Article:
Meng, Shuyang, Curran, Fionnuala, Senno, Gabriel et al. (2024) Maximal intrinsic randomness of a quantum state. Physical Review A. L010403. ISSN: 1094-1622

https://doi.org/10.1103/PhysRevA.110.L010403

Maximal intrinsic randomness of a quantum state

Shuyang Meng,[1] Fionnuala Curran,[2] Gabriel Senno,[3] Victoria J. Wright,[2] Máté Farkas,[4,2] Valerio Scarani,[5,1] and Antonio Acín[2,6]

[1] *Department of Physics, National University of Singapore, 2 Science Drive 3, Singapore 117542*
[2] *ICFO-Institut de Ciències Fotòniques, The Barcelona Institute of Science and Technology, Av. Carl Friedrich Gauss 3, 08860 Castelldefels (Barcelona), Spain*
[3] *Quside Technologies S.L., C/Esteve Terradas 1, 08860 Castelldefels, Barcelona, Spain*
[4] *Department of Mathematics, University of York, Heslington, York, YO10 5DD, United Kingdom*
[5] *Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543*
[6] *ICREA - Institució Catalana de Recerca i Estudis Avançats, 08010 Barcelona, Spain*

(Dated: July 12, 2024)

One of the most counterintuitive aspects of quantum theory is its claim that there is 'intrinsic' randomness in the physical world. Quantum information science has greatly progressed in the study of intrinsic, or secret, quantum randomness in the past decade. With much emphasis on device-independent and semi-device-independent bounds, one of the most basic questions has escaped attention: how much intrinsic randomness can be extracted from a given state ρ, and what measurements achieve this bound? We answer this question for three different randomness quantifiers: the conditional min-entropy, the conditional von Neumann entropy and the conditional max-entropy. For the first, we solve the min-max problem of finding the projective measurement that minimises the maximal guessing probability of an eavesdropper. The result is that one can guarantee an amount of conditional min-entropy $H^*_{\min} = -\log_2 P^*_{\text{guess}}(\rho)$ with $P^*_{\text{guess}}(\rho) = \frac{1}{d}\left(\text{tr}\sqrt{\rho}\right)^2$ by performing suitable projective measurements. For the conditional von Neumann entropy, we find that the maximal value is $H^* = \log_2 d - S(\rho)$, with $S(\rho)$ the von Neumann entropy of ρ, while for the conditional max-entropy, we find the maximal value $H^*_{\max} = \log_2 d + \log_2 \lambda_{\max}(\rho)$, where $\lambda_{\max}(\rho)$ is the largest eigenvalue of ρ. Optimal values for $H^*_{\min}$, H^* and $H^*_{\max}$ are achieved by measuring in any basis that is unbiased with respect to the eigenbasis of ρ, as well as by other, less intuitive, measurements.

I. INTRODUCTION

One of the core differences between classical and quantum physics is the latter's probabilistic character, which is irreducible to ignorance of underlying variables. This difference has fundamental implications for our worldview, but it is also attractive as a natural source of randomness for practical uses. Indeed, Geiger counting was already used as a source of physical randomness in the second half of the 20th century. In the past two decades, with the development of quantum information science, a large number of quantum random number generators (QRNGs) have been designed, and many have been implemented, usually with light (see [1, 2] for comprehensive reviews). The *amount of randomness* is naturally captured by the *guessing probability* P_{guess}: the higher the probability that the random variable is guessed, the smaller the randomness. This intuitive characterisation was found to have operational meaning: the *min-entropy* $H_{\min} = -\log_2 P_{\text{guess}}$ quantifies (informally) the fraction of perfect coin tosses that can be extracted from a string generated by the available source. But randomness is not an absolute notion: one has to specify *for whom* the source should be partly unpredictable. For mere sampling purposes, it might be sufficient to take the observed probabilities at face value; for cryptographic applications, however, one needs to estimate the probability that *an adversary, Eve,* guesses the outcomes. The resulting randomness is called *secret randomness*, or *intrinsic randomness*.

The computation of intrinsic randomness using quantum resources and against a quantum adversary has been studied from different perspectives. When considering a user with classical data correlated with quantum information in the hands of an adversary, the min-entropy quantifies the amount of perfect random bits that the user can establish [3]. The question was also addressed for the task of quantum key distribution, which is the extraction of secret *shared* randomness. It was in this context that the idea of device-independent certification was born: the possibility of bounding the amount of randomness in a black-box setting, based on the observation of Bell-nonlocal correlations [4]. Next, it was noticed that device-independent certification can be performed for randomness as well [5, 6], providing the first disruptive case for quantum randomness in a non-shared setting [7]. This breakthrough happened as the race to demonstrate loophole-free Bell tests was taking up speed. There followed an explosion of designs and implementations of QRNGs certifiable under various assumptions, from device-independent (disruptive, but hard to implement), to semi-device-independent in various forms, to fully characterised (practical and fast, but requiring a precise modelling of the setups). For these developments, we refer to the reviews [1, 2, 8, 9].

In this flurry of activity, one of the most basic questions was somehow left out: *how much secret randomness can be extracted from a known state ρ.* In this paper, we solve this problem for three of the most natural and op-

erational measures of randomness: the conditional min-entropy, the conditional von Neumann entropy and the conditional max-entropy. For the first, we show that the answer is $H^*_{\min} = -\log_2 P^*_{\text{guess}}(\rho)$, with

$$P^*_{\text{guess}}(\rho) = \frac{1}{d}\left(\operatorname{tr}\sqrt{\rho}\right)^2, \tag{1}$$

where d is the dimension of the Hilbert space of the system, assumed to be finite. We find a family of measurements that generate this amount of randomness, which is closely related to the concept of 'pretty good measurements' [10], originally used as a close-to-optimal way to distinguish an ensemble of states. For the second, we find the maximal value

$$H^* = \log_2 d - S(\rho), \tag{2}$$

where $S(\rho) = -\operatorname{tr}\left(\rho\log_2\rho\right)$ is the von Neumman entropy of ρ, while for the third, we find

$$H^*_{\max} = \log_2 d + \log_2 \lambda_{\max}(\rho), \tag{3}$$

where $\lambda_{\max}(\rho)$ is the largest eigenvalue of ρ. Interestingly, for $d > 2$, we find that some measurements maximise one of $H_{\min}$, H and $H_{\max}$, but not the other two.

II. QUBIT EXAMPLE

A case study will help to introduce the main ideas. Alice has a source that produces a qubit. She has characterised its state to the best of her knowledge and found it to be

$$\rho = \frac{1}{2}(\mathbb{1} + m\sigma_z) = \frac{1+m}{2}\,|0\rangle\langle 0| + \frac{1-m}{2}\,|1\rangle\langle 1| \tag{4}$$

for some $0 \le m \le 1$. If she measures σ_x, her observed statistics will be those of a perfect unbiased coin: $P_A(+1) = P_A(-1) = \frac{1}{2}$. Suppose now that what the source really does is produce a pure state in each round, specifically half of the rounds $|\chi_+\rangle$ and half of the rounds $|\chi_-\rangle$, with

$$|\chi_\pm\rangle = \sqrt{\frac{1 \pm \sqrt{1-m^2}}{2}}\,|+x\rangle + \sqrt{\frac{1 \mp \sqrt{1-m^2}}{2}}\,|-x\rangle \tag{5}$$

(indeed, $\frac{1}{2}\,|\chi_+\rangle\langle\chi_+| + \frac{1}{2}\,|\chi_-\rangle\langle\chi_-| = \rho$). If Eve knows the working of the source exactly, she will guess $i = +1$ ($i = -1$) in the rounds when the source sent out $|\chi_+\rangle$ ($|\chi_-\rangle$). Her guess will then be correct with probability

$$P_{\text{guess}} = \frac{1}{2}\left(1 + \sqrt{1-m^2}\right), \tag{6}$$

which is strictly larger than $\frac{1}{2}$ when $m < 1$ (i.e. when ρ is mixed). Thus, the intrinsic randomness of Alice's protocol is less than her apparent perfect randomness. In particular, there is no secret randomness in the state $\rho = \frac{1}{2}\mathbb{1}$, since $P_{\text{guess}} = 1$ for $m = 0$.

As will be expanded on in what follows, two things are already known about this case study and its generalisation to higher dimensions. First: we presented this example with Eve having perfect classical information about the source, in the sense that she knows at each instance which state has been prepared and accordingly makes her guess on Alice's measurement outcome. However, the result is unchanged if Eve holds quantum side-information. Eve then holds a purification of Alice's state, and she measures her own system to guess Alice's result. Since the two scenarios are equivalent in terms of the guessing probability, we will move from one to the other when convenient for the argumentation. Second: having fixed Alice's protocol (both the state and the measurement), the maximisation of P_{guess} over all decompositions of ρ is a known semidefinite program (SDP) [11]; in the case study, we have presented the optimal decomposition. What is not known is whether σ_x is the best measurement for Alice, even in the presence of Eve: could another measurement on the same state ρ decrease Eve's guessing probability, at the expense of biasing the observed P_A? We set out to solve this min-max problem, and thus determine the maximal amount of secret randomness that can be extracted from ρ.

III. SETTING OF THE PROBLEM

Alice holds a quantum state ρ from a Hilbert space of dimension d. We want to determine how much intrinsic randomness she can extract from ρ and which measurement achieves this maximum. We consider only measurements $\mathcal{M} = \{M_i\}_i$ which are projective, i.e. $M_i M_j = \delta_{ij} M_i$, where δ_{ij} is the Kronecker delta (we discuss general POVMs at the end of this section). To quantify how intrinsically random, that is, how unpredictable, Alice's measurement outcome is, one considers the existence of an eavesdropper, Eve, who has a more detailed knowledge than Alice about the process, but cannot actively influence it (she is 'outside the lab'). Concretely, in every round, Eve knows the true state ρ_c produced by the source. Given this knowledge, she guesses the most likely outcome $i = i(c)$ for that round. Without loss of generality, we can group together all of Eve's states that lead to the same guessed outcome, since Eve does not gain anything in treating them as distinct. We denote by ρ_i the states seen by Eve, sub-normalised such that $q_i = \operatorname{tr}\rho_i$ is the probability that Eve's most likely outcome is i. These states must satisfy $\sum_i \rho_i = \rho$.

