

# Operational relevance of resource theories of quantum measurements

Michał Oszmaniec and Tanmoy Biswas

Institute of Theoretical Physics and Astrophysics, National Quantum Information Centre, Faculty of Mathematics, Physics and Informatics, University of Gdansk, Wita Stwosza 57, 80-308 Gdańsk, Poland

April 17, 2019

For any resource theory it is essential to identify tasks for which resource objects offer advantage over free objects. We show that this identification can always be accomplished for resource theories of quantum measurements in which free objects form a convex subset of measurements on a given Hilbert space. To this aim we prove that every resourceful measurement offers advantage for some quantum state discrimination task. Moreover, we give an operational interpretation of robustness, which quantifies the minimal amount of noise that must be added to a measurement to make it free. Specifically, we show that this geometric quantity is related to the maximal relative advantage that a resourceful measurement offers in a class of minimal-error state discrimination (MESD) problems. Finally, we apply our results to two classes of free measurements: incoherent measurements (measurements that are diagonal in the fixed basis) and separable measurements (measurements whose effects are separable operators). For both of these scenarios we find, in the asymptotic setting in which the dimension or the number of particles increase to infinity, the maximal relative advantage that resourceful measurements offer for state discrimination tasks.

## 1 Introduction

Resource theories [1] constitute a powerful toolbox to study physical systems in the presence of limitations resulting from experimental or operational constraints on the ability to address and manipulate physical systems. This mathematical framework is general enough to encompass both classical and quantum physics, or even more general theories [2]. In recent years it has been successfully applied to classical and quantum thermodynamics [3, 4], pro-

cessing of quantum information in distributed scenarios [5, 6], contextuality [7], nonlocality [8], steering [9] and, last but not least, magic state distillation paradigm of quantum computation [10]. On the general level, all resource theories are defined by specifying *free objects* and *free operations*. Free objects form a subset of the set of all objects relevant for the physical situation in question. Likewise, free operations form a subclass of all relevant physical transformations. Typically objects that are not free are called *resource objects*. The specific choice of free objects and free operations depends on the physical context. For example, in entanglement theory the relevant resource theory is based on the principle of locality: set of free objects consists of separable states while free operations are all local operations assisted by classical communication (LOCC).

For any resource theory it is desirable to give operational interpretation of resource objects i.e to identify a task for which a given resource object would prove advantageous over all free objects. The main purpose of this work is to provide such interpretations for resource theories in which a set of free objects consists of convex subset of the set of quantum measurements (POVMs) on a relevant Hilbert space. We realize this goal by showing, that in this setting for every resourceful measurement  $\mathbf{M}$  there an instance of minimal-error quantum state discrimination game for which  $\mathbf{M}$  gives greater success probability than all free measurements. Quantum state discrimination [11, 12] is a popular quantum information subroutine that finds applications in different areas of quantum information science including quantum communication [13], quantum metrology [14], nonlocality [15] or quantum computation [16]. We push our operational interpretation further by proving the quantitative relation between the relative advantage of resourceful measurement for state discrimination and the geometric measure of resourcefulness called robustness [17, 18], which quantifies the minimal amount of noise that has to be added to a POVM to make it free. Previously, the measure robustness was introduced [19] to math-

Michał Oszmaniec: [michal.oszmaniec@gmail.com](mailto:michal.oszmaniec@gmail.com)

ematically quantify entanglement. Later, it was employed in the general resource-theoretic framework [1, 20, 21].

We apply our general results to two classes of free measurements: incoherent measurement and separable measurements. Incoherent measurements are POVM analogues of incoherent states [22–24] and can be understood as measurements originating from a single projective measurement (in the fixed basis) followed by arbitrary classical post-processing. Separable measurements [25–28], on the other hand, are POVMs on composite quantum systems whose effects are separable operators. This class contains the set of LOCC measurements [29, 30] i.e. measurements that can be implemented via local measurements and LOCC operations. For both incoherent and separable measurements we focus on the asymptotic setting in which the dimension of the system or the number of particles involved go to infinity. In this regime we identify, for both classes of measurements, the maximal relative advantage that resourceful measurements can offer for quantum state discrimination tasks.

Traditionally, operational interpretations of quantum resources were developed case-by-case for different resource theories [31–36]. A recent paper [18] gave an unified treatment of this problem for all quantum resource theories in which the set of free objects is convex subset of the set of quantum states. Specifically, the authors of this work showed that in this context all resource states offer advantage for some sub-channel discrimination problem (this result was previously obtained for resource theories of entanglement [37], coherence [38], steering [39] and asymmetry [40]). Our results are complementary to [18], as they give the operational interpretation of *all convex resource theories of quantum measurements*. Moreover, our work greatly generalizes some of the results of [17], where the analogous analysis was presented for the case where free measurements consisted of maximally uninformative measurements i.e. measurements that do not recover any information about the measured quantum states.

So far, research in quantum resource theories focussed mainly on quantum states while quantum measurements, despite their importance, did not receive much attention. Previously, resource-theoretic perspective was applied in the context of measurement incompatibility [41, 42], measurement simulability via projective measurements [43, 44], and simulability in the more general scenarios [42, 45]. We believe that our results, especially previously unexplored quantitative relation between state dis-

crimination and robustness, provide new quantitative tools to study the restricted classes of POVMs and, more generally, quantum resource theories concentrated around quantum measurements.

## 2 Notation and main concepts

Throughout the paper we will be interested in POVMs on finite dimensional Hilbert space  $\mathcal{H} \approx \mathbb{C}^d$ . Such generalized measurements (POVMs) can be understood as tuples  $\mathbf{M} = (M_1, \dots, M_n)$  of non-negative operators on  $\mathbb{C}^d$  satisfying  $\sum_{i=1}^n M_i = \mathbb{1}$ , where  $n$  is the number of outcomes and  $\mathbb{1}$  is the identity on  $\mathbb{C}^d$ . The operators  $M_i$  are called the effects of POVM  $\mathbf{M}$ . According to Born's rule, when a POVM  $\mathbf{M}$  is measured on a quantum state  $\rho$  the probability of obtaining the outcome  $i$  is given by  $p_i = \text{tr}(M_i \rho)$ . We denote the set of POVMs on  $\mathbb{C}^d$  with  $n$  outcomes by  $\mathcal{P}(d, n)$ . This set has a natural convex structure [46]: given two POVMs  $\mathbf{M}, \mathbf{N} \in \mathcal{P}(d, n)$ , their convex combination  $p\mathbf{M} + (1-p)\mathbf{N}$  is the POVM with  $i$ -th effect given by  $[p\mathbf{M} + (1-p)\mathbf{N}]_i := pM_i + (1-p)N_i$ . The operation of taking convex combinations of measurements can be operationally realized as performing POVMs  $\mathbf{M}$  and  $\mathbf{N}$  with certain probabilities and then combining the outcomes.

Our operational interpretation of *resourcefulness* of quantum measurements will be based on the task of minimal-error state discrimination [11, 12]. The purpose of this task is to optimally distinguish quantum states generated by the ensemble  $\mathcal{E} = \{q_i, \rho_i\}_{i=1}^n$ , where  $\{q_i\}_{i=1}^n$  is a probability distribution and  $\{\rho_i\}_{i=1}^n$  is a collection of quantum states. The success probability of identifying the states generated by  $\mathcal{E}$  via a measurement  $\mathbf{M} \in \mathcal{P}(d, n)$  is given by

$$p_{\text{succ}}(\mathcal{E}, \mathbf{M}) = \sum_{i=1}^n q_i \text{tr}(M_i \rho_i). \quad (1)$$

One is often interested in choosing the measurement  $\mathbf{M}$ , such that  $p_{\text{succ}}(\mathcal{E}, \mathbf{M})$  is maximized. However, if only a restricted class of measurements  $\mathcal{F} \subset \mathcal{P}(d, n)$  is allowed, the maximum  $p_{\text{succ}}$  may not be achieved.

