

# From information tasks to exclusivity and Bell bounds

Samuel Mausberg

Independent Researcher  
3 October 2026

**Can possible information tasks explain the quantum bound on Bell correlations? We obtain a conditional bound, without assuming a quantum state space, and identify the additional task requirements that recover the quantum value. Homogeneous frequency limits and common calibration domains imply consistent exclusivity. Completing two-copy parity measurements with an “other” outcome gives a Bell bound controlled by the weights of these residual records. Marletto’s T2 supplies four aligned parity sectors. A missing domain then requires an extra distinguishable label. An upper bound on readable labels, stable under adjoining spectators, and a reversible setting interchange remove the residual records. No-hypersignaling supplies this capacity bound when the relevant preparations are accessible and local signaling dimensions match the label counts. The result separates measurement existence, domain coverage, and readout weights.**

## 1 Introduction

A Bell experiment has two separated laboratories, each choosing between two binary measurements. Quantum theory bounds the Clauser–Horne–Shimony–Holt (CHSH) correlation expression by  $2\sqrt{2}$ , although its algebraic maximum is four. Can constructor theory explain the smaller number through laws about possible and impossible information tasks [1, 2]? We prove a conditional bound and isolate what its domain premise must accomplish. The issue is not the existence of a discriminator for joint parity labels. It is whether those labels cover the preparation used in the experiment.

The answer can be stated operationally. Add an “other” record to each of the aligned and crossed two-copy parity discriminators. If their

respective weights are  $\varepsilon_0$  and  $\varepsilon_1$ , then our task premises imply

$$B^2 \leq 8 + 4(\varepsilon_0 + \varepsilon_1) + 288\kappa, \quad (1)$$

where  $\kappa$  bounds discrepancies between readouts with the same calibration. Homogