental that it was one of the first problems considered in the field of quantum information theory [13]. Although closed-form solutions for minimum-error quantum state discrimination are known only for a few sets of states [14–18], for the binary case, it had already been solved by Helstrom [19] and Holevo [20].

When one consider the asymptotic scenario, where a large number of samples are available, the situation becomes more complicated. In the classical case, Chernoff proved his famous bound in a seminal paper [21], which states that the minimal probability of error $P_{e,n}^{\text{opt}}$ in discriminating two probability distributions decreases exponentially in the number of tests n that one can perform, that is,

$$P_{e,n}^{\text{opt}} \sim \exp(-n\xi_{\text{CB}}).$$

The optimal exponent ξ_{CB} arising in the asymptotic limit goes under several alternative names; for consistency in terminology we refer to it as Chernoff information [22]. For finite tests n this bound is a rather crude approximation. However, as n grows larger one finds better and better agreement, and the exponent ξ_{CB} becomes meaningful in the asymptotic limit.

The Chernoff information ξ_{CB} enjoys a compact expression. Specifically, for two discrete probability distributions p and q, this asymptotic exponent can be expressed by

$$\xi_{\text{CB}}(p,q) := -\ln\left(\inf_{0\leqslant\alpha\leqslant 1}\sum_i p(i)^{(1-\alpha)}q(i)^{\alpha}\right), \tag{1}$$

which is of closed form but for a single variable minimization [22]. The Chernoff information ξ_{CB} yields a very natural distance measure between probability distributions. It is essentially the unique distinguishability measure in the situation of independent and identically distributed random variables. Indeed, the distance measure for a particular pair of probability distributions gives a meaningful indication of how well these two distributions can be distinguished. This is especially meaningful when the applied strategy is the optimal one.

As the classical Chernoff bound has proved to be extremely useful in many branches of science, it is desirable to consider the quantum generalization of Chernoff's result [13]. Indeed, given the large amount of experimental effort in the context of quantum information processing to prepare and measure quantum states, it is of fundamental importance to have a theory that allows one to discriminate different quantum states in a meaningful way. After considerable effort [1,19,23], where the optimal Holevo-Helstrom strategy for discriminating between the two states is applied, the quantum generalization of the Chernoff bound was eventually settled by the combined work of [24,25].

*Contact author: xxieww@ustc.edu.cn

In particular, the quantum Chernoff bound [24] shows that the minimal error probability $P_{E,N}^{\rm opt}$ of discrimination of two quantum states decays exponentially with sample number N. In other words, one prepares N independent copies of a quantum system in an unknown state, which is either ρ_0 or ρ_1, and performs an optimal measurement to discriminate them. We assume that the quantum systems are finite, implying that the states are associated with density operators on a finite-dimensional complex Hilbert space. The combined N copies correspond to an N-fold tensor product density operator $\rho_i^{\otimes N}$, $i = 0, 1$. The minimal error probability $P_{E,N}^{\rm opt}$ of state discrimination satisfies

$$P_{E,N}^{\rm opt} \sim e^{-N\xi_{\rm QCB}(\rho_0,\rho_1)}, \tag{2}$$

where the asymptotic error exponent

$$\begin{aligned}\xi_{\rm QCB}(\rho_0, \rho_1) &= -\lim_{N\to\infty}\frac{1}{N}\ln P_{E,N}^{\rm opt}\\ &= -\ln\Big[\inf_{0\leqslant\alpha\leqslant1}\mathrm{Tr}\big(\rho_0^\alpha\rho_1^{1-\alpha}\big)\Big]\end{aligned} \tag{3}$$

is called the quantum Chernoff information. For simplicity, we write hereinafter ξ_Q for the quantum Chernoff information; ξ_Q is a natural distinguishability measure between quantum states because of its clear operational meaning and has several applications in quantum information theory [26–29]. Remarkably, the quantum Chernoff information (3) looks like an almost naive generalization of the classical expression (1).

While the quantum Chernoff bound provides a complete solution to the asymptotic setting, it is not completely satisfying from a practical point of view [30]. In fact, in realistic scenarios one has access only to finitely many copies of a system; the quantum Chernoff bound only provide some insightful bound for finite copies. Nevertheless, by virtue of the quantum Chernoff bound, one can assess the discrimination capabilities of different measurement strategies and evaluate the achievable error rates for a given set of states. This information is valuable for designing optimal discrimination protocols and understanding the fundamental limits of finite-size quantum state discrimination.

However, it is important to note that calculating the quantum Chernoff bound can be a challenging task, especially for large quantum systems. Various techniques, including semidefinite programming and convex optimization [31–34], have been developed to compute or approximate the quantum Chernoff bound in different settings, but when one deal with high-dimensional systems, such as many-body systems or systems with a large number of qubits, the calculation of the Chernoff bound becomes computationally demanding. The density matrices involved in the calculation can be high dimensional, requiring efficient numerical methods and computational resources.

To simplify the calculation, one may resort to approximations or bounds on the quantum Chernoff bound. It is the purpose of this paper to demonstrate. an approach to estimate the quantum Chernoff bound efficiently from the viewpoint of differential geometry.

Indeed, quantum Chernoff information induces a monotone Riemannian metric that gives a geometrical structure to the state space [35], which coincides with the Wigner-Yanase metric [36,37]. Since the Wigner-Yanase metric is a natural result of the Wigner-Yanase skew information [38], which is a celebrated quantity in quantum information theory [1,39–46], it is desirable to investigate the relationship between the quantum Chernoff bound and the Wigner-Yanase skew information.

In this paper we first disclose an intimate relationship between the quantum Chernoff information and Wigner-Yanase skew information from a geometrical viewpoint. By virtue of this observation, we show that the Wigner-Yanase skew information can be used to give a pretty good approximation of the quantum Chernoff information. We then present some illustrated examples to demonstrate this efficiency.

