Conditions for optimal input states for discrimination of quantum channels

Anna Jenčová and Martin Plávala

Conditions for optimal input states for discrimination of quantum channels

Anna Jenčová and Martin Plávala

Mathematical Institute, Slovak Academy of Sciences, Bratislava 81104, Slovakia

We find optimality conditions for testers in discrimination of quantum channels. These conditions are obtained using semidefinite programming and are similar to optimality conditions for discrimination of quantum states. We get a simple condition for existence of an optimal tester with any given input state with maximal Schmidt rank, in particular with a maximally entangled input state. In the case when maximally entangled state is not optimal, an upper bound on the optimal success probability is obtained. The results for discrimination of two channels are applied to covariant channels, qubit channels, unitary channels, and simple projective measurements. *Published by AIP Publishing.* [http://dx.doi.org/10.1063/1.4972286]

I. INTRODUCTION

The problem of multiple hypothesis testing in the setting of quantum channels can be formulated as follows. Assume that Φ is an unknown quantum channel, but some *a priori* information is available, in the sense that Φ is one of the given channels $\Phi_1, \ldots, \Phi_n$, with probabilities $\lambda_1, \ldots, \lambda_n$. The task is to find a procedure that determines the true channel, with the greatest possible probability of success.

For quantum states, this problem was formulated by Helstrom[14] and since then it has been the subject of active research, see, e.g., Ref. 1 for an overview and further references. Here, an ensemble $\{\lambda_i, \rho_i\}_{i=1}^n$ is given, where $\rho_1, \ldots, \rho_n$ are quantum states with prior probabilities $\lambda_1, \ldots, \lambda_n$, with a similar interpretation as above. A testing procedure, or a measurement, for this problem is described by a positive operator valued measure (POVM) M, defined as a collection of positive operators $M_1, \ldots, M_n$ summing up to the identity operator I. The value $\operatorname{Tr} M_i \rho_j$ is interpreted as the probability that the procedure chooses ρ_i while the true state is ρ_j. The task is to maximize the average success probability

$$p(M) = \sum_i \lambda_i \operatorname{Tr} M_i \rho_i$$

over all POVMs. In the case $n = 2$, it is well known that the optimal POVM is projection valued, given by the projections onto the positive and negative parts of the operator $\lambda \rho_1 - (1 - \lambda)\rho_2$.[14] For $n > 2$, there is no explicit expression for the optimal POVM in general, but it is known that a POVM M is optimal if and only if it satisfies

$$\sum_i \lambda_i \rho_i M_i \geq \lambda_j \rho_j, \qquad \forall j. \tag{1}$$

This condition was obtained in Refs. 15 and 26 using the methods of semidefinite programming (SDP).

In the case of quantum channels, a most general measurement scheme is described by a triple $(\mathcal{H}_0, \rho, M)$, where $\mathcal{H}_0$ is an ancilla, ρ a (pure) state on $\mathcal{H} \otimes \mathcal{H}_0$ and $M = \{M_1, \ldots, M_n\}$ is a POVM on $\mathcal{K} \otimes \mathcal{H}_0$. For $i, j \in \{1, \ldots, n\}$, the value

$$\operatorname{Tr} M_i (\Phi_j \otimes id)(\rho)$$

is interpreted as the probability that Φ_i is chosen when the true channel is Φ_j. The average success probability is then

$$p(M, \rho) = \sum_i \lambda_i \text{Tr} \, M_i (\Phi_i \otimes id)(\rho). \tag{2}$$

The task is to maximize this value over all triples $(\mathcal{H}_0, \rho, M)$.

It was observed[17,5,21,20] that using entangled input states may give greater success probability and it was shown in Ref. 19 that every entangled state is useful for some channel discrimination problem. However, there are situations when, e.g., the maximally entangled input state does not give an optimal success probability. It is therefore important to find out whether an optimal scheme with a given input state exists.

In the broader context of generalized decision problems, conditions for existence of an optimal scheme with an input state having maximal Schmidt rank were found in Ref. 16; a related problem was studied in Ref. 18. In the present paper, we show that these conditions can be obtained using the methods of semidefinite programming. Such methods were already applied before in the context of discrimination of quantum channels, see Refs. 7, 25, and 12. Compared to these works, we are more concerned with the choice of an optimal input state. It is an easy observation that if a given scheme $(\mathcal{H}_0, \rho, M)$ is optimal, then M must be an optimal measurement for the ensemble $\{\lambda_i, (\Phi_i \otimes id)(\rho)\}$. We show that, at least in the case that the input state is assumed to have maximal Schmidt rank, the optimality condition for a channel measurement can be divided into the condition (1) for this ensemble and an additional condition that ensures optimality of the input state. If the Schmidt rank of the input state is not maximal we obtain a comparably weaker result, but show an example where the use of such an input state is required.

As an important special case, we get a necessary and sufficient condition for existence of an optimal scheme with a maximally entangled input state. If this condition is not fulfilled, we give an upper bound on the optimal success probability. For discrimination of two channels, we use the known form of an optimal POVM for two states to obtain a relatively simple condition in terms of Choi matrices of the involved channels, which we call the Maximally Entangleg Input (MEI) condition. We also derive an upper bound on the diamond norm, which is tighter than the previously known bound given, e.g., in Ref. 4, see Remark 2. The results are applied to discrimination of covariant channels, qubit channels, unitary channels, and simple projective measurements.

The paper is organized as follows: in Sec. II we rewrite the problem as a problem of SDP from which we obtain necessary and sufficient conditions for optimal solution and derive an upper bound on the optimal success probability. In Section III we investigate the (MEI) condition and the related bounds. In Secs. IV and V, we study special cases of channels and present some examples demonstrating the results.

II. OPTIMALITY CONDITIONS

Let $\mathcal{H}$ be a finite dimensional Hilbert space. We denote by $B(\mathcal{H})^+$ the set of positive operators and by $\mathfrak{S}(\mathcal{H})$ the set of states, that is, positive operators of unit trace. A completely positive trace preserving map $\Phi : B(\mathcal{H}) \to B(\mathcal{K})$ is called a channel; we will denote the set of all channels by $C(\mathcal{H}, \mathcal{K})$. Any linear map $\Phi : B(\mathcal{H}) \to B(\mathcal{K})$ is represented by its Choi matrix $C(\Phi) \in B(\mathcal{K} \otimes \mathcal{H})$, defined in Ref. 9 as

$$C(\Phi) := (\Phi \otimes id)(|\psi_{\mathcal{H}}\rangle\langle\psi_{\mathcal{H}}|), \quad |\psi_{\mathcal{H}}\rangle = \sum_i |i\rangle \otimes |i\rangle,$$

where $\{|i\rangle\}$ is a fixed orthonormal basis of $\mathcal{H}$. Note that $\Phi \in C(\mathcal{H}, \mathcal{K})$ if and only if $C(\Phi)$ is positive and $\text{Tr}_{\mathcal{K}} C(\Phi) = I_{\mathcal{H}}$.

An alternative description of a channel measurement is given in terms of process POVMs[27] (or testers,[8] see also Ref. 13). A process POVM is a collection $F = \{F_1, \dots, F_n\}$ of positive operators in $B(\mathcal{K} \otimes \mathcal{H})$ with $\sum_i F_i = I \otimes \sigma$ for some state $\sigma \in \mathfrak{S}(\mathcal{H})$. For any triple $(\mathcal{H}_0, \rho, M)$, there is a process POVM F such that for all $\Phi \in C(\mathcal{H}, \mathcal{K})$ and $i = 1, \dots, n$,

$$\text{Tr} \, M_i (\Phi \otimes id)(\rho) = \text{Tr} \, C(\Phi) F_i. \tag{3}$$

Conversely, for any process POVM F, one can find some ancilla $\mathcal{H}_0$, a pure state ρ and a POVM M such that (3) holds.[27] To see this, let $\rho = |\psi\rangle\langle\psi|$, for $\psi \in \mathcal{H} \otimes \mathcal{H}_0$ and observe that by Schmidt

decomposition, we have $|\psi\rangle = \sum_i \lambda_i(I \otimes U)|i\rangle \otimes |i\rangle$ for some unitary operator U and $\lambda_i \geq 0$ for every i. Denoting $A = U \sum_i \lambda_i |i\rangle\langle i|$, we get

$$|\psi\rangle = (I \otimes A)|\psi_{\mathcal{H}}\rangle.$$

Since the channel in (3) acts only on the first part of the system we get

$$\mathrm{Tr}\,(\Phi \otimes id)(\rho)M_i = \mathrm{Tr}\,C(\Phi)(I \otimes A^*)M_i(I \otimes A) \tag{4}$$

and (3) holds with $F_i = (I \otimes A^*)M_i(I \otimes A)$. Conversely, let $\sum_i F_i = I \otimes \sigma$ for $\sigma \in \mathfrak{S}(\mathcal{H})$. Let $\sigma^{-1/2}$ be defined on the support of σ and 0 elsewhere, then $M_i = (I \otimes \sigma^{-1/2})F_i(I \otimes \sigma^{-1/2})$ is a POVM on $\mathcal{K} \otimes \mathcal{H}_0$ where now $\mathcal{H}_0 = \mathrm{supp}(\sigma)$ and (3) holds as before.

Using the description by process POVMs, we will show that the maximization of the success probability can be written as a problem of semidefinite programming,

$$\max_{F \in B(\mathbb{C}^n \otimes \mathcal{K} \otimes \mathcal{H})} \mathrm{Tr}\,CF$$

$$\text{s.t. } \mathrm{Tr}\,F = \dim(\mathcal{K}),$$

$$\mathrm{Tr}\,(I \otimes X_i)F = 0, \quad i = 1,\ldots,m,$$

$$F \geq 0.$$

Here $C = \sum_{i=1}^n |e_i^n\rangle\langle e_i^n| \otimes \lambda_i C(\Phi_i)$, $\{|e_i^n\rangle\}$ is the canonical basis of $\mathbb{C}^n$ and $X_1,\ldots,X_m$ is any basis of the (real) linear subspace

$$\mathcal{L} := \{X = X^* \in B(\mathcal{K} \otimes \mathcal{H}), \mathrm{Tr}_{\mathcal{K}}X = 0\}.$$

To see this, note that according to (3), Equation (2) can be rewritten as

$$p(M,\rho) = \sum_i \lambda_i \mathrm{Tr}\,C(\Phi_i)F_i.$$

Put $F := \sum_{i=1}^n |e_i^n\rangle\langle e_i^n| \otimes F_i \in B(\mathbb{C}^n \otimes \mathcal{K} \otimes \mathcal{H})$. Then

$$\mathrm{Tr}\,CF = \sum_i \lambda_i \mathrm{Tr}\,C(\Phi_i)F_i$$

and the problem of maximizing $p(M,\rho)$ can be understood as the problem of maximizing $\mathrm{Tr}\,CF$. We have $\mathrm{Tr}_{\mathbb{C}^n}F = \sum_i F_i = I \otimes \sigma$ and from $F_i \geq 0$ it follows $F \geq 0$. Note also that since C is block-diagonal, we may extend the maximization over all positive elements $F \in B(\mathbb{C}^n \otimes \mathcal{K} \otimes \mathcal{H})$ with $\mathrm{Tr}_{\mathbb{C}^n}F = I \otimes \sigma$, $\sigma \in \mathfrak{S}(\mathcal{H})$ (and not only over block-diagonal ones).

