Cost of Simulating Entanglement in Steering Scenarios

Yujie Zhang[1], Jiaxuan Zhang[1], and Eric Chitambar[2]

[1]Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[2] Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

arXiv:2302.09060v4 [quant-ph] 27 Oct 2025

Quantum entanglement is a fundamental feature of quantum mechanics, yet certain entangled states that are unsteerable can be classically simulated in steering scenarios, making them unable to exhibit quantum steering. Despite their significance, a systematic comparison of such entangled states has not been explored. In this work, we quantify the resource content of unsteerable quantum states in terms of the amount of shared randomness required to simulate the assemblages they generate in the steering scenario. We rigorously demonstrate that the simulation cost is unbounded even for certain unsteerable two-qubit states. Moreover, the simulation cost of entangled two-qubit states is always strictly larger than that for any separable state.

A significant portion of our results rests on the relationship between the simulation cost of two-qubit Werner states and that of noisy spin measurements. Using noisy spin measurements as our central example, we also investigate the minimum number of outcomes a parent measurement requires to simulate a given set of compatible measurements. Although certain continuous measurement families admit a finite-outcome parent measurement, we identify scenarios where the simulation cost is unbounded. Our results establish previously unknown lower bounds and upper bounds on the shared randomness simulation cost, supported by connections between the simulation cost of noisy spin measurements and various geometric inequalities, including ones from the zonotope approximation problem in Banach space theory.

Yujie Zhang: yujie4physics@gmail.com
Eric Chitambar: echitamb@illinois.edu

1 Introduction

In quantum information theory, entanglement is recognized as a resource because it enables certain dynamical processes (i.e., channels) that would otherwise be impossible, or at least more costly, to perform. For instance, shared entanglement can be used in conjunction with some noisy classical channel to increase its zero-error transmission [1, 2]. Another dramatic effect is quantum teleportation, in which a classical channel is combined with shared entanglement to generate a perfect quantum channel [3]. In these examples, entanglement is used for the purpose of channel coding as it helps transform a noisy channel into a noiseless one. The converse task, analogous to the "Reverse Shannon" problem [4, 5], considers how much noiseless resource is needed to simulate a noisy entanglement-enhanced setup. Here the goal is not communication but rather simulation. From this perspective, one can quantify the resource value of an entangled state by how much classical resources is needed to simulate the dynamics that it generates.

Beyond the standard entanglement-assisted paradigm, there are various other ways in which a bipartite entangled state can be used to build a channel [6], and each of them will typically have a different simulation cost. In this paper, we focus on the type of channels that emerge in quantum steering setups [7]. These are classical-to-classical-quantum (c-to-cq) channels that are generated when Alice measures her half of the state ρ_{AB} in some chosen manner, thereby preparing Bob's system into one of many possible post-measurement states. From a fundamental perspective, these channels are vital to the study of quantum nonlocality [8]. But more practically they have important applications in semi-device-independent entanglement witnessing [9] and operational games [10].

As shown in Fig. 1, each input x to the chan-

nel corresponds to a local measurement choice for Alice, which we collectively represent by a family of positive operator-valued measures (POVMs) $\{M_{a|x}\}_{a,x}$. She receives a classical output a while Bob receives the (unnormalized) quantum output

$$\sigma_{a|x} = \mathrm{Tr}_A[(M_{a|x} \otimes \mathbb{I})\rho_{AB}] \tag{1}$$

with probability $p(a|x) = \mathrm{Tr}[\sigma_{a|x}]$. It is convenient to denote a c-to-cq channel by the set of labeled quantum outputs $\mathcal{A} = \{\sigma_{a|x}\}$, called a state assemblage [7], and we write $\rho_{AB} \mapsto \mathcal{A}$ when Eq. (1) holds.

To quantify the classical simulation cost of entanglement, many previous works adopt the feed-forward classical communication cost (measured in bits) as their figure of merit. Notable examples include the trade-off between ebits and cbits in quantum-channel simulation [11], the communication cost of reproducing Bell nonlocal correlations [12, 13] and the communication cost of reproducing quantum steering [14]. However, any entangled state admitting a local hidden variable (LHV) model or a local hidden state (LHS) model can already be simulated with zero forward communication in its corresponding scenario. This causes communication-based metrics to classify such states as "free", even though Bell-local states or unsteerable states that are entangled remain useful in many quantum applications[15, 16]. Moreover, implementing an LHV model or an LHS model generally demands a nontrivial amount of shared randomness, whose distribution itself, in fact, must invoke some sort of prior communication. Therefore, a complete quantification of classical simulation cost should account not only for the communication cost but also for the shared randomness cost needed to reproduce the desired correlations.

In this paper, we study the simulation of quantum correlations using only shared randomness—with no forward communication. While shared randomness alone cannot produce Bell nonlocal correlations or demonstrate quantum steering, it can reproduce the measurement statistics of a broad class of entangled states [17, 18]. There is currently no general method to quantify this share randomness cost, particularly to obtain tight lower bounds on the required amount of shared randomness. In this work, we focus on simulation cost problem in the steering scenario. Although less commonly considered than the Bell scenario, its comparatively simple characterization lets us uncover several key features of the shared randomness simulation cost, including its unboundedness and its operational relevance for separating separable from entangled states.

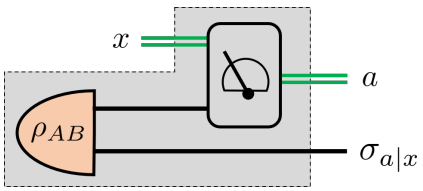

Figure 1: The bipartite state ρ_{AB} induces a c-to-cq channel represented by the assemblage $\mathcal{A} = \{\sigma_{a|x}\}$.

The benefit of a systematic study of shared randomness simulation cost is twofold. On the one hand, unbounded shared randomness has long been invoked in LHV or LHS models for steering or Bell nonlocality [19, 8] and is often believed to be necessary. Yet, a rigorous proof of the necessity of such unbounded shared randomness has been lacking; our paper provides the first concrete demonstration. On the other hand, our findings provide a systematic way to quantify simulation cost across different entangled states and different sets of compatible measurements. This reveals complexity-based advantages of quantum resources and offers a complementary operational understanding of how quantum resources can be harnessed beyond existing paradigms. For instance, implementing incompatible measurements often requires additional quantum memory [20]. Although compatible measurements can always be implemented with classical memory, those with high simulation cost may still serve as valuable resources because their classical simulation requires high-dimensional—or even unbounded—classical memory [21].

2 Preliminary

For a given quantum state ρ_{AB}, we wish to quantify the amount of classical resources needed to simulate each assemblage $\mathcal{A}$ for which $\rho_{AB} \mapsto \mathcal{A}$. For this simulation task, it is natural to quantify the classical resource in terms of how much shared randomness Alice and Bob need to reproduce $\mathcal{A}$ using a local hidden state (LHS) model. An LHS model for assemblage $\mathcal{A} = \{\sigma_{a|x}\}$, should it exist, is given by a shared random variable λ with

distribution $p(\lambda) > 0$, a family of quantum states $\{\rho_\lambda\}$ for Bob, and a classical channel $p(a|x,\lambda)$ for Alice such that

$$\sigma_{a|x} = \sum_{\lambda\in\Lambda} p(a|x,\lambda)p(\lambda)\rho_\lambda \qquad \sigma_{a|x} \in \mathcal{A},\ \forall a,x. \tag{2}$$

Deploying this model requires Alice and Bob to have shared randomness of size $|\Lambda| \in [1,+\infty]$, and we define $\gamma(\mathcal{A}) := \inf|\Lambda|$ as the smallest possible cardinality of Λ for which Eq. (2) holds, and we refer to this as the *simulation cost* of state assemblage $\mathcal{A}$. For LHS models using a continuous shared variable, the sum in Eq. (2) is replaced with an integral, and thus the simulation cost will exist but could be unbounded.

If an assemblage $\mathcal{A}$ does not admit an LHS model, then the resource transformation $\rho_{AB} \mapsto \mathcal{A}$ constitutes the celebrated effect of quantum steering [8], specifically in the direction from Alice to Bob. Although entanglement is a necessary condition for steering, not all entangled states are steerable. For such states, the next natural question then is to look *beyond steering* and ask how much shared randomness is needed to simulate all the assemblages $\mathcal{A}$ that the state can generate. For an unsteerable ρ_{AB}, we define

$$\gamma(\rho_{AB}) := \sup_{\mathcal{A}} \gamma(\mathcal{A}) \tag{3}$$

as the highest simulation cost to simulate any $\mathcal{A}$ for which $\rho_{AB} \mapsto \mathcal{A}$.

To be fully comprehensive, one should also consider the simulation cost $\gamma(\mathcal{B})$ for any assemblage $\mathcal{B}$ generated from ρ_{AB} by locally measuring on Bob's side since quantum steering is asymmetric between the two parties [22, 23]. However, the specific examples we study in this paper involve symmetric quantum states (under the exchange of parties). Consequently, our results remain general by just considering assemblages generated by local measurements on Alice's side. As an interesting special case that will be discussed in the paper, when Alice and Bob hold a qubit-qubit state, we also define

$$\gamma^p(\rho_{AB}) := \sup_{\mathcal{A}^p} \gamma(\mathcal{A}^p) \tag{4}$$

as the simulation cost of simulating any $\mathcal{A}^p$ generated by measurements that are confined to the x-z plane, i.e., $\{M_{a|x} := \alpha_{a|x}(\mathbb{I} + \hat{n}_{a|x}\cdot\vec{\sigma})\}$ in Eq. (1) with Bloch vectors $\{\hat{n}_{a|x}\}$ lying in the same plane [24].

One important question we consider is whether the simulation cost $\gamma(\rho_{AB})$ satisfies some general dimensionality bound. In fact, we always have $\gamma(\rho_{AB}) \leq (\dim(A))^2(\dim(B))^2$ for every separable state ρ_{AB} due to the existence of minimal separable decompositions [25]. Does this bound also hold for unsteerable entangled states? This question was first raised and left unanswered by Brunner et al. in a study that established an upper bound on the simulation cost by explicitly constructing local hidden state models [18].

In this work, we answer the question negatively: the simulation cost is unbounded for some unsteerable entangled states. We do so by deriving lower bounds on simulation cost that have not been previously discussed. Specifically, we prove that $\gamma(\rho_{AB})$ can be unbounded by demonstrating that its lower bound diverges (tends to infinity), even for certain two-qubit states.

To develop the analysis, we leverage the close link between steering and measurement incompatibility [26, 27, 28]. We now briefly review the notion of measurement compatibility.

A family of measurements $\{M_{a|x}\}_{a,x}$ is called compatible or jointly measurable if each measurement in the family can be simulated by a single 'parent' POVM $\{\Pi_\lambda\}_{\lambda\in\Lambda}$ with index set Λ as:

$$M_{a|x} = \sum_{\lambda\in\Lambda} p(a|x,\lambda)\Pi_\lambda \qquad \forall a,x \tag{5}$$

for some classical channel $p(a|x,\lambda)$ [29, 30]. In the simple case of quantum observables, compatibility amounts to pairwise commutativity, but the latter is not necessary for general POVMs.

In parallel with the simulation cost of steering assemblages and bipartite quantum states $\gamma(\rho_{AB})$, Skrzypczyk *et al.* introduced the notion of *compatible complexity*, which, for a family of compatible measurements $\mathcal{M} := \{M_{a|x}\}$, is the smallest number of outcomes $|\Lambda|$ such that Eq. (5) holds [31]. We interpret this as the simulation cost of a set of compatible measurements $\mathcal{M}$ and denote it by $N(\mathcal{M})$.

In this paper, we first prove that $N(\mathcal{M})$ can be unbounded for a natural class of compatible measurements—namely, the noisy spin measurements introduced below. Using the connection between measurement incompatibility and EPR steering (Propositions 1–2), this yields an unbounded simulation cost $\gamma(\rho_{AB})$ for two-qubit Werner states and, more generally, for a broader class of en-

Singlet Weight	SR Cost $\gamma(\omega_r)$	State type
$r = 0$	$\gamma = 1$	product state
$0 < r \leq \frac{1}{3}$	$\gamma = 4$	separable
$\frac{1}{3} < r \leq \frac{1}{2}$	$\gamma > 4$; $\gamma = \Omega\big((\frac{1}{2} - r)^{-\frac{2}{5}}\big)$	entangled and unsteerable
$\frac{1}{2} < r < 1$	—	steerable

Table 1: Simulation cost of two-qubit Werner states ω_r.

tangled yet unsteerable states via reductions to Werner states.

3 Simulation Cost of Werner States and Noisy Spin Measurements

The bulk of our analysis focuses on:
(1):the two-qubit family of Werner states.

$$\omega_r = r|\Psi^-\rangle\langle\Psi^-| + (1-r)\frac{\mathbb{I}\otimes\mathbb{I}}{4}, \tag{6}$$

where $|\Psi^-\rangle = \sqrt{1/2}(|01\rangle - |10\rangle)$ is the singlet state;
(2): the family of noisy spin measurements (i.e, noisy qubit projective-valued measure):

$$\mathcal{P}_r = \{\{M^{(r)}_{+|\hat{n}}, M^{(r)}_{-|\hat{n}}\}\}_{\hat{n}}, \quad M^{(r)}_{\pm|\hat{n}} = \tfrac{1}{2}(\mathbb{I} \pm r\hat{n}\cdot\vec{\sigma}) \tag{7}$$

with $\hat{n}$ being a normalized Bloch vector, and the noise parameter $r \in [0,1]$ effectively shrinks the Bloch sphere down to radius r. We similarly write $\mathcal{P}^p_r$ for the subset of all noisy spin measurements with $\hat{n}$ confined to the x-z plane.

The Werner states and the noisy spin measurements provide paradigmatic cases for the study of quantum steering, and measurement incompatibility, respectively[19, 8, 27, 28, 32]. It is known that $\mathcal{P}_r$ is a compatible family if and only if $r \in [0, \frac{1}{2}]$ [8, 32], whereas the compatibility range for $\mathcal{P}^p_r$ (the planar spin measurements) is $r \in [0, \frac{2}{\pi}]$ [33] (see also Appendix E for explicit construction). Similarly, it was shown recently that ω_r is unsteerable if $r \in [0, \frac{1}{2}]$[34, 35] even when one considers all POVMs, and is unsteerable when $r \in [0, \frac{2}{\pi}]$ under planar POVMs (see also Proposition 2).

In this paper, however, rather than asking whether two-qubit Werner state or the set of noisy spin measurements is unsteerable or compatible, we focus on their simulation cost within their corresponding parameter regimes. We write $\gamma(\omega_r)$ and $\gamma^p(\omega_r)$ for the simulation cost of the Werner state under all POVMs and under planar POVMs, respectively. This simulation cost is a special case of those for general bipartite state ρ_{AB}, more formally:

Definition 1. The simulation cost $\gamma(\omega_r)$ is the smallest n such that every assemblage $\mathcal{A} := \{\sigma_{a|x}\}$ generated from ω_r by Eq. 1 admits an LHS model with at most n hidden states.

$$\gamma(\omega_r) := \inf\{n : \text{cost for simulating every } \mathcal{A} \text{ of } \omega_r\}, \tag{8}$$

Likewise, we write $N(r)$ and $N^p(r)$ for the simulation costs (compatible complexities) of the measurement families $\mathcal{P}_r$ and $\mathcal{P}^p_r$ within their compatibility regions, this simulation cost is a special case of those for general set of measurements $\mathcal{M} = \{M_{a|x}\}$, more formally:

Definition 2. The simulation cost $N(r)$ of the noisy spin measurements $\mathcal{P}_r$ is the smallest n for which there exists an n-outcome parent POVM $\{\Pi_\lambda\}$ such that every $\{M^{(r)}_{\pm|\hat{n}}\} \in \mathcal{P}_r$ can be simulated by Eq. (5).

$$N(r) := \inf\{n : \text{cost for simulating } \mathcal{P}_r\} \tag{9}$$

Our restriction to this simple family of states and sets of measurements may seem like a limitation. However, on the contrary, we show that many interesting features of simulation cost can already be demonstrated with these examples, and for the family of Werner states, they are summarized in Table 1. Furthermore, as we explain in this paper, the lower bounds we prove on $\gamma(\omega_r)$ can be translated into lower bounds on $\gamma(\rho_{AB})$ for *any* unsteerable state ρ_{AB}.

Another reason for focusing on Werner states and noisy spin measurements is their deep connection: the unsteerability of Werner states and the compatibility of noisy spin measurements are two sides of the same phenomenon, a link studied extensively in [27, 26]. Here we extend this connection to the level of simulation cost, which partially enables bidirectional translation of results.

Lemma 1. $\mathcal{P}_r$ can be simulated with an n-element POVM if and only if the simulation cost of Werner states ω_r under *projective measurements* equals n. $\gamma_{\mathrm{PM}}(\omega_r) = N(r)$.

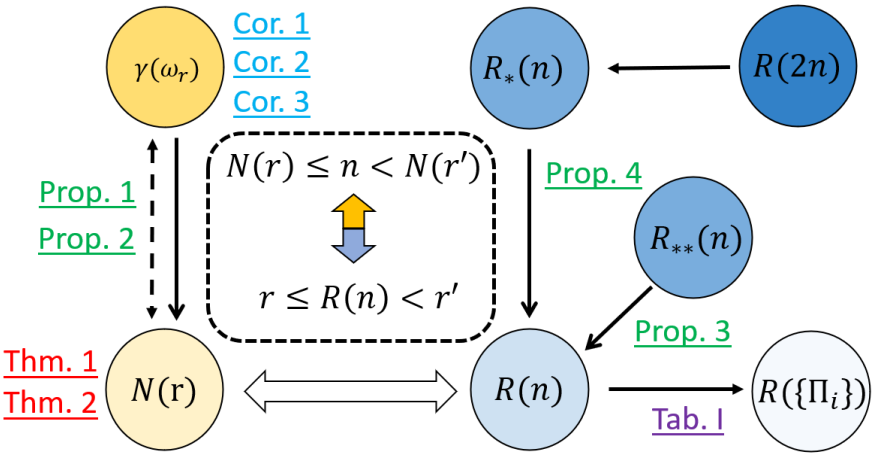

Figure 2: Schematic showing the connections between different quantities in this paper: (a) Left: Relation between simulation cost of Werner states $\gamma(\omega_r)$ and simulation cost (compatible complexity) $N(r)$ for the set of all noisy spin measurements ; (b) Right: Compatible radius for noisy spin measurements $R(n)$ and its derivatives, i.e., various lower and upper bounds; (c) The left and right parts are connected by the relation in the middle box (see Eq. 12); (d) $\rightarrow$(dashed): Upper bound from left to right (in special cases).

where $\gamma_{\mathrm{PM}}(\omega_r)$ is the simulation cost of ω_r when only projective-valued measures are considered.

The proof of Lemma 1 is detailed in Appendix F, this connection allows us to directly write the following propositions.

Proposition 1. $\gamma(\omega_r) \geq N(r)$.

