

Optimal Quantum Measurements with respect to the Fidelity

Datong Chen^{1,2,3} and Huangjun Zhu^{1,2,3,*}

¹State Key Laboratory of Surface Physics, Department of Physics,
and Center for Field Theory and Particle Physics, Fudan University, Shanghai 200433, China

²Institute for Nanoelectronic Devices and Quantum Computing, Fudan University, Shanghai 200433, China

³Shanghai Research Center for Quantum Sciences, Shanghai 201315, China

(Dated: December 1, 2025)

Fidelity is the standard measure for quantifying the similarity between two quantum states. It is equal to the square of the minimum Bhattacharyya coefficient between the probability distributions induced by quantum measurements on the two states. Though established for over thirty years, the structure of fidelity-optimal quantum measurements remains unclear when the two density operators are singular (not invertible). Here we address this gap, with a focus on minimal optimal measurements, which admit no nontrivial coarse graining that is still optimal. We show that there exists either a unique minimal optimal measurement or infinitely many inequivalent choices. Moreover, the first case holds if and only if the two density operators satisfy a weak commutativity condition. In addition, we provide a complete characterization of all minimal optimal measurements when one state is pure, leveraging geometric insights from the Bloch-sphere representation. The connections with quantum incompatibility, operator pencils, and geometric means are highlighted.

Introduction—Nonorthogonal quantum states cannot be perfectly distinguished; this fundamental fact is tied to the no-cloning theorem and has profound implications for quantum information processing. Constructing optimal quantum measurements for distinguishing such states is instrumental to many practical applications, including quantum communication, quantum cryptography, quantum metrology, and quantum computation. Not surprisingly, the optimal measurements depend on the specific measure employed to quantify the distinguishability or similarity between quantum states. Fidelity is one of the most popular similarity measures because of its simplicity and appealing properties [1–5]. Notably, the fidelity between two pure states is equal to the transition probability, and the fidelity between two mixed states is equal to the maximum fidelity between their purifications thanks to Uhlmann and Jozsa [6, 7]. Moreover, the fidelity is equal to the square of the minimum Bhattacharyya coefficient [8, 9] between probability distributions induced by quantum measurements on the two states thanks to Fuchs and Caves [1, 10]. Although these facts have been known for more than three decades, the structure of optimal measurements is still quite elusive when the two density operators are singular.

In this work, we clarify the structure of optimal quantum measurements for distinguishing two general quantum states with respect to the fidelity. To this end, we introduce the concept of minimal optimal measurements as those optimal measurements that admit no nontrivial coarse graining preserving optimality. By virtue of connections with operator pencils [11, 12] and geometric means [13, 14], we demonstrate a sharp dichotomy: there exists either a unique minimal optimal measurement or infinitely many inequivalent choices. Notably, the first

case holds if and only if (iff) the two density operators satisfy a weak commutativity condition (assuming that their sum is nonsingular): the projectors onto their supports commute, but the density operators do not have to commute. In addition, we show that there exists a unique minimal optimal measurement that commutes with the projector onto the support of one density operator. Furthermore, we completely characterize all minimal optimal measurements when one quantum state is pure and provide geometric intuition via Bloch-sphere representation in the case of two pure states. In the course of study, we derive a number of results on operator pencils and geometric means, which are of independent interest.

Preliminaries—Let $\mathcal{H}$ be a d -dimensional Hilbert space, and denote by $\mathbb{1}$ the identity operator on $\mathcal{H}$. Given a linear operator A on $\mathcal{H}$, denote by $\text{Null}(A)$ the null space of A and by $\text{supp}(A)$ the support of A , that is, the orthogonal complement of $\text{Null}(A)$. Quantum states on $\mathcal{H}$ are usually described by density operators. Recall that a density operator ρ on $\mathcal{H}$ is a positive (semidefinite) operator of unit trace. Quantum measurements on $\mathcal{H}$ can be characterized by positive operator-valued measures (POVMs) [3, 15] since we are not interested in post-measurement quantum states in this work. Recall that a POVM $\mathcal{E} = \{E_m\}_m$ is composed of a collection of positive operators that sum up to the identity operator $\mathbb{1}$, that is, $E_m \geq 0$ for all m and $\sum_m E_m = \mathbb{1}$. A POVM is a projector-valued measure (PVM) if all its POVM elements are mutually orthogonal projectors. If we perform the POVM $\mathcal{E} = \{E_m\}_m$ on ρ , then the probability of obtaining outcome m reads $p_m = \text{tr}(E_m \rho)$. In this way, the POVM $\mathcal{E}$ induces a map from quantum states on $\mathcal{H}$ to probability distributions. The probability distribution associated with ρ and $\mathcal{E}$ is denoted by

$$\Lambda_{\mathcal{E}}(\rho) := \{\text{tr}(E_m \rho)\}_m. \quad (1)$$

Next, we introduce an order relation on POVMs from the perspective of data processing [16–18]. Given two

* zhuangjun@fudan.edu.cn

POVMs $\mathcal{A} = \{A_j\}_j$ and $\mathcal{B} = \{B_k\}_k$ on $\mathcal{H}$, the POVM $\mathcal{A}$ is a *coarse graining* of $\mathcal{B}$, if each POVM element of $\mathcal{A}$ can be expressed as follows:

$$A_j = \sum_k S_{jk} B_k, \quad (2)$$

where S is a stochastic matrix. Accordingly, $\mathcal{B}$ is a *refinement* of $\mathcal{A}$. Two POVMs are *equivalent* if they are coarse grainings of each other. A coarse graining or refinement of $\mathcal{A}$ is trivial if it is equivalent to $\mathcal{A}$ and non-trivial otherwise. A POVM is *simple* if no POVM element is proportional to another POVM element. Every POVM is equivalent to a simple POVM; in addition, two simple POVMs are equivalent iff they are identical up to relabeling (or permutation of POVM elements). In this paper, we are mainly interested in simple POVMs.

Next, $\mathcal{A}$ commutes with $\mathcal{B}$ if all POVM elements in $\mathcal{A}$ commute with all POVM elements in $\mathcal{B}$. Notably, any coarse graining or refinement of a PVM commutes with the PVM. By contrast, $\mathcal{A}$ and $\mathcal{B}$ are *compatible* or *jointly measurable* if they admit a common refinement and *incompatible* otherwise [18–20]. Commuting POVMs are compatible, but the converse is not guaranteed in general. Nevertheless, two PVMs are compatible iff they commute [20, 21].

Optimal measurements with respect to the fidelity—Let $\mathbf{p} = \{p_m\}_m$ and $\mathbf{q} = \{q_m\}_m$ be two probability distributions indexed by a same set. The classical fidelity (also known as the square of the Bhattacharyya coefficient) between $\mathbf{p}$ and $\mathbf{q}$ reads [8, 9]

$$F(\mathbf{p}, \mathbf{q}) := \left(\sum_m \sqrt{p_m q_m} \right)^2, \quad (3)$$

which is symmetric in $\mathbf{p}$ and $\mathbf{q}$ by definition. As a generalization, the fidelity between two density operators ρ and σ on $\mathcal{H}$ reads [6, 7]

$$F(\rho, \sigma) := \left(\text{tr} \sqrt{\sqrt{\rho} \sigma \sqrt{\rho}} \right)^2 = \left(\text{tr} \sqrt{\sqrt{\sigma} \rho \sqrt{\sigma}} \right)^2, \quad (4)$$

where the second equality holds because $\sqrt{\rho} \sqrt{\sigma}$ and $\sqrt{\sigma} \sqrt{\rho}$ share the same singular values, including multiplicities. So $F(\rho, \sigma)$ is symmetric in ρ and σ , which is analogous to the classical counterpart.

To understand the connection between the quantum and classical fidelities, given a POVM $\mathcal{E}$ on $\mathcal{H}$, denote by $F_{\mathcal{E}}(\rho, \sigma)$ the classical fidelity between $\Lambda_{\mathcal{E}}(\rho)$ and $\Lambda_{\mathcal{E}}(\sigma)$ as shown in Fig. 1. Then $F_{\mathcal{E}}(\rho, \sigma) \geq F(\rho, \sigma)$ [1, 3, 10], and this lower bound is tight. The POVM $\mathcal{E}$ is *F-optimal* for distinguishing ρ and σ if this bound is saturated, that is, $F_{\mathcal{E}}(\rho, \sigma) = F(\rho, \sigma)$. In addition, the induced classical fidelity is monotonic under data processing as shown in the following proposition, which is a corollary of Theorem 9 in Ref. [22]. See Appendices H and I for proofs of results presented in the main text.

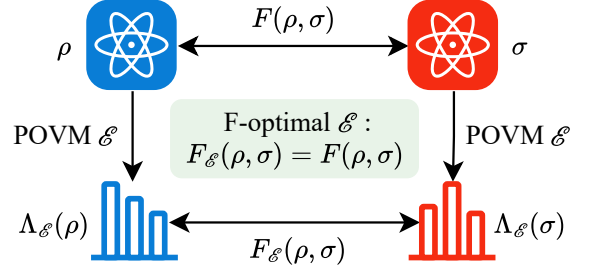


FIG. 1. Schematic diagram of the fidelity $F(\rho, \sigma)$ and the classical fidelity $F_{\mathcal{E}}(\rho, \sigma)$ induced by the POVM $\mathcal{E}$. The POVM $\mathcal{E}$ is F-optimal if $F_{\mathcal{E}}(\rho, \sigma) = F(\rho, \sigma)$.

Proposition 1. Suppose ρ, σ are two quantum states on $\mathcal{H}$ and $\mathcal{A}, \mathcal{B}$ are two POVMs on $\mathcal{H}$, where $\mathcal{A}$ is a coarse graining of $\mathcal{B}$. Then

$$F_{\mathcal{A}}(\rho, \sigma) \geq F_{\mathcal{B}}(\rho, \sigma). \quad (5)$$

As a simple corollary of Proposition 1, equivalent POVMs induce the same classical fidelity as expected. In addition, any refinement of an F-optimal POVM is still F-optimal. In view of this fact, an F-optimal POVM $\mathcal{E}$ is *minimal* if it is simple and any nontrivial coarse graining is not F-optimal; in other words, any coarse graining of $\mathcal{E}$ that preserves optimality is equivalent to $\mathcal{E}$. Such an optimal POVM extracts the least information and causes the least disturbance and is thus of special interest. In addition, the probability distributions induced by $\mathcal{E}$ only depend on its restriction on $\text{supp}(\rho + \sigma)$. In the rest of this paper, unless otherwise stated, we assume that ρ and σ are two distinct quantum states on $\mathcal{H}$, $\rho + \sigma$ is nonsingular, and $\mathcal{E} = \{E_m\}_m$ is a simple POVM on $\mathcal{H}$.

Insights from operator pencils—To clarify the conditions for constructing F-optimal POVMs, we need to introduce an additional concept. Two operators K and L on $\mathcal{H}$ are *parallel* if

$$L = 0 \quad \text{or} \quad K = \lambda L, \quad (6)$$

where λ is a nonnegative number. This condition is important to us because it arises naturally in saturating operator Cauchy-Schwarz inequalities tied to F-optimal POVMs. The following proposition offers a necessary and sufficient condition for a POVM to be F-optimal, which refines a result in Ref. [1].

Proposition 2. Suppose U is any unitary operator that satisfies $\sqrt{\rho} \sqrt{\sigma} U = \sqrt{\sqrt{\rho} \sigma \sqrt{\rho}}$. Then $\mathcal{E}$ is F-optimal for distinguishing ρ and σ iff $\sqrt{\sigma} \sqrt{E_m}$ and $U \sqrt{\rho} \sqrt{E_m}$ are parallel for each POVM element E_m in $\mathcal{E}$, that is,

$$\sqrt{\rho} \sqrt{E_m} = 0 \quad \text{or} \quad \sqrt{\sigma} \sqrt{E_m} = \kappa_m U \sqrt{\rho} \sqrt{E_m}, \quad (7)$$

where κ_m is a nonnegative number.

Equation (7) above may be regarded as a generalized eigenequation. To get a deeper understanding about this

condition, it is instructive to introduce the concept of linear operator pencils or pencils for short, which are useful in a number of research areas, including entanglement classification [23, 24] and quantum learning of many-body systems [25, 26]. Here a *pencil* refers to a pair (K, L) of linear operators on $\mathcal{H}$ [12] (see Appendix G for more details). A complex number λ is an eigenvalue of (K, L) if $K - \lambda L$ is singular. In that case, $\text{Null}(K - \lambda L)$ is called the eigenspace of (K, L) with eigenvalue λ , and any nonzero vector in this space is called an eigenvector. If $L = \mathbb{1}$, then the eigenvalues (eigenspaces) of (K, L) coincide with the counterparts of K . If instead L is singular, then ∞ is also regarded as an eigenvalue of (K, L) , and $\text{Null}(L)$ is regarded as the corresponding eigenspace. Using the language of operator pencils, we can provide precise characterizations of F-optimal and minimal F-optimal POVMs.

Proposition 3. *Suppose U is any unitary operator on $\mathcal{H}$ that satisfies $\sqrt{\rho}\sqrt{\sigma}U = \sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}$. Then $\mathcal{E}$ is F-optimal for distinguishing ρ and σ iff each POVM element E_m is supported in an eigenspace of $(\sqrt{\sigma}, U\sqrt{\rho})$ with a nonnegative eigenvalue (including ∞). If $\mathcal{E}$ is F-optimal, then $\mathcal{E}$ is minimal iff no two POVM elements are supported in the same eigenspace of $(\sqrt{\sigma}, U\sqrt{\rho})$.*

Suppose $\mathcal{E}$ is F-optimal for distinguishing ρ and σ . Then, as a simple corollary of Proposition 3, there exists a unique coarse graining of $\mathcal{E}$ up to relabeling that is F-optimal and minimal: this POVM is constructed by merging the POVM elements of $\mathcal{E}$ that are supported in each eigenspace, where merging is a special coarse graining that replaces two or more POVM elements with their sum.

