Measurement incompatibility and quantum advantage in communication

Debashis Saha [1,2] Debarshi Das [1,3] Arun Kumar Das [1] Bihalan Bhattacharya [1,4] and A. S. Majumdar [1]

[1] *S. N. Bose National Centre for Basic Sciences, Block JD, Sector III, Salt Lake, Kolkata 700106, India*
[2] *School of Physics, Indian Institute of Science Education and Research Thiruvananthapuram, Kerala 695551, India*
[3] *Department of Physics and Astronomy, University College London, Gower Street, WC1E 6BT London, England, United Kingdom*
[4] *Department of Mathematical Sciences, Indian Institute of Science Education & Research (IISER) Berhampur,*
Transit Campus, Govt. ITI, NH 59, Berhampur 760 010, Ganjam, Odisha, India

Measurement incompatibility stipulates the existence of quantum measurements that cannot be carried out simultaneously on single systems. We show that the set of input-output probabilities obtained from d-dimensional classical systems assisted with shared randomness is the same as the set obtained from d-dimensional quantum strategies restricted to compatible measurements with shared randomness in any communication scenario. Thus, measurement incompatibility is necessary for quantum advantage in communication, and *any* quantum advantage (with or without shared randomness) in communication acts as a witness to the incompatibility of the measurements at the receiver's end in a semi-device-independent way. We introduce a class of communication tasks—a general version of random access codes—to witness incompatibility of an arbitrary number of quantum measurements with arbitrary outcomes acting on d-dimensional systems and provide generic upper bounds on the success metric of these tasks for compatible measurements. We identify all sets of three incompatible rank-one projective qubit measurements that random access codes can witness. Finally, we present the generic relationship between different sets of probability distributions—classical, quantum with or without shared randomness, and quantum restricted to compatible measurements with or without shared randomness—produced in communication scenarios.

I. INTRODUCTION

In quantum theory, a set of quantum measurements is called incompatible if these measurements cannot be performed simultaneously on a single copy of a quantum system [1]. The best-known example of quantum incompatibility pertains to the measurements of position and momentum of a quantum mechanical particle that cannot be measured simultaneously with arbitrary precision. The notion of measurement incompatibility is an inherent property of quantum theory that differentiates it from classical physics. Quantum measurement incompatibility is at the root of demonstrating various fundamental quantum aspects such as Bell nonlocality [2,3], Einstein-Podolsky-Rosen steering [4–8], measurement uncertainty relations [9–11], quantum contextuality [12,13], quantum violation of macrorealism [14,15], and temporal and channel steering [16–18].

Bell violation is the most compelling operational witness of incompatible measurements since it relies only on the input-output statistics of bipartite systems [3,19,20]. Measurement incompatibility can also be witnessed through Einstein-Podolsky-Rosen steering [4,5,7,21,22]. These protocols, however, rely on entanglement. Recently, witnessing of quantum measurement incompatibility in the prepare-and-measure scenario based on a state discrimination task [23] has been proposed. It is particularly noteworthy that measurement incompatibility is necessary but not sufficient for Bell violations employing fully untrusted devices [24,25], whereas incompatibility is shown to be necessary as well as sufficient

in steering with one-sided trusted devices [5,7] and in a state discrimination task with fully trusted preparations [26] (see, also, [27–29]).

Nonetheless, the generic link between measurement incompatibility and nonclassical correlations in the simplest prepare-and-measure scenario is still not fully explored. The present article is motivated towards filling this crucial gap in the literature. Moreover, the results presented here address the issue as to whether incompatible quantum measurements are necessary for probing quantum advantage in any communication scenario. Apart from answering this fundamental question in the affirmative, this work further aims to provide an operational witness of incompatibility for any set of quantum measurements of an arbitrary setting—any set of arbitrary number of measurements acting on an arbitrary but finite dimension wherein different measurements have different arbitrary number of outcomes.

Specifically, we consider the communication scenario consisting of two players, say, Alice (sender) and Bob (receiver). Alice and Bob are given inputs such that each player does not know the input of the other player. Alice, upon receiving her input, sends classical or quantum communication to Bob. Bob, upon receiving his input and the communication sent by Alice, produces the outputs. In this scenario, we show that the input-output statistics obtained from d-dimensional classical systems assisted with unlimited shared randomness are the same as those obtained from the d-dimensional quantum states and compatible quantum measurements with unlimited shared randomness. Therefore, any quantum advantage (with

or without shared randomness) in a communication task in the prepare-and-measure scenario serves as a tool to witness measurement incompatibility in a semi-device-independent way—assuming nothing about the quantum states and measurements except their dimension. Furthermore, we point out that whenever the figure of merit of any task is a convex function of the input-output statistics, its maximum value in classical communication is the same as in quantum communication with compatible measurements.

Subsequently, we focus on random access codes (RACs), a specific quantum communication task in the prepare-and-measure scenario. Based on the operational figure of merit of this task, we propose a witness of measurement incompatibility of a set of arbitrary number of quantum measurements having arbitrary number of outcomes acting on an arbitrary but finite-dimensional state. We derive upper bounds (or exact values in specific cases) of the average success probability of RACs assisted by the best classical strategy or, equivalently, the best quantum strategy involving compatible measurements by the receiver. It follows that given any set of quantum measurements, if the average success probability of the RAC involving the given measurements by the receiver exceeds the above bound, then we can certify that the given measurements are incompatible. Here, it should be noted that RACs, being one of the fundamental quantum communication protocols, have been implemented in a series of experiments [30–35]. Hence, the results presented in this work reveal the unrecognized fact that measurement incompatibility has been witnessed in these experiments. Further, we identify all sets of three incompatible rank-one projective qubit measurements that can be witnessed by RACs.

Finally, we present the generic relationship between different types of input-output statistics—classical, quantum with shared randomness, quantum without shared randomness, quantum restricted to compatible measurements with shared randomness, and quantum restricted to compatible measurements without any shared randomness.

We arrange the rest of this paper as follows. In the next Sec. II, we introduce the formal definitions of incompatible quantum measurements. In Sec. III, we show that incompatible measurements by the receiver are necessary for quantum advantage in any communication task. Next, considering RACs, we present an operational witness of measurement incompatibility of an arbitrary set of quantum measurements in Sec. IV. Relationships between different types of probability distributions in communication tasks are depicted in Sec. V. Finally, we conclude with a short discussion in Sec. VI.

II. QUANTUM MEASUREMENT INCOMPATIBILITY

An arbitrary measurement is conceptualized by some positive operator-valued measure (POVM) defined as $E_y \equiv \{M_{b_y|y}\}_{b_y}$, with $M_{b_y|y} \geqslant 0$ for all b_y and $\sum_{b_y} M_{b_y|y} = \mathbb{1}$. Here, y corresponds to the choice of measurement and b_y denotes the outcomes of measurement y.

A set of measurements, $\{E_y\}_y$, with $y \in [n]$ (here we use the notation $[k] := \{1, \ldots, k\}$), is compatible [1] if there exists a parent POVM $\{G_\kappa : G_\kappa \geqslant 0 \,\forall \kappa, \sum_\kappa G_\kappa = \mathbb{1}\}$ and classical

postprocessing for each y given by $\{P_y(b_y|\kappa)\}$ such that

$$\forall b_y, y, \quad M_{b_y|y} = \sum_\kappa P_y(b_y|\kappa)G_\kappa. \quad (1)$$

Postprocessing for each y is defined by $\{P_y(b_y|\kappa)\}$ such that

$$P_y(b_y|\kappa) \geqslant 0 \,\forall y, b_y, \kappa; \quad \sum_{b_y} P_y(b_y|\kappa) = 1 \,\forall y, \kappa. \quad (2)$$

We note that the above definition of compatibility is equivalent to the existence of a parent POVM $\{G_\kappa\}$ whose appropriate marginals give rise to all the individual measurement effects $\{M_{b_y|y}\}_{b_y,y}$ [1].

III. INCOMPATIBILITY IS NECESSARY FOR QUANTUM ADVANTAGE IN COMMUNICATION TASKS

Now, we will show that incompatible measurements are necessary for showing quantum advantage in any communication task. Before proceeding, let us briefly describe a generic communication scenario consisting of two players—Alice and Bob. Alice and Bob are given inputs $x \in [l]$ and $y \in [n]$, respectively. Further, initially neither player has any idea about the other player's input. Alice, upon receiving the input x, sends a d-dimensional classical or quantum system to Bob, where d is finite. Bob, upon receiving the input y and the message (which is d-dimensional classical or quantum system) sent by Alice, outputs $b_y \in [d_y]$, wherein d_y is the number of outcomes for input y. The figure of merit of this communication task is determined by the set of probability distributions, $\{p(b_y|x, y)\}$.

In classical communication, they can use unlimited preshared randomness λ with some probability distribution $\pi(\lambda)$, and, therefore, any typical probability can be expressed as

$$p(b_y|x, y) = \sum_{m=1}^d \int_\lambda \pi(\lambda)p_a(m|x, \lambda)p_b(b_y|y, m, \lambda) \, d\lambda. \quad (3)$$

Here, $\{p_a(m|x, \lambda)\}, \{p_b(b_y|y, m, \lambda)\}$ are encoding and decoding functions by Alice and Bob, satisfying non-negativity and

$$\sum_m p_a(m|x, \lambda) = \sum_{b_y} p_b(b_y|y, m, \lambda) = 1, \quad (4)$$

while in quantum communication, one can consider the set of probabilities with or without the preshared classical randomness. Suppose $\mathcal{B}(\mathbb{C}^d)$ stands for the space of all operators acting on d-dimensional complex Hilbert space. In the former scenario with preshared randomness,

$$p(b_y|x, y) = \int_\lambda \pi(\lambda)\mathrm{Tr}(\rho_{x,\lambda}M_{b_y|y,\lambda})d\lambda,$$
$$\rho_{x,\lambda}, \, M_{b_y|y,\lambda} \in \mathcal{B}(\mathbb{C}^d), \quad (5)$$

where $\{\rho_{x,\lambda}\}$ is the quantum state sent by Alice upon input x and random variable λ, and $\{M_{b_y|y,\lambda}\}$ is the measurement executed by Bob for his input y and random variable λ. Without the preshared randomness, the expression of the probabilities reduces to

$$p(b_y|x, y) = \mathrm{Tr}(\rho_x M_{b_y|y}), \quad \rho_x, M_{b_y|y} \in \mathcal{B}(\mathbb{C}^d). \quad (6)$$

A communication scenario is specified by a set of natural numbers l, n, and $\vec{d} = (d_1, \ldots, d_n)$ such that $x \in [l]$, $y \in [n]$, $b_y \in [d_y]$. Given a scenario, we define the set of all probabilities obtainable by *d-dimensional classical communication*,

$$\mathcal{C}_d := \{p(b_y|x, y)\}, \tag{7}$$

where $p(b_y|x, y)$ is given by Eq. (3); the set of all probabilities in *d-dimensional quantum communication*,

$$\mathcal{Q}_d := \{p(b_y|x, y)\}, \tag{8}$$

where $p(b_y|x, y)$ is given by Eq. (5); and the set of all probabilities in *d-dimensional quantum communication without randomness*,

$$\overline{\mathcal{Q}}_d := \{p(b_y|x, y)\}, \tag{9}$$

where $p(b_y|x, y)$ is given by Eq. (6). In this work, we are interested in another two sets of probabilities. First, the set of all probabilities in *d-dimensional quantum communication restricted to compatible measurements*,

$$\mathcal{Q}_d^C := \{p(b_y|x, y)\}, \tag{10}$$

where $p(b_y|x, y)$ is given by Eq. (5) such that the set of measurements acting on d-dimensional quantum states used by Bob, $\{M_{b_y|y,\lambda}\}$, is compatible, that is, there exists parent POVM $\{G_\kappa\}$ such that

$$\forall b_y, y, \lambda, \quad M_{b_y|y,\lambda} = \sum_\kappa P_{y,\lambda}(b_y|\kappa) \, G_\kappa \tag{11}$$

and

$$P_{y,\lambda}(b_y|\kappa) \geqslant 0 \; \forall y, \lambda, b_y, \kappa; \; \sum_{b_y} P_{y,\lambda}(b_y|\kappa) = 1 \; \forall y, \lambda, \kappa. \tag{12}$$

And second, the set of all probabilities in *d-dimensional quantum communication restricted to compatible measurements without shared randomness*,

$$\overline{\mathcal{Q}}_d^C := \{p(b_y|x, y)\}, \tag{13}$$

where $p(b_y|x, y)$ is given by Eq. (6), such that the set of measurements, $\{M_{b_y|y}\}$, is compatible according to (1).