## 3 Convex resource theories of measurements and measurement robustness

We now give a minimal formulation of a resource theory of measurements, under the assumption of convexity. In our treatment we will focus on a set of quantum measurements  $\mathcal{F}$  and a class of free op-

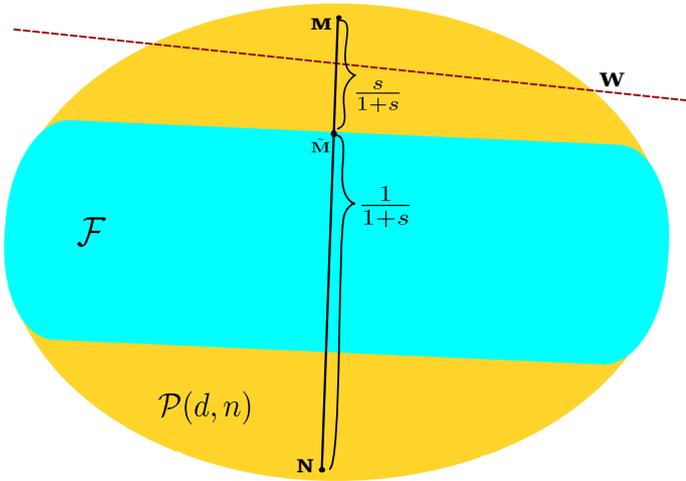


Figure 1: A figure demonstrating some of the key ideas of the paper. The blue part denotes the set of all free measurements  $\mathcal{F}$  while orange part denotes the set of all  $n$  outcome measurements on  $\mathbb{C}^d$  that are not free. (i) Illustration of Theorem 1: for any  $\mathbf{M} \notin \mathcal{F}$  there always exists a hyperplane  $\mathbf{W}$  which separates any POVM  $\mathbf{M} \notin \mathcal{F}$  in a sense  $\langle \mathbf{M}, \mathbf{W} \rangle > \langle \mathbf{F}, \mathbf{W} \rangle$  for all  $\mathbf{F} \in \mathcal{F}$ . By Theorem 1, every separating hyperplane can be rewritten in terms of a state discrimination problem for the ensemble of quantum states  $\mathcal{E}$  such that for all free measurements  $\mathbf{F}$  we have  $p_{\text{succ}}(\mathcal{E}, \mathbf{M}) > p_{\text{succ}}(\mathcal{E}, \mathbf{F})$ . (ii) Geometric explanation of Robustness: the straight line represents convex combinations  $\frac{\mathbf{M} + s\mathbf{N}}{1+s}$ . For a given POVMs  $\mathbf{M}$  and  $\mathbf{N}$  there exist the minimal value  $s_*$  (depending on  $\mathbf{N}$  for a fixed  $\mathbf{M}$ ) such that  $\frac{\mathbf{M} + s\mathbf{N}}{1+s}$  becomes free. The robustness  $R_{\mathcal{F}}(\mathbf{M})$  is defined as the minimum  $s_*$ , considering all possible POVM  $\mathbf{N} \in \mathcal{P}(d, n)$ .

erations <sup>1</sup>  $\mathcal{O}$ . We first assume (i) that the set of free measurements  $\mathcal{F}$  is a convex and closed subset of  $n$ -outcome measurements on  $\mathbb{C}^d$  i.e.  $\mathcal{F} \subset \mathcal{P}(d, n)$ , for some suitable Hilbert space  $\mathcal{H} \approx \mathbb{C}^d$ . Second, we assume that the class of free operations  $\mathcal{O}$  consists of mappings  $\varphi : \mathcal{P}(d, n) \rightarrow \mathcal{P}(d, n)$  that (ii) preserve the set of free measurements i.e. for all  $\mathbf{N} \in \mathcal{F}$  we have  $\varphi(\mathbf{N}) \in \mathcal{F}$  and (iii) are convex-linear i.e.  $\varphi(p\mathbf{M} + (1-p)\mathbf{M}') = p\varphi(\mathbf{M}) + (1-p)\varphi(\mathbf{M}')$  for all measurements  $\mathbf{M}, \mathbf{M}' \in \mathcal{P}(d, n)$ .

*Justification of assumptions:* The convexity of  $\mathcal{F}$  can be justified by the fact that, as argued before, convex combination of measurements can be realised by a purely classical process. The compactness of  $\mathcal{F}$  is a technical assumption that allows us to formally state our results. However, in basically any physically interesting case the class of free measurements  $\mathcal{F}$  can be viewed as a set of convex combinations

<sup>1</sup>Of course, the full treatment of resource theory of measurements requires to take into account more complicated aspects such as composability. We plan to address such questions in the future work.

(convex hull) of a compact set of "primitive" measurements  $\mathcal{F}'$ , and hence is naturally compact. Finally, convex linearity of free operations follows from linearity of quantum mechanics while preservation of the  $\mathcal{F}$  by free operation is also natural - operations that are considered free should not be able to create resourceful measurement. In particular, in many cases the set  $\mathcal{F}$  contains (a) classical post-processing i.e. operation of the form  $\mathcal{Q}(\mathbf{M}) = \mathbf{M}'$ , where  $M'_i = \sum_j q(i|j)M_j$ , for some stochastic matrix  $\mathcal{Q}_{ij} = q(i|j)$ , and (b) unitary symmetries of the set  $\mathcal{F}$ .

For every set of free measurements  $\mathcal{F}$  it is possible to introduce a geometric measure quantifying the minimum amount of noise which has to be added to a given POVM  $\mathbf{M}$  to make it free. In analogy to previous works [17–19] we will call this quantifier *robustness* and denote it by  $R_{\mathcal{F}}$ . It is formally defined in the following way (See Fig. 1):

$$R_{\mathcal{F}}(\mathbf{M}) := \min \left\{ s \mid \exists \mathbf{N} \text{ such that } \frac{\mathbf{M} + s\mathbf{N}}{1+s} \in \mathcal{F} \right\}. \quad (2)$$

The robustness enjoys a number of natural properties as stated by the following proposition.

**Proposition 1** (Properties of robustness). *Let  $\mathcal{F} \subset \mathcal{P}(d, n)$  be a closed convex set of free measurements. The robustness  $R_{\mathcal{F}}$  satisfies the following properties: (i) Faithfulness:  $R_{\mathcal{F}}(\mathbf{M}) \geq 0$  and  $R_{\mathcal{F}}(\mathbf{M}) > 0$  iff  $\mathbf{M} \in \mathcal{F}$ , (ii) Convexity:  $R_{\mathcal{F}}(p\mathbf{M}_1 + (1-p)\mathbf{M}') \leq pR_{\mathcal{F}}(\mathbf{M}_1) + (1-p)R_{\mathcal{F}}(\mathbf{M}')$ , (iii) Monotonicity under free operations:  $R_{\mathcal{F}}[\varphi(\mathbf{M})] \leq R_{\mathcal{F}}(\mathbf{M})$  for all free operations  $\varphi \in \mathcal{O}$ .*

The proof of the above follows directly from the definitions of the concepts involved and we defer it to the Appendix.

## 4 Main results

In this part we give our main results concerning the operational interpretation of resourceful measurements for a class of quantum state discrimination tasks.

**Theorem 1** (Minimal-error state discrimination can certify resource character of measurements). *For any resourceful POVM  $\mathbf{M} \notin \mathcal{F}$ , there exist an ensemble  $\mathcal{E}_{\mathbf{M}} = \{p_i, \rho_i\}$  such that*

$$p_{\text{succ}}(\mathcal{E}_{\mathbf{M}}, \mathbf{M}) > \max_{\mathbf{N} \in \mathcal{F}} p_{\text{succ}}(\mathcal{E}_{\mathbf{M}}, \mathbf{N}). \quad (3)$$

*Proof.* We begin the proof by noting that  $\mathcal{P}(d, n)$  is a convex subset of  $(\text{Herm}(\mathcal{H}))^{\times n}$  where  $\mathcal{H} \approx \mathbb{C}^d$

is a  $d$ - dimensional Hilbert space and  $\text{Herm}(\mathcal{H})$  is the set of hermitian operator from  $\mathcal{H}$  to itself. The space  $(\text{Herm}(\mathcal{H}))^{\times n}$  can be regarded as  $nd^2$  dimensional real vector space with the inner product defined by  $\langle \mathbf{A}, \mathbf{B} \rangle = \sum_{i=1}^n \text{tr}(A_i B_i)$ , where  $\mathbf{A}, \mathbf{B} \in (\text{Herm}(\mathcal{H}))^{\times n}$ . We have assumed that  $\mathcal{F}$  is a compact set and hence, due to hyperplane separation theorem (see Fig. 1), for any  $\mathbf{M} \notin \mathcal{F}$ , there exists a tuple of Hermitian operators  $\mathbf{W} = (W_1, W_2, \dots, W_n)$  such that  $\sum_{i=1}^n \text{tr}(W_i M_i) > 0$  while for all  $\mathbf{N} \in \mathcal{F}$  we have  $\sum_i \text{tr}(W_i N_i) \leq 0$ . We now define  $\tilde{W}_i := W_i + |\lambda| \mathbb{I}$ , where  $\lambda$  is the smallest eigenvalue of operators  $\{W_i\}$ . We note that for every  $i$  the operator  $\tilde{W}_i$  is positive-semi-definite and moreover  $\sum_i \text{tr}(\tilde{W}_i N_i) < \sum_i \text{tr}(\tilde{W}_i M_i)$ . Using this we construct the ensemble  $\mathcal{E}_{\mathbf{M}} = \left\{ \frac{\text{tr}(\tilde{W}_i)}{\sum_i \text{tr}(\tilde{W}_i)}, \frac{\tilde{W}_i}{\text{tr}(\tilde{W}_i)} \right\}$ , for which the inequality (3) holds.  $\square$

The above result shows the qualitative advantage of resourceful measurements for a class of state-discrimination tasks. In what follows we show a natural way to quantify this advantage. In Theorem 2 below we show that under the same assumptions as above, the robustness  $\mathbb{R}_{\mathcal{F}}(\mathbf{M})$  is related to the maximal relative advantage that a resourceful measurement  $\mathbf{M}$  offers for MESD. To prove this result we will use the following technical Lemma that can be regarded as dual characterisation of measurement robustness. Its proof relies on duality of semi-infinite convex optimization programs [47] and we present it in the Appendix.