The rest of the paper is organized as follows. In Sec. II we present a complete mathematical formulation of the quantum Chernoff bound and its related properties. We demonstrate an intimate relationship between the quantum Chernoff information and Wigner-Yanase skew information from the viewpoint of differential geometry in Sec. III. By virtue of this key observation, in Sec. IV we present some illustrative examples to show that the skew information can be used to approximate the quantum Chernoff bound efficiently. In Sec. V we conclude the paper with a discussion.

II. FORMULATION OF THE QUANTUM CHERNOFF BOUND

In this section we review some basic results about the quantum Chernoff bound and its related properties. To begin with, we consider the simplest but nontrivial problem: two-state minimal-error discrimination. Recall that a state of a quantum system with finite-dimensional Hilbert space $\mathcal{H}$ is given by a density matrix ρ, that is, a non-negative operator on $\mathcal{H}$ with unit trace $\mathrm{Tr}(\rho) = 1$. We denote by $\mathcal{E}(\mathcal{H})$ the convex cone formed by all states on $\mathcal{H}$.

For two given quantum states $\rho_i \in \mathcal{E}(\mathcal{H})$, $i = 0, 1$, occurring with the *a priori* probabilities η_i, $i = 0, 1$, respectively, the discrimination problem consists in finding the optimum measurement strategy that minimizes the probability of errors. The measurement can be formally described by a positive-operator-valued measure (POVM) $\{M_j\}_{j=0}^1$, where $M_j \geqslant 0$, $j = 0, 1$. They are defined in such a way that $\mathrm{Tr}(\rho M_j)$ is the probability to infer the system is in the state ρ_j if it has been prepared in a state ρ. The overall probability P_E to make an erroneous guess for any of the incoming states is given by

$$\begin{aligned}P_E &= 1 - \sum_{j=0}^{1}\eta_j\mathrm{Tr}(\rho_j M_j)\\ &= \eta_0\mathrm{Tr}(\rho_0 M_1) + \eta_1\mathrm{Tr}(\rho_1 M_0),\end{aligned}$$

with $\eta_0 + \eta_1 = 1$. In order to find the minimum-error measurement strategy, one has to determine the value of P_E under the constraint that

$$M_0 + M_1 = I.$$

The minimum error probability $P_E^{\rm opt}$ can be formulated as

$$P_E^{\rm opt} = \min_{\{M_j\}} P_E,$$

where the minimum is over all possible POVMs $\{M_j\}_{j=0}^1$. By introducing the Hermitian operator

$$\Lambda = \eta_0\rho_0 - \eta_1\rho_1,$$

the error probability P_E can be alternatively expressed as

$$P_E = \eta_0 - \mathrm{Tr}(\Lambda M_0) = \eta_1 + \mathrm{Tr}(\Lambda M_1).$$

Our optimization task now consists in determining the specific operators M_0 or M_1, respectively, that minimize P_E.

Recall that the absolute value $|\Lambda|$ is defined as

$$|\Lambda| = \sqrt{\Lambda^\dagger\Lambda}$$

and the Jordan-decomposition of Λ is given by

$$\Lambda = \Lambda_+ - \Lambda_-,$$

where

$$\Lambda_+ = \frac{|\Lambda| + \Lambda}{2}, \quad \Lambda_- = \frac{|\Lambda| - \Lambda}{2}$$

are the positive part and negative part of Λ, respectively. Both Λ_+ and Λ_- are positive and of orthogonal support.

It follows that the minimal error probability P_E^{opt} can be achieved when M_1 is the projector Π on the support of Λ_+. Consequently, the P_E^{opt} is given by the Helstrom formula

$$P_E^{\mathrm{opt}}\big(\{\eta_i, \rho_i\}_{i=0}^1\big) = \tfrac{1}{2}(1 - \mathrm{Tr}|\Lambda|),$$

and the optimal measurement is a von Neumann measurement $\{\Pi, 1 - \Pi\}$ [19,20,47].

A single copy of the quantum system is not enough for a good decision; it is therefore desirable to consider a setting where large numbers of quantum states are available. In this case, the two quantum states are represented as N-fold tensor product states $\rho_0^{\otimes N}$ and $\rho_1^{\otimes N}$, where N is the sample number. The measurement to discriminate $\rho_0^{\otimes N}$ and $\rho_1^{\otimes N}$ is performed on composite systems; this fact enforces that the optimal measurement is a collective measurement. However, the particular permutational symmetry of N-copy states guarantees that the optimal collective measurement can be implemented efficiently (with a polynomial-size circuit) [48], and hence the minimum probability of error is achievable with a reasonable amount of resources. When the optimal strategy for discriminating the state is used, the corresponding minimal error probability $P_{E,N}^{\mathrm{opt}}$ can be rephrased as

$$P_{E,N}^{\mathrm{opt}}\big(\{\eta_i, \rho_i^{\otimes N}\}_{i=0}^1\big) = \tfrac{1}{2}(1 - \mathrm{Tr}|\Lambda_N|),$$

where

$$\Lambda_N = \eta_0\rho_0^{\otimes N} - \eta_1\rho_1^{\otimes N}.$$

One is typically interested in the limit of a large number of samples, i.e., $N \to \infty$. The quantum Chernoff bound [24,25] says that the minimal error probability $P_{E,N}^{\mathrm{opt}}$ decreases exponentially in N. As noted in the Introduction, when $N \to \infty$ it turns out that

$$P_{E,N}^{\mathrm{opt}} \sim e^{-N\xi_Q(\rho_0,\rho_1)},$$

with the optimal asymptotic rate exponent $\xi_Q(\rho_0, \rho_1)$ defined in Eq. (3).