This inequality follows immediately from the definitions: $\gamma(\omega_r)$ is the simulation cost of the Werner state under arbitrary POVMs on Alice's side, which include projective measurements as a special case. However, strong numerical evidence suggests that simulating all noisy POVMs is strictly more costly than simulating noisy spin PVMs. We discuss this in Appendix G and summarize it in the following conjecture.

Conjecture 1. The set of noisy qubit POVMs and noisy qubit PVMs (spin measurements) have different simulation costs.

Thus, it appears the above inequality might not be tight in general.

Nevertheless, in a special case, we further show in the appendix G that this inequality can be saturated for planar measurements.

Proposition 2. $\gamma^p(\omega_r) = N^p(r)$ for $r \in [0, \frac{2}{\pi}]$.

In what follows, we begin by studying the simulation cost $N(r)$ of noisy spin measurements. The connections made in Proposition 1 and Proposition 2 will then help us establish results for the simulation cost $\gamma(\omega_r)$ of two-qubit Werner states. More general results on the simulation cost of arbitrary bipartite states will be discussed in the last part of the paper.

4 Compatibility Radius and Its Geometric Characterization

Propositions 1 and 2 directly translate the problem of simulating ω_r into the problem of simulating noisy spin measurements $\mathcal{P}_r$, and accordingly the simulation cost $\gamma(\omega_r)$ to the simulation cost $N(r)$. However, Because $N(r)$ is an integer-valued step function of r, it is awkward to analyze directly. We therefore invert the viewpoint: instead of fixing the visibility r and asking for the simulation cost $N(r)$ for $\mathcal{P}_r$. Here we fix n and define $R(n)$ as the largest r for which there exists an n-outcome POVM that can simulate $\mathcal{P}_r$, i.e.,

Definition 3. The *compatibility radius* $R(n)$ is the largest visibility r such that the entire family of noisy spin measurements $\mathcal{P}_r$ can be simulated via Eq. (5) by some n-outcome POVM $\{\Pi_\lambda\}_{\lambda=1}^n$.

$$\begin{aligned} N(r) &= \inf\{n : \text{cost for simulating } \mathcal{P}_r\} \quad (10) \\ \Rightarrow R(n) &= \sup\{r : \mathcal{P}_r \text{ is simulable with cost } n\}. \end{aligned}$$

That is,

$$\begin{aligned} N(r) &:= \inf\{n : r \leq R(n)\} \quad (11) \\ \Rightarrow R(n) &:= \sup\{r : N(r) \leq n\}, \end{aligned}$$

thus, for any $r < r' \in [0, \frac{1}{2}]$ and integer n, we have the relationship:

$$r \leq R(n) < r' \quad \Leftrightarrow \quad N(r) \leq n < N(r'). \tag{12}$$

The same applies in the planar case with $N^p(r)$ and $R^p(n)$, i.e., $r \leq R^p(n) < r' \Leftrightarrow N^p(r) \leq n < N^p(r')$[1].

From Propositions 1 and 2, bounds on $\gamma(\omega_r)$ and $\gamma^p(\omega_r)$ can therefore be obtained by studying $R(n)$ and $R^p(n)$. This inversion allows us study each discrete n directly, and to characterize $R(n)$ and $R^p(n)$ both at finite n and in the asymptotic regime with $n \to \infty$; the detailed relation among all these quantitites are summarized in Fig. 1.

symbol	Explanation
$\{\Pi_\lambda\}_{\lambda\in\Lambda}$	POVM with effects $\{\Pi_\lambda\}$
$\mathrm{sym}\{\Pi_\lambda\}_{\lambda\in\Lambda}$	Symmetric extended POVM
Λ	Index set of outcomes
$\{M_{a\|x}\}_{a,x}$	Set of Children POVMs
$\gamma(\omega_r)$	Simulation cost of Werner states
$N(r)$	Simulation cost of noisy spin measurements
$\mathfrak{m}_{\{\Pi_\lambda\}}$	Compatible region of $\{\Pi_\lambda\}$
$\mathfrak{m}^*_{\{\Pi_\lambda\}}$	Compatible region of $\mathrm{sym}\{\Pi_\lambda\}$
$R(\{\Pi_\lambda\})$	Compatible radius for $\{\Pi_\lambda\}$
$R(n)$	$R(\{\Pi_\lambda\})$ for all n-outcome POVM
$R_*(n)$	Symmetric upper bound of $R(n)$
$R_{**}(n)$	Simplex upper bound of $R(n)$
X^p	Quantity for the planar case

Table 2: Notation used throughout the paper

We note that $R(n)$ and $R^p(n)$ are also of great interest on their own. Previously, studies on $R(n)$ and $R^p(n)$ have only been implicitly made in [18, 17, 36] by constructing finite-share-randomness LHS or LHV models with fixed outcome number n. Instead of optimizing over all n-outcome POVMs, these Such constructions typically use symmetric parent POVMs n-outcome POVMs $\{\Pi_\lambda\}$, which yield only lower bounds on $R(n)$ for specific finite n and offering neither extension to arbitrary n nor meaningful upper bounds. In contrast, here we present a thorough investigation of $R(n)$, deriving both lower and upper bounds in the finite and asymptotic regimes by exploring different geometric characterizations and related inequalities.

To study $R(n)$, we first fix a parent POVM $\{\Pi_\lambda\}$ and define its *compatibility radius* as the largest visibility r for which it simulates all noisy spin measurements $\mathcal{P}_r$:

$$R(\{\Pi_\lambda\}) = \sup\{r : \mathcal{P}_r \text{ is simulable by } \{\Pi_\lambda\}\}. \tag{13}$$

The n-outcome compatible radius $R(n)$ in definition 3 is then obtained by optimizing over all n-outcomes POVMs:

$$R(n) = \sup_{\{\Pi_\lambda\}} R(\{\Pi_\lambda\}). \tag{14}$$

We now characterize $R(\{\Pi_\lambda\})$ and $R(n)$ in qubit case, where the geometry is especially useful.

In qubit case, each POVM effect can be written in the Pauli basis as:

$$\Pi_\lambda = \alpha_\lambda(\mathbb{I} + \eta_\lambda \hat{n}_\lambda \cdot \vec{\sigma}), \tag{15}$$

with $0 \leq \eta_\lambda, \alpha_\lambda \leq 1$, and normalization constraints $\sum_\lambda \alpha_\lambda = 1$ and $\sum_\lambda \alpha_\lambda \eta_\lambda \hat{n}_\lambda = \vec{0}$[2].

Then, a noisy spin measurement $\{M^{(r)}_{\pm|\hat{n}} = \frac{1}{2}(\mathbb{I} \pm r\hat{n} \cdot \vec{\sigma})\}$ is simulated by the parent POVM $\{\Pi_\lambda = \alpha_\lambda(\mathbb{I} + \eta_\lambda \hat{n}_\lambda \cdot \vec{\sigma})\}$ precisely when there exist post-processing coefficients $q_\lambda \in [0, 1]$ such that one of its effect $M^{(r)}_{+|\hat{n}} = \sum_\lambda q_\lambda \Pi_\lambda$. This yields the natural geometric conditions:

$$r\hat{n} = 2\sum_\lambda q_\lambda \alpha_\lambda \eta_\lambda \hat{n}_\lambda \tag{16a}$$

$$\frac{1}{2} = \sum_\lambda q_\lambda \alpha_\lambda \tag{16b}$$

These relations underlie our geometric characterizations of $R(\{\Pi_\lambda\})$, specifically, we will have:

$$\begin{aligned} R(\{\Pi_\lambda\}) :&= \sup\{r : B(0, r) \subseteq \mathfrak{m}_{\{\Pi_\lambda\}}\} \\ &= \mathrm{inr}(\mathfrak{m}_{\{\Pi_\lambda\}}), \end{aligned} \tag{17}$$

where $B(0, r) = \{x : \|x\| \leq r\}$, $\mathrm{inr}(S)$ denotes inradius, and

$$\mathfrak{m}_{\{\Pi_\lambda\}} = \{2\textstyle\sum_\lambda q_\lambda \alpha_\lambda \eta_\lambda \hat{n}_\lambda \mid 0 \leq q_\lambda \leq 1, \sum_\lambda q_\lambda \alpha_\lambda = \frac{1}{2}\}. \tag{18}$$

[1] The equality cannot, in general, be taken on the right-hand side because $R(n)$ is defined as a supremum and is only right-continuous by definition.

[2] In the appendix D, we further show that it is sufficient to consider $\eta_\lambda = 1$ when optimizing over all $R(\{\Pi_\lambda\})$

is a convex 'constrained zonotope'[3] obtained by intersecting the zonotope generated by $\{2\alpha_\lambda\eta_\lambda\hat{n}_\lambda\}$ [37, 38], namely:

$$\mathfrak{m}^*_{\{\Pi_\lambda\}} = \{2\textstyle\sum_\lambda q_\lambda\alpha_\lambda\eta_\lambda\hat{n}_\lambda \mid 0 \le q_\lambda \le 1\} \quad (19)$$

with the affine hyperplane $\sum_\lambda q_\lambda\alpha_\lambda = \frac{1}{2}$ from Eq. (16b). Here the set $\mathfrak{m}^*_{\{\Pi_\lambda\}}$ is an important outer approximation of $\mathfrak{m}_{\{\Pi_\lambda\}}$, which coincides with the "constrained zonotope" associated with the $2n$-outcome POVM $\mathbf{sym}\{\Pi_\lambda\}_\lambda$:

$$\mathbf{sym}\{\Pi_\lambda\}_\lambda := \{\tfrac{\alpha_\lambda}{2}(\mathbb{I}+\eta_\lambda\hat{n}_\lambda\cdot\vec{\sigma}), \tfrac{\alpha_\lambda}{2}(\mathbb{I}-\eta_\lambda\hat{n}_\lambda\cdot\vec{\sigma})\}_\lambda.$$

as discussed in the appendix A.

By definition, $R(n)$ is obtained by taking the supremum $\mathrm{inr}(\mathfrak{m}_{\{\Pi_\lambda\}})$. But optimizing this quantity directly is challenging. In contrast, the outer approximation $\mathfrak{m}^*_{\{\Pi_\lambda\}}$-a zonotope- admits a simpler characterization and is related to $R(n)$ via the following chain of inequalities:

$$\begin{aligned} R(n) &= \sup_{\{\Pi_\lambda\}_{i=1}^n} \mathrm{inr}(\mathfrak{m}_{\{\Pi_\lambda\}}) \quad (20)\\ &\le \sup_{\{\Pi_\lambda\}_{i=1}^n} \mathrm{inr}(\mathfrak{m}^*_{\{\Pi_\lambda\}}) =: R_*(n) \le R(2n). \end{aligned}$$

The first inequality holds because $\mathfrak{m}^*_{\{\Pi_\lambda\}}$ is an outer approximation of $\mathfrak{m}_{\{\Pi_\lambda\}}$, while the last inequality follows since $R_*(n)$ is the greatest compatible radius achievable by symmetric $2n$-outcome measurement, which form a strict subset of all $2n$-outcome measurement.

In what follows, many of our results hinge on the characterization of the inradius of the zonotope $\mathrm{inr}(\mathfrak{m}^*_{\{\Pi_\lambda\}})$. These analysis yield new upper bounds on $R(n)$ via $R(n) \le R_*(n)$. Moreover, since $R_*(n)$ lower-bounds $R(2n)$, the same analysis provides lower bounds on $R(2n)$, in particular, it allows us to obtain results in the asymptotic regime with $n \to \infty$, which have never been explored before.

5 Lower Bound on Simulation Cost

In the previous section, we introduced the compatibility radius $R(n)$ as a dual perspective on simulation cost $N(r)$ for noisy spin measurements $\mathcal{P}_r$. And proposition 1 allows us to connect lower bound of $N(r)$ with lower bound on $\gamma(\omega_r)$. We now show that these links allow us to obtain lower bounds on simulation costs from upper bounds on $R(n)$ and to reinterpret the problem geometrically in terms of various geometric inequalities.

We begin with the simplest cases $n = 3$ and $n = 4$, where tight upper bounds follow from elementary geometric inequalities.

Proposition 3. $R^p(3) = \frac{1}{2}$ and $R(4) = \frac{1}{3}$.

Proof. The proof relies on the well-known circumradius–inradius inequality, which states that the inradius of an n-simplex (e.g., a triangle is a 2-simplex) is at least n times smaller than its circumradius [39].

To apply this, consider any three-outcome POVM $\{\Pi_\lambda\}_{\lambda=1}^3$. Observe that the set $\mathfrak{m}_{\{\Pi_\lambda\}}$ is contained within the 2-simplex:

$$\mathfrak{m}_{\{\Pi_\lambda\}} \subseteq \mathfrak{m}^{**}_{\{\Pi_\lambda\}} = \{\sum_{\lambda=1}^3 p_\lambda\hat{n}_\lambda | \sum_{\lambda=1}^3 p_\lambda = 1, p_\lambda \ge 0\} \quad (21)$$

where $\mathfrak{m}^{**}_{\{\Pi_\lambda\}}$ is simply obtained from Eq. 18 by setting $p_\lambda = 2q_\lambda\eta_\lambda\alpha_\lambda$ and dropping the constraints $q_\lambda \le 1$.),

Since $\{\hat{n}_\lambda\}$ are unit vectors, the circumradius of $\mathfrak{m}^{**}_{\{\Pi_\lambda\}}$ is at most one. The circumradius–inradius inequality [39] (outlined in Appendix D) then implies:

$$R^p(3) \le \sup_{\{\Pi_\lambda\}_{i=1}^3} \mathrm{inr}(\mathfrak{m}^{**}_{\{\Pi_\lambda\}}) =: R^p_{**}(3) \le \frac{1}{2}.$$

An analogous argument for four-outcome POVMs (corresponds to a 3-simplex) on the full Bloch sphere follows directly, which implies that $R(4) \le \frac{1}{3}$.

Combining these with the lower bounds for $R(n)$ and $R^p(n)$ in the next section (see Table 3) establishes the proposition. □

Corollary 1. For any $r > \frac{1}{3}$, we have $\gamma(\omega_r) > 4$. Thus, the simulation cost of any entangled two-qubit Werner state (with $r > \frac{1}{3}$[19]) is strictly greater than that of any separable two-qubit Werner state.

[3] As discussed in Appendix A, the set $\mathfrak{m}_{\{\Pi\lambda\}}$ is precisely the region of Bloch vectors corresponding to all unbiased two-outcome measurements simulable by $\{\Pi_\lambda\}$ (here "unbiased" means $\mathrm{Tr}[M_1] = \mathrm{Tr}[M_2]$ for a measurement M_1, M_2). Thereforefore, we will sometimes simply refer to $\mathfrak{m}_{\{\Pi\lambda\}}$ as the compatibility region.

Unfortunately, the outer approximation $R_{**}(n)$ is too loose for arbitrary n. In the following, we seek tighter outer approximations of $\mathsf{m}_{\{\Pi_\lambda\}}$ and better geometric inequalities.

We begin again with the planar case. Using an isoperimetric inequality for polygons, we obtain a sharper upper bound on $R^p(n)$:

Proposition 4. For n-outcome planar measurement $R^p(n) \leq R^p_*(n) = \frac{1}{n}\cot(\frac{\pi}{2n})$.

Proof. Let $\{\Pi_\lambda\}_{\lambda=1}^n$ be an arbitrary n-outcome planar POVM. According to Eq. (20), its compatible radius will be upper bounded by the inscribed radius of $\mathsf{m}^*_{\{\Pi_\lambda\}}$ in Eq. (19), i.e.,

$$R^p(n) = \sup_{\{\Pi_\lambda\}} \mathbf{inr}(\mathsf{m}^*_{\{\Pi_\lambda\}}) \leq \sup_{\{2\alpha_\lambda \hat{n}_\lambda\}} \frac{2A}{L}, \quad (22)$$

where $\mathsf{m}^*_{\{\Pi_\lambda\}}$ is a $2n$-sided zonogon generated by $\{2\alpha_\lambda \eta_\lambda \hat{n}_\lambda\}$, with area A and perimeter $L = 4\sum_\lambda \eta_\lambda \alpha_\lambda \leq 4$ (note that the polygon consists of n pairs of sides with length equals to its generators). Here the last inequalities follows simply by triangulating the convex polygon from its incenter, which relates the area A, perimeter L and inradius of the polygon.

The isoperimetric inequality of a such a $2n$-sided polygon [40] stipulates that

$$A \leq \frac{L^2}{8n\tan(\frac{\pi}{2n})}. \quad (23)$$

Therefore, we have the upper bound

$$R^p(n) \leq \frac{1}{n}\cot(\frac{\pi}{2n}) < \frac{2}{\pi} \quad (24)$$

□

Expanding $x\cot x$ about $x = 0$ using the Taylor expansion, we obtain

$$R^p(n) \leq \frac{1}{n}\cot(\frac{\pi}{2n}) < \frac{2}{\pi}\left(1 - \frac{1}{3}(\frac{\pi}{2n})^2\right) \quad (25)$$

Combining this bound with Eq. (12) and Proposition 2 then yields Corollary 2.

Corollary 2. For any $r \in [0, \frac{2}{\pi}]$,

$$\gamma^p(\omega_r) = N^p(r) > \sqrt{\frac{\pi}{6}}\left(\frac{2}{\pi} - r\right)^{-\frac{1}{2}}. \quad (26)$$

We now move beyond the planar case to the full Bloch sphere. Here we rely on stronger isoperimetric inequalities for zonotopes. The culminating result is stated in the following theorem and discussed in detail in the appendix D.

Theorem 1. For any $r \in [0, \frac{1}{2}]$,

$$R(n) < \frac{1}{2} - cn^{-\frac{5}{2}} \quad (27)$$

$$\gamma(\omega_r) \geq N(r) > c'\left(\frac{1}{2} - r\right)^{-\frac{2}{5}} \quad (28)$$

for some positive constant c and $c' = c^{\frac{2}{5}}$.

Both Corollary 2 and Theorem 1 establish our primary claim that the simulation cost of an unsteerable entangled state diverges at $r = \frac{1}{2}$ (resp. $\frac{2}{\pi}$) when considering general measurements (resp. planar measurements). We also note that the constants above trace back to those appearing in the original zonotope isoperimetric inequalities [41, 42]; see Appendix D for details.

6 Upper bound on the simulation cost

We now turn to upper bounds on the simulation cost of unsteerable Werner states $\gamma(\omega_r)$ and noisy spin measurements $N(r)$. We note, however, due to the inequality $\gamma(\omega_r) \geq N(r)$ in propositions 1, upper bounds on $N(r)$ do not translate into upper bounds on $\gamma(\omega_r)$ (except the planar case). Nevertheless, the simulation cost results $N(r)$ on noisy spin measurements remain valid on its own.

As before, we use Eq. (12) and turn the problem of upper bounding $N(r)$ into lower bounding $R(n)$. The latter is further lower bounded by $R(\{\Pi_\lambda\})$ for any fixed choice of POVM $\{\Pi_\lambda\}$. The same reasoning holds for the planar case.