Proposition 4. *Suppose $\mathcal{E}_1$ and $\mathcal{E}_2$ are two inequivalent minimal F-optimal POVMs for distinguishing ρ and σ . Then $\mathcal{E}_1$ and $\mathcal{E}_2$ are not compatible. Given $0 \leq p \leq 1$, let $\mathcal{E}(p)$ be the coarse graining of $p\mathcal{E}_1 \sqcup (1-p)\mathcal{E}_2$ that is F-optimal and minimal. Then $\mathcal{E}(p)$ for any two different values of p are not equivalent and thus not compatible.*

Notably, if the minimal F-optimal POVM for distinguishing ρ and σ is not unique, then there exist infinitely many (uncountable) minimal F-optimal POVMs, as illustrated in Fig. 2.

Insights from geometric means—Propositions 2 and 3 are instructive for understanding (minimal) F-optimal POVMs, but eigenspaces of an operator pencil are less tangible than the eigenspaces of a single operator. To simplify these criteria and to construct concrete minimal F-optimal POVMs, we need to introduce the concept of operator geometric means (see Appendix F), which are also helpful for studying quantum communication [27–29]. When ρ is nonsingular, the geometric mean of ρ^{-1} and σ is defined as follows [13, 14]:

$$\mathcal{M}(\rho^{-1}, \sigma) := \sqrt{\rho^{-1}} \sqrt{\sqrt{\rho}\sigma\sqrt{\rho}} \sqrt{\rho^{-1}}. \quad (8)$$

Here we do not assume that σ is nonsingular, although this is a common assumption in some literature. If ρ

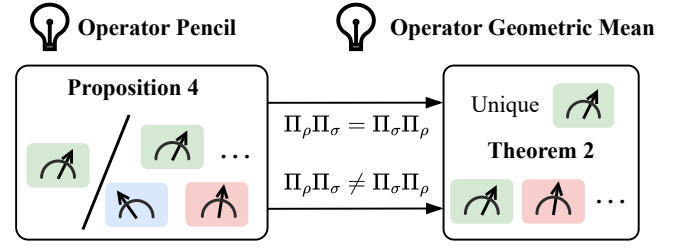


FIG. 2. Weak-commutativity condition and dichotomy between a unique minimal F-optimal POVM and infinitely many inequivalent choices as represented by different colors.

and σ commute, then $\mathcal{M}(\rho^{-1}, \sigma) = \sqrt{\rho^{-1}}\sqrt{\sigma}$ may be regarded as the quotient of $\sqrt{\sigma}$ over $\sqrt{\rho}$, and the unitary operator U in Proposition 3 can be chosen to be the identity. In this case, there is a simple connection between the eigenvalues (eigenspaces) of $\mathcal{M}(\rho^{-1}, \sigma)$ and the counterparts of $(\sqrt{\sigma}, U\sqrt{\rho})$, which intuitively explains why $\mathcal{M}(\rho^{-1}, \sigma)$ is relevant to studying F-optimal POVMs.

When ρ is singular, the inverse in Eq. (8) can be replaced by the Moore-Penrose generalized inverse [30, 31] of ρ , denoted by ρ^+ henceforth:

$$\mathcal{M}(\rho^+, \sigma) := \sqrt{\rho^+} \sqrt{\sqrt{\rho}\sigma\sqrt{\rho}} \sqrt{\rho^+}. \quad (9)$$

Nevertheless, little is known about this generalization in the literature. Our technical contribution on this topic as presented in Appendix F is crucial to understanding F-optimal POVMs. By definition we can deduce that $\mathcal{M}(\rho^+, \sigma)$ is a positive operator with

$$\text{supp}(\mathcal{M}(\rho^+, \sigma)) = \text{supp}(\Pi_\rho \Pi_\sigma \Pi_\rho) \leq \text{supp}(\rho), \quad (10)$$

where Π_ρ and Π_σ are the projectors onto $\text{supp}(\rho)$ and $\text{supp}(\sigma)$, respectively.

Let $\Pi_{\rho, \sigma}$ be the projector onto $\text{supp}(\Pi_\rho \Pi_\sigma \Pi_\rho)$. Define $\mathcal{M}(\rho, \sigma)$ as the PVM constructed from $\mathbb{1} - \Pi_{\rho, \sigma}$, $\Pi_\rho - \Pi_{\rho, \sigma}$, and the eigenprojectors of $\mathcal{M}(\rho^+, \sigma)$ with positive eigenvalues; here the element $\mathbb{1} - \Pi_{\rho, \sigma}$ ($\Pi_\rho - \Pi_{\rho, \sigma}$) will be omitted if its rank is 0, so that $\mathcal{M}(\rho, \sigma)$ is a simple PVM. Note that a POVM is a refinement of $\mathcal{M}(\rho, \sigma)$ iff each POVM element is supported in $\text{Null}(\rho)$ or an eigenspace of $\mathcal{M}(\rho^+, \sigma)$ within $\text{supp}(\rho)$. The geometric mean $\mathcal{M}(\sigma^+, \rho)$ and the simple PVM $\mathcal{M}(\sigma, \rho)$ can be defined in a similar way. As we will see shortly, both $\mathcal{M}(\rho, \sigma)$ and $\mathcal{M}(\sigma, \rho)$ are minimal F-optimal POVMs for distinguishing ρ and σ , which can also be deduced from Ref. [10]. To clarify the significance of the geometric mean $\mathcal{M}(\rho^+, \sigma)$ and the PVM $\mathcal{M}(\rho, \sigma)$, we need to introduce an auxiliary lemma.

Lemma 1. *Suppose P is a positive operator on $\mathcal{H}$ that commutes with Π_ρ , and U is any unitary operator that satisfies $\sqrt{\rho}\sqrt{\sigma}U = \sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}$. Then P is supported in an eigenspace of $(\sqrt{\sigma}, U\sqrt{\rho})$ iff P is supported in $\text{Null}(\rho)$ or an eigenspace of $\mathcal{M}(\rho^+, \sigma)$ within $\text{supp}(\rho)$.*

Note that any coarse graining or refinement of $\mathcal{M}(\rho, \sigma)$ commutes with Π_ρ . In conjunction with Proposition 3

and Lemma 1, we can immediately derive the following theorem, which implies that $\mathcal{M}(\rho, \sigma)$ [$\mathcal{M}(\sigma, \rho)$] is the unique minimal F-optimal POVM up to relabeling that commutes with Π_ρ (Π_σ).

Theorem 1. *The PVM $\mathcal{M}(\rho, \sigma)$ is F-optimal and minimal for distinguishing ρ and σ . Suppose $\mathcal{E}$ commutes with Π_ρ ; then $\mathcal{E}$ is F-optimal iff it is a refinement of $\mathcal{M}(\rho, \sigma)$. The same conclusions hold if ρ and σ are exchanged.*

Next, we present our main result on the connection between the uniqueness of the minimal F-optimal POVM and the weak commutativity condition. It turns out the same condition is also necessary and sufficient to guarantee that $\mathcal{M}(\rho, \sigma)$ and $\mathcal{M}(\sigma, \rho)$ are commuting, compatible, and equivalent.

Theorem 2. *The five statements below are equivalent:*

1. Π_ρ and Π_σ commute;
2. $\mathcal{M}(\rho, \sigma)$ and $\mathcal{M}(\sigma, \rho)$ commute;
3. $\mathcal{M}(\rho, \sigma)$ and $\mathcal{M}(\sigma, \rho)$ are compatible;
4. $\mathcal{M}(\rho, \sigma)$ and $\mathcal{M}(\sigma, \rho)$ are equivalent;
5. $\mathcal{M}(\rho, \sigma)$ is the unique minimal F-optimal POVM for distinguishing ρ and σ up to relabeling.

Theorem 2 unifies five statements on commutativity, compatibility, equivalence, and uniqueness, which link algebraic and operational properties (see also Lemma 7 in Appendix F). When Π_ρ and Π_σ commute, which means each principal angle between $\text{supp}(\rho)$ and $\text{supp}(\sigma)$ equals either 0 or $\pi/2$, all F-optimal POVMs for distinguishing ρ and σ are refinements of $\mathcal{M}(\rho, \sigma)$. Otherwise, by contrast, $\mathcal{M}(\rho, \sigma)$ and $\mathcal{M}(\sigma, \rho)$ are two inequivalent minimal F-optimal POVMs, and we can construct infinitely many inequivalent alternative choices by Proposition 4.

When ρ is nonsingular, $\mathcal{M}(\rho, \sigma)$ is composed of the eigenprojectors of $\mathcal{M}(\rho^{-1}, \sigma)$. By virtue of Theorem 1 (and also Theorem 2) we can immediately deduce the following corollary, which is also implicit in Ref. [10].

Corollary 1. *Suppose ρ is nonsingular. Then $\mathcal{E}$ is F-optimal for distinguishing ρ and σ iff each POVM element is supported in an eigenspace of $\mathcal{M}(\rho^{-1}, \sigma)$. It is F-optimal and minimal iff it is composed of the eigenprojectors of $\mathcal{M}(\rho^{-1}, \sigma)$.*

So far, by virtue of operator pencils and geometric means, we have clarified the basic structure of minimal F-optimal POVMs, as illustrated in Fig. 2. In the rest of this paper, we offer additional insights on a special case in which all minimal F-optimal POVMs can be characterized completely.

F-optimal measurements for distinguishing a pure state and a mixed state—Here we focus on the case in which one of the two states, say σ , is pure, assuming that $\rho + \sigma$ is nonsingular as before. Now, the minimal F-optimal PVM $\mathcal{M}(\sigma, \rho) = \{\mathbb{1} - \sigma, \sigma\}$ for distinguishing ρ and σ is binary,

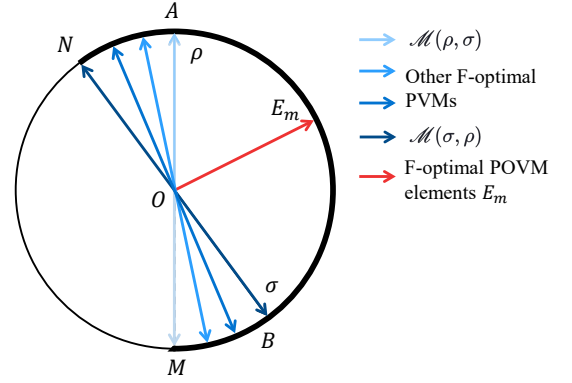


FIG. 3. F-optimal POVMs for distinguishing two pure states ρ and σ illustrated on the Bloch sphere. Here A and B denote the two pure states ρ and σ , while M and N denote their respective antipodal points. Each normalized POVM element of any F-optimal POVM is represented by a point on the major arc $\widehat{NABM}$.

while the minimal F-optimal PVM $\mathcal{M}(\rho, \sigma)$ is either binary or ternary. If σ commutes with Π_ρ , then $\mathcal{M}(\rho, \sigma)$ is equivalent to $\mathcal{M}(\sigma, \rho)$, and every F-optimal POVM is a refinement of $\mathcal{M}(\sigma, \rho)$ according to Theorem 2.

Next, we delve into the more interesting situation in which the pure state σ does not commute with Π_ρ . Then ρ is necessarily singular and $\dim(\text{Null}(\rho)) = 1$. In addition, $\mathcal{M}(\rho, \sigma)$ and $\mathcal{M}(\sigma, \rho)$ are not equivalent by Theorem 2, and there exist infinitely many inequivalent minimal F-optimal POVMs according to Proposition 4. Is there a simple characterization of all such POVMs? To resolve this problem, here we introduce a simpler necessary and sufficient criterion on F-optimal POVMs (cf. Proposition 3).

Proposition 5. *Suppose σ is a pure state and $\rho\sigma \neq 0$. Then $\mathcal{E}$ is F-optimal for distinguishing ρ and σ iff each POVM element is supported in an eigenspace of the pencil $(\Pi_\rho\sigma, \Pi_\rho)$ with a nonnegative eigenvalue.*

Surprisingly, F-optimal POVMs for ρ and σ are independent of ρ once Π_ρ and σ are fixed. Incidentally, the eigenspaces of $(\Pi_\rho\sigma, \Pi_\rho)$ with eigenvalues 0 and 1 coincide with $\text{Null}(\Pi_\rho\sigma) = \text{Null}(\sigma)$ and $\text{supp}(\sigma)$, respectively. The two eigenspaces are tied to the minimal F-optimal PVM $\mathcal{M}(\sigma, \rho) = \{\mathbb{1} - \sigma, \sigma\}$.

As an application of Proposition 5, first we analyze a simple but crucial case. Suppose ρ is also a pure state, then $\dim(\mathcal{H}) = 2$, and both $\mathcal{M}(\rho, \sigma)$ and $\mathcal{M}(\sigma, \rho)$ are rank-1 PVMs. In addition, each quantum state or normalized POVM element can be represented by a point on the Bloch sphere. Based on this picture, we can derive a more explicit criterion, which has a simple graphical representation on the Bloch sphere, as illustrated in Fig. 3.

Proposition 6. *Suppose $\dim \mathcal{H} = 2$ and ρ, σ are two pure states on $\mathcal{H}$ with noncommuting density operators represented by points A and B on the Bloch sphere; let*

M and N be their respective antipodal points. Then $\mathcal{E}$ is F -optimal for distinguishing ρ and σ iff each POVM element is proportional to a pure state whose representative point lies on the major arc $\widehat{NABM}$. An F -optimal POVM is minimal iff it is simple.

In the case of a qubit, if ρ and σ commute or if ρ or σ is nonsingular, then all simple F -optimal POVMs for ρ and σ are identical to $\mathcal{M}(\sigma, \rho)$ up to relabeling by Theorem 2. In conjunction with Proposition 6 we can completely characterize F -optimal PVMs and POVMs for distinguishing two general single-qubit states.

Next, suppose $\text{rank}(\rho) \geq 2$ and the pure state σ does not commute with Π_ρ , which means $\dim(\mathcal{H}) \geq 3$. As shown in Appendix B, we can establish a one-to-one correspondence between F -optimal POVMs for (distinguishing) ρ and σ and F -optimal POVMs for ϱ and σ , where $\varrho = \Pi_2 \Pi_\rho \Pi_2$ is a pure state and Π_2 is the projector onto $\text{Null}(\rho) + \text{supp}(\sigma)$. So this case is no more difficult than the case of two pure states.