Result 1. Given any scenario specified by $(l, n, \vec{d})$,

$$\overline{\mathcal{Q}}_d^C \subseteq \mathcal{Q}_d^C = \mathcal{C}_d. \tag{14}$$

Thus, measurement incompatibility is necessary for any quantum advantage (with or without shared randomness) over classical communication.

Proof. The first relation, $\overline{\mathcal{Q}}_d^C \subseteq \mathcal{Q}_d^C$, follows from the definition of these two sets. The nontrivial part of this result is proving the equality.

Consider the case where Bob performs POVM measurements $\{G_\kappa\}$, which is the parent POVM of the measurement set $\{M_{b_y|y,\lambda}\}$. The Frenkel-Weiner theorem [36] implies that the set of input-output probabilities $p(\kappa|x)$ with a single quantum measurement on d-dimensional quantum states can always be reproduced by a suitable classical d-dimensional communication in the presence of shared randomness. In other words, $\forall \rho_{x,\lambda}$ there exists classical strategy

$\tilde{\pi}(\tilde{\lambda})$, $p_a(m|x, \lambda, \tilde{\lambda})$, $p_b(\kappa|m, \tilde{\lambda})$ such that

$$\text{Tr}(\rho_{x,\lambda} G_\kappa) = \sum_{m=1}^d \int_{\tilde{\lambda}} \tilde{\pi}(\tilde{\lambda}) p_a(m|x, \lambda, \tilde{\lambda}) p_b(\kappa|m, \tilde{\lambda}) \, d\tilde{\lambda}. \tag{15}$$

Here, note that in the scenario considered by Frenkel-Weiner [36], Bob does not receive any input y therein. That is why Bob's output κ depends only on the message m sent by Alice and classical shared randomness $\tilde{\lambda}$.

Let us now focus on the set of probabilities $\mathcal{C}_d$ wherein $(\lambda, \tilde{\lambda})$ is the preshared randomness. Take into account the following decoding function:

$$p_b(b_y|y, m, \lambda, \tilde{\lambda}) = \sum_\kappa P_{y,\lambda}(b_y|\kappa) p_b(\kappa|m, \tilde{\lambda}), \tag{16}$$

where $\{P_{y,\lambda}(b_y|\kappa)\}$ is the postprocessing defined in (11) and $p_b(\kappa|m, \tilde{\lambda})$ is given in (15). One can check that this is indeed a valid decoding function.

Next, we show that an arbitrary $p(b_y|x, y) \in \mathcal{Q}_d^C$ can always be reproduced by a suitable classical strategy (3) involving preshared randomness $(\lambda, \tilde{\lambda})$ and the decoding function (16). With the help of (11), (15), and (16) in a subsequent order, we find

$$\int_\lambda \pi(\lambda) \text{Tr}(\rho_{x,\lambda} M_{b_y|y,\lambda}) \, d\lambda$$

$$= \sum_\kappa \int_\lambda P_{y,\lambda}(b_y|\kappa) \pi(\lambda) \text{Tr}(\rho_{x,\lambda} G_\kappa) \, d\lambda$$

$$= \sum_m \int_\lambda \int_{\tilde{\lambda}} \pi(\lambda) \tilde{\pi}(\tilde{\lambda}) p_a(m|x, \lambda, \tilde{\lambda})$$

$$\times \left[\sum_\kappa P_{y,\lambda}(b_y|\kappa) p_b(\kappa|m, \tilde{\lambda}) \right] d\lambda \, d\tilde{\lambda}$$

$$= \sum_m \iint_{\lambda,\tilde{\lambda}} \pi(\lambda) \tilde{\pi}(\tilde{\lambda}) p_a(m|x, \lambda, \tilde{\lambda}) p_b(b_y|y, m, \lambda, \tilde{\lambda}) d\lambda d\tilde{\lambda}. \tag{17}$$

Therefore, an arbitrary probability distribution $p(b_y|x, y)$ obtainable from a compatible set of measurements can be reproduced by a suitable classical strategy, implying that $\mathcal{Q}_d^C \subseteq \mathcal{C}_d$. On the other hand, any classical strategy is always realized by quantum strategy with compatible measurements, i.e., $\mathcal{C}_d \subseteq \mathcal{Q}_d^C$. Therefore, these two sets are identical.

Finally, we remark that the figure of merit of any communication task is some arbitrary function of the probabilities $p(b_y|x, y)$; we can infer that any quantum advantage (with or without shared randomness) in such tasks over classical communication can be attained only if the set of measurements is incompatible. This completes the proof. ∎

Note that a weaker version of Result 1, $\overline{\mathcal{Q}}_d^C \subseteq \mathcal{C}_d$, has been proposed in Ref. [37]. However, from the arguments presented therein, it is not clear how their Eq. (4) follows, which raises questions upon the status of the last paragraph of the proof of their Theorem 1 presented in Appendix A of [37].

One profound implication of Result 1 is that any communication task can serve as a witness of measurement incompatibility. However, in general, *measurement incompatibility is not sufficient for quantum advantage without*

preshared randomness [38]—there exist incompatible qubit measurements such that the set of probabilities given by (6) for arbitrary quantum states is within $\mathcal{C}_2$ (see Sec. IV A of [38]).

Another useful implication of Result 1 appears when we are interested in linear functions of $\{p(b_y|x, y)\}$. Communication tasks for which the figures of merit are linear functions of $\{p(b_y|x, y)\}$ are widespread due to their practical importance in quantum communication complexity tasks [39–41], quantum key distribution [42], quantum randomness generation [43,44], quantum random access codes [45,46], oblivious transfer [30,41], and many other applications.

Fact 1. Given any scenario specified by $(l, n, \vec{d})$, the maximum values of any linear function of $\{p(b_y|x, y)\}$ obtained within the three different sets $\mathcal{C}_d$, $\mathcal{Q}_d^C$, and $\overline{\mathcal{Q}}_d^C$ are the same.

Proof. To find the optimum value of any linear function of $\{p(b_y|x, y)\}$, it is sufficient to consider classical strategy without shared randomness (see Lemma 1 in Appendix B for a detailed explanation). Now, all probability distributions $\{p(b_y|x, y)\}$, which are obtained from classical strategy without shared randomness, can always be realized within $\overline{\mathcal{Q}}_d^C$. Upon receiving the input x, Alice sends the quantum state ρ_x such that ρ_x is diagonal in the computational basis. Bob, upon receiving the input y and the diagonal state ρ_x, performs a fixed measurement $\{G_\kappa\}$ in the computational basis followed by some postprocessing depending on y. This observation together with the relation $\overline{\mathcal{Q}}_d^C \subseteq \mathcal{Q}_d^C = \mathcal{C}_d$ implies the above fact. ∎

The generic relation among the sets $\overline{\mathcal{Q}}_d^C, \mathcal{Q}_d^C, \overline{\mathcal{Q}}_d, \mathcal{Q}_d, \mathcal{C}_d$ is further analyzed in Sec. V. Next, we will propose incompatibility witness for an arbitrary set of measurements using a family of communication tasks, namely, the general version of RACs [47].

IV. INCOMPATIBILITY WITNESS FOR ANY SET OF MEASUREMENTS OF ARBITRARY SETTING

Take the most general form of a set of measurements. There are n measurements, defined by $\{M_{b_y|y}\}$, where $y \in [n]$ each of which has different outcomes, say, measurement y has d_y outcomes, that is, $b_y \in [d_y]$. These measurements are acting on d-dimensional quantum states where d is finite (see Fig. 1). In order to witness the incompatibility of this set, we introduce the most general form of random access codes with Bob having this set of measurements. Alice gets a string of n dits $x = x_1 x_2 \cdots x_n$ randomly from the set of all possible strings in which $x_y \in [d_y]$ for all $y \in [n]$. While Alice communicates a d-dimensional classical or quantum system to encode the information about the obtained string, the task for Bob is to guess the yth dit when y is chosen randomly. The figure of merit is the average success probability defined by the following linear function:

$$S(n, \vec{d}, d) = \frac{1}{n \prod_y d_y} \sum_{x,y} p(b_y = x_y | x, y), \quad (18)$$

which is fully specified by n, $\vec{d} = (d_1, d_2, \ldots, d_n)$, and d. Since Eq. (18) is a linear function of $p(b_y|x, y)$, by Fact 1, the maximum value over $\mathcal{C}_d$, $\mathcal{Q}_d^C$, and $\overline{\mathcal{Q}}_d^C$ is the same and denoted
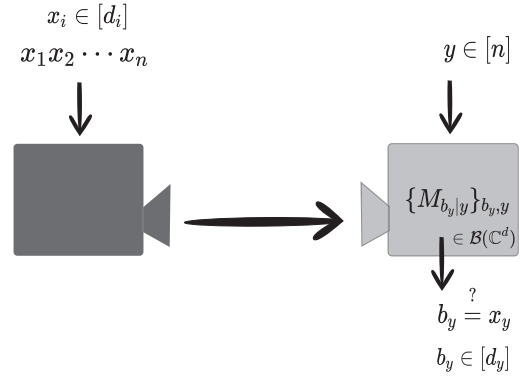

FIG. 1. An unknown measurement set of arbitrary settings, $\{M_{b_y|y}\}_{b_y,y}$, is provided; we only know the dimension (d) on which this set of measurements acts. Our task is to certify the incompatibility of this set of measurements. Here, the notation $[k] := \{1, \cdots, k\}$ for any natural number k.

by $S_c(n, \vec{d}, d)$. Precisely,

$$S_c(n, \vec{d}, d) = \max_{\{p(b_y|x,y)\} \in \mathcal{C}_d} S(n, \vec{d}, d)$$

$$= \max_{\{p(b_y|x,y)\} \in \mathcal{Q}_d^C} S(n, \vec{d}, d). \quad (19)$$

Hence, $S_c(n, \vec{d}, d)$ can be evaluated by maximizing the average success probability either over all classical strategies or over all quantum strategies involving compatible measurements only.