For qubits, the compatible radius $R(\{\Pi_\lambda\})$ (with $\{\Pi_\lambda\}$ written in Pauli basis as in Eq. (15)) can actually be efficiently computed using the following optimization-based criterion:

Proposition 5. For a given POVM Π_λ, the compatible radius $R(\{\Pi_\lambda\})$ is given by:

$$R(\{\Pi_\lambda\}) = \inf_{\substack{|\vec{c}|=1 \\ -1\leq c_0\leq 1}} \sum_\lambda |\alpha_\lambda(c_0 + \eta_\lambda \vec{c}\cdot\hat{n}_\lambda)|. \quad (29)$$

Thus, when taking the supremum over all n-outcome POVMs Π_λ, the compatibility radius $R(n)$ can be expressed as:

$$R(n) = \sup_{\{\Pi_\lambda\}} \inf_{\substack{|\vec{c}|=1 \\ -1\leq c_0\leq 1}} \sum_\lambda |\alpha_\lambda(c_0 + \eta_\lambda \vec{c}\cdot\hat{n}_\lambda)|, \quad (30)$$

However, in general, the optimization problem is difficult to solve, and the POVMs that maximize the compatibility radius are typically asymmetric (see Table 3 for a brief overview). Nevertheless, the following corollary provides a useful simplification:

Corollary 3. $R(n)$ is maximized by rank-1 POVM $\{\Pi_n\}$, i.e., $\eta_\lambda = 1$ for $\lambda \in [n]$.

Proposition 5 and Corollary 3 are proved in detail in Appendix C; here we summarize its implications. Equation (29) shows that for any qubit POVM $\Pi_\lambda = \alpha_\lambda(\mathbb{I} + \eta_\lambda \hat{n}_\lambda \cdot \vec{\sigma})\}$, we can explicitly compute its compatible radius $R(\{\Pi_\lambda\})$ via a concrete minimization that depends only on the parameters $\alpha_\lambda, \eta_\lambda$ and $\hat{n}_\lambda$ of $\{\Pi_\lambda\}$. Intuitively, this minimization arises by recasting the problem of finding $R(\{\Pi_\lambda\})$ into a inclusion problem in Eq (17). Equation (30) then takes the supremum of these individual compatible radius over all n-outcome POVMs.

6.1 Qubit planar measurements

In general, to obtain the exact value of $R(n)$, one must optimize over all n-outcome measurements $\{\Pi_\lambda\}_{\lambda=1}^n$. However, some well-chosen n-outcome measurements can typically provide a relatively tight lower bound on $R(n)$.

Intuitively, POVMs $\{\Pi_\lambda\}_{\lambda=1}^n$ with a large compatible radius are those that are "symmetric". For example, consider qubit planar measurements. We first examine the rotationally symmetric planar POVM $\{\Pi_\lambda^{\text{rot}} = \frac{1}{n}(\mathbb{I} + \hat{n}_\lambda \cdot \vec{\sigma})\}_{\lambda=1}^n$, where $\hat{n}_\lambda = \left(\cos(\frac{2\pi\lambda}{n}), 0, \sin(\frac{2\pi\lambda}{n})\right)^T$ [24]. In the appendix E, we analytically determine their compatibility radius to be

$$R^p(\{\Pi_\lambda^{\text{rot}}\}) = \begin{cases} \frac{1}{n}\cot(\frac{\pi}{2n})\cos(\frac{\pi}{2n}) & \text{if } n \text{ is odd} \\ \frac{2}{n}\cot(\frac{\pi}{n}) & \text{if } n \text{ is even} \end{cases}. \tag{31}$$

Similar to the proof of Corollary 2, one can check the asymptotic limit of the above compatibility radius via Taylor expansion

$$R^p(\{\Pi_\lambda^{\text{rot}}\}) = \begin{cases} \frac{2}{\pi} - \frac{5\pi}{12n^2} + O(\frac{1}{n^4}) & \text{if } n \text{ is odd} \\ \frac{2}{\pi} - \frac{2\pi}{3n^2} + O(\frac{1}{n^4})) & \text{if } n \text{ is even.} \end{cases} \tag{32}$$

Since $\frac{1}{n}\cot(\frac{\pi}{2n})\cos(\frac{\pi}{2n})$ is much tighter comparing to $\frac{2}{n}\cot(\frac{\pi}{n})$ in the asymptotical region, we will always lower bound $R^p(n)$ using the odd n strategy (with a "$n-1$" shift to obtain lower bounds for both the even and odd cases.):

$$R^p(n) \geq R^p(\{\Pi_\lambda^{\text{rot}}\}) > \frac{2}{\pi} - \frac{5\pi}{12(n-1)^2}, \tag{33}$$

Combing Eq. (33) with Eq. (12) and Proposition 2, we can hence obtain the following upper bound of $N^p(r)$ and $\gamma^p(\omega_r)$:

Corollary 4. For any $r \in [0, \frac{2}{\pi}]$,

$$\gamma^p(\omega_r) = N^p(r) \leq \sqrt{\frac{5\pi}{12}}\left(\frac{2}{\pi} - r\right)^{-\frac{1}{2}} + 1, \tag{34}$$

A noteworthy and somewhat counterintuitive features of $R^p(\{\Pi_\lambda^{\text{rot}}\})$ is its non-monotonic in n, which implies that symmetric POVMs do not always have the largest compatibility radius (see also Table 3), contrary to common intuition [43, 19], where symmetric strategy were typically used to construct hidden variable models.

Nevertheless, a comparison of Corollaries 2 and 4 demonstrates that $\{\Pi_\lambda^{\text{rot}}\}_{\lambda=1}^n$ is asymptotically optimal for simulating noisy spin measurements in the x-z plane.

	Planar $R^p(n)$		General $R(n)$	
n	Symmetric	Numerics	Thomson	Numerics
3	0.5	0.5	0	0
4	0.5	0.5274	0.3333	0.3333
5	0.5854	0.5854	0.3464	0.3718
6	0.5774	0.5927	0.3333	0.4004
7	0.6102	0.6102	0.2857	0.4060
8	0.6035	0.6111	0.4392	
9	0.6206	0.6206	0.4446	
10	0.6155	0.6213	0.4376	
11	0.6259	0.6259	–	
12	0.6220	0.6265	0.4588	

Table 3: Lower bounds on the compatible radii $R^p(n)$ and $R(n)$. For the planar case, the first column describes the largest radius achievable by symmetric POVMs, as given by Eq. (31). For the general case, the third column provides a lower bound on $R(n)$ using solutions to the Thomson problem, while tighter numerical bounds are obtained by random sampling the parent POVM $\{\Pi_\lambda\}$[44].

6.2 General qubit measurements

Next, we turn to simulating all noisy spin measurements. In this general scenario, the rotationally symmetric planar POVM $\{\Pi_\lambda^{\text{rot}}\}$ cannot be directly generalized, as noisy spin measurements are no longer represented by a shrunken 2D disk but rather a 3D shrunken ball in the Bloch sphere representation. To address this, we consider alternative symmetric constructions for general qubit measurements, based on the idea of "evenly distributing" points on a sphere.

First, we utilize the equally-spaced points obtained from the Thompson problem [45]. The numerical results for $n \in \{4, \cdots, 12\}$ are shown in Table 3. Second, we construct a parent POVM by leveraging the symmetries of Platonic solids. A detailed analysis of this approach is provided in the appendix E, with results summarized in Table 4.

n	POVM $\{\Pi_\lambda\}$	$R(\{\Pi_\lambda\})$
4	tetrahedron	$\frac{1}{3}$
6	Octahedron	$\frac{1}{3}$
8	Cube	$\frac{\sqrt{6}}{6}$
12	Icosahedron	$\frac{\phi^3\sqrt{1+(1-\phi)^2}}{3(1+\phi^2)}$
20	Dodecahedron	$\sqrt{\frac{5}{6}\frac{\phi^2}{5}}$

Table 4: Parent POVM with platonic configuration and their associated compatible radius $R(\{\Pi_\lambda\})$, where $\phi = \frac{1+\sqrt{5}}{2}$ is the golden ratio.

The final and most powerful construction leverages results from Refs. [41, 46] on zonotope approximations. These works discuss how to achieve the best approximation of a unit ball using a zonotope generated by a finite number of generators. This is particularly relevant because the compatible region $\mathsf{m}^*_{\{\Pi_\lambda\}}$ mentioned earlier in the preliminaries is, in fact, a zonotope.

A detailed discussion on zonotopes is provided in the appendix D, and the key result is summarized in the following.

Theorem 2. For any $r \in [0, \frac{1}{2}]$,

$$R(2n) \geq \frac{1}{2} - Cn^{-\frac{5}{4}}. \tag{35}$$

$$N(r) \leq C' \left(\frac{1}{2} - r\right)^{-\frac{4}{5}} \tag{36}$$

for some positive constants C and $C' = 2C^{\frac{4}{5}}$ originated from the study of zonotope approximation problem [41, 46], and will be briefly discussed in Appendix D.

However, this does not directly provide an upper bound for the simulation cost of $\gamma(\omega_r)$, as Proposition 1 is not an equality. 2

7 Simulating Cost of General Unsteerable States

Corollary 1 shows that the separability threshold for two-qubit Werner states coincides with a jump in shared randomness cost from 4 to 5 (i.e., 2 bits to $\log_2 5$ bits). In fact, this connection between separability and simulation cost holds for all two-qubit states ρ_{AB}. That is, $\gamma(\rho_{AB}) > 4$ if and only if ρ_{AB} is entangled. As explained below, this previously unrecognized result actually follows directly from two different results in quantum information theory. First, it happens that for two-qubit separable states, the number of product states needed in a separable decomposition is always at most 4 [47, 48]. Second, Jevtic *et al.* have provided a 'nested tetrahedron' condition [49], which, when combined with the 'four-packable' condition by Nguyen and Vu [50], says the following.

Lemma 2 ([49])**.** A two-qubit state ρ_{AB} is separable if and only if its steering ellipsoid $\mathcal{E}_B$ fits inside a tetrahedron that fits inside the Bloch sphere

To be more precise, the steering ellipsoid $\mathcal{E}_B$ for a bipartite state ρ_{AB} is defined by first transforming ρ_{AB} to $\tilde{\rho}_{AB} \propto [(\rho_A)^{-1/2} \otimes \mathbb{I}]\rho_{AB}[(\rho_A)^{-1/2} \otimes \mathbb{I}]$ with $\rho_A = \text{Tr}_B[\rho_{AB}]$ (assumes ρ_A has full-rank, otherwise, it is just a product state). The possible unnormalized states that it can be steered to using the measurement effects $M_{\pm|\hat{n}} = \frac{1}{2}(\mathbb{I} \pm \hat{n} \cdot \vec{\sigma})$ as in Eq. 1 are given as

$$\tilde{\rho}_{AB} = \begin{pmatrix} 1 & \vec{b} \\ \vec{0} & \tilde{T} \end{pmatrix} \xrightarrow{M_{\pm|\hat{n}}} \sigma_{\pm|\hat{n}} = \frac{1}{4}\begin{pmatrix} 1 \\ \vec{b} \pm \tilde{T}\hat{n} \end{pmatrix}, \tag{37}$$

where we have written down both $\tilde{\rho}_{AB}$ the unnormalized state $\sigma_{\pm|\hat{n}}$ as four-dimensional vectors in

Pauli basis $\{\mathbb{I}, \sigma_x, \sigma_y, \sigma_z\}$, and the steering ellipsoid is then defined as the set

$$\mathcal{E}_B = \{\vec{\tilde{b}} + \tilde{T}\hat{n}\}_{\hat{n}} \tag{38}$$

Proposition 6. A two-qubit state ρ_{AB} is entangled if and only if $\gamma(\rho_{AB}) > 4$.

Proof. The "only if" part follows directly from the existence of minimal separable decompositions for two-qubit state [47, 48], which is no more than 4.

The "if" part can also be simply proved by contradiction. Assume ρ_{AB} is entangled (thus must have full-rank marginal) but has a simulation cost less than or equal to 4, then the filtered state $\tilde{\rho}_{AB} \propto [(\rho_A)^{-1/2} \otimes \mathbb{I}]\rho_{AB}[(\rho_A)^{-1/2} \otimes \mathbb{I}]$ must also be entangled, and with simulation cost less than or equal to 4. This is because the full-rank local filter on Alice is invertible and completely positive; it preserves separability/entanglement and induces a bijection between LHS models.

Therefore, there exists a set of local hidden states $\{\rho_\lambda = \frac{1}{2}(\mathbb{I} + \vec{n}_\lambda \cdot \vec{\sigma})\}_{\lambda=1}^4$ such that for measurement effects $M_{\pm|\hat{n}} = \frac{1}{2}(\mathbb{I} \pm \hat{n} \cdot \vec{\sigma})$ acting on $\tilde{\rho}_{AB}$, the unnormalized state $\sigma_{\pm|\hat{n}}$ can be simulated by

$$\begin{aligned} \sigma_{\pm|\hat{n}} &= \sum_{\lambda=1}^{4} p(\pm|\hat{n}, \lambda)p(\lambda)\rho_\lambda \\ \Rightarrow \mathcal{E}_B &= \{\vec{\tilde{b}} + \tilde{T}\hat{n}\}_{\hat{n}} \subseteq \mathrm{Conv}[\vec{n}_\lambda]_{\lambda=1}^4. \end{aligned} \tag{39}$$

This essentially shows that there exists a tetrahedron defined by the Bloch vectors $\{\vec{n}_\lambda\}_{\lambda=1}^4$ of $\{\rho_\lambda\}_{\lambda=1}^4$ that contain the ellipsoid $\mathcal{E}_B$. Thus, $\tilde{\rho}_{AB}$ is separable from Lemma 2, and there is a contradiction.

Thus, we conclude that a two-qubit ρ_{AB} is entangled if and only if $\gamma(\rho_{AB}) > 4$. □

Remark. For two qubits the same threshold holds symmetrically if one considers steering Alice by measuring Bob; choosing which side is measured does not affect the conclusion.

We now discuss ways in which our results can be applied to unsteerable states beyond two-qubit Werner family. Observe that any unsteerable state ρ_{AB} will satisfy $\gamma(\rho) \geq \gamma(\mathcal{L}(\rho))/m$, i.e., $\log_2(\gamma(\rho)) \geq \log_2(\gamma(\mathcal{L}(\rho))) - \log_2 m$ under any map of the form $\mathcal{L} = \sum_{i=1}^m \mathcal{N}_i \otimes \mathcal{E}_i$ (one-way LOCC map from Bob to Alice); where the $\mathcal{N}_i$ are completely positive and trace-preserving (CPTP) maps while the $\mathcal{E}_i$ are completely positive and $\sum_i \mathcal{E}_i$ being trace-preserving. Indeed, suppose that $\mathrm{Tr}_A[M_{a|x} \otimes \mathbb{I}\rho_{AB}] = \sum_{\lambda=1}^t p(a|x, \{M_{a|x}\}, \lambda)p(\lambda)\rho_\lambda$ is a LHS model for ρ_{AB}, with $\mathcal{M} = \{M_{a|x}\}$ being an arbitrary family of measurements. Then for any CP map $\mathcal{E}$ on Bob's system, we also have $\mathrm{Tr}_A[(M_{a|x} \otimes \mathbb{I})\mathcal{E}_B(\rho_{AB})] = \sum_{\lambda=1}^t p(a|x, \{M_{a|x}\}, \lambda)p(\lambda)\mathcal{E}(\rho_\lambda)$, and so

$$\begin{aligned} &\mathrm{Tr}_A[(M_{a|x} \otimes \mathbb{I})\mathcal{L}(\rho_{AB})] \\ &= \sum_{i=1}^m \mathrm{Tr}_A[(M_{a|x} \otimes \mathbb{I})\mathcal{N}_i \otimes \mathcal{E}_i(\rho_{AB})] \\ &= \sum_{i=1}^m \mathrm{Tr}_A[(\mathcal{N}_i^\dagger(M_{a|x}) \otimes \mathbb{I})\mathrm{id} \otimes \mathcal{E}_i(\rho_{AB})] \\ &= \sum_{i=1}^m \sum_{\lambda=1}^t p(a|x, \{\mathcal{N}_i^\dagger(M_{a|x})\}, \lambda)p(\lambda)\mathcal{E}_i(\rho_\lambda) \\ &= \sum_{i=1}^m \sum_{\lambda=1}^t q(a|x, \{M_{a|x}\}, \lambda, i)q(\lambda, i)\sigma_{\lambda,i} \end{aligned} \tag{40}$$

where $q(a|x, \{M_{a|x}\}, \lambda, i) = p(a|x, \{\mathcal{N}_i^\dagger(M_{a|x})\}, \lambda)$, $q(\lambda, i) = p(\lambda)\mathrm{Tr}[\mathcal{E}_i(\rho_\lambda)]$, and $\sigma_{\lambda_i} = \mathcal{E}_i(\rho_\lambda)/\mathrm{Tr}[\mathcal{E}_i(\rho_\lambda)]$. Hence, we can lower bound the simulation cost of ρ_{AB} by first transforming it to a two-qubit state using the map above.

Secondly, it is well-known that every bipartite state can be converted into a Werner state when Alice and Bob simultaneously apply one of twelve random Clifford gates to their system (a so-called "twirling map") [51]. Moreover, the noise parameter r in the resulting state is given by the singlet fraction, i.e. $\rho \mapsto \rho_W(r)$ with $r = \frac{4\langle\Phi^-|\rho|\Phi^-\rangle - 1}{3}$. Since the twirling map will increase the shared randomness in an LHS model by a factor of twelve, we conclude that any unsteerable state ρ_{AB} with $\langle\Phi^-|\mathcal{L}(\rho)|\Phi^-\rangle > \frac{3R(n)+1}{4}$ will require an LHS model with at least $\log_2 n - \log_2(12m)$ bits of shared randomness, by the same reasoning used in Eq. (40). Otherwise, one can use $\log_2(12m)$ bit share randomness to turn the state into a two-qubit Werner state and find a better way to simulate two-qubit Werner state. From Theorem 1, this amount grows unbounded as $\langle\Phi^-|\mathcal{L}(\rho)|\Phi^-\rangle \to \frac{5}{8}$, which corresponds to the threshold $r = \frac{1}{2}$ for Werner states.

Conclusions – In this work we studied the shared randomness cost $\gamma(\rho_{AB})$ for simulating c-to-cq channels built using ρ_{AB}. We character-

ized the simulation cost and showed that it can be unbounded. The mathematical methods we developed to study this problem used the correspondence between steerability and measurement incompatibility and their similar geometric picture. Analogous to any other type of channel simulation cost, the quantity $\gamma(\rho_{AB})$ provides one measure of operational resourcefulness for every unsteerable state ρ. Furthermore, understanding the simulation cost of different assemblages can be used for semi-device-independent entanglement verification [9] when Alice and Bob are known to have a limited amount of shared randomness, even in noisy environments where quantum steering is not possible.

The results presented here pertain to LHS models. As future work, it would be interesting if similar techniques can be applied to bound the shared randomness needed for local hidden variable (LHV) models. Since LHS models are more demanding than LHV models, the upper bounds presented in this work will still apply.