**Lemma 1.** *Let  $\mathbf{M} \in \mathcal{P}(d, n)$  and  $\mathcal{F} \subset \mathcal{P}(d, n)$  be a compact convex set of free measurements. Then, the generalized robustness  $R_{\mathcal{F}}(\mathbf{M})$  can be expressed as the following optimisation problem.*

$$\begin{aligned} & \text{maximize} && \sum_{i=1}^n \text{tr}(Z_i M_i) - 1 \\ & \text{subject to} && Z_i \geq 0, \quad i = 1, \dots, n, \\ & && \sum_{i=1}^n \text{Tr}(Z_i N_i) \leq 1 \quad \forall \mathbf{N} \in \mathcal{F} \end{aligned} \quad (4)$$

We are now ready to state and prove our main result.

**Theorem 2** (Operational interpretation of robustness in terms of quantum state discrimination). *For any class of free measurements  $\mathcal{F}$  and any POVM  $\mathbf{M} \in \mathcal{P}(d, n)$  the generalized robustness can be expressed as*

$$R_{\mathcal{F}}(\mathbf{M}) = \max_{\mathcal{E}} \frac{p_{\text{succ}}(\mathcal{E}, \mathbf{M})}{\max_{\mathbf{N} \in \mathcal{F}} p_{\text{succ}}(\mathcal{E}, \mathbf{N})} - 1, \quad (5)$$

where the outer optimisation is over  $n$  element ensembles of quantum states.

*Proof.* We begin by showing that  $\text{LHS}(5) \geq \text{RHS}(5)$ . From the definition of generalized robustness from (2) it follows that for each effect  $M_i$  of POVM  $\mathbf{M} \in \mathcal{P}(d, n)$  it holds that  $M_i + \mathcal{R}_{\mathcal{F}}(\mathbf{M})M'_i = (1 + \mathcal{R}_{\mathcal{F}}(\mathbf{M}))N'_i$ , where  $M'_i$  are the effects of POVM  $\mathbf{M}' \in \mathcal{P}(d, n)$  and  $N'_i$  are the effects of POVM  $\mathbf{N}' \in \mathcal{F}$ . Therefore, we have  $M_i \leq (1 + \mathcal{R}_{\mathcal{F}}(\mathbf{M}))N'_i$  for every  $i \in \{1, n\}$  and therefore

$$\begin{aligned} p_{\text{succ}}(\mathcal{E}, \mathbf{M}) &\leq (1 + \mathcal{R}_{\mathcal{F}}(\mathbf{M}))p_{\text{succ}}(\mathcal{E}, \mathbf{N}') \\ &\leq (1 + \mathcal{R}_{\mathcal{F}}(\mathbf{M})) \max_{\mathbf{N} \in \mathcal{F}} p_{\text{succ}}(\mathcal{E}, \mathbf{N}), \end{aligned} \quad (6)$$

where the second inequality follows from  $\mathbf{N}' \in \mathcal{F}$ .

To show  $\text{LHS}(5) \leq \text{RHS}(5)$ , we use the dual characterization of the generalized robustness from Lemma 1. Let  $(Z_1^*, \dots, Z_n^*)$  be the optimal tuple for which (4) is maximized. We now define

$$\mathcal{E}_* := \left\{ \frac{\text{tr}(Z_i^*)}{\sum_i \text{tr}(Z_i^*)}, \frac{Z_i^*}{\text{tr}(Z_i^*)} \right\}. \quad (7)$$

Simple algebraic manipulations give  $\mathcal{R}_{\mathcal{F}}(\mathbf{M}) + 1 = \left( \sum_{i=1}^n \text{tr}(Z_i^*) \right) p_{\text{succ}}(\mathcal{E}_*, \mathbf{M})$ . On the other hand from (4) it follows that  $\forall \mathbf{N} \in \mathcal{F}$  we have  $p_{\text{succ}}(\mathcal{E}_*, \mathbf{N}) \leq \frac{1}{\sum_i \text{tr}(Z_i^*)}$ . Using these two properties we obtain

$$\frac{p_{\text{succ}}(\mathcal{E}_*, \mathbf{M})}{\max_{\mathbf{N} \in \mathcal{F}} p_{\text{succ}}(\mathcal{E}_*, \mathbf{N})} \geq \mathcal{R}_{\mathcal{F}}(\mathbf{M}) + 1. \quad (8)$$

By combining (6) and (8) we complete the proof.  $\square$

Importantly, Theorems 1 and 2 are valid for all convex and compact sets of free measurements  $\mathcal{F}$ . In particular, we can directly apply them to the measurements that are simulable by projective measurements [43, 44] or  $k$ -outcome measurements [42, 45]. In what follows, we will consider application of our results to two other classes of free measurements: incoherent measurements and separable measurements. We will focus on obtaining the maximal relative advantage for MESD that can be obtained in these scenarios by resourceful measurements.

#### 4.1 Incoherent measurements

Quantum coherence or superposition is one of the salient nonclassical aspects of quantum theory. Recently, this feature of quantum theory was investigated thoroughly from the resource-theoretic perspective [23, 24]. Moreover, quantum coherence has been identified as resource for a number of quantum

information processing tasks like quantum channel discrimination [37], designing quantum algorithms [48], quantum metrology [14] and quantum thermodynamics [3]. Finally, there is a plethora of works on quantification of quantum coherence of quantum states [23, 49] and operations [50, 51] and on giving operational interpretation to some quantifiers [35, 36].

Here we propose to introduce a concept of coherence to the realm of POVMs. Specifically, we define the class of *incoherent measurements*  $\mathcal{IC}(d, n)$  consisting of  $n$ -outcome measurements  $\mathbf{S}$  on  $\mathbb{C}^d$ , whose effects are diagonal in a particular orthogonal basis  $\{|i\rangle\}$ . In what follows, for the sake of simplicity, we assume  $n \geq d$ . Incoherent measurements constitute a POVM analogues of the set of *incoherent states*. More formally, the effects of incoherent measurements  $\mathbf{S}$  can be expressed as  $S_i = \sum_{j=1}^d q(i|j)|j\rangle\langle j|$ , where  $i \in [n]$ . It is now easy to see that incoherent measurements can always be obtained via classical post-processing of projective measurements in the standard basis.

One way to quantify the coherence present in a POVM is the robustness  $R_{\mathcal{IC}}(\mathbf{M})$  with respect to the set of incoherent measurements. Importantly,  $R_{\mathcal{IC}}(\mathbf{M})$  satisfies a number of natural properties stated in Lemma 1. Here, the class of free operations  $\mathcal{O}_{\mathcal{IC}}$  can be chosen to contain all diagonal unitaries and classical post-processing<sup>2</sup>. By applying Theorem 2 directly to this setting we get that robustness  $R_{\mathcal{IC}}(\mathbf{M})$ , can be identified as the maximal relative advantage of  $\mathbf{M}$  for a suitable state discrimination problem. In the following we will take the complementary approach based on semi-definite programming [52]. From the definition of incoherent measurements it follows that  $R_{\mathcal{IC}}(\mathbf{M})$  can be casted as the following semi-definite program (SDP)

$$\begin{aligned} & \text{minimize } s \\ & \text{s.t } \frac{M_a + sN_a}{1 + s} = \sum_i q(a|i)|i\rangle\langle i| \quad \forall a, \\ & N_a \geq 0, \quad a = 1, \dots, n. \quad \sum_a N_a = \mathbb{I}, \\ & \sum_a q(a|i)(1 + s) = b \quad \forall i. \end{aligned} \quad (9)$$

Taking the dual problem we get the following auxiliary result whose proof we defer to the Appendix.

**Lemma 2** (Dual characterization of robustness w.r.t. incoherent measurement). *Robustness of a*

<sup>2</sup>In fact all quantum channels  $\Lambda$  such that  $\Lambda^*$  is an incoherent channel also preserve  $\mathcal{IC}(n, d)$ .

POVM  $\mathbf{M} \in \mathcal{P}(d, n)$  w.r.t set of incoherent measurement  $\mathcal{IC}(d, n)$ , i.e,  $R_{\mathcal{IC}}(\mathbf{M})$  can be expressed as the solution of optimization problem given by

$$\begin{aligned} & \text{maximize } \sum_{a=1}^n \text{tr}(Z_a M_a) - 1 \\ & \text{subject to } \forall i, a \quad Z_a \geq 0 \quad \langle i|Z_a|i\rangle = \langle i|Z_n|i\rangle, \\ & \quad \text{tr}(Z_n) = 1. \end{aligned} \quad (10)$$

From the above we get a physical characterization of ensembles of states that suffice to capture the maximal advantage of coherent measurements over the incoherent ones.