To rewrite this to the more usable form stated above, we need to note that $\mathrm{Tr}_{\mathbb{C}^n}F = I \otimes \sigma$ with $\mathrm{Tr}\,\sigma = 1$ if and only if $\mathrm{Tr}\,F(I \otimes X) = 0$ for all $X \in \mathcal{L}$ and $\mathrm{Tr}\,F = \dim(\mathcal{K})$. To prove this statement, let us first assume that $\mathrm{Tr}_{\mathbb{C}^n}F = I \otimes \sigma$, then for any X,

$$\mathrm{Tr}\,F(I \otimes X) = \mathrm{Tr}\,X\mathrm{Tr}_{\mathbb{C}^n}F = \mathrm{Tr}\,X(I \otimes \sigma)$$
$$= \mathrm{Tr}\,\sigma\mathrm{Tr}_{\mathcal{K}}X$$

and $\mathrm{Tr}\,F = \mathrm{Tr}\,\mathrm{Tr}_{\mathbb{C}^n}F = \mathrm{Tr}\,I \otimes \sigma = \dim(\mathcal{K})$.

Conversely, assume that $\mathrm{Tr}\,F(I \otimes X) = 0$ for all $X \in \mathcal{L}$ and $\mathrm{Tr}\,F = \dim(\mathcal{K})$. Consider $B(\mathcal{K})$ as a Hilbert space with Hilbert-Schmidt inner product, then there is an orthonormal basis $\{(\dim(\mathcal{K}))^{-1/2} I, \chi_{\mathcal{K},1},\ldots,\chi_{\mathcal{K},N}\}$ in $B(\mathcal{K})$, where each $\chi_{\mathcal{K},j}$ is a self-adjoint operator such that $\mathrm{Tr}\,\chi_{\mathcal{K},j} = 0$. With respect to this basis, each $X \in \mathcal{L}$ can be expressed as

$$X = I \otimes X_{\mathcal{H},0} + \sum_j \chi_{\mathcal{K},j} \otimes X_{\mathcal{H},j}$$

with some $X_{\mathcal{H},i} \in B(\mathcal{H})$. From the condition $\mathrm{Tr}_{\mathcal{K}}X = 0$ we obtain $X_{\mathcal{H},0} = 0$. Expressing $\mathrm{Tr}_{\mathbb{C}^n}F = I \otimes F_{\mathcal{H},0} + \sum_j \chi_{\mathcal{K},j} \otimes F_{\mathcal{H},j}$ and using the condition $\mathrm{Tr}\,F(I \otimes X) = 0$, we get

$$\mathrm{Tr}\,X_{\mathcal{H},j}F_{\mathcal{H},j} = 0, \qquad \forall j > 0.$$

Since there is no restriction on $X_{\mathcal{H},j}$ for $j > 0$, we must have $F_{\mathcal{H},j} = 0$ for all $j > 0$, and hence $\mathrm{Tr}_{\mathbb{C}^n}F = I \otimes F_{\mathcal{H},0} = I \otimes \sigma$. To conclude the proof, from the condition $\mathrm{Tr}\,F = \dim(\mathcal{K})$ we get

$\text{Tr} \, \sigma = 1$. Moreover, it is worth realizing that from the condition $F \geq 0$ we get $\sigma \geq 0$, hence $\sigma \in \mathfrak{S}(\mathcal{H})$.

The following result is obtained using standard methods of semidefinite programming (see, e.g., Ref. 3). The expression for maximal success probability was obtained also in Ref. 7, in a more general setting.

Theorem 1. *Let $\hat{F}$ be a process POVM. Then $\hat{F}$ is optimal if and only if there is some $\lambda_0 \geq 0$ and some $\Phi_0 \in C(\mathcal{H}, \mathcal{K})$, such that for all i,*

$$\lambda_i C(\Phi_i) \leq \lambda_0 C(\Phi_0)$$

and

$$(\lambda_0 C(\Phi_0) - \lambda_i C(\Phi_i))\hat{F}_i = 0, \qquad \forall i.$$

Moreover, in this case, the maximal success probability is

$$\text{Tr} \, \hat{F} C = \max_F \text{Tr} \, F C$$
$$= \min_{\Phi \in C(\mathcal{H}, \mathcal{K})} \min\{\lambda, \lambda_i C(\Phi_i) \leq \lambda C(\Phi), \forall i\}.$$

Proof. As first, we will formulate the dual problem. Let X_i, $i = 1, \ldots, m$ be some basis of $\mathcal{L}$ and let $y = (y_1, \ldots, y_m) \in \mathbb{R}^m$, then dual problem is

$$\min_{\lambda \in \mathbb{R}, \, y \in \mathbb{R}^m} \lambda$$

$$\text{s.t.} \ \sum_{i=1}^m y_i(I \otimes X_i) + \frac{\lambda}{\dim(\mathcal{K})} I \geq C.$$

Let $\lambda, y_1, \ldots, y_m$ be dual feasible, then since $\text{Tr} \, X_i = 0$ and $\text{Tr} \, C > 0$, we must have $\lambda > 0$. If we denote $\sum_{i=1}^m \frac{y_i}{\lambda} X_i + \frac{1}{\dim(\mathcal{K})} I =: C'$, then

$$\text{Tr}_{\mathcal{K}} C' = I,$$

and from $(I \otimes C') \geq \lambda^{-1} C \geq 0$, we obtain $C' \geq 0$. Hence there is some channel $\Phi \in C(\mathcal{K}, \mathcal{H})$, such that $C' = C(\Phi)$. From the condition $\lambda(I \otimes C(\Phi)) \geq C$, we obtain

$$\lambda C(\Phi) \geq \lambda_i C(\Phi_i), \tag{5}$$

for all i. From here we see that the dual problem may be formulated as

$$\min_{\Phi \in C(\mathcal{H}, \mathcal{K})} \min\{\lambda, \lambda_i C(\Phi_i) \leq \lambda C(\Phi), \forall i\}. \tag{6}$$

Now let $F' = \frac{1}{n \dim(\mathcal{H})} I$, then F' is a primal feasible plan. Moreover F' belongs to the interior of the cone of positive operators, therefore by Slater's condition we obtain that the duality gap is zero, in other words $\max \text{Tr} \, CF = \min \lambda$ or $\text{Tr} \, C\hat{F} = \lambda_0$, where by $\hat{F}$ we denote the primal optimal plan and by λ_0, y_0, Φ_0 we denote the dual optimal plan. Since $\hat{F}$ is feasible, we have $\lambda_0 = \sum_{i=1}^m y_{0,i} \text{Tr} \, \hat{F}(I \otimes X_i) + \frac{\lambda_0}{\dim(\mathcal{K})} \text{Tr} \, \hat{F} = \lambda_0 \text{Tr} \, (I \otimes C(\Phi_0)) \hat{F}$ and we get

$$\sum_i \text{Tr} \, (\lambda_0 C(\Phi_0) - \lambda_i C(\Phi_i)) \hat{F}_i = 0. \tag{7}$$

As $\lambda_0 C(\Phi_0) - \lambda_i C(\Phi_i) \geq 0$ and $\hat{F}_i \geq 0$, the sum may be zero if and only if all summands are zero. Moreover, trace of the product of two positive matrices is zero if and only if their product is zero. To see this let $A, B \geq 0$ and $\text{Tr} \, (AB) = 0$. We have $\text{Tr} \, AB = \text{Tr} \, ((A^{\frac{1}{2}} B^{\frac{1}{2}})^* A^{\frac{1}{2}} B^{\frac{1}{2}}) = 0$ and hence $A^{\frac{1}{2}} B^{\frac{1}{2}} = 0$ and $AB = 0$.

By the above argumentation, we get from (7)

$$(\lambda_0 C(\Phi_0) - \lambda_i C(\Phi_i)) \, \hat{F}_i = 0, \qquad \forall i. \tag{8}$$

On the other hand, condition (5) must hold for any dual feasible plan, but if for some primal and dual feasible plans condition (8) holds, then the duality gap for these plans is zero and they are optimal. This concludes the proof. □

Using this result, we can characterize optimality of measurement schemes with input states of maximal Schmidt rank.

Corollary 1. Let $\rho \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{H})$ be a pure state such that $\mathrm{Tr}_1 \rho =: \rho_2$ is invertible. Then a measurement scheme $(\mathcal{H}, \rho, M)$ is optimal if and only if

(i) $Z := \sum_i \lambda_i (\Phi_i \otimes id)(\rho) M_i$ *majorizes* $\lambda_i (\Phi_i \otimes id)(\rho)$ *for all i,*

(ii) $\mathrm{Tr}_{\mathcal{K}} Z \propto \rho_2$.

Proof. Let $\rho = |\psi\rangle\langle\psi|$ and let $A \in B(\mathcal{H})$ be the operator such that $|\psi\rangle = (I \otimes A)|\psi_{\mathcal{H}}\rangle$, so that the process POVM corresponding to $(\mathcal{H}, \rho, M)$ is given by $\hat{F}_i = (I \otimes A^*)M_i(I \otimes A)$, see (4). Note that by our assumptions, A is invertible, $\rho_2 = AA^*$ and $\sum_i \hat{F}_i = I \otimes A^*A$.

Assume that $\hat{F}$ is optimal, then by Theorem 1, there must be some $\lambda_0 > 0$ and $\Phi_0 \in C(\mathcal{H}, \mathcal{K})$ such that $\lambda_0 C(\Phi_0) \geq \lambda_i C(\Phi_i)$ and

$$(\lambda_0 C(\Phi_0) - \lambda_i C(\Phi_i))\hat{F}_i = 0, \qquad \forall i.$$

Summing up over i, we obtain

$$\lambda_0 C(\Phi_0)(I \otimes A^*A) = \sum_i \lambda_i C(\Phi_i)\hat{F}_i.$$

Multiplying the above equality by $(I \otimes A)$ from the left and by $(I \otimes A^{-1})$ from the right, we get using the above expression for $\hat{F}_i$,

$$\lambda_0 (I \otimes A)C(\Phi_0)(I \otimes A^*)$$
$$= \sum_i \lambda_i (I \otimes A)C(\Phi_i)(I \otimes A^*)M_i = Z.$$

The two conditions follow easily from this equality.