Having set this stage, Eve's average guessing probability is $P_{\text{guess}}\big(\{\rho_i\}, \mathcal{M}\big) = \sum_i \operatorname{tr}(M_i\rho_i)$. Since we don't know the true states ρ_i, we need to consider the worst case scenario, i.e. the decomposition that maximises Eve's guessing probability,

$$P_{\text{guess}}(\rho, \mathcal{M}) \;=\; \max_{\{\rho_i\}} \sum_i \operatorname{tr}(M_i\rho_i) \tag{7}$$

$$\text{s.t.}\;\; \rho_i \geq 0\,,\;\; \sum_i \rho_i = \rho\,.$$

This optimisation is an SDP, and so can be solved efficiently. In order to determine, *the maximal amount of secret randomness that can be extracted from the known state* ρ, one need to optimize Eq. (7) over Alice's measurement, i.e. compute

$$P^*_{\text{guess}}(\rho) = \min_{\mathcal{M}\in\Pi} P_{\text{guess}}(\rho,\mathcal{M})\,, \tag{8}$$

where Π is the set of all projective measurements. Our main result is to show that Eq. (1) is the solution to the optimisation (8).

The search for an optimal measurement could have been extended to the larger set of Positive Operator-Valued Measures (POVMs), but the operational interpretation in our context is unclear. Recall that our goal is to quantify the secret randomness *in the state* ρ. When implementing a POVM, however, the projective measurement acts on the given state ρ plus an auxiliary system, so part of the obtained randomness may come from the latter. In fact, for extremal measurements minimising the guessing probability, the auxiliary system has to be in a pure state, say $|a\rangle$, of dimension d_A [12]. It follows from our main result that the maximal amount of randomness obtained when implementing a projective measurement on the global state is $P^*_{\text{guess}}\big(\rho\otimes|a\rangle\langle a|\,\big)$. It is easy, however, to see that $P^*_{\text{guess}}\big(\rho\otimes|a\rangle\langle a|\,\big) = \frac{1}{d_A}P^*_{\text{guess}}(\rho)$, that is, the optimal guessing probability is equal to that obtained by performing the corresponding optimal projective measurements independently on the system and the auxiliary.

Thus, the extra randomness supplied by using the optimal POVM is exactly equal to the intrinsic randomness of the auxiliary system used to implement the POVM, so we view it as arising from the auxiliary rather than from ρ itself.

IV. MAIN RESULTS

Theorem 1. *The maximal amount of secret randomness that can be extracted from a quantum state* ρ *using a projective measurement is given by* $H^*_{min} = -\log_2 P^*_{\text{guess}}(\rho)$ *with* $P^*_{\text{guess}}(\rho) = \frac{1}{d}\left(\text{tr}\,\sqrt{\rho}\right)^2$.

Without loss of generality, one can restrict the optimisation to rank-one projective measurements (see Appendix B 1). In what follows, we outline a proof that uses notions from state discrimination and the resource theory of coherence, with full details in Appendix C. An alternative proof using properties of the min-entropy and semidefinite programming is provided in Appendix D. We prove the theorem by first proving the lower bound $P^*_{\text{guess}}(\rho) \geq \frac{1}{d}\left(\text{tr}\,\sqrt{\rho}\right)^2$ (Lemma 1) and then showing that there exist measurements that achieve that bound (Lemma 2).

Lemma 1. *The lower bound* $P^*_{\text{guess}}(\rho) \geq \frac{1}{d}\left(\text{tr}\,\sqrt{\rho}\right)^2$ *holds for every state* ρ.

Proof. Using the fact that rank-one measurements are optimal for Alice, from [13, Theorem 1, (iii)], we find

$$P_{\text{guess}}(\rho,\mathcal{M}) = \max_{\{\sigma\in\mathcal{I}_{\mathcal{M}}\}} F(\rho,\sigma)\,, \tag{9}$$

where F is the Uhlmann fidelity and $\mathcal{I}_{\mathcal{M}}$ is the set of states that are diagonal in the measurement basis $\{|m_i\rangle\}$. Notice that $\mathbb{1}/d \in \mathcal{I}_{\mathcal{M}}$ for all $\mathcal{M}\in\Pi$, so $P_{\text{guess}}(\rho,\mathcal{M}) \geq F(\rho,\mathbb{1}/d) = \frac{1}{d}(\text{tr}\,\sqrt{\rho})^2$ for all $\mathcal{M}$. Hence, $P^*_{\text{guess}}(\rho)$ cannot be smaller than $\frac{1}{d}(\text{tr}\,\sqrt{\rho})^2$. □

Lemma 2. *A projective measurement* $\mathcal{M}$ *in the basis* $\{|m_i\rangle\}$ *achieves the bound* $P_{\text{guess}}(\rho,\mathcal{M}) = \frac{1}{d}(\text{tr}\,\sqrt{\rho})^2$ *if and only if* $\langle m_i|\sqrt{\rho}|m_i\rangle = \frac{1}{d}\,\text{tr}\,\sqrt{\rho}$ *for all* $i = 1,...,d$.

Proof. The details missing here are provided in Appendix C 1. In the quantum side-information scenario, any k-outcome rank-one measurement, $\mathcal{M}$, by Alice steers k pure states on Eve. To optimise her guess, Eve has to measure her system to optimally discriminate among these k states. It is known [14] that the best discrimination of a set of pure states is obtained with rank-one measurements: thus, Eve will also perform a rank-one measurement. In turn, Eve's measurement defines an ensemble realising the mixed state ρ. This ensemble consists of k pure states $\rho_i = \left|\tilde{\psi}_i\right\rangle\!\left\langle\tilde{\psi}_i\right|$. If $\mathcal{M}$ is projective, we have $k = d$ and it follows from [15] that any decomposition of ρ in d pure states is defined by the choice of an orthonormal basis $\{|i\rangle\}$ through

$$|\tilde{\psi}_i\rangle = \sqrt{\rho}\,|i\rangle \quad \text{with} \quad \langle i|i'\rangle = \delta_{ii'}\,,\; i,i' = 1,...,d\,. \tag{10}$$

Inserting all these observations in (7), we obtain

$$P_{\text{guess}}(\rho,\mathcal{M}) = \max_{\{|i\rangle\}} \sum_i |\,\langle m_i|\sqrt{\rho}|i\rangle\,|^2\,. \tag{11}$$

The r.h.s. has been called the geometric coherence of ρ [16] and was shown in [17] to be equivalent to $\max_{\{\sigma\in\mathcal{I}_{\mathcal{M}}\}} F(\rho,\sigma)$, with $\mathcal{I}_{\mathcal{M}}$ the set of states diagonal in the basis $\{|m_i\rangle\}$. If we rewrite (11) as

$$\begin{aligned} P_{\text{guess}}(\rho,\mathcal{M}) = \quad & \max_{\{\Pi_i\}_i} \sum_i \text{tr}\left(\Pi_i\,|\tilde{\gamma}_i\rangle\langle\tilde{\gamma}_i|\,\right) \\ \text{s.t.} \quad & \Pi_i \geq 0\,,\; \sum_i \Pi_i = \mathbb{1}\,, \end{aligned} \tag{12}$$

the r.h.s defines the optimal discrimination of the subnormalised states $|\tilde{\gamma}_i\rangle := \sqrt{\rho}\,|m_i\rangle$ with a projective measurement $\Pi_i = |i\rangle\langle i|$. One then checks (see Appendix C 2) that, under the assumption that

$$\langle m_i|\sqrt{\rho}|m_i\rangle = \frac{1}{d}\,\text{tr}\,\sqrt{\rho} \;\text{ for all }\; i = 1,...,d\,, \tag{13}$$

the choice $|i\rangle = |m_i\rangle$ fulfills all the conditions for optimal discrimination of the $|\tilde{\gamma}_i\rangle$ [18–20]. Thus, for measurements satisfying (13), it holds that $P_{\text{guess}}(\rho,\mathcal{M}) =$

$\sum_i \left|\langle m_i|\sqrt{\rho}|m_i\rangle\right|^2 = \frac{1}{d}(\mathrm{tr}\sqrt{\rho})^2$. Furthermore, we prove (see Appendix C 3) that the condition (13) is also necessary for a projective measurement to achieve the optimal guessing probability. What remains to be proven is that there exist measurements satisfying condition (13). An example of such a measurement valid for any state is the one defined by a basis $\{|m_i\rangle\}$ that is unbiased to the eigenbasis of ρ, that is, all moduli of inner products between elements of the two different bases equal $\frac{1}{\sqrt{d}}$. However, as we discuss in Section V, one can find other measurements satisfying condition (13) when $d > 2$. □

Notice that, when Alice uses measurements satisfying (13), the decomposition (10) that is optimal for Eve is $|\tilde{\psi}\rangle_i = \sqrt{\rho}\,|m_i\rangle$. If ρ is full rank, $M_i = \rho^{-1/2}\rho_i\rho^{-1/2}$ is the 'pretty good measurement' [10] for the ensemble $\{q_i, \rho_i/q_i\}$ steered by Eve. This measurement is known to be optimal when special symmetries like (13) are present in the problem [21] (in the notation of that work, the Gram matrix has entries $G_{ij} = \langle m_i|\rho|m_j\rangle$). Moreover, when Alice's measurement satisfies (13) and when ρ is full-rank, we can show (see Appendix C 4) that Eve's optimal measurement to discriminate her local states is also a 'pretty good' measurement.

After solving the problem for the guessing probability, we now move to the von Neumann entropy of the measurement outcomes conditioned on Eve's side information, a quantity of relevance in the multi-round setting [22–24].

Theorem 2. *The maximal conditional entropy that can be extracted from a quantum state ρ using a projective measurement is $H^* = \log_2 d - S(\rho)$, where $S(\rho) = -\,\mathrm{tr}\,\rho\log_2\rho$ is the von Neumann entropy.*

Proof. From [13, Theorem 1, (i)], we have that the entropy $H(Z|E)$ of Alice's measurement outcomes Z conditioned on Eve's side information E is

$$H(Z|E) = D\big(\rho\,||\sum_z M_z\rho M_z\big)\,, \tag{14}$$

where $\{M_z\}_z$ is Alice's projective measurement and $D(\rho||\sigma)$ is the quantum relative entropy between the states ρ and σ,

$$D(\rho||\sigma) = \mathrm{tr}\left(\rho\big(\log_2\rho - \log_2\sigma\big)\right), \tag{15}$$

which is defined when the support of ρ is contained within the support of σ. In Appendix B 2, we show that: 1) a rank-one measurement is optimal for Alice to maximise $H(Z|E)$ for a given ρ, and 2) that

$$D\big(\rho\,||\sum_z M_z\rho M_z\big) = S\big(\sum_z M_z\rho M_z\big) - S(\rho)\,. \tag{16}$$

In [25], the r.h.s. is shown to be equivalent to the relative entropy of coherence of ρ with respect to the measurement basis, which is used as a quantifier of randomness.