Finally, we close by conjecturing a more fundamental role for the simulation cost $\gamma(\rho_{AB})$ in the study of nonlocality and entanglement. The notion of steerability provides a clean partitioning in the set of bipartite quantum states between those that are steerable versus those that are not. However, this partition does not coincide with the separation between entangled and separable states [52, 8]. In contrast, we have found here that the for two-qubit states the jump from separable to entangled-unsteerable coincides with a jump in the simulation cost. Hence, by moving beyond steering and focusing on simulation cost, we recover the separable/entangled boundary. Perhaps such a relationship holds for higher-dimensional bipartite states as well.

Acknowledgements— This work was supported by NSF Award No. 1839177. We thank Virginia Lorenz and Marius Junge for helpful discussions during the preparation of this manuscript. We also thank the anonymous reviewer for carefully reading the manuscript and for pointing out errors and typos in earlier versions.

References

[1] Toby S. Cubitt, Debbie Leung, William Matthews, and Andreas Winter. "Improving zero-error classical communication with entanglement". Phys. Rev. Lett. **104**, 230503 (2010).

[2] Debbie Leung, Laura Mancinska, William Matthews, Maris Ozols, and Aidan Roy. "Entanglement can increase asymptotic rates of zero-error classical communication over classical channels". Communications in Mathematical Physics **311**, 97–111 (2012).

[3] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters. "Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels". Phys. Rev. Lett. **70**, 1895–1899 (1993).

[4] C.H. Bennett, P.W. Shor, J.A. Smolin, and A.V. Thapliyal. "Entanglement-assisted capacity of a quantum channel and the reverse shannon theorem". IEEE Transactions on Information Theory **48**, 2637–2655 (2002).

[5] Toby S. Cubitt, Debbie Leung, William Matthews, and Andreas Winter. "Zero-error channel capacity and simulation assisted by non-local correlations". IEEE Transactions on Information Theory **57**, 5509–5523 (2011).

[6] David Schmid, Denis Rosset, and Francesco Buscemi. "The type-independent resource theory of local operations and shared randomness". Quantum **4**, 262 (2020).

[7] Roope Uola, Ana C. S. Costa, H. Chau Nguyen, and Otfried Gühne. "Quantum steering". Rev. Mod. Phys. **92**, 015001 (2020).

[8] H. M. Wiseman, S. J. Jones, and A. C. Doherty. "Steering, entanglement, nonlocality, and the einstein-podolsky-rosen paradox". Phys. Rev. Lett. **98**, 140402 (2007).

[9] D Cavalcanti and P Skrzypczyk. "Quantum steering: a review with focus on semidefinite programming". Reports on Progress in Physics **80**, 024001 (2016).

[10] Marco Piani and John Watrous. "Necessary and sufficient quantum information characterization of einstein-podolsky-rosen steering". Phys. Rev. Lett. **114**, 060404 (2015).

[11] Charles H. Bennett, Igor Devetak, Aram W. Harrow, Peter W. Shor, and Andreas Winter. "The quantum reverse shannon theorem

and resource tradeoffs for simulating quantum channels". IEEE Transactions on Information Theory **60**, 2926–2959 (2014).

[12] B. F. Toner and D. Bacon. "Communication cost of simulating bell correlations". Phys. Rev. Lett. **91**, 187904 (2003).

[13] Martin J. Renner, Armin Tavakoli, and Marco Túlio Quintino. "Classical cost of transmitting a qubit". Phys. Rev. Lett. **130**, 120801 (2023).

[14] Ana Belén Sainz, Leandro Aolita, Nicolas Brunner, Rodrigo Gallego, and Paul Skrzypczyk. "Classical communication cost of quantum steering". Phys. Rev. A **94**, 012308 (2016).

[15] Daniel Cavalcanti, Paul Skrzypczyk, and Ivan Šupić. "All entangled states can demonstrate nonclassical teleportation". Phys. Rev. Lett. **119**, 110501 (2017).

[16] Marco Piani and John Watrous. "All entangled states are useful for channel discrimination". Phys. Rev. Lett. **102**, 250501 (2009).

[17] Flavien Hirsch, Marco Túlio Quintino, Tamás Vértesi, Matthew F. Pusey, and Nicolas Brunner. "Algorithmic construction of local hidden variable models for entangled quantum states". Phys. Rev. Lett. **117**, 190402 (2016).

[18] Joseph Bowles, Flavien Hirsch, Marco Túlio Quintino, and Nicolas Brunner. "Local hidden variable models for entangled quantum states using finite shared randomness". Phys. Rev. Lett. **114**, 120401 (2015).

[19] Reinhard F. Werner. "Quantum states with einstein-podolsky-rosen correlations admitting a hidden-variable model". Phys. Rev. A **40**, 4277–4281 (1989).

[20] Francesco Buscemi, Eric Chitambar, and Wenbin Zhou. "Complete resource theory of quantum incompatibility as quantum programmability". Phys. Rev. Lett. **124**, 120401 (2020).

[21] Yujie Zhang. "Towards quantum networks: Theory, experiment, and applications". Ph.d. dissertation. University of Illinois at Urbana–Champaign. Urbana, IL (2024). url: https://hdl.handle.net/2142/124212.

[22] Joseph Bowles, Tamás Vértesi, Marco Túlio Quintino, and Nicolas Brunner. "One-way einstein-podolsky-rosen steering". Phys. Rev. Lett. **112**, 200402 (2014).

[23] Joseph Bowles, Flavien Hirsch, Marco Túlio Quintino, and Nicolas Brunner. "Sufficient criterion for guaranteeing that a two-qubit state is unsteerable". Phys. Rev. A **93**, 022121 (2016).

[24] Jessica Bavaresco, Marco Túlio Quintino, Leonardo Guerini, Thiago O. Maciel, Daniel Cavalcanti, and Marcelo Terra Cunha. "Most incompatible measurements for robust steering tests". Phys. Rev. A **96**, 022110 (2017).

[25] Armin Uhlmann. "Entropy and optimal decompositions of states relative to a maximal commutative subalgebra". Open Systems & Information Dynamics **5**, 209–228 (1998).

[26] Marco Túlio Quintino, Tamás Vértesi, and Nicolas Brunner. "Joint measurability, einstein-podolsky-rosen steering, and bell nonlocality". Phys. Rev. Lett. **113**, 160402 (2014).

[27] Roope Uola, Tobias Moroder, and Otfried Gühne. "Joint measurability of generalized measurements implies classicality". Phys. Rev. Lett. **113**, 160403 (2014).

[28] Nicolas Brunner, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner. "Bell nonlocality". Rev. Mod. Phys. **86**, 419–478 (2014).

[29] Teiko Heinosaari, Takayuki Miyadera, and Mário Ziman. "An invitation to quantum incompatibility". Journal of Physics A: Mathematical and Theoretical **49**, 123001 (2016).

[30] Leonardo Guerini, Jessica Bavaresco, Marcelo Terra Cunha, and Antonio Acín. "Operational framework for quantum measurement simulability". Journal of Mathematical Physics **58**, 092102 (2017).

[31] Paul Skrzypczyk, Matty J. Hoban, Ana Belén Sainz, and Noah Linden. "Complexity of compatible measurements". Phys. Rev. Research **2**, 023292 (2020).

[32] Teiko Heinosaari, Jukka Kiukas, Daniel Reitzner, and Jussi Schultz. "Incompatibility breaking quantum channels". Journal of Physics A: Mathematical and Theoretical **48**, 435301 (2015).

[33] Roope Uola, Kimmo Luoma, Tobias Moroder, and Teiko Heinosaari. "Adaptive strategy for joint measurements". Phys. Rev. A **94**, 022109 (2016).

[34] Yujie Zhang and Eric Chitambar. "Exact steering bound for two-qubit werner states". Phys. Rev. Lett. **132**, 250201 (2024).

[35] Martin J. Renner. "Compatibility of generalized noisy qubit measurements". Phys. Rev. Lett. **132**, 250202 (2024).

[36] D Cavalcanti and P Skrzypczyk. "Quantum steering: a review with focus on semidefinite programming". Reports on Progress in Physics **80**, 024001 (2016).

[37] Peter McMullen. "On zonotopes". Transactions of the American Mathematical Society **159**, 91–109 (1971).

[38] Günter M. Ziegler. "Lectures on polytopes". Volume 152 of Graduate Texts in Mathematics. Springer. New York, NY (2012).

[39] Murray S. Klamkin and George A. Tsintsifas. "The circumradius-inradius inequality for a simplex". Mathematics Magazine **52**, 20–22 (1979). url: http://www.jstor.org/stable/2689968.

[40] Viktor Blåsjö. "The isoperimetric problem". The American Mathematical Monthly **112**, 526–566 (2005).

[41] J. Bourgain and J. Lindenstrauss. "Distribution of points on spheres and approximation by zonotopes". Israel Journal of Mathematics **64**, 25–31 (1988).

[42] Jean Bourgain and Joram Lindenstrauss. "Approximating the ball by a minkowski sum of segments with equal length". Discrete & Computational Geometry **9**, 131–144 (1993).

[43] H. Chau Nguyen, Huy-Viet Nguyen, and Otfried Gühne. "Geometry of einstein-podolsky-rosen correlations". Phys. Rev. Lett. **122**, 240401 (2019).

[44] Yujie Zhang. "Compatible-radius-calculation". https://github.com/yujie4phy/Compatible-radius-calculation (2023). GitHub repository.

[45] David J. Wales and Sidika Ulker. "Structure and dynamics of spherical crystals characterized for the thomson problem". Phys. Rev. B **74**, 212101 (2006).

[46] Jonathan W. Siegel. "Optimal approximation of zonoids and uniform approximation by shallow neural networks". Constructive Approximation **62**, 441–469 (2025).

[47] Anna Sanpera, Rolf Tarrach, and Guifré Vidal. "Local description of quantum inseparability". Phys. Rev. A **58**, 826–830 (1998).

[48] William K. Wootters. "Entanglement of formation of an arbitrary state of two qubits". Phys. Rev. Lett. **80**, 2245–2248 (1998).

[49] Sania Jevtic, Matthew Pusey, David Jennings, and Terry Rudolph. "Quantum steering ellipsoids". Phys. Rev. Lett. **113**, 020402 (2014).

[50] H. Chau Nguyen and Thanh Vu. "Nonseparability and steerability of two-qubit states from the geometry of steering outcomes". Phys. Rev. A **94**, 012114 (2016).

[51] D.P. DiVincenzo, D.W. Leung, and B.M. Terhal. "Quantum data hiding". IEEE Transactions on Information Theory **48**, 580–598 (2002).

[52] Jonathan Barrett. "Nonsequential positive-operator-valued measurements on entangled mixed states do not always violate a bell inequality". Phys. Rev. A **65**, 042302 (2002).

[53] Joseph K. Scott, Davide Martino Raimondo, Giuseppe Roberto Marseglia, and Richard D. Braatz. "Constrained zonotopes: A new tool for set-based estimation and fault detection". Automatica **69**, 126–136 (2016).

[54] J. Bourgain, J. Lindenstrauss, and V. Milman. "Approximation of zonoids by zonotopes". Acta Mathematica **162**, 73 – 141 (1989).

Appendix

A Geometry of compatible region **15**

B Proof of proposition 3, and Corollary 1 **17**

C Proof of Theorem 1, Theorem 2 and connection to Zonotope approximation **18**

D Criterion for compatible radius $R(\{\Pi_\lambda\})$ **20**

E Example of $R(\{\Pi_\lambda\})$ for symmetric POVMs, and compatible models **22**

F Simulation cost of entangled state and compatible measurements **26**

G Discussion on prop. 1, prop. 2 and conj. 1 **27**

1 Inequivalence between simulation cost of PVMs and POVMs 27

2 Equivalence between simulation cost of PVMs and POVMs for qubit planar measurements 29

Here we provide some technical details that complements the main manuscript.
In section A we establish a detailed geometrical connection between our problem and the analysis of zonotopes, a special type of convex polytope.
In section B, we prove prop. 3 and cor. 1 of the main text related to the first geometric inequality - the circumradius–inradius inequality of simplex.
In section C, we connect our problem to the zonotope approximation problems and thm 1 and thm 2 are proven with connection to geometric inequality therein, which allows us to study simulation cost in the asymptotic region (with cost $n \to \infty$).
In section D, we provide an optimization-based criterion for the computation of the compatibility radius $R(\{\Pi_\lambda\})$ for a given $\{\Pi_\lambda\}$, which can be used as lower bound for compatible radius $R(n)$, and thus for obtaining upper bound for $N(r)$.
In section E, we give special examples of symmetric POVMs and calculate their compatibility region $\mathfrak{m}_{\{\Pi_\lambda\}}$ and compatibility radius $R(\{\Pi_\lambda\})$ explicitly.
In section F, we establish the connection shared randomness cost in unsteerable state and compatible measurements.
Finally, in section G, we provide evidence and conjecture that the simulation cost of noisy spin measurements (PVMs) might be strictly smaller than the simulation cost of noisy POVMs.

A Geometry of compatible region

.

Definition 4. A zonotope is a set of points in d-dimensional space constructed from vectors $\{\vec{v}_\lambda\}$ by taking the Minkowski sum of line segments:

$$\mathcal{Z} = \left\{ \sum_\lambda x_\lambda \vec{v}_\lambda | 0 \leq x_\lambda \leq 1 \right\},$$

where the set of vectors $\{\vec{v}_i\}$ is called the generator of the zonotope.

With the above notation, we introduce the zonotope defined by $\{\Pi_\lambda\}$

$$\mathcal{M}_{\{\Pi_\lambda\}} := \left\{ \sum_\lambda q_\lambda \Pi_\lambda \middle| 0 \leq q_\lambda \leq 1 \right\}, \tag{41}$$

which is the collection of all effects that can be simulated by $\{\Pi_\lambda\}$ (or in other words, representing all dichotomic measurement that can be simulated by $\{\Pi_\lambda\}$). This region can be viewed as a zonotope in an d^2-dimensional space of bounded operators [37, 38], where d is the dimension of the Hilbert space on which the measurement effects act.

For qubit measurements, we can simplify the problem by parametrizing a quantum measurement $\{\Pi_\lambda\}$ in the Pauli basis, which provides a natural way to represent each effect $\Pi_\lambda = \alpha_\lambda(\mathbb{I} + \eta_\lambda \hat{n}_\lambda \cdot \vec{\sigma})$ as a 4-dimensional vector:

$$\vec{\pi}_\lambda = (\alpha_\lambda, \alpha_\lambda \eta_\lambda \hat{n}_\lambda)^T = (\alpha_\lambda, \alpha_\lambda \eta_\lambda \hat{n}_\lambda^x, \alpha_\lambda \eta_\lambda \hat{n}_\lambda^y, \alpha_\lambda \eta_\lambda \hat{n}_\lambda^z)^T. \tag{42}$$

The normalization and positivity constraints become $\sum_\lambda \vec{\pi}_\lambda = (1,0,0,0)$ and $\alpha_\lambda \geq 0$. Thus, we can geometrically visualize the compatible region for any given measurement in a 4-dimensional Euclidean space:

$$\mathfrak{M}_{\{\Pi_\lambda\}} := \left\{ 2 \sum_\lambda q_\lambda \vec{\pi}_\lambda \middle| 0 \leq q_\lambda \leq 1 \right\}, \tag{43}$$

where the factor of 2 is introduced for convenience.

Compared to the 4-dimensional compatible region, there are two 3-dimensional subsets of it that play a crucial role in the study of the simulation cost of noisy spin measurements: $\mathfrak{m}_{\{\Pi_\lambda\}}$ and $\mathfrak{m}^*_{\{\Pi_\lambda\}}$ defined in the main text. Both subsets can be derived from this 4-dimensional zonotope and are explained as follows:

Definition 5. The *constrained zonotope* $\mathfrak{m}_{\{\Pi_\lambda\}} := \left\{ 2\sum_\lambda q_\lambda \alpha_\lambda \eta_\lambda \hat{n}_\lambda \middle| 0 \leq q_\lambda \leq 1, \sum_\lambda q_\lambda \alpha_\lambda = \frac{1}{2} \right\}$ can be the three-dimensional cross-section of $\mathfrak{M}_{\{\Pi_\lambda\}}$ on the plane defined by $\sum_\lambda q_\lambda \alpha_\lambda = \frac{1}{2}$. However, in general, it need not be a zonotope[53]. The set $\mathfrak{m}_{\{\Pi_\lambda\}}$ represents the Bloch vectors of all unbiased two-outcome measurements (with each effect having the same trace) that can be simulated by $\{\Pi_\lambda\}$.

Definition 6. The *projected zonotope* $\mathfrak{m}^*_{\{\Pi_\lambda\}} := \left\{ 2\sum_\lambda q_\lambda \mathbb{P}\vec{\pi}_\lambda \middle| 0 \leq q_\lambda \leq 1 \right\} = \left\{ 2\sum_\lambda q_\lambda \alpha_\lambda \eta_\lambda \hat{n}_\lambda \middle| 0 \leq q_\lambda \leq 1 \right\}$ is a zonotope, where $\mathbb{P}$ projects vector $(x, \vec{y})^T \in \mathbb{R}^4$ onto $(\vec{y})^T \in \mathbb{R}^3$. By definition, $\mathfrak{m}_{\{\Pi_\lambda\}} \subseteq \mathfrak{m}^*_{\{\Pi_\lambda\}}$ with equality holding if and only if $\{\mathbb{P}\vec{\pi}_\lambda\} = \{(\alpha_\lambda \eta_\lambda \hat{n}_\lambda)^T\}$ is centrally symmetric. Centrally symmetric measurements are the POVMs $\mathbf{sym}\{\Pi_\lambda\}$ introduced in the main text.

The projected zonotope is meaningful both geometrically and analytically:

- Geometrically, given any arbitrary POVM $\{\Pi_\lambda\}$, we can symmetrically extend it to a new POVM $\mathbf{sym}\{\Pi_\lambda\} = \{\frac{\Pi_\lambda}{2}, \mathrm{Tr}[\Pi_\lambda]\mathbb{I} - \frac{\Pi_\lambda}{2})\}$. The projected zonotope for the original POVM is actually the constrained zonotope of the new symmetric-extended POVM $\mathbf{sym}\{\Pi_\lambda\}$.
- Analytically, when computing the compatibility radius $R(\{\Pi_\lambda\}) = \inf_{|\vec{c}|=1, -1 \leq c_0 \leq 1} \sum_\lambda |\alpha_\lambda(c_0 + \vec{c} \cdot \eta_\lambda \hat{n}_\lambda)|$ for the constrained zonotope (see Appendix D), an upper bound is given by the compatible radius for the projected zonotope: $R(\{\Pi_\lambda\}) \leq R_*(\{\Pi_\lambda\}) := \inf_{|\vec{c}|=1} \sum_\lambda |\alpha_\lambda(\vec{c} \cdot \eta_\lambda \hat{n}_\lambda)|$, which appears to be easier to characterize in many cases.