Summary—We clarified the structure of optimal and minimal optimal quantum measurements for distinguishing two general quantum states with respect to the fidelity. When the projectors onto the two density operators commute, the minimal optimal measurement is uniquely determined by a geometric mean tied to the

two density operators, and all optimal measurements are refinements of this special projective measurement. Otherwise, there exist infinitely many inequivalent minimal optimal measurements determined by the eigenspaces of a special operator pencil. Surprisingly, this dichotomy is tied to such a weak commutativity condition rather than the commutativity condition between the two density operators themselves as one might expect a priori. In addition, we showed that there exists a unique minimal optimal measurement that commutes with the projector onto the support of one of the two density operators. We also provided a complete characterization of all minimal optimal measurements in the case where one quantum state is pure. Our work sheds valuable insights on one of the most important similarity measures in quantum information science and a number of related topics, including quantum measurements and quantum incompatibility. Our results are amenable to experimental demonstration with current technologies in various platforms; notably, it would be interesting to demonstrate inequivalent minimal optimal measurements in experiments. During our investigation, we derived a series of results on operator pencils and geometric means, which are of independent interest.

-
- [1] C. A. Fuchs, *Distinguishability and Accessible Information in Quantum Theory*, Ph.D. thesis, The University of New Mexico, Albuquerque (1996).
 - [2] C. A. Fuchs and J. van De Graaf, Cryptographic distinguishability measures for quantum-mechanical states, *IEEE Trans. Inf. Theor.* **45**, 1216 (1999).
 - [3] M. A. Nielsen and I. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, Cambridge, UK, 2010).
 - [4] M. M. Wilde, *Quantum Information Theory* (Cambridge University Press, Cambridge, UK, 2013).
 - [5] Y.-C. Liang, Y.-H. Yeh, P. E. M. F. Mendonça, R. Y. Teh, M. D. Reid, and P. D. Drummond, Quantum fidelity measures for mixed states, *Rep. Prog. Phys.* **82**, 076001 (2019).
 - [6] A. Uhlmann, The “transition probability” in the state space of a*-algebra, *Rep. Math. Phys.* **9**, 273 (1976).
 - [7] R. Jozsa, Fidelity for mixed quantum states, *J. Mod. Opt.* **41**, 2315 (1994).
 - [8] A. Bhattacharyya, On a measure of divergence between two statistical populations defined by their probability distributions, *Bull. Calcutta Math. Soc.* **35**, 99 (1943).
 - [9] W. K. Wootters, Statistical distance and Hilbert space, *Phys. Rev. D* **23**, 357 (1981).
 - [10] C. A. Fuchs and C. M. Caves, Mathematical techniques for quantum communication theory, *Open Syst. Inf. Dyn.* **3**, 345 (1995).
 - [11] K. D. Ikramov, Matrix pencils: Theory, applications, and numerical methods, *J. Math. Sci.* **64**, 783 (1993).
 - [12] G. H. Golub and C. F. Van Loan, *Matrix computations* (Johns Hopkins University Press, Baltimore, 2013).
 - [13] J. D. Lawson and Y. Lim, The geometric mean, matrices, metrics, and more, *Am. Math. Mon.* **108**, 797 (2001).
 - [14] R. Bhatia, *Positive Definite Matrices* (Princeton University Press, Princeton, 2007).
 - [15] J. von Neumann, *Mathematical Foundations of Quantum Mechanics: New Edition* (Princeton University Press, Princeton, 2018).
 - [16] H. Martens and W. M. de Muynck, Nonideal quantum measurements, *Found. Phys.* **20**, 255 (1990).
 - [17] Y. Kuramochi, Minimal sufficient positive-operator valued measure on a separable Hilbert space, *J. Math. Phys.* **56**, 102205 (2015).
 - [18] H. Zhu, Quantum Measurements in the Light of Quantum State Estimation, *PRX Quantum* **3**, 030306 (2022).
 - [19] P. Busch, Unsharp reality and joint measurements for spin observables, *Phys. Rev. D* **33**, 2253 (1986).
 - [20] O. Gühne, E. Haapasalo, T. Kraft, J.-P. Pellonpää, and R. Uola, Colloquium: Incompatible measurements in quantum information science, *Rev. Mod. Phys.* **95**, 011003 (2023).
 - [21] P. Lahti, Coexistence and joint measurability in quantum mechanics, *Int. J. Theor. Phys.* **42**, 893 (2003).
 - [22] T. van Erven and P. Harremoës, Rényi Divergence and Kullback-Leibler Divergence, *IEEE Trans. Inf. Theor.* **60**, 3797 (2014).
 - [23] E. Chitambar, C. A. Miller, and Y. Shi, Matrix pencils and entanglement classification, *J. Math. Phys.* **51**, 072205 (2010).
 - [24] O. Słowik, M. Hebenstreit, B. Kraus, and A. Sawicki, A link between symmetries of critical states and the structure of SLOCC classes in multipartite systems, *Quantum*

- 4, 300 (2020).
- [25] A. Steffens, C. A. Riofrío, R. Hübener, and J. Eisert, Quantum field tomography, *New J. Phys.* **16**, 123010 (2014).
 - [26] R. Kaneko, M. Imada, Y. Kabashima, and T. Ohtsuki, Forecasting long-time dynamics in quantum many-body systems by dynamic mode decomposition, *Phys. Rev. Res.* **7**, 013085 (2025).
 - [27] M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr, and M. Tomamichel, On quantum Rényi entropies: A new generalization and some properties, *J. Math. Phys.* **54**, 122203 (2013).
 - [28] K. Matsumoto, A new quantum version of f-divergence, in *Nagoya Winter Workshop: Reality and Measurement in Algebraic Quantum Theory* (Springer, 2015) pp. 229–273.
 - [29] K. Fang and H. Fawzi, Geometric Rényi divergence and its applications in quantum channel capacities, *Commun. Math. Phys.* **384**, 1615 (2021).
 - [30] J. C. A. Barata and M. S. Hussein, The Moore–Penrose pseudoinverse: A tutorial review of the theory, *Braz. J. Phys.* **42**, 146 (2012).
 - [31] A. Ben-Israel and T. N. Greville, *Generalized inverses: theory and applications* (Springer, New York, 2003).
 - [32] C. W. Helstrom, Quantum detection and estimation theory, *J. Stat. Phys.* **1**, 231 (1969).

CONTENTS

A. An example to support Theorem 2	6
B. F-optimal measurements for distinguishing a pure state and a mixed state	7
C. F-optimal measurements when $\rho + \sigma$ is singular	7
D. Optimal measurements with respect to the trace distance	8
1. Basic properties	8
2. Proofs of Propositions 8-10	9
E. Auxiliary lemmas	9
F. Properties of operator geometric means	10
1. Basic properties	10
2. Proofs of Lemmas 5-9	11
G. Properties of operator pencils	12
1. Basic properties	12
2. Proofs of Lemmas 12-16	13
H. Proofs of general results on F-optimal measurements	14
1. Proof of Proposition 1	14
2. Proof of Proposition 2	15
3. Proof of Proposition 3	15
4. Proof of Proposition 4	15
5. Proof of Lemma 1	15
6. Proof of Theorem 2	15

I. Proofs of Propositions 5 and 6

16

In this appendix, first we offer an illustrative example to support Theorem 2 and provide more details on F-optimal measurements for distinguishing a pure state σ and a mixed state ρ . Then we clarify F-optimal and minimal F-optimal measurements when $\rho + \sigma$ is singular. Next, we clarify minimal optimal measurements with respect to the trace distance for comparison. Subsequently, in preparation for proving our main results, we derive a number of auxiliary results on operator geometric means and operator pencils, which are of independent interest. Finally, we prove the results on F-optimal measurements as presented in the main text.

As in the main text, $\mathcal{H}$ denotes a d -dimensional Hilbert space, and $\mathbb{1}$ denotes the identity operator on $\mathcal{H}$. For any linear operator A on $\mathcal{H}$, $\text{Null}(A)$ and $\text{supp}(A)$ denote the null space and support of A , respectively. When A is a positive operator, Π_A denotes the projector onto the support of A . In addition, we denote by $\mathcal{P}(\mathcal{H})$ the cone of positive (semidefinite) operators on $\mathcal{H}$.

Appendix A: An example to support Theorem 2

When σ commutes with Π_ρ , it is not difficult to find a unitary operator U on $\mathcal{H}$ that commutes with Π_ρ and satisfies the condition $\sqrt{\rho}\sqrt{\sigma}U = \sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}$ (see Lemma 4 in Appendix E). Then it is straightforward to clarify the eigenvalues and eigenspaces of the pencil $(\sqrt{\sigma}, U\sqrt{\rho})$, from which Theorem 2 can be proved. When instead Π_σ commutes with Π_ρ , but σ does not commute with Π_ρ , some equivalence relations in Theorem 2 are highly non-trivial. Here we offer a concrete example to help build intuition.

Consider the following two qutrit states:

$$\begin{aligned}\rho &= \frac{1}{4}(|2\rangle\langle 0| + |0\rangle\langle 1| + |1\rangle\langle 0| + 2|1\rangle\langle 1|), \\ \sigma &= \frac{1}{4}(|2\rangle\langle 1| + |1\rangle\langle 2| + |2\rangle\langle 1| + 2|2\rangle\langle 2|),\end{aligned}\tag{A1}$$

with $\Pi_\rho = |0\rangle\langle 0| + |1\rangle\langle 1|$ and $\Pi_\sigma = |1\rangle\langle 1| + |2\rangle\langle 2|$. Note that Π_ρ and Π_σ commute with each other, and Statement 1 in Theorem 2 holds as desired. In addition, both $\mathcal{M}(\rho^+, \sigma)$ and $\mathcal{M}(\sigma^+, \rho)$ have an extremely simple form: $\mathcal{M}(\rho^+, \sigma) = \mathcal{M}(\sigma^+, \rho) = |1\rangle\langle 1|$, which means $\mathcal{M}(\rho, \sigma) = \mathcal{M}(\sigma, \rho) = \{|0\rangle\langle 0|, |1\rangle\langle 1|, |2\rangle\langle 2|\}$ up to relabeling. So the first four statements in Theorem 2 hold.

Next, we choose a specific unitary operator U that satisfies $\sqrt{\rho}\sqrt{\sigma}U = \sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}$. Let $a = \sqrt{\sqrt{3}+2}$ and $b = 3\sqrt{3}+5$, then a desired unitary can be expressed with respect to the computational basis as follows:

$$U = \frac{1}{4a^2} \begin{pmatrix} -2a^3 & 2a & 0 \\ a^2 & a^4 & -2a \\ 1 & a^2 & 2a^3 \end{pmatrix}.\tag{A2}$$

In addition, we have

$$U\sqrt{\rho} = \frac{1}{8a^2} \begin{pmatrix} -2\sqrt{6}a^2 & 0 & 0 \\ b & 2b & 0 \\ \sqrt{2}a & 2\sqrt{2}a & 0 \end{pmatrix}. \quad (\text{A3})$$

Now, it is straightforward to determine all eigenvalues of $(\sqrt{\sigma}, U\sqrt{\rho})$ and the corresponding normalized eigenvectors (up to overall phase factors), with the result

$$\begin{aligned} \lambda_0 &= 0, & |\lambda_0\rangle &= |0\rangle, \\ \lambda_1 &= 1, & |\lambda_1\rangle &= |1\rangle, \\ \lambda_2 &= \infty, & |\lambda_2\rangle &= |2\rangle. \end{aligned} \quad (\text{A4})$$

According to Proposition 3, every F-optimal POVM for distinguishing ρ and σ is equivalent to and also a refinement of $\mathcal{M}(\rho, \sigma)$. So $\mathcal{M}(\rho, \sigma)$ is the unique minimal F-optimal POVM for distinguishing ρ and σ up to relabeling, which confirms Theorem 2 for this specific case.

Appendix B: F-optimal measurements for distinguishing a pure state and a mixed state

Here we provide more details on F-optimal measurements for distinguishing a pure state σ and a mixed state ρ with $\text{rank}(\rho) \geq 2$, where σ does not commute with Π_ρ , which means $\dim(\mathcal{H}) \geq 3$. We will reduce this case to the case of two pure states clarified in the main text.

Let $\mathcal{V} = \text{Null}(\rho) + \text{supp}(\sigma)$ and Π_2 be the projector onto $\mathcal{V}$; then $\dim(\mathcal{V}) = 2$ and Π_2 commutes with Π_ρ . Let $\varrho = \Pi_2 \Pi_\rho \Pi_2$; then ϱ is a pure state supported in $\mathcal{V}$ and satisfies $\varrho\sigma = \Pi_\rho\sigma$. Based on this observation, we can construct F-optimal POVMs for distinguishing ρ and σ from the counterpart for ϱ and σ , and vice versa.

Specifically, if $\mathcal{E}$ is an F-optimal POVM for ρ and σ , then $\Pi_2 \mathcal{E} \Pi_2$ is an F-optimal POVM for ϱ and σ . If in addition $\mathcal{E}$ is minimal, then $\Pi_2 \mathcal{E} \Pi_2$ is minimal after deleting a zero POVM element if necessary. Conversely, if $\mathcal{E}'$ is a minimal F-optimal POVM for ϱ and σ , then $\mathcal{E}' \cup \{\mathbb{1} - \Pi_2\}$ is a minimal F-optimal POVM for distinguishing ρ and σ after merging POVM elements supported in $\text{Null}(\sigma)$ if necessary. In this way, we establish a one-to-one correspondence between minimal F-optimal POVMs for ρ and σ and the counterpart for ϱ and σ (and similarly for PVMs). Since F-optimal POVMs for distinguishing ϱ and σ are completely characterized by Proposition 6, the structure of F-optimal POVMs for ρ and σ is also clear. In contrast with the qubit case, now $\mathcal{M}(\sigma, \rho)$ is the unique binary F-optimal POVM up to relabeling; any other F-optimal POVM, including $\mathcal{M}(\rho, \sigma)$, contains at least three POVM elements.