The maximum von Neumann entropy of a state of dimension d is $\log_2 d$ and is achieved only for maximally mixed states, so we can upper bound Eq. (16) with

$$H(Z|E) \leq \log_2 d - S(\rho)\,, \tag{17}$$

with equality reached if and only if Alice's measurement basis $\{|m_z\rangle\}_z$ leaves her system in the maximally mixed state, i.e. if the condition

$$\langle m_z|\rho|m_z\rangle = \frac{1}{d} \quad \text{for all} \quad z = 1, ..., d \tag{18}$$

is satisfied. □

As in the case for the condition (13) for $H^*_{\min}$, suitable measurements satisfying (18) include bases $\{|m_z\rangle\}$ that are unbiased to the eigenbasis of ρ, implying the tightness of (17). However, when $d > 2$ we can find other suitable measurements, as discussed in Section V. The quantity $\log_2 d - S(\rho)$ is defined in [26] as the total information of ρ and it is used in [27] as a measure of the objective information of ρ.

We now consider the conditional max-entropy of the measurement outcomes conditioned on Eve's side information. This quantity has been interpreted as the security of Alice's measurement outcomes when used as a secret key [3].

Theorem 3. *The maximal conditional max-entropy that can be extracted from a quantum state ρ using a projective measurement is $H^*_{\max} = \log_2 d + \log_2\lambda_{\max}(\rho)$, where $\lambda_{\max}(\rho)$ is the largest eigenvalue of ρ.*

Proof. The details missing here are given in Appendix E. Without loss of generality, we restrict Alice to performing rank-one projective measurements (see Appendix B 1). In the case where Alice makes a rank-one projective measurement, the conditional max-entropy of her outcomes conditioned on Eve can be formulated [3] as

$$H_{\max}(A|E) = \log_2 p_{\mathrm{secr}}\,, \tag{19}$$

where

$$p_{\mathrm{secr}} = \max_\sigma \left(\sum_x \sqrt{p_x\,\mathrm{tr}\left(\sigma\,|\psi^E_x\rangle\langle\psi^E_x|\right)}\right)^2 \tag{20}$$

$$\text{s.t.} \quad \sigma \geq 0\,,\ \mathrm{tr}\,\sigma = 1\,, \tag{21}$$

where $\{|\psi^E_x\rangle\}$ are Eve's post-measurement states and $p_x = \langle m_x|\rho|m_x\rangle$. By applying the Cauchy-Schwartz inequality and identifying the semidefinite optimisation problem for the maximum eigenvalue of a quantum state, we find

$$p_{\mathrm{secr}} \leq d\lambda_{\max}(\rho)\,. \tag{22}$$

In the case where the largest eigenvalue of ρ is unique, the bound (22) is reached if and only if the condition

$$|\langle m_x|u_{\max}\rangle|^2 = \frac{1}{d} \quad \text{for all} \quad x = 1, ..., d \tag{23}$$

is satisfied, where $|u_{\max}\rangle$ is the eigenvector of ρ corresponding to its largest eigenvalue. The optimal measurements in the case where the maximum eigenvalue of ρ is degenerate are discussed in Appendix E. □

As in the case of $H^*_{\min}$ and H^*, suitable measurements satisfying (23) include bases $\{|m_x\rangle\}$ that are unbiased to the eigenbasis of ρ, but, as before, when $d > 2$ we can find other suitable measurements, as discussed in Section V.

V. TWO CASE STUDIES

Let us now study measurements that satisfy (13), (18) or (23) but which are not unbiased to the eigenbasis of ρ. For one qubit, it is quickly verified that all measurements that satisfy (13) are unbiased, so our first case study is for *one qutrit*. Consider $\rho = \sum_{i=1}^{3} \lambda_i |i\rangle\langle i|$ with $\lambda_1 \geq \lambda_2 \geq \lambda_3$, and the measurement basis $\{M_i = |m_i\rangle\langle m_i|\}_{i=1,2,3}$, with

$$
\begin{aligned}
|m_1\rangle &= \sqrt{\frac{1+a}{3}}\,|1\rangle + \sqrt{\frac{1+b}{3}}\,|2\rangle + \sqrt{\frac{1+c}{3}}\,|3\rangle\ , \\
|m_2\rangle &= \sqrt{\frac{1+a}{3}}\,e^{i\theta_1}\,|1\rangle + \sqrt{\frac{1+b}{3}}\,|2\rangle + \sqrt{\frac{1+c}{3}}\,e^{i\theta_2}\,|3\rangle\ ,
\end{aligned}
\tag{24}
$$

and $|m_3\rangle$ defined by the normalization condition $\sum_i M_i = \mathbb{1}$, where $a = -(\gamma_2 - \gamma_3)k$, $b = (\gamma_1 - \gamma_3)k$, $c = -(\gamma_1 - \gamma_2)k$, $k \in \mathbb{R}$ and each $\gamma_i \geq 0$ with $\gamma_1 \geq \gamma_2 \geq \gamma_3$. We show in Appendix F 1 that suitable parameters θ_1 and θ_2 can always be chosen such that this is a valid rank-one projective measurement when k is in the range $-\frac{1}{2} \leq k \leq \frac{1}{2}$.

This measurement basis is not in general unbiased to the eigenbasis of ρ, except when ρ is maximally mixed. When we set $\{\gamma_i\} = \{\sqrt{\lambda_i}\}$, it is straightforward to show that the condition (18) for the measurement to maximise $H_{\min}$ is satisfied. Similarly, if we set $\{\gamma_i\} = \{\lambda_i\}$, we see that the condition (18) for maximal H is satisfied. Finally, the condition (23) for maximal $H_{\max}$ is satisfied when $\gamma_2 = \gamma_3$, so we see that, for qutrits at least, there exist non-unbiased measurements that achieve maximal randomness for every ρ for all three of our quantifiers of randomness. Interestingly, though, these three conditions are inequivalent in general, so one can choose parameters $\{\gamma_i\}$ such that the measurement maximises any one of the entropies but not the other two.

The second case study uses *two qubits*. It is based on the observation (proved in Appendix F 2) that there is no product basis unbiased to the basis

$$
\begin{aligned}
|\psi_1\rangle &= |00\rangle \\
|\psi_2\rangle &= \frac{1}{\sqrt{3}}(|01\rangle + |10\rangle + |11\rangle) \\
|\psi_3\rangle &= \frac{1}{\sqrt{3}}(|01\rangle + \omega\,|10\rangle + \omega^2\,|11\rangle) \\
|\psi_4\rangle &= \frac{1}{\sqrt{3}}(|01\rangle + \omega^2\,|10\rangle + \omega\,|11\rangle)\,,
\end{aligned}
\tag{25}
$$

where $\omega = e^{i\,2\pi/3}$. Consider a state $\rho = \sum_{k=1}^{4} \lambda_k\,|\psi_k\rangle\langle\psi_k|$ diagonal in this basis. To extract the maximal randomness with an unbiased measurement, one must be able to perform entangled measurements. This is not a conceptual problem in our setting, since there is no reason why the two qubits should be far apart; nonetheless, such measurements may be more challenging to perform than basic single-qubit measurements. The question is: can one extract maximal randomness from ρ by using a product basis? The answer seems to be positive. While we do not have an analytical proof, for a large number of choices of λ, we performed a heuristic optimisation over product bases, both general ($\{|a,b\rangle, |a,b^\perp\rangle, |a^\perp,c\rangle, |a^\perp,c^\perp\rangle\}$, with six free parameters) and restricted to proper product measurements ($\{|a,b\rangle, |a,b^\perp\rangle, |a^\perp,b\rangle, |a^\perp,b^\perp\rangle\}$, with four free parameters). In both cases and for all states that we probed, we numerically found measurements satisfying $\sum_i \left(\langle m_i|\sqrt{\rho}|m_i\rangle - \frac{\mathrm{tr}\sqrt{\rho}}{4}\right)^2 \leq 10^{-15}$, $\sum_i \left(\langle m_i|\rho|m_i\rangle - \frac{1}{4}\right)^2 \leq 10^{-15}$ or $\sum_i \left(|\langle m_i|u_{\max}\rangle|^2 - \frac{1}{4}\right)^2 \leq 10^{-15}$, which suggests that there exist product measurements satisfying the conditions (13), (18) and (23), respectively. In this family of examples, therefore, the freedom to choose a measurement basis that is not unbiased may lead to a practical advantage: it allows one to obtain maximal randomness with product measurements.

VI. CONCLUSION

It is well known that quantum physics contains an intrinsic form of randomness, but, somewhat surprisingly, given a quantum state, it is unknown what is the optimal measurement to extract from it the maximum amount of such randomness. In this work, we concentrate on three different quantifiers of the amount of randomness in a measurement's outcomes conditioned on an adversary's side information: the conditional min-entropy, the conditonal von Neumann entropy and the conditional max-entropy. As one might have expected, all measurements in a basis that is unbiased to the eigenbasis of ρ maximise all three of these conditional entropies. However, we also find other measurements that achieve the optimal values, providing a flexibility that may have practical implications, as in the second case study reported. In fact, beyond its fundamental motivation, our analysis is also

relevant for the design of device-dependent QRNGs, for which the quantum state is fully characterised. Interestingly, we find measurements in the qutrit case that maximise one of the three conditional entropies considered, but which are not optimal for the other two.

ACKNOWLEDGMENTS

We thank Siddhartha Das for pointing out to us the use of Eq. (2) in other contexts [26, 27]. This work is supported by the National Research Foundation, Singapore and A*STAR under its CQT Bridging Grant, the Government of Spain (Severo Ochoa CEX2019-000910-S, Torres Quevedo PTQ2021-011870, TRANQI and European Union NextGenerationEU PRTR-C17.I1), Fundació Cellex, Fundació Mir-Puig, Generalitat de Catalunya (CERCA program), the European Union (QSNP, 101114043 and Quantera project Veriqtas), the ERC AdG CERQUTE, the AXA Chair in Quantum Information Science, and the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 754510.