To see $\mathfrak{m}_{\mathbf{sym}\{\Pi_\lambda\}} = \mathfrak{m}^*_{\{\Pi_\lambda\}}$, we note that, $\mathfrak{m}_{\mathbf{sym}\{\Pi_\lambda\}\}}$ for POVM $\mathbf{sym}\{\Pi_\lambda\} = \left\{\frac{\alpha_\lambda}{2}(\mathbb{I} \pm \eta_\lambda \hat{n}_\lambda \cdot \vec{\sigma})\right\}_\lambda$ is defined as:

$$\begin{aligned}
\mathfrak{m}_{\mathbf{sym}\{\Pi_\lambda\}} &:= \left\{ \sum_{\lambda, t\in\{\pm\}} t q_{\lambda,t} \alpha_\lambda \eta_\lambda \hat{n}_\lambda \middle| 0 \leq q_{\lambda,t} \leq 1, \sum_{\lambda, t\in\{\pm\}} q_{\lambda,t} \frac{\alpha_\lambda}{2} = \frac{1}{2} \right\} \\
&= \left\{ \sum_\lambda (q_{\lambda,+} - q_{\lambda,-}) \alpha_\lambda \eta_\lambda \hat{n}_\lambda \middle| 0 \leq q_{\lambda,t} \leq 1, \sum_\lambda \frac{(q_{\lambda,+} + q_{\lambda,-})}{2} \alpha_\lambda = \frac{1}{2} \right\}
\end{aligned} \tag{44}$$

In order to show that $\mathfrak{m}_{\mathbf{sym}\{\Pi_\lambda\}} \supseteq \mathfrak{m}^*_{\{\Pi_\lambda\}}$. Given any $q_\lambda \in [0,1]$, we can choose $q_{\lambda,\pm} := \frac{1\pm(2q_\lambda-1)}{2} \in [0,1]$, which ensures that $\sum_\lambda \frac{\alpha_\lambda}{2}(q_{\lambda,+}+q_{\lambda,-}) = \frac{1}{2}$ and

$$\sum_\lambda (q_{\lambda,+}-q_{\lambda,-})\,\alpha_\lambda\eta_\lambda\hat{n}_\lambda = \sum_\lambda (2q_\lambda-1)\,\alpha_\lambda\eta_\lambda\hat{n}_\lambda = 2\sum_\lambda q_\lambda\,\alpha_\lambda\eta_\lambda\hat{n}_\lambda, \tag{45}$$

where in the last equality, we use the fact that $\sum_\lambda \alpha_\lambda\eta_\lambda\hat{n}_\lambda = \vec{0}$. Hence every point of $\mathfrak{m}^*_{\{\Pi_\lambda\}}$ lies in $\mathfrak{m}_{\mathbf{sym}\{\Pi_\lambda\}}$.

Similarly, to show that $\mathfrak{m}_{\mathbf{sym}\{\Pi_\lambda\}} \subseteq \mathfrak{m}^*_{\{\Pi_\lambda\}}$. Given any feasible $q_{\lambda,\pm} \in [0,1]$ with $\sum_\lambda \frac{\alpha_\lambda}{2}(q_{\lambda,+}+q_{\lambda,-}) = \frac{1}{2}$, we can define

$$q_\lambda := \frac{1+(q_{\lambda,+}-q_{\lambda,-})}{2} \in [0,1]. \tag{46}$$

Then

$$\sum_\lambda (q_{\lambda,+}-q_{\lambda,-})\,\alpha_\lambda\eta_\lambda\hat{n}_\lambda = \sum_\lambda (2q_\lambda-1)\,\alpha_\lambda\eta_\lambda\hat{n}_\lambda = 2\sum_\lambda q_\lambda\,\alpha_\lambda\eta_\lambda\hat{n}_\lambda, \tag{47}$$

again using $\sum_\lambda \alpha_\lambda\eta_\lambda\hat{n}_\lambda = \vec{0}$. Thus every point of $\mathfrak{m}_{\mathbf{sym}\{\Pi_\lambda\}}$ lies in $\mathfrak{m}^*_{\{\Pi_\lambda\}}$.

B Proof of proposition 3, and Corollary 1

Here we sketch the proof of the circumradius–inradius inequality for an n-simplex, following Murray [39]. As noted in the remark below, the result for $n = 2, 3$ can also be viewed as a direct consequence of the other identities/inequalities collected there.

Lemma 3 (The circumradius-inradius inequality[39])**.** The inradius r of an arbitrary n-simplex is at least n times less than it circumradius R. The inequality saturates when the n-simplex is regular.

Proof. Let A_i and F_i (for $i = 1, 2, ..., n+1$) denote, respectively, the vertices and its opposite $(n-1)$-dimensional faces of an n-dimensional simplex of volume V. Also, let h_i and η_λ denote the distances from A_i and circumcenter O to the F_i, respectively.

Then $R + \eta_\lambda \geq h_i$. Moreover, the volume of the n-dimensional simplex is given by h_iF_i/n, and so by evaluating the volume of the simplex in three different ways, we get:

$$nV = h_iF_i = \sum \eta_\lambda F_i = r\sum F_i.$$

Therefore,

$$\sum (R+\eta_\lambda)F_i = (R+r)\sum F_i \geq \sum h_iF_i = (n+1)r\sum F_i \to R \geq nr$$

□

Remark. For $n = 2$, the Euler's theorem in geometry states that, the inradius r, the circumradius R and the distance d of the centers of the circumcicle and inscribed cicle are related by:

$$0 \leq d^2 = R(R-2r) \Rightarrow R \geq 2r$$

For $n = 3$, the Grace–Danielsson inequality states that, the inradius r, the circumradius R and the distance d of the centers of the circumcicle and inscribed cicle are related by:

$$0 \leq d^2 \leq (R+r)(R-3r) \Rightarrow R \geq 3r$$

Proposition 3. $R^p(3) = \frac{1}{2}$ and $R(4) = \frac{1}{3}$.

Proof. To apply Lemma 3, we begin with the planar case. Given any three-outcome POVM $\{\Pi_\lambda\}_{i=1}^3$, observe that the set $\mathfrak{m}_{\{\Pi_\lambda\}}$ is contained in the 2-simplex $\mathfrak{m}^{**}_{\{\Pi_\lambda\}} = \{2\sum_{i=1}^3 p_\lambda \hat{n}_\lambda \mid \sum_{\lambda=1}^3 p_\lambda = 1, p_\lambda \geq 0\}$ in two dimensions. To see that, one can just replace $p_\lambda = 2q_\lambda\alpha_\lambda$ in the constrained zonotope, and the constrained zonotope has more constrain on $0 \leq q_\lambda \geq 1$. The 2-simplex has circumradius one because each $\hat{n}_i$ is a unit vector. By the circumradius–inradius inequality,

$$R^p(3) \leq \sup_{\{\Pi_\lambda\}_{i=1}^3} \mathrm{inr}(\mathfrak{m}^{**}_{\{\Pi_\lambda\}}) =: R^p_{**}(3) \leq \frac{1}{2}.$$

The same reasoning for four-outcome POVMs on the Bloch sphere gives $R(4) \leq \frac{1}{3}$, where now the compatible region $\mathfrak{m}_{\{\Pi_\lambda\}}$ and three-simplex $\mathfrak{m}^{**}_{\{\Pi_\lambda\}}$ are in $\mathbb{R}^3$. Combined with the lower bounds from Table I in the main text, we have the stated equalities. □

Corollary 1. For any $r > \frac{1}{3}$, we have $\gamma(\omega_r) > 4$. The simulation cost of any entangled Werner state is strictly greater than that of a separable Werner state.

C Proof of Theorem 1, Theorem 2 and connection to Zonotope approximation

In this section, we present the technical details showing how the well-studied results on approximating Euclidean balls with zonotopes apply to our problem.

Lemma 4. [proposition 6.6 [54]] For any zonotope generated by $2n$ line segments $\{\pm\alpha_\lambda\hat{n}_\lambda\}$ with $\sum_\lambda \alpha_\lambda = 1$, There exists a positive constant c_d depending only on dimension d such that:

$$\left\| \sum_\lambda \alpha_\lambda |\langle \hat{n}_\lambda, \hat{x}\rangle| - \beta_d \right\|_{L^2(\sigma_d)} \geq c_d n^{-\frac{d+2}{2(d-1)}}, \tag{48}$$

where $\beta_3 = \frac{1}{2}$, $\beta_2 = \frac{2}{\pi}$ and $\|f\|_{L^2(\sigma_d)} = \sqrt{\int_{S_{d-1}} |f(\hat{x})|^2 d\sigma_d(\hat{x})}$ is the L^2 norm on the integral of function over the $d-1$ dimensional unit sphere S_{d-1}. $\sigma_d(\hat{x})$ is the noramlized rotation invariant measure on S_{d-1} with

$$\beta_d := \int_{S_{d-1}} |\langle \hat{n}, \hat{x}\rangle| d\sigma_d(\hat{x}) \qquad \forall\ \hat{n}. \tag{49}$$

The proof is given in detail in [54] based on inequalities obtained from spherical harmonic expansion of these quantities.

Theorem 1. For any n-outcome POVM $\{\Pi_\lambda\}$, the compatible radius is upper bounded by

$$\begin{aligned} R(n) &\leq \frac{1}{2} - \frac{1}{2}c_3^2 n^{-\frac{5}{2}} \Rightarrow \gamma(\omega_r) \geq c_3' |\frac{1}{2} - r|^{-\frac{2}{5}} \\ R^p(n) &\leq \frac{2}{\pi} - \frac{1}{2}c_2^2 n^{-4} \Rightarrow \gamma^p(\omega_r) \geq c_2' |\frac{2}{\pi} - r|^{-\frac{1}{4}} \end{aligned} \tag{50}$$

for some positive constant $c_d' := (\frac{1}{2}c_d^2)^{\frac{d-1}{d+2}}$

Proof. For any n-outcome POVM $\{\Pi_\lambda\}$, we have a chain of inequality given as:

$$R(n) = \sup_{\{\Pi_\lambda\}} R(\{\Pi_\lambda\}) \leq \sup_{\{\Pi_\lambda\}} R(\mathbf{sym}\{\Pi_\lambda\}) = \sup_{\{2\alpha_\lambda\hat{n}_\lambda\}} \mathrm{inr}(\mathfrak{m}^*_{\{\Pi_\lambda\}}) = \sup_{\{2\alpha_\lambda\hat{n}_\lambda\}} \inf_{\hat{x}} \sum_\lambda \alpha_\lambda |\langle \hat{n}_\lambda, \hat{x}\rangle| := R_*(n) \tag{51}$$

(see Eq. (67) for the derivation of the last equality). From Lemma 4, we have a lower bound on $\|\sum_\lambda \alpha_\lambda |\langle \hat{n}_\lambda, \hat{x}\rangle| - \beta_d\|_{L^2(\sigma_d)}$. Using Holder's inequality and the fact that

$\|\sum_\lambda \alpha_\lambda |\langle \hat{n}_\lambda, \hat{x}\rangle| - \beta_d\|_{L^\infty(\sigma_d)} \leq 1$ (we note this is a quite loose relaxation, but is enough to establish our claim.),

$$\begin{aligned}\left\|\sum_\lambda \alpha_\lambda |\langle \hat{n}_\lambda, \hat{x}\rangle| - \beta_d\right\|^2_{L^2(\sigma_d)} &\leq \left\|\sum_\lambda \alpha_\lambda |\langle \hat{n}_\lambda, \hat{x}\rangle| - \beta_d\right\|_{L^1(\sigma_d)} \left\|\sum_\lambda \alpha_\lambda |\langle \hat{n}_\lambda, \hat{x}\rangle| - \beta_d\right\|_{L^\infty(\sigma_d)} \\ &\leq \left\|\sum_\lambda \alpha_\lambda |\langle \hat{n}_\lambda, \hat{x}\rangle| - \beta_d\right\|_{L^1(\sigma_d)} \end{aligned} \tag{52}$$

Where L^1, and L^∞ stand for the L^1 and L^∞ norm of the integral of measurable function. Additionally, since $\int(\sum_\lambda \alpha_\lambda |\langle \hat{n}_\lambda, \hat{x}\rangle| - \beta_d) d\sigma_d(\hat{x}) = 0$, we have

$$\int_{\sum_\lambda \alpha_\lambda |\langle \hat{n}_\lambda, \hat{x}\rangle| < \beta_d} \left|\sum_\lambda \alpha_\lambda |\langle \hat{n}_\lambda, \hat{x}\rangle| - \beta_d\right| d\sigma_d(\hat{x}) = \frac{1}{2}\left\|\sum_\lambda \alpha_\lambda |\langle \hat{n}_\lambda, \hat{x}\rangle| - \beta_d\right\|_{L^1(\sigma_d)}, \tag{53}$$

we therefore have

$$\sup_{\hat{x}} \left(\beta_d - \sum_\lambda \alpha_\lambda |\langle \hat{n}_\lambda, \hat{x}\rangle|\right) \geq \frac{1}{2}\left\|\sum_\lambda \alpha_\lambda |\langle \hat{n}_\lambda, \hat{x}\rangle| - \beta_d\right\|_{L^1(\sigma_d)} \geq \frac{1}{2}\left\|\sum_\lambda \alpha_\lambda |\langle \hat{n}_\lambda, \hat{x}\rangle| - \beta_d\right\|^2_{L^2(\sigma_d)} \geq \frac{1}{2} c_d^2 n^{-\frac{d+2}{d-1}}. \tag{54}$$

In the end we have

$$\sup_{\{2\alpha_\lambda \hat{n}_\lambda\}} \inf_{\hat{x}} \sum_\lambda \alpha_\lambda |\langle \hat{n}_\lambda, \hat{x}\rangle| \leq \beta_d - \frac{1}{2} c_d^2 n^{-\frac{d+2}{d-1}}, \tag{55}$$

from which we conclude:

$$\begin{aligned} R(n) \leq R_*(n) \leq \frac{1}{2} - \frac{1}{2} c_3^2 n^{-\frac{5}{2}} &\Rightarrow \gamma(\omega_r) \geq c_3' |\frac{1}{2} - r|^{-\frac{2}{5}} \\ R^p(n) \leq R^p_*(n) \leq \frac{2}{\pi} - \frac{1}{2} c_2^2 n^{-4} &\Rightarrow \gamma^p(\omega_r) \geq c_2' |\frac{2}{\pi} - r|^{-\frac{1}{4}} \end{aligned} \tag{56}$$

for some positive constant $c_d' := (\frac{1}{2} c_d^2)^{\frac{d-1}{d+2}}$. □

Remark. The bound here for planar measurements ($d = 2$) is less tight than the upper bound we obtained in corollary 4.

Lemma 5 (Theorem 1 [46])**.** There exists a positive constant C_d and zonotope in $\mathbb{R}^d$ with $2n$ generators $\{\pm\alpha_\lambda \hat{n}_\lambda\}_{\lambda=1}^n$ with $\sum \alpha_\lambda = 1$ such that

$$\left|\sum_{\lambda=1}^n \alpha_\lambda |\hat{n}_\lambda \cdot \hat{x}| - \beta_d\right| < C_d n^{-\frac{d+2}{2(d-1)}} \qquad \forall \hat{x} \in S_{d-1}, \tag{57}$$

where $\beta_d := \int_{S_{d-1}} |\langle \hat{n}, \hat{x}\rangle| d\sigma_d(\hat{x})$, e.g., $\beta_2 = \frac{2}{\pi}$ and $\beta_3 = \frac{1}{2}$.

Since $\inf_{\hat{x}} \sum_{\lambda=1}^n \alpha_\lambda |(\hat{x} \cdot \hat{n}_\lambda)| \leq \sum_{\lambda=1}^n \alpha_\lambda \int |(\hat{x} \cdot \hat{n}_\lambda)| d\sigma_d(\hat{x}) = \beta_d \sum_{\lambda=1}^n \alpha_\lambda = \beta_d$ by the definition in Eq. (49), from lemma 5,

$$R(2n) \geq R(\mathrm{sym}\{\Pi_\lambda\}) = \inf_{\hat{x}} \sum_{\lambda=1}^n \alpha_\lambda |(\hat{x} \cdot \hat{n}_\lambda)| > \beta_d - C_d n^{-\frac{d+2}{2(d-1)}}. \tag{58}$$

Thus we immediately obtain an upper bound on the simulation cost:

Theorem 2. For $r < \beta_3 = \frac{1}{2}$, the simulation cost is upper bounded by:

$$N(r) \leq C_3' |\frac{1}{2} - r|^{-4/5} \tag{59}$$

for some positive constants $C_3' = 2(C_3)^{\frac{4}{5}}$.

D Criterion for compatible radius $R(\{\Pi_\lambda\})$

The following proposition is closely related to theorem 1 in [43] for quantum steering for bipartite state.

Proposition 5. For a given qubit POVM $\{\Pi_\lambda = \alpha_\lambda(\mathbb{I} + \eta_\lambda \hat{n}_\lambda \cdot \vec{\sigma})\}$, the compatible radius $R(\{\Pi_\lambda\})$ is given by:

$$R(\{\Pi_\lambda\}) = \inf_{\substack{|\vec{c}|=1 \\ -1\le c_0\le 1}} \sum_\lambda |\alpha_\lambda(c_0 + \eta_\lambda \vec{c}\cdot\hat{n}_\lambda)|. \tag{60}$$

Proof. We start by noticing that the set $\mathcal{M}_{\{\Pi_\lambda\}}$ is convex for any given $\{\Pi_\lambda\}$. Therefore, we can always find a set of tight inequalities that bound $\mathcal{M}_{\{\Pi_\lambda\}}$. Such an inequality can be represented using an operator $C = c_0\mathbb{I} + \vec{c}\cdot\vec{\sigma}$. Let $N \in \mathcal{M}_{\{\Pi_\lambda\}}$, then for any C we can write the inequality as

$$\langle C, N\rangle \le \sup_{M\in\mathcal{M}_{\{\Pi_\lambda\}}} \langle C, M\rangle \tag{61}$$

where $\langle X, Y\rangle = \mathrm{Tr}[XY]$.

Given a POVM $N = n_0(\mathbb{I} + \vec{n}\cdot\vec{\sigma})$, we can proceed to simplify the above inequality:

$$\begin{aligned} n_0(c_o + \vec{c}\cdot\vec{n}) &\le \sum_\lambda \sup_{0\le x_\lambda\le 1} \alpha_\lambda x_\lambda(c_0 + \eta_\lambda\vec{c}\cdot\hat{n}_\lambda) \\ &= \sum_\lambda \max[\alpha_\lambda(c_0 + \eta_\lambda\vec{c}\cdot\hat{n}_\lambda), 0], \end{aligned} \tag{62}$$

where the last equality always holds by setting $x_\lambda = 0$ whenever $c_0 + \eta_\lambda\vec{c}\cdot\hat{n}_\lambda \le 0$ and $x_\lambda = 1$ otherwise. Moreover for any unbiased measurement $N \in \mathcal{M}_{\{\Pi_\lambda\}}$ with $n_0 = \frac{1}{2}$ (which takes noisy spin measurements as a special case), we can simplify the above expression as:

$$\begin{aligned} \vec{c}\cdot\vec{n} \le 2\sum_\lambda \max[\alpha_\lambda(c_0 + \eta_\lambda\vec{c}\cdot\hat{n}_\lambda), 0] - c_0 &= 2\sum_\lambda \max[\alpha_\lambda(c_0 + \eta_\lambda\vec{c}\cdot\hat{n}_\lambda), 0] - \sum_\lambda \alpha_\lambda c_0 - \sum_\lambda \alpha_\lambda\eta_\lambda\vec{c}\cdot\hat{n}_\lambda \\ &= \sum_\lambda |\alpha_\lambda(c_0 + \eta_\lambda\vec{c}\cdot\hat{n}_\lambda)|, \end{aligned} \tag{63}$$

where we use the identity $2\max[t, 0] - t = |t|$, $\sum_\lambda \alpha_\lambda = 1$ and $\sum_\lambda \alpha_\lambda\eta_\lambda\hat{n}_\lambda = \vec{0}$ in the first equality.