**Proposition 2** (Alternative characterisation of robustness for incoherent measurements). *Let  $\mathbf{M} \in \mathcal{P}(d, n)$ . We have the following equality*

$$R_{\mathcal{IC}}(\mathbf{M}) = \max_{\mathcal{E}_0} \frac{p_{\text{succ}}(\mathcal{E}_0, \mathbf{M})}{\max_{\mathbf{N} \in \mathcal{IC}(d, n)} p_{\text{succ}}(\mathcal{E}_0, \mathbf{N})} - 1, \quad (11)$$

where the outer optimisation is over ensembles  $\mathcal{E}_0 = \{1/n, \sigma_i\}_{i=1}^n$  consisting of states that cannot be classically distinguished [53] (that is, for all  $j \in [d]$  and for all  $i \in [n]$  we have  $\langle j|\sigma_i|j\rangle = 1/d$ ). Due to this property of  $\sigma_i$  for all incoherent measurements  $\mathbf{N}$  we have  $p_{\text{succ}}(\mathcal{E}_0, \mathbf{N}) = 1/n$ .

*Proof.* We observe that due to the assumed property of states  $\sigma_i$  (for all  $k$   $\langle k|\sigma_i|k\rangle = 1/d$ ), for any POVM  $\mathbf{N} \in \mathcal{IC}(d, n)$ ,  $\max_{\mathbf{N} \in \mathcal{IC}(d, n)} p_{\text{succ}}(\mathcal{E}_0, \mathbf{N}) = \frac{1}{n}$ . Then RHS of (11) is exactly the same as optimization problem given in (10) which completes the proof.  $\square$

The above operational characterisation of  $R_{\mathcal{IC}}$  allows us to identify measurements with the greatest possible robustness (hence relative advantage for MESD) in a given dimension  $d$ .

**Proposition 3** (Maximal advantage of coherent measurements over incoherent measurements). *The following equality holds*

$$\max_{\mathbf{M} \in \mathcal{P}(d, n)} R_{\mathcal{IC}}(\mathbf{M}) = \min\{d, n\} - 1. \quad (12)$$

Moreover, when  $n \geq d$  the measurement  $\mathbf{M}$  attains the maximal value of  $\mathbb{R}_{\mathcal{F}}$  if and only if its effects are proportional to maximally coherent states i.e pure states of the form  $|\psi\rangle = (1/\sqrt{d}) \sum_{j=1}^d \exp(i\theta_j)|j\rangle$ .

*Proof.* From Lemma 2 it follows that  $R_{\mathcal{IC}}(\mathbf{M}) + 1 = \sum_i \text{tr}(M_i \sigma_i)$ , where  $\sigma_i$  cannot be classically distinguished. Since for any  $\rho$  we have  $\text{tr}(M_i \rho) \leq \text{tr}(M_i)$

we obtain  $R_{\mathcal{IC}}(\mathbf{M}) + 1 \leq \sum_i \text{tr}(M_i) = d$ . Moreover, if  $n \geq d$  this bound can be easily saturated for projective measurements in the Fourier basis  $|\tilde{j}\rangle = \frac{1}{\sqrt{d}} \sum_{k=0}^{d-1} \exp\left(\frac{i2\pi jk}{d}\right) |k\rangle$  with ensemble  $\mathcal{E}_0 = \left\{\frac{1}{d}, |\tilde{j}\rangle\right\}_{j=1}^d$ .

For  $n < d$ , we will use (10) to get the missing upper bound. We first note that since operators  $M_a$  are upper bounded by  $\mathbb{1}$  we have  $\text{tr}(Z_a M_a) \leq \text{tr}(Z_a) = 1$ . Therefore, from (10) we can conclude that  $R_{\mathcal{IC}}(\mathbf{M}) = \sum_a \text{tr}(Z_a M_a) \leq \sum_a \text{tr}(Z_a) = n$ , hence  $R_{\mathcal{IC}}(\mathbf{M}) \leq n - 1$ . Combining this with the previous case we conclude  $R_{\mathcal{IC}}(\mathbf{M}) = \min\{d, n\} - 1$ . Just like in the case  $n \geq d$  we will proof the saturation of this bound using Fourier transform of the original basis  $|\tilde{j}\rangle = \frac{1}{\sqrt{d}} \sum_{k=0}^{d-1} e^{\frac{i2\pi jk}{d}} |k\rangle$ . To prove the equality for the case  $n < d$ , we choose an ensemble  $\tilde{\mathcal{E}}$  uniformly distributed over first  $n$  orthonormal states from ensemble  $\mathcal{E}_0$ , i.e.,  $\tilde{\mathcal{E}} = \left\{\frac{1}{n}, |\tilde{j}\rangle\right\}_{j=1}^n$  and an  $n$  outcome POVM  $\tilde{\mathbf{M}} := \left\{\left\{|\tilde{j}\rangle\langle\tilde{j}|\right\}_{i=1}^{n-1}, \mathbb{I} - \sum_{j=1}^{n-1} |\tilde{j}\rangle\langle\tilde{j}|\right\}$ .

For this particular case, again RHS of (11) reduce to  $n - 1$ . Combining these two cases we obtain  $\max_{\mathbf{M} \in \mathcal{P}(d, n)} R_{\mathcal{IC}}(\mathbf{M}) = \min\{d, n\} - 1$ .

Now we will show that for  $n \geq d$  any POVM for which (11) holds, have the effects directly proportional to  $|\psi\rangle\langle\psi|$  where  $|\psi\rangle = (1/\sqrt{d}) \sum_{j=1}^d \exp(i\theta_j) |j\rangle$ . This follows from the simple observation that for  $n \geq d$ , maximization of (10)  $\forall a$ ,  $\text{tr}(Z_a M_a) = \text{tr}(M_a)$  requires that  $Z_a = P_{\text{Supp}(M_a)} + B_a$ , where  $\text{tr}(B_a P_{\text{Supp}(M_a)}) = 0$ . Moreover, to satisfy the constraint of (10),  $P_{\text{Supp}(M_a)}$  has to be one dimensional. This implies only feasible solution will be

$$M_a = \alpha_a \Pi_a \quad Z_a = \Pi_a, \quad (13)$$

where  $\Pi_a$  is rank one projector with trace 1. Then to be a valid POVM,

$$\sum_a \alpha_a \Pi_a = \mathbb{1} \Rightarrow \sum_a \alpha_a \langle i | \Pi_a | i \rangle = 1. \quad (14)$$

Combining (14) and (10) we obtain  $\langle i | \Pi_a | i \rangle = \frac{1}{d}$  for every  $a$ . This leads to the conclusion that the POVM  $\mathbf{M}$  with  $n \geq d$  outcomes and satisfying  $R_{\mathcal{IC}}(\mathbf{M}) = d - 1$  must have effects proportional to  $|\psi\rangle\langle\psi|$  where  $|\psi\rangle = (1/\sqrt{d}) \sum_{j=1}^d \exp(i\theta_j) |j\rangle$ .  $\square$

## 4.2 Separable measurements

We now turn our attention to the measurements that can be performed on a composite quantum systems. In what follows we will study the maximal advantage that entangled measurements can offer for MESD over so-called separable measurements on

multiparticle quantum systems. Separable measurements contain the set of LOCC measurements and are relatively easy to characterize. For this reason the maximal advantage with respect to separable measurements is in general smaller than the advantage with respect to LOCC measurements.

Consider a bipartite quantum system  $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ . We define a POVM  $\mathbf{F}$  as separable if all its effects  $\{F_k\}_{k=1}^n$  are separable operators i.e.  $F_i = \sum_k Q_i^k \otimes \tilde{Q}_i^k$  where  $\{Q_i^k\}_{i=1}^n$  and  $\{\tilde{Q}_i^k\}_{i=1}^n$  are positive operators on  $\mathcal{H}_A$  and  $\mathcal{H}_B$  respectively. The above definition can be easily extended to system of multiple qubits  $\mathcal{H}_N = (\mathbb{C}^2)^{\otimes N}$ . We denote the set of bipartite separable measurements by  $\text{Sep}(AB)$  and separable multiqubit measurements by  $\text{Sep}(N)$ . Recall that we made the dependence on number of outcomes  $n$  implicit. In what follows we will assume that  $n$  is fixed and equals at least the dimension of the relevant Hilbert space ( $d_A d_B$  and  $2^N$  respectively). It easily follows that  $\text{Sep}(AB)$  and  $\text{Sep}(N)$  are convex compact subsets of  $\mathcal{P}(d, n)$  for suitable Hilbert spaces. The following Lemma gives the comprehensive answer to the question about the maximal advantage for MESD that entangled measurements can offer over separable measurements in both bipartite and multiparty scenario.