[1] M. Herrero-Collantes and J. C. Garcia-Escartin, Quantum random number generators, Rev. Mod. Phys. **89**, 015004 (2017).
[2] A. P. Vaisakh Mannalath, Sandeep Mishra, A comprehensive review of quantum random number generators: Concepts, classification and the origin of randomness (2022), arXiv:arXiv:2203.00261 [quant-ph].
[3] R. Konig, R. Renner, and C. Schaffner, The operational meaning of min- and max-entropy, IEEE Transactions on Information Theory **55**, 4337 (2009).
[4] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Device-independent security of quantum cryptography against collective attacks, Phys. Rev. Lett. **98**, 230501 (2007).
[5] R. Colbeck, *Quantum And Relativistic Protocols For Secure Multi-Party Computation*, Ph.D. thesis, University of Cambridge (2006).
[6] S. Pironio, A. Acín, S. Massar, A. B. de la Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe, Random numbers certified by Bell's theorem, Nature **464**, 1021 (2010).
[7] As long as *process randomness* requires characterised devices, classical and quantum RNGs compete on the same grounds for speed, stability, practicality etc. But, given an alleged RNG as a black box, on classical devices, one can only test product randomness with statistical tests. While process randomness implies product randomness, the opposite is certainly not true: one could have recorded a long enough list of random numbers, and the device under study may just be reading deterministically from that record.
[8] M. N. Bera, A. Acín, M. Kuś, M. W. Mitchell, and M. Lewenstein, Randomness in quantum mechanics: philosophy, physics and technology, Reports on Progress in Physics **80**, 124001 (2017).
[9] A. Acín and L. Masanes, Certified randomness in quantum physics, Nature **540**, 213 (2016).
[10] P. Hausladen and W. K. Wootters, A 'pretty good' measurement for distinguishing quantum states, Journal of Modern Optics **41**, 2385 (1994).
[11] Y. Z. Law, L. P. Thinh, J.-D. Bancal, and V. Scarani, Quantum randomness extraction for various levels of characterization of the devices, Journal of Physics A: Mathematical and Theoretical **47**, 424028 (2014).
[12] G. Senno, T. Strohm, and A. Acín, Quantifying the intrinsic randomness of quantum measurements, Physical Review Letters **131**, 10.1103/physrevlett.131.130202 (2023).
[13] P. J. Coles, Unification of different views of decoherence and discord, Physical Review A **85**, 10.1103/physreva.85.042103 (2012).
[14] Y. C. Eldar, A. Megretski, and G. C. Verghese, Designing optimal quantum detectors via semidefinite programming, IEEE Transactions on Information Theory **49**, 1007 (2003).
[15] L. P. Hughston, R. Jozsa, and W. K. Wootters, A complete classification of quantum ensembles having a given density matrix, Physics Letters A **183**, 14 (1993).
[16] A. Streltsov, U. Singh, H. S. Dhar, M. N. Bera, and G. Adesso, Measuring quantum coherence with entanglement, Physical review letters **115**, 020403 (2015).
[17] C. Xiong and J. Wu, Geometric coherence and quantum state discrimination, Journal of Physics A: Mathematical and Theoretical **51**, 414005 (2018).
[18] A. S. Holevo, Statistical decision theory for quantum systems, Journal of multivariate analysis **3**, 337 (1973).
[19] C. W. Helstrom, Quantum detection and estimation theory, Journal of Statistical Physics **1**, 231 (1969).
[20] H. Yuen, R. Kennedy, and M. Lax, Optimum testing of multiple hypotheses in quantum detection theory, IEEE transactions on information theory **21**, 125 (1975).
[21] N. Dalla Pozza and G. Pierobon, Optimality of square-root measurements in quantum state discrimination, Physical Review A **91**, 042334 (2015).
[22] M. Tomamichel, R. Colbeck, and R. Renner, A fully quantum asymptotic equipartition property, IEEE Transactions on information theory **55**, 5840 (2009).
[23] F. Dupuis, O. Fawzi, and R. Renner, Entropy accumulation, Communications in Mathematical Physics **379**, 867 (2020).
[24] H. Dai, B. Chen, X. Zhang, and X. Ma, Intrinsic randomness under general quantum measurements, Physical Review Research **5**, 033081 (2023).
[25] X. Yuan, Q. Zhao, D. Girolami, and X. Ma, Quantum coherence and intrinsic randomness, Advanced Quantum Technologies **2**, 1900053 (2019).
[26] W. H. Zurek, Information transfer in quantum measurements: irreversibility and amplification (2001), arXiv:quant-ph/0111137 [quant-ph].
[27] M. Horodecki, P. Horodecki, and J. Oppenheim, Reversible transformations from pure to mixed states and the unique measure of information, Physical Review A

67, 10.1103/physreva.67.062104 (2003).
[28] R. Renner, Security of quantum key distribution (2006), arXiv:quant-ph/0512258 [quant-ph].
[29] K. Modi, A. Brodutch, H. Cable, T. Paterek, and V. Vedral, The classical-quantum boundary for correlations: Discord and related measures, Reviews of Modern Physics **84**, 1655 (2012).
[30] M. Tomamichel, *Quantum Information Processing with Finite Resources* (Springer International Publishing, 2016).
[31] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).
[32] In [28, Lemma 3.1.13], on which Lemma 5 is based, Eq. (D1) is stated in terms of $H_{\min}(\rho^{AE}|\sigma^{E})$ and $\geq H_{\min}(\tilde{\rho}^{AE}|\sigma^{E})$, and for any σ_E (c.f. Eq. (A4)). It is easy to see that (D1) also holds by letting σ_E be one state achieving the maximum in $H_{\min}(A|E)_{\tilde{\rho}^{AE}}$.
[33] P. Skrzypczyk and D. Cavalcanti, *Semidefinite Programming in Quantum Information Science*, 2053-2563 (IOP Publishing, 2023).

Appendix A: Entropy definitions

Since we will use them in more than one section of this Appendix, here we state the definitons of the von Neumann conditional entropy and the min- and max-entropies of bipartite states ρ_{AE} [28].

Definition 1. *The conditional entropy of ρ_{AE} is defined by*

$$H(A|E)_{\rho_{AE}} := S(\rho_{AE}) - S(\rho_E)\,, \tag{A1}$$

where $S(\rho) = -\operatorname{tr}\left(\rho \log_2 \rho\right)$ is the von Neumann entropy of ρ and $\rho_E := \operatorname{tr}_A \rho_{AE}$.

Definition 2. *The conditional min-entropy of ρ_{AE} is defined by*

$$H_{\min}(A|E)_{\rho_{AE}} := \max_{\sigma_E}\ H_{\min}(\rho_{AE}|\sigma_E) \tag{A2}$$

with

$$H_{\min}(\rho_{AE}|\sigma_E) := -\min\{\lambda \mid 2^{\lambda}(\mathbb{1}_A \otimes \sigma_E) \geq \rho_{AE}\}\,. \tag{A3}$$

Definition 3. *The conditional max-entropy of ρ_{AE} is*

$$H_{\max}(A|E)_{\rho_{AE}} := \max_{\sigma_E}\ \log_2 \operatorname{tr}\left((\mathbb{1}_A \otimes \sigma_E)\Pi_{AE}\right), \tag{A4}$$

where Π_{AE} is the projector onto the support of ρ_{AE}.

Notice that when the system E is trivial (i.e. its Hilbert space is one-dimensional), $H_{\max}(A|E)_{\rho_{AE}} = H_{\max}(A)_{\rho_A} = \log_2 \operatorname{rank}(\rho_A)$. In the following, when the state ρ_{AE} to which we refer is clear from the context, we will drop the corresponding subscript in the notation for the conditional entropies.

In our state discrimination scenario where Alice and Eve share a bipartite state ρ_{AE}, given any POVM $\mathcal{M} = \{M_x\}_x$ for Alice, one can define a classical-quantum state (cq-state) ρ_{XE} to model the correlations between the measurement outcomes (classical information) and the corresponding post-measurement states ρ_E^x on Eve's subsystem,

$$\rho_{XE} = \sum_x p_x \left|x\right\rangle\!\left\langle x\right| \otimes \rho_E^x\,, \tag{A5}$$

where $p_x = \operatorname{tr}[M_x \rho_A]$, $\rho_E^x = \operatorname{tr}_A[(M_x \otimes \mathbb{1}_E)\rho_{AE}]/p_x$ and $\{|x\rangle\}_x$ is some orthonormal basis representing Alice's outcomes. When the POVM $\mathcal{M}$ is extremal, we have the following relation between $H_{\min}(X|E)$ and Eve's optimal guessing probability given ρ and $\mathcal{M}$ [3, 12]:

$$H_{\min}(X|E) = -\log_2 P_{\text{guess}}(\rho, \mathcal{M})\,. \tag{A6}$$

Appendix B: Proof of optimality of rank-one measurement operators

In this appendix we prove the optimality of rank-one measurements for $H_{\min}^{*}$ and H^{*}. Similar results have been obtained for other information-theoretic quantities (see, e.g., [29, Section II.I] in the context of quantum discord).

1. Optimality of rank-one measurements for $H_{\min}^{*}$ and $H_{\max}^{*}$

We show that Alice's optimal measurement can be assumed to be rank-one. First, notice that any measurement can be obtained by coarse-graining a rank-one measurement, since, given some coarse-grained (i.e. not rank-one) measurement $\mathcal{M}_{\text{coarse}} = \{M_i\}_i$, we can represent each of its elements in its spectral decomposition as $M_i = \sum_j \lambda_{ij} \left|f_j^i\right\rangle\!\left\langle f_j^i\right|$, with $\lambda_{ij} \geq 0$ for all i, j. $\mathcal{M}_{\text{coarse}}$ can then be seen as a coarse-graining of the fine-grained measurement $\mathcal{M}_{\text{fine}} = \{\left|f_j^i\right\rangle\!\left\langle f_j^i\right|\}_{i,j}$. If we restrict Alice to performing some projective measurement $\mathcal{M}_{\text{coarse}}$, the corresponding $\mathcal{M}_{\text{fine}}$ will also be projective.