The limit (3) can be identified by a lower and an upper estimate, that is,

$$\xi_Q \leqslant \liminf -\frac{1}{N}\ln P_{E,N}^{\mathrm{opt}} \tag{4}$$

and

$$\xi_Q \geqslant \limsup -\frac{1}{N}\ln P_{E,N}^{\mathrm{opt}}. \tag{5}$$

The proof of (4), first appearing in [25], is called the optimality part; it shows that the best discrimination is specified by the quantum Chernoff information. The very essential point in the proof is the following construction (6) of probability distributions p and q from density matrices ρ_0 and ρ_1. Let the spectral decompositions of ρ_0 and ρ_1 be given by

$$\rho_0 = \sum_i \lambda_i|x_i\rangle\langle x_i|, \quad \rho_1 = \sum_j \mu_j|y_j\rangle\langle y_j|.$$

We can define two probability distributions

$$p(i, j) = \lambda_i|\langle x_i|y_j\rangle|^2, \quad q(i, j) = \mu_j|\langle x_i|y_j\rangle|^2. \tag{6}$$

The limit (4) is obtained by the application of the classical Chernoff theorem (1) on probability distributions p and q.

The justification of (5) is the achievability part of the quantum Chernoff bound, which states that the error rate limit is actually equal to the quantum Chernoff information. To this end, we need the matrix inequality

$$\mathrm{Tr}(A^\alpha B^{1-\alpha}) \geqslant \mathrm{Tr}(A + B - |A - B|)/2 \tag{7}$$

for positive operators A and B and for all $0 \leqslant \alpha \leqslant 1$. The inequality (5) follows from the inequality (7) by taking $A = \eta_0\rho_0^{\otimes N}$ and $B = \eta_1\rho_1^{\otimes N}$.

The inequality (7), which first appeared in [24], not only plays a central role in the proof of the quantum Chernoff bound, but also has some application in the achievability of the quantum Hoeffding bound [13]. On its own it is very interesting from a purely matrix analysis point of view, as it relates the trace norm to a multiplicative quantity that is highly nontrivial. The original proof of (7) is not very transparent; a simple proof due to Ozawa was reported in [47,49].

The quantum Chernoff information $\xi_Q(\rho_0, \rho_1)$ defines a jointly convex function, which is contractive under quantum operations. More specifically, for any state ensembles $\{p_i, \rho_i\}$ and $\{p_i, \sigma_i\}$ with the same probabilities p_i, the joint convexity yields

$$\xi_Q\left(\sum_i p_i\rho_i, \sum_i p_i\sigma_i\right) \leqslant \sum_i p_i\xi_Q(\rho_i, \sigma_i).$$

The contractivity of ξ_Q means that for any pair of states ρ and σ, it holds that

$$\xi_Q(\rho, \sigma) \geqslant \xi_Q(\Phi(\rho), \Phi(\sigma))$$

for any quantum operation Φ.

In fact, one can see that the expression of ξ_Q is intimately related to the Petz-Rényi relative entropy [50]

$$D_\alpha(\rho, \sigma) = \frac{1}{\alpha - 1}\ln[\mathrm{Tr}(\rho^\alpha\sigma^{1-\alpha})], \quad \alpha \geqslant 0,\ \alpha \neq 1.$$

The joint convexity of ξ_Q results from the joint convexity of the relative entropy $D_\alpha(\rho, \sigma)$ for $\alpha \in (0, 1)$, which follows from the Lieb concavity theorem [51]. One then gets the contractivity of ξ_Q with respect to quantum operations from the joint convexity of ξ_Q [47].

The infimum in the definition of ξ_Q (3) is attained for a unique $\alpha \in (0,1)$ satisfying

$$\mathrm{Tr}\big[\rho_0^\alpha \rho_1^{1-\alpha}(\ln\rho_0 - \ln\rho_1)\big] = 0.$$

Actually, for any fixed ρ and σ, the function

$$Q_\alpha(\rho,\sigma) = \mathrm{Tr}(\rho^\alpha \sigma^{1-\alpha})$$

is convex on α, which is a simple consequence of the convexity of function

$$F(\alpha) = p^\alpha q^{1-\alpha}$$

for $p, q > 0$. The convexity means that the minimization has only one local minimum and therefore this local minimum is automatically the global minimum. By the Holder inequality, one has

$$Q_\alpha(\rho,\sigma) \leqslant Q_{0,1}(\rho,\sigma) = 1.$$

The convexity of Q_α provides an important benefit in actual calculations of ξ_Q.

III. RELATION BETWEEN THE QUANTUM CHERNOFF DISTANCE AND WIGNER-YANASE SKEW INFORMATION

However, it is important to note that calculating the quantum Chernoff information can be a challenging task, especially for large quantum systems. It is desirable to develop more efficient methods to obtain the precise calculation for the quantum Chernoff distance. The main contribution of this paper is that we demonstrate a remarkable relation between the quantum Chernoff bound and the Wigner-Yanase skew information from the viewpoint of differential geometry. Consequently, the quantum Chernoff information can be estimated efficiently by virtue of the skew information.

Note that an intriguing property of $\xi_Q(\rho_0,\rho_1)$ is that it induces a Riemannian metric [24,25], which is coincides with the Wigner-Yanase metric [36,52]. Recall that a Riemannian metric on $\mathcal{E}(\mathcal{H})$ is a map g which associates with each $\rho \in \mathcal{E}(\mathcal{H})$ a scalar product g_ρ on the tangent space $T_\rho\mathcal{E}(\mathcal{H})$ to $\mathcal{E}(\mathcal{H})$ at ρ. For any state ρ on $\mathcal{H}$, the tangent space $T_\rho\mathcal{E}(\mathcal{H})$ can be identified with the (real) vector space $\mathcal{B}(\mathcal{H})^0_{\mathrm{sa}}$ of self-adjoint operators on $\mathcal{H}$ with zero trace. A metric g defines a Riemannian distance d, which is such that the square distance

$$ds^2 = d(\rho, \rho + d\rho)^2$$

between two infinitesimally close states ρ and $\rho + d\rho$ is given by

$$ds^2 = g_\rho(d\rho, d\rho).$$

A curve γ in $\mathcal{E}(\mathcal{H})$ joining two states ρ_0 and ρ_1 is a (continuously differentiable) map

$$\gamma : t \in [0,1] \mapsto \rho(t) \in \mathcal{E}(\mathcal{H}),$$

with $\gamma(0) = \rho_0$ and $\gamma(1) = \rho_1$. Its length $L(\gamma)$ is defined as

$$L(\gamma) = \int_\gamma ds = \int_0^1 dt\sqrt{g_{\rho(t)}(\dot\rho(t), \dot\rho(t))},$$

where $\dot\rho(t)$ stands for the time derivative $d\rho/dt$.