When varying over all choices of operator C (thus, all inequalities that bound the convex set $\mathcal{M}_{\{\Pi_\lambda\}}$), we finally arrive at the criteria:

$$|\vec{n}| \le \inf_{c_0,\vec{c}} \frac{\sum_\lambda |\alpha_\lambda(c_0 + \eta_\lambda\vec{c}\cdot\hat{n}_\lambda)|}{|\vec{c}|}. \tag{64}$$

Since we are always allowed to scale c_0 and $|\vec{c}|$, the above infimum can be further simplified with constraint $|\vec{c}| = 1$ and $-1 \le c_0 \le 1$ (the last constraint follows from the fact that minimal of any sum of sign functions $\inf_x \sum_i |x + b_i|$ is achieved with $|x| \le \sup_i |b_i|$):

$$R(\{\Pi_\lambda\}) = \inf_{\substack{|\vec{c}|=1 \\ -1\le c_0\le 1}} \sum_\lambda |\alpha_\lambda(c_0 + \eta_\lambda\vec{c}\cdot\hat{n}_\lambda)| \tag{65}$$

The compatibility radius $R(n)$ can then be expressed as

$$R(n) = \sup_{\{\Pi_\lambda\}} \inf_{\substack{|\vec{c}|=1 \\ -1\le c_0\le 1}} \sum_\lambda |\alpha_\lambda(c_0 + \eta_\lambda\vec{c}\cdot\hat{n}_\lambda)|, \tag{66}$$

where the supremum is taken over all n-element POVMs. □

Corollary 3. $R(n)$ is maximized by rank-1 POVM $\{\Pi_n\}$, i.e., $\eta_\lambda = 1$ for $\lambda \in [n]$.

Proof. Given any POVM $\{\Pi_\lambda\}$ with $\Pi_\lambda = \alpha_\lambda(\mathbb{I} + \eta_\lambda \hat{n}_\lambda \cdot \vec{\sigma})$, we can define POVM $\{\Pi'_\lambda\}$ with:

$$\Pi'_\lambda = \beta_\lambda(\mathbb{I} + \hat{n}_\lambda \cdot \vec{\sigma})$$

where $\beta_\lambda = \frac{\alpha_\lambda \eta_\lambda}{\sum_j \alpha_j \eta_j}$. Let c_0^* and $\vec{c^*}$ be such that

$$R(\{\Pi'_\lambda\}) = \inf_{\substack{|\vec{c}|=1 \\ -1 \le c_0 \le 1}} \sum_\lambda |\beta_\lambda(c_0 + \vec{c} \cdot \hat{n}_\lambda)| = \sum_\lambda |\beta_\lambda(c_0^* + \vec{c^*} \cdot \hat{n}_\lambda)| = \frac{1}{\sum_j \alpha_j \eta_j} \sum_\lambda |\alpha_\lambda \eta_\lambda (c_0^* + \vec{c^*} \cdot \hat{n}_\lambda)|$$

For each individual term above,

$$|\alpha_\lambda \eta_\lambda (c_0^* + \vec{c^*} \cdot \hat{n}_\lambda)| + |(1 - \eta_\lambda)\alpha_\lambda c_0^*| \ge |\alpha_\lambda (c_0^* + \eta_\lambda \vec{c^*} \cdot \hat{n}_\lambda)|.$$

Summing over λ, we have

$$\sum_\lambda |\alpha_\lambda \eta_\lambda (c_0^* + \vec{c^*} \cdot \hat{n}_\lambda)| + \sum_\lambda |(1 - \eta_\lambda)\alpha_\lambda c_0^*| \ge \sum_\lambda |\alpha_\lambda (c_0^* + \eta_\lambda \vec{c^*} \cdot \hat{n}_\lambda)|,$$

which is equivalent to

$$\sum_j \alpha_j \eta_j R(\{\Pi'_\lambda\}) + (1 - \sum_j \alpha_j \eta_j)|c_0^*| \ge \sum_\lambda |\alpha_\lambda (c_0^* + \eta_\lambda \vec{c^*} \cdot \hat{n}_\lambda)| \ge R(\{\Pi_\lambda\}).$$

The last inequality holds because $(c_0^*, \vec{c^*})$ is optimal for $\{\Pi'_\lambda\}$ but might not be the optimal choice for $\{\Pi_\lambda\}$.

The last piece of the proof relies on showing $R(\{\Pi'_\lambda\}) \ge |c_0^*|$. This holds because:

$$R(\{\Pi'_\lambda\}) = \sum_\lambda |\beta_\lambda(c_0^* + \vec{c^*} \cdot \hat{n}_\lambda)| \ge |\sum_\lambda \beta_\lambda (c_0^* + \vec{c^*} \cdot \hat{n}_\lambda)| = |c_0^*|,$$

where the last equality holds since $\sum_\lambda \Pi_\lambda = \mathbb{I}$ (i.e., $\sum_\lambda \beta_\lambda \hat{n}_\lambda = \vec{0}$ and $\sum_\lambda \beta_\lambda = 1$). Therefore, we finally have:

$$R(\{\Pi'_\lambda\}) \ge R(\{\Pi_\lambda\}).$$

□

Remark. For symmetric POVM $\mathrm{sym}\{\Pi_\lambda\}$, the compatible radius can be computed as

$$R(\mathrm{sym}\{\Pi_\lambda\}) = \inf_{\hat{c}} \sum_{\lambda=1}^{n} \alpha_\lambda |(\hat{c} \cdot \eta_\lambda \hat{n}_\lambda)| \tag{67}$$

To show this is true for all symmetrically symmetric POVM, notice $\inf_x(|x+y| + |x-y|) = 2|y|$, thus

$$R(\mathrm{sym}\{\Pi_\lambda\}) = \inf_{\substack{|\vec{c}|=1 \\ -1 \le c_0 \le 1}} \frac{1}{2}[\sum_\lambda |\alpha_\lambda (c_0 + \eta_\lambda \vec{c} \cdot \hat{n}_\lambda)| + \sum_\lambda |\alpha_\lambda (c_0 - \eta_\lambda \vec{c} \cdot \hat{n}_\lambda)|] = \inf_{\hat{c}} \sum_{\lambda=1}^{n} \alpha_\lambda |(\hat{c} \cdot \eta_\lambda \hat{n}_\lambda)| \tag{68}$$

In the following we assume $\eta_\lambda = 1$ for convenience.

Corollary 4. The Compatible radius $R(\{\Pi_\lambda\})$ can be computed by considering a finite set of less than $\binom{n}{3}$ operators $C = c_0 \mathbb{I} + \vec{c} \cdot \vec{\sigma}$, where $(c_0, \vec{c})^T$ is defined by the norm vector of facets of $\mathfrak{M}_{\{\Pi_\lambda\}}$.

Proof. As discussed in Section A, the compatible region is a convex zonotope. Instead of running over all hyperplane boundaries given by operator $C = c_0\mathbb{I} + \vec{c}\cdot\vec{\sigma}$, it suffices to consider a finite set of the operator C that define the normal vectors of those facets. A similar idea was discussed in [43].

From the property of a zonotope, every edge of the compatible region $\mathfrak{M}_{\{\Pi_\lambda\}}$ is always parallel to one of its generator vectors proportional to $(1, \hat{n}_\lambda)$, while each facet is defined by $d-1=3$ edges of the set. Therefore, we just have to consider at most $\binom{n}{3}$ different choices of normal vectors. To be more specific, the vector $(c_0, \vec{c})$ should be perpendicular to three different vectors $(1, \hat{n}_\lambda)$. Hence, we can solve the equations $(c_0, \vec{c})\cdot(1, \hat{n}_\lambda)^T = 0$ and obtain:

$$\vec{c} = \frac{(\hat{n}_x - \hat{n}_y)\times(\hat{n}_x - \hat{n}_z)}{|(\hat{n}_x - \hat{n}_y)\times(\hat{n}_x - \hat{n}_z)|}, \qquad c_0 = -\vec{c}\cdot\hat{n}_x \tag{69}$$

where $\hat{n}_x, \hat{n}_y$ and $\hat{n}_z$ are any choices of distinct vectors associated to POVM $\{\Pi_\lambda\}$. Therefore, we can now simplify the criteria above by calculating the minimum value over a finite set of size $\binom{n}{3}$.

Similarly, for the case with planar measurements, the vector $(c_0, \vec{c})$ should be perpendicular to two different vectors $(1, \hat{n}_\lambda)^T$. Hence we can write:

$$(c_0, \vec{c})^T = (1, \hat{n}_x)^T \times (1, \hat{n}_y)^T \tag{70}$$

where $\vec{c}, \hat{n}_\lambda$ are written as vectors in $\mathbb{R}^2$. □

E Example of $R(\{\Pi_\lambda\})$ for symmetric POVMs, and compatible models

In this section, we compute the compatible radius $R(\{\Pi_\lambda\})$ for different classes of symmetric POVMs, including: **Result 1**: the compatible radius for qubit rotationally symmetric planar measurements;
Result 2: a compatible model based on qubit rotationally symmetric planar measurements;
Result 3: the compatible radius for qubit POVMs with regular polyhedron configuration.

These results are relatively easy to compute thanks to the remark mentioned in the last section, i.e., for symmetric POVM $\mathrm{sym}\{\Pi_\lambda\}$, the compatible radius can be computed as

$$R(\mathrm{sym}\{\Pi_\lambda\}) = \inf_{\hat{c}} \sum_{\lambda=1}^{n} \alpha_\lambda |(\hat{c}\cdot\eta_\lambda\hat{n}_\lambda)| \tag{71}$$

Result 1. For equally spaced planar measurement $\{\Pi_\lambda^{\mathrm{rot}} = \frac{1}{n}(\mathbb{I} + \hat{n}_\lambda\cdot\sigma)\}$ with $\hat{n}_\lambda = [\cos(\frac{2\pi\lambda}{n}), 0, \sin(\frac{2\pi\lambda}{n})]^T$, we have:

$$R^p(\{\Pi_\lambda^{\mathrm{rot}}\}) = \begin{cases} \frac{1}{n}\cot(\frac{\pi}{2n})\cos(\frac{\pi}{2n}) & \text{if } n \text{ is odd} \\ \frac{2}{n}\cot(\frac{\pi}{n}) & \text{if } n \text{ is even.} \end{cases} \tag{72}$$

Proof. If n is even, the original vectors $\hat{n}_\lambda$ form a regular polygon with n sides. To get the extreme point $M_i^n \in \mathcal{M}_{\{\Pi_\lambda\}}$, we can simply check that, they are obtained by adding up half of $\{\Pi_\lambda^{\mathrm{rot}}\}$:

$$M_i^n = \sum_{j=i}^{j=n/2+i-1} \Pi_j^{\mathrm{rot}} = \frac{1}{2}(\mathbb{I} + \vec{m}_i\cdot\sigma) \tag{73}$$

where $\{\vec{m}_i\}$ form exactly the same regular polygon with $|\vec{m}_i| = \frac{2}{n\sin(\frac{\pi}{n})}$.

Given an n-sided regular polygon, the ratio between its inscribed radius and circumscribed radius is $\frac{r_i}{r_c} = \cos(\frac{\pi}{n})$ (standard regular-polygon formula). Therefore, the circle contained in the n-sided polygon defined by $\vec{m}_i$ has radius

$$r_n = \frac{r_i}{r_c}|\vec{m}_i| = \cos(\frac{\pi}{n})\cdot\frac{2}{n\sin(\frac{\pi}{n})} = \frac{2}{n}\cot(\frac{\pi}{n}). \tag{74}$$

If n is odd, the case is slightly different; instead, the extreme points can be enumerated as

$$\begin{aligned} M_i^n &= \sum_{j=i}^{j=(n-1)/2+i-1} \Pi_j^{\text{rot}} + \frac{1}{2}\Pi_{(n+1)/2+i-1}^{\text{rot}} = \frac{1}{2}(\mathbb{I} + \vec{m}_i \cdot \sigma) \\ W_i^n &= \sum_{j=i}^{j=(n-1)/2+i-1} \Pi_j^{\text{rot}} + \frac{1}{2}\Pi_{i-1}^{\text{rot}} = \frac{1}{2}(\mathbb{I} + \vec{w}_i \cdot \vec{\sigma}). \end{aligned} \tag{75}$$

In total there are $2n$ vectors $\{\vec{m}_i\} \cup \{\vec{w}_i\}$ which forms a $2n$-sided regular polygon with $|\vec{m}_i| = |\vec{w}_i| = \frac{1}{n}\cot(\frac{\pi}{2n})$ and radius

$$r_n = \frac{r_i}{r_c}|\vec{m}_i| = \cos(\frac{\pi}{2n}) \cdot \frac{1}{n}\cot(\frac{\pi}{2n}) = \frac{1}{n}\cot(\frac{\pi}{2n})\cos(\frac{\pi}{2n}). \tag{76}$$

□

From our Table I in the main text, we observe that when n is odd, the rotationally symmetric scheme appears to be optimal. However, for even n , it is consistently suboptimal. This can be quantitatively explained by the fact that the unbiased compatible region $\mathfrak{m}_{\{\Pi_\lambda\}}$ is an n-sided polygon for even n, but a $2n$-sided polygon for odd n. More specifically, when n is even, there is an inherent trade-off between maximizing the number of sides of $\mathfrak{m}_{\{\Pi_\lambda\}}$ and maintaining its symmetry. The rotationally symmetric scheme exhibits the highest degree of symmetry; however, in this case, the unbiased compatible region has only n sides. In contrast, by adopting a slightly asymmetric POVM, we can construct a $2n$-sided $\mathfrak{m}_{\{\Pi_\lambda\}}$, which, while not a regular polygon, results in a larger inradius.

Result 2. Here we provide explicit compatible model (equivalently a local hidden state model) when the parent POVM is rotationally symmetric.

For finite n, with parent POVM $\Pi_\lambda^{\text{rot}} = \frac{1}{n}(\mathbb{I}+\hat{n}_\lambda\cdot\vec{\sigma})$, any arbitrary child POVM $T = \frac{1}{2}(\mathbb{I}+R(\{\Pi_\lambda^{\text{rot}}\})\hat{t}\cdot\vec{\sigma})$ with $\hat{t} = (\cos(\theta), 0, \sin(\theta))^T$ can be simulated as

$$T = \sum_\lambda p_\lambda \Pi_\lambda^{\text{rot}}$$

For even n:

$$p_\lambda = \frac{1 + \frac{\text{sign}(\hat{m}_1\cdot\hat{n}_\lambda)\sin(\frac{\pi}{n}-x)+\text{sign}(\hat{m}_2\cdot\hat{n}_\lambda)\sin(\frac{\pi}{n}+x)}{2\sin(\frac{\pi}{n})}}{2}$$

with $x = \theta + \frac{2\pi k}{n} \in [-\frac{\pi}{n}, \frac{\pi}{n}]$, $\hat{m}_1 = [\cos(\frac{2\pi(k-\frac{1}{2})}{n}), 0, \sin(\frac{2\pi(k-\frac{1}{2})}{n})]^T$, $\hat{m}_2 = [\cos(\frac{2\pi(k+\frac{1}{2})}{n}), 0, \sin(\frac{2\pi(k+\frac{1}{2})}{n})]^T$ and $k \in \mathbb{Z}$.

For odd n:

$$p_\lambda = \frac{1 + \frac{\text{sign}(\hat{m}_1\cdot\hat{n}_\lambda)\sin(\frac{\pi}{2n}-x)+\text{sign}(\hat{m}_2\cdot\hat{n}_\lambda)\sin(\frac{\pi}{2n}+x)}{2\sin(\frac{\pi}{2n})}}{2}$$

with $x = \theta + \frac{\pi k}{n} \in [-\frac{\pi}{2n}, \frac{\pi}{2n}]$, $\hat{m}_1 = [\cos(\frac{\pi(k-\frac{1}{2})}{n}), 0, \sin(\frac{\pi(k-\frac{1}{2})}{n})]^T$, $\hat{m}_2 = [\cos(\frac{\pi(k+\frac{1}{2})}{n}), 0, \sin(\frac{\pi(k+\frac{1}{2})}{n})]^T$ and $k \in \mathbb{Z}$.
When taking $n \to \infty$, we have $\hat{m}_1 = \hat{m}_2 = \hat{t}$ and $\sin(f(x)) \to f(x)$ when $f(x) \to 0$. The above model becomes:

$$p_\lambda = \frac{1 + \frac{\text{sign}(\hat{m}_1\cdot\hat{n}_\lambda)\sin(\frac{\pi}{n}-x)+\text{sign}(\hat{m}_2\cdot\hat{n}_\lambda)\sin(\frac{\pi}{n}+x)}{2\sin(\frac{\pi}{n})}}{2} \longrightarrow \frac{1 + \frac{\text{sign}(\hat{t}\cdot\hat{n}_\lambda)(\frac{2\pi}{n}-x)+\text{sign}(\hat{t}\cdot\hat{n}_\lambda)x}{2\frac{\pi}{n}}}{2} = \frac{1+\text{sign}(\hat{t}\cdot\hat{n}_\lambda)}{2} \tag{77}$$

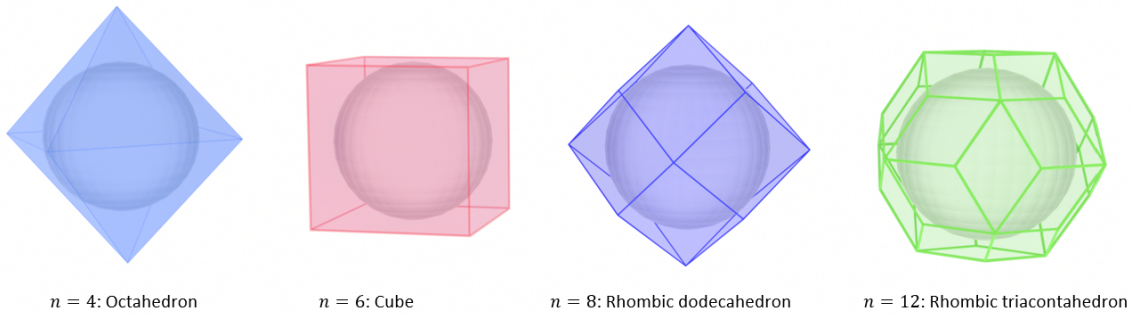

Figure 3: Compatible region for Platonic solid with $n = 4, 6, 8, 12$. From left to right: Octahedron, Cube, Rhombic dodecahedron and Rhombic triacontahedron .