The coarse-graining of $\mathcal{M}_{\text{coarse}}$ can be seen as a deterministic post-processing of the outcomes of $\mathcal{M}_{\text{fine}}$. Consider the classical-quantum state (A5) formed by the classical information of Alice's measurement outcomes and Eve's quantum states in the case of $\mathcal{M}_{\text{fine}}$. The cq-state for any coarse-graining of $\mathcal{M}_{\text{fine}}$ can be found by applying a deterministic function $f(x)$ to the classical register X,

$$\rho_{XE}^{\text{coarse}} = \sum_x p_x \left|f(x)\right\rangle\!\left\langle f(x)\right| \otimes \rho_E^x\,. \tag{B1}$$

Applying a function to a classical register cannot increase either the conditional min-entropy or the conditional max-entropy (see, e.g., [30, Proposition 6.20]). Therefore,

$$H_{\min}(X|E)_{\rho_{XE}^{\text{coarse}}} \leq H_{\min}(X|E)_{\rho_{XE}^{\text{fine}}}\,, \tag{B2}$$

$$H_{\max}(X|E)_{\rho_{XE}^{\text{coarse}}} \leq H_{\max}(X|E)_{\rho_{XE}^{\text{fine}}}\,. \tag{B3}$$

Then, by (A6), it is optimal for Alice to choose a rank-one measurement in order to minimise Eve's guessing probability and to maximise the conditional max-entropy of her measurement outcomes.

2. Optimality of rank-one measurements for H^*

We know from [13, Theorem 1, (i)] that the conditional entropy of the classical-quantum state formed by Alice's outcomes Z and Eve's post-measurement states is

$$H(Z|E) = D\big(\rho\,||\sum_i M_i\rho M_i\big)\,, \tag{B4}$$

where $\{M_i\}_i$ is Alice's projective measurement and $D(\rho||\sigma)$ is the quantum relative entropy between the states ρ and σ,

$$D(\rho||\sigma) = \text{tr}\left(\rho\big(\log_2\rho - \log_2\sigma\big)\right), \tag{B5}$$

and is defined when the support of ρ is contained within the support of σ. $D\big(\rho\,||\sum_i M_i\rho M_i\big)$ can be written as

$$D\big(\rho\,||\sum_i M_i\rho M_i\big) = -S(\rho) + \text{tr}\left(\rho\log_2\big(\sum_i M_i\rho M_i\big)\right). \tag{B6}$$

Since $\{M_i\}_i$ is projective, we have $\log_2\big(\sum_i M_i\rho M_i\big) = \sum_i \log_2\big(M_i\rho M_i\big)$ and $\Big(\log_2\big(M_i\rho M_i\big)\Big)M_j = \delta_{ij}\log_2\big(M_i\rho M_i\big)$, so the second term on the r.h.s. of (B6) is

$$\begin{aligned}\text{tr}\left(\rho\sum_i\log_2\big(M_i\rho M_i\big)\right) &= \text{tr}\left(\sum_j M_j\rho M_j\sum_i\log_2\big(M_i\rho M_i\big)\right)\\ &= -S\big(\sum_i M_i\rho M_i\big)\,,\end{aligned} \tag{B7}$$

and we recover Equation (16). To show that it is optimal for Alice to perform a rank-one measurement $\{M_{ij}^{\text{fine}} = |f_j^i\rangle\langle f_j^i|\}_{i,j}$ rather than a coarse-grained one $\{M_i^{\text{coarse}} = \sum_j \lambda_{ij}\,|f_j^i\rangle\langle f_j^i|\}_i$, denoting $\rho_{\text{fine}} = \sum_{i,j} M_{ij}\rho M_{ij}$ and $\rho_{\text{coarse}} = \sum_i M_i^{\text{coarse}}\rho M_i^{\text{coarse}}$, it is sufficient to show that

$$S(\rho_{\text{fine}}) \geq S(\rho_{\text{coarse}})\,. \tag{B8}$$

Note that the state ρ^{fine} is the average state after performing the measurement $\{M_{ij}^{\text{fine}}\}_{i,j}$ on ρ_{coarse}. Projective measurements cannot increase the von Neumann of a state (see, e.g., [31, Theorem 11.9]), so the inequality (B8) holds and it is optimal for Alice to perform a rank-one measurement to maximise the conditional entropy.

Appendix C: Technical steps in the proof of Theorem 1

1. Proof of Lemma 2 sketched in the main text

The main steps of the proof of Lemma 2 sketched in the text are given here in the form of two lemmas.

Lemma 3. *Since $\mathcal{M}$ is rank-one, the d states ρ_i can be taken pure, $\rho_i = q_i\,|\psi_i\rangle\langle\psi_i| := |\tilde{\psi}_i\rangle\langle\tilde{\psi}_i|$. Hence,*

$$P_{\text{guess}}(\rho,\mathcal{M}) = \max_{\{|\tilde{\psi}_i\rangle\}}\sum_i |\langle m_i|\tilde{\psi}_i\rangle|^2\,, \tag{C1}$$

with the constraint that $\rho = \sum_{i=1}^d |\tilde{\psi}_i\rangle\langle\tilde{\psi}_i|$.

Proof. We consider that Alice and Eve share a pure state $|\Phi\rangle_{AE}$ such that $\text{Tr}_E\,|\Phi\rangle\langle\Phi| = \rho_A := \rho$, and Eve performs a measurement $\mathcal{N} = \{N_i\}_i$. If assume that Eve performs her measurement before Alice, we find that she steers Alice's state ρ_i with probability $p_i = \langle\Phi|\mathbb{1}\otimes N_i|\Phi\rangle$ and we recover (7). But, because this is a no-signalling scenario, the order of their measurements does not matter, so we could equally think of Alice as measuring first. When Alice measures outcome i, since every M_i is rank-one, she steers the pure state $|\phi_i\rangle \propto \text{tr}_A(M_i\otimes\mathbb{1}\,|\Phi\rangle)$. Thus

$$\begin{aligned}P_{\text{guess}}(\rho,\mathcal{M}) \;&=\; \max_{\mathcal{N}}\sum_i \text{Tr}(\rho M_i)\,\langle\phi_i|N_i|\phi_i\rangle \\ &\text{s.t.}\;\; N_i\geq 0\,,\;\sum_i N_i = \mathbb{1}\,.\end{aligned} \tag{C2}$$

The SDP (C2) describes the optimal discrimination by Eve of an ensemble compatible with her reduced state. It was shown in [14] that the optimal measurement to distinguish an ensemble of pure states is made of d rank-one operators. Thus, Eve will also be performing a rank-one measurement on her system, and the states ρ_i steered on Alice's side will be pure. □

Lemma 4. *We will show that one can always write $|\tilde{\psi}\rangle_i = \sqrt{\rho}\,|i\rangle$, with $\{|i\rangle\}$ an orthonormal basis [Eq. (10) of the main text], and that Eq. (11) follows as a consequence, namely*

$$P_{\text{guess}}(\rho,\mathcal{M}) \;=\; \max_{\{|i\rangle\}}\sum_i |\,\langle m_i|\sqrt{\rho}|i\rangle|^2\,. \tag{C3}$$

Proof. Following [15], given two density matrices ρ' and ρ, there is a one-to-one map

$$|\tilde{\psi}_i'\rangle = \rho'^{1/2}\rho^{-1/2}|\tilde{\psi}_i\rangle\,,\;\; i=1,...,n \tag{C4}$$

between any two sub-normalised decompositions $\{|\tilde{\psi}_i'\rangle\}$ and $\{|\tilde{\psi}_i\rangle\}$ into the same number n of pure states. Choosing $\rho' = \mathbb{1}/d$, we obtain $|\tilde{\psi}_i\rangle = \rho^{1/2}|\tilde{i}\rangle$, where $\sum_i |\tilde{i}\rangle\langle\tilde{i}| = \mathbb{1}$. Since we start with a decomposition of ρ into $n=d$ pure states, and since any decomposition of $\mathbb{1}$ into d pure states defines an orthonormal basis, we find that

$$|\tilde{\psi}_i\rangle = \sqrt{\rho}\,|i\rangle \tag{C5}$$

for some $\{|i\rangle\}$ forming an orthonormal basis. Plugging this into (C1), we prove (C3). □

2. Optimisation (12) in the proof of Lemma 2

The optimisation (12) is a special case of the following: given two orthonormal bases $\{|i\rangle\}_{i=1,...,d}$ and $\{|m_i\rangle\}_{i=1,...,d}$ and an operator A ($\sqrt{\rho}$ in our case) satisfying $A = A^\dagger \geq 0$ such that

$$\langle m_i|A|m_i\rangle = \frac{1}{d}\operatorname{tr}(A) \text{ for all } i = 1, ..., d\,, \tag{C6}$$

we want to compute

$$P \;=\; \max_{\{|i\rangle\}} \sum_i |\,\langle i|A|m_i\rangle\,|^2\,. \tag{C7}$$

This expression is equivalent to the probability of correct state discrimination of an ensemble $\sigma_i = |\tilde{\gamma}_i\rangle\langle\tilde{\gamma}_i|$, with $|\tilde{\gamma}_i\rangle = A\,|m_i\rangle$, using the measurement $\{\Pi_i\}_{i=1,...,d}$. A measurement $\{\Pi_i\}_{i=1,...,d}$ is optimal to distinguish a set of states $\{\sigma_i\}_{i=1,...d}$ if and only if $Y \equiv \sum_i \sigma_i\Pi_i \geq \sigma_j$ for all $j = 1, ..., d$ [18–20]. In our case, $\{|i\rangle\}$ achieves the maximisation (C7) if and only if

$$\sum_i A\,|m_i\rangle\langle m_i|\,A\,|i\rangle\langle i| - A\,|m_j\rangle\langle m_j|\,A \geq 0 \text{ for all } j\,. \tag{C8}$$

Let us make the *guess* that, under the condition (C6), the optimal measurement is given by $|i\rangle = |m_i\rangle$. With this, conditions (C8) read

$$\frac{\operatorname{tr}(A)}{d}\,A \;-\; A\,|m_j\rangle\langle m_j|\,A := B_j \geq 0 \;\text{ for all } j\,. \tag{C9}$$

It is clear that $B_j = B_j^\dagger$. Moreover, using (C6) we see that

$$\langle\phi|B_j|\phi\rangle \;=\; \langle m_j|A|m_j\rangle\,\langle\phi|A|\phi\rangle - \langle\phi|A|m_j\rangle\,\langle m_j|A|\phi\rangle$$

for any vector $|\phi\rangle$. The r.h.s. is always non-negative when $A = A^\dagger \geq 0$, because of the Cauchy-Schartz inequality applied to the vectors $\sqrt{A}\,|\phi\rangle$ and $\sqrt{A}\,|m_j\rangle$. This proves that $B_j \geq 0$ for all j, which is the desired condition, vindicating our guess.