Assume conversely that the conditions (i) and (ii) are satisfied. Put $Z_0 = (I \otimes A^{-1})Z(I \otimes (A^*)^{-1})$, then (i) and (ii) imply that $Z_0 \geq 0$ and $\mathrm{Tr}_{\mathcal{K}} Z_0 \propto I$. It follows that there is some positive number λ_0 and $\Phi_0 \in C(\mathcal{H}, \mathcal{K})$ such that $Z_0 = \lambda_0 C(\Phi_0)$. Moreover, (i) implies that $\lambda_0 C(\Phi_0) \geq \lambda_i C(\Phi_i)$ for all i and

$$\lambda_0 C(\Phi_0)(I \otimes (A^*A)) = \sum_i \lambda_i C(\Phi_i)\hat{F}_i.$$

It follows that $\sum_i (\lambda_0 C(\Phi_0) - \lambda_i C(\Phi_i))\hat{F}_i = 0$ and this implies the optimality condition of Theorem 1, exactly as in its proof. □

Note that (i) is the optimality condition (1) for a POVM in discrimination of the ensemble $\{\lambda_i, \rho_i\}$, where $\rho_i = (\Phi_i \otimes id)(\rho)$. In other words, if $\hat{M}$ is an optimal POVM for this ensemble and

$$\hat{Z} := \sum_i \lambda_i (\Phi_i \otimes id)(\rho)\hat{M}_i,$$

the majorization $\hat{Z} \geq \lambda_i (\Phi_i \otimes id)(\rho)$ is satisfied. It follows that the existence of an optimal scheme with the given input state is equivalent to the condition (ii). Clearly, in this case, $(\mathcal{H}, \rho, \hat{M})$ is the optimal scheme and the optimal success probability is $p_{\mathrm{opt}} = \mathrm{Tr}\,\hat{Z}$.

Next, we show that the conditions of Corollary 1 are necessary for a general pure input state.

Corollary 2. Let $\rho \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{H})$ be a pure state such that $\mathrm{Tr}_1 \rho =: \rho_2$. Then a measurement scheme $(\mathcal{H}, \rho, M)$ is optimal only if the conditions (i) and (ii) from the previous corollary hold.

Proof. We will show that the measurement scheme is optimal for some problem with reduced input space. Let us denote by $\mathcal{H}_2$ the support of ρ_2. Since ρ is pure, it must be of the form $\rho = |\psi\rangle\langle\psi| = \sum_{i,j} \sqrt{\xi_i \xi_j}|i\rangle\langle j| \otimes |\tilde{i}\rangle\langle\tilde{j}|$ for some Schmidt decomposition of $|\psi\rangle$. From here we see that $|\psi\rangle \in \mathcal{H}_2' \otimes \mathcal{H}_2$, where $\mathcal{H}_2'$ is a subspace isomorphic to $\mathcal{H}_2$. Let Φ_i' be the restriction of Φ_i to $B(\mathcal{H}_2')$ and let P be the projection onto $\mathcal{K} \otimes \mathcal{H}_2$, then it is clear that $\Phi_i' \in C(\mathcal{H}_2', \mathcal{K})$; moreover, $(\mathcal{H}_2, \rho, PMP)$ defines an optimal measurement scheme for the reduced channels, with full Schmidt rank input state. The rest follows from the previous corollary. □

In general, the opposite implication does not hold. That is because if we limit the problem to some subspace $\mathcal{H}_{sub}$ of the original Hilbert space $\mathcal{H}$, then in general we do not have a guarantee that the optimal input state will be supported on a subspace of the form $\mathcal{H}_{sub} \otimes \mathcal{H}_{anc}$, or in other words we would have to maximize the average success probability over all choices of the subspace $\mathcal{H}_{sub}$. We demonstrate this by the following simple example.

Example 1. Let $\rho = |\psi\rangle\langle\psi| \otimes |\varphi\rangle\langle\varphi|$, where $|\psi\rangle, |\varphi\rangle \in \mathcal{H}$, and let $M_i = \tilde{M}_i \otimes I$, where $\tilde{M}_i$ is the optimal POVM for discrimination of the ensemble $\{\lambda_i, \Phi_i(|\psi\rangle\langle\psi|)\}$. By (1) we have

$$\tilde{Z} = \sum_i \lambda_i \Phi_i(|\psi\rangle\langle\psi|)\tilde{M}_i \geq \lambda_i \Phi_i(|\psi\rangle\langle\psi|)$$

and $Z = \tilde{Z} \otimes |\varphi\rangle\langle\varphi|$. It is easy to see that both conditions (i) and (ii) are satisfied, but as argued in Ref. 19, there are cases when entangled input states give strictly larger probability of success than any separable state, so that a scheme of the form $(\mathcal{H}, \rho, M)$ cannot be optimal.

It seems that optimality of input states strongly depends on the structure of the channels. In some cases it is even necessary to use an input state with lower Schmidt rank, because using maximal Schmidt rank input state would "waste" some normalization of the input state on parts of the channels where it is unnecessary, as will be demonstrated in Example 4. It is an open question whether some stronger conditions for general input states can be obtained. See also Refs. 22 and 20 for a discussion of a similar problem in the case of qubit Pauli channels.

We will next present an upper bound for p_{opt} in the case that condition (ii) is violated. We assume that the input state ρ is maximally entangled, but a similar bound can be obtained for any input state having a maximal Schmidt rank.

Theorem 2. *Let M be an optimal POVM for discrimination of the ensemble $\{\lambda_i, \dim(\mathcal{H})^{-1} C(\Phi_i)\}$ and let $Z = \sum_i \lambda_i C(\Phi_i) M_i$, $p_{MEI} = \dim(\mathcal{H})^{-1}\mathrm{Tr}\, Z$. Let $\|\cdot\|$ denote the operator norm. Then the optimal success probability p_{opt} satisfies*

$$p_{MEI} \leq p_{opt} \leq \|\mathrm{Tr}_{\mathcal{K}} Z\|.$$

Proof. Note that p_{MEI} is the largest success probability that can be obtained by the maximally entangled input state; this implies the first inequality. Further, note that we have $\lambda_i C(\Phi_i) \leq Z$ by optimality of the POVM M. If now $\lambda > 0$ and $\Phi \in C(\mathcal{H}, \mathcal{K})$ are such that $Z \leq \lambda C(\Phi)$, then λ, Φ correspond to a dual feasible plan, hence $p_{opt} \leq \lambda$ by (6). To obtain the tightest upper bound in this way, we put

$$\lambda_0' := \inf_{\Phi \in C(\mathcal{H},\mathcal{K})} \inf\{\lambda > 0, Z \leq \lambda C(\Phi)\}.$$

By the Choi isomorphism, there is some completely positive map $\xi : B(\mathcal{H}) \to B(\mathcal{K})$, such that $Z = C(\xi)$. As it was shown in Ref. 16 (see Corollary 2 and Section 3.1), $\lambda_0' = \|\xi\|_\diamond$, where the diamond norm is defined as

$$\|\xi\|_\diamond = \sup_{\tau \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{H})} \|(\xi \otimes id)(\tau)\|_1.$$

Moreover, since ξ is completely positive, this norm simplifies to

$$\lambda_0' = \|\xi\|_\diamond = \sup_{\psi \in \mathfrak{S}(\mathcal{H})} \mathrm{Tr}\, \xi(\psi) = \sup_{\psi \in \mathfrak{S}(\mathcal{H})} \mathrm{Tr}\, Z(I \otimes \psi)$$

$$= \sup_{\psi \in \mathfrak{S}(\mathcal{H})} \mathrm{Tr}\, \mathrm{Tr}_{\mathcal{K}}[Z]\psi = \|\mathrm{Tr}_{\mathcal{K}} Z\|.$$

$\square$

In general, the bound that we obtain in this way does not have to be meaningful, that is, it may happen that $\|\mathrm{Tr}_{\mathcal{K}} Z\| > 1$. But, as will be demonstrated by the examples in Sec. V, there are cases when the bound is meaningful, or even tight.

Remark 1. Note that if $\mathrm{Tr}_{\mathcal{K}} Z = cI$, then $c = p_{MEI}$ and the value of $\epsilon := \|p_{MEI}^{-1}\mathrm{Tr}_{\mathcal{K}} Z - I\|$ indicates how much the condition (ii) is violated. It is easy to see that $\|\mathrm{Tr}_{\mathcal{K}} Z\| \leq (1 + \epsilon)p_{MEI}$; this shows that if ϵ is small, the maximally entangled state is close to optimal.

III. DISCRIMINATION OF TWO CHANNELS BY MAXIMALLY ENTANGLED INPUT STATES

Let $n = 2$ and $\Phi_1, \Phi_2 \in C(\mathcal{H}, \mathcal{K})$. The following notation will be used throughout. Let $\lambda \in (0, 1)$, then we put

$$\Phi_\lambda = \lambda\Phi_1 - (1 - \lambda)\Phi_2$$

and

$$\Delta_\lambda = \lambda C(\Phi_1) - (1 - \lambda)C(\Phi_2) = C(\Phi_\lambda). \tag{9}$$

Let $\rho = \dim(\mathcal{H})^{-1}|\psi_\mathcal{H}\rangle\langle\psi_\mathcal{H}|$ be the maximally entangled state and consider any two-outcome POVM on $\mathcal{K} \otimes \mathcal{H}$, given by $\{M, I - M\}$ for some operator $0 \leq M \leq I$ on $\mathcal{K} \otimes \mathcal{H}$. The average success probability for the triple $(\mathcal{H}, \rho, M)$ as defined by Equation (2) is

$$p(M, \rho) = \frac{1}{\dim(\mathcal{H})} \text{Tr}\,\Delta_\lambda M + (1 - \lambda),$$

The optimal POVM is obtained if M is the projection onto the support of the positive part of Δ_λ. In this case,

$$Z = \sum_i \lambda_i C(\Phi_i) M_i = (1 - \lambda)C(\Phi_2) + (\Delta_\lambda)_+$$

and

$$p_{\text{MEI}} = \dim(\mathcal{H})^{-1}\text{Tr}\,Z = \frac{1}{2}(1 + \dim(\mathcal{H})^{-1}\text{Tr}\,|\Delta_\lambda|).$$

Corollary 3. An optimal measurement scheme $(\mathcal{H}, \rho, M)$ with a pure maximally entangled input state ρ exists if and only if the Choi operators satisfy

$$\text{Tr}_\mathcal{K}|\Delta_\lambda| \propto I. \tag{MEI}$$

Proof. By the remarks below Corollary 1, such a scheme exists if and only if $\text{Tr}_\mathcal{K} Z \propto I$, equivalently, $\text{Tr}_\mathcal{K}(\Delta_\lambda)_+ \propto I$. Since we always have $\text{Tr}_\mathcal{K}\Delta_\lambda \propto I$ and

$$(\Delta_\lambda)_+ = \frac{1}{2}(\Delta_\lambda + |\Delta_\lambda|),$$

the condition can be rewritten as stated.   □

The following corollary describes the upper bound of the optimal probability.

Corollary 4. We have the following bounds:

$$p_{\text{MEI}} \leq p_{\text{opt}} \leq \frac{1}{2}\left(1 + \|\text{Tr}_\mathcal{K}|\Delta_\lambda|\|\right).$$

If the condition (MEI) is satisfied, the inequalities become equalities.