Denoting by $C(\rho,\sigma)$ the set of all possible curves that joins two states ρ and σ,

$$C(\rho,\sigma) = \{\gamma \in C^1([0,1]), \mathcal{E}(\mathcal{H}) : \gamma(0) = \rho, \gamma(1) = \sigma\},$$

the map from a curve to its length

$$\gamma \in C(\rho,\sigma) \to L(\gamma)$$

has a stationary point. Consequently, the distance between two states ρ and σ can be defined as the length of the shortest geodesic joining these two states, that is,

$$d(\rho,\sigma) = \min_{\gamma\in C(\rho,\sigma)} L(\gamma).$$

Due to this formula, a distance d on $\mathcal{E}(\mathcal{H})$ can be associated with any metric g.

Conversely, one can associate a metric g with a distance d if the following condition is satisfied: For any $\rho \in \mathcal{E}(\rho)$ and $\dot\rho \in \mathcal{B}(\mathcal{H})^0_{\mathrm{sa}}$, the square distance between ρ and $\rho + \dot\rho t$ has a small-time Taylor expansion of the form

$$ds^2 = d(\rho, \rho + \dot\rho t)^2 = g_\rho(\dot\rho, \dot\rho)t^2 + O(t^3).$$

Needless to say, determining the metric induced by a given distance d is much simpler than finding an explicit formula for $d(\rho,\sigma)$ for arbitrary states $\rho, \sigma \in \mathcal{E}(\mathcal{H})$ from the expression of the metric g.

We then consider the Riemannian metric induced by the quantum Chernoff information $\xi_Q(\rho_0,\rho_1)$. For the fixed state ρ and its infinitesimally close state $\rho + d\rho$, it turns out that [24,36]

$$\xi_Q(\rho, \rho + d\rho) = (g_{\mathrm{WY}})_\rho(d\rho, d\rho),$$

where [47,52,53]

$$(g_{\mathrm{WY}})_\rho(A, A) = \sum_{j,k} \frac{|\langle j|A|k\rangle|^2}{2(\sqrt{\lambda_j} + \sqrt{\lambda_k})^2}. \tag{8}$$

Here $\rho = \sum_j \lambda_j |j\rangle\langle j|$ is the spectral decomposition of ρ and $A \in \mathcal{B}(\mathcal{H})^0_{\mathrm{sa}}$ is a tangent vector in the tangent space associated with ρ. In this paper we refer to the metric g_{WY} as the Wigner-Yanase metric, which will be clear after the following discussion. This relation builds a bridge between the quantum Chernoff bound and the Wigner-Yanase skew information.

Indeed, since any nontrivial element in the tangent space of the Riemannian manifold $\mathcal{E}(\mathcal{H})$ is of the form $A = i[\rho, K]$ [36], where K is a Hermitian operator related to ρ, for an arbitrary derivative $A = i[\rho, K]$ we have

$$(g_{\mathrm{WY}})_\rho(A, A) = (g_{\mathrm{WY}})_\rho(i[\rho, K], i[\rho, K]) = I(\rho, K), \tag{9}$$

where

$$I(\rho, K) = \mathrm{Tr}\rho K^2 - \mathrm{Tr}\sqrt{\rho}K\sqrt{\rho}K$$

is the celebrated Wigner-Yanase skew information [38] and g_{WY} is the appropriate name for the Wigner-Yanase metric.

Note that the very Hermitian operator K is also called the generator of translations and can be derived using the methods introduced in [54,55]. There the quantum estimation problem of spatial deformation was investigated by virtue of the quantum Fisher information as the figure of merit. This fact suggested that one can use the Wigner-Yanase skew information to study some particular quantum metrology information processing tasks.

Equation (9) is our starting point for establishing efficient estimation of the quantum Chernoff information. In particular, we consider the case of two states ρ and σ that can be connected by a smooth curve in the manifold $\mathcal{E}(\mathcal{H})$. More specifically, consider a smooth curve

$$\gamma : \theta \in [0, \tau] \to \rho(\theta) \in \mathcal{E}(\mathcal{H}), \quad \tau < 1.$$
$$\gamma(0) = \rho, \quad \gamma(\theta) = \rho_\theta, \quad \gamma(\tau) = \sigma.$$

For any ρ_θ, by the above discussion there must exist a Hermitian operator K_θ such that

$$d\rho_\theta = i[\rho_\theta, K_\theta]d\theta.$$

From Eq. (9) it must hold that

$$\begin{aligned}\xi_Q(\rho_\theta, \rho_\theta + d\rho_\theta) &= (g_{\mathrm{WY}})_{\rho_\theta} d(\rho_\theta, d\rho_\theta)\\ &= I(\rho_\theta, K_\theta)d\theta^2. \end{aligned} \quad (10)$$