Next, we clarify the reason underlying the above one-to-one correspondence. If we regard $(\varrho\sigma, \varrho)$ as a pencil on $\mathcal{V}$, then $(\varrho\sigma, \varrho)$ and $(\Pi_\rho\sigma, \Pi_\rho)$ have the same eigenvalues. Given a complex number λ (including ∞), let Π_λ be the eigenprojector of $(\Pi_\rho\sigma, \Pi_\rho)$ with eigenvalue λ and let Π'_λ be the counterpart of $(\varrho\sigma, \varrho)$. Then we have the following

simple correspondence:

$$\Pi_\lambda = \Pi'_\lambda + (\mathbb{1} - \Pi_2)\delta_{\lambda,0} = \begin{cases} \mathbb{1} - \sigma & \lambda = 0, \\ \Pi'_\lambda & \lambda \neq 0. \end{cases} \quad (\text{B1})$$

Notably, if $\lambda \neq 0$, then the eigenspace of $(\Pi_\rho\sigma, \Pi_\rho)$ with eigenvalue λ is a one-dimensional subspace of $\mathcal{V}$ and coincides with the eigenspace of $(\varrho\sigma, \varrho)$ with eigenvalue λ .

Appendix C: F-optimal measurements when $\rho + \sigma$ is singular

Now, ρ and σ can also be regarded as density operators on $\text{supp}(\rho + \sigma)$. Not surprisingly, a POVM on $\mathcal{H}$ is F-optimal for distinguishing ρ and σ iff its restriction on $\text{supp}(\rho + \sigma)$ is F-optimal, as shown in the following proposition.

Proposition 7. *Suppose ρ, σ are two quantum states on $\mathcal{H}$ and $\mathcal{E} = \{E_m\}_m$ is a POVM on $\mathcal{H}$. Let Π be the projector onto $\text{supp}(\rho + \sigma)$ and*

$$\mathcal{E}' = \Pi \mathcal{E} \Pi := \{\Pi E_m \Pi\}_m, \quad (\text{C1})$$

regarded as a POVM on the support of $\rho + \sigma$. Then $\mathcal{E}$ is F-optimal for distinguishing ρ and σ iff $\mathcal{E}'$ is F-optimal for distinguishing ρ and σ . In addition, if $\mathcal{E}$ is F-optimal and minimal, then $\mathcal{E}'$ is F-optimal and minimal.

When $\rho + \sigma$ is singular, Proposition 2 in the main text still holds. In addition, the two PVMs $\mathcal{M}(\rho, \sigma)$ and $\mathcal{M}(\sigma, \rho)$ are still F-optimal for distinguishing ρ and σ , but they cannot be equivalent. If $\text{supp}(\sigma) \leq \text{supp}(\rho)$, then $\mathcal{M}(\rho, \sigma)$ is not minimal; if $\text{supp}(\rho) \leq \text{supp}(\sigma)$, then $\mathcal{M}(\sigma, \rho)$ is not minimal. If neither $\text{supp}(\sigma) \leq \text{supp}(\rho)$ nor $\text{supp}(\rho) \leq \text{supp}(\sigma)$, then both $\mathcal{M}(\rho, \sigma)$ and $\mathcal{M}(\sigma, \rho)$ are F-optimal and minimal, but they are not equivalent.

Whenever $\rho + \sigma$ is singular, there exist infinitely many inequivalent minimal F-optimal POVMs for distinguishing ρ and σ . To be concrete, let $\mathcal{C} = \{C_j\}_j$ be a minimal F-optimal POVM on $\text{supp}(\rho + \sigma)$ and let Γ_0 be the projector onto $\text{Null}(\rho + \sigma)$. Then, for any probability distribution $\{a_j\}_j$ on $\mathcal{C}$, the POVM $\{C_j + a_j \Gamma_0\}_j$ on $\mathcal{H}$ is F-optimal and minimal; moreover, any two distinct POVMs in this family are inequivalent. On the other hand, all these POVMs lead to the same classical probability distribution when performed on ρ or σ , so they are essentially the same for distinguishing ρ and σ .

In the rest of this appendix, we prove Proposition 7.

Proof of Proposition 7. By assumption, it is straightforward to verify that $F_{\mathcal{E}'}(\rho, \sigma) = F_{\mathcal{E}}(\rho, \sigma)$. Therefore, $\mathcal{E}$ is F-optimal for distinguishing ρ and σ iff $\mathcal{E}'$ is F-optimal for distinguishing ρ and σ .

Next, suppose $\mathcal{E} = \{E_m\}_m$ is F-optimal and minimal and is thus automatically simple. Then $\mathcal{E}' = \{E'_m\}_m$ with $E'_m = \Pi E_m \Pi$ is F-optimal as shown above. In addition, $\mathcal{E}'$ is simple. To prove this point, suppose, by way of contradiction, that $\mathcal{E}'$ is not simple. Then there

exist two POVM elements in $\mathcal{E}'$, say E'_1 and E'_2 , that are linearly dependent. Let $\mathcal{C}$ be the POVM constructed from $\mathcal{E}$ by replacing the two POVM elements E_1 and E_2 with their sum $E_1 + E_2$. Then $\mathcal{C}$ is a nontrivial coarse graining of $\mathcal{E}$ that is also F-optimal, which contradicts the assumption that $\mathcal{E}$ is F-optimal and minimal. This contradiction shows that $\mathcal{E}'$ is simple.

Now, let $\mathcal{B}' = \{B'_j\}_j$ be a coarse graining of $\mathcal{E}'$ that is F-optimal for distinguishing ρ and σ . Then B'_j can be expressed as follows:

$$B'_j = \sum_k S_{jk} E'_k = \sum_k S_{jk} \Pi E_k \Pi = \Pi \sum_k S_{jk} E_k \Pi, \quad (\text{C2})$$

where S is a stochastic matrix. Construct the POVM $\mathcal{B} = \{B_j\}_j$ with $B_j = \sum_k S_{jk} E_k$; then $\mathcal{B}$ is a coarse graining of $\mathcal{E}$. By definition, we have

$$F_{\mathcal{B}}(\rho, \sigma) = F_{\mathcal{B}'}(\rho, \sigma) = F(\rho, \sigma), \quad (\text{C3})$$

so the POVM $\mathcal{B}$ is also F-optimal for distinguishing ρ and σ . Since $\mathcal{E}$ is F-optimal and minimal by assumption, $\mathcal{B}$ is necessarily equivalent to $\mathcal{E}$, which means $\mathcal{B}'$ is equivalent to $\mathcal{E}'$. Therefore, any nontrivial coarse graining of $\mathcal{E}'$ is not F-optimal. It follows that $\mathcal{E}'$ is F-optimal and minimal given that $\mathcal{E}'$ is F-optimal and simple as shown above. This observation completes the proof of Proposition 7. $\square$

Appendix D: Optimal measurements with respect to the trace distance

The trace distance is a popular measure for quantifying the distinction between two quantum states [1–3]. To complement the results presented in the main text, here we briefly discuss optimal and minimal optimal measurements with respect to the trace distance.

1. Basic properties

Let $\mathbf{p} \equiv \{p_m\}_m$ and $\mathbf{q} \equiv \{q_m\}_m$ be two probability distributions indexed by a same set. The (classical) trace distance between $\mathbf{p}$ and $\mathbf{q}$ reads

$$D(\mathbf{p}, \mathbf{q}) = \frac{1}{2} \sum_m |p_m - q_m| = \max_{\mathbb{S}} \sum_{m \in \mathbb{S}} (p_m - q_m), \quad (\text{D1})$$

where the maximization is taken over all subsets of the index set. Suppose the two probability distributions are chosen with probability 1/2 each, then the maximum probability of successfully distinguishing between the two distributions is given by

$$P_s = \frac{1 + D(\mathbf{p}, \mathbf{q})}{2}. \quad (\text{D2})$$

A larger trace distance corresponds to a higher success probability, highlighting the operational significance of the trace distance.

Next, we turn to the quantum analog. Suppose ρ, σ are two quantum states on $\mathcal{H}$ and $\mathcal{A}, \mathcal{B}, \mathcal{E} = \{E_m\}_m$ are three POVMs on $\mathcal{H}$. The trace distance between ρ and σ reads [3]

$$D(\rho, \sigma) = \frac{1}{2} \text{tr}|\rho - \sigma|. \quad (\text{D3})$$

The trace distance and fidelity are closely tied to each other and satisfy the well-known Fuchs-van de Graaf inequalities [2, 3]:

$$1 - \sqrt{F(\rho, \sigma)} \leq D(\rho, \sigma) \leq \sqrt{1 - F(\rho, \sigma)}. \quad (\text{D4})$$

To understand the connection with the classical counterpart, denote by $D_{\mathcal{E}}(\rho, \sigma)$ the classical trace distance between the probability distributions $\Lambda_{\mathcal{E}}(\rho)$ and $\Lambda_{\mathcal{E}}(\sigma)$ induced by the POVM $\mathcal{E}$. Then $D_{\mathcal{E}}(\rho, \sigma) \leq D(\rho, \sigma)$ according to Helstrom-Holevo theorem, and the inequality is tight [2, 3, 32]. So the trace distance $D(\rho, \sigma)$ determines how well the two states ρ and σ can be distinguished. The POVM $\mathcal{E}$ is *T-optimal* for distinguishing ρ and σ if the inequality is saturated, that is, $D_{\mathcal{E}}(\rho, \sigma) = D(\rho, \sigma)$. Similar to the fidelity, the induced classical trace distance is monotonic under data processing.

Proposition 8. Suppose $\mathcal{A}$ is a coarse graining of $\mathcal{B}$. Then

$$D_{\mathcal{A}}(\rho, \sigma) \leq D_{\mathcal{B}}(\rho, \sigma). \quad (\text{D5})$$

A T-optimal POVM is *minimal* if it is simple and any nontrivial coarse graining is not T-optimal.

To characterize T-optimal POVMs, note that $\rho - \sigma$ can be expressed as

$$\rho - \sigma = Q_+ - Q_-, \quad (\text{D6})$$

where Q_+ and Q_- are two positive operators that are mutually orthogonal, that is, $Q_+ Q_- = 0$, and are called the positive part and negative part of $\rho - \sigma$, respectively. Let Π_+ and Π_- be the projectors onto the supports of Q_+ and Q_- , respectively, and let Π_0 be the projector onto the null space of $\rho - \sigma$. The following proposition clarifies T-optimal POVMs for distinguishing ρ and σ [3].

Proposition 9. The POVM $\mathcal{E}$ is T-optimal for distinguishing ρ and σ iff each POVM element E_m is supported in $\text{Null}(Q_+)$ or $\text{Null}(Q_-)$, that is,

$$Q_+ E_m = 0 \quad \text{or} \quad Q_- E_m = 0 \quad \forall m, \quad (\text{D7})$$

where Q_+ and Q_- denote the positive part and negative part of $\rho - \sigma$, respectively.

When ρ and σ are distinct single-qubit states, both Π_+ and Π_- are rank-1 eigenprojectors of $\rho - \sigma$, and Q_+ and Q_- are proportional to Π_+ and Π_- , respectively. Therefore, all simple T-optimal POVMs for distinguishing ρ and σ are equivalent to the PVM $\{\Pi_+, \Pi_-\}$.

Proposition 10. *The POVM $\mathcal{E}$ is T-optimal and minimal for distinguishing ρ and σ iff it has two POVM elements and has the form $\{\Pi_+ + Q_0, \mathbb{1} - \Pi_+ - Q_0\}$ up to relabeling, where Π_+ and Π_- are the projectors onto the supports of the positive part and negative part of $\rho - \sigma$, respectively, and $0 \leq Q_0 \leq \mathbb{1}$ is supported in $\text{Null}(\rho - \sigma)$.*

According to Proposition 10, if $\rho - \sigma$ is nonsingular, then $\{\Pi_+, \Pi_-\}$ is the unique minimal T-optimal POVM up to relabeling for distinguishing ρ and σ ; otherwise, there exist infinitely many inequivalent minimal T-optimal POVMs. For instance, the POVM

$$\mathcal{A} = \{\Pi_+ + \lambda\Pi_0, \Pi_- + (1 - \lambda)\Pi_0\} \quad (\text{D8})$$

is T-optimal and minimal for each λ in the interval $[0, 1]$. Moreover, any two distinct POVMs in this family are inequivalent. However, these POVMs commute and are compatible, in sharp contrast with the results on minimal F-optimal POVMs (cf. Proposition 4 and Theorem 2).

2. Proofs of Propositions 8-10

Proof of Proposition 8. Suppose $\mathcal{A} = \{A_j\}_j$ and $\mathcal{B} = \{B_k\}_k$. Denote the probability vectors $\Lambda_{\mathcal{A}}(\rho)$, $\Lambda_{\mathcal{A}}(\sigma)$, $\Lambda_{\mathcal{B}}(\rho)$, $\Lambda_{\mathcal{B}}(\sigma)$ by $\mathbf{p}, \mathbf{q}, \mathbf{p}', \mathbf{q}'$, respectively. By assumption, $\mathcal{A}$ is a coarse graining of $\mathcal{B}$, so each POVM element A_j of $\mathcal{A}$ can be expressed as $A_j = \sum_k S_{jk} B_k$, where S is a stochastic matrix. Accordingly, we have

$$p_j = \sum_k S_{jk} p'_k, \quad q_j = \sum_k S_{jk} q'_k. \quad (\text{D9})$$

Therefore,

$$\begin{aligned} D_{\mathcal{A}}(\rho, \sigma) &= \frac{1}{2} \sum_j |p_j - q_j| = \frac{1}{2} \sum_j \left| \sum_k S_{jk} (p'_k - q'_k) \right| \\ &\leq \frac{1}{2} \sum_{j,k} S_{jk} |p'_k - q'_k| = D_{\mathcal{B}}(\rho, \sigma), \end{aligned} \quad (\text{D10})$$

where the inequality follows from the triangle inequality and the last equality holds because S is a stochastic matrix, which means $\sum_j S_{jk} = 1$ for all k . The above equation confirms Eq. (D5) and completes the proof of Proposition 8. $\square$

Proof of Proposition 9. By definition, we can deduce that

$$\begin{aligned} D_{\mathcal{E}}(\rho, \sigma) &= \frac{1}{2} \sum_m |\text{tr}[E_m(\rho - \sigma)]| \\ &= \frac{1}{2} \sum_m |\text{tr}[E_m(Q_+ - Q_-)]| \leq \frac{1}{2} \sum_m \text{tr}[E_m(Q_+ + Q_-)] \\ &= \frac{1}{2} \sum_m \text{tr}(E_m |\rho - \sigma|) = \frac{1}{2} \text{tr} |\rho - \sigma| = D(\rho, \sigma), \end{aligned} \quad (\text{D11})$$

where the first equality holds because $\rho - \sigma = Q_+ - Q_-$. The inequality follows from the triangle inequality and is saturated iff $\text{tr}(Q_+ E_m) = 0$ or $\text{tr}(Q_- E_m) = 0$ for each m , which is the case iff Eq. (D7) holds, that is, each POVM element E_m is supported in $\text{Null}(Q_+)$ or $\text{Null}(Q_-)$. This observation completes the proof of Proposition 9. $\square$

Proof of Proposition 10. First, by way of contradiction, we will prove that the number of POVM elements in any minimal T-optimal POVM cannot surpass 2. Suppose $\mathcal{E}$ is a minimal T-optimal POVM for distinguishing ρ and σ and has at least three POVM elements. According to Proposition 9, each POVM element in $\mathcal{E}$ is supported in either $\text{Null}(Q_+)$ or $\text{Null}(Q_-)$, and there exist at least two elements that are supported in $\text{Null}(Q_+)$ simultaneously or in $\text{Null}(Q_-)$ simultaneously. Therefore, we can construct a nontrivial coarse graining of $\mathcal{E}$ by merging these POVM elements. The resulting POVM is still T-optimal, which contradicts the assumption that $\mathcal{E}$ is T-optimal and minimal. This contradiction shows that any minimal T-optimal POVM has only two POVM elements.