Letting $\lim_{n\to\infty}\sum_{i=1}^{n}(\cdot)\frac{1}{n} = \int_0^{2\pi}(\cdot)\frac{d\psi}{2\pi}$, we can define $\Pi_\psi = \frac{1}{2\pi}(\mathbb{I}+\hat{n}_\psi\cdot\vec{\sigma})$ and $p_\psi = \frac{1+\text{sign}(\hat{t}\cdot\hat{n}_\psi)}{2}$ from the discrete one above to a continuous one, where $\hat{n}_\psi = (\cos(\psi), 0, \sin(\psi))^T$. Therefore, any child POVM $T = \frac{1}{2}(\mathbb{I} + \frac{2}{\pi}\hat{t}\cdot\vec{\sigma})$ can be written as:

$$T = \int_0^{2\pi} d\theta p_\psi \Pi_\phi = \int_0^{2\pi} d\psi \left(\frac{1+\text{sign}(\hat{t}\cdot\hat{n}_\psi)}{2}\right)\frac{1}{2\pi}(\mathbb{I}+\hat{n}_\psi\cdot\vec{\sigma}) = \frac{1}{2}(\mathbb{I}+\frac{2}{\pi}\hat{t}\cdot\vec{\sigma}) \tag{78}$$

Result 3. For a parent POVM $\{\Pi_\lambda\}$ with platonic-solid configuration, we can compute the $R(\{\Pi_\lambda\})$ similarly using Result 1. The polyhedron $\mathcal{M}_{\{\Pi_\lambda\}}$ associated with different platonic configurations are summarized below:

Complexity	POVM $\{\Pi_\lambda\}$	Compatible region $\mathsf{m}_{\{\Pi_\lambda\}}$	compatible radius $R(\{\Pi_\lambda\})$
4	tetrahedron	Octahedron	$\frac{1}{3}$
6	Octahedron	Cube	$\frac{1}{3}$
8	Cube	Rhombic dodecahedron	$\frac{\sqrt{6}}{6}$
12	Icosahedron	Rhombic triacontahedron	$\frac{\phi^3\sqrt{1+(1-\phi)^2}}{3(1+\phi^2)}$
20	Dodecahedron	Rhombic enneacontahedron	$\sqrt{\frac{5}{6}}\frac{\phi^2}{5}$

Table 5: Parent POVM with platonic configuration and their associated unbiased compatible region $\mathsf{m}_{\{\Pi_\lambda\}}$, where $\phi = \frac{1+\sqrt{5}}{2}$ is the golden ratio.

Tetrahedron: With the vertices of an Tetrahedron, we can construct a 4-element parent POVM as follows:

$$\Pi_\lambda^T = \frac{1}{4}(\mathbb{I}+\hat{n}_\lambda\cdot\vec{\sigma}) \quad \text{with} \quad \hat{n}_\lambda \in \left\{\frac{1}{\sqrt{3}}(1,1,1), \frac{1}{\sqrt{3}}(1,-1,-1), \frac{1}{\sqrt{3}}(-1,1,-1), \frac{1}{\sqrt{3}}(-1,-1,1)\right\} \tag{79}$$

The extreme points of the associated compatible region $\mathsf{m}_{\{\Pi_\lambda\}}$ can be computed as:

$$M_\lambda^T = \frac{1}{2}(\mathbb{I}+\vec{m}_\lambda\cdot\vec{\sigma}) \quad \text{with} \quad \vec{m}_\lambda \in \left\{\frac{1}{\sqrt{3}}(\pm1,0,0), \frac{1}{\sqrt{3}}(0,\pm1,0), \frac{1}{\sqrt{3}}(0,0,\pm1)\right\} \tag{80}$$

These 6 vertices altogether form a convex polyhedron – Octahedron, to calculate the inscribed radius, we have:

$$r^T = |\vec{m}_\lambda|\cdot\frac{1}{\sqrt{3}} = \frac{1}{3}$$

Octahedron: With the vertices of a Octahedron, we can construct a 6-element parent POVM as follows:

$$\Pi_\lambda^O = \frac{1}{6}(\mathbb{I} + \hat{n}_\lambda \cdot \vec{\sigma}) \quad \vec{n}_\lambda \in \{(\pm 1, 0, 0), (0, \pm 1, 0), (0, 0, \pm 1)\}, \tag{81}$$

The extreme points of the associate compatible region $\mathsf{m}_{\{\Pi_\lambda^O\}}$ can be computed as:

$$M_\lambda^O = \frac{1}{2}(\mathbb{I} + \vec{m}_\lambda \cdot \sigma) \quad \vec{m}_\lambda \in \{\frac{1}{3}(\pm 1, \pm 1, \pm 1)\}, \tag{82}$$

These 8 vertices altogether form a convex polyhedron – Cube, to calculate the inscribed radius, we have:

$$r^O = |\vec{m}_\lambda| \cdot \frac{1}{\sqrt{3}} = \frac{1}{3}$$

Cube: With the vertices of a Cube, we can construct an 8-element parent POVM as follows:

$$\Pi_\lambda^C = \frac{1}{8}(\mathbb{I} + \hat{n}_\lambda \cdot \vec{\sigma}) \quad \vec{n}_\lambda \in \{\frac{1}{\sqrt{3}}(\pm 1, \pm 1, \pm 1)\} \tag{83}$$

The extreme point of the associated compatible region $\mathsf{m}_{\{\Pi_\lambda^C\}}$ can be computed as:

$$M_\lambda^C = \frac{1}{2}(\mathbb{I} + \vec{m}_\lambda \cdot \sigma) \quad \text{with} \quad \vec{m}_\lambda \in \{\frac{1}{2\sqrt{3}}(\pm 1, \pm 1, \pm 1), \frac{1}{2\sqrt{3}}(\pm 2, 0, 0), \frac{1}{2\sqrt{3}}(0, \pm 2, 0), \frac{1}{2\sqrt{3}}(0, 0, \pm 2)\} \tag{84}$$

These 14 vertices altogether form a convex polyhedron –Rhombic dodecahedron with edge length $a = |\vec{m}_\lambda - \vec{m}_j| = \frac{1}{2}$

$$r^C = \frac{\sqrt{6}}{3} \cdot \frac{1}{2} = \frac{\sqrt{6}}{6} \approx 0.408$$

Icosahedron: With the vertices of an icosahedron, we can construct a 12-element parent POVM as following:

$$\Pi_\lambda^I = \frac{1}{12}(\mathbb{I} + \hat{n}_\lambda \cdot \vec{\sigma}) \quad \text{with} \quad \hat{n}_\lambda \in \left\{\frac{1}{\sqrt{1+\phi^2}}(0, \pm 1, \pm\phi), \frac{1}{\sqrt{1+\phi^2}}(\pm\phi, 0, \pm 1), \frac{1}{\sqrt{1+\phi^2}}(\pm 1, \pm\phi, 0)\right\} \tag{85}$$

where $\phi = \frac{1+\sqrt{5}}{2}$. This time the extreme points of the associate compatible region $\mathsf{m}_{\{\Pi_\lambda\}}$ can be computed as:

$$M_\lambda^I = \frac{1}{2}(\mathbb{I} + \hat{m}_\lambda \cdot \vec{\sigma}) \quad \text{with} \quad \hat{m}_i \in \left\{\frac{\phi}{3\sqrt{1+\phi^2}}(0, \pm 1, \pm\phi), \frac{\phi}{3\sqrt{1+\phi^2}}(\pm\phi, 0, \pm 1), \frac{\phi}{3\sqrt{1+\phi^2}}(\pm 1, \pm\phi, 0)\right\}$$
$$\bigcup \left\{\frac{\phi}{3\sqrt{1+\phi^2}}(\pm 1, \pm 1, \pm 1), \frac{\phi}{3\sqrt{1+\phi^2}}(0, \pm\phi, \pm 1/\phi), \frac{\phi}{3\sqrt{1+\phi^2}}(\pm 1/\phi, 0, \pm\phi), \frac{\phi}{3\sqrt{1+\phi^2}}(\pm\phi, \pm 1/\phi, 0)\right\} \tag{86}$$

These 32 vertices altogether form a convex polyhedron – Rhombic triacontahedron, to calculate the inscribed radius, we first notice that the edge length of Rhombic triacontahedron given as $a = |\vec{m}_\lambda - \vec{m}_j| = \frac{\phi}{3\sqrt{1+\phi^2}}\sqrt{1+(1-\phi)^2}$, and the inscribed radius is $r = \frac{\phi^2}{\sqrt{1+\phi^2}}a$. Therefore, the radius of the shrinking Bloch sphere built by parent POVM $\{\Pi_i^I\}$ will be

$$r^I = \frac{\phi}{3\sqrt{1+\phi^2}}\sqrt{1+(1-\phi)^2} \cdot \frac{\phi^2}{\sqrt{1+\phi^2}} \approx 0.4588$$

Dodecahedron: With the vertices of a dodecahedron, we can construct a 20-element parent POVM as follows:

$$\begin{aligned}\Pi_\lambda^D = \frac{1}{20}(\mathbb{I} + \hat{n}_\lambda \cdot \vec{\sigma}) \quad &\text{with} \quad \hat{n}_\lambda \in \left\{\frac{1}{\sqrt{3}}(\pm 1, \pm 1, \pm 1)\right\} \\ \text{or with} \quad &\hat{n}_\lambda \in \left\{\frac{1}{\sqrt{1+\phi^4}}(0, \pm 1, \pm\phi^2), \frac{1}{\sqrt{1+\phi^4}}(\pm\phi^2, 0, \pm 1), \frac{1}{\sqrt{1+\phi^4}}(\pm 1, \pm\phi^2, 0)\right\}\end{aligned} \tag{87}$$

The extreme points of the associate compatible region $\mathsf{m}_{\{\Pi_\lambda^D\}}$ can be computed as:

$$
\begin{aligned}
M_\lambda^D = \frac{1}{2}(\mathbb{I} + \hat{m}_\lambda \cdot \vec{\sigma}) \quad \text{with} \quad \hat{m}_i \in & \left\{ \frac{2(\phi^2+1)}{10\sqrt{3}\phi}(0, \pm\phi, \pm 1), \frac{2(\phi^2+1)}{10\sqrt{3}\phi}(\pm\phi, \pm 1, 0), \frac{2(\phi^2+1)}{10\sqrt{3}\phi}(0 \pm 1, 0, \pm\phi) \right\} \\
& \bigcup \left\{ \frac{2\phi^2}{10\sqrt{3}}(\pm 1, \pm 1, \pm 1), \frac{2\phi^2}{10\sqrt{3}}(0, \pm 1/\phi, \pm\phi), \frac{2\phi^2}{10\sqrt{3}}(\pm\phi, 0, \pm 1/\phi), \frac{2\phi^2}{10\sqrt{3}}(\pm 1/\phi, \pm\phi, 0) \right\} \\
& \bigcup \left\{ \frac{2\phi^2}{10\sqrt{3}}(\pm 2/\phi, \pm 1/\phi^2, \pm 1), \frac{2\phi^2}{10\sqrt{3}}(\pm 1/\phi^2, \pm 1, \pm 2/\phi), \frac{2\phi^2}{10\sqrt{3}}(\pm 1, \pm 2/\phi, \pm 1/\phi^2) \right\} \\
& \bigcup \left\{ \frac{2\phi^2}{10\sqrt{3}}(\pm\phi, \pm 1/\phi, \pm 1/\phi), \frac{2\phi^2}{10\sqrt{3}}(\pm 1/\phi, \pm\phi, \pm 1/\phi), \frac{2\phi^2}{10\sqrt{3}}(\pm 1/\phi, \pm 1/\phi, \pm\phi) \right\} \\
& \bigcup \left\{ \frac{1}{10\sqrt{3}}(\pm 2(\phi^2+1), \pm 2(\phi-1), 0), \frac{1}{10\sqrt{3}}(\pm 2(\phi-1), \pm 2(\phi^2+1), 0), \frac{1}{10\sqrt{3}}(0, \pm 2(\phi^2+1), \pm 2(\phi-1)) \right\}
\end{aligned}
\tag{88}
$$

These 92 vertices altogether form a convex polyhedron – Rhombic enneacontahedron, the inradius of which can be computed as:

$$r^D = \sqrt{\frac{5}{6}}\frac{\phi^2}{5} = 0.4780$$

F Simulation cost of entangled state and compatible measurements

It is well known that a noisy two-qubit Werner state is unsteerable under projective measurements (PVMs) iff the corresponding family of noisy qubit spin observables with the same visibility is jointly measurable [27, 26]. We establish an analogous correspondence here: (i) an LHS model for the noisy two-qubit Werner state under PVMs is equivalent to (ii) a compatibility model (parent POVM plus classical post-processing) for all noisy qubit spin measurements with the same noise parameter, under a fixed shared randomness budget. The proof is similar to that of [27, 26]; we include it below for completeness.

Lemma 1. $\mathcal{P}_r$ can be simulated with an n-element POVM if and only if the simulation cost of Werner states ω_r under *projective measurements* equals n.

Proof. Assume $\mathcal{P}_r$ is simulated by an n-outcome parent POVM $\{\Pi_\lambda\}_{\lambda=1}^n$ with classical post-processing, then any noisy spin measurement $M_{\pm|\hat{n}}^{(r)} = rM_{\pm|\hat{n}} + \frac{1}{2}(1-r)\mathbb{I}$ can be simulated as:

$$M_{\pm|\hat{n}}^{(r)} = \sum_{\lambda=1}^{n} p(\pm|\hat{n}, \lambda)\, \Pi_\lambda, \quad \sum_{\pm} p(\pm|\hat{n}, \lambda) = 1. \tag{89}$$

Therefore, for any noisy Werner state $\omega_r = r|\Psi^-\rangle\langle\Psi^-| + \frac{1}{4}(1-r)\mathbb{I}\otimes\mathbb{I}$ the assemblage on B prepared with noisy spin measurement is

$$\sigma_{\pm|\hat{n}}^{(r)} = \mathrm{Tr}_A[(M_{\pm|\hat{n}}\otimes\mathbb{I})\omega_r] = \mathrm{Tr}_A[(M_{\pm|\hat{n}}^{(r)}\otimes\mathbb{I})|\Psi^-\rangle\langle\Psi^-|] = \sum_{\lambda=1}^{n} p(\pm|\hat{n},\lambda)\,\mathrm{Tr}_A[(\Pi_\lambda\otimes\mathbb{I})|\Psi^-\rangle\langle\Psi^-|]. \tag{90}$$

Let $\rho_\lambda := \mathrm{Tr}_A[(\Pi_\lambda\otimes\mathbb{I})|\Psi^-\rangle\langle\Psi^-|]$. Then $\sigma_{\pm|\hat{n}}^{(r)} = \sum_\lambda p(\pm|\hat{n},\lambda)\rho_\lambda$, so $\{\rho_\lambda\}$ is a valid LHS ensemble. Hence the Werner assemblage admits an LHS model with n hidden states (shared randomness cost $\log_2 n$ bits).

Conversely, suppose the Werner assemblage admits an LHS model with n hidden states: $\sigma_{\pm|\hat{n}}^{(r)} = \sum_\lambda p(\pm|\hat{n},\lambda)\,\rho_\lambda$, with $\rho_\lambda \geq 0$ and $\sum_\lambda \rho_\lambda = \frac{1}{2}\mathbb{I}$. From the identity $\mathrm{Tr}_A[(X\otimes\mathbb{I})\,|\Psi^-\rangle\langle\Psi^-|] = \frac{1}{2}\sigma_y X^T \sigma_y$, we have

$$\frac{1}{2}\sigma_y (M_{\pm|\hat{n}}^{(r)})^T \sigma_y = \sum_\lambda p(\pm|\hat{n},\lambda)\,\rho_\lambda \Rightarrow \frac{1}{2}M_{\pm|\hat{n}}^{(r)} = \sum_\lambda p(\pm|\hat{n},\lambda)(\sigma_y\rho_\lambda\sigma_y)^T. \tag{91}$$

Therefore, we can define $\Pi_\lambda := 2\left(\sigma_y \rho_\lambda \sigma_y\right)^T \geq 0$, which satisfies

$$M^{(r)}_{\pm|\hat{n}} = \sum_\lambda p(\pm|\hat{n}, \lambda)\Pi_\lambda, \qquad \sum_\lambda \Pi_\lambda = 2\Big(\sum_\lambda \sigma_y\, \rho_\lambda\, \sigma_y\Big)^{\mathsf{T}} = \mathbb{I},$$

so $\{\Pi_\lambda\}$ is an n-outcome parent POVM that simulates $\mathcal{P}_r$.

Hence the minimal number of parent outcomes needed to simulate $\mathcal{P}_r$ equals the minimal number of hidden states in an LHS model for ω_r under PVMs; equivalently, both simulation costs are n (or $\log_2 n$ shared bits). □

G Discussion on prop. 1, prop. 2 and conj. 1

1 Inequivalence between simulation cost of PVMs and POVMs

In an earlier paper[34, 35], the equivalence of POVMs and PVMs in the context of quantum steering was shown, which is by establishing that a compatible model exists for all noisy POVMs whenever such a model exists for all noisy PVMs at the same noise threshold. However, this elegant equivalence disappears when considering scenarios that involve only finite-shared randomness and specific parent POVMs (or specific local hidden states).

In the main text, we conduct an extensive discussion on the feasibility of a compatible model for noisy PVMs with finite shared randomness. Notably, one intriguing observation we make there is that the optimal parent POVM for simulating the entire set of noisy PVMs may not necessarily be centrally symmetric, even if it does exist. The realization that asymmetry has alue in simulating all PVMs provides insight into the potential divergence between POVMs and PVMs within this restricted compatible model.