**Proposition 4** (Maximal advantage of entangled measurements over separable measurements). *Let  $d_A$  and  $d_B$  denote the dimensions of local spaces in the bipartite Hilbert space  $\mathcal{H}_{AB}$ . Then, the following inequalities hold*

$$\min\{d_A, d_B\} - 1 \leq \max_{\mathbf{M}} R_{\text{Sep}(AB)}(\mathbf{M}) \leq \min\{d_A, d_B\}, \quad (15)$$

where the maximization is over all measurements on  $\mathcal{H}_{AB}$ .

Moreover, the maximal robustness of entangled qubit measurement increases exponentially with the size of the system. Specifically, for sufficiently large  $N$  we have

$$c \frac{2^N}{8N^2} - 1 \leq \max_{\mathbf{M}} R_{\text{Sep}(N)}(\mathbf{M}) \leq 2^{\frac{3}{2}N-1} - 1, \quad (16)$$

where the maximization is over all measurements on  $\mathcal{H}_N$  and  $c > 0.7$ .

*Sketch of the proof.* The upper bounds in (15) and (16) follow from application of explicit simulation strategy based on (global) depolarizing map  $\Phi_t$ . This map, transforms effects of a measurement  $\mathbf{M}$  as follows:  $\Phi_t(M_i) = tM_i + (1-t)\frac{\text{tr}(M_i)}{d}\mathbb{1}$ . It is easy to check that the resulting operators still form a POVM on a total space. Because this map preserves traces of the effects we can write  $\Phi_t(M_i) =$

$\text{tr}(\mathbf{M}_i) (\Phi_t(\rho_i))$ , where  $\rho_i = M_i/\text{tr}(M_i)$  is a quantum state. To get an upper bound on the robustness of  $\mathbf{M}$  it suffices to find lower bound on  $t_*$  defined as the number for which *all states*  $\rho_i$  constructed from the effects of  $M_i$  are separable. To find the upper bound we can directly use the results from entanglement theory. Concretely, the problem of how much "white noise" can be added to a quantum state before it becomes separable quantum states has been studied before. Specifically the maximal value of "random robustness" (minimal  $s = 1/t - 1$  such that  $\Phi_t(\rho)$  is separable) equals  $\min\{d_A, d_B\} - 1$  in the biparticle scenario [19] and can be upper bounded by  $2^{\frac{3}{2}N-1} - 1$  for the multipartite case [54]. To obtain the lower bounds we use characterization of the robustness MESD given in Eq.(5) by showing advantage of entangled measurements over separable ones for particular uniform ensembles of quantum states. for biparticle case we consider ensemble consisting of orthogonal maximally states [29], while for multipartite case we consider ensemble consisting of  $2^N$  iid random quantum states taken from the Haar measure on  $(\mathbb{C}^2)^{\otimes N}$  [55] (see Appendix for details).  $\square$

*Remark.* The above results complement existing results [29, 56, 57] on the relative power of LOCC separable, and global state discrimination. Our focus was on the advantage in discrimination of arbitrary ensembles of quantum states while previously the emphasis was put mostly on ensembles consisting of two states (for example in the context of quantum data hiding).

## 5 Open problems

We finish with giving possible directions of further study. First, it is tempting to ask if non-free character of measurements can be certified in a device-independent manner in prepare-and measure scenarios [58, 59]. Second interesting problem concerns finding the maximal that resourceful POVMs advantage for state discrimination problems for other important classes of measurements such as LOCC measurements [29, 30] or projective-simulable measurements [43, 44]. The potential lack of the increasing separation (as  $d \rightarrow \infty$ ) between projective and general POVMs will likely have strong consequences for quantum information processing and Bell nonlocality. Finally, it is natural to ask if analogous operational interpretations hold also for other classes of quantum objects such as quantum channels or quantum instruments.

## Acknowledgments

We thank Zbigniew Puchała, Michał Horodecki, Paul Skrzypczyk and Sibasish Ghosh, for interesting and stimulating discussions. We are grateful to Filip Maciejewski for careful reading of the manuscript and his valuable comments. We acknowledge the support of Homing programme of the Foundation for Polish Science co-financed by the European Union under the European Regional Development Fund. Note added— Upon the completion of this work we became aware of two works [60, 61] that independently obtained results overlapping with Theorem 2 presented here. Moreover, similar methods were used in the context of measurement incompatibility [62, 63].

## References

- [1] E. Chitambar and G. Gour, [arXiv:1806.06107 \(2018\)](#).
- [2] B. Coecke, T. Fritz, and R. W. Spekkens, [Information and Computation](#) **250**, 59 (2016).
- [3] F. G. S. L. Brandão, M. Horodecki, J. Oppenheim, J. M. Renes, and R. W. Spekkens, [Phys. Rev. Lett.](#) **111**, 250404 (2013).
- [4] J. Goold, M. Huber, A. Riera, L. del Rio, and P. Skrzypczyk, [Journal of Physics A: Mathematical and Theoretical](#) **49**, 143001 (2016).
- [5] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, [Phys. Rev. Lett.](#) **78**, 2275 (1997).
- [6] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, [Rev. Mod. Phys.](#) **81**, 865 (2009).
- [7] A. Grudka, K. Horodecki, M. Horodecki, P. Horodecki, R. Horodecki, P. Joshi, W. Kłobus, and A. Wójcik, [Phys. Rev. Lett.](#) **112**, 120401 (2014).
- [8] J. Barrett, [Phys. Rev. A](#) **75**, 032304 (2007).
- [9] R. Gallego and L. Aolita, [Phys. Rev. X](#) **5**, 041008 (2015).
- [10] V. Veitch, C. Ferrie, D. Gross, and J. Emerson, [New Journal of Physics](#) **14**, 113011 (2012).
- [11] S. M. Barnett and S. Croke, [Adv. Opt. Photon.](#) **1**, 238 (2009).
- [12] J. Bae and L.-C. Kwek, [J. Phys. A: Math. Gen.](#) **48**, 083001 (2015).
- [13] N. Brunner, M. Navascués, and T. Vértesi, [Phys. Rev. Lett.](#) **110**, 150501 (2013).
- [14] I. Marvian and R. W. Spekkens, [Phys. Rev. A](#) **94**, 052324 (2016).
- [15] J. Bae, W.-Y. Hwang, and Y.-D. Han, [Phys. Rev. Lett.](#) **107**, 170403 (2011).
- [16] C. Lomont, [arxiv:0411037 \(2005\)](#).