Next, suppose $\mathcal{E}$ is a binary POVM that is T-optimal and minimal. Then, according to Proposition 9, one POVM element is supported in $\text{Null}(Q_-)$, and the other is supported in $\text{Null}(Q_+)$. So $\mathcal{E}$ necessarily has the form $\mathcal{E} = \{\Pi_+ + Q_0, \mathbb{1} - \Pi_+ - Q_0\}$ up to relabeling, where Q_+ and Q_- denote the positive part and negative part of $\rho - \sigma$, respectively, and Q_0 is a positive operator supported in $\text{Null}(\rho - \sigma)$ and satisfies $0 \leq Q_0 \leq \mathbb{1}$. This observation completes the proof of Proposition 10. $\square$

Appendix E: Auxiliary lemmas

Lemma 2. *Suppose P and Q are two projectors on $\mathcal{H}$, and Π_{PQP} and Π_{QPQ} are projectors onto PQP and QPQ , respectively. Then the following four statements are equivalent.*

1. P and Q commute;
2. $\text{supp}(PQP) = \text{supp}(P) \cap \text{supp}(Q)$;
3. $\text{Null}(PQP) = \text{Null}(P) + \text{Null}(Q)$;
4. PQP and QPQ have the same support.
5. PQP and QPQ commute;
6. Π_{PQP} and Π_{QPQ} commute.

Lemma 3. *Suppose $A, B \in \mathcal{P}(\mathcal{H})$. Then $\text{supp}(AB) \leq \text{supp}(A)$ iff B commutes with Π_A . In addition, AB and BA have the same support iff A commutes with Π_B and B commutes with Π_A .*

Lemma 4. *Suppose $A, B \in \mathcal{P}(\mathcal{H})$. Then there exists a unitary operator U on $\mathcal{H}$ that commutes with Π_A and satisfies $\sqrt{A}\sqrt{B}U = \sqrt{\sqrt{A}B\sqrt{A}}$ iff B commutes with Π_A .*

Proof of Lemma 2. For any projectors P and Q , there exists two orthogonal basis $\{|\psi_i\rangle\}_i$ and $\{|\varphi_i\rangle\}_i$ such that

$$P = \sum_{i=1}^{\text{rank } P} |\psi_i\rangle\langle\psi_i|, \quad Q = \sum_{j=1}^{\text{rank } Q} |\varphi_j\rangle\langle\varphi_j|, \quad (\text{E1})$$

$$|\langle\psi_i|\varphi_j\rangle| = q_i\delta_{i,j} \quad \forall i \leq \text{rank } P, j \leq \text{rank } Q.$$

Therefore, it suffices to prove the case when $\text{rank } P = \text{rank } Q = 1$. In this case, it is direct to show that the six statements in Lemma 2 are equivalent. $\square$

Proof of Lemma 3. If $\text{supp}(AB) \leq \text{supp}(A)$, then we have $AB\Pi_A = AB$. After multiplying both sides of the equation by A^+ from the left side, we can deduce that

$$\Pi_A B \Pi_A = \Pi_A B = B \Pi_A, \quad (\text{E2})$$

so B commutes with Π_A . Here the last equality holds because $\Pi_A B \Pi_A$ is Hermitian. Conversely, if B commutes with Π_A , then $AB\Pi_A = AB$, which implies that $\text{supp}(AB) \leq \text{supp}(A)$. Therefore, $\text{supp}(AB) \leq \text{supp}(A)$ iff B commutes with Π_A . By symmetry we can deduce that $\text{supp}(BA) \leq \text{supp}(B)$ iff A commutes with Π_B .

Next, we turn to the second statement in Lemma 3. If AB and BA have the same support, then $\text{supp}(AB) \leq \text{supp}(A)$ and $\text{supp}(BA) \leq \text{supp}(B)$. So A commutes with Π_B and B commutes with Π_A according to the above discussion. Conversely, if A commutes with Π_B and B commutes with Π_A , then

$$\text{supp}(AB) = \text{supp}(BA) = \text{supp}(A) \cap \text{supp}(B), \quad (\text{E3})$$

which completes the proof of Lemma 3. $\square$

Proof of Lemma 4. By assumption we can deduce that $\text{supp}(\sqrt{A}B\sqrt{A}) \leq \text{supp}(\sqrt{A}) = \text{supp}(A)$. If B commutes with Π_A , then $\sqrt{B}$ also commutes with Π_A , and $\text{supp}(\sqrt{A}\sqrt{B}) \leq \text{supp}(A)$ by Lemma 3. So there exists a unitary operator U on $\mathcal{H}$ that commutes with Π_A and satisfies $\sqrt{A}\sqrt{B}U = \sqrt{\sqrt{A}B\sqrt{A}}$.

Conversely, if there exists a unitary operator U that commutes with Π_A and satisfies $\sqrt{A}\sqrt{B}U = \sqrt{\sqrt{A}B\sqrt{A}}$, then $U^\dagger$ also commutes with Π_A and we can deduce that $\text{supp}(\sqrt{A}\sqrt{B}) \leq \text{supp}(A)$. Therefore, $\sqrt{B}$ commutes with $\Pi_{\sqrt{A}}$, which means B commutes with Π_A and completes the proof of Lemma 4. $\square$

Appendix F: Properties of operator geometric means

Before proving the results on the optimal measurements with respect to the fidelity as presented in the main text, here we summarize the key properties of the geometric mean of two positive operators A and B on $\mathcal{H}$. When A, B are nonsingular (or positive-definite), their geometric mean, denoted by $\mathcal{M}(A, B)$, is defined as

$$\mathcal{M}(A, B) := \sqrt{A}\sqrt{\sqrt{A^{-1}}B\sqrt{A^{-1}}}\sqrt{A}. \quad (\text{F1})$$

This geometric mean has the following properties [13, 14]:

1. $\mathcal{M}(A, B)$ is positive definite.
2. $\mathcal{M}(A, B)^{-1} = \mathcal{M}(A^{-1}, B^{-1})$.
3. $\mathcal{M}(A, B) = \mathcal{M}(B, A)$.
4. $\mathcal{M}(A, B)$ is the unique positive solution of the matrix equation $XA^{-1}X = B$. Notably, $\mathcal{M}(A, B) = \mathbb{1}$ iff $A = B^{-1}$.

Note that Eq. (F1) is still well-defined when B is singular as long as A is nonsingular. When A is singular, we can extend the above definition by virtue of the Moore-Penrose generalized inverse (MPGI) [30, 31].

Generally, for any linear operator O on $\mathcal{H}$, the Moore-Penrose generalized inverse of O , denoted by O^+ , is uniquely determined by the following four conditions:

1. $OO^+O = O$;
2. $O^+OO^+ = O^+$;
3. $(OO^+)^\dagger = OO^+$;
4. $(O^+O)^\dagger = O^+O$.

By definition we have $(O^+)^+ = O$; in addition, OO^+ and O^+O are projectors onto $\text{range}(O)$ and $\text{range}(O^+)$, respectively. If O is invertible, then the MPGI of O coincides with its inverse O^{-1} . If O is normal, that is, $O^\dagger O = OO^\dagger$, then O^+ has the same support as O and is the inverse of O in $\text{supp}(O)$, that is, $OO^+ = O^+O = \Pi_O$, where Π_O is the projector onto the support of O . If in addition O is positive semidefinite, then $(O^x)^+ = (O^+)^x$ and $(O^+)^x O^x = O^x (O^+)^x = \Pi_O$ for $x > 0$.

Now, the (generalized) geometric mean of two general positive operators A, B on $\mathcal{H}$ can be defined as follows:

$$\mathcal{M}(A, B) := \sqrt{A}\sqrt{\sqrt{A^+}B\sqrt{A^+}}\sqrt{A}, \quad (\text{F2})$$

which is a positive operator by definition. Next, we summarize the basic properties of the geometric mean of two positive operators.

1. Basic properties

Lemma 5. Suppose $A, B \in \mathcal{P}(\mathcal{H})$, Π_A and Π_B are the projectors onto the supports of A and B , respectively, and $\mathcal{M} = \mathcal{M}(A, B)$. Then

$$\text{Null}(\mathcal{M}) = \text{Null}(\Pi_A \Pi_B \Pi_A), \quad (\text{F3})$$

$$\text{supp}(\mathcal{M}) = \text{supp}(\Pi_A \Pi_B \Pi_A), \quad (\text{F4})$$

$$\text{Null}(A) \leq \text{Null}(\mathcal{M}) \leq \text{Null}(A) + \text{Null}(B), \quad (\text{F5})$$

$$\text{supp}(A) \cap \text{supp}(B) \leq \text{supp}(\mathcal{M}) \leq \text{supp}(A). \quad (\text{F6})$$

Lemma 6. Suppose $A, B, P \in \mathcal{P}(\mathcal{H})$ and P is supported in $\text{supp}(A) \cap \text{supp}(B)$ and an eigenspace of $\mathcal{M}(A, B)$ with eigenvalue λ . Then $\lambda > 0$ and P is supported in the eigenspace of $\mathcal{M}(B^+, A^+)$ with eigenvalue $1/\lambda$. The same result holds if $\mathcal{M}(A, B)$ and $\mathcal{M}(B^+, A^+)$ are exchanged.

Lemma 7. Suppose $A, B \in \mathcal{P}(\mathcal{H})$, and Π_A and Π_B are the projectors onto the supports of A and B , respectively. Then the following nine statements are equivalent:

1. Π_A and Π_B commute;
2. $\text{supp}(\mathcal{M}(A, B)) = \text{supp}(A) \cap \text{supp}(B)$;
3. $\text{supp}(\mathcal{M}(A, B)) \leq \text{supp}(B)$;
4. $\text{Null}(\mathcal{M}(A, B)) = \text{Null}(A) + \text{Null}(B)$;
5. $\text{Null}(B) \leq \text{Null}(\mathcal{M}(A, B))$;
6. $\mathcal{M}(B, A)$ and $\mathcal{M}(A, B)$ have the same support;
7. $\mathcal{M}(B^+, A^+)$ and $\mathcal{M}(A, B)$ have the same support;
8. $\mathcal{M}(B^+, A^+)$ is the MPGI of $\mathcal{M}(A, B)$.
9. $\mathcal{M}(B^+, A^+)$ and $\mathcal{M}(A, B)$ commute.

Lemma 8. Suppose $A, B \in \mathcal{P}(\mathcal{H})$. Then $\mathcal{M}(A, B) = \mathcal{M}(B, A)$ iff $\mathcal{M}(B^+, A^+) = \mathcal{M}(A^+, B^+)$.

Lemma 9. Suppose $A, B \in \mathcal{P}(\mathcal{H})$, Π_A and Π_B are the projectors onto the supports of A and B , respectively, A commutes with Π_B , and B commutes with Π_A . Then $\mathcal{M}(A, B) = \mathcal{M}(B, A)$ and $\mathcal{M}(B^+, A^+) = \mathcal{M}(A^+, B^+)$.

Next, we turn to the special case in which one operator, say A , is nonsingular. The following two lemmas are simple corollaries of Lemmas 5 and 7, respectively.

Lemma 10. Suppose $A, B \in \mathcal{P}(\mathcal{H})$ and A is nonsingular. Then the null spaces (supports) of $\mathcal{M}(A, B)$, $\mathcal{M}(B, A)$, $\mathcal{M}(A, B^+)$, and $\mathcal{M}(B^+, A)$ all coincide with the null space (support) of B , and the same results still hold if A is replaced by A^{-1} .

Lemma 11. Suppose $A, B \in \mathcal{P}(\mathcal{H})$ and A is nonsingular. Then $\mathcal{M}(B^+, A^{-1})$ and $\mathcal{M}(A^{-1}, B^+)$ are the MPGIs of $\mathcal{M}(A, B)$ and $\mathcal{M}(B, A)$, respectively.

2. Proofs of Lemmas 5-9

Proof of Lemma 5. Let $R = \sqrt{A^+} B \sqrt{A^+}$, then

$$\mathcal{M} = \sqrt{A} \sqrt{R} \sqrt{A} = (R^{1/4} \sqrt{A})^\dagger R^{1/4} \sqrt{A}. \quad (\text{F7})$$

Therefore,

$$\begin{aligned} \text{Null}(\mathcal{M}) &= \text{Null}(R^{1/4} \sqrt{A}) = \text{Null}(R \sqrt{A}) \\ &= \text{Null}(\sqrt{A^+} B \Pi_A) = \text{Null}(\Pi_A B \Pi_A) = \text{Null}(\sqrt{B} \Pi_A) \\ &= \text{Null}(\Pi_B \Pi_A) = \text{Null}(\Pi_A \Pi_B \Pi_A), \end{aligned} \quad (\text{F8})$$

which confirms Eq. (F3) and implies Eq. (F4). Equations (F5) and (F6) are simple corollaries of Eqs. (F3) and (F4). $\square$

Proof of Lemma 6. Note that $\mathcal{M}(A, B)$ is a positive operator and $\text{supp}(A) \cap \text{supp}(B) \leq \text{supp}(\mathcal{M}(A, B))$ according to Lemma 5. In addition, P is supported in $\text{supp}(A) \cap \text{supp}(B)$ by assumption and is thus supported in $\text{supp}(\mathcal{M}(A, B))$. Now, suppose P is supported in an eigenspace of $\mathcal{M}(A, B)$ with eigenvalue λ , then we have $\lambda > 0$ because this eigenspace is necessarily contained in $\text{supp}(\mathcal{M}(A, B))$.