3. Necessary condition for P^*_{guess}

Here we show that satisfying condition (13) is necessary for Alice's measurement basis $\{|m_i\rangle\}$ to achieve the optimal guessing probability. From (11), choosing the orthonormal basis $\{|i\rangle\} = \{|m_i\rangle\}$, we have the bound

$$P_{\text{guess}}(\rho, \mathcal{M}) \geq \sum_i |\,\langle m_i|\sqrt{\rho}|m_i\rangle|^2\,. \tag{C10}$$

Define the set of real numbers $\varepsilon_i := \frac{\operatorname{tr}\sqrt{\rho}}{d} - \langle m_i|\sqrt{\rho}|m_i\rangle$. Note that the constraint

$$\operatorname{tr}\sqrt{\rho} = \sum_i \langle m_i|\sqrt{\rho}|m_i\rangle = \operatorname{tr}\sqrt{\rho} - \sum_i \varepsilon_i \tag{C11}$$

implies that $\sum_i \varepsilon_i = 0$. In terms of $\{\varepsilon_i\}$, the bound in (C10) is

$$\begin{aligned} P_{\text{guess}}(\rho, \mathcal{M}) &\geq \sum_i \left(\frac{\operatorname{tr}\sqrt{\rho}}{d} - \varepsilon_i\right)^2 \\ &= \frac{(\operatorname{tr}\sqrt{\rho})^2}{d} - \frac{2\operatorname{tr}\sqrt{\rho}}{d}\sum_i \varepsilon_i + \sum_i \varepsilon_i^2 \\ &= P^*_{\text{guess}}(\rho) + \sum_i \varepsilon_i^2\,, \end{aligned} \tag{C12}$$

so the optimal $P^*_{\text{guess}}(\rho)$ cannot be achieved if any $\varepsilon_i \neq 0$, i.e. if the condition

$$\langle m_i|\sqrt{\rho}|m_i\rangle = \frac{1}{d}\operatorname{tr}\sqrt{\rho} \;\text{ for all }\; i = 1, ..., d \tag{C13}$$

is not satisfied.

4. Eve's optimal measurement for $H^*_{\min}$

Denote the pure state shared by Alice and Eve by $|\Psi\rangle_{AE}$. Its Schmidt decomposition is

$$|\Psi\rangle_{AE} = \sum_{k=1}^{d} \sqrt{\lambda_k}\,|u_k\rangle_A\,|v_k\rangle_E\,, \tag{C14}$$

where $\rho = \sum_{k=1}^d \lambda_k\,|u_k\rangle\langle u_k|$ and $\{|v_k\rangle\}$ is an orthonormal basis in which Eve's reduced state in diagonal. We can assume without loss of generality that Eve's subsystem has the same dimension as Alice's subsystem, as Eve does not gain any advantage in discriminating Alice's states by holding a system of a higher dimension. Let Eve's local eigenbasis $\{|v_k\rangle\}$ be related to that of Alice by the unitary $|v_k\rangle = U\,|u_i\rangle$. Then we can write Eve's reduced state as $\sigma_E = U\rho U^\dagger$. After Alice performs the rank-one measurement in the basis $\{|m_i\rangle\}$, Eve receives the subnormalised pure states $|\tilde{\gamma}_i\rangle = \sum_{k=1}^d \sqrt{\lambda_k}\langle m_i|u_k\rangle\,|v_k\rangle$, which can also be represented as

$$|\tilde{\gamma}_i\rangle = \sum_{k=1}^{d} \sqrt{\lambda_k}\langle m_i|u_k\rangle\,U\,|u_k\rangle = U\sqrt{\rho}\,|m_i^*\rangle\,, \tag{C15}$$

where $|m_i^*\rangle$ is the complex conjugate of $|m_i\rangle$ in the eigenbasis of ρ. Eve should then use an optimal rank-one measurement to discriminate between the possible states, $|\tilde{\gamma}_i\rangle$, of her system. Note that, since Alice chooses an optimal measurement,

$$\frac{\operatorname{tr}\sqrt{\rho}}{d} = \langle m_i|\sqrt{\rho}|m_i\rangle = \langle m_i^*|\sqrt{\rho}|m_i^*\rangle\,.$$

In Appendix C 2, we showed that a measurement in an orthogonal basis $\{|n_i\rangle\}$ is optimal to discriminate an ensemble of states $\{|\tilde{\gamma}_i\rangle = A\,|n_i\rangle\}$ if the condition $\langle n_i|A|n_i\rangle = \frac{1}{d}\operatorname{tr} A$ holds for all i, where $A \geq 0$. Here Eve receives the states $|\tilde{\gamma}_i\rangle = U\sqrt{\rho}\,|m_i^*\rangle = A\,|n_i\rangle$, where $A := U\sqrt{\rho}\,U^\dagger \geq 0$ and $|n_i\rangle := U\,|m_i^*\rangle$. We have $\frac{1}{d}\operatorname{tr} A = \frac{1}{d}\operatorname{tr}\left(U\sqrt{\rho}\,U^\dagger\right) = \frac{1}{d}\operatorname{tr}\sqrt{\rho}$ and

$$\langle n_i|A|n_i\rangle = \langle m_i^*|\sqrt{\rho}|m_i^*\rangle = \frac{1}{d}\operatorname{tr}\sqrt{\rho} = \frac{1}{d}\operatorname{tr} A\,, \tag{C16}$$

so it is optimal for Eve to measure in the basis $|n_i\rangle = U\,|m_i^*\rangle$. In the case where ρ is full-rank, we have

$$U\,|m_i^*\rangle = U\rho^{-1/2}U^\dagger\,|\tilde{\gamma}_i\rangle = \sigma_E^{-1/2}\,|\tilde{\gamma}_i\rangle\,, \tag{C17}$$

so Eve is performing a 'pretty good' measurement.

Appendix D: Alternative proof of Theorem 1

1. Proof of the bound (1) on P_{guess}

Here we provide an alternative derivation of the lower bound (1) for Eve's optimal guessing probability. We make use of the min- and max-entropies defined in Appendix A and of the following lemma, which is a slight variation [32] of [28, Lemma 3.1.13]:

Lemma 5. *Let $\{|x\rangle\}_x$ be an orthonormal basis on $\mathcal{H}_X$, let $\{|\psi^x\rangle\}_x$ be a family of unnormalized vectors on $\mathcal{H}_A \otimes \mathcal{H}_E$, and define*

$$\rho_{AE} := |\psi\rangle\langle\psi| \;\text{ with }\; |\psi\rangle = \sum_x |\psi_x\rangle\,,$$

$$\tilde{\rho}_{XAE} := \sum_x |x\rangle\langle x| \otimes |\psi_x\rangle\langle\psi_x|\,.$$

Then,

$$H_{\min}(A|E)_{\rho_{AE}} \geq H_{\min}(A|E)_{\tilde{\rho}_{AE}} - H_{\max}(X)\,. \tag{D1}$$

Consider a pure bipartite state $\rho_{AE} = |\psi\rangle\,\langle\psi|$. Given any orthogonal basis $\{|u_x\rangle\}_x$ for $\mathcal{H}_A$, some orthogonal basis $\{|v_y\rangle\}_y$ of $\mathcal{H}_E$ can always be found such that the state vector $|\psi\rangle$ can be written as

$$|\psi\rangle = \sum_{x,y}\alpha_{xy}\,|u_x\rangle\,|v_y\rangle = \sum_x |u_x\rangle\sum_y \alpha_{xy}\,|v_j\rangle := \sum_x |\psi^x\rangle\,. \tag{D2}$$

If the rank-one projective measurement $\{|u_x\rangle\langle u_x|\}_x$ is performed on A and the outcome x is obtained, the (unnormalized) post-measurement state of AE is

$$\Big(\,|u_x\rangle\langle u_x| \otimes \mathbb{1}\Big)\rho^{AE}\Big(\,|u_x\rangle\langle u_x| \otimes \mathbb{1}\Big) = |\psi^x\rangle\langle\psi^x|\,.$$

Therefore, the cq-state representing the correlations between the outcomes and the post-measurement states on $\mathcal{H}_A \otimes \mathcal{H}_E$ is, precisely as in Lemma 5,

$$\tilde{\rho}_{AEX} = \sum_x |\psi^x\rangle\langle\psi^x| \otimes |x\rangle\langle x|\;..$$

To prove our target bound (1), we let ρ_{AE} be the initial pure state shared by Alice and Eve and $\mathcal{M} = \{|u_x\rangle\langle u_x|\}_x$ be Alice's rank-one projective measurement. Noting that $H_{\max}(X) = \log_2 \operatorname{rank}(\tilde{\rho}_X) \leq \log_2 d$ and that for any pure state ρ_{AE}

$$H_{\min}(A|E) = -2\log_2\sqrt{\rho_A}\,, \tag{D3}$$

returning to our usual notation with $\rho_A = \rho$ and $d_A = d$, we get from Lemma 5 the following upper bound on $H_{\min}(A|E)_{\tilde{\rho}_{AE}}$,

$$\log_2 d - 2\log_2\operatorname{tr}\sqrt{\rho} \geq H_{\min}(A|E)_{\tilde{\rho}_{AE}}\,. \tag{D4}$$

Finally, notice that, since Alice's post-measurement states form an orthonormal basis, the states $\tilde{\rho}_{XE}$ and $\tilde{\rho}_{AE}$ are related by a unitary on the first subsystem. Therefore, using twice the data processing inequality for conditional min-entropies [28], one can replace $H_{\min}(A|E)_{\tilde{\rho}_{AE}}$ with $H_{\min}(X|E)_{\tilde{\rho}_{XE}}$ in (D4) and, together with (A6), obtain

$$P_{\text{guess}}(\rho,\mathcal{M}) \geq \frac{1}{d}\Big(\operatorname{tr}\sqrt{\rho}\Big)^2\,. \tag{D5}$$