- [17] P. Skrzypczyk and N. Linden, *Phys. Rev. Lett.* **122**, 140403 (2019).
- [18] R. Takagi, B. Regula, K. Bu, Z.-W. Liu, and G. Adesso, *Phys. Rev. Lett.* **122**, 140402 (2019).
- [19] G. Vidal and R. Tarrach, *Phys. Rev. A* **59**, 141 (1999).
- [20] F. G. S. L. Brandão and G. Gour, *Phys. Rev. Lett.* **115**, 070503 (2015).
- [21] B. Regula, *Journal of Physics A: Mathematical and Theoretical* **51**, 045303 (2017).
- [22] J. Aberg, [arXiv:quant-ph/0612146](https://arxiv.org/abs/quant-ph/0612146) (2006).
- [23] T. Baumgratz, M. Cramer, and M. B. Plenio, *Phys. Rev. Lett.* **113**, 140401 (2014).
- [24] A. Streltsov, G. Adesso, and M. B. Plenio, *Rev. Mod. Phys.* **89**, 041003 (2017).
- [25] S. Bandyopadhyay and M. Nathanson, *Phys. Rev. A* **88**, 052313 (2013).
- [26] J. Watrous, *The Theory of Quantum Information* (Cambridge University Press, 2018).
- [27] E. Bagan, M. A. Ballester, R. Muñoz Tapia, and O. Romero-Isart, *Phys. Rev. Lett.* **95**, 110504 (2005).
- [28] V. Kargin, *The Annals of Statistics* **33**, 959 (2004).
- [29] M. Nathanson, *J. Math. Phys.* **46**, 141 (2005).
- [30] E. Chitambar, D. Leung, L. Mančinska, M. Ozols, and A. Winter, *Communications in Mathematical Physics* **328**, 303 (2014).
- [31] G. M. D’Ariano, P. Lo Presti, and M. G. A. Paris, *Phys. Rev. Lett.* **87**, 270404 (2001).
- [32] M. Horodecki and J. Oppenheim, *Nat. Commun.* **4** (2013), 10.1038/ncomms3059.
- [33] R. W. Spekkens, D. H. Buzacott, A. J. Keehn, B. Toner, and G. J. Pryde, *Phys. Rev. Lett.* **102**, 010401 (2009).
- [34] D. Cavalcanti, P. Skrzypczyk, and I. Šupić, *Phys. Rev. Lett.* **119**, 110501 (2017).
- [35] A. Winter and D. Yang, *Phys. Rev. Lett.* **116**, 120404 (2016).
- [36] T. Biswas, M. G. Diaz, and A. Winter, *Proc. R. Soc. A* **473** (2017), 10.1098/rspa.2017.0170.
- [37] M. Piani, M. Cianciaruso, T. R. Bromley, C. Napoli, N. Johnston, and G. Adesso, *Phys. Rev. A* **93**, 042107 (2016).
- [38] C. Napoli, T. R. Bromley, M. Cianciaruso, M. Piani, N. Johnston, and G. Adesso, *Phys. Rev. Lett.* **116**, 150502 (2016).
- [39] M. Piani and J. Watrous, *Phys. Rev. Lett.* **114**, 060404 (2015).
- [40] M. Piani, M. Cianciaruso, T. R. Bromley, C. Napoli, N. Johnston, and G. Adesso, *Phys. Rev. A* **93**, 042107 (2016).
- [41] T. Heinosaari, J. Kiukas, and D. Reitzner, *Phys. Rev. A* **92**, 022115 (2015).
- [42] L. Guerini, J. Bavaresco, M. Terra Cunha, and A. Acín, *J. Math. Phys.* **58**, 092102 (2017).
- [43] M. Oszmaniec, L. Guerini, P. Wittek, and A. Acín, *Phys. Rev. Lett.* **119**, 190501 (2017).
- [44] M. Oszmaniec, F. B. Maciejewski, and Z. Puchała, [arXiv:1807.08449](https://arxiv.org/abs/1807.08449) (2018).
- [45] M. Kleinmann and A. Cabello, *Phys. Rev. Lett.* **117**, 150401 (2016).
- [46] G. M. D’Ariano, P. L. Presti, and P. Perinotti, *J. Phys. A: Math. Gen.* **38**, 5979 (2005).
- [47] F. G. Brandao, *Phys. Rev. A* **72**, 022310 (2005).
- [48] M. Hillery, *Phys. Rev. A* **93**, 012111 (2016).
- [49] F. Bischof, H. Kampermann, and D. Bruss, [arXiv:1812.00018](https://arxiv.org/abs/1812.00018) (2018).
- [50] T. Theurer, D. Egloff, L. Zhang, and M. B. Plenio, [arXiv:1806.07332](https://arxiv.org/abs/1806.07332) (2018).
- [51] M. Bera, [arXiv:1809.02578](https://arxiv.org/abs/1809.02578) (2018).
- [52] L. Vandenberghe and S. Boyd, *SIAM Review* **38**, 49 (1996).
- [53] K. Korzekwa, S. Czachorski, Z. Puchała, and K. Życzkowski, [arXiv:1812.09083](https://arxiv.org/abs/1812.09083) (2018).
- [54] L. Gurvits and H. Barnum, *Phys. Rev. A* **68**, 042312 (2003).
- [55] A. Montanaro, *Commun Math Phys* **273**, 619 (2007).
- [56] P. Hayden, D. Leung, P. W. Shor, and A. Winter, *Communications in Mathematical Physics* **250**, 371 (2004).
- [57] W. Matthews, S. Wehner, and A. Winter, *Communications in Mathematical Physics* **291**, 813 (2009).
- [58] A. Tavakoli, M. Smania, T. Vértesi, D. Rosset, and N. Brunner, [arXiv:1811.12712](https://arxiv.org/abs/1811.12712) (2018).
- [59] P. Mironowicz and M. Pawłowski, [arXiv:1811.12872](https://arxiv.org/abs/1811.12872) (2018).
- [60] R. Uola, T. Kraft, J. Shang, X.-D. Yu, and O. Guhne, *Phys. Rev. Lett.* **122**, 130404 (2019).
- [61] R. Takagi and B. Regula, [arXiv:1901.08127](https://arxiv.org/abs/1901.08127) (2019).
- [62] C. Carmeli, T. Heinosaari, and A. Toigo, *Phys. Rev. Lett.* **122**, 13040 (2019).
- [63] P. Skrzypczyk, I. Šupić, and D. Cavalcanti, *Phys. Rev. Lett.* **122**, 130403 (2019).
- [64] D. Cavalcanti and P. Skrzypczyk, *Reports on Progress in Physics* **80**, 024001 (2016).
- [65] D. Gross, S. T. Flammia, and J. Eisert, *Phys. Rev. Lett.* **102**, 190501 (2009).

# Appendix

In what follows we give proofs that were omitted in the main part of the manuscript. In Part A we give proofs concerning the general properties of the robustness  $R_{\mathcal{F}}$  and its dual formulation. In Part B we give justifications of results specific to particular scenarios of incoherent and separable measurements.

## A Proofs of general results

For the convenience of the reader we first recall the definition of robustness for general convex and compact subset of free measurements  $\mathcal{F} \subset \mathcal{P}(d, n)$

$$R_{\mathcal{F}}(\mathbf{M}) := \min \left\{ s \mid \exists \mathbf{N} \text{ such that } \frac{\mathbf{M} + s\mathbf{N}}{1 + s} \in \mathcal{F} \right\}. \quad (17)$$

**Proposition 1** (Properties of robustness). *Let  $\mathcal{F} \subset \mathcal{P}(d, n)$  be a compact convex set of free measurements. The robustness  $R_{\mathcal{F}}$  satisfies the following properties: (i) Faithfulness:  $R_{\mathcal{F}}(\mathbf{M}) \geq 0$  and  $R_{\mathcal{F}}(\mathbf{M}) > 0$  iff  $\mathbf{M} \notin \mathcal{F}$ , (ii) Convexity:  $R_{\mathcal{F}}(p\mathbf{M} + (1-p)\mathbf{M}') \leq pR_{\mathcal{F}}(\mathbf{M}) + (1-p)R_{\mathcal{F}}(\mathbf{M}')$ , (iii) Monotonicity under free operations:  $R_{\mathcal{F}}[\varphi(\mathbf{M})] \leq R_{\mathcal{F}}(\mathbf{M})$  for all free operations  $\varphi \in \mathcal{O}$ .*

*Proof.*

(i) *Faithfulness.* This property follows directly from the definition given in Eq.(17).

(ii) *Convexity.* From the definition of the robustness (17) it follows that for two POVMs  $\mathbf{M}, \mathbf{M}'$  there exist POVMs  $\mathbf{N}, \mathbf{N}' \in \mathcal{P}(d, n)$ , such that

$$\mathbf{M} = (1 + R_{\mathcal{F}}(\mathbf{M}))\tilde{\mathbf{M}} - R_{\mathcal{F}}(\mathbf{M})\mathbf{N}, \quad (18)$$

$$\mathbf{M}' = (1 + R_{\mathcal{F}}(\mathbf{M}'))\tilde{\mathbf{M}}' - R_{\mathcal{F}}(\mathbf{M}')\mathbf{N}', \quad (19)$$

where  $\mathbf{N}, \mathbf{N}' \in \mathcal{F}$ . We therefore get that for all  $p \in [0, 1]$

$$\begin{aligned} p\mathbf{M} + (1-p)\mathbf{M}' &= p[(1 + R_{\mathcal{F}}(\mathbf{M}))\tilde{\mathbf{M}} - R_{\mathcal{F}}(\mathbf{M})\mathbf{N}] \\ &\quad + (1-p)[(1 + R_{\mathcal{F}}(\mathbf{M}'))\tilde{\mathbf{M}}' - R_{\mathcal{F}}(\mathbf{M}')\mathbf{N}']. \end{aligned} \quad (20)$$

After defining

$$\tilde{r} := pR_{\mathcal{F}}(\mathbf{M}) + (1-p)R_{\mathcal{F}}(\mathbf{M}'), \quad (21)$$

$$\tilde{\mathbf{N}} := \frac{pR_{\mathcal{F}}(\mathbf{M})\mathbf{N} + (1-p)R_{\mathcal{F}}(\mathbf{M}')\mathbf{N}'}{\tilde{r}}, \quad (22)$$

$$\mathbf{T} := \frac{p(\tilde{\mathbf{M}} + R_{\mathcal{F}}(\mathbf{M})\tilde{\mathbf{M}}) + (1-p)(\tilde{\mathbf{M}}' + R_{\mathcal{F}}(\mathbf{M}')\tilde{\mathbf{M}}')}{1 + \tilde{r}}, \quad (23)$$

we note that  $\tilde{\mathbf{N}} \in \mathcal{P}(d, n)$ , while  $\mathbf{T} \in \mathcal{F}$ . We finally observe that

$$p\mathbf{M} + (1-p)\mathbf{M}' = (1 + \tilde{r})\mathbf{T} - \tilde{r}\tilde{\mathbf{N}}, \quad (24)$$

which directly implies  $\tilde{r} = pR_{\mathcal{F}}(\mathbf{M}) + (1-p)R_{\mathcal{F}}(\mathbf{M}') \geq R_{\mathcal{F}}(p\mathbf{M} + (1-p)\mathbf{M}')$ .