Whenever a set of measurements in the scenario specified by n, $\vec{d}$, d gives $S(n, \vec{d}, d) > S_c(n, \vec{d}, d)$ in the above-introduced general version of RACs, we can conclude that the measurements are incompatible. Hence, in order to witness measurement incompatibility, we need to know $S_c(n, \vec{d}, d)$. Now we present an upper bound on $S_c(n, \vec{d}, d)$ for arbitrary $n, \vec{d}, d$.

Result 2. The following relation holds true for arbitrary $(n, \vec{d}, d)$:

$$S_c(n, \vec{d}, d) \leqslant \frac{1}{n} \times \min \left\{ 1 + \sum_{\substack{i,j \\ i<j}} \frac{d}{d_i d_j} , \ n - 1 + \frac{d}{\prod_y d_y} \right\}.$$

$$(20)$$

This upper bound in Eq. (20) is obtained for $\mathcal{Q}_d^C$, that is, by taking the existence of a parent POVM of the measurements $\{M_{b_y|y}\}_{b_y,y}$ performed by Bob. The proof of this result is presented in Appendix A. When the outcome of all the measurements is the same, which is $d_y = \tilde{d}$ for all y, the above bound simplifies to

$$S_c(n, \tilde{d}, d) \leqslant \frac{1}{n} \times \min \left\{ 1 + \frac{n(n-1)d}{2\tilde{d}^2} , \ n - 1 + \frac{d}{\tilde{d}^n} \right\}.$$

$$(21)$$

Hence, in different types of RACs involving an arbitrary set of quantum measurements by Bob, if the average success probability exceeds the aforementioned upper bounds on S_c, then we can conclude that the measurements by Bob are incompatible.

On the other hand, whenever

$$d \leqslant \min_y d_y, \qquad (22)$$

we find out the exact value of $S_c(n, \vec{d}, d)$. Say, k_i is the number of sets among $[d_1], \cdots, [d_n]$ such that dit $i \in [d_y]$. For example, consider the RACs with $n = 4$ and $d_1 = 2$, $d_2 = 3$, $d_3 = 4$, and $d_4 = 3$. That is, Alice gets a string of four dits, $x = x_1 x_2 x_3 x_4$, randomly, where $x_1 \in [2]$, $x_2 \in [3]$, $x_3 \in [4]$, and $x_4 \in [3]$. In this case, $k_1 = 4$, $k_2 = 4$, $k_3 = 3$, and $k_4 = 1$. Also, we denote $d_{\max} = \max_y d_y$.

Result 3. If (22) holds, then

$$S_c(n, \vec{d}, d) = \frac{1}{n \prod_y d_y} \sum \left[\left(\prod_{j=1}^{d_{\max}} C_{n_j}^{\alpha_j} \right) \max_{i=1,\cdots,d} \{n_i\} \right], \quad (23)$$

with

$$\alpha_j = k_j - \sum_{i=j+1}^{d_{\max}} n_i, \quad C_{n_j}^{\alpha_j} = \frac{\alpha_j(\alpha_j - 1) \cdots (\alpha_j - n_j + 1)}{n_j(n_j - 1) \cdots 1},$$

and where the summation is taken over all possible integer solutions of the following equation:

$$\sum_{i=1}^{d_{\max}} n_i = n, \qquad (24)$$

such that $n_i \leqslant k_i$ for all i.

Note here that (23) is obtained for $\mathcal{C}_d$ by considering classical strategies. The detailed proof is given in Appendix B. For a particular case of Result 3 wherein $d_y = \tilde{d} = d$ for all y, the proof is previously given in [45]. Hence, when $d \leqslant \min_y d_y$, a necessary criteria for a set of measurements to be compatible is given by

$$S(n, \vec{d}, d) \leqslant S_c(n, \vec{d}, d), \qquad (25)$$

where $S_c(n, \vec{d}, d)$ is given by (23).

For $n = 2$, $d_y = \tilde{d}$ for all y, and $d \leqslant \tilde{d}$, the expression (23) simplifies to (for details, see Appendix C)

$$S_c(2, \tilde{d}, d) = \frac{1}{2\tilde{d}^2}(d + 2d\tilde{d} - d^2). \qquad (26)$$

And for $n = 3$, $d_y = \tilde{d}$ for all y, and $d \leqslant \tilde{d}$, the expression (23) simplifies to (for details, see Appendix C)

$$S_c(3, \tilde{d}, d) = \frac{d}{3\tilde{d}^3}[d^2 - 1 + 3\tilde{d}(\tilde{d} + 1 - d)]. \qquad (27)$$

The particular case of Result 2 for $n = 2$ can be found in [46] and, moreover, it is shown that any pair of rank-one projective measurements that is incompatible provides advantage in RACs [11]. In order to showcase the generic applicability of Results 2 and 3, we consider an arbitrary set of three rank-one projective qubit measurements, which using some unitary can be expressed as

$$M_{x_1|1} = (1/2)U[\mathbb{1} + (-1)^{x_1}\sigma_z]U^\dagger,$$

$$M_{x_2|2} = (1/2)U[\mathbb{1} + (-1)^{x_2}(\alpha\sigma_z + \sqrt{1 - \alpha^2}\sigma_x)]U^\dagger,$$

$$M_{x_3|3} = (1/2)U[\mathbb{1} + (-1)^{x_3}(\beta\sigma_z + \gamma\sqrt{1 - \beta^2}\sigma_x$$
$$\pm\sqrt{1 - \beta^2}\sqrt{1 - \gamma^2}\sigma_y)]U^\dagger, \qquad (28)$$

where $x_1, x_2, x_3 \in [2]$, the variables $\alpha, \beta, \gamma \in [-1, 1]$, and U can be an arbitrary unitary operator acting on $\mathbb{C}^2$. We obtain the following result:

Result 4. The figure of merit (18) of RACs for $n = 3$, $\tilde{d} = 2$, $d = 2$ can witness any set of three incompatible rank-one projective qubit measurements, except for the sets defined by (28) with

$$(\alpha, \beta, \gamma) = \{(\pm 1/2, \pm 1/2, -1), (\pm 1/2, \mp 1/2, 1)\}.$$

This result is proved with the help of numerical optimizations and the proof is available in Appendix D.

V. GENERIC RELATIONS BETWEEN PROBABILITY SETS

We now point out a few generic relations between the sets in order to get a generic perspective. It is trivial that $\overline{\mathcal{Q}}_d \not\subset \mathcal{C}_d$ since we observe the quantum advantage for $S(n, \vec{d}, d)$. We also observe the following:

Fact 2. In general, $\mathcal{C}_d \not\subset \overline{\mathcal{Q}}_d$ and thus $\overline{\mathcal{Q}}_d^C \subsetneq \mathcal{C}_d$. In other words, there exist probabilities that belong to $\mathcal{C}_d$ but do not belong to $\overline{\mathcal{Q}}_d$. Moreover, in general, $\mathcal{Q}_d \setminus (\mathcal{C}_d \cup \overline{\mathcal{Q}}_d) \neq \varnothing$. In other words, there exists a probability distribution that belongs to $\mathcal{Q}_d$ but does not belongs to $\mathcal{C}_d$ and $\overline{\mathcal{Q}}_d$.

Proof. Once again, reckon the RAC task and, instead of the average success probability (18), let us consider the figure of merit to be the worst-case success probability,

$$W(n, \tilde{d}, d) = \min_{x,y}\{p(b_y = x_y|x, y)\}. \qquad (29)$$

It follows from Yao's principle that the average success probability $S(n, 2, 2)$ is the same as $W(n, 2, 2)$ when preshared randomness is available. See Lemma 1 in [48] for the proof. Therefore, $W(4, 2, 2)$ in $\mathcal{C}_2$ is the same as $S_c(4, 2, 2) = 11/16$ [48]. However, it was proven that the best value of $W(4, 2, 2)$ in $\overline{\mathcal{Q}}_2$ is $1/2$ [49], implying $\mathcal{C}_d \not\subset \overline{\mathcal{Q}}_d$. Moreover, since $\overline{\mathcal{Q}}_d^C$ is a subset of both $\mathcal{C}_d$ and $\overline{\mathcal{Q}}_d$, it must be a proper subset.

Besides, there exists a quantum strategy in $\mathcal{Q}_2$ that achieves $W(4, 2, 2) = 0.74$ (Sec. 4.1.2 in [48]). This means that both $\mathcal{C}_d$ and $\overline{\mathcal{Q}}_d$ are proper subsets of $\mathcal{Q}_d$. ∎

Next, we present another observation that helps to completely understand the relationship between different sets of probability distributions.

Fact 3. In general, $(\mathcal{C}_d \cap \overline{\mathcal{Q}}_d) \setminus \overline{\mathcal{Q}}_d^C \neq \varnothing$. In other words, there exists a set of probabilities that does not belong to $\overline{\mathcal{Q}}_d^C$ at the same time inside both of the sets $\mathcal{C}_d$ and $\overline{\mathcal{Q}}_d$.

Proof. In $\mathcal{C}_d$, $W(3, 2, 2) = S_c(3, 2, 2) = 3/4$ [48]. It is also well known that there exist two-dimensional quantum states and measurements that lead to $W(3, 2, 2) = 1/2 + 1/(2\sqrt{3}) \approx 0.79$, without preshared randomness. An example of such quantum states is given in Example 2 of Ref. [49]. If we consider the noisy version of those quantum states $\nu\rho + (1 - \nu)\mathbb{1}/2$ such that $\nu = \sqrt{3}/2$, then it is easy to see that the value of $W(3, 2, 2)$ reduces to $3/4$. Therefore, $W(3, 2, 2) = 3/4$ is obtained in both of the sets $\mathcal{C}_2$ and $\overline{\mathcal{Q}}_2$. On the other hand, a numerical analysis using semidefinite programming shows that $W(3, 2, 2)$ is, at most, $2/3$ in $\overline{\mathcal{Q}}_2^C$, which implies $(\mathcal{C}_2 \cap \overline{\mathcal{Q}}_2) \setminus \overline{\mathcal{Q}}_2^C \neq \varnothing$. ∎
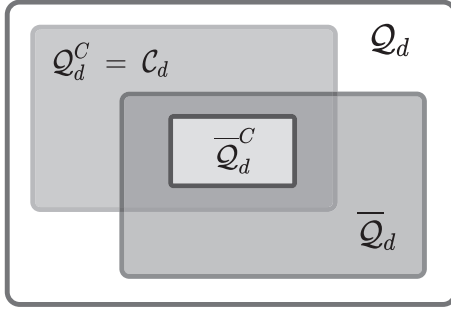
FIG. 2. Generic relation between different sets of probabilities. The equivalence between $\mathcal{Q}_d^C$ and $\mathcal{C}_d$ is shown in Result 1. There exist many examples such that $\overline{\mathcal{Q}}_d$ is not a subset of $\mathcal{C}_d$. On the other hand, Fact 2 points out that $\mathcal{C}_d$ is not a subset of $\overline{\mathcal{Q}}_d$ and both $\mathcal{C}_d$ as well as $\overline{\mathcal{Q}}_d$ are proper subsets of $\mathcal{Q}_d$. While it is immediate that $\overline{\mathcal{Q}}_d^C$ is a subset of both $\mathcal{C}_d$ and $\overline{\mathcal{Q}}_d$, Fact 3 shows that $\overline{\mathcal{Q}}_d^C$ is not the intersection of these two sets.