Let U be any unitary on $\mathcal{H}$ that satisfies the condition

$$\sqrt{A^+} \sqrt{B} U = U^\dagger \sqrt{B} \sqrt{A^+} = \sqrt{\sqrt{A^+} B \sqrt{A^+}}. \quad (\text{F9})$$

Then

$$\begin{aligned} U^\dagger \sqrt{B} P &= \sqrt{\sqrt{A^+} B \sqrt{A^+}} \sqrt{A} P \\ &= \sqrt{A^+} \sqrt{A} \sqrt{\sqrt{A^+} B \sqrt{A^+}} \sqrt{A} P \\ &= \sqrt{A^+} \mathcal{M}(A, B) P = \lambda \sqrt{A^+} P, \end{aligned} \quad (\text{F10})$$

given that P is supported in $\text{supp}(A) \cap \text{supp}(B)$ and an eigenspace of $\mathcal{M}(A, B)$ with eigenvalue $\lambda > 0$. In addition, $\sqrt{A} \sqrt{A^+}$ and $\sqrt{A^+} \sqrt{A}$ coincide with the projector onto $\text{supp}(A)$; similarly, $\sqrt{B} \sqrt{B^+}$ and $\sqrt{B^+} \sqrt{B}$ coincide with the projector onto $\text{supp}(B)$.

Next, multiplying Eq. (F10) with $\sqrt{B^+} U$ (on the left) yields

$$\begin{aligned} P &= \lambda \sqrt{B^+} U \sqrt{A^+} P = \lambda \sqrt{B^+} U \sqrt{A^+} \sqrt{B} \sqrt{B^+} P \\ &= \lambda \mathcal{M}(B^+, A^+) P, \end{aligned} \quad (\text{F11})$$

where the last equality follows from the equation below,

$$\sqrt{B} \sqrt{A^+} U^\dagger = U \sqrt{A^+} \sqrt{B} = \sqrt{\sqrt{B} A^+ \sqrt{B}}, \quad (\text{F12})$$

which, in turn, follows from Eq. (F9). So P is supported in the eigenspace of $\mathcal{M}(B^+, A^+)$ with eigenvalue $1/\lambda$.

Finally, suppose instead P is supported in the eigenspace of $\mathcal{M}(B^+, A^+)$ with eigenvalue λ . Note that A^+ and B^+ have the same supports as A and B , respectively; in addition, $(A^+)^+ = A$ and $(B^+)^+ = B$. Therefore, $\lambda > 0$ and P is supported in the eigenspace of $\mathcal{M}(A, B)$ with eigenvalue $1/\lambda$ according to a similar reasoning employed above, which completes the proof of Lemma 6. $\square$

Proof of Lemma 7. By assumption A and B are positive operators, so A^+ and B^+ have the same supports as A and B , respectively. In conjunction with Lemma 5 we can deduce that

$$\begin{aligned} \text{supp}(\mathcal{M}(A, B)) &= \text{supp}(\Pi_A \Pi_B \Pi_A), \\ \text{supp}(\mathcal{M}(B^+, A^+)) &= \text{supp}(\mathcal{M}(B, A)) \\ &= \text{supp}(\Pi_B \Pi_A \Pi_B). \end{aligned} \quad (\text{F13})$$

Therefore, the equivalence of Statements 1-7 in Lemma 7 follows from Lemma 2.

Next, suppose the first seven statements in Lemma 7 hold. Then the supports of $\mathcal{M}(B^+, A^+)$ and $\mathcal{M}(A, B)$

coincide with $\text{supp}(A) \cap \text{supp}(B)$. Moreover, according to Lemma 6, $\mathcal{M}(B^+, A^+)$ is the inverse of $\mathcal{M}(A, B)$ when restricted to $\text{supp}(A) \cap \text{supp}(B)$. So $\mathcal{M}(A^+, B^+)$ is the MPGI of $\mathcal{M}(A, B)$, which confirms Statement 8. Conversely, if Statement 8 holds, then $\mathcal{M}(B^+, A^+)$ and $\mathcal{M}(A, B)$ necessarily have the same support, which confirms Statement 7, given that $\mathcal{M}(B^+, A^+)$ and $\mathcal{M}(A, B)$ are positive operators. So the first eight statements in Lemma 7 are equivalent.

Finally, suppose the first eight statements in Lemma 7 hold. Then $\mathcal{M}(B^+, A^+)$ is the MPGI of $\mathcal{M}(A, B)$ and thus commutes with $\mathcal{M}(A, B)$, which confirms Statement 9. Conversely, if $\mathcal{M}(B^+, A^+)$ and $\mathcal{M}(A, B)$ commute, then the projectors onto their supports necessarily commute. In conjunction with Eq. (F13) and Lemma 2 we can deduce that Π_A and Π_B commute, which confirms Statement 1. In summary, all nine statements in Lemma 7 are equivalent, which completes the proof. $\square$

Proof of Lemma 8. Suppose $\mathcal{M}(A, B) = \mathcal{M}(B, A)$; then $\mathcal{M}(A, B)$ and $\mathcal{M}(B, A)$ have the same supports. According to Lemma 7, $\mathcal{M}(B^+, A^+)$ and $\mathcal{M}(A^+, B^+)$ are the MPGIs of $\mathcal{M}(A, B)$ and $\mathcal{M}(B, A)$, respectively, which implies that $\mathcal{M}(B^+, A^+) = \mathcal{M}(A^+, B^+)$. Conversely, if $\mathcal{M}(B^+, A^+) = \mathcal{M}(A^+, B^+)$, then $\mathcal{M}(A, B) = \mathcal{M}(B, A)$ by a similar reasoning given that $(A^+)^+ = A$ and $(B^+)^+ = B$. $\square$

Proof of Lemma 9. By assumption, B can be expressed in the form $B = B_0 + B_1$, where B_0 and B_1 are positive operators supported in $\text{Null}(A)$ and $\text{supp}(A)$, respectively. In addition, $\Pi_B = \Pi_{B_0} + \Pi_{B_1}$ and A commutes with both Π_{B_0} and Π_{B_1} . It follows that A can be expressed as $A = A_0 + A_1$, where A_0 and A_1 are positive operators supported in $\text{supp}(A) \cap \text{Null}(B_1)$ and $\text{supp}(B_1) = \text{supp}(A) \cap \text{supp}(B)$, respectively; meanwhile, A_1 has full rank within $\text{supp}(B_1)$. Therefore,

$$\mathcal{M}(A, B) = \mathcal{M}(A_1, B_1) = \mathcal{M}(B_1, A_1) = \mathcal{M}(B, A), \quad (\text{F14})$$

where the second equality holds because the MPGI is symmetric in the two operators when both operators have the same support and are nonsingular within the support.

In addition, A^+ and B^+ have the same supports and the same sets of eigenspaces as A and B , respectively. Furthermore, Π_B commutes with A iff it commutes with A^+ ; similarly, Π_A commutes with B iff it commutes with B^+ . Therefore, $\mathcal{M}(B^+, A^+) = \mathcal{M}(A^+, B^+)$ by a similar reasoning employed above. Alternatively, this conclusion follows from the equality $\mathcal{M}(A, B) = \mathcal{M}(B, A)$ proved above thanks to Lemma 8. $\square$

Appendix G: Properties of operator pencils

In this section we clarify the basic properties of linear operator pencils that are helpful for understanding F-optimal measurements. Here we assume that A and B

are nonzero positive operators in $\mathcal{P}(\mathcal{H})$ and $A + B$ is nonsingular. In Lemmas 12-15 and Corollary 2 below, we further assume that U is a unitary operator that satisfies the condition $\sqrt{A}\sqrt{B}U = \sqrt{\sqrt{A}B\sqrt{A}}$.

1. Basic properties

Although eigenvectors of the pencil $(\sqrt{B}, U\sqrt{A})$ associated with distinct eigenvalues are not necessarily linearly independent, any two such eigenvectors are linearly independent according to the following lemma.

Lemma 12. *Any two distinct eigenspaces of the pencil $(\sqrt{B}, U\sqrt{A})$ have a trivial intersection. If $P, Q \in \mathcal{P}(\mathcal{H})$ are nonzero positive operators supported in distinct eigenspaces of $(\sqrt{B}, U\sqrt{A})$. Then $P + Q$ is not supported on any eigenspace.*

Lemma 13. *Suppose $P \in \mathcal{P}(\mathcal{H})$ is a positive operator that commutes with Π_A and is supported in an eigenspace of $(\sqrt{B}, U\sqrt{A})$. Then P is supported in either $\text{supp}(A)$ or $\text{Null}(A)$.*

Next, we clarify the relations between the eigenspaces of the pencil $(\sqrt{B}, U\sqrt{A})$ and the eigenspaces of the geometric mean $\mathcal{M} = \mathcal{M}(A^+, B)$. Recall that the eigenspaces of $(\sqrt{B}, U\sqrt{A})$ with eigenvalues ∞ and 0 coincide with $\text{Null}(A)$ and $\text{Null}(B)$, respectively, and $\text{Null}(\mathcal{M}) = \text{Null}(\Pi_A \Pi_B \Pi_A)$ by Lemma 5. If Π_A and Π_B commute, then $\text{Null}(A) \perp \text{Null}(B)$ and $\text{Null}(\mathcal{M}) = \text{Null}(A) \oplus \text{Null}(B)$ by Lemma 7, so $\text{Null}(\mathcal{M})$ is the direct sum of the two eigenspaces of $(\sqrt{B}, U\sqrt{A})$ with eigenvalues ∞ and 0.

Lemma 14. *Any vector in $\text{supp}(A)$ is an eigenstate of $(\sqrt{B}, U\sqrt{A})$ iff it is an eigenstate of $\mathcal{M}(A^+, B)$ with the same eigenvalue.*

If A is nonsingular, then $\text{supp}(A) = \mathcal{H}$ and all eigenvalues of $(\sqrt{B}, U\sqrt{A})$ are finite. In this case, Lemma 14 implies the following result as a simple corollary.

Corollary 2. *Suppose A is nonsingular. Then any vector in $\mathcal{H}$ is an eigenstate of $(\sqrt{B}, U\sqrt{A})$ iff it is an eigenstate of $\mathcal{M}(A^{-1}, B)$ with the same eigenvalue.*

Lemma 15. *Suppose B and U commute with Π_A . Then every eigenstate of $(\sqrt{B}, U\sqrt{A})$ belongs to either $\text{Null}(A)$ or $\text{supp}(A)$.*

When B commutes with Π_A , thanks to Lemma 4, there exists a unitary operator U on $\mathcal{H}$ that commutes with Π_A and satisfies $\sqrt{A}\sqrt{B}U = \sqrt{\sqrt{A}B\sqrt{A}}$. Furthermore, the commutativity condition between B and Π_A in Lemma 15 can be relaxed as shown in the following lemma.

Lemma 16. *Suppose Π_A and Π_B commute. Then there exists a unitary operator U on $\mathcal{H}$ that satisfies $\sqrt{A}\sqrt{B}U = \sqrt{\sqrt{A}B\sqrt{A}}$, and every eigenvector of $(\sqrt{B}, U\sqrt{A})$ belongs to either $\text{Null}(A)$ or $\text{supp}(A)$.*

The following lemma is a simple corollary of Lemmas 14 and 16.

Lemma 17. *Suppose Π_A and Π_B commute. Then there exists a unitary operator U on $\mathcal{H}$ that satisfies $\sqrt{A}\sqrt{B}U = \sqrt{\sqrt{A}B\sqrt{A}}$, and any nonzero vector in $\mathcal{H}$ is an eigenstate of $(\sqrt{B}, U\sqrt{A})$ iff it belongs to $\text{Null}(A)$ or an eigenspace of $\mathcal{M}(A^+, B)$ within $\text{supp}(A)$.*

If in addition B commutes with Π_A , then there exists a unitary operator U that commutes with Π_A and satisfies the condition $\sqrt{A}\sqrt{B}U = \sqrt{\sqrt{A}B\sqrt{A}}$. Moreover, such a unitary operator automatically satisfies the requirement in Lemma 17 according to Lemmas 14 and 15.

2. Proofs of Lemmas 12-16

Proof of Lemma 12. Suppose, by way of contradiction, that $\mathcal{V}_1$ and $\mathcal{V}_2$ are two eigenspaces of $(\sqrt{B}, U\sqrt{A})$ with different eigenvalues λ_1 and λ_2 that have a nontrivial intersection. Let $|\psi\rangle$ be a normalized vector in $\mathcal{V}_1 \cap \mathcal{V}_2$. If $\lambda_1 = \infty$, then $\lambda_2 \neq \infty$ and

$$\begin{aligned} |\psi\rangle &\in \text{Null}(U\sqrt{A}) = \text{Null}(\sqrt{A}) = \text{Null}(A), \\ |\psi\rangle &\in \text{Null}(\sqrt{B} - \lambda_2 U\sqrt{A}), \end{aligned} \quad (\text{G1})$$

which means $|\psi\rangle \in \text{Null}(A) \cap \text{Null}(B) = \text{Null}(A+B)$, in contradiction with the assumption that $A+B$ is nonsingular. This contradiction shows that $\lambda_1 \neq \infty$ and $\lambda_2 \neq \infty$ by symmetry. Consequently,

$$|\psi\rangle \in \text{Null}(\sqrt{B} - \lambda_1 U\sqrt{A}) \cap \text{Null}(\sqrt{B} - \lambda_2 U\sqrt{A}), \quad (\text{G2})$$

which means $|\psi\rangle \in \text{Null}(\sqrt{A}) \cap \text{Null}(\sqrt{B}) = \text{Null}(A+B)$. This contradiction shows that any two distinct eigenspaces of $(\sqrt{B}, U\sqrt{A})$ have a trivial intersection.