One of the remarkable properties of the Wigner-Yanase skew information is that it remains invariant under unitary transformations, that is,

$$I(\rho, K) = I(U\rho U^\dagger, UKU^\dagger)$$

for any unitary operator U. It is natural to consider the case where two states are connected by a one-parameter unitary transformation, which is ubiquitous in quantum information theory. In this case, the curve can be written as

$$\gamma(\theta) = \rho_\theta = U_\theta \rho U_\theta^\dagger, \quad \theta \in [0, \tau],$$

where

$$U_\theta = \exp(-iH\theta), \quad (11)$$

As a result, the derivative reduces to

$$I(\rho_\theta, K_\theta) = I(\rho, H),$$

which shows that the skew information is independent of θ. Equation (10) can be written as

$$\xi_Q(\rho_\theta, \rho_\theta + d\rho_\theta) = I(\rho, H)d\theta^2.$$

Consequently, if ρ is close to σ, that is, $\tau \to 0$, it must hold that

$$\xi_Q(\rho, \sigma) = I(\rho, H)\tau^2. \quad (12)$$

Equation (12) is our first significant result in this work. In the situation that we want to discriminate ρ and σ, which are connected by (11), τ is determined by ρ and σ. The calculation of the quantum Chernoff information $\xi_Q(\rho, \sigma)$ can be safely replaced by determining the value of the Wigner-Yanase skew information $I(\rho, H)$. The latter is much easier than the former. If we are only interested in an estimation of $\xi_Q(\rho, \sigma)$, the above discussion shows that we can confine our attention to the estimation of the value of τ. The precision of the interval τ is related to another important issue in quantum information theory, namely, quantum metrology.

Additionally, if we only focus on the case where the quantum Chernoff information is upper bounded by a small number ε, that is,

$$\xi(\rho, \sigma) \leqslant \varepsilon,$$

then by Eq. (12) we have that

$$\tau \leqslant \sqrt{\frac{\varepsilon}{I(\rho, H)}}.$$

This result provides a pretty good bound for quantum parameter estimation.

Next we investigate an alternative method to approximate ξ_Q. Recall that the definition of the quantum Chernoff distance

$$\begin{aligned}\xi_Q(\rho, \sigma) &= -\inf_{\alpha\in(0,1)}\{\ln[\mathrm{Tr}(\rho^\alpha \sigma^{1-\alpha})]\}\\ &= \sup_{\alpha\in(0,1)}\{-\ln[\mathrm{Tr}(\rho^\alpha \sigma^{1-\alpha})]\}.\end{aligned}$$

To obtain an efficient estimation of ξ_Q, one can consider a special value $\alpha = \frac{1}{2}$ and then the following inequality holds:

$$\begin{aligned}\xi_Q(\rho, \sigma) &\geqslant -\ln \mathrm{Tr}(\sqrt{\rho}\sqrt{\sigma})\\ &= -\ln A(\rho, \sigma).\end{aligned}$$

Recall that $A(\rho, \sigma) = \mathrm{Tr}(\sqrt{\rho}\sqrt{\sigma})$ is exactly the quantum affinity [56], which has an intimate relationship with the skew information. It follows that

$$\ln[A(\rho, \sigma)] = \ln\{1 - [1 - A(\rho, \sigma)]\}.$$

By virtue of the basic inequality

$$\ln(1 + x) \leqslant x,$$

it must hold that

$$\ln[A(\rho, \sigma)] \leqslant -[1 - A(\rho, \sigma)]$$

and hence

$$-\ln A(\rho, \sigma) \geqslant 1 - A(\rho, \sigma)$$

We then arrive at

$$\xi_Q(\rho, \sigma) \geqslant 1 - A(\rho, \sigma).$$

In the situation where two states are connected by a unitary operator (11),

$$\sigma = \rho_\tau = U_\tau \rho U_\tau^\dagger, \quad (13)$$

by the Taylor expansion of quantum affinity [43],

$$A(\rho, \sigma) = 1 - I(\rho, H)\tau^2 + o(\tau^2),$$

that is,

$$1 - A(\rho, \sigma) = I(\rho, H)\tau^2 + o(\tau^2),$$

we obtain the inequality

$$\xi_Q(\rho, \sigma) \geqslant I(\rho, H)\tau^2 + o(\tau^2).$$

Therefore, the minimal error probability can be written as

$$P_{E,N}^{\mathrm{opt}} \sim e^{-N\xi_Q(\rho,\sigma)} \leqslant e^{-NI(\rho,H)\tau^2} e^{-No(\tau^2)}.$$

For fixed N, when

$$\tau \to 0, \quad e^{-No(\tau^2)} \to 1,$$

we can ignore the factor $e^{-No(\tau^2)}$ safely and we acquire an approximation of error probability

$$P_{E,N}^{\text{opt}} \sim e^{-N\xi_Q(\rho,\sigma)} \leqslant e^{-NI(\rho,H)\tau^2}. \tag{14}$$

Hence, for fixed ρ and σ, we can determine τ. If $I(\rho, H)$ is large, we then need only a small number of samples N and the minimal error probability can be controlled efficiently.

In summary, for a pair of states ρ and σ which can be connected via a unitary transformation (11), the length of the curve τ is fixed, the minimal error probability for discriminating $\rho^{\otimes N}$ and $\sigma^{\otimes N}$ can decrease exponentially with increasing N, and the rate of decay can be bounded by the skew information $I(\rho, H)$.

Equation (14) is our second main result in this paper. This observation can serve as a new operational interpretation for the Wigner-Yanase skew information in quantum information theory.