Based on the above observations, the generic relationship between different sets of probabilities can be understood as depicted in Fig. 2.

VI. CONCLUDING REMARKS

By characterizing the set of quantum correlations in prepare-and-measure scenarios produced from any set of compatible measurements, we have shown in this article that incompatible measurements at the receiver's end are necessary to demonstrate a quantum advantage in any communication task. Based on this result, we have presented a semi-device-independent witness of measurement incompatibility invoking generalized random access codes. Further, we have completely characterized the sets of three incompatible projective qubit measurements that can be detected using our proposed witness. The relationship between the classical probability distributions and different types of quantum probability distributions produced in an arbitrary communication task has also been presented.

It might be noted that some of the results derived in [11,46] appear as natural corollaries of the results obtained here. It has recently been shown that the measurement statistics produced in any communication task involving compatible measurements by the receiver can be reproduced by classical communication, where the dimension of the classical communication is of the order of the number of outcomes of the parent POVM [29]. Significantly, our results much further reduce the dimension of the classical communication involved.

The importance of the analysis presented here lies in the fact that the classical bound of the success metric of any communication task becomes an upper bound on the metric of the task under a compatible set of measurements. Consequently, violating the classical bound of any communication task can be used as a sufficient criteria to witness measurement incompatibility. These bounds are tight whenever the dimension on which the measurements act is not larger than the number of outcomes of any of the measurements. The present study establishes that measurement incompatibility is the fundamental

quantum resource for nonclassicality in any communication task or, more generally, in prepare-and-measure scenarios.

Our analysis paves the way for investigations of several open questions. First, deriving more efficient incompatibility witnesses based on different communication tasks is worthwhile for future studies. Second, our results may be generalized to propose semi-device-independent witnesses for incompatible quantum channels [50] and quantum instruments [51,52]. Finally, proposing operational witnesses for all incompatible extremal POVMs [53] is another fundamentally motivated open problem.

APPENDIX A: PROOF OF RESULT 2

Proof. For a given set of measurements, $\{M_{b_y|y}\}$, on Bob's side, the maximum average success probability (18) in quantum theory is given by [11]

$$
\begin{aligned}
\max_{\{\rho_x\}} \quad & \frac{1}{n \prod_y d_y} \sum_{x_1 x_2 \cdots x_n} \sum_y \mathrm{Tr}(\rho_x M_{b_y = x_y|y}) \\
= & \frac{1}{n \prod_y d_y} \sum_{x_1 x_2 \cdots x_n} \max_{\{\rho_x\}} \mathrm{Tr}\left[\rho_x \left(\sum_y M_{b_y = x_y|y} \right) \right] \\
= & \frac{1}{n \prod_y d_y} \sum_{x_1 x_2 \cdots x_n} ||\chi||,
\end{aligned}
\tag{A1}
$$

where

$$
\chi = M_{x_1|1} + M_{x_2|2} + \cdots + M_{x_n|n}.
\tag{A2}
$$

Here, $||\chi||$ denotes the operator norm of χ, which is simply the maximum eigenvalue of the operator χ. Our aim is to obtain an upper bound on expression (A1) when $\{M_{b_y|y}\}$ are compatible. An alternative definition of measurement incompatibility, which is equivalent to the standard one (1), is associated with the existence of a parent POVM whose appropriate marginals give rise to all the individual measurements [1]. Precisely, if the measurements $\{M_{b_y|y}\}$ are compatible, then there exists a parent measurement, $G \equiv \{G(b_1, \cdots, b_n)\}$, with $\prod_{y=1}^n d_y$ elements from which all the measurement operators can be reconstructed by taking marginals as follows:

$$
M_{x_y|y} = \sum_{b_1, \cdots, b_{y-1}, b_{y+1}, \cdots, b_n} G(b_1, \cdots, b_{y_1}, x_y, b_{y+1}, \cdots, b_n),
\tag{A3}
$$

where

$$\sum_{b_1,\cdots,b_n} G(b_1, \cdots, b_{y_1}, b_y, b_{y+1}, \cdots, b_n) = \mathbb{1}_{d \times d}. \tag{A4}$$

Let us first expand χ in terms of the parent POVM using (A3),

$$\chi = \sum_{b_2,b_3,\cdots,b_n} G(x_1, b_2, b_3 \cdots, b_n) + \sum_{b_1,b_3,\cdots,b_n} G(b_1, x_2, b_3, \cdots, b_n) + \cdots + \sum_{b_1,b_2,b_3,\cdots,b_{n-1}} G(b_1, b_2, b_3, \cdots, b_{n-1}, x_n). \tag{A5}$$

Each term in the above expansion can be split into two terms in the following way:

$$\chi = \sum_{b_3,b_4,\cdots,b_n} G(x_1, x_2, b_3, b_4, \cdots, b_n) + \sum_{\substack{b_2,\cdots,b_n \\ b_2 \neq x_2}} G(x_1, b_2, \cdots, b_n) + \sum_{b_1,b_4,\cdots,b_n} G(b_1, x_2, x_3, b_4, \cdots, b_n)$$

$$+ \sum_{\substack{b_1,b_3,\cdots,b_n \\ b_3 \neq x_3}} G(b_1, x_2, b_3, \cdots, b_n) + \cdots + \sum_{b_2,b_3,\cdots,b_{n-1}} G(x_1, b_2, b_3, \cdots, b_{n-1}, x_n) + \sum_{\substack{b_1,\cdots,b_{n-1} \\ b_1 \neq x_1}} G(b_1, \cdots, b_{n-1}, x_n). \tag{A6}$$

In the above Eq. (A6), there are two sums in each line and there is a total of n lines. Let us denote the first sum and the second sum in the ith line by $\mathcal{S}_1^i$ and $\mathcal{S}_2^i$, respectively, where $i \in \{1, \cdots, n\}$. Hence, Eq. (A6) can be expressed as

$$\chi = \sum_{i=1}^{n} \left(\mathcal{S}_1^i + \mathcal{S}_2^i \right), \tag{A7}$$

where

$$\mathcal{S}_1^i = \sum_{b_1,\cdots,b_{i-1},b_{i+2},\cdots,b_n} G(b_1, \cdots, b_{i-1}, x_i, x_{i+1}, b_{i+2}, \cdots, b_n) \tag{A8}$$

and

$$\mathcal{S}_2^i = \sum_{\substack{b_1,\cdots,b_{i-1},b_{i+1},\cdots,b_n \\ b_{i+1} \neq x_{i+1}}} G(b_1, \cdots, b_{i-1}, x_i, b_{i+1}, \cdots, b_n). \tag{A9}$$

Here the index i is taken to be modulo n. Each $G(\cdots)$ in the above sums will be termed as an element.

Let us now make an observation that there is no common element between $\mathcal{S}_2^i$ and $\mathcal{S}_2^{i+1}$. The common element between $\mathcal{S}_2^i$ and $\mathcal{S}_2^j$ with $i, j \in \{1, \cdots, n\}$ and $j > i + 1$ is

$$\sum_{\substack{b_1,\cdots,b_n \\ b_{i+1} \neq x_{i+1} \\ b_{j+1} \neq x_{j+1}}} G(b_1, \cdots, b_{i-1}, x_i, b_{i+1}, \cdots, b_{j-1}, x_j, b_{j+1}, \cdots, b_n) \leqslant \sum_{\substack{k=1 \\ k \neq i,j}}^{n} \sum_{b_k} G(b_1, \cdots, b_{i-1}, x_i, b_{i+1}, \cdots, b_{j-1}, x_j, b_{j+1}, \cdots, b_n), \tag{A10}$$

where the index i, j is taken to be modulo n. Hence, we have

$$\sum_{i=1}^{n} \mathcal{S}_2^i \leqslant \sum_{\substack{i,j \in \{1,\cdots,n\} \\ j > i+1}} \left[\sum_{\substack{k=1 \\ k \neq i,j}}^{n} \sum_{b_k} G(b_1, \cdots, b_{i-1}, x_i, b_{i+1}, \cdots, b_{j-1}, x_j, b_{j+1}, \cdots, b_n) \right] + \text{other terms with no common element.} \tag{A11}$$

Next, let us focus on $\mathcal{S}_1^i$. It can be checked that

$$\sum_{i=1}^{n} \mathcal{S}_1^i = \sum_{i=1}^{n} \left[\sum_{\substack{k=1 \\ k \neq i,i+1}}^{n} \sum_{b_k} G(b_1, \cdots, b_{i-1}, x_i, x_{i+1}, b_{i+2}, \cdots, b_n) \right]. \tag{A12}$$

Replacing $\mathcal{S}_2^i$ and $\mathcal{S}_1^i$ using (A11) and (A12) in (A7), we have

$$\chi \leqslant \sum_{\substack{i,j \in \{1,\cdots,n\} \\ i < j}} \left[\sum_{\substack{k=1 \\ k \neq i,j}}^{n} \sum_{b_k} G(b_1, \cdots, b_{i-1}, x_i, b_{i+1}, \cdots, b_{j-1}, x_j, b_{j+1}, \cdots, b_n) \right] + \text{other terms with no common element}$$

$$\leqslant \sum_{\substack{i,j\in\{1,\cdots,n\}\\i<j}} \left[\sum_{\substack{k=1\\k\neq i,j}}^{n} \sum_{b_k} G(b_1,\cdots,b_{i-1},x_i,b_{i+1},\cdots,b_{j-1},x_j,b_{j+1},\cdots,b_n) \right] + \sum_{b_1,\cdots,b_n} G(b_1,\cdots,b_n). \tag{A13}$$

Substituting the above expression into (A1) and employing the triangle inequality for the norm, we find that

$$S_c(n,\vec{d},d) \leqslant \frac{1}{n\prod_y d_y} \left[\sum_{x_1,\cdots,x_n} \sum_{\substack{i,j\in\{1,\cdots,n\}\\i<j}} \left\| \sum_{\substack{k=1\\k\neq i,j}}^{n} \sum_{b_k} G(b_1,\cdots,b_{i-1},x_i,b_{i+1},\cdots,b_{j-1},x_j,b_{j+1},\cdots,b_n) \right\| \right.$$
$$\left. + \sum_{x_1,\cdots,x_n} \left\| \sum_{b_1,\cdots,b_n} G(b_1,\cdots,b_n) \right\| \right]. \tag{A14}$$

Due to (A4), the second term of the above expression can be evaluated as

$$\sum_{x_1,\cdots,x_n} \left\| \sum_{b_1,\cdots,b_n} G(b_1,\cdots,b_n) \right\| = \sum_{x_1,\cdots,x_n} \left\| \mathbb{1}_{d\times d} \right\| = \prod_{y=1}^{n} d_y. \tag{A15}$$