Proof. We only have to note that if the (MEI) condition is satisfied, then $\dim(\mathcal{H})^{-1}\text{Tr}\,|\Delta_\lambda| = \|\text{Tr}_\mathcal{K}|\Delta_\lambda|\|$.   □

Remark 2. It is well known that p_{opt} is related to the diamond norm as $p_{opt} = \frac{1}{2}(1 + \|\Phi_\lambda\|_\diamond)$. To our knowledge, the only known bounds on the diamond norm in terms of the Choi matrices are the following

$$\dim(\mathcal{H})^{-1}\|C(\Phi_\lambda)\|_1 \leq \|\Phi_\lambda\|_\diamond \leq \|C(\Phi_\lambda)\|_1, \tag{10}$$

(see, e.g., Ref. 4 [Lemma 6]) which is quite coarse. As in Remark 1, we obtain from Corollary 4 the following new upper bound:

$$\|\Phi_\lambda\|_\diamond \leq \|\text{Tr}_\mathcal{K}|C(\Phi_\lambda)|\|$$
$$\leq (1 + \epsilon')\dim(\mathcal{H})^{-1}\|C(\Phi_\lambda)\|_1, \tag{11}$$

where $\epsilon' = \|\frac{\dim(\mathcal{H})}{\|C(\Phi_\lambda)\|_1}\text{Tr}_\mathcal{K}|C(\Phi_\lambda)| - I\|$. This shows that if (MEI) is nearly satisfied, the above bounds are quite precise.

To show that the upper bound given by (11) is better than the bound (10) we will show that in general

$$(1 + \epsilon') \dim(\mathcal{H})^{-1} \le 1.$$

We have

$$\epsilon' = \dim(\mathcal{H}) \left\| \frac{\mathrm{Tr}_{\mathcal{K}} |C(\Phi_\lambda)|}{\|C(\Phi_\lambda)\|_1} - \frac{1}{\dim(\mathcal{H})} I \right\|$$

$$\le \dim(\mathcal{H}) - 1$$

since $\frac{\mathrm{Tr}_{\mathcal{K}} |C(\Phi_\lambda)|}{\|C(\Phi_\lambda)\|_1}$ is a state. This implies the above inequality. We also see that this inequality is strict unless $\mathrm{Tr}_{\mathcal{K}} |C(\Phi_\lambda)|$ is of rank 1.

IV. APPLICATIONS

We apply the results of Sec. III to the problem of discrimination of covariant channels, unitary channels, qubit channels, and measurements. In the case of covariant channels and unital qubit channels, similar results were obtained in Ref. 18 for more general decision problems on families of quantum channels.

A. Covariant channels

Let $\mathcal{U}(\mathcal{H})$ denote the unitary group of $\mathcal{H}$. For $U \in \mathcal{U}(\mathcal{H})$, let

$$Ad_U(A) := UAU^*, \quad A \in B(\mathcal{H}).$$

Let G be a group and let $g \mapsto U_g \in \mathcal{U}(\mathcal{H})$ and $g \mapsto V_g \in \mathcal{U}(\mathcal{K})$ be unitary representations. Assume that Φ_1 and Φ_2 are covariant channels, that is,

$$\Phi_i \circ Ad_{U_g} = Ad_{V_g} \circ \Phi_i, \qquad i = 1, 2, \; g \in G. \tag{12}$$

Irreducibility of $g \mapsto U_g$ plays a strong role, as we will see. In this case, the only non-zero projection that commutes with all U_g is I, see, e.g., Ref. 2.

Proposition 1. Let Φ_1, Φ_2 be channels satisfying (12). Assume that the representation $g \mapsto U_g$ is irreducible. Then the condition (MEI) is satisfied for any $\lambda \in (0, 1)$.

Proof. Let U^t denote the transpose of U with respect to the fixed basis $\{|i\rangle\}$. Let $\lambda \in (0, 1)$ and let $\Delta_\lambda = C(\Phi_\lambda)$ be as in (9). We will prove the proposition by showing that $\mathrm{Tr}_{\mathcal{K}} |\Delta_\lambda|$ is invariant under $Ad_{U_g^t}$ and by the discussion above, this implies that $\mathrm{Tr}_{\mathcal{K}} |\Delta_\lambda| \propto I$. For every $g \in G$ we have

$$Ad_{U_g^t}(\mathrm{Tr}_{\mathcal{K}} |\Delta_\lambda|) =$$

$$= \mathrm{Tr}_{\mathcal{K}} (id \otimes Ad_{U_g^t})(|\Delta_\lambda|)$$

$$= \mathrm{Tr}_{\mathcal{K}} |(\Phi_\lambda \otimes Ad_{U_g^t})(|\psi_\mathcal{H}\rangle\langle\psi_\mathcal{H}|)|$$

$$= \mathrm{Tr}_{\mathcal{K}} |(\Phi_\lambda \circ Ad_{U_g} \otimes id)(|\psi_\mathcal{H}\rangle\langle\psi_\mathcal{H}|)|$$

$$= \mathrm{Tr}_{\mathcal{K}} (Ad_{V_g} \otimes id)(|\Delta_\lambda|) = \mathrm{Tr}_{\mathcal{K}} |\Delta_\lambda|.$$

$\square$

In case the representation U is reducible, let us sketch an upper bound of the optimal probability. By the previous proof, we have $Ad_{U_g^t}(\mathrm{Tr}_{\mathcal{K}} |\Delta_\lambda|)$, hence

$$\mathrm{Tr}_{\mathcal{K}} |\Delta_\lambda| = \sum_i k_i P_i^t,$$

where $k_i \in \mathbb{R}$ and P_i are projections onto the subspaces of the irreducible representations, orthogonal sum of which is U. Let $t_i = \mathrm{Tr}\, P_i$ then $p_{\mathrm{MEI}} = \frac{1}{2}(1 + \dim(\mathcal{H})^{-1} \sum_i t_i k_i)$ and we have

$$p_{\mathrm{opt}} \le \frac{1}{2} \left(1 + \max_i k_i \right).$$

B. Qubit channels

Let $\mathcal{H} = \mathcal{K} = \mathbb{C}^2$ and let us denote $\psi_{\mathcal{H}} =: \psi_2$. Let $\Gamma(X) = (\mathrm{Tr}\,X)I - X^t$ be the Werner-Holevo channel, where X^t denotes the transpose map with respect to the canonical basis $|0\rangle, |1\rangle$. Then Γ is a unitary channel, given by the unitary U such that

$$U|0\rangle = -|1\rangle, \qquad U|1\rangle = |0\rangle.$$

It can be easily checked that $\Gamma \circ \Gamma = id$ and $(id \otimes \Gamma)(|\psi_2\rangle\langle\psi_2|) = (\Gamma \otimes id)(|\psi_2\rangle\langle\psi_2|)$. If $\phi : B(\mathbb{C}^2) \to B(\mathbb{C}^2)$ is a linear map such that there is some $a \in \mathbb{R}$, satisfying

$$\mathrm{Tr}\,\phi(X) = a\mathrm{Tr}\,X, \qquad X \in B(\mathbb{C}^2), \tag{13}$$

then

$$\phi \circ \Gamma(X) = \Gamma \circ \phi^t(X) + (\mathrm{Tr}\,X)(\phi(I) - aI),$$

where $\phi^t(X) = \phi(X^t)^t$. Moreover, for a self-adjoint $X \in B(\mathbb{C}^2)$, $\Gamma(X) = X^t$ if and only if $X \propto I$.

Let Φ_1 and Φ_2 be two qubit channels and let $\Delta_\lambda = C(\Phi_\lambda)$ for $\lambda \in (0, 1)$ as before. By the previous remarks, the condition (MEI) is equivalent to $\Gamma(\mathrm{Tr}_{\mathcal{K}}|\Delta_\lambda|) = (\mathrm{Tr}_{\mathcal{K}}|\Delta_\lambda|)^t$. We are now going to investigate this equality.

Note that Φ_λ satisfies (13) with $a = 2\lambda - 1$. Since Γ is a unitary channel, we have

$$\Gamma(\mathrm{Tr}_{\mathcal{K}}|\Delta_\lambda|) = \mathrm{Tr}_{\mathcal{K}}(id \otimes \Gamma)(|\Delta_\lambda|)$$

$$= \mathrm{Tr}_{\mathcal{K}}|(\Phi_\lambda \otimes \Gamma)(|\psi_2\rangle\langle\psi_2|)|.$$

Using further properties of Γ and Φ_λ, we get

$$\mathrm{Tr}_{\mathcal{K}}\,|(\Phi_\lambda \otimes \Gamma)(|\psi_2\rangle\langle\psi_2|)$$

$$= \mathrm{Tr}_{\mathcal{K}}|(\Phi_\lambda \circ \Gamma \otimes id)(|\psi_2\rangle\langle\psi_2|)|$$

$$= \mathrm{Tr}_{\mathcal{K}}|(\Gamma \circ \Phi_\lambda^t \otimes id)(|\psi_2\rangle\langle\psi_2|)$$

$$\quad + (\Gamma \circ \Gamma)(\Phi_\lambda(I) - (2\lambda - 1)I) \otimes I|$$

$$= \mathrm{Tr}_{\mathcal{K}}|(\Phi_\lambda^t \otimes id)(|\psi_2\rangle\langle\psi_2|)$$

$$\quad + ((2\lambda - 1)I - \Phi_\lambda^t(I)) \otimes I|$$

$$= [\mathrm{Tr}_{\mathcal{K}}|\Delta_\lambda + ((2\lambda - 1)I - \Phi_\lambda(I)) \otimes I|]^t.$$

The last equality follows from the fact that $C(\Phi_\lambda^t) = C(\Phi_\lambda)^{t \otimes t}$, where $t \otimes t$ denotes transpose with respect to the product basis $|i\rangle \otimes |j\rangle$, and that $\mathrm{Tr}_{\mathcal{K}}|X^{t \otimes t}| = (\mathrm{Tr}_{\mathcal{K}}|X|)^t$ for any $X = X^* \in B(\mathbb{C}^2 \otimes \mathbb{C}^2)$. Thus we have proved.

Proposition 2. For a pair of qubit channels, the condition (MEI) holds if and only if

$$\mathrm{Tr}_{\mathcal{K}}|\Delta_\lambda + ((2\lambda - 1)I - \Phi_\lambda(I)) \otimes I| = \mathrm{Tr}_{\mathcal{K}}|\Delta_\lambda|.$$

In particular, this is true if $\Phi_\lambda(I) = (2\lambda - 1)I$. If both channels are unital, this holds for any λ, hence maximally entangled input state is optimal, as it was already observed in Ref. 18 and in Ref. 22 in the case of qubit Pauli channels. If $\lambda = \frac{1}{2}$, then the condition (MEI) is satisfied if $\Phi_1(I) = \Phi_2(I)$, even if the channels are not unital.

C. Unitary channels

Let $U, V \in \mathcal{U}(\mathcal{H})$ and let $\Phi_1 = Ad_U$, $\Phi_2 = Ad_V$ be the corresponding unitary channels. As it was proved in Ref. 6, it is not necessary to use entangled inputs for optimal discrimination of two unitary channels. Nevertheless, it is an interesting question whether a maximally entangled state is also optimal; this will be addressed in this paragraph.