(iii) *Monotonicity under free operations.* From the definition we have that for every  $\mathbf{M}$  there exist  $\mathbf{N} \in \mathcal{P}(d, n)$  and  $\tilde{\mathbf{M}} \in \mathcal{F}$  such that

$$\mathbf{M} = (1 + R_{\mathcal{F}}(\mathbf{M}))\tilde{\mathbf{M}} - R_{\mathcal{F}}(\mathbf{M})\mathbf{N}. \quad (25)$$

Applying  $\varphi$  to both sides of the above equality we obtain

$$\varphi(\mathbf{M}) = (1 + R_{\mathcal{F}}(\mathbf{M}))\varphi(\tilde{\mathbf{M}}) - R_{\mathcal{F}}(\mathbf{M})\varphi(\mathbf{N}). \quad (26)$$

Since, by assumption,  $\varphi(\tilde{\mathbf{M}}) \in \mathcal{F}$ , equation (26) directly implies  $R_{\mathcal{F}}[\varphi(\mathbf{M})] \leq R_{\mathcal{F}}(\mathbf{M})$ .  $\square$

**Lemma 1.** Let  $\mathbf{M} \in \mathcal{P}(d, n)$  and  $\mathcal{F} \subset \mathcal{P}(d, n)$  be a compact convex set of free measurements. Then, the generalized robustness  $R_{\mathcal{F}}(\mathbf{M})$  can be expressed as the following optimisation problem.

$$\begin{aligned} & \text{maximize} && \sum_{i=1}^n \text{tr}(Z_i M_i) - 1 \\ & \text{subject to} && Z_i \geq 0, \quad i = 1, \dots, n, \\ & && \sum_{i=1}^n \text{Tr}(Z_i N_i) \leq 1 \quad \forall \mathbf{N} \in \mathcal{F}. \end{aligned} \quad (27)$$

*Proof.* Our proof strategy is the following. The problem in (27) belongs to the class of semi-infinite optimisation problems [47]. We will compute its dual and argue that there is no duality gap. Finally, the dual problem will turn out to be equivalent to the original definition of robustness from (17). After simple algebraic manipulations we obtain that (27) is equivalent to

$$\begin{aligned} & - \min_{\tilde{\mathbf{X}}} \left\{ \sum_i \text{tr}(X_i M_i) - 1 : \forall i, X_i \leq 0, \forall \mathbf{N} \in \mathcal{F} \quad \sum_i \text{Tr}(X_i N_i) \geq -1 \right\} \\ & = - \min_{\tilde{\mathbf{X}}} \left\{ \sum_i \text{tr}(\tilde{X}_i M_i) : \forall i, X_i \leq \frac{\mathbb{I}}{d} \quad \forall \mathbf{N} \in \mathcal{F} \quad \sum_i \text{Tr}(\tilde{X}_i N_i) \geq 0 \right\}. \end{aligned} \quad (28)$$

The *Lagrangian* corresponding to the last problem in (28) reads

$$\mathcal{L}(\tilde{\mathbf{X}}, \mathbf{G}, \{k(\mathbf{N})\}) = \sum_{i=1}^n \left( \text{tr}(\tilde{X}_i M_i) + \text{tr}(X_i - \frac{\mathbb{I}}{d}) G_i - \int d\mathbf{N} k(\mathbf{N}) \text{tr}(N_i X_i) \right), \quad (29)$$

where  $\mathbf{G} = (G_1, \dots, G_n)$  and  $k(\mathbf{N})$  are lagrange multipliers which satisfy  $G_i \geq 0 \forall i$  and  $k(\mathbf{N}) \geq 0$  for almost all  $\mathbf{N} \in \mathcal{F}$ . To construct the dual problem to (29) we minimize  $\mathcal{L}(\tilde{\mathbf{X}}, \mathbf{G}, \{k(\mathbf{N})\})$  over primal variables  $\tilde{\mathbf{X}}$ . Since the *Lagrangian* is affine function of  $\tilde{\mathbf{X}}$ , we obtain minimum equal to  $\infty$  unless  $\mathbf{G} + \mathbf{M} = \int k(\mathbf{N}) \mathbf{N} d\mathbf{N}$ . If the condition holds then we obtain  $-\frac{1}{d} \sum_{i=1}^n \text{tr}(G_i)$ . Hence we see that the dual of the problem (27) can be written as

$$\begin{aligned} & \text{minimize} && \frac{1}{d} \sum_{i=1}^n \text{tr}(G_i) \\ & \text{subject to} && \mathbf{G} + \mathbf{M} = \int k(\mathbf{N}) \mathbf{N} d\mathbf{N}, \\ & && G_i \geq 0 \quad \forall i, \\ & && k(\mathbf{N}) \geq 0 \quad \forall \mathbf{N} \in \mathcal{F}. \end{aligned} \quad (30)$$

We note that the primal problem (27) possesses a strictly feasible points (for example  $\mathbf{Z} = (\frac{\lambda \mathbb{I}}{d}, \dots, \frac{\lambda \mathbb{I}}{d})$ , for  $\lambda \in (0, 1)$ ), which follows from the assumption of the compactness of  $\mathcal{F}$  and the form of the constrains in (27). Finally, we conclude the proof by noting that  $\mathbf{G} + \mathbf{M} = \int k(\mathbf{N}) \mathbf{N} d\mathbf{N}$  iff  $\mathbf{G} = s\mathbf{M}'$  where  $\mathbf{M}' \in \mathcal{P}(d, n)$  and  $s = \sum_i \frac{\text{tr} G_i}{d}$ .  $\square$

## B Proofs of results that concern particular classes of free measurements

### B.1 Incoherent measurements

From the definition of robustness (see Eq.(17)) and the definition of incoherent measurements we obtain that  $R_{\mathcal{IC}}(\mathbf{M})$  can be cast as the following SDP program

$$\begin{aligned} & \text{minimize} && s \\ & \text{s.t} && \frac{M_a + s N_a}{1 + s} = \sum_i q(a|i) |i\rangle \langle i| \quad \forall a, \\ & && N_a \geq 0, \quad a = 1, \dots, n. \quad \sum_a N_a = \mathbb{I}, \\ & && \sum_a q(a|i) (1 + s) = b \quad \forall i. \end{aligned} \quad (31)$$

**Lemma 2** (Dual characterization of robustness w.r.t. incoherent measurement). *Robustness of a POVM  $\mathbf{M} \in \mathcal{P}(d, n)$  w.r.t set of incoherent measurement  $\mathcal{IC}(d, n)$ , i.e,  $R_{\mathcal{IC}}(\mathbf{M})$  can be expressed as the solution of optimization problem given by*

$$\begin{aligned} & \text{maximize} && \sum_{a=1}^n \text{tr}(Z_a M_a) - 1 \\ & \text{subject to} && \forall i, a \quad Z_a \geq 0 \quad \langle i|Z_a|i \rangle = \langle i|Z_n|i \rangle, \\ & && \text{tr}(Z_n) = 1. \end{aligned} \tag{32}$$

*Proof.* Introducing variables  $\tilde{q}(a|i) := (1+s)q(a|i) \forall i$  and  $\forall a \in [1, (n-1)]$  we obtain an equivalent form of problem from (31)

$$\begin{aligned} & \text{minimize} && b - 1 \\ & \text{subject to} && - \sum_i \tilde{q}(a|i)|i\rangle\langle i| + M_a \leq 0 \quad \forall a \in [1, (n-1)], \\ & && - b\mathbb{I} + \sum_i \sum_{a=1}^{n-1} \tilde{q}(a|i)|i\rangle\langle i| + M_n \leq 0. \end{aligned} \tag{33}$$

To construct the dual of the optimization problem [52, 64] given in (33), we start by writing down the *Lagrangian* associated with primal problem

$$\mathcal{L}(\mathbf{Z}, \tilde{q}(a|i)) := (b-1) + \text{tr}(Z_n(-b\mathbb{I} + \sum_i \sum_{a=1}^{n-1} \tilde{q}(a|i)|i\rangle\langle i| + M_n)) + \sum_{a=1}^{n-1} \text{tr}(Z_a(- \sum_i \tilde{q}(a|i)|i\rangle\langle i| + M_a)), \tag{34}$$

where we have introduced  $\mathbf{Z} = (Z_1, Z_2, \dots, Z_n)$  as dual variable (Lagrange multipliers). To ensure that the *Lagrangian* is always smaller than the objective function whenever the constraints of the (33) are satisfied, we impose  $Z_j \geq 0$  and  $\langle i|Z_j|i \rangle = \langle i|Z_n|i \rangle \forall j \in [1, n]$ . To ensure that the *Lagrangian* is independent of primal variables we have  $\text{tr}(Z_n) = 1$ . Finally, the corresponding dual problem which sets the upper bound for primal optimization problem is given by

$$\begin{aligned} & \text{maximize} && \sum_{a=1}^n \text{tr}(Z_a M_a) - 1 \\ & \text{subject to} && \forall i, a \quad Z_a \geq 0 \quad \langle i|Z_a|i \rangle = \langle i|Z_n|i \rangle, \\ & && \text{tr}(Z_n) = 1. \end{aligned} \tag{35}$$

By inspection one can conclude that strong duality holds, since  $\mathbf{Z}_0 = (\frac{\mathbb{I}}{d}, \dots, \frac{\mathbb{I}}{d})$  satisfy the constraints in (35) and is positive-definite. Therefore, the solutions (33) and (35) are equal.  $\square$

## B.2 Separable measurements

**Proposition 4** (Maximal advantage of entangled measurements over separable measurements). *Let  $d_A$  and  $d_B$  denote the dimensions of local spaces in the biparticle Hilbert space  $\mathcal{H}_{AB}$ . Then, the following inequalities hold*

$$\min\{d_A, d_B\} - 1 \leq \max_{\mathbf{M}} R_{\text{Sep}(AB)}(\mathbf{M}) \leq \min\{d_A, d_B\}, \tag{36}$$

where the maximization is over all measurements on  $\mathcal{H}_{AB}$ .