In addition to the above discussion, we can make a comparison between the skew information and geodesic distance. In the Riemannian metric space induced by the Wigner-Yanase metric, the geodesic distance between ρ and σ can be calculated as [36]

$$L(\rho, \sigma) = L = \arccos A(\rho, \sigma).$$

By taking the Taylor expansion of the cosine function, it must hold that

$$1 - A(\rho, \sigma) = \frac{L^2}{2} + o(L^2),$$

and hence

$$\xi(\rho, \sigma) \geqslant 1 - A(\rho, \sigma) \sim \frac{L^2}{2}.$$

Thus the minimal error probability can be written as

$$P_{E,N}^{\text{opt}} \sim e^{-N\xi_Q(\rho,\sigma)} \leqslant e^{-NL^2/2}.$$

Consider again the case where two quantum states are connected by (11). The length of the curve γ is

$$\int_0^\tau \sqrt{I(\rho, H)}d\theta = \sqrt{I(\rho, H)}\tau.$$

Since the geodesic distance is the minimal distance joining two points, we have

$$\frac{L^2}{2} \leqslant L^2 \leqslant I(\rho, H)\tau^2.$$

Therefore,

$$P_{E,N}^{\text{opt}} \sim e^{-N\xi_Q(\rho,\sigma)} \leqslant^{-NI(\rho,H)\tau^2} \leqslant e^{-NL^2/2}. \tag{15}$$

Equation (15) indicates that for the estimation of the quantum Chernoff information, the skew information is more efficient than the geodesic distance in this special case. This is our third important finding in this work.

IV. EXAMPLES

In this section we consider several illustrated examples to exhibit the links between quantum Chernoff information and the Wigner-Yanase skew information.

The first example is the simplest one. For $\lambda_0 + \lambda_1 = 1$ and $\lambda_j > 0$, consider the qubit state

$$\rho = \lambda_0|0\rangle\langle 0| + \lambda_1|1\rangle\langle 1|$$

and the Hermitian operator

$$H = X = |+\rangle\langle +| - |-\rangle\langle -|.$$

The unitary transformation reads

$$U_\theta = \exp(-iH\theta) = e^{-i\theta}|+\rangle\langle +| + e^{i\theta}|-\rangle\langle -|$$

and

$$\rho_\theta = U_\theta \rho U_\theta^\dagger, \quad \rho_\theta^\alpha = U_\theta \rho^\alpha U_\theta^\dagger.$$

The quantum Chernoff information turns out to be

$$\xi_Q(\rho, \rho_\theta) = -\inf_{0\leqslant\alpha\leqslant 1} \ln \text{Tr}\rho^{1-\alpha}\rho_\theta^\alpha.$$

Next we calculate the quantity

$$Q_\alpha = \text{Tr}\rho^{1-\alpha}\rho_\theta^\alpha.$$

Direct calculation shows that

$$Q_\alpha = \cos^2\theta + \sin^2\theta\big(\lambda_0^\alpha\lambda_1^{1-\alpha} + \lambda_0^{1-\alpha}\lambda_1^\alpha\big).$$

We define the function $f(\alpha)$ as

$$f(\alpha) = \lambda_0^\alpha\lambda_1^{1-\alpha} + \lambda_0^{1-\alpha}\lambda_1^\alpha$$

and then

$$f'(\alpha) = \big(\lambda_0^\alpha\lambda_1^{1-\alpha} - \lambda_0^{1-\alpha}\lambda_1^\alpha\big)(\ln\lambda_0 - \ln\lambda_1).$$

The case that $\lambda_0 = \lambda_1$ is trivial and we confine our attention to the case that $\lambda_0 \neq \lambda_1$. Then to find the extreme point we have

$$\lambda_0^\alpha\lambda_1^{1-\alpha} = \lambda_0^{1-\alpha}\lambda_1^\alpha,$$

which implies that $\alpha = \frac{1}{2}$. Hence

$$\xi_Q(\rho, \rho_\theta) = -\ln\text{Tr}(\sqrt{\rho}\sqrt{\rho_\theta}) \geqslant I(\rho, H)\theta^2 + o(\theta^2).$$

The Wigner-Yanase skew information in this case is

$$I(\rho, H) = 1 - 2\sqrt{\lambda_0\lambda_1}.$$

Therefore, the quantum Chernoff information can be approximately written as

$$\xi_Q \sim I(\rho, H)\theta^2$$

up to second order.

More generally, we consider the general single-qubit unitary operator

$$U_\theta = R_{\hat{n}}(\theta) = \exp(-iH\theta) = \cos(\theta)I - i\sin(\theta)H,$$

where

$$H = \hat{n}\cdot\vec{\sigma} = n_xX + n_yY + n_zZ$$

and

$$\hat{n} = (n_x, n_y, n_z), \quad n_x^2 + n_y^2 + n_z^2 = 1.$$

The quantity Q_α reads

$$Q_\alpha = \cos^2(\theta) + \sin^2(\theta)n_z^2 + f(\alpha)\sin^2(\theta)\big(n_x^2 + n_y^2\big).$$

The minimum of $f(\alpha)$, as well as Q_α, is also attained when $\alpha = \frac{1}{2}$.

We conclude that for the qubit case the quantum Chernoff information is attained by taking $\alpha = \frac{1}{2}$ and the skew information is applied. In this case, the skew information reduces to

$$I(\rho, H) = 1 - 2\sqrt{\lambda_0\lambda_1}\left(n_x^2 + n_y^2\right) - n_z^2$$

and

$$\xi_Q \sim I(\rho, H)\theta^2 \tag{16}$$

up to order 2. One can see that instead of calculating the quantum Chernoff information involving an optimization, the skew information (7) is much more convenient and efficient.

We then study the three-dimensional quantum states. For $\lambda_0 + \lambda_1 + \lambda_2 = 1$ and $\lambda_j > 0$, consider the state

$$\rho = \begin{pmatrix} \lambda_0 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}$$

and the Gell-Mann matrix

$$H = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

When $U_\theta = \exp(-iH\theta)$ is applied, we have

$$Q_\alpha = (\lambda_0 + \lambda_2)\cos^2(\theta) + g(\alpha)\sin^2(\theta) + \lambda_1,$$

where

$$g(\alpha) = \lambda_0^\alpha \lambda_2^{1-\alpha} + \lambda_0^{1-\alpha}\lambda_2^\alpha.$$

The optimum is attained also in $\alpha = \frac{1}{2}$.