Let $W = UV^*$. Since any input state ρ may be replaced by $(V^* \otimes I)\rho(V \otimes I)$, it is clear that discrimination of Ad_U and Ad_V is equivalent to discrimination of Ad_W and the identity channel, and that a maximally entangled input state is optimal for one problem if and only if it is optimal for the other. We may therefore assume that $\Phi_1 = Ad_W$ and $\Phi_2 = id$. Since the unitaries are given only up to a phase, we may also assume that $\operatorname{Tr} W \in \mathbb{R}$. Put

$$|\phi\rangle = \sum_i W|i\rangle \otimes |i\rangle,$$

$$|\psi\rangle = \sum_i |i\rangle \otimes |i\rangle,$$

so that $|\phi\rangle\langle\phi| = C_1$, $|\psi\rangle\langle\psi| = C_2$ are the Choi matrices of the unitary channel Ad_W and identity. By the results of the Appendix it is clear that $\operatorname{Tr}_1|\Delta_\lambda| \propto I$ if and only if $z\operatorname{Tr}_1(|\phi\rangle\langle\psi| + |\psi\rangle\langle\phi|) \propto I$, where

$$z = \langle\phi,\psi\rangle = \operatorname{Tr} W^* = \operatorname{Tr} W$$

and

$$\operatorname{Tr}_1|\phi\rangle\langle\psi| = W^t.$$

Since the transpose is a linear map and $I^t = I$, we see that (MEI) is equivalent to

$$(\operatorname{Tr} W)(W + W^*) \propto I.$$

If $\operatorname{Tr} W = 0$ this condition is obviously satisfied. If $\operatorname{Tr} W \neq 0$ it is equivalent to

$$W + W^* \propto I. \tag{14}$$

The unitary W has a spectral decomposition

$$W = \sum_{j=1}^{\dim(\mathcal{H})} (\cos(\alpha_j) + i\sin(\alpha_j))|\xi_j\rangle\langle\xi_j|,$$

where $\alpha_1, \ldots, \alpha_{\dim(\mathcal{H})} \in [-\pi, \pi]$. From the condition (14) we see that $\cos(\alpha_j)$ must be constant with respect to j, or in other words there must exist $\beta \in [0, \pi]$ and numbers $\eta_j \in \{0, 1\}$ such that $\alpha_j = (-1)^{\eta_j}\beta$ for every j. By the assumption $\operatorname{Tr} W \in \mathbb{R}$ we must have

$$\sum_{j=1}^{\dim(\mathcal{H})} \sin(\alpha_j) = \sin(\beta) \sum_{j=1}^{\dim(\mathcal{H})} (-1)^{\eta_j} = 0.$$

This implies that either $\beta = 0$ and $W = I$ or W has exactly two eigenvalues, each of the same multiplicity. The fact that in our calculation the eigenvalues are complex conjugate of each other is simply caused by the choice $\operatorname{Tr} W \in \mathbb{R}$ and does not have to be generally required. We have proved the following.

Proposition 3. Let $\Phi_1 = Ad_U$, $\Phi_2 = Ad_V$ be unitary channels. Put $W = UV^*$ and let $\lambda \in (0, 1)$. Then (MEI) holds if and only if either $\operatorname{Tr} W = 0$ or W has at most two different eigenvalues, each of the same multiplicity.

Note that if $\dim(\mathcal{H})$ is odd, (MEI) holds iff $\operatorname{Tr} W = \langle\psi,\phi\rangle = 0$, in which case the two channels are perfectly distinguishable.

D. Simple projective measurements

A special case of a channel is a measurement, which is given by a POVM $M = \{M_1, \ldots, M_m\}$. One shot discrimination of quantum measurements was investigated in Ref. 23, where it was proved that entangled input states are necessary in some cases.

The corresponding channel $\Phi_M : B(\mathcal{H}) \to B(\mathbb{C}^m)$ is defined as

$$A \mapsto \sum_i (\operatorname{Tr} M_i A)|i\rangle\langle i|$$

and the Choi matrix has the form $C(\Phi_M) = \sum_i |i\rangle\langle i| \otimes M_i^t$. Let $\Phi_1 = \Phi_M$, $\Phi_2 = \Phi_N$ for two POVM's M, N with m outcomes. In this case, the condition (MEI) has the form

$$\sum_i |\lambda M_i - (1-\lambda)N_i| \propto I.$$

We will further investigate simple projective measurements. Let $\{|\xi_i\rangle\}$ and $\{|\eta_i\rangle\}$ be two orthonormal bases in $\mathcal{H}$ and let $M_i = P_{\xi_i} := |\xi_i\rangle\langle\xi_i|$ and $N_i = P_{\eta_i} := |\eta_i\rangle\langle\eta_i|$. We will also assume that $\lambda = \frac{1}{2}$.

Proposition 4. Assume that $P_{\xi_j} = P_{\eta_j}$ for some j. Then the condition (MEI) is satisfied if and only if $M = N$.

Proof. In this case, the condition is

$$\sum_i |P_{\xi_i} - P_{\eta_i}| = \sum_{i\neq j} |P_{\xi_i} - P_{\eta_i}| \propto I.$$

Since $|P_{\xi_i} - P_{\eta_i}| \leq I - P_{\xi_j} = I - P_{\eta_j}$ for all $i \neq j$, this can be true only if $\sum |P_{\xi_i} - P_{\eta_i}| = 0$, that is, $M = N$. □

From now on we will always assume that $|\langle\xi_i,\eta_i\rangle| < 1$, equivalently, $P_{\xi_i} \neq P_{\eta_i}$, for all i. Then $|P_{\xi_i} - P_{\eta_i}| = c_i P_{\xi_i,\eta_i}$, where $c_i = (1 - |\langle\xi_i,\eta_i\rangle|^2)^{1/2}$ and P_{ξ_i,η_i} is the projection onto span$\{\xi_i,\eta_i\}$, so that the condition (MEI) becomes

$$\sum_i c_i P_{\xi_i,\eta_i} = dI, \quad d = 2\dim(\mathcal{H})^{-1}\sum_i c_i. \tag{15}$$

Note that if $\dim(\mathcal{H}) = 2$, then $P_{\xi_i,\eta_i} = I$, so that the condition trivially holds. In this case, Φ_M and Φ_N are unital qubit channels, so that this follows also by previous results. Put

$$|\xi_i^\perp\rangle = c_i^{-1}(|\eta_i\rangle - \langle\xi_i,\eta_i\rangle|\xi_i\rangle), \quad i = 1,\ldots,\dim(\mathcal{H}).$$

Then $P_i := P_{\xi_i,\eta_i} = P_{\xi_i,\xi_i^\perp} = P_{\xi_i} + P_{\xi_i^\perp}$. The condition (15) is equivalent to

$$d|\xi_j\rangle = \sum_i c_i P_i|\xi_j\rangle = c_j|\xi_j\rangle + \sum_{i\neq j} c_i\langle\xi_i^\perp,\xi_j\rangle|\xi_i^\perp\rangle$$

for all j, or

$$(d - c_j)\langle\xi_k,\xi_j\rangle = \sum_i c_i\langle\xi_i^\perp,\xi_j\rangle\langle\xi_k,\xi_i^\perp\rangle, \tag{16}$$

for all j,k. Note that the diagonal matrix $dI - C$, where $C = diag(c_1,\ldots,c_n)$, is invertible. Indeed, $d - c_j = 0$ for some j implies that

$$d - c_j = \sum_{i\neq j} c_i^{-1}|\langle\eta_i,\xi_j\rangle|^2 = 0,$$

so that $\langle\eta_i,\xi_j\rangle = 0$ for all $i \neq j$. But then $|\langle\eta_j,\xi_j\rangle| = 1$, which is a contradiction. Hence $d - c_j > 0$ for all j and $dI - C$ is positive definite.

Let us begin from (16). Let us denote $W_{ij} = \langle\xi_i,\eta_j\rangle$. Clearly W is a unitary matrix. It is straightforward to see that

$$\langle\xi_i,\xi_j^\perp\rangle = (1 - \delta_{ij})W_{ij}c_j^{-1},$$

so the condition (16) becomes

$$(d - c_j)\delta_{kj} = \sum_i (1 - \delta_{ki})W_{ki}c_i^{-1}(1 - \delta_{ij})W_{ij}^*, \tag{17}$$

which can be written as a matrix equation of the form

$$dI - C = (W - diag(W))C^{-1}(W^* - diag(W^*)). \tag{18}$$

At this point we are ready to settle the case $\dim(\mathcal{H}) = 3$.

Proposition 5. Let $\dim(\mathcal{H}) = 3$*, then the ([MEI](#)) condition holds if and only if the matrix W, defined as above, has of one of the following forms*

$$W_1 = \begin{pmatrix} 0 & 0 & e^{i\varphi_1} \\ e^{i\varphi_2} & 0 & 0 \\ 0 & e^{i\varphi_3} & 0 \end{pmatrix},$$

$$W_2 = \begin{pmatrix} 0 & e^{i\varphi_1} & 0 \\ 0 & 0 & e^{i\varphi_2} \\ e^{i\varphi_3} & 0 & 0 \end{pmatrix},$$

for $\varphi_i \in \mathbb{R}$*. In other words, one POVM is a cyclic permutation of the other.*

Proof. Consider Equation ([17](#)) and let $j \neq k$, then we obtain

$$0 = W_{ki} c_i^{-1} \bar{W}_{ji}, \tag{19}$$

for $i \neq j \neq k \neq i$. By putting $i = 1, 2, 3$, we obtain $0 = W_{31}\bar{W}_{21} = W_{12}\bar{W}_{23} = W_{23}\bar{W}_{13}$. It follows that some off-diagonal elements of the matrix W must be zero. On the other hand, for $k = j$ Equation ([17](#)) becomes

$$d - c_j = \sum_{i \neq j} W_{ji} c_i^{-1} \bar{W}_{ji} \tag{20}$$

which shows that some off-diagonal elements must be non-zero. Putting, say, $W_{12} = 0$, the above equalities imply that also $W_{23} = W_{31} = 0$ and all other off-diagonal elements are nonzero. From the condition $WW^* = I$, we obtain that $W_{11} = W_{22} = W_{33} = 0$ and $|W_{21}| = |W_{32}| = |W_{13}| = 1$, this implies that W is of the form W_1 and the basis $|\xi_i\rangle$ and $|\eta_j\rangle$ are just cyclic permutations of each other, modulo phase change. Similarly, assuming that $W_{12} \neq 0$, we obtain that W is of the form W_2, which is just the other possible cyclic permutation of the basis, modulo complex phase.

For the converse, it is easy to check that both W_1 and W_2 satisfy ([18](#)). See also the remark following the proof of Proposition 7. □

The basis vectors $\{|\xi_i\rangle\}$ and $|\eta_i\rangle$ are given only up to a phase, note that also the condition ([18](#)) is invariant under phase changes. We may therefore assume that $\langle \xi_i, \eta_i \rangle \geq 0$ for all i. Then the matrix $diag(W)$ has only real non-negative entries, hence it is self-adjoint.

From $c_i = \sqrt{1 - |\langle \xi_i, \eta_i \rangle|^2} = \sqrt{1 - W_{ii}^2}$, we can express $W_{ii} = \sqrt{1 - c_i^2}$ and hence $diag(W) = \sqrt{I - C^2}$. Since both C and $diag(W)$ are self-adjoint matrices, their functions are well defined by the spectral theorem. Multiplying out the right hand side of Equation ([18](#)) and using the obtained expression for $diag(W)$ we obtain

$$dI = WC^{-1}W^* + C^{-1} - \left(C^{-1}\sqrt{I - C^2}W^* + WC^{-1}\sqrt{I - C^2} \right). \tag{21}$$

Clearly the operators $WC^{-1}W^* + C^{-1}$ and $C^{-1}\sqrt{I - C^2}W^* + WC^{-1}\sqrt{I - C^2}$ are self-adjoint. Let us denote $\mathcal{L}_S(\mathcal{H})$ the real linear space of self-adjoint operators on $\mathcal{H}$, endowed with a Hilbert-Schmidt scalar product, which we denote as $(\cdot, \cdot)_{HS}$. Let us choose an orthonormal basis on $\mathcal{L}_S$ comprised of the operator $\dim(\mathcal{H})^{-\frac{1}{2}}I$ and of operators X_i, such that $\text{Tr}(X_i) = 0$.