2. Proof that the bound (1) can be reached

Here, we provide an alternative proof using semidefinite programming that the lower bound (1) for Eve's optimal guessing probability can be reached when Alice performs measurements satisfying (13). We consider the maximisation problem (7) for the guessing probability and reduce our study to the set of projective rank-one measurements $M_i = |m_i\rangle\langle m_i|$ in dimension d that satisfy $\operatorname{tr}\left(\,|m_i\rangle\langle m_i|\,\sqrt{\rho}\right) = \frac{1}{d}\operatorname{tr}\sqrt{\rho}$ for all $|m_i\rangle$ (the set is certainly not empty, since all measurements in a basis that is unbiased with the eigenbasis of ρ satisfy this condition). From hereon, we assume without loss of generality that ρ is full-rank, as we note that only the projection of the measurement elements M_i onto the support of ρ plays a role in (7). More precisely, defining Π_ρ as the orthogonal projector onto the support of ρ and $M_i' := \Pi_\rho M_i \Pi_\rho$, we have

$$\begin{aligned}\max_{\{\rho_i\}}\sum_i \operatorname{tr}(M_i\rho_i) &= \max_{\{\rho_i\}}\sum_i \operatorname{tr}(M_i\,\Pi_\rho\rho_i\Pi_\rho)\\ &= \max_{\{\rho_i\}}\sum_i \operatorname{tr}(M_i'\rho_i)\,.\end{aligned} \tag{D6}$$

Note too that

$$\operatorname{tr}\left(M_i\sqrt{\rho}\right) = \operatorname{tr}\left(M_i'\sqrt{\rho}\right) = \frac{1}{d}\operatorname{tr}\sqrt{\rho}\,. \tag{D7}$$

Then, in the case where ρ is not rank-one, we can consider the problem projected onto the support of ρ, using the rank-one (not necessarily projective) measurement with elements $M_i' = |m_i'\rangle\langle m_i'|$ in (7), where M_i' is in dimension

$r = \mathrm{rank}(\rho)$, has d outcomes and satisfies $\mathrm{tr}\left(M_i'\sqrt{\rho}\right) = \frac{1}{d}\,\mathrm{tr}\,\sqrt{\rho}$.

Now with the full-rank ρ assumption, we note that, since (7) is a semidefinite programming problem, we can define its corresponding minimisation (or dual) problem, which is given by

$$\beta(\rho, \mathcal{M}) = \min_X \mathrm{tr}\left(X\rho\right), \qquad \text{s.t. } X \geq M_i\,. \tag{D8}$$

The set of states $\{\rho_i = \rho/d\}$ and the matrix $X = 2\mathbb{1}$ define strongly feasible points (i.e. points that satisfy the necessary constraints with strict inequalities) on the dual and primal problems respectively, so (7) and (D8) both return $P_{\text{guess}}(\rho, \mathcal{M})$. For further details on semidefinite programming problems, see, for example, [33]. Given the measurement $\mathcal{M}$, we can use the primal and dual problems to set upper and lower bounds respectively on Eve's optimal guessing probability,

$$\sum_i \mathrm{tr}\left(\rho_i M_i\right) \leq P_{\text{guess}}(\rho, \mathcal{M}) \leq \mathrm{tr}\left(X\rho\right), \tag{D9}$$

where $\{\rho_i\}$ is any set of subnormalised states satisfying $\sum_i \rho_i = \rho$ and X is any positive-semidefinite matrix satisfying $X - M_i \geq 0$ for all M_i. The set of states $\rho_i = \sqrt{\rho}\,|m_i\rangle\langle m_i|\,\sqrt{\rho}$ recovers the lower bound (1). We now show that the matrix $X = \frac{\mathrm{tr}\sqrt{\rho}}{d}\rho^{-\frac{1}{2}}$ satisfies

$$X - |m_i\rangle\langle m_i| = \frac{\mathrm{tr}\,\sqrt{\rho}}{d}\rho^{-\frac{1}{2}} - |m_i\rangle\langle m_i| \geq 0 \qquad \forall\, i\,. \tag{D10}$$

We can define any vector in dimension d as $\sqrt{\rho}\,|\phi\rangle$ for some $|\phi\rangle$. To prove (D10), it suffices to show that

$$\langle m_i|\sqrt{\rho}|m_i\rangle\ \langle\phi|\sqrt{\rho}|\phi\rangle - \langle\phi|\sqrt{\rho}|m_i\rangle\ \langle m_i|\sqrt{\rho}|\phi\rangle \geq 0\,. \tag{D11}$$

In analogy with the proof in C 2, we see that (D11) is true by applying the Cauchy–Schwarz inequality on the vectors $\rho^{\frac{1}{4}}\,|m_i\rangle$ and $\rho^{\frac{1}{4}}\,|\phi\rangle$.

Using $X = \frac{\mathrm{tr}\sqrt{\rho}}{d}\rho^{-\frac{1}{2}}$ in (D9), we find

$$\frac{1}{d}\Big(\mathrm{tr}\,\sqrt{\rho}\Big)^2 \leq P_{\text{guess}}(\rho, \mathcal{M}) \leq \frac{1}{d}\Big(\mathrm{tr}\,\sqrt{\rho}\Big)^2. \tag{D12}$$

This shows that the bound (1) can be saturated by measurements satisfying (13), so we have that the optimal guessing probability for Alice given the state ρ is

$$P^*_{\text{guess}}(\rho) = \frac{1}{d}\Big(\mathrm{tr}\,\sqrt{\rho}\Big)^2. \tag{D13}$$

Appendix E: Technical details in the proof of Theorem 3

We denote the Schmidt decomposition of the pure state $|\Psi\rangle_{AE}$ shared by Alice and Eve by

$$|\Psi\rangle_{AE} = \sum_{k=1}^{d} \sqrt{\lambda_k}\,|u_k\rangle_A\,|v_k\rangle_E\,, \tag{E1}$$

where $\rho = \sum_{k=1}^{d}\lambda_k\,|u_k\rangle\langle u_k|$ and $\{|v_k\rangle\}$ is an orthonormal basis in which Eve's reduced state in diagonal. When Alice measures in an orthonormal basis $\{|m_x\rangle\}$ with dimension d, we can write Eve's post-measurement states $\{|\psi^E_x\rangle\}$ as

$$|\psi^E_x\rangle = \frac{1}{\sqrt{p_x}}\langle m_x|\Psi_{AE}\rangle\,, \quad p_x = \langle m_x|\rho|m_x\rangle\,. \tag{E2}$$

Denoting Eve's average state post-measurement as $\tilde{\rho}_E$, i.e.

$$\tilde{\rho}_E = \sum_x p_x\,|\psi^E_x\rangle\langle\psi^E_x|\,, \tag{E3}$$

we note that

$$\tilde{\rho}_E = \sum_x \langle m_x|\Psi_{AE}\rangle\langle\Psi_{AE}|m_x\rangle = \mathrm{tr}_A\,|\Psi_{AE}\rangle\langle\Psi_{AE}| := \rho_E\,, \tag{E4}$$

where ρ_E is Eve's local state before the measurement. The following expression for $H_{\max}(X|E)$ in the special case of classical-quantum states is given in [3],

$$H_{\max}(A|E) = \log_2 p_{\text{secr}}\,, \tag{E5}$$

where

$$p_{\text{secr}} = \max_\sigma \left(\sum_x \sqrt{p_x} F\left(\rho^E_x, \sigma\right)\right)^2, \tag{E6}$$

where σ is a quantum state ($\sigma \geq 0$ and $\mathrm{tr}\,\sigma = 1$) and

$$F(\rho, \sigma) = \mathrm{tr}\,\sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}} \tag{E7}$$

(note that $F(\rho, \sigma)$ is symmetric in ρ and σ). Since, in our case, the states ρ^E_x are pure, the expression (E6) reduces to

$$p_{\text{secr}} = \max_\sigma \left(\sum_x \sqrt{p_x\langle\psi^E_x|\sigma|\psi^E_x\rangle}\right)^2 \tag{E8}$$

$$= \max_\sigma \left(\sum_x \sqrt{p_x\,\mathrm{tr}\left(\sigma\,|\psi^E_x\rangle\langle\psi^E_x|\right)}\right)^2. \tag{E9}$$

We define by σ^* the state σ that achieves the maximisation in (E6). By the Cauchy-Schwartz inequality, we have

$$p_{\text{secr}} = \left(\sum_x \sqrt{p_x\,\mathrm{tr}\left(\sigma^*\,|\psi^E_x\rangle\langle\psi^E_x|\right)}\right)^2 \tag{E10}$$

$$\leq d\sum_x p_x\,\mathrm{tr}\left(\sigma^*\,|\psi^E_x\rangle\langle\psi^E_x|\right) = d\,\mathrm{tr}\left(\sigma^*\rho_E\right). \tag{E11}$$

The inequality is saturated if and only if all of the terms inside the sum are identical,

$$p_x\,\mathrm{tr}\left(\sigma^*\,|\psi^E_x\rangle\langle\psi^E_x|\right) = \frac{\mathrm{tr}\left(\sigma^*\rho_E\right)}{d} \quad \text{for all } \ x = 1, ..., d\,. \tag{E12}$$

We find a second inequality for p_{secr} by noting that

$$\text{tr}\left(\sigma^*\rho_E\right) \leq \max_{\sigma} \text{tr}\left(\sigma\rho_E\right) = \lambda_{\max}(\rho_E)\,, \tag{E13}$$

where the second term is a known SDP that returns the largest eigenvalue of ρ_E. In the case where the largest eigenvalue of ρ_E is non-degenerate, the maximisation is achieved only by $\sigma = |v_{\max}\rangle\langle v_{\max}|$, where $|v_{\max}\rangle$ is the vector from Eve's local basis $\{|v_i\rangle\}$ corresponding to the largest eigenvalue. Restricting for now to the case where the largest eigenvalue is non-degenerate, we see that the bound (E13) is reached if and only if $\sigma^* = |v_{\max}\rangle\langle v_{\max}|$ and (E12) is satisfied. We are free to combine these conditions such that the necessary and sufficient conditions for the measurement basis $\{|m_x\rangle\}$ to achieve the bound (E13) are, from (E12),

$$|\langle m_x|u_{\max}\rangle|^2 = \frac{1}{d} \text{ for all } x = 1, ..., d\,, \tag{E14}$$

where $|u_{\max}\rangle$ is the vector from Alice's local basis $\{|u_i\rangle\}$ corresponding to the largest eigenvalue. In the case where the largest eigenvalue of ρ is degenerate, denote by $\{|{v_{\max}}^{(i)}\rangle\}$ the set of vectors in Eve's local basis corresponding to the largest eigenvalue. The optimal σ in (E13) is now of the form

$$\sigma = \sum_i \gamma_i \left|{v_{\max}}^{(i)}\right\rangle\!\left\langle{v_{\max}}^{(i)}\right|\,, \quad \gamma_i \geq 0\,, \quad \sum_i \gamma_i = 1\,. \tag{E15}$$

Now the necessary and sufficient conditions for $\{|m_x\rangle\}$ to achieve the bound are

$$\sum_i \gamma_i \left|\langle m_x|{u_{\max}}^{(i)}\rangle\right|^2 = \frac{1}{d} \text{ for all } x = 1, ..., d\,, \tag{E16}$$

where $\{\gamma_i\}$ is any set of non-negative numbers summing to 1 and $\{|{u_{\max}}^{(i)}\rangle\}$ is the set of vectors in Alice's local basis corresponding to the maximum eigenvalue.