Next, suppose $P, Q \in \mathcal{P}(\mathcal{H})$ are nonzero and supported in distinct eigenspaces of $(\sqrt{B}, U\sqrt{A})$. Suppose, by way of contradiction, that $P+Q$ is supported in an eigenspace of $(\sqrt{B}, U\sqrt{A})$. Then P and Q are supported in the same eigenspace, which contradicts the conclusion proved above. This contradiction completes the proof of Lemma 12. $\square$

Proof of Lemma 13. If P is supported in $\text{Null}(A)$, then there is nothing to prove, so we assume that P is not supported in $\text{Null}(A)$ in this proof. By assumption P commutes with Π_A and thus can be expressed as $P = P_0 + P_1$, where P_0 and $P_1 \neq 0$ are supported in $\text{Null}(A)$ and $\text{supp}(A)$, respectively, so that $\sqrt{A}P = \sqrt{A}P_1 \neq 0$. In addition, by assumption $\sqrt{B}P$ and $U\sqrt{A}P$ are parallel, which means

$$\sqrt{B}P = \sqrt{B}P_0 + \sqrt{B}P_1 = \lambda U\sqrt{A}P = \lambda U\sqrt{A}P_1, \quad (\text{G3})$$

where λ is some nonnegative constant. This equation can hold only if $\sqrt{B}P_0 = 0$, so P_0 is necessarily supported in

$\text{Null}(A) \cap \text{Null}(B) = \text{Null}(A+B)$, which means $P_0 = 0$ given that $A+B$ is nonsingular by assumption. Therefore, $P = P_1$ is supported in $\text{supp}(A)$, which completes the proof of Lemma 13. $\square$

Proof of Lemma 14. Without loss of generality, let $|\psi\rangle$ be a normalized vector in $\text{supp}(A)$. Then $\Pi_A|\psi\rangle = |\psi\rangle$, where Π_A is the projector onto $\text{supp}(A)$ and can be expressed as $\Pi_A = \sqrt{A}\sqrt{A^+} = \sqrt{A^+}\sqrt{A}$. In conjunction with the assumption $\sqrt{A}\sqrt{B}U = \sqrt{\sqrt{A}B\sqrt{A}}$ we can deduce that

$$\begin{aligned} \sqrt{\sqrt{A}B\sqrt{A}} &= U^\dagger \sqrt{B}\sqrt{A}, \\ \sqrt{\sqrt{A}B\sqrt{A}}\sqrt{A^+}|\psi\rangle &= U^\dagger \sqrt{B}\Pi_A|\psi\rangle = U^\dagger \sqrt{B}|\psi\rangle, \\ \mathcal{M}(A^+, B)|\psi\rangle &= \sqrt{A^+}U^\dagger \sqrt{B}|\psi\rangle, \end{aligned} \quad (\text{G4})$$

Suppose $|\psi\rangle$ is an eigenstate of $(\sqrt{B}, U\sqrt{A})$ with eigenvalue λ . Then $\lambda \neq \infty$ given that $|\psi\rangle \in \text{supp}(A)$ by assumption, and we have

$$U^\dagger \sqrt{B}|\psi\rangle = \lambda \sqrt{A}|\psi\rangle. \quad (\text{G5})$$

The above equations together imply that

$$\mathcal{M}(A^+, B)|\psi\rangle = \lambda \sqrt{A^+}\sqrt{A}|\psi\rangle = \lambda|\psi\rangle. \quad (\text{G6})$$

So $|\psi\rangle$ is an eigenstate of $\mathcal{M}(A^+, B)$ with eigenvalue λ .

Conversely, suppose $|\psi\rangle$ is an eigenstate of $\mathcal{M}(A^+, B)$ with eigenvalue λ , then $0 \leq \lambda < \infty$, given that $\mathcal{M}(A^+, B)$ is a positive operator. In conjunction with Eq. (G4) we can deduce that

$$\begin{aligned} U^\dagger \sqrt{B}|\psi\rangle &= \sqrt{A}\sqrt{A^+}\sqrt{\sqrt{A}B\sqrt{A}}\sqrt{A^+}|\psi\rangle \\ &= \sqrt{A}\mathcal{M}(A^+, B)|\psi\rangle = \lambda \sqrt{A}|\psi\rangle. \end{aligned} \quad (\text{G7})$$

So $|\psi\rangle$ is an eigenstate of $(\sqrt{B}, U\sqrt{A})$ with eigenvalue λ , which completes the proof of Lemma 14. $\square$

Proof of Lemma 15. Let $\bar{\Pi}_A = \mathbb{1} - \Pi_A$ be the projector onto $\text{Null}(A)$. By assumption B and U commute with Π_A and $\bar{\Pi}_A$, so $\sqrt{B}$ also commutes with Π_A and $\bar{\Pi}_A$.

Suppose $|\psi\rangle$ is a normalized eigenstate of $(\sqrt{B}, U\sqrt{A})$ with eigenvalue λ , and let $|\psi_0\rangle = \bar{\Pi}_A|\psi\rangle$, which is not necessarily normalized. If $|\psi\rangle \in \text{Null}(A)$, then there is nothing to prove, so we assume that $|\psi\rangle \notin \text{Null}(A)$, which means $\lambda \neq \infty$ and

$$\sqrt{B}|\psi\rangle = \lambda U\sqrt{A}|\psi\rangle. \quad (\text{G8})$$

Multiplying both sides of the above equation by the projector $\bar{\Pi}_A$ yields $\sqrt{B}|\psi_0\rangle = 0$, which implies that $|\psi_0\rangle \in \text{Null}(A) \cap \text{Null}(B) = \text{Null}(A+B)$. Since $A+B$ is nonsingular by assumption, we can conclude that $|\psi_0\rangle = 0$ and $|\psi\rangle \in \text{supp}(A)$, which completes the proof of Lemma 15. $\square$

Proof of Lemma 16. By assumption A and B are positive operators on $\mathcal{H}$, $A+B$ is nonsingular, and Π_A commutes with Π_B . If A is nonsingular, then the conclusion in Lemma 16 is obvious. So we assume that A is nonsingular in the following proof. Then $\mathcal{H}$ can be expressed as follows:

$$\begin{aligned}\mathcal{H} &= \text{Null}(A) \oplus \text{supp}(A) = \text{Null}(B) \oplus \text{supp}(B) \\ &= \text{Null}(A) \oplus \text{Null}(B) \oplus [\text{supp}(A) \cap \text{supp}(B)],\end{aligned}\quad (\text{G9})$$

where $\text{Null}(A)$, $\text{Null}(B)$, and $\text{supp}(A) \cap \text{supp}(B)$ are mutually orthogonal, and $\text{Null}(B)$ can be omitted if B is nonsingular. Let $\bar{\Pi}_A = \mathbb{1} - \Pi_A$ be the projector onto $\text{Null}(A)$, $r = \text{tr}(\Pi_A \Pi_B) = \dim(\text{supp}(A) \cap \text{supp}(B))$ and $\mathcal{M} = \mathcal{M}(A^+, B)$. By virtue of Lemma 7 we can deduce that $\mathcal{M}^+ = \mathcal{M}(B^+, A)$ and

$$\text{rank}(\sqrt{A}\sqrt{B}) = \text{rank}(\mathcal{M}) = \text{rank}(\Pi_A \Pi_B) = r. \quad (\text{G10})$$

Next, we turn to the singular-value decomposition of $\sqrt{A}\sqrt{B}$ in order to find a desired unitary operator U featured in the polar decomposition. In view of Eq. (G10), $\sqrt{A}\sqrt{B}$ has a singular-value decomposition of the form

$$\sqrt{A}\sqrt{B} = \sum_{i=1}^r s_i |\alpha_i\rangle \langle \beta_i|, \quad (\text{G11})$$

where $\{s_i\}_{i=1}^r$ is the set of nonzero singular values, $\{|\alpha_i\rangle\}_{i=1}^r$ is a set of orthonormal vectors in $\text{supp}(A)$, and $\{|\beta_i\rangle\}_{i=1}^r$ is a set of orthonormal vectors in $\text{supp}(B)$. Thanks to Eq. (G9), we can extend $\{|\alpha_i\rangle\}_{i=1}^r$ and $\{|\beta_i\rangle\}_{i=1}^r$ into orthonormal bases $\{|\alpha_i\rangle\}_{i=1}^d$ and $\{|\beta_i\rangle\}_{i=1}^d$ on $\mathcal{H}$ with the following properties:

$$\begin{aligned}|\alpha_i\rangle &\in \text{Null}(A), \quad |\beta_i\rangle \in \text{supp}(B) \quad \text{if } r < i \leq r_B, \\ |\alpha_i\rangle &\in \text{supp}(A), \quad |\beta_i\rangle \in \text{Null}(B) \quad \text{if } r_B < i \leq d,\end{aligned}\quad (\text{G12})$$

where $r_B = \text{rank}(B)$ and the second case can occur only if B is singular. By construction we can deduce the following results:

$$\begin{aligned}\text{supp}(B) &= \text{span}\{|\beta_i\rangle\}_{i=1}^{r_B} = \text{span}\{\sqrt{B^+}|\beta_i\rangle\}_{i=1}^{r_B}, \\ \text{supp}(\sqrt{\sqrt{B}A\sqrt{B}}) &= \text{supp}(\sqrt{B}A\sqrt{B}) = \text{span}\{|\beta_i\rangle\}_{i=1}^r, \\ \text{supp}(\Pi_A \Pi_B) &= \text{supp}(\mathcal{M}^+) = \text{span}\{\sqrt{B^+}|\beta_i\rangle\}_{i=1}^r, \\ \text{Null}(A) &= \text{span}\{\sqrt{B^+}|\beta_i\rangle\}_{i=r+1}^{r_B}.\end{aligned}\quad (\text{G13})$$

Now, a desired unitary operator U that satisfies the condition $\sqrt{A}\sqrt{B}U = \sqrt{\sqrt{A}B\sqrt{A}}$ can be constructed as follows:

$$U = \sum_{i=1}^d |\beta_i\rangle \langle \alpha_i|. \quad (\text{G14})$$

Suppose $|\psi\rangle$ is an eigenvector of $(\sqrt{B}, U\sqrt{A})$ with eigenvalue λ . If $\lambda = \infty$, then $|\psi\rangle \in \text{Null}(A)$ by definition.

Otherwise, $|\psi\rangle \notin \text{Null}(A)$ and $(\sqrt{B} - \lambda U\sqrt{A})|\psi\rangle = 0$. Consequently,

$$\begin{aligned}0 &= (U^\dagger \sqrt{B} - \lambda \sqrt{A})|\psi\rangle \\ &= (U^\dagger \sqrt{B}\sqrt{A}\sqrt{A^+} - \lambda \sqrt{A})|\psi\rangle + U^\dagger \sqrt{B}\bar{\Pi}_A|\psi\rangle \\ &= \sqrt{A}(\mathcal{M} - \lambda \mathbb{1})|\psi\rangle + U^\dagger \sqrt{B}\bar{\Pi}_A|\psi\rangle,\end{aligned}\quad (\text{G15})$$

where the second equality holds because $\mathbb{1} = \Pi_A + \bar{\Pi}_A$ and $\Pi_A = \sqrt{A}\sqrt{A^+}$, and the third equality holds because

$$\begin{aligned}U^\dagger \sqrt{B}\sqrt{A}\sqrt{A^+} &= \sqrt{\sqrt{A}B\sqrt{A^+}} \\ &= \sqrt{A}\sqrt{A^+}\sqrt{\sqrt{B}A\sqrt{B}}\sqrt{A^+} = \sqrt{A}\mathcal{M}.\end{aligned}\quad (\text{G16})$$

According to Eq. (G13), $\bar{\Pi}_A|\psi\rangle$ in Eq. (G15) can be expressed as follows:

$$\bar{\Pi}_A|\psi\rangle = \sum_{i=r+1}^{r_B} a_i \sqrt{B^+}|\beta_i\rangle, \quad (\text{G17})$$

where a_i are complex coefficients. In conjunction with Eq. (G12) we can deduce that

$$U^\dagger \sqrt{B}\bar{\Pi}_A|\psi\rangle = U^\dagger \sum_{i=r+1}^{r_B} a_i |\beta_i\rangle = \sum_{i=r+1}^{r_B} a_i |\alpha_i\rangle \in \text{Null}(A), \quad (\text{G18})$$

where the first equality holds because $\sqrt{B}\sqrt{B^+} = \Pi_B$ and $|\beta_i\rangle \in \text{supp}(B)$ for all $i = 1, 2, \dots, r_B$. Meanwhile, $\sqrt{A}(\mathcal{M} - \lambda \mathbb{1})|\psi\rangle \in \text{supp}(A)$. So Eq. (G15) implies that $U^\dagger \sqrt{B}\bar{\Pi}_A|\psi\rangle$, which means $a_i = 0$ for $r+1 \leq i \leq r_B$ and $\bar{\Pi}_A|\psi\rangle = 0$. Therefore, $|\psi\rangle \in \text{supp}(A)$, which completes the proof of Lemma 16. $\square$

Appendix H: Proofs of general results on F-optimal measurements

1. Proof of Proposition 1

Proof. Suppose $\mathcal{A} = \{A_j\}_j$ and $\mathcal{B} = \{B_k\}_k$. Denote the probability vectors $\Lambda_{\mathcal{A}}(\rho)$, $\Lambda_{\mathcal{A}}(\sigma)$, $\Lambda_{\mathcal{B}}(\rho)$, $\Lambda_{\mathcal{B}}(\sigma)$ by $\mathbf{p}, \mathbf{q}, \mathbf{p}', \mathbf{q}'$, respectively. By assumption, $\mathcal{A}$ is a coarse graining of $\mathcal{B}$. So we have

$$p_j = \sum_k S_{jk} p'_k, \quad q_j = \sum_k S_{jk} q'_k, \quad (\text{H1})$$

where S is a stochastic matrix. Therefore,

$$\begin{aligned}\sqrt{F_{\mathcal{A}}(\rho, \sigma)} &= \sum_j \sqrt{p_j q_j} = \sum_j \sqrt{\sum_k S_{jk} p'_k \sum_k S_{jk} q'_k} \\ &\geq \sum_j \sum_k S_{jk} \sqrt{p'_k q'_k} = \sum_k \sqrt{p'_k q'_k} = \sqrt{F_{\mathcal{B}}(\rho, \sigma)},\end{aligned}\quad (\text{H2})$$

which confirms Eq. (5). Here the inequality follows from the Cauchy-Schwarz inequality and the third equality holds because S is a stochastic matrix, which means $\sum_j S_{jk} = 1$ for all k . $\square$