However, things will not be perfect when we consider some linear combination of Gell-mann matrices, for example,

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix},$$

The quantum Chernoff information does not need to be attained when $\alpha = \frac{1}{2}$. However, the skew information can be computed easily as

$$I(\rho, H) = \frac{1}{2} + \frac{\sqrt{\lambda_2}}{2}(\sqrt{\lambda_2} - 2\sqrt{\lambda_0} - 2\sqrt{\lambda_1}).$$

This instance suggests that, to obtain an estimation of the quantum Chernoff information, instead of determining the exact value of α, one can consider the calculation of the Wigner-Yanase skew information.

V. DISCUSSION

The quantum Chernoff bound is a famous result in quantum state discrimination problem and it often is addressed in the setting of a symmetric quantum hypothesis test. The quantum Chernoff information, appearing as the asymptotic exponential rate at which the error probability tends to zero, was identified by virtue of an interesting matrix inequality. One of the intriguing properties of the quantum Chernoff information is that it induces a Riemannian metric, which coincides with the Wigner-Yanase metric. It is desirable to investigate the properties of the quantum Chernoff information from the geometric perspective. In this paper we first revealed an intimate relationship between the quantum Chernoff information and Wigner-Yanase skew information from the viewpoint of differential geometry. Specifically, when two quantum states are close to each other, the quantum Chernoff information can be replaced by the Wigner-Yanase skew information multiplying a factor. Although the quantum Chernoff bound is of essential action in the theory of quantum hypothesis testing, the difficulty of calculation discounts its applications in practice. By virtue of this link, the Wigner-Yanase skew information, however, gives a pretty good approximation of the quantum Chernoff information, which can be widely used in quantum state discrimination and quantum estimation theory. In the special case where two quantum states are connected by a unitary evolution, the skew information with respect to the Hamiltonian provides an efficient estimation of the quantum Chernoff information. We presented several illustrative examples to demonstrated our results. In particular, we showed that in the qubit case, the optimum quantum Chernoff information is attained when the parameter $\alpha = \frac{1}{2}$ and the skew information plays a prominent role in estimating the quantum Chernoff bound. As for the case where the optimum does not need to be approached by $\alpha = \frac{1}{2}$, the skew information also provides an approximation of the quantum Chernoff information.

Although the quantum Chernoff bound is an important quantity in the setting of quantum state discrimination, our results show that it can also play a role in the problem of quantum parameter estimation. The fact that the Wigner-Yanase skew information is a type of special quantum Fisher information [35] suggests that the existing tools developed for quantum estimation theory can be applied to the problem of quantum state discrimination [57]. We hope that our results may motivate further interest in the investigation of quantum information from the geometry standpoint.

ACKNOWLEDGMENTS

This work was partially supported by the National Natural Science Foundation of China (Grant No. 62102388) and Innovation Program for Quantum Science and Technology (Grant No. 2021ZD0302901).

[1] M. Hayashi, *Quantum Information Theory: Mathematical Foundation* (Springer, Berlin, 2016).

[2] S.-H. Tan, B. I. Erkmen, V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, S. Pirandola, and J. H. Shapiro, Quantum illumination with Gaussian states, Phys. Rev. Lett. **101**, 253601 (2008).

[3] S. Pirandola, Quantum reading of a classical digital memory, Phys. Rev. Lett. **106**, 090504 (2011).

[4] R. Nair and B. J. Yen, Optimal quantum states for image sensing in loss, Phys. Rev. Lett. **107**, 193602 (2011).

[5] S. Lloyd, V. Giovannetti, and L. Maccone, Sequential projective measurements for channel decoding, Phys. Rev. Lett. **106**, 250501 (2011).

[6] J. Bae and L.-C. Kwek, Quantum state discrimination and its applications, J. Phys. A: Math. Theor. **48**, 083001 (2015).

[7] M. A. Solís-Prosser, M. F. Fernandes, O. Jiménez, A. Delgado, and L. Neves, Experimental minimum-error quantum-state discrimination in high dimensions, Phys. Rev. Lett. **118**, 100501 (2017).

[8] J. S. Sidhu, M. S. Bullock, S. Guha, and C. Lupo, Linear optics and photodetection achieve near-optimal unambiguous coherent state discrimination, Quantum **7**, 1025 (2023).

[9] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2010).

[10] A. Chefles, Quantum state discrimination, Contemp. Phys. **41**, 401 (2000).

[11] S. M. Barnett and S. Croke, Quantum state discrimination, Adv. Opt. Photon. **1**, 238 (2009).

[12] J. A. Bergou, Discrimination of quantum states, J. Mod. Opt. **57**, 160 (2010).

[13] K. M. Audenaert, M. Nussbaum, A. Szkoła, and F. Verstraete, Asymptotic error rates in quantum hypothesis testing, Commun. Math. Phys. **279**, 251 (2008).

[14] 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).

[15] Y. C. Eldar and G. D. Forney, On quantum detection and the square-root measurement, IEEE Trans. Inf. Theory **47**, 858 (2001).

[16] C.-L. Chou and L.-Y. Hsu, Minimum-error discrimination between symmetric mixed quantum states, Phys. Rev. A **68**, 042305 (2003).

[17] Y. C. Eldar, A. Megretski, and G. C. Verghese, Optimal detection of symmetric mixed quantum states, IEEE Trans. Inf. Theory **50**, 1198 (2004).

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

[19] C. W. Helstrom, Quantum detection and estimation theory, J. Stat. Phys. **1**, 231 (1969).

[20] A. S. Kholevo, On asymptotically optimal hypotheses testing in quantum statistics, Teor. Veroyatn. Ee Primen. **23**, 429 (1978).

[21] H. Chernoff, A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations, Ann. Math. Stat. **23**, 493 (1952).

[22] T. M. Cover, *Elements of Information Theory* (Wiley, New York, 1999).