From Equation ([21](#)) we can see that

$$WC^{-1}W^* + C^{-1} = \beta \dim(\mathcal{H})^{-\frac{1}{2}}I + L, \tag{22}$$

$$C^{-1}\sqrt{I - C^2}W^* + WC^{-1}\sqrt{I - C^2} = b \dim(\mathcal{H})^{-\frac{1}{2}}I + L, \tag{23}$$

where L is some real linear combination of X_i, hence $\text{Tr } L = 0$ and $\beta, b \in \mathbb{R}$. It is easy to compute β as follows:

$$\beta = (\dim(\mathcal{H})^{-\frac{1}{2}}I, WC^{-1}W^* + C^{-1})_{HS} = 2\text{Tr}\left(\dim(\mathcal{H})^{-\frac{1}{2}}C^{-1}\right) = 2\dim(\mathcal{H})^{-\frac{1}{2}}\sum_i c_i^{-1}.$$

Now it is easy to see, that:

$$b = 2\dim(\mathcal{H})^{-\frac{1}{2}} \sum_i \left(c_i^{-1} - c_i\right)$$

because from Equation (21) we see that $\beta - b = 2\dim(\mathcal{H})^{-\frac{1}{2}} \sum_i c_i$ as $d = 2(\dim(\mathcal{H}))^{-1} \sum_i c_i$.

We have proved the following.

Proposition 6. Let $M_i = P_{\xi_i}$ and $N_i = P_{\eta_i}$ for some orthonormal bases $|\xi_i\rangle$ and $|\eta_j\rangle$. Then (MEI) condition is equivalent to Equations (22) and (23).

The obtained equations look rather complicated, but they yield some results in specific cases.

Proposition 7. With the above notations, assume that $c_i = c$ for all i. Then

(i)   *if $c \neq 1$ and $\dim(\mathcal{H})$ is odd, then (MEI) cannot be satisfied;*
(ii)   *if $c \neq 1$ and $\dim(\mathcal{H})$ is even, then (MEI) is satisfied if and only if $W = \sqrt{1 - c^2}I + icG$; where G is a symmetric unitary matrix with zero diagonal.*
(iii) *if $c = 1$, then (MEI) always holds.*

Proof. Let us assume that $c_i = c \neq 1$ for all i. Immediately we see that $C^{-1} = c^{-1}I$ and $\sqrt{I - C^2} = \sqrt{1 - c^2}I$. As a direct consequence of this, we obtain $L = 0$ from Equation (22). Using this as our advantage in Equation (23) we obtain that in this case, (MEI) is equivalent to

$$W^* + W = 2\sqrt{1 - c^2}I. \tag{24}$$

As in the proof of Proposition 3, we see that W has exactly two eigenvalues with the same multiplicity, this also implies (i). If $\dim(\mathcal{H})$ is even, it is clear that these eigenvalues must be equal to $\lambda_\pm := \sqrt{1 - c^2} \pm ic$. Let P be the eigenprojection corresponding to λ_+, then $W = \sqrt{1 - c^2}I + icG$, with $G = 2P - I$. Conversely, it is easy to see that if W is of this form, then (24) holds; this finishes the proof of (ii).

Now assume that $c_i = 1$, which means that $\langle \xi_i, \eta_i \rangle = 0$ for all i. Again we immediately see that conditions (22) and (23) are satisfied, with $L = 0$. This implies (iii).   □

Note that we can conclude from the last statement that any two bases, such that the unitary matrix mapping one basis to the other is hollow satisfy the condition (MEI). Especially, this happens if the basis $|\eta_j\rangle$ is a permutation of $|\xi_j\rangle$, leaving no element fixed.

A particular case of the situation described in Proposition 7 is when the bases are mutually unbiased, that is when $|\langle \xi_i, \eta_j \rangle| = \dim(\mathcal{H})^{-1/2}$ for all i, j. Then $H = \sqrt{dim(\mathcal{H})}W$ is a Hadamard matrix, see, e.g., Ref. 24 for more information on complex Hadamard matrices. As we have seen, such bases can satisfy the condition (MEI) only in even-dimensional Hilbert spaces. Below, we provide a further result for $\dim(\mathcal{H}) = 4$. Recall that two Hadamard matrices are equivalent if one can be turned into the other by multiplication by diagonal unitaries and permutations. Up to equivalence, any Hadamard matrix can be turned into a dephased form, with all elements in the first row and column equal to unity. In this way, any four-dimensional Hadamard matrix is equivalent to a member of a one-parameter family, containing a unique matrix $H_{\mathbb{R}}$ with real entries. Since the vectors $|\xi_i\rangle$ and $\eta_i\rangle$ are given only up to a phase, the next result shows that this is the only case when (MEI) is satisfied.

Proposition 8. Let $\dim(\mathcal{H}) = 4$, and let the bases $|\xi_j\rangle$ and $|\eta_j\rangle$ be mutually unbiased. Then the condition (MEI) is satisfied if and only if the corresponding Hadamard matrix satisfies $2W = D_1 H_{\mathbb{R}} D_2$, where D_1, D_2 are diagonal unitaries and

$$H_{\mathbb{R}} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \end{pmatrix}.$$

Proof. It is clear that $c_i = c = \sqrt{3}/2$. By Proposition 7, we see that we need to search for a 4×4 hollow symmetric unitary matrix G such that each off-diagonal element has modulus $1/\sqrt{3}$. A general form for such a matrix can be found by a straightforward computation. We find that

$$2W = I + i\sqrt{3}G = \begin{pmatrix} 1 & \bar{b}c & i\bar{b} & i\bar{a} \\ -b\bar{c} & 1 & i\bar{c} & -i\bar{a}b\bar{c} \\ ib & ic & 1 & \bar{a}b \\ ia & -ia\bar{b}c & -a\bar{b} & 1 \end{pmatrix} \tag{25}$$

for $|a| = |b| = |c| = 1$, so that $2W = D_1 H_{\mathbb{R}} D_2$, where $D_1 = diag(1, \bar{b}c, i\bar{b}, i\bar{a})$ and $D_2 = diag(1, -b\bar{c}, ib, ia)$.

Conversely, suppose that $2W = D_1 H_{\mathbb{R}} D_2$ for diagonal unitaries D_1 and D_2. Since we also assume that all diagonal elements of W are equal to 1, we obtain that we must have $D_1 = diag(d_1, d_2, d_3, d_4)$ and $D_2 = diag(\bar{d}_1, -\bar{d}_2, -\bar{d}_3, -\bar{d}_4)$. It is now easy to check that $2W$ has the form (25), with $a = -i\bar{d}_1 d_4$, $b = -i\bar{d}_1 d_3$ and $c = i\bar{d}_2 d_3$. □

V. EXAMPLES

In this section, we present examples based on the results of Secs. I–IV.

A. Qubit channels

To underline how maximally entangled input states may be used for discrimination of two qubit channels, we will present an example of discrimination of identity channel ϕ_{id} and amplitude damping channel ϕ_{AD}. The amplitude damping channels are not unital and hence (MEI) is not necessarily satisfied, and we will see that the maximally entangled state is indeed not optimal. The identity channel may be replaced by a unitary channel with some changes to the following calculations. We will set $\lambda = \frac{1}{2}$.

Example 2. Let $\mathcal{H}$ denote complex Hilbert space with $\dim(\mathcal{H}) = 2$ and $|0\rangle, |1\rangle$ will denote some orthonormal basis of $\mathcal{H}$. The amplitude damping channel is represented by Kraus operators A_θ, B_θ, defined as

$$A_\theta = |0\rangle\langle 0| + \sqrt{1-\theta}|1\rangle\langle 1|,$$
$$B_\theta = \sqrt{\theta}|0\rangle\langle 1|,$$

where $\theta \in [0,1]$ is a parameter. Note that for $\theta = 0$ the amplitude damping channel becomes the identity. We will proceed as follows: we will find the Choi matrices corresponding to ϕ_{id} and ϕ_{AD} to obtain p_{MEI} and the upper bound on the optimal success probability. Then we will find p_{opt} to compare it with the upper bound.

Let $C_{id} = C(\phi_{id})$ and $C_{AD} = C(\phi_{AD})$. We will be interested in the matrix $\Delta_{\frac{1}{2}} = \frac{1}{2}(C_{id} - C_{AD})$ which is of the form

$$\Delta_{\frac{1}{2}} = \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 & 1-\sqrt{1-\theta} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -\theta & 0 \\ 1-\sqrt{1-\theta} & 0 & 0 & \theta \end{pmatrix}.$$

Since $p_{\text{MEI}} = \frac{1}{2}(1 + \frac{1}{2}\text{Tr}|\Delta_{\frac{1}{2}}|)$, we can already find p_{MEI} as a function of θ. It is easy to see that 0 and $-\frac{1}{2}\theta$ are eigenvalues of $\Delta_{\frac{1}{2}}$. Finding the other two eigenvalues is easy since the problem reduces to finding eigenvalues of a 2×2 matrix, they are $\frac{1}{4}(\theta + \sqrt{\theta^2 + 4(1-\sqrt{1-\theta})^2})$ and

$\frac{1}{4}\big(\theta - \sqrt{\theta^2 + 4(1 - \sqrt{1-\theta})^2}\,\big)$. Let us denote

$$\lambda_1 = \frac{1}{2}\big(\theta + \sqrt{\theta^2 + 4(1 - \sqrt{1-\theta})^2}\,\big),$$

$$\lambda_2 = \frac{1}{2}\big(\theta - \sqrt{\theta^2 + 4(1 - \sqrt{1-\theta})^2}\,\big),$$

then we have $p_{\mathrm{MEI}} = \frac{1}{2}(1 + \frac{1}{2}\lambda_1)$. Moreover after a tedious calculation it can be seen that

$$\mathrm{Tr}_1 |\Delta_{\frac{1}{2}}| = \frac{\lambda_1}{2} I - \frac{\lambda_1^2 + \lambda_1\lambda_2}{2(\lambda_1 - \lambda_2)} |0\rangle\langle 0| + \frac{\lambda_1^2 + \lambda_1\lambda_2}{2(\lambda_1 - \lambda_2)} |1\rangle\langle 1|,$$

and we have $p_{\mathrm{opt}} \le \frac{1}{2} \| I + \mathrm{Tr}_1 |\Delta_{\frac{1}{2}}| \|$. Since $\lambda_1 \ge 0$ and $\lambda_1 \pm \lambda_2 \ge 0$, we have

$$p_{\mathrm{opt}} \le \frac{1}{2} + \frac{\lambda_1}{4}\left(1 + \frac{\theta}{\sqrt{\theta^2 + 4(1 - \sqrt{1-\theta})^2}}\right).$$

To verify the upper bound we will find p_{opt}, for which we only need to find the value of $\frac{1}{2}\|\phi_{id} - \phi_{AD}\|_{\diamond}$, which we have done by numeric methods. According to Ref. 25 the problem of computing diamond norm can be formulated the following SDP problem:

$$\max_{X} \frac{1}{2}\big(\mathrm{Tr}\,(\Delta_{\frac{1}{2}}^* X) + \mathrm{Tr}\,(\Delta_{\frac{1}{2}} X)\big)$$

$$\text{s.t.} \begin{pmatrix} I \otimes \rho_0 & X \\ X^* & I \otimes \rho_1 \end{pmatrix} \ge 0,$$

where $X \in B(\mathcal{H} \otimes \mathcal{H})$, $\rho_0, \rho_1 \in \mathfrak{S}(\mathcal{H})$. We used MATLAB and the package CVX for solving the convex program.[10,11] We have computed the numerical values for 100 values of the parameter θ, homogeneously distributed on the interval $[0,1]$. Between these points straight lines were drawn, hence the figure looks like a continuous line.