Next, consider the first term in (A14), given by

$$\sum_{x_1,\cdots,x_n} \sum_{\substack{i,j\in\{1,\cdots,n\}\\i<j}} \left\| \sum_{\substack{k=1\\k\neq i,j}}^{n} \sum_{b_k} G(b_1,\cdots,b_{i-1},x_i,b_{i+1},\cdots,b_{j-1},x_j,b_{j+1},\cdots,b_n) \right\| = \sum_{\substack{i,j\in\{1,\cdots,n\}\\i<j}} \beta_{i,j}, \tag{A16}$$

where

$$\beta_{i,j} = \sum_{x_1,\cdots,x_n} \left\| \sum_{\substack{k=1\\k\neq i,j}}^{n} \sum_{b_k} G(b_1,\cdots,b_{i-1},x_i,b_{i+1},\cdots,b_{j-1},x_j,b_{j+1},\cdots,b_n) \right\|$$
$$\leqslant \sum_{x_1,\cdots,x_n} \mathrm{Tr}\left[\sum_{\substack{k=1\\k\neq i,j}}^{n} \sum_{b_k} G(b_1,\cdots,b_{i-1},x_i,b_{i+1},\cdots,b_{j-1},x_j,b_{j+1},\cdots,b_n) \right]$$
$$= \sum_{\substack{r=1\\r\neq i,j}}^{n} \sum_{x_r} \mathrm{Tr}\left[\sum_{x_i,x_j} \sum_{\substack{k=1\\k\neq i,j}}^{n} \sum_{b_k} G(b_1,\cdots,b_{i-1},x_i,b_{i+1},\cdots,b_{j-1},x_j,b_{j+1},\cdots,b_n) \right]$$
$$= \sum_{\substack{r=1\\r\neq i,j}}^{n} \sum_{x_r} \mathrm{Tr}\left[\sum_{b_1,\cdots,b_n} G(b_1,\cdots,b_n) \right]$$
$$= \sum_{\substack{r=1\\r\neq i,j}}^{n} \sum_{x_r} \mathrm{Tr}(\mathbb{1}_{d\times d})$$
$$= d \prod_{\substack{y=1\\y\neq i,j}}^{n} d_y. \tag{A17}$$

Hence, we get the first term in (A14), given by

$$\sum_{x_1,\cdots,x_n} \sum_{\substack{i,j\in\{1,\cdots,n\}\\i<j}} \left\| \sum_{\substack{k=1\\k\neq i,j}}^{n} \sum_{b_k} G(b_1,\cdots,b_{i-1},x_i,b_{i+1},\cdots,b_{j-1},x_j,b_{j+1},\cdots,b_n) \right\| = d \sum_{\substack{i,j\in\{1,\cdots,n\}\\i<j}} \prod_{\substack{y=1\\y\neq i,j}}^{n} d_y. \tag{A18}$$

By substituting the bounds from (A15)–(A18) into (A14), we obtain

$$S_c(n, \vec{d}, d) \leqslant \frac{1}{n \prod_y d_y} \left[d \sum_{\substack{i,j \in \{1,\cdots,n\} \\ i<j}} \left(\prod_{\substack{y=1 \\ y \neq i,j}}^{n} d_y \right) + \prod_{y=1}^{n} d_y \right], \tag{A19}$$

which reduces to the first expression of Eq. (20).

For the other bound, let us use the fact that in (A5), only the term $G(x_1, x_2, \cdots, x_n)$ occurs n times and all the other terms can occur, at most, $(n-1)$ times to get an upper bound on χ as follows:

$$\chi \leqslant G(x_1, x_2, \cdots, x_n) + (n-1) \sum_{b_1,\cdots,b_n} G(b_1, \cdots, b_n). \tag{A20}$$

Replacing this bound into (A4) and employing the triangle inequality for the norm, we get

$$S_c(n, \vec{d}, d) \leqslant \frac{1}{n \prod_y d_y} \left[\sum_{x_1,\cdots,x_n} \left\| G(x_1, \cdots, x_n) \right\| + (n-1) \sum_{x_1,\cdots,x_n} \left\| \sum_{b_1,\cdots,b_n} G(b_1, \cdots, b_n) \right\| \right]. \tag{A21}$$

We already have a bound given by (A15) on the second sum in the above equation. The first term is bounded by d since

$$\sum_{x_1,\cdots,x_n} \left\| G(x_1, \cdots, x_n) \right\| \leqslant \sum_{x_1,\cdots,x_n} \mathrm{Tr}[G(x_1, \cdots, x_n)]$$

$$= \mathrm{Tr}\left[\sum_{x_1,\cdots,x_n} G(x_1, \cdots, x_n) \right]$$

$$= \mathrm{Tr}(\mathbb{1}_{d \times d})$$

$$= d. \tag{A22}$$

Therefore, we arrive at

$$S_c(n, \vec{d}, d) \leqslant \frac{1}{n \prod_y d_y} \left[d + (n-1) \prod_y d_y \right], \tag{A23}$$

which reduces to the second expression of Eq. (20). This completes the proof. ∎

APPENDIX B: PROOF OF RESULT 3

In order to provide a detailed proof of Result 3, we first state a general feature of communication tasks.

Lemma 1. Consider a general form of a linear function of $\{p(b_y|x, y)\}$,

$$S = \sum_{x,y,b_y} c_{x,y,b_y} p(b_y|x, y). \tag{B1}$$

The maximum value of S within $\mathcal{C}_d$, which we denote by S_c, is obtained by deterministic strategies and can be written only in terms of decoding function $\{p_b(b_y|y, m)\}$.

Proof. Replacing the expression of $p(b_y|x, y)$ for classical communication given by Eq. (3) into (B1), we see that

$$S_c = \max_{\substack{\{p_a(m|x,\lambda)\} \\ \{p_b(b_y|y,m,\lambda)\} \\ \{\pi(\lambda)\}}} \sum_x \left[\sum_{y,b_y} c_{x,y,b_y} p(b_y|x, y) \right]$$

$$= \max_{\substack{\{p_a(m|x,\lambda)\} \\ \{p_b(b_y|y,m,\lambda)\} \\ \{\pi(\lambda)\}}} \int_\lambda \pi(\lambda) \left\{ \sum_x \left[\sum_m p_a(m|x, \lambda) \left(\sum_{y,b_y} c_{x,y,b_y} p_b(b_y|y, m, \lambda) \right) \right] \right\} d\lambda$$

$$= \max_{\substack{\{p_b(b_y|y,m,\lambda)\} \\ \{\pi(\lambda)\}}} \int_\lambda \pi(\lambda) \left[\sum_x \max_m \left(\sum_{y,b_y} c_{x,y,b_y} p_b(b_y|y, m, \lambda) \right) \right] d\lambda. \tag{B2}$$

This is achieved when $p_a[m^*(\lambda, x)|x, \lambda] = 1$, where, for each λ, $m^*(\lambda, x)$ is defined as follows:

$$\sum_{y,b_y} c_{x,y,b_y} p_b[b_y|y, m^*(\lambda, x), \lambda]$$

$$\geqslant \sum_{y,b_y} c_{x,y,b_y} p_b(b_y|y, m, \lambda) \ \forall \ m \in [d]. \tag{B3}$$

Now, the above expression (B2) is a convex sum with respect to $\pi(\lambda)$ and thus we can omit the dependence of λ by taking the best value of $\{\sum_x \max_m[\sum_{y,b_y} c_{x,y,b_y} p_b(b_y|y, m, \lambda)]\}$ over different choices of λ as follows:

$$S_c = \max_{\{p_b(b_y|y,m)\}} \left[\sum_x \max_m \left(\sum_{y,b_y} c_{x,y,b_y} p_b(b_y|y, m) \right) \right]. \tag{B4}$$

Therefore, it is sufficient to consider deterministic decoding, that is, $p_b(b_y|y, m) \in \{0, 1\}$, to achieve S_c. Moreover, given any decoding strategy $\{p_b(b_y|y, m)\}$, the best encoding function is

$$p_a(m^*|x) = 1 \quad \text{where} \quad \sum_{y,b_y} c_{x,y,b_y} p_b(b_y|y, m^*)$$

$$\geqslant \sum_{y,b_y} c_{x,y,b_y} p_b(b_y|y, m) \ \forall \ m \in [d]. \quad \text{(B5)}$$

This completes the proof. ∎

Proof of Result 3. The proof is essentially a generalization of the proof given in Sec. II A of [45], which was restricted for the particular case where $d = d_y$ for all y. We know from the above lemma that the optimal encoding and decoding functions are deterministic. Thus, this can be written in a functional form as

$$E(x_1 \cdots x_n) = m \ \text{ if } p_a(m|x) = 1 \quad \text{(B6)}$$

and

$$D_y(m) = b_y \ \text{ if } \ p_b(b_y|y, m) = 1. \quad \text{(B7)}$$

Here, $E(x_1 \cdots x_n)$ is a function whose domain is the set of inputs, $x = x_1 \cdots x_n$, and range is the set of messages, $[d]$. Also, $D_y(m)$ is a function whose domain is the set of messages, $[d]$, and range is the set $[d_y]$. We say the decoding strategy is "identity decoding," denoted by $\{\tilde{D}_y\}$, if

$$\forall y, \ \tilde{D}_y(m) = m. \quad \text{(B8)}$$

We want to show that without loss of generality, we can take $\{\tilde{D}_y\}$ for the maximum success probability. Consider an encoding $E(x)$ (B6) and a decoding $\{D_y\}$ (B7) that may not be $\{\tilde{D}_y\}$, that is, there may exist y such that $D_y(m) \neq m$. Let $D_y^{\leftarrow}(b_y)$ be the pre-image of b_y, that is, $D_y^{\leftarrow}(b_y) = \{m \in [d] : D_y(m) = b_y\}$.

Subsequently, we consider the following quantity:

$$D_1^{\leftarrow}(b_1) \cdots D_n^{\leftarrow}(b_n)$$
$$= \{m_1 \cdots m_n : D_1(m_1) = b_1, \cdots, D_n(m_n) = b_n\}, \quad \text{(B9)}$$

which is simply the set of dit-string $\{m_1 \cdots m_n\}$ that is mapped to the string $b_1 \cdots b_n$. We define another encoding function $\{\tilde{E}_x\}$ as follows:

$$\tilde{E}[D_1^{\leftarrow}(x_1)D_2^{\leftarrow}(x_2) \cdots D_n^{\leftarrow}(x_n)] = m \ \text{ if } E(x_1 \cdots x_n) = m. \quad \text{(B10)}$$

The above definition of $\tilde{E}$ is not complete since it is not defined if $x_i \notin [d]$ since $D_y^{\leftarrow}(x_i) \in [d]$. In those cases, we take any encoding strategy. Now, we note that $\tilde{E}$ is a well-defined encoding function. Also note that $\tilde{E}$ is a valid encoding for the random access codes considered by us only if $d \leqslant \min_y d_y$. That is because if $d > d_y$ for some y, then the domain of $\tilde{E}$ may have a string of n dits that does not belong to x.