In the following sections, we provide a rigorous proof for the aforementioned assertion using Farka's lemma. To cast the simulation problem as a feasibility test, represent each rank-1 qubit parent effect $\Pi_\lambda = \alpha_\lambda\left(\mathbb{I} + \hat{n}_\lambda \cdot \vec{\sigma}\right)$ by the Pauli-basis 4-vector $\vec{\pi}_\lambda = (\alpha_\lambda, \alpha_\lambda \hat{n}_\lambda)^\top \in \mathbb{R}^4$, and each target (child) effect $M_a = \mu_a\left(\mathbb{I} + r\hat{m}_a \cdot \vec{\sigma}\right)$ by $\vec{m}_a = \mu_a(1, r\hat{m}_a)^\top$. Joint simulability with post-processing coefficients $p_{a|\lambda} \geq 0$ is exactly the set of linear constraints

$$\sum_{\lambda=1}^{n} p_{a|\lambda}\, \vec{\pi}_\lambda = \vec{m}_a \quad (\forall a), \qquad \sum_{a=1}^{m} p_{a|\lambda} = 1 \quad (\forall \lambda).$$

Stacking all outcomes yields a single system $A\mathbf{p} = \mathbf{b}$ with $A \in \mathbb{R}^{(4m+n)\times(mn)}$ formed by m blocks $[\vec{\pi}_1 \ \cdots \ \vec{\pi}_n]$ (one per a) and n normalization rows, $\mathbf{p} = \left(p_{1|1}, \ldots, p_{m|1}, p_{1|2}, \ldots, p_{m|n}\right)^\top \in \mathbb{R}^{mn}$, and $\mathbf{b} = (\vec{m}_1, \ldots, \vec{m}_m, 1, \ldots, 1)^\top \in \mathbb{R}^{4m+n}$. Specifically:

$$A = \begin{pmatrix} \vec{\pi}_1 & \cdots & \vec{\pi}_n & 0 & \cdots & 0 & 0 & 0 & 0 \\ & \ddots & & & \ddots & & & \ddots & \\ 0 & 0 & 0 & 0 & \cdots & 0 & \vec{\pi}_1 & \cdots & \vec{\pi}_n \\ 1 & \cdots & 0 & 1 & \cdots & 0 & 1 & \cdots & 0 \\ & \ddots & & & \ddots & & & \ddots & \\ 0 & \cdots & 1 & 0 & \cdots & 1 & 0 & \cdots & 1 \end{pmatrix}, \qquad \boldsymbol{p} = \begin{pmatrix} p_{1|1} \\ \vdots \\ p_{1|n} \\ p_{2|n} \\ \vdots \\ p_{m|n} \end{pmatrix}, \qquad \boldsymbol{b} = \begin{pmatrix} \vec{m}_1 \\ \vdots \\ \vec{m}_m \\ 1 \\ \vdots \\ 1 \end{pmatrix}. \tag{92}$$

Lemma 2 (Farkas' Lemma)**.**

$$\nexists \boldsymbol{p} \geq \mathbf{0} \text{ s.t. } A \cdot \boldsymbol{p} = \boldsymbol{b} \iff \exists y \text{ s.t. } A^T \cdot \boldsymbol{y} \geq \mathbf{0} \text{ and } \boldsymbol{b}^T \cdot \boldsymbol{y} < 0 \tag{93}$$

where $\boldsymbol{y} \in \mathbb{R}^{4m+n}$ as $y = (\vec{y}_1, \cdots, \vec{y}_m, z_1, \cdots, z_n)^T$.

Therefore, to demonstrate the non-existence of a positive, normalized response function $p_{a|i}$ for simulating $\{\widetilde{M}_a\}$ with $\{\Pi_\lambda\}_{\lambda=1}^n$, it suffices to find a vector $\boldsymbol{y}$ such that $A^T \cdot \boldsymbol{y} \geq \boldsymbol{0}$ and $\boldsymbol{b}^T \cdot \boldsymbol{y} < 0$ holds at the same time.

Proposition 6. There exists a five-effect parent POVM $\{\Pi_\lambda\}_{\lambda=1}^5$ with $R(\{\Pi_\lambda\}) \approx 0.3714$, that is, it can simulate all noisy spin measurements with radius $r \leq 0.3714$, whereas a compatible model fails to exist for some three-outcome noisy POVMs with radius $r > 0.3220$

Proof. The 5-effect POVM $\{\Pi_\lambda = \alpha_\lambda(\mathbb{I} + \hat{n}_\lambda \cdot \sigma)\}$ to be considered is of the form of:

$$\begin{aligned}
&\vec{\pi}_1 = 0.220(1, \hat{n}_3)^T \quad \hat{n}_1 = (-\frac{12}{55}, \sqrt{1-(\frac{12}{55})^2}, 0)^T \\
&\vec{\pi}_2 = 0.220(1, \hat{n}_4)^T \quad \hat{n}_2 = (-\frac{12}{55}, -\frac{1}{2}\sqrt{1-(\frac{12}{55})^2}, \frac{\sqrt{3}}{2}\sqrt{1-(\frac{12}{55})^2})^T \\
&\vec{\pi}_3 = 0.220(1, \hat{n}_5)^T \quad \hat{n}_3 = (-\frac{12}{55}, -\frac{1}{2}\sqrt{1-(\frac{12}{55})^2}, -\frac{\sqrt{3}}{2}\sqrt{1-(\frac{12}{55})^2})^T \\
&\vec{\pi}_4 = 0.242(1, \hat{n}_1)^T \quad \hat{n}_4 = (1, 0, 0)^T \\
&\vec{\pi}_5 = 0.098(1, \hat{n}_2)^T \quad \hat{n}_5 = (-1, 0, 0)^T
\end{aligned} \tag{94}$$

First, we used the criteria for computing the compatible radius $R(\{\Pi_\lambda\})$ in appendix D to obtain a close form expression:

$$R(\{\Pi_\lambda\}) = 0.34 + 0.144\frac{12}{55} \approx 0.3714. \tag{95}$$

Therefore, any PVMs with visibility $r \leq 0.3714$ can be simulated by the corresponding 5-outcome POVM.

Now, we give examples showing the infeasibility of simulating three-outcome POVMs $\{M_a^r\}$ with this five-effect parent POVM for an even smaller threshold $r < R(\{\Pi_\lambda\})$. The specific three-outcome POVM $\{M_a^r\}$ we consider is given by the four-vectors

$$\begin{aligned}
&\vec{m}_1 = \frac{1}{3}(1, r\hat{m}_1)^T \quad \hat{m}_1 = (0, -1, 0)^T \\
&\vec{m}_2 = \frac{1}{3}(1, r\hat{m}_2)^T \quad \hat{m}_2 = (0, \frac{1}{2}, \frac{\sqrt{3}}{2})^T \\
&\vec{m}_3 = \frac{1}{3}(1, r\hat{m}_3)^T \quad \hat{m}_3 = (0, \frac{1}{2}, -\frac{\sqrt{3}}{2})^T.
\end{aligned} \tag{96}$$

By considering vector $\boldsymbol{y} \in \mathbb{R}^{4\times 3+5}$ as $y = (\vec{y}_1, \vec{y}_2, \vec{y}_3, z_1, z_2, z_3, z_4, z_5)^T$, where:

$$\begin{aligned}
&\vec{y}_1 = (0, -\hat{m}_1)^T, \quad \vec{y}_2 = (0, -\hat{m}_2)^T \quad \vec{y}_3 = (0, -\hat{m}_3)^T \\
&z_1 = z_2 = z_3 = 0.110\sqrt{1-(\frac{12}{55})^2}, \quad z_4 = z_5 = 0,
\end{aligned} \tag{97}$$

and $A^T \cdot \boldsymbol{y} \geq 0$, one can easily verify that with $\boldsymbol{b} = (\vec{m}_1, \vec{m}_2, \vec{m}_3, 1, 1, 1, 1, 1)^T$:

$$\boldsymbol{b}^T \cdot \boldsymbol{y} = -r + 0.330\sqrt{1-(\frac{12}{55})^2} < 0 \Longrightarrow r > 0.330\sqrt{1-(\frac{12}{55})^2} \approx 0.3220 \tag{98}$$

Therefore, from farkas' lemma, for $r > 0.3220$, the three-outcome POVM defined in Eq. (96) can never be simulated by the five-effect POVM $\{\Pi_\lambda\}$, whereas all PVMs can be simulated by it with $r < 0.3714$.

□

Conjecture 1. The 5-outcome POVM in Eq. (94) is near optimal in terms of simulating PVMs, i.e, close to the optimal $\{\Pi_\lambda\}$ that has the largest compatible radius $R(\{\Pi_\lambda\}) \approx 0.3718$ reported in the main text. Therefore, we conjecture that the set of noisy POVMs and noisy PVMs has different simulation costs.

However, as noted in Proposition 2 in the main text, there exist special cases where the simulation cost of a set of noisy qubit PVMs and a set of noisy qubit POVMs is actually the same.

To continue our discussion, we define a few quantities generalized from the main text:
(1) $N^p_{\mathrm{POVM}}(r)$: the simulation cost of noisy planar qubit POVMs with visibility r ;
(2) $R^p_{\mathrm{POVM}}(n)$ he compatibility radius for simulating noisy planar qubit POVMs with n-outcome measurements.

The following lemma, which will be critical for the subsequent proposition, is stated and proved below:

Lemma 3. $R^p_{\mathrm{POVM}}(3) = R^p(3) = \frac{1}{2} \Leftrightarrow \gamma^p(\omega_{\frac{1}{2}}) = N^p(\frac{1}{2}) = 3$

Proof. We first notice that, $R^p_{\mathrm{POVM}}(3)$ has to be obtained by 3-outcome parent $\{\Pi_\lambda\}$ with linear independent effects since $R^p(3) = 0$ otherwise.

$R^p_{\mathrm{POVM}}(3) \leq R^p(3)$ is obvious since PVMs form a subset of POVMs.

To show $R^p_{\mathrm{POVM}}(3) \geq R^p(3)$, for planar POVM $\{M^r_a\}$, it suffices to consider three outcome measurements since they are extreme in the planar case, given a fixed noise threshold r, if all noisy 3-outcome qubit planar measurement can be simulated, then all noisy qubit planar POVMs can be simulated. Each of its effects M^r_a can be simulated individually by $\{\Pi_\lambda\}$ for all $r \leq R^p(3)$, therefore there exist $0 \leq p_{a|\lambda} \leq 1$ such that $M^r_a = \sum_{\lambda=1}^3 p_{a|\lambda}\Pi_\lambda$.

Now considering the two normalization condition: $\sum_{a=1}^3 M^r_a = \sum_{\lambda=1}^3 \Pi_\lambda = \mathbb{I}$, we have:

$$\sum_{\lambda=1}^{3}(1 - \sum_{a} p_{a|\lambda})\Pi_\lambda = 0 \tag{99}$$

The linearly independence of $\{\Pi_\lambda\}$ implies $\sum_a p_{a|\lambda} = 1$, thus $\{M^r_a\}_a$ can be simultaneous simulated by $\{\Pi_\lambda\}$, and therefore, $R^p_{\mathrm{POVM}}(3) \geq R^p(3)$. We thus have $R^p_{\mathrm{POVM}}(3) = R^p(3) = \frac{1}{2}$ which implies $\gamma(\rho_W(\frac{1}{2})) = N^p(\frac{1}{2})$ using duality between compatibility of noisy spin measurements and steerability of Werner state as discussed in section F. □

Similarly one could show that

Corollary 5. $R_{\mathrm{POVM}}(4) = R(4) = \frac{1}{3} \Leftrightarrow \gamma(\rho_W(\rho_W(\frac{1}{3})) = N(\frac{1}{3}) = 4$

Proposition 7. $R^p_{\mathrm{POVM}}(n) = R^p(n) \Leftrightarrow \gamma^p(\omega_r) = N^p(r)$.

Proof. Without loss of generality, we always assume $\{\Pi_\lambda\}_{\lambda=1}^n$ is rank-1 and with pairwise distinct effects. Consequently, any three of its effects are linearly independent. Moreover, since these effect are planar effect (i.e. with Bloch vector on the xz plane), any four of them must be linear dependent.

Suppose each effect of the spin measurement (an instance of 2-outcome measurement) $\{M^{(r)}_a, \mathbb{I} - M^{(r)}_a\}$ admits a simulation

$$M^{(r)}_a = \sum_{\lambda=1}^{n} p_{a|\lambda}\,\Pi_\lambda, \qquad p_{a|\lambda} \in [0,1],$$

with no assumption that $\sum_a p_{a|\lambda} = 1$. Nevertheless, since $\sum_{\lambda=1}^n \Pi_\lambda = \sum_{a=1}^3 M^{(r)}_a = \mathbb{I}$, we always have the operator identity

$$\sum_{\lambda=1}^{n}(s_\lambda - 1)\Pi_\lambda = 0, \qquad \text{where} \quad s_\lambda := \sum_{a=1}^{3} p_{a|\lambda}.$$

Thus it suffices to adjust the coefficients to obtain a single column-stochastic post-processing $p_{a|\lambda}$ with $\sum_a p_{a|\lambda} = 1$ for all λ while keeping each $M^{(r)}_a$ unchanged. (Here λ indexes columns, and a indexes raw.)

Case $n = 3$**.** By linear independence of $\{\Pi_\lambda\}_{\lambda=1}^3$ the equality $\sum_\lambda (s_\lambda - 1)\Pi_\lambda = 0$ forces $s_\lambda = 1$ for each λ. Hence $p_{a|\lambda}$ is already normalized (column-stochastic)

Case $n = 4$**.** There must exist a non-trivial linear dependent relation

$$\sum_{\lambda=1}^{4} q_\lambda \, \Pi_\lambda = 0 \quad \text{with} \quad q_\lambda := s_\lambda - 1 = \sum_a p_{a|\lambda} - 1. \tag{100}$$

Since $\sum_\lambda q_\lambda \operatorname{tr}(\Pi_\lambda) = 0$ and $\operatorname{tr}(\Pi_\lambda) > 0$, both $\{q_\lambda > 0\}$ and $\{q_\lambda < 0\}$ are nonempty. Moreover, each set should have exactly two elements, otherwise, one effect equals the mixture of the other three and can not be rank-1. Without loss of generality, we assume $q_\lambda > 0$ (or equivalently $s_\lambda > 1$) for $\lambda \in \{1, 2\}$.

Because there are only three raw $a \in \{1, 2, 3\}$, the condition $s_\lambda = \sum_a p_{a|\lambda} > 1$ for $\lambda \in \{1, 2\}$ implies the existence of a raw a such that $p_{a|1}, p_{a|2} > 0$ (as $0 \le p_{a|\lambda} \le 1$).

Now, we can apply the linear dependent relation in Eq 100 to shift the element $p_{a|\lambda}$ in row a by:

$$p'_{a|\lambda}(\beta) := p_{a|\lambda} - \beta \, q_\lambda, \qquad \beta \ge 0$$

to drive $s'_\lambda := \sum_a p'_{a|\lambda}(\beta)$ for $\lambda \in \{1, 2\}$ towards 1.

Because $\sum_\lambda q_\lambda \Pi_\lambda = 0$, we have that $\sum_\lambda p'_{a|\lambda}(\beta)\Pi_\lambda = \sum_\lambda p_{a|\lambda}\Pi_\lambda = M_a^{(r)}$, so the simulation of effect $M_a^{(r)}$ is preserved. Let $\beta^\star$ be the largest value for which $p'_{a|\lambda}(\beta) \ge 0$ for all λ. At $\beta^\star$ one of the following occurs:

(i) For at least one λ with $q_\lambda > 0$, the column sum hits 1: $s'_\lambda(\beta^\star) = s_\lambda - \beta^\star q_\lambda = 1$; or

(ii) Some entry in row a hits zero: $p'_{a|\lambda}(\beta^\star) = 0$ for some λ.

In case (i) it strictly fixes one column ($s_\lambda = 1$) and problem reduce to the case with $n = 3$, thus the column stochastic condition is met.

In case (ii) we have reduced the support of row a while preserving all effects. Now, pick another row a' that $p_{a'|\lambda} > 0$ for $\lambda \in \{1, 2\}$ (such a row must exist, otherwise s_1 and s_2 could not exceed 1), and repeat the same update with row a'.

Each iteration either fixes a column or strictly reduces a row's support, so the process terminates only with $s_\lambda = 1$ for all λ. The final coefficients $p_{a|\lambda}$ are column-stochastic and still reproduce each $M_a^{(r)}$.

Case $n > 5$**.**: Assuming $\sum_a p_{a|\lambda} \ne 1$ for all λ. We can divide these effects in two indexes set $\mathcal{J}$ and $\mathcal{K}$ based on the sign of q_λ, and thus there must exist two pairs of Π_λ such that

$$\sum_\lambda q_\lambda \Pi_\lambda = \sum_{j \in \mathcal{J}} q_j \Pi_j - \sum_{k \in \mathcal{K}} q_k \Pi_k = t_{j_1} \Pi_{j_1}^p + t_{j_2} \Pi_{j_2}^p - t_{k_1} \Pi_{k_1}^p - t_{k_2} \Pi_{k_2}^p = 0,$$

with $t_\nu > 0$. This is possible since $\{\Pi_\lambda\}$ is assumed to be a rank-1 POVM with distinct effects, and we explained this in more detail in the lemma below.

With this linear dependent relation, we can follow the same procedure as in $n = 4$ iteratively and smooth $\sum_a p_{a|i} = 1$ one by one until we get $n = 3$

□

Lemma 4. Let $\{\Pi_\lambda = \alpha_\lambda(\mathbb{I} + \hat{n}_\lambda \cdot \vec{\sigma})\}$ be rank-1 qubit effects whose Bloch directions $\{\hat{n}_\lambda\}$ are coplanar and pairwise distinct. Let $\mathcal{J}, \mathcal{K}$ be disjoint triples of indices and suppose there exist strictly positive weights q_λ such that

$$\sum_{j \in \mathcal{J}} q_j \, \Pi_j \;=\; \sum_{k \in \mathcal{K}} q_k \, \Pi_k. \tag{101}$$

Then there exist distinct $j_1 \ne j_2 \in \mathcal{J}$, $k_1 \ne k_2 \in \mathcal{K}$ and $t_{j_1}, t_{j_2}, t_{k_1}, t_{k_2} > 0$ with

$$t_{j_1} \, \Pi_{j_1} + t_{j_2} \, \Pi_{j_2} \;=\; t_{k_1} \, \Pi_{k_1} + t_{k_2} \, \Pi_{k_2}, \tag{102}$$

Proof. From (101), equality of identity and Bloch parts gives

$$\sum_{j\in\mathcal{J}} q_j\alpha_j = \sum_{k\in\mathcal{K}} q_k\alpha_k \quad \text{and} \quad \sum_{j\in\mathcal{J}} q_j\alpha_j\hat{n}_j = \sum_{k\in\mathcal{K}} q_k\alpha_k\hat{n}_k \;=:\; \vec{v}. \tag{103}$$

Thus $\frac{\vec{v}}{\sum_{j\in\mathcal{J}} q_j\alpha_j} \in \mathrm{conv}\{\hat{n}_j : j \in \mathcal{J}\} \cap \mathrm{conv}\{\hat{n}_k : k \in \mathcal{K}\}$. Each of these convex hulls is a polygon inscribed in the same Bloch circle. Since all vertices lie on the circle and the vertex sets are disjoint, no vertex of one polygon is contained in the boundary or (open) interior of the other. Therefore every vertex of the intersection $\mathrm{conv}\{\hat{n}_j : j \in \mathcal{J}\} \cap \mathrm{conv}\{\hat{n}_k : k \in \mathcal{K}\}$ is an intersection of an edge from the first polygon with an edge from the second. Pick such a vertex and denote it by $\vec{w}$. Then there exist distinct $j_1 \neq j_2 \in \mathcal{J}$, $k_1 \neq k_2 \in \mathcal{K}$ and parameters $t, s \in (0, 1)$ with

$$\vec{w} \;=\; t\hat{n}_{j_1} + (1-t)\hat{n}_{j_2} \;=\; s\hat{n}_{k_1} + (1-s)\hat{n}_{k_2}. \tag{104}$$

Setting $t_{j_1} = \frac{t}{\alpha_{j_1}}$, $t_{j_2} = \frac{1-t}{\alpha_{j_2}}$, $t_{k_1} = \frac{s}{\alpha_{k_1}}$, $t_{k_2} = \frac{1-s}{\alpha_{k_2}}$ (all > 0), both sides of (102) equal $\mathbb{I} + \vec{w}\cdot\vec{\sigma}$. □