Note, finally, that $\lambda_{\max}(\rho_E) = \lambda_{\max}(\rho)$, so in terms of Alice's state ρ we have

$${p_{\text{secr}}}^* = d\lambda_{\max}(\rho) \tag{E17}$$

and

$${H_{\max}}^*(A|E) = \log_2 d + \log_2 \lambda_{\max}(\rho)\,. \tag{E18}$$

Appendix F: Additional details from Section V

1. Parameters for qutrit measurement (24)

For convenience, we restate the measurement (24) in the following. Consider the measurement basis $\{M_i = |m_i\rangle\langle m_i|\}_{i=1,2,3}$, with

$$|m_1\rangle = \sqrt{\frac{1+a}{3}}\,|1\rangle + \sqrt{\frac{1+b}{3}}\,|2\rangle + \sqrt{\frac{1+c}{3}}\,|3\rangle\,,$$

$$|m_2\rangle = \sqrt{\frac{1+a}{3}}\,e^{i\theta_1}\,|1\rangle + \sqrt{\frac{1+b}{3}}\,|2\rangle + \sqrt{\frac{1+c}{3}}\,e^{i\theta_2}\,|3\rangle\,, \tag{F1}$$

and $|m_3\rangle$ defined by the normalization condition $\sum_i M_i = \mathbb{1}$, where $a = -(\gamma_2 - \gamma_3)k$, $b = (\gamma_1 - \gamma_3)k$, $c = -(\gamma_1 - \gamma_2)k$, $k \in \mathbb{R}$ and each $\gamma_i \geq 0$ with $\gamma_1 \geq \gamma_2 \geq \gamma_3$. To ensure that the square root terms in the coefficients of $|m_1\rangle$ and $|m_2\rangle$ are well-defined, we need that $1 + x \geq 0$, $x \in \{a, b, c\}$, which imposes the following constraint on k,

$$-\frac{1}{\gamma_1 - \gamma_3} \leq k \leq \frac{1}{\max\{\gamma_1 - \gamma_2, \gamma_2 - \gamma_3\}}\,. \tag{F2}$$

Furthermore, imposing $\langle m_1|m_2\rangle = 0$ sets the following restriction,

$$\frac{1+a}{3}e^{i\theta_1} + \frac{1+b}{3} + \frac{1+c}{3}e^{i\theta_2} = 0\,. \tag{F3}$$

Considering each of the terms in the sum as vectors in the complex plane, the sum $\frac{1+a}{3}e^{i\theta_1} + \frac{1+b}{3}$ can take any absolute value in the range $\frac{1}{3}|b-a|$ to $\frac{1}{3}|2+a+b|$ by taking an appropriate choice of θ_1. We then see that in order to satisfy (F3), $\frac{1+c}{3}$ must fall in this range, i.e.

$$|b - a| \leq 1 + c \leq 2 + a + b\,. \tag{F4}$$

This gives the following restriction on k,

$$-\frac{1}{2}\frac{1}{\gamma_1 - \gamma_3} \leq k \leq \frac{1}{2}\frac{1}{\gamma_1 - \gamma_3}\,, \tag{F5}$$

which is a tighter bound than the constraint (F2). Notice that the tightest constraint is $-\frac{1}{2} \leq k \leq \frac{1}{2}$ and is obtained for pure states.

For $k \neq 0$, this measurement basis is unbiased to the eigenbasis of ρ if and only if $a = b = c = 0$ i.e. when ρ is the maximally mixed state. When we set $\{\gamma_i\} = \{\sqrt{\lambda_i}\}$, it is easy to show that $\langle m_i|\sqrt{\rho}|m_i\rangle = \frac{1}{3}\,\text{tr}\,\sqrt{\rho}$ for $i = 1, 2, 3$, so condition (13) for maximal $H_{\min}$ is satisfied. However, this measurement does not satisfy the necessary and sufficient condition (18) to maximise the conditional entropy, except in the special case where

$$\lambda_1(\sqrt{\lambda_2} - \sqrt{\lambda_3}) - \lambda_2(\sqrt{\lambda_1} + \sqrt{\lambda_3}) - \lambda_3(\sqrt{\lambda_1} - \sqrt{\lambda_2}) = 0\,. \tag{F6}$$

Moreover, we can also consider this qutrit measurement taking $\{\gamma_i\} = \{\lambda_i\}$. Similarly, it is straightforward to see that this measurement satisfies (18) and thus achieves the maximal conditional entropy H^*. It does not, however, satisfy the necessary and sufficient condition (13) to achieve $H^*_{\min}$ except in the special case where

$$\sqrt{\lambda_1}(\lambda_2 - \lambda_3) - \sqrt{\lambda_2}(\lambda_1 - \lambda_3) + \sqrt{\lambda_3}(\lambda_1 - \lambda_2) = 0\,. \tag{F7}$$

Finally, when $\lambda_1 > \lambda_2$ (i.e. the largest eigenvalue of ρ is not degenerate), the condition (23) for the measurement (24) to produce maximal $H_{\max}$ is satisfied if and only if $a = 0$, so $\gamma_2 = \gamma_3$. This condition is, in general, inequivalent to the conditions for maximal $H_{\min}$ and H.

2. A two-qubit basis that has no unbiased product basis

We will show that there is no orthonormal basis of two-qubit product states that is unbiased to the basis:

$$\begin{aligned}
|\psi_1\rangle &= |00\rangle \\
|\psi_2\rangle &= \frac{1}{\sqrt{3}}(|01\rangle + |10\rangle + |11\rangle) \\
|\psi_3\rangle &= \frac{1}{\sqrt{3}}(|01\rangle + \omega\,|10\rangle + \omega^2\,|11\rangle) \\
|\psi_4\rangle &= \frac{1}{\sqrt{3}}(|01\rangle + \omega^2\,|10\rangle + \omega\,|11\rangle),
\end{aligned} \tag{F8}$$

where $\omega = e^{2\pi i/3}$. Firstly, we note that any orthonormal two-qubit product basis can be expressed as either

(i) $\left\{|a\rangle\,|A\rangle\,, |a\rangle\,\left|A^\perp\right\rangle, \left|a^\perp\right\rangle|B\rangle\,, \left|a^\perp\right\rangle\left|B^\perp\right\rangle\right\}$ or

(ii) $\left\{|a\rangle\,|A\rangle\,, \left|a^\perp\right\rangle|A\rangle\,, |b\rangle\,\left|A^\perp\right\rangle, \left|b^\perp\right\rangle\left|A^\perp\right\rangle\right\}$

for some single qubit states $|a\rangle\,, |A\rangle\,, |b\rangle\,, |B\rangle$, where $\perp$ denotes the unique orthogonal state to a given qubit state. Consider Case (i). If this basis is unbiased to the vectors (F8), we have

$$|\langle 00|aA\rangle|^2 = \left|\langle 00\left|aA^\perp\right\rangle\right|^2 = \frac{1}{4}. \tag{F9}$$

Letting $|A\rangle = A_0\,|0\rangle + A_1\,|1\rangle$, etc, with $A_0 \in \mathbb{R}$, $A_1 \in \mathbb{C}$, it follows that $A_0 = A_0^\perp = 1/2a_0$. In order for $|A\rangle$ and $\left|A^\perp\right\rangle$ to be orthogonal and have the same coefficient of $|0\rangle$, both states must lie on the XY-plane of the Bloch sphere, i.e. $|A\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{iy}\,|1\rangle)$ and $\left|A^\perp\right\rangle = \frac{1}{\sqrt{2}}(|0\rangle - e^{iy}\,|1\rangle)$ for some $0 \le y < 2\pi$. The relation between a_0 and A_0 implies that $|a\rangle$ also lies on this plane and we can take $|a\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{ix}\,|1\rangle)$ for some $0 \le x < 2\pi$. Similar reasoning shows that all the states $|a\rangle\,, |A\rangle\,, |b\rangle\,, |B\rangle$ must lie in the XY-plane, in both Cases (i) and (ii).

We now show the states $|aA\rangle$ and $\left|aA^\perp\right\rangle$ cannot be mutually unbiased to both $|\psi_2\rangle$ and $|\psi_3\rangle$. Firstly, the equalities $|\langle\psi_2|aA\rangle|^2 = \left|\langle\psi_2\left|aA^\perp\right\rangle\right|^2 = 1/4$ give

$$\begin{aligned}
&\cos(x) + \cos(y) + \cos(x-y) = \\
&\cos(x) - \cos(y) - \cos(x-y) = 0\,.
\end{aligned} \tag{F10}$$

Thus, we find $\cos(x) = 0$, giving four possibilities for the values of (x, y): either $(\pi/2, 3\pi/4)$, $(\pi/2, 7\pi/4)$, $(3\pi/2, \pi/4)$, or $(3\pi/2, 5\pi/4)$. Substitution shows that none of these pairs give $|\langle\psi_3|aA\rangle|^2 = 1/4$. A similar argument shows that bases of the form in Case (ii) also cannot be mutually unbiased to the vectors (F8).