Suppose, for any input pair $x_1 \cdots x_n$, y so that the encoding E and decoding $\{D_y\}$ guess the correct dit x_y. Hence, if the encoding strategy is given by $E(x = x_1 \cdots x_n) = m$, then the decoding strategy is given by $D_y(m) = b_y = x_y$. Therefore, we have $D_y^{\leftarrow}(x_y) = m$. As a consequence, the new encoding $\tilde{E}_x$ and the "identify decoding" $\{\tilde{D}_y\}$ also provide the correct answer for at least one input pair from

$\{D_1^{\leftarrow}(x_1)D_2^{\leftarrow}(x_2) \cdots D_n^{\leftarrow}(x_n)\}$, y. Hence, the average success probability for the strategy consisting of the encoding $\tilde{E}_x$ and the identify decoding $\{\tilde{D}_y\}$ is greater than or equal to that for the strategy with encoding E and decoding $\{D_y\}$. Therefore, we can consider identity decoding without loss of generality.

Next, from Eq. (B4), the expression for S_c pertaining to the random access codes for identity decoding can be written as

$$S_c = \frac{1}{n \prod_y d_y} \sum_x \max_m \left[\sum_y P(b_y = x_y|y, m) \right]$$

$$= \frac{1}{n \prod_y d_y} \sum_x \max_m \left(\sum_y \delta_{x_y, m} \right), \quad \text{(B11)}$$

and for the identity decoding, the best encoding can be determined from (B5) as follows:

$$p_a(m^*|x) = 1 \quad \text{where} \quad \sum_y \delta_{x_y, m^*} \geqslant \sum_y \delta_{x_y, m} \ \forall \ m \in [d]. \quad \text{(B12)}$$

Hence, the best encoding pertaining to the identity decoding is just sending the dit that belongs to $[d]$ and occurs maximum times in the input string $x_1 \cdots x_n$.

Finally, we provide an expression for S_c for the best classical strategy derived above. In an input string $x_1 \cdots x_n$, say, the dit i occurs n_i number of times and the maximum value of a dit can be $\max_y d_y$. Alice sends message m such that $n_m = \max_{i=1,\cdots,d} n_i$. As a result, out of n different values of y, they get success $(\max_{i=1,\cdots,d} n_i)$ times. As the total number of dits is n, the set of values of n_i should satisfy

$$\sum_{i=1}^{d_{\max}} n_i = n, \quad \text{(B13)}$$

where $d_{\max} = \max_y d_y$. Moreover, dit i may not belong to all $[d_y]$ and thus n_i cannot take all the solutions of the above equation. Say, k_i is the number of sets among $[d_1], \cdots, [d_n]$ such that dit $i \in [d_y]$. Therefore, we are only interested in those solutions where $n_i \leqslant k_i$.

Given such a solution of $\{n_i\}$, there will be many possible numbers of input dit strings x having that $\{n_i\}$. Next, let us evaluate the number of input dit strings x that can have an arbitrary $\{n_i\}$. In any input string, at most $k_{d_{\max}}$ number of input dits can have the value $d_{\max}$. In the given set of input dit strings having $\{n_i\}$, the dit $d_{\max}$ occurs $n_{d_{\max}}$ number of times. Hence, the dit $d_{\max}$ can be arranged in $C_{n_{d_{\max}}}^{k_{d_{\max}}}$ different possible ways. Next, in any input string, at most $k_{d_{\max}-1}$ number of input dits can have the value $(d_{\max} - 1)$. However, among these $k_{d_{\max}-1}$ number of input dits, $n_{d_{\max}}$ number of dits have already taken the value $d_{\max}$ in the case of the given set of input strings. Also, in the given set of input dit strings having $\{n_i\}$, the dit $(d_{\max} - 1)$ occurs $n_{d_{\max}-1}$ number of times. Therefore, for any of the above-mentioned arrangements of the dit $d_{\max}$, the dit $(d_{\max} - 1)$ can be arranged in $C_{n_{d_{\max}-1}}^{k_{d_{\max}-1}-n_{d_{\max}}}$ different possible ways. Proceeding in this way, it can be shown that an arbitrary dit j can be arranged in $C_{n_j}^{\alpha_j}$ different possible ways with $\alpha_j = k_j - \sum_{i=j+1}^{d_{\max}} n_i$ for any arrangement of the dits, $d_{\max}$, $(d_{\max} - 1), \cdots, j + 1$. Therefore, given any $\{n_i\}$, there will

be $(\prod_{j=1}^{d_{\max}} C_{n_j}^{\alpha_j})$ (with $\alpha_j = k_j - \sum_{i=j+1}^{d_{\max}} n_i$) number of input dit strings having that $\{n_i\}$. Combining these facts, we obtain Eq. (23). ∎

APPENDIX C: DERIVATION OF EQ. (26) AND EQ. (27)

From Result 3, we can write the following for $n = 2$, $d_y = \tilde{d}$ for all y, and $d \leqslant \tilde{d}$:

$$S_c(2, \tilde{d}, d) = \frac{1}{2\tilde{d}^2} \sum_{\{n_i\} \in \mathcal{S}} \left[N_{\{n_i\}} \max_{i=1,\cdots,d} \{n_i\} \right], \quad \text{(C1)}$$

where $N_{\{n_i\}}$ is the number of input dit strings having a given $\{n_i\}$ and $\mathcal{S}$ denotes the set of $\{n_i\}$ satisfying

$$\sum_{i=1}^{\tilde{d}} n_i = 2, \quad \text{(C2)}$$

such that $n_i \leqslant 2$ for all i.

Next, let us characterize the set $\mathcal{S}$. It can be noted that there are the following two types of $\{n_i\} \in \mathcal{S}$:

(1) For each $i \in [\tilde{d}]$, $n_i = 2$ and $n_j = 0$ for all $j \neq i$ and $j \in [\tilde{d}]$.

There are $\tilde{d}$ number of such $\{n_i\} \in \mathcal{S}$. However, $\max_{i=1,\cdots,d}\{n_i\} = 0$ for each of those $\{n_i\} \in \mathcal{S}$ satisfying $n_i = 2$ for any i such that $i \in \{d+1, \cdots, \tilde{d}\}$ and $n_j = 0$ for all $j \in [\tilde{d}]$ and $j \neq i$. Hence, only d number of $\{n_i\} \in \mathcal{S}$ belonging to this class contribute to the sum of (C1). It is straightforward to check that for each of these d number of $\{n_i\} \in \mathcal{S}$, $N_{\{n_i\}} = 1$ and $\max_{i=1,\cdots,d}\{n_i\} = 2$.

(2) For each $i, j \in [\tilde{d}]$ with $i \neq j$, $n_i = n_j = 1$ and $n_k = 0$ for all $k \notin \{i, j\}$ with $k \in [\tilde{d}]$.

There are $C_2^{\tilde{d}}$ number of such $\{n_i\} \in \mathcal{S}$. However, $\max_{i=1,\cdots,d}\{n_i\} = 0$ for each of those $\{n_i\} \in \mathcal{S}$ satisfying $n_i = n_j = 1$ for any i, j with $i \neq j$, $i, j \in \{d+1, \cdots, \tilde{d}\}$ and $n_k = 0$ for all $k \in [\tilde{d}]$, $k \neq i$, $k \neq j$. There are $C_2^{(\tilde{d}-d)}$ number of such $\{n_i\} \in \mathcal{S}$ satisfying this. Hence, only $C_2^{\tilde{d}} - C_2^{(\tilde{d}-d)}$ number of $\{n_i\} \in \mathcal{S}$ belonging to this second class contribute to the sum of (C1). It can be checked that for each of these $C_2^{\tilde{d}} - C_2^{(\tilde{d}-d)}$ number of $\{n_i\} \in \mathcal{S}$, $N_{\{n_i\}} = 2$ and $\max_{i=1,\cdots,d}\{n_i\} = 1$.

Therefore, we have, from Eq. (C1),

$$S_c(2, \tilde{d}, d) = \frac{1}{2\tilde{d}^2} \left[2d + 2 \left(C_2^{\tilde{d}} - C_2^{(\tilde{d}-d)} \right) \right]$$

$$= \frac{1}{2\tilde{d}^2} [d + 2d\tilde{d} - d^2]. \quad \text{(C3)}$$

Similarly, following the same analysis as above, we can get the expression for $n = 3$, $d_y = \tilde{d}$ for all y, and $d \leqslant \tilde{d}$,

$$S_c(3, \tilde{d}, d) = \frac{1}{3\tilde{d}^3} \sum_{\{n_i\} \in \mathcal{S}} \left[N_{\{n_i\}} \max_{i=1,\cdots,d} \{n_i\} \right], \quad \text{(C4)}$$

where $N_{\{n_i\}}$ is the number of input dit strings with a given $\{n_i\}$ and $\mathcal{S}$ denotes the set of $\{n_i\}$ satisfying

$$\sum_{i=1}^{\tilde{d}} n_i = 3, \quad \text{(C5)}$$

such that $n_i \leqslant 3$ for all i. Now there are three cases that satisfy Eq. (C5):

(1) For each $i \in [\tilde{d}]$, $n_i = 3$ and $n_j = 0$ for all $j \neq i$ and $j \in [\tilde{d}]$.

There are $\tilde{d}$ number of such $\{n_i\} \in \mathcal{S}$. However, $\max_{i=1,\cdots,d}\{n_i\} = 0$ for each of those $\{n_i\} \in \mathcal{S}$ satisfying $n_i = 3$ for any i with $i \in \{d+1, \cdots, \tilde{d}\}$ and $n_j = 0$ for all $j \in [\tilde{d}]$ and $j \neq i$. Hence, only d number of $\{n_i\} \in \mathcal{S}$ belonging to this class contribute to the sum of (C4). For each of these d number of $\{n_i\} \in \mathcal{S}$, we have that $N_{\{n_i\}} = 1$ and $\max_{i=1,\cdots,d}\{n_i\} = 3$. Hence, the contribution to the sum is $3d$.

(2) For each $i, j, k \in [\tilde{d}]$ with $i \notin \{j, k\}$, $j \notin \{i, k\}$, $k \notin \{i, j\}$, $n_i = n_j = n_k = 1$ and $n_l = 0$ for all $l \notin \{i, j, k\}$ and $l \in [\tilde{d}]$.

There are $C_3^{\tilde{d}}$ number of such $\{n_i\} \in \mathcal{S}$. Moreover, $\max_{i=1,\cdots,d}\{n_i\} = 0$ for each of those $\{n_i\} \in \mathcal{S}$ satisfying $n_i = n_j = n_k = 1$ for any choice of i, j, k with $i, j, k \in \{d+1, \cdots, \tilde{d}\}$, $i \notin \{j, k\}$, $j \notin \{i, k\}$, $k \notin \{i, j\}$ and $n_l = 0$ for all $l \in [\tilde{d}]$ and $l \notin \{i, j, k\}$. There are $C_3^{(\tilde{d}-d)}$ number of such $\{n_i\} \in \mathcal{S}$ satisfying this. Thus, only $C_3^{\tilde{d}} - C_3^{(\tilde{d}-d)}$ number of $\{n_i\} \in \mathcal{S}$ belonging to this class contribute to the sum of (C4). It can be checked that for each of these $C_3^{\tilde{d}} - C_3^{(\tilde{d}-d)}$ number of $\{n_i\} \in \mathcal{S}$, $N_{\{n_i\}} = 3!$ and $\max_{i=1,\cdots,d}\{n_i\} = 1$. Therefore, the contribution to the sum will be $(3!)(C_3^{\tilde{d}} - C_3^{(\tilde{d}-d)})$.