The resulting expressions of p_{MEI}, $\|\mathrm{Tr}_{\mathcal{K}} Z\|$, and numerical data of p_{opt} as functions of θ are plotted in the Fig. 1. Even though it shows that maximally entangled input state is not optimal, notice that the upper bound is close to p_{opt} for small values of the parameter θ.
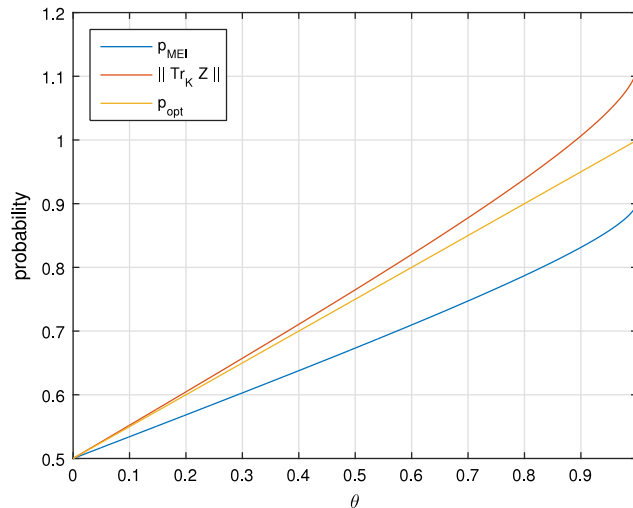

FIG. 1. The dependence of success probability with maximally entangled input state p_{MEI}, optimal success probability p_{opt}, and the upper bound $\|\mathrm{Tr}_{\mathcal{K}} Z\|$ given by Thm. 2 on the parameter θ in discrimination of the amplitude damping channel ϕ_{AD} and the identity channel (Example 2).

B. Unitary channels

Example 3. Let $\lambda = 1/2$ and $\dim(\mathcal{H}) = 3$. Without loss of generality, we may always assume that we discriminate the identity channel against a unitary Ad_W. As our first example, we consider the unitary matrix

$$W_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & e^{i\xi} \end{pmatrix}. \tag{26}$$

At this point, it is easy to compute p_{MEI} and $\|\mathrm{Tr}_{\mathcal{K}} Z\|$, as we can express them as functions of W_1, which is only dependent on ξ, so the bounds are functions of ξ. Moreover it is possible to compute p_{opt} by numerical methods described in Example 2. The expressions for p_{MEI} and $\|\mathrm{Tr}_{\mathcal{K}} Z\|$ are long and messy; they are plotted in Fig. 2 as well as the computed values of p_{opt}. From the figure it is once again clear that on one hand as p_{MEI} rises the upper bound becomes meaningless, but on the other hand if p_{MEI} is small then the upper bound and p_{opt} is small as well.

Another case we consider is matrix of the form

$$W_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}}(1+i) & 0 \\ 0 & 0 & e^{i\xi} \end{pmatrix}. \tag{27}$$

Again, it is straightforward to obtain p_{MEI} and $\|\mathrm{Tr}_{\mathcal{K}} Z\|$ and to compute p_{opt} numerically, the obtained functions are plotted in Fig. 3. Again it can be nicely seen that as p_{MEI} rises towards 1 the upper bound becomes meaningless.

C. Simple projective measurements

Let us demonstrate our results once more, this time for simple projective measurements.

Example 4. Let $\lambda = \frac{1}{2}$, $\dim(\mathcal{H}) \geq 3$, $|\eta_1\rangle = \frac{1}{\sqrt{2}}(|\xi_1\rangle + |\xi_2\rangle)$, $|\eta_2\rangle = \frac{1}{\sqrt{2}}(|\xi_1\rangle - |\xi_2\rangle)$, $|\eta_j\rangle = |\xi_j\rangle$ for $j \geq 3$. According to our previous results, the maximally entangled input state is not
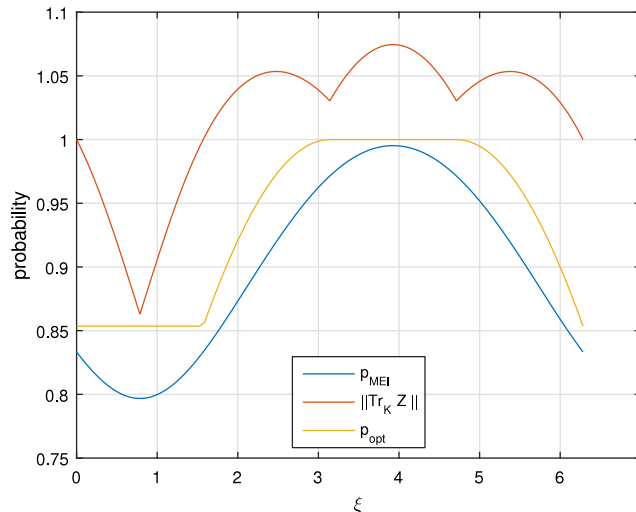

FIG. 2. The dependence of success probability with maximally entangled input state p_{MEI}, optimal success probability p_{opt}, and the upper bound $\|\mathrm{Tr}_{\mathcal{K}} Z\|$ given by Thm. 2 on the parameter ξ in discrimination of the unitary channel (26) and the identity (Example 3).
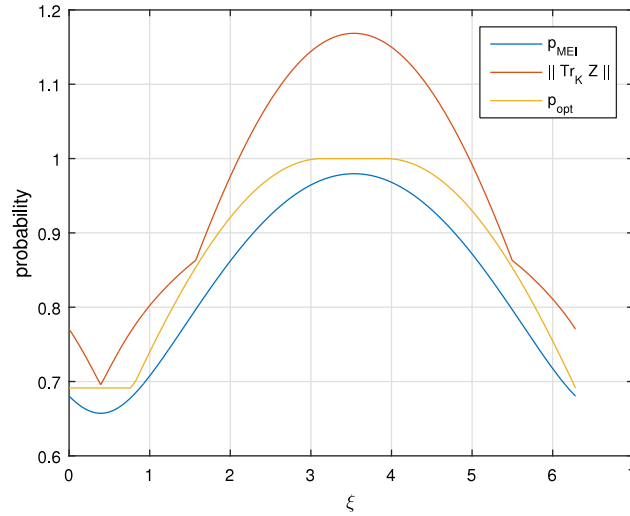
FIG. 3. The dependence of success probability with maximally entangled input state p_{MEI}, optimal success probability p_{opt}, and the upper bound $\|\text{Tr}_{\mathcal{K}} Z\|$ given by Thm. 2 on the parameter ξ in discrimination of the unitary channel (27) and the identity (Example 3).

optimal for discriminating these two measurements. As first we need to find $\text{Tr}_{\mathcal{K}}|\Delta_{\frac{1}{2}}| = \frac{1}{2} \sum_i |P_{\xi_i} - P_{\eta_i}|$. We get

$$|P_{\xi_1} - P_{\eta_1}| = \frac{1}{2}|P_{\xi_1} - P_{\xi_2} - |\xi_1\rangle\langle\xi_2| - |\xi_2\rangle\langle\xi_1||,$$

$$|P_{\xi_2} - P_{\eta_2}| = \frac{1}{2}|P_{\xi_2} - P_{\xi_1} + |\xi_1\rangle\langle\xi_2| + |\xi_2\rangle\langle\xi_1||,$$

from which we see, that $|P_{\xi_1} - P_{\eta_1}| = |P_{\xi_2} - P_{\eta_2}|$. The operator $P_{\xi_2} - P_{\eta_2}$ is diagonal in the orthonormal basis $|\varphi_i\rangle$, $i = 1, \ldots, \dim(\mathcal{H})$, where

$$|\varphi_1\rangle = \frac{1}{\sqrt{2}\sqrt{2-\sqrt{2}}} \left(|\xi_1\rangle + (1-\sqrt{2})|\xi_2\rangle\right),$$

$$|\varphi_2\rangle = \frac{1}{\sqrt{2}\sqrt{2+\sqrt{2}}} \left(|\xi_1\rangle + (1+\sqrt{2})|\xi_2\rangle\right),$$

$$|\varphi_j\rangle = |\xi_j\rangle,$$

where $j \geq 3$. In this basis it holds that

$$P_{\xi_1} - P_{\eta_1} = \frac{1}{\sqrt{2}}P_{\varphi_1} - \frac{1}{\sqrt{2}}P_{\varphi_2}.$$

The following calculation is straightforward. We obtain

$$\text{Tr}_{\mathcal{K}}|\Delta_{\frac{1}{2}}| = \frac{1}{\sqrt{2}}(P_{\varphi_1} + P_{\varphi_2}).$$

Now it is easy to see that

$$p_{\text{MEI}} = \frac{1}{2} + \frac{1}{\sqrt{2}\dim(\mathcal{H})},$$

$$\|\text{Tr}_{\mathcal{K}} Z\| = \frac{2+\sqrt{2}}{4} \approx 0.8535\ldots,$$

i.e., the bound is meaningful and the same for all $\dim(\mathcal{H})$, even though p_{MEI} tends to $\frac{1}{2}$ from above in the formal limit $\dim(\mathcal{H}) \to \infty$.