Moreover, the maximal robustness of entangled qubit measurement increases exponentially with the size of the system. Specifically, for sufficiently large  $N$  we have

$$c \frac{2^N}{8N^2} \leq \max_{\mathbf{M}} R_{\text{Sep}(N)}(\mathbf{M}) \leq 2^{\frac{3}{2}N-1} - 1, \tag{37}$$

where the maximization is over all measurements on  $\mathcal{H}_N$  and  $c > 0.7$ .

*Proof of lower bounds in Eq.(36) and Eq.(37).* Our strategy to prove the lower bounds is to construct ensembles  $\mathcal{E}$  for which the relative advantage

$$\frac{p_{\text{succ}}(\mathcal{E}, \mathbf{M})}{\max_{\mathbf{N} \in \text{Sep}(\mathcal{H}_A: \mathcal{H}_B)} p_{\text{succ}}(\mathcal{E}, \mathbf{N})} \quad (38)$$

of some entangled measurement  $\mathbf{M}$  over the separable ones will be large. (i) *Biparticle case* Let  $D = \min\{d_A, d_B\}$ . Following Nathanson [29], we consider an uniform ensemble  $\mathcal{E}_0 = \{\frac{1}{D^2}, \Psi_{nm}\}_{n,m=1}^d$  consisting orthogonal maximally entangled states from  $\mathcal{H}_A \otimes \mathcal{H}_B$ ,

$$|\Psi_{nm}\rangle = \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} e^{\frac{i2\pi jn}{d}} |j\rangle \otimes |(j+m) \bmod d\rangle . \quad (39)$$

Since vectors  $|\Psi_{nm}\rangle$  are orthogonal, they can be distinguished perfectly by measurement  $\mathbf{M}_0$  whose effects are of the form  $M_{nm} = \Psi_{nm} + R_{nm}$ , where  $R_{nm}$  are operators orthogonal to states  $|\Psi_{nm}\rangle$ . On the other hand there is a limit of distinguishability of states from  $E_0$  via separable measurements because of the well-known property (see for example Lemma 2 of [25]) stating that for all biparticle quantum states  $|Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$  and positive separable operators  $T$

$$\text{tr}(\Psi T) \leq \lambda^2 \text{tr}(T) , \quad (40)$$

where  $\lambda$  is the greatest Schmidt coefficient of  $|\Psi\rangle$ . Let  $\mathbb{P}_0$  be a projector onto a subspace  $\tilde{H}_{AB} \subset \mathcal{H}_{AB}$  spanned by vectors  $|i\rangle|j\rangle$ , where  $i, j = 0, \dots, D-1$ . Importantly,  $\mathbb{P}_0$  is a separable operator. Applying inequality (40) to our problem we obtain that for all  $D$ -outcome separable measurements  $\mathbf{N}$  we can upper bound  $p_{\text{succ}}$  as follows

$$p_{\text{succ}}(\mathcal{E}_0, \mathbf{N}) = \frac{1}{D^2} \sum_{n,m=1}^D \text{tr}(\Psi_{nm} N_{nm}) = \frac{1}{D^2} \sum_{n,m=1}^D \text{tr}(\Psi_{nm} \mathbb{P}_0 N_{nm} \mathbb{P}_0) \leq \frac{1}{D} \frac{1}{D^2} \sum_{n,m=1}^D \text{tr}(\mathbb{P}_0 N_{nm} \mathbb{P}_0) = \frac{1}{D} , \quad (41)$$

where we have used the separability of operators  $\mathbb{P}_0 N_{nm} \mathbb{P}_0$  and the fact that states  $\Psi_{nm}$  are maximally entangled and hence  $\lambda = 1/\sqrt{D}$ . Using the operational interpretation of the robustness  $R_{\text{Sep}(AB)}$  we conclude the proof by observing that

$$R_{\text{Sep}(AB)}(\mathbf{M}_0) \geq \frac{p_{\text{succ}}(\mathcal{E}_0, \mathbf{M}_0)}{\max_{N \in \text{Sep}(AB)} p_{\text{succ}}(\mathcal{E}_0, \mathbf{N})} - 1 \geq D - 1 . \quad (42)$$

(ii) *Multiparticle case* We proceed analogously to the biparticle case. Consider a random uniform ensemble of  $M = 2^N$  Haar random, independent and identically distributed pure states in  $\mathcal{H}_N$ ,

$$\mathcal{E}_{\text{Haar}} = \left\{ \frac{1}{2^N}, \Psi_i \right\}_{i=1}^{2^N} . \quad (43)$$

The above ensemble is random as states composing it are independent random pure states i.e.  $\Psi_i = U_i \Psi_0 U_i^\dagger$ , for  $U_i$  distributed according to Haar measure on the group  $U(\mathcal{H}_N)$ . In [55] it was shown that in the limit of large  $N$ , for typical realizations of the ensemble  $\mathcal{E}_{\text{Haar}}$  the so-called pretty good measurement  $\mathbf{M}_{\text{PGM}}(\mathcal{E}_{\text{Haar}})$  is capable to distinguish states from  $\mathcal{E}_{\text{Haar}}$  with success probability bounded away from zero

$$p_{\text{succ}}(\mathcal{E}_{\text{Haar}}, \mathbf{M}_{\text{PGM}}(\mathcal{E}_{\text{Haar}})) \geq 0.7 . \quad (44)$$

Importantly, since states  $\Psi_i$  are Haar random, they are typically highly entangled and this causes them to be indistinguishable by separable measurements. To show this we will use a result that can be found in the proof of Theorem 2 from [65], which states that for Haar random pure state  $\Psi$  with probability  $P \leq \exp(-N^2)$

$$\max_{\phi \in \text{SEP}} \text{tr}(\phi \Psi) \geq \frac{8N^2}{2^N} , \quad (45)$$

where the maximum is over all separable pure states on  $\mathcal{H}_N$ . Now, since states  $\Psi_i$  are iid Haar-random, by the union bound, we get that with probability  $q \geq 1 - 2^N \exp(-N^2)$ , in the limit of large dimensions

$$\forall_i \max_{\phi \in \text{SEP}} \text{tr}(\phi \Psi_i) \leq \frac{8N^2}{2^N} . \quad (46)$$

The above bound (which holds with high probability as  $N \rightarrow \infty$ ) can be used to upper bound the success probability of discrimination of states from the ensemble. Note that for every separable effect  $N_i$  we have  $\text{tr}(\Psi_i N_i) \leq \text{tr}(N_i) \max_{\phi \in \text{SEP}} \text{tr}(\phi \Psi_i) \leq \text{tr}(N_i) \frac{8N^2}{2^N}$ . Using this we obtain

$$p_{\text{succ}}(\mathcal{E}_{\text{Haar}}, \mathbf{N}) \leq \frac{1}{2^N} \sum_{i=1}^{2^N} \text{tr}(N_i) \frac{2^N}{8N^2} = \frac{8N^2}{2^N} . \quad (47)$$

Finally, combining (44), (47), and the operational characterisation of  $R_{\text{Sep}(N)}$  we obtain the desired lower bound by realizing that there exist a realization of a random ensemble  $\mathcal{E}_{\text{Haar}}$  such that

$$\max_{\mathbf{M}} R_{\text{Sep}(N)}(\mathbf{M}) \geq \frac{p_{\text{succ}}(\mathcal{E}_{\text{Haar}}, \mathbf{M}_{\text{PGM}}(\mathcal{E}_{\text{Haar}}))}{\max_{\mathbf{N} \in \text{Sep}(N)} p_{\text{succ}}(\mathcal{E}_{\text{Haar}}, \mathbf{N})} - 1 \geq c \frac{2^N}{8N^2} - 1 . \quad (48)$$

□