(3) For each $i, j \in [\tilde{d}]$ with $i \neq j$, $n_i = 2$, $n_j = 1$ and $n_k = 0$ for all $k \notin \{i, j\}$ with $k \in [\tilde{d}]$.

The feasible solutions of Eq. (C5) that contribute to Eq. (C4) are of two types:

(A) $i \in [d]$ and $j \in [\tilde{d}] - \{i\}$. The number of possible such $\{n_i\} \in \mathcal{S}$ is given by $d(\tilde{d} - 1)$. Also, for each such $\{n_i\}$, we have that $N_{\{n_i\}} = 3$ and $\max_{i=1,\cdots,d}\{n_i\} = 2$. Therefore, the contribution to the sum appearing in Eq. (C4) by this case is $6 d(\tilde{d} - 1)$.

(B) $i \in \{d+1, \cdots, \tilde{d}\}$ and $j \in [d]$. The number of possible such $\{n_i\} \in \mathcal{S}$ is given by $d(\tilde{d} - d)$. And for each such $\{n_i\}$, we have that $N_{\{n_i\}} = 3$ and $\max_{i=1,\cdots,d}\{n_i\} = 1$. Hence, the contribution to the sum appearing in Eq. (C4) for this case is given by $3 d(\tilde{d} - d)$.

Therefore, the total contribution to the sum of Eq. (C4) is given by $6 d(\tilde{d} - 1) + 3 d(\tilde{d} - d) = 3 d(3\tilde{d} - d - 2)$.

Therefore, we have

$$S_c(3, \tilde{d}, d) = \frac{1}{3\tilde{d}^3} \left[(3!) \left(C_3^{\tilde{d}} - C_3^{\tilde{d}-d} \right) + 3 d(3\tilde{d} - d - 2) + 3d \right]$$

$$= \frac{d}{3\tilde{d}^3} [d^2 - 1 + 3\tilde{d}(\tilde{d} + 1 - d)]. \quad \text{(C6)}$$

APPENDIX D: PROOF OF RESULT 4

Let us take three arbitrary orthonormal bases $\{|\psi_1^1\rangle, |\psi_2^1\rangle\}$, $\{|\psi_1^2\rangle, |\psi_2^2\rangle\}$, and $\{|\psi_1^3\rangle, |\psi_2^3\rangle\}$ in $\mathbb{C}^2$ such that $M_{x_y|y} = |\psi_{x_y}^y\rangle\langle\psi_{x_y}^y|$ with $x_y \in [2]$ for all $y \in \{1, 2, 3\}$. A unitary can always be applied to these three measurements. Therefore, without any loss of generality, we can assume that

$$|\psi_{x_1}^1\rangle\langle\psi_{x_1}^1| = \tfrac{1}{2}[\mathbb{1} + (-1)^{x_1}\sigma_z] \quad \text{with } x_1 \in [2], \quad \text{(D1)}$$

$$|\psi_{x_2}^2\rangle\langle\psi_{x_2}^2| = \tfrac{1}{2}[\mathbb{1} + (-1)^{x_2}(\alpha\sigma_z + \sqrt{1-\alpha^2}\sigma_x)] \quad \text{with } x_2 \in [2], \quad \text{(D2)}$$

$$|\psi_{x_3}^3\rangle\langle\psi_{x_3}^3| = \tfrac{1}{2}[\mathbb{1} + (-1)^{x_3}(\beta\sigma_z + \gamma\sqrt{1-\beta^2}\sigma_x$$
$$\pm \sqrt{1-\beta^2}\sqrt{1-\gamma^2}\sigma_y)] \quad \text{with } x_3 \in [2], \quad \text{(D3)}$$

where $-1 \leqslant \alpha, \beta, \gamma \leqslant 1$.

Due to the same reasoning as for (A1), the maximum average success probability for the above-mentioned given set of three rank-one projective qubit measurements is given by

$$S_c(n=3, \tilde{d}=2, d=2) = \frac{1}{24} \sum_{x_1,x_2,x_3=1}^{2} ||M_{x_1|1}+M_{x_2|2}+M_{x_3|3}||.$$

(D4)

By definition, $||M_{x_1|1} + M_{x_2|2} + M_{x_3|3}||$ is the maximum eigenvalue of $(M_{x_1|1} + M_{x_2|2} + M_{x_3|3})$, which can be evaluated easily. Subsequently, it can be checked that

$$\sum_{x_1,x_2,x_3=1}^{2} ||M_{x_1|1} + M_{x_2|2} + M_{x_3|3}||$$

$$= 12 + \sqrt{3 + 2\alpha - 2\beta - 2\alpha\beta - 2\gamma\sqrt{1-\alpha^2}\sqrt{1-\beta^2}}$$

$$+ \sqrt{3 - 2\alpha + 2\beta - 2\alpha\beta - 2\gamma\sqrt{1-\alpha^2}\sqrt{1-\beta^2}}$$

$$+ \sqrt{3 - 2\alpha - 2\beta + 2\alpha\beta + 2\gamma\sqrt{1-\alpha^2}\sqrt{1-\beta^2}}$$

$$+ \sqrt{3 + 2\alpha + 2\beta + 2\alpha\beta + 2\gamma\sqrt{1-\alpha^2}\sqrt{1-\beta^2}}.$$

(D5)

We have found the minimum of the above expression (D5) by performing numerical optimization. It turns out that

$$\min_{\alpha,\beta,\gamma\in[-1,1]} \left(\sum_{x_1,x_2,x_3=1}^{2} ||M_{x_1|1} + M_{x_2|2} + M_{x_3|3}|| \right) = 18.$$ (D6)

In other words,

$$\min_{\alpha,\beta,\gamma\in[-1,1]} (\xi_1 + \xi_2 + \xi_3 + \xi_4) = 6,$$ (D7)

where

$$\xi_1 = \sqrt{3 + 2\alpha - 2\beta - 2\alpha\beta - 2\gamma\sqrt{1-\alpha^2}\sqrt{1-\beta^2}},$$

$$\xi_2 = \sqrt{3 - 2\alpha + 2\beta - 2\alpha\beta - 2\gamma\sqrt{1-\alpha^2}\sqrt{1-\beta^2}},$$

$$\xi_3 = \sqrt{3 - 2\alpha - 2\beta + 2\alpha\beta + 2\gamma\sqrt{1-\alpha^2}\sqrt{1-\beta^2}},$$

$$\xi_4 = \sqrt{3 + 2\alpha + 2\beta + 2\alpha\beta + 2\gamma\sqrt{1-\alpha^2}\sqrt{1-\beta^2}}.$$

In order to prove Result 4, it is sufficient to show that $(\xi_1 + \xi_2 + \xi_3 + \xi_4) = 6$ only if the three projective measurements are compatible or $(\alpha, \beta, \gamma) = \{(\pm 1/2, \pm 1/2, -1), (\pm 1/2, \mp 1/2, 1)\}$.

Since $-1 \leqslant \alpha, \beta, \gamma \leqslant 1$, we divide the regions of α, β, and γ into the following subregions:

(i) $\alpha, \beta, \gamma \in [0, 1]$,
(ii) $\alpha \in [-1, 0]; \beta, \gamma \in [0, 1]$,
(iii) $\alpha, \beta \in [-1, 0]; \gamma \in [0, 1]$,
(iv) $\alpha, \beta, \gamma \in [-1, 0]$,
(v) $\alpha, \gamma \in [0, 1]; \beta \in [-1, 0]$,
(vi) $\alpha, \gamma \in [-1, 0]; \beta \in [0, 1]$,
(vii) $\alpha, \beta \in [0, 1]; \gamma \in [-1, 0]$,
(viii) $\alpha \in [0, 1]; \beta, \gamma \in [-1, 0]$.

We start by considering the above-mentioned subregion (i), i.e., $\alpha, \beta, \gamma \in [0, 1]$. In this case, we note the following holds from numerical evaluation:

$$\min_{\alpha,\beta,\gamma\in[0,1]} (\xi_1 + \xi_2) = 2$$ (D8)

and

$$\min_{\alpha,\beta,\gamma\in[0,1]} (\xi_3) \geqslant \min_{\alpha,\beta\in[0,1]} \sqrt{3 - 2\alpha - 2\beta + 2\alpha\beta} = 1,$$ (D9)

since the derivative of the above expression is zero at $\alpha = 1$ or/and $\beta = 1$.

Next, we evaluate the maximum as well as minimum of $(\xi_1 + \xi_2 + \xi_3)$ numerically under the constraint that $(\xi_1 + \xi_2 + \xi_3 + \xi_4) = 6$. It is obtained that

$$\min_{\alpha,\beta,\gamma\in[0,1]} (\xi_1 + \xi_2 + \xi_3)$$

$$= \max_{\alpha,\beta,\gamma\in[0,1]} (\xi_1+\xi_2+\xi_3) = 3 \text{ when } \xi_1+\xi_2 + \xi_3 + \xi_4 = 6.$$

(D10)

Therefore, we have that

$$\xi_1 + \xi_2 + \xi_3 = 3 \text{ when } \xi_1 + \xi_2 + \xi_3 + \xi_4 = 6.$$ (D11)

Hence, the following is implied from (D8), (D9), and (D11):

$$\xi_1 + \xi_2 = 2, \quad \xi_3 = 1, \quad \text{and} \quad \xi_4 = 3 \text{ when}$$
$$\xi_1 + \xi_2 + \xi_3 + \xi_4 = 6.$$ (D12)

Next, it can be checked that $\xi_3 = 1$ only if $\alpha = 1$ or/and $\beta = 1$. Now, when $\alpha = 1$, then $\xi_4 = 3$ implies that $\beta = 1$. Similarly, when $\beta = 1$, then $\xi_4 = 3$ implies that $\alpha = 1$. Therefore, when $\alpha, \beta, \gamma \in [0, 1]$, then $(\xi_1 + \xi_2 + \xi_3 + \xi_4) = 6$ holds only if $\alpha = \beta = 1$.

Next, consider the subregion (iii), i.e., for $\alpha, \beta \in [-1, 0]$; $\gamma \in [0, 1]$. We note that if $\alpha \to -\alpha$ and $\beta \to -\beta$, then the four expressions ξ_i interchange among themselves as we can readily verify, $\xi_1 \to \xi_2, \xi_2 \to \xi_1, \xi_3 \to \xi_4$, and $\xi_4 \to \xi_3$. Thus, following a similar calculation as for subregion (i), we find that $\xi_1 + \xi_2 + \xi_3 + \xi_4 = 6$ holds only if $\alpha = \beta = -1$. Similarly, for subregions (vi) and (viii), one can show that $\xi_1 + \xi_2 + \xi_3 + \xi_4 = 6$ holds only if $-\alpha = \beta = 1$ and $\alpha = -\beta = 1$, respectively.