To underline the correctness of the upper bound we will find the optimal state for discrimination of the channels. Let us denote Φ_ξ, Φ_η the channel corresponding to simple projective measurement corresponding to the set of projectors $\{P_{\xi_i}\}$, $\{P_{\eta_i}\}$, respectively, and denote $P_2 = P_{\xi_1} + P_{\xi_2}$. Notice that the channels Φ_ξ and Φ_η can be separated as follows:

$$\Phi_\xi(\rho) = \sum_{i=1}^{2} \mathrm{Tr}\,(P_{\xi_i}\rho)P_{\xi_i} + \sum_{i=3}^{\dim(\mathcal{H})} \mathrm{Tr}\,(P_{\xi_i}\rho)P_{\xi_i}$$
$$= \phi_\xi(\rho) + \chi(\rho),$$
$$\Phi_\eta(\rho) = \sum_{i=1}^{2} \mathrm{Tr}\,(P_{\eta_i}\rho)P_{\xi_i} + \sum_{i=3}^{\dim(\mathcal{H})} \mathrm{Tr}\,(P_{\xi_i}\rho)P_{\xi_i}$$
$$= \phi_\eta(\rho) + \chi(\rho),$$

where ρ is a state on $\mathcal{H}$ and $\phi_\xi = \sum_{i=1}^{2} \mathrm{Tr}\,(P_{\xi_i}\rho)P_{\xi_i}$, $\phi_\eta = \sum_{i=1}^{2} \mathrm{Tr}\,(P_{\eta_i}\rho)P_{\xi_i}$ and $\chi = \sum_{i=3}^{\dim(\mathcal{H})} \mathrm{Tr}\,(P_{\xi_i}\rho)P_{\xi_i}$. Notice that

$$\phi_\xi(\rho) = \phi_\xi(P_2\rho P_2), \qquad \phi_\xi(P_2^\perp \rho P_2^\perp) = 0,$$
$$\phi_\eta(\rho) = \phi_\eta(P_2\rho P_2), \qquad \phi_\eta(P_2^\perp \rho P_2^\perp) = 0,$$
$$\chi(\rho) = \chi(P_2^\perp \rho P_2^\perp), \qquad \chi(P_2\rho P_2) = 0,$$

where $P_2^\perp = I - P_2$. This leads to the following:

$$\Phi_\xi(\rho) = \Phi_\xi(P_2\rho P_2 + P_2^\perp \rho P_2^\perp),$$
$$\Phi_\eta(\rho) = \Phi_\eta(P_2\rho P_2 + P_2^\perp \rho P_2^\perp).$$

In other words, it is sufficient to consider only input states of the form $\rho = \lambda'\rho_2 + (1 - \lambda')\rho_\perp$, where ρ_2 and $\rho^\perp$ are states, such that $P_2\rho_2 P_2 = \rho_2$, $P_2^\perp \rho_\perp P_2^\perp = \rho_\perp$ and $0 \le \lambda' \le 1$. Since

$$p_{\mathrm{opt}} = \max\left(\frac{1}{2}\mathrm{Tr}\,((\Phi_\xi \otimes id)(\rho)M) + \frac{1}{2}\mathrm{Tr}\,((\Phi_\eta \otimes id)(\rho)(I - M))\right)$$
$$= \max\left(\frac{\lambda}{2}\left(1 + \mathrm{Tr}\,(((\phi_\xi \otimes id)(\rho_2) - (\phi_\eta \otimes id)(\rho_2))_+)\right) + \frac{1 - \lambda'}{2}\right),$$

where $M, I - M$ is the POVM we use for discrimination of the channels; it is obvious that for p_{opt} to be maximal, we also have to set $\lambda = 1$ and according to corollary 2 the problem reduces to discriminating the channels ϕ_ξ and ϕ_η. Both ϕ_ξ and ϕ_η are unital qubit channels and by our previous results the optimal input state ρ_{opt} is

$$\rho_{opt} = \frac{1}{2} \sum_{i,j=1}^{2} |\xi_i\rangle\langle\xi_j| \otimes |\xi_i\rangle\langle\xi_j|.$$

In the product basis given by $|\xi_1\rangle, |\xi_2\rangle$, we have

$$(\phi_\xi \otimes id)(\rho_{opt}) - (\phi_\eta \otimes id)(\rho_{opt}) = \frac{1}{4}\begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

This matrix has two eigenvalues $\frac{1}{2\sqrt{2}}$ and $-\frac{1}{2\sqrt{2}}$ each with multiplicity 2. We get $\mathrm{Tr}\,(((\phi_\xi \otimes id)(\rho_2) - (\phi_\eta \otimes id)(\rho_2))_+) = \frac{1}{\sqrt{2}}$ and

$$p_{\mathrm{opt}} = \frac{1}{2}\left(1 + \frac{1}{\sqrt{2}}\right) = \frac{2 + \sqrt{2}}{4}$$

which is exactly the same as our upper bound.

VI. CONCLUSIONS

We presented necessary and sufficient conditions for optimality of a process POVM in channel discrimination, especially for a process POVM corresponding to a measurement scheme with full Schmidt rank input state. In particular, a necessary and sufficient condition for existence of an optimal measurement scheme with a given full Schmidt rank input state was found. In the case of maximally entangled input states, an upper bound of the optimal success probability was given if the optimality condition is not satisfied. For discrimination of two channels, we obtained a simple condition in terms of the Choi matrices of the channels and a new upper bound on the diamond norm.

The results were applied to discrimination of four types of channels. For covariant channels, known results for the irreducible case were extended to some reducible cases and an upper bound on the optimal success probability was found. For qubit channels, the obtained condition generalizes previously known results to some pairs of non-unital channels. We proved that for discrimination of unitary channels, maximally entangled input states are optimal only in some very special cases. To our best knowledge, the results obtained for unitary channels and simple projective measurements are new.

An interesting open question is whether it is possible to obtain a similar condition for input states with lower Schmidt rank. As it was shown, there are cases when such input states are optimal and there may even be no optimal full rank input states. The question is not only what the Schmidt rank of the optimal input state may be but also how to select the subspace of the input Hilbert space that will form the support of the partial trace of the input state. Another possible future directions of research is to investigate optimal discrimination of more complex quantum processes.

ACKNOWLEDGMENTS

APPENDIX: COMPUTATION OF THE ABSOLUTE VALUE OF THE DIFFERENCE OF POSITIVE RANK-1 OPERATORS

Let $|\phi\rangle, |\psi\rangle$ be unit vectors, $P_\phi = |\phi\rangle\langle\phi|$, $P_\psi = |\psi\rangle\langle\psi|$. Let $\lambda \in (0,1)$ and

$$D_\lambda := \lambda P_\psi - (1-\lambda)P_\phi.$$

Then

$$|D_\lambda| = |\mu_1|P_{\xi_1} + |\mu_2|P_{\xi_2},$$

where $\mu_1, \mu_2 \in \mathbb{R}$ are the eigenvalues and ξ_1, ξ_2 the corresponding eigenvectors of D_λ. Moreover, we have

$$\xi_i = \alpha_i \phi + \beta_i \psi, \qquad i = 1,2$$

and since we do not worry about a phase, we may suppose that $\alpha_i \geq 0$. From $D_\lambda |\xi_i\rangle = \mu_i |\xi_i\rangle$, we obtain

$$\lambda(\alpha_i \bar{z} + \beta_i)|\psi\rangle - (1-\lambda)(\alpha_i + \beta_i z)|\phi\rangle = \mu_i \alpha_i |\phi\rangle + \mu_i \beta_i |\psi\rangle,$$

where $z = \langle \phi, \psi \rangle$. It follows that $\lambda(\alpha_i \bar{z} + \beta_i) = \mu_i \beta_i$, so that

$$\beta_i = k_i \bar{z}, \qquad i = 1,2,$$

where $k_i = \frac{\lambda}{\mu_i - \lambda} \alpha_i \in \mathbb{R}$. We obtain

$$
\begin{aligned}
|D_\lambda| &= \sum_{i=1}^{2} |\mu_i| |\alpha_i \phi + k_i \bar{z} \psi \rangle \langle \alpha_i \phi + k_i \bar{z} \psi| \\
&= \sum_{i=1}^{2} |\mu_i| (\alpha_i^2 P_\phi + k_i^2 |z|^2 P_\psi \\
&\quad + \alpha_i k_i z |\phi\rangle\langle\psi| + \alpha_i k_i \bar{z} |\psi\rangle\langle\phi|) \\
&= (|\mu_1| \alpha_1^2 + |\mu_2| \alpha_2^2) P_\phi \\
&\quad + (|\mu_1| k_1^2 + |\mu_2| k_2^2) |z|^2 P_\psi + (\alpha_1 k_1 |\mu_1| \\
&\quad + \alpha_2 k_2 |\mu_2|)(z|\phi\rangle\langle\psi| + \bar{z}|\psi\rangle\langle\phi|).
\end{aligned}
$$

[1] S. M. Barnett and S. Croke, "Quantum state discrimination," Adv. Opt. Photonics **1**(2), 238–278 (2009).

[2] A. O. Barut and R. Rączka, *Theory of Group Represetations and Applications* (PWN - Polish Scientific Publishers, Warszava, 1980).

[3] A. Barvinok, *A Course in Convexity* (American Mathematical Society, 2002).

[4] F. G. S. L. Brandao, M. Piani, and P. Horodecki, Nat. Commun. **6**, 7908 (2015).

[5] G. M. D'Ariano, P. Lo Presti, and M. G. A. Paris, Phys. Rev. Lett. **87**, 270404 (2001).

[6] G. M. D'Ariano, P. Lo Presti, and M. G. A. Paris, J. Opt. B: Quantum Semiclassical Opt. **4**, 273 (2002).

[7] G. Chiribella, "Optimal networks for quantum metrology: Semidefinite programs and product rules," New J. Phys. **14**, 1–19 (2012).

[8] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Phys. Rev. Lett. **101**, 180501 (2008).

[9] M. D. Choi, "Completely positive maps on complex matrices," Linear Algebra Appl. **10**, 285–290 (1975).

[10] CVX Research, Inc., CVX: Matlab software for disciplined convex programming, version 2.1, http://cvxr.com/cvx, June 2015.

[11] M. Grant and S. Boyd, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control (a tribute to M. Vidyasagar)*, Lecture Notes in Control and Information Sciences, edited by V. Blondel, S. Boyd, and H. Kimura (Springer, 2008), pp. 95–110.

[12] G. Gutoski, J. Math. Phys. **53**(3), 032202 (2012).

[13] G. Gutoski and J. Watrous, "Toward a general theory of quantum games," in *Proceedings of STOC 2007* (ACM New York, NY, USA, 2007), pp. 565–574, http://dl.acm.org/citation.cfm?id=1250873.

[14] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, 1976).

[15] A. S. Holevo, J. Multivar. Anal. **3**, 337–394 (1973).

[16] A. Jenčová, J. Math. Phys. **55**, 022201 (2014).

[17] A. Kitaev, "Quantum computations: Algorithms and error correction," Russ. Math. Surv. **52**, 1191–1249 (1997).

[18] K. Matsumoto, e-print arXiv:1209.2392.

[19] M. Piani and J. Watrous, Phys. Rev. Lett. **102**(25), 250501 (2009).

[20] M. F. Sacchi, "Optimal discrimination of quantum operations," Phys. Rev. A **71**, 062340 (2005).

[21] M. F. Sacchi, "Entanglement can enhance the distinguishability of entanglement-breaking channels," Phys. Rev. A: At., Mol., Opt. Phys. **72**(1), 1–2 (2005).

[22] M. F. Sacchi, "Minimum error discrimination of Pauli channels," J. Opt. B: Quantum Semiclassical Opt. **7**(10), S333–S336 (2005).

[23] M. Sedlák and M. Ziman, Phys. Rev. A **90**, 052312 (2014).

[24] W. Tadej and K. Zyczkowski, Open Syst. Inf. Dyn. **13**, 133–177 (2006).

[25] J. Watrous, Chicago J. Theor. Comput. Sci. **19**, 1–19 (2013).

[26] H. Yuen, R. Kennedy, and M. Lax, IEEE Trans. Inf. Theory **21**, 125–134 (1975).

[27] M. Ziman, Phys. Rev. A **77**, 062112 (2008).