2. Proof of Proposition 2

Proof of Proposition 2. By definition, we have

$$\begin{aligned}
\sqrt{F_{\mathcal{E}}(\rho, \sigma)} &= \sum_m \sqrt{\text{tr}(E_m \rho) \text{tr}(E_m \sigma)} \\
&\geq \sum_m \Re \text{tr}[(U \sqrt{\rho} \sqrt{E_m})(\sqrt{\sigma} \sqrt{E_m})^\dagger] \\
&= \sum_m \Re \text{tr}(\sqrt{\rho} E_m \sqrt{\sigma} U) = \text{tr}(\sqrt{\rho} \sqrt{\sigma} U) \\
&= \|\sqrt{\rho} \sqrt{\sigma}\|_1 = \sqrt{F(\rho, \sigma)}. \tag{H3}
\end{aligned}$$

Here the inequality follows from the Cauchy-Schwarz inequality, and it is saturated iff each POVM element E_m satisfies the following condition:

$$\sqrt{\rho} \sqrt{E_m} = 0 \quad \text{or} \quad \sqrt{\sigma} \sqrt{E_m} = \kappa_m U \sqrt{\rho} \sqrt{E_m}, \tag{H4}$$

where κ_m is a nonnegative constant, that is, $\sqrt{\sigma} \sqrt{E_m}$ and $U \sqrt{\rho} \sqrt{E_m}$ are parallel. This observation completes the proof of Proposition 2. $\square$

3. Proof of Proposition 3

Proof of Proposition 3. Note that the two operators $\sqrt{\sigma} \sqrt{E_m}$ and $U \sqrt{\rho} \sqrt{E_m}$ are parallel iff E_m is supported in an eigenspace of $(\sqrt{\sigma}, U \sqrt{\rho})$ with a nonnegative eigenvalue. Therefore, $\mathcal{E}$ is F-optimal for distinguishing ρ and σ iff each POVM element E_m is supported in an eigenspace of $(\sqrt{\sigma}, U \sqrt{\rho})$ with a nonnegative eigenvalue according to Proposition 2.

Next, suppose $\mathcal{E}$ is F-optimal and there exist two POVM elements supported in a same eigenspace of $(\sqrt{\sigma}, U \sqrt{\rho})$. Then we can construct a nontrivial coarse graining of $\mathcal{E}$ by merging the two POVM elements, and the resulting POVM is still F-optimal according to the conclusion in the previous paragraph, so $\mathcal{E}$ is not minimal. In other words, if $\mathcal{E}$ is F-optimal and minimal, then no two POVM elements of $\mathcal{E}$ are supported in a same eigenspace of $(\sqrt{\sigma}, U \sqrt{\rho})$.

Conversely, suppose $\mathcal{E}$ is F-optimal and no two POVM elements are supported in a same eigenspace of $(\sqrt{\sigma}, U \sqrt{\rho})$. Then any nontrivial coarse graining of $\mathcal{E}$ will contain a POVM element not supported in any eigenspace of $(\sqrt{\sigma}, U \sqrt{\rho})$ by Lemma 12. Therefore, $\mathcal{E}$ is F-optimal and minimal. This observation completes the proof of Proposition 3. $\square$

4. Proof of Proposition 4

Proof of Proposition 4. To prove the first statement in Proposition 4, suppose, by way of contradiction, that $\mathcal{E}_1$ and $\mathcal{E}_2$ are compatible. Then they admit a common refinement, denoted by $\mathcal{C}$ henceforth. Now, the minimal F-optimal POVMs $\mathcal{E}_1$ and $\mathcal{E}_2$ are both coarse graining

of $\mathcal{C}$ and are thus equivalent according to Proposition 3. This contradiction means $\mathcal{E}_1$ and $\mathcal{E}_2$ cannot be compatible.

Next, we turn to the second statement in Proposition 4. Let U be any unitary operator on $\mathcal{H}$ that satisfies $\sqrt{\rho} \sqrt{\sigma} U = \sqrt{\sqrt{\rho} \sigma \sqrt{\rho}}$. Then each POVM element in $\mathcal{E}(p)$ or $\mathcal{E}(q)$ is supported in an eigenspace of $(\sqrt{\sigma}, U \sqrt{\rho})$ according to Proposition 3. Suppose, by way of contradiction, that $\mathcal{E}(p)$ and $\mathcal{E}(q)$ with $0 \leq p < q \leq 1$ are equivalent. Then $\mathcal{E}(p)$ and $\mathcal{E}(q)$ are identical up to relabeling given that they are F-optimal and minimal and are thus automatically simple. Therefore, the POVM elements of $\mathcal{E}(p)$ and $\mathcal{E}(q)$ supported in the same eigenspace must coincide. Accordingly, the POVM elements of $\mathcal{E}_1$ and $\mathcal{E}_2$ supported in the same eigenspace must coincide, which means $\mathcal{E}_1$ and $\mathcal{E}_2$ are equivalent. This contradiction shows that $\mathcal{E}(p)$ and $\mathcal{E}(q)$ are not equivalent and are thus not compatible according to the first statement in Proposition 4 proved above. This observation completes the proof of Proposition 4. $\square$

5. Proof of Lemma 1

Proof of Lemma 1. If P is supported in an eigenspace of $(\sqrt{\sigma}, U \sqrt{\rho})$, then P is supported in either $\text{Null}(\rho)$ or $\text{supp}(\rho)$ by Lemma 13 and, in the later case, P is supported in an eigenspace of $\mathcal{M}(\rho^+, \sigma)$ by Lemma 14.

Conversely, if P is supported in $\text{Null}(\rho)$, then it is supported in the eigenspace of $(\sqrt{\sigma}, U \sqrt{\rho})$ with eigenvalue ∞ . If P is supported in an eigenspace of $\mathcal{M}(\rho^+, \sigma)$ within $\text{supp}(\rho)$, then P is supported in an eigenspace of $(\sqrt{\sigma}, U \sqrt{\rho})$ thanks to Lemma 14 again. This observation completes the proof of Lemma 1. $\square$

6. Proof of Theorem 2

Proof of Theorem 2. First, we prove that Statements 1-4 in Theorem 2 are equivalent. If Π_ρ and Π_σ commute, then, by virtue of Lemma 5, we can deduce that

$$\begin{aligned}
&\text{Null}(\rho) \perp \text{Null}(\sigma), \\
&\text{Null}(\mathcal{M}(\rho^+, \sigma)) = \text{Null}(\rho) \oplus \text{Null}(\sigma), \tag{H5} \\
&\mathcal{M}(\rho^+, \sigma) = \mathcal{M}(\sigma^+, \rho)^+.
\end{aligned}$$

so both PVMs $\mathcal{M}(\rho, \sigma)$ and $\mathcal{M}(\sigma, \rho)$ are composed of the projectors onto $\text{Null}(\rho)$, $\text{Null}(\sigma)$, and the eigenprojectors of $\mathcal{M}(\rho^+, \sigma)$ with positive eigenvalues (any null projector is deleted by default). It follows that $\mathcal{M}(\rho, \sigma)$ and $\mathcal{M}(\sigma, \rho)$ are commuting, compatible, and equivalent, which confirms the implications $1 \Rightarrow 2, 3, 4$.

By construction Π_ρ (Π_σ) is a sum of POVM elements in $\mathcal{M}(\rho, \sigma)$ [$\mathcal{M}(\sigma, \rho)$]. If $\mathcal{M}(\rho, \sigma)$ and $\mathcal{M}(\sigma, \rho)$ commute, then Π_ρ and Π_σ necessarily commute, which confirms the implication $2 \Rightarrow 1$. If instead $\mathcal{M}(\rho, \sigma)$ and $\mathcal{M}(\sigma, \rho)$ are equivalent, then both Π_ρ and Π_σ can be expressed as

sums of POVM elements in $\mathcal{M}(\rho, \sigma)$, so Π_ρ and Π_σ commute as before, which confirms the implication $4 \Rightarrow 1$. In addition, two PVMs are compatible iff they are commuting [20, 21], which confirms the equivalence of Statements 2 and 3. In conjunction with the above discussions we conclude that Statements 1-4 in Theorem 2 are equivalent.

Next, we will prove that Statement 5 is equivalent to Statements 1-4. Suppose $\mathcal{M}(\rho, \sigma)$ is the unique minimal F-optimal POVM. Then all F-optimal POVMs for distinguishing ρ and σ are refinements of $\mathcal{M}(\rho, \sigma)$, so $\mathcal{M}(\rho, \sigma)$ is a coarse graining of and equivalent to $\mathcal{M}(\sigma, \rho)$ according to Theorem 1. This result confirms the implication $5 \Rightarrow 4$, which further means $5 \Rightarrow 1, 2, 3, 4$ given that Statements 1-4 are equivalent.

Conversely, suppose Π_ρ and Π_σ commute. According to Lemma 17, there exists a unitary operator U that satisfies $\sqrt{\rho}\sqrt{\sigma}U = \sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}$, and each eigenspace of $(\sqrt{\sigma}, U\sqrt{\rho})$ coincides with either $\text{Null}(\rho)$ or an eigenspace of $\mathcal{M}(\rho^+, \sigma)$ within $\text{supp}(\rho)$. If $\mathcal{E}$ is an F-optimal POVM for distinguishing ρ and σ , then, according to Proposition 3, each POVM element of $\mathcal{E}$ is supported in an eigenspace of $(\sqrt{\sigma}, U\sqrt{\rho})$, that is, either $\text{Null}(\rho)$ or an eigenspace of $\mathcal{M}(\rho^+, \sigma)$ within $\text{supp}(\rho)$. This is the case iff $\mathcal{E}$ is a refinement of $\mathcal{M}(\rho, \sigma)$. Therefore, all F-optimal POVMs are refinements of $\mathcal{M}(\rho, \sigma)$, and $\mathcal{M}(\rho, \sigma)$ is the unique minimal F-optimal POVM, which confirms the implication $1 \Rightarrow 5$ and completes the proof of Theorem 2. $\square$

Appendix I: Proofs of Propositions 5 and 6

In this section we prove Propositions 5 and 6, which focus on F-optimal measurements for distinguishing a pure state and a mixed state.

Proof of Proposition 5. By assumption σ is a pure state, $\sqrt{\sigma} = \sigma$, and $\rho\sigma \neq 0$. Let $F = \text{tr}(\rho\sigma)$ be the fidelity between ρ and σ ; then $F > 0$ and $\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}} = \sqrt{\rho}\sigma\sqrt{\rho}/\sqrt{F}$. Therefore, we can find a unitary U on $\mathcal{H}$ that satisfies

$$\sqrt{\rho}\sigma U = \sqrt{\sqrt{\rho}\sigma\sqrt{\rho}} = \frac{1}{\sqrt{F}}\sqrt{\rho}\sigma\sqrt{\rho}, \quad (\text{I1})$$

which means $U^\dagger\sigma = \sqrt{\rho}\sigma/\sqrt{F}$. In turn this result implies

that

$$\begin{aligned} \text{Null}(\sigma - \lambda U\sqrt{\rho}) &= \text{Null}(U^\dagger\sigma - \lambda\sqrt{\rho}) \\ &= \text{Null}(\sqrt{\rho}\sigma - \lambda\sqrt{F}\sqrt{\rho}) = \text{Null}(\Pi_\rho\sigma - \lambda\sqrt{F}\Pi_\rho), \end{aligned} \quad (\text{I2})$$

where the last equality holds because $\sqrt{\rho}$ and Π_ρ have the same null space (and support).

Next, according to Proposition 3, $\mathcal{E}$ is F-optimal iff each POVM element is supported in an eigenspace of $(\sigma, U\sqrt{\rho})$ with a nonnegative eigenvalue, which is the case iff each POVM element is supported in an eigenspace of $(\Pi_\rho\sigma, \Pi_\rho)$ with a nonnegative eigenvalue thanks to Eq. (I2). This result completes the proof of Proposition 5. $\square$

Proof of Proposition 6. Proposition 6 is a simple corollary of Proposition 5 proved above and Lemma 18 below. $\square$

Lemma 18. Suppose $\dim \mathcal{H} = 2$, and ρ, σ are two pure states on $\mathcal{H}$ with non-commuting density operators. Let A and B be the representative points of ρ and σ on the Bloch sphere and let M and N be their antipodal points. Then a pure state is an eigenstate of $(\rho\sigma, \rho)$ with a non-negative eigenvalue iff its representative point lies on the major arc $\widehat{NABM}$ that connects M and N .

Proof of Lemma 18. Without loss of generality, ρ and σ can be expressed as $\rho = |\psi\rangle\langle\psi|$ and $\sigma = |\varphi\rangle\langle\varphi|$, where

$$|\psi\rangle = |0\rangle, \quad |\varphi\rangle = \cos\frac{\theta}{2}|0\rangle + \sin\frac{\theta}{2}|1\rangle, \quad \theta \in (0, \pi); \quad (\text{I3})$$

then $\langle\psi|\varphi\rangle = \cos(\theta/2) > 0$. Now, it is clear that any complex number λ (including ∞) is a nondegenerate eigenvalue of $(\rho\sigma, \rho)$, and any pure state is an eigenstate of $(\rho\sigma, \rho)$. When $\lambda \neq \infty$, the eigenstate associated with eigenvalue λ spans $\text{Null}(\rho\sigma - \lambda\rho)$ and is orthogonal to the unnormalized state $|\varphi\rangle\langle\varphi|\psi\rangle - \lambda|\psi\rangle$. The eigenstate of $(\rho\sigma, \rho)$ corresponding to the eigenvalue ∞ spans $\text{Null}(\rho)$ and is represented by M on the Bloch sphere (the antipodal point of the representative point of ρ).

Next, suppose λ is a nonnegative. As λ increases from 0 to ∞ , the representative point of $|\varphi\rangle\langle\varphi|\psi\rangle - \lambda|\psi\rangle$ (after normalization) traces the major arc $\widehat{BMNA}$. Accordingly, the representative point of the normalized eigenstate of $(\rho\sigma, \rho)$ with eigenvalue λ traces the major arc $\widehat{NABM}$. This observation completes the proof of Lemma 18. $\square$