[23] T. Ogawa and M. Hayashi, On error exponents in quantum hypothesis testing, IEEE Trans. Inf. Theory **50**, 1368 (2004).

[24] K. M. R. Audenaert, J. Calsamiglia, R. Muñoz-Tapia, E. Bagan, L. Masanes, A. Acin, and F. Verstraete, Discriminating states: The quantum Chernoff bound, Phys. Rev. Lett. **98**, 160501 (2007).

[25] M. Nussbaum and A. Szkoła, The Chernoff lower bound for symmetric quantum hypothesis testing, Ann. Stat. **37**, 1040 (2009).

[26] D. F. Abasto, N. T. Jacobson, and P. Zanardi, Quantum Chernoff bound metric for the XY model at finite temperature, Phys. Rev. A **77**, 022327 (2008).

[27] M. Zwolak, C. J. Riedel, and W. H. Zurek, Amplification, redundancy, and quantum Chernoff information, Phys. Rev. Lett. **112**, 140406 (2014).

[28] J. Calsamiglia, R. Muñoz-Tapia, L. Masanes, A. Acin, and E. Bagan, Quantum Chernoff bound as a measure of distinguishability between density matrices: Application to qubit and Gaussian states, Phys. Rev. A **77**, 032311 (2008).

[29] M. Boca, I. Ghiu, P. Marian, and T. A. Marian, Quantum Chernoff bound as a measure of nonclassicality for one-mode Gaussian states, Phys. Rev. A **79**, 014302 (2009).

[30] K. M. R. Audenaert, M. Mosonyi, and F. Verstraete, Quantum state discrimination bounds for finite sample size, J. Math. Phys. **53**, 122205 (2012).

[31] Y. C. Eldar, A semidefinite programming approach to optimal unambiguous discrimination of quantum states, IEEE Trans. Inf. Theory **49**, 446 (2003).

[32] M. Ježek, J. Řeháček, and J. Fiurášek, Finding optimal strategies for minimum-error quantum-state discrimination, Phys. Rev. A **65**, 060301(R) (2002).

[33] P. Mironowicz, Semi-definite programming and quantum information, J. Phys. A: Math. Theor. **57**, 163002 (2024).

[34] P. Skrzypczyk and D. Cavalcanti, *Semidefinite Programming in Quantum Information Science* (IOP, Bristol, 2023).

[35] D. Petz, Monotone metrics on matrix spaces, Linear Algebra Appl. **244**, 81 (1996).

[36] P. Gibilisco and T. Isola, Wigner–Yanase information on quantum state space: The geometric approach, J. Math. Phys. **44**, 3752 (2003).

[37] W. Zhong, J. Ma, J. Liu, and X.-G. Wang, Derivation of quantum Chernoff metric with perturbation expansion method, Chin. Phys. B **23**, 090305 (2014).

[38] E. P. Wigner and M. M. Yanase, Information contents of distributions, Proc. Natl. Acad. Sci. USA **49**, 910 (1963).

[39] S. Luo, Wigner-Yanase skew information and uncertainty relations, Phys. Rev. Lett. **91**, 180403 (2003).

[40] S. Luo, Wigner-Yanase skew information vs. quantum Fisher information, Proc. Am. Math. Soc. **132**, 885 (2004).

[41] S. Luo and Q. Zhang, On skew information, IEEE Trans. Inf. Theory **50**, 1778 (2004).

[42] S. Luo, Quantum uncertainty of mixed states based on skew information, Phys. Rev. A **73**, 022324 (2006).

[43] S. Luo and Q. Zhang, Skew information decreases under quantum measurements, Theor. Math. Phys. **151**, 529 (2007).

[44] S. Luo, Notes on superadditivity of Wigner–Yanase–Dyson information, J. Stat. Phys. **128**, 1177 (2007).

[45] S. Luo and Y. Sun, Quantum coherence versus quantum uncertainty, Phys. Rev. A **96**, 022130 (2017).

[46] D. Petz, *Quantum Information Theory and Quantum Statistics* (Springer Science + Business Media, New York, 2007).

[47] D. Spehner, Quantum correlations and distinguishability of quantum states, J. Math. Phys. **55**, 075211 (2014).

[48] D. Bacon, I. L. Chuang, and A. W. Harrow, Efficient quantum circuits for Schur and Clebsch-Gordan transforms, Phys. Rev. Lett. **97**, 170502 (2006).

[49] V. Jakšić, Y. Ogata, C.-A. Pillet, and R. Seiringer, Quantum hypothesis testing and non-equilibrium statistical mechanics, Rev. Math. Phys. **24**, 1230002 (2012).

[50] D. Petz, Quasi-entropies for finite quantum systems, Rep. Math. Phys. **23**, 57 (1986).

[51] E. H. Lieb, Convex trace functions and the Wigner-Yanase-Dyson conjecture, Adv. Math. **11**, 267 (1973).

[52] F. Hansen, Metric adjusted skew information, Proc. Natl. Acad. Sci. USA **105**, 9909 (2008).

[53] D. Spehner, F. Illuminati, M. Orszag, and W. Roga, in *Lectures on General Quantum Correlations and their Applications*, edited by F. F. Fanchini, D. de Oliveira Soares Pinto, and G. Adesso (Springer, Cham, 2017), pp. 105–157.

[54] J. S. Sidhu and P. Kok, Quantum metrology of spatial deformation using arrays of classical and quantum light emitters, Phys. Rev. A **95**, 063829 (2017).

[55] J. S. Sidhu and P. Kok, Quantum Fisher information for general spatial deformations of quantum emitters, arXiv:1802.01601.

[56] S. Luo and Q. Zhang, Informational distance on quantum-state space, Phys. Rev. A **69**, 032106 (2004).

[57] J. S. Sidhu and P. Kok, Geometric perspective on quantum parameter estimation, AVS Quantum Sci. **2**, 014701 (2020).