Next, let us focus on the subregion (iv), i.e., when $\alpha, \beta, \gamma \in [-1, 0]$. In this case, we obtain the following by performing numerical optimizations:

$$\min_{\alpha,\beta,\gamma\in[-1,0]} (\xi_1 + \xi_4) = \max_{\alpha,\beta,\gamma\in[-1,0]} (\xi_1 + \xi_4) = 2$$
$$\text{when } \xi_1 + \xi_2 + \xi_3 + \xi_4 = 6,$$

$$\min_{\alpha,\beta,\gamma\in[-1,0]} (\xi_2 + \xi_4) = \max_{\alpha,\beta,\gamma\in[-1,0]} (\xi_2 + \xi_4) = 2$$
$$\text{when } \xi_1 + \xi_2 + \xi_3 + \xi_4 = 6,$$

$$\min_{\alpha,\beta,\gamma\in[-1,0]} (\xi_1 + \xi_3) = \max_{\alpha,\beta,\gamma\in[-1,0]} (\xi_1 + \xi_3) = 4$$
$$\text{when } \xi_1 + \xi_2 + \xi_3 + \xi_4 = 6,$$

$$\min_{\alpha,\beta,\gamma\in[-1,0]} (\xi_2 + \xi_3) = \max_{\alpha,\beta,\gamma\in[-1,0]} (\xi_2 + \xi_3) = 4$$
$$\text{when } \xi_1 + \xi_2 + \xi_3 + \xi_4 = 6.$$

Hence, we can infer that whenever $\xi_1 + \xi_2 + \xi_3 + \xi_4 = 6$,

$$\xi_1 + \xi_4 = \xi_2 + \xi_4 = 2, \quad \xi_1 + \xi_3 = \xi_2 + \xi_3 = 4. \quad \text{(D13)}$$

Therefore, we have $\xi_1 = \xi_2$ if $\xi_1 + \xi_2 + \xi_3 + \xi_4 = 6$. Also, it can be easily checked from the expressions of ξ_1, ξ_2, ξ_3, and ξ_4 that

$$\xi_1^2 + \xi_2^2 + \xi_3^2 + \xi_4^2 = 12. \quad \text{(D14)}$$

By putting $\xi_1 = \xi_2 = \xi, \xi_3 = 4 - \xi, \xi_4 = 2 - \xi$, we get from (D14) that

$$2\xi^2 + (4 - \xi)^2 + (2 - \xi)^2 = 12. \quad \text{(D15)}$$

The possible solutions of the above equation are $\xi = 1$ and $\xi = 2$.

Before proceeding, let us point out the following observations that can be checked numerically:

$$\min_{\alpha, \beta, \gamma \in [-1, 0]} (\xi_1) = \min_{\alpha, \beta, \gamma \in [-1, 0]} (\xi_2) = 1. \quad \text{(D16)}$$

First we take $\xi = 1$. It can be shown that $\alpha, \beta, \gamma \in [-1, 0]$, $\xi_1 = \xi_2 = 1$ only if $\alpha = -1$ and $\beta = -1$. Next, let us take $\xi = 2$. Consequently, we have that $\xi_1 = \xi_2 = 2, \xi_3 = 2, \xi_4 = 0$. It can be checked that the unique solution of these four equations is given by $\alpha = -1/2, \beta = -1/2, \gamma = -1$.

Next, we remark that for the remaining subregions (ii), (v), and (vii) wherein the variables α, β, γ changes their signs with respect to the subregion (iv) where $\alpha, \beta, \gamma \in [-1, 0]$, the four expressions ξ_i interchange among themselves. Thus, a similar calculation applies to these three regions and, consequently, the solution for $\xi_1 + \xi_2 + \xi_3 + \xi_4 = 6$ is the same with the appropriate signs.

Finally, let us note that there are, in general, two cases where we do not observe any advantage. First, $\alpha = \pm 1$ and $\beta = \pm 1$, which implies that the three measurements $\{M_{x_1|1}\}$, $\{M_{x_2|2}\}$, and $\{M_{x_3|3}\}$ are compatible. Second, $(\alpha, \beta, \gamma) = \{(\pm 1/2, \pm 1/2, -1), (\pm 1/2, \mp 1/2, 1)\}$, which are obtained in subregions (iv), (ii), (v), and (vii), implies that the three measurements are incompatible. This completes the proof.

---

[1] O. Gühne, E. Haapasalo, T. Kraft, J.-P. Pellonpää, and R. Uola, Rev. Mod. Phys. **95**, 011003 (2023).

[2] A. Fine, J. Math. Phys. **23**, 1306 (1982).

[3] M. M. Wolf, D. Perez-Garcia, and C. Fernandez, Phys. Rev. Lett. **103**, 230402 (2009).

[4] R. Uola, T. Moroder, and O. Gühne, Phys. Rev. Lett. **113**, 160403 (2014).

[5] M. T. Quintino, T. Vértesi, and N. Brunner, Phys. Rev. Lett. **113**, 160402 (2014).

[6] T. Pramanik, M. Kaplan, and A. S. Majumdar, Phys. Rev. A **90**, 050305(R) (2014).

[7] R. Uola, C. Budroni, O. Gühne, and J.-P. Pellonpää, Phys. Rev. Lett. **115**, 230402 (2015).

[8] P. Chowdhury, T. Pramanik, and A. S. Majumdar, Phys. Rev. A **92**, 042317 (2015).

[9] P. Busch, P. Lahti, and R. F. Werner, J. Math. Phys. **55**, 042111 (2014).

[10] A. G. Maity, S. Datta, and A. S. Majumdar, Phys. Rev. A **96**, 052326 (2017).

[11] D. Saha, M. Oszmaniec, L. Czekaj, M. Horodecki, and R. Horodecki, Phys. Rev. A **101**, 052104 (2020).

[12] Y.-C. Liang, R. W. Spekkens, and H. M. Wiseman, Phys. Rep. **506**, 1 (2011).

[13] Z.-P. Xu and A. Cabello, Phys. Rev. A **99**, 020103(R) (2019).

[14] S. Mal and A. Majumdar, Phys. Lett. A **380**, 2265 (2016).

[15] R. Uola, G. Vitagliano, and C. Budroni, Phys. Rev. A **100**, 042117 (2019).

[16] H. S. Karthik, J. P. Tej, A. R. U. Devi, and A. K. Rajagopal, J. Opt. Soc. Am. B **32**, A34 (2015).

[17] M. Banik, S. Das, and A. S. Majumdar, Phys. Rev. A **91**, 062124 (2015).

[18] R. Uola, F. Lever, O. Gühne, and J.-P. Pellonpää, Phys. Rev. A **97**, 032301 (2018).

[19] M. Banik, M. R. Gazi, S. Ghosh, and G. Kar, Phys. Rev. A **87**, 052125 (2013).

[20] G. Kar, S. Ghosh, S. K. Choudhary, and M. Banik, Mathematics **4**, 52 (2016).

[21] S.-L. Chen, C. Budroni, Y.-C. Liang, and Y.-N. Chen, Phys. Rev. Lett. **116**, 240401 (2016).

[22] S. Sarkar, D. Saha, and R. Augusiak, Phys. Rev. A **106**, L040402 (2022).

[23] C. Carmeli, T. Heinosaari, and A. Toigo, Phys. Rev. Lett. **122**, 130402 (2019).

[24] E. Bene and T. Vértesi, New J. Phys. **20**, 013021 (2018).

[25] F. Hirsch, M. T. Quintino, and N. Brunner, Phys. Rev. A **97**, 012129 (2018).

[26] P. Skrzypczyk, I. Šupić, and D. Cavalcanti, Phys. Rev. Lett. **122**, 130403 (2019).

[27] C. Carmeli, T. Heinosaari, and A. Toigo, Phys. Rev. A **98**, 012126 (2018).

[28] L. Guerini, M. T. Quintino, and L. Aolita, Phys. Rev. A **100**, 042308 (2019).

[29] M. Ioannou, P. Sekatski, S. Designolle, B. D. M. Jones, R. Uola, and N. Brunner, Phys. Rev. Lett. **129**, 190401 (2022).

[30] R. W. Spekkens, D. H. Buzacott, A. J. Keehn, B. Toner, and G. J. Pryde, Phys. Rev. Lett. **102**, 010401 (2009).

[31] X.-R. Wang, L.-Y. Wu, C.-X. Liu, T.-J. Liu, J. Li, and Q. Wang, Phys. Rev. A **99**, 052313 (2019).

[32] A. Tavakoli, A. Hameedi, B. Marques, and M. Bourennane, Phys. Rev. Lett. **114**, 170502 (2015).

[33] E. A. Aguilar, M. Farkas, D. Martínez, M. Alvarado, J. Cariñe, G. B. Xavier, J. F. Barra, G. Cañas, M. Pawłowski, and G. Lima, Phys. Rev. Lett. **120**, 230503 (2018).

[34] A. Hameedi, D. Saha, P. Mironowicz, M. Pawłowski, and M. Bourennane, Phys. Rev. A **95**, 052345 (2017).

[35] A. Hameedi, B. Marques, P. Mironowicz, D. Saha, M. Pawłowski, and M. Bourennane, Phys. Rev. A **102**, 032621 (2020).

[36] P. E. Frenkel and M. Weiner, Commun. Math. Phys. **340**, 563 (2015).

[37] C. Vieira, C. de Gois, L. Pollyceno, and R. Rabelo, arXiv:2205.05171 v2.

[38] C. de Gois, G. Moreno, R. Nery, S. Brito, R. Chaves, and R. Rabelo, PRX Quantum **2**, 030311 (2021).

[39] H. Buhrman, R. Cleve, S. Massar, and R. de Wolf, Rev. Mod. Phys. **82**, 665 (2010).

[40] R. de Wolf, Theor. Comput. Sci. **287**, 337 (2002).

[41] D. Saha, P. Horodecki, and M. Pawłowski, New J. Phys. **21**, 093057 (2019).

[42] M. Pawłowski and N. Brunner, Phys. Rev. A **84**, 010302(R) (2011).

[43] H.-W. Li, M. Pawłowski, Z.-Q. Yin, G.-C. Guo, and Z.-F. Han, Phys. Rev. A **85**, 052308 (2012).

[44] T. Lunghi, J. B. Brask, C. Ci Wen Lim, Q. Lavigne, J. Bowles, A. Martin, H. Zbinden, and N. Brunner, Phys. Rev. Lett. **114**, 150501 (2015).

[45] M. Czechlewski, D. Saha, A. Tavakoli, and M. Pawłowski, Phys. Rev. A **98**, 062305 (2018).

[46] C. Carmeli, T. Heinosaari, and A. Toigo, Europhys. Lett. **130**, 50001 (2020).

[47] A. Ambainis, A. Nayak, A. Ta-Shma, and U. Vazirani, J. ACM **49**, 496 (2002).

[48] A. Ambainis, D. Leung, L. Mancinska, and M. Ozols, arXiv:0810.2937.

[49] M. Hayashi, K. Iwama, H. Nishimura, R. Raymond, and S. Yamashita, New J. Phys. **8**, 129 (2006).

[50] C. Carmeli, T. Heinosaari, T. Miyadera, and A. Toigo, J. Math. Phys. **60**, 122202 (2019).

[51] T. Heinosaari, T. Miyadera, and D. Reitzner, Found. Phys. **44**, 34 (2014).

[52] A. Mitra and M. Farkas, Phys. Rev. A **105**, 052202 (2022).

[53] G. Sentís, B. Gendra, S. D. Bartlett, and A. C. Doherty, J. Phys. A: Math. Theor. **46**, 375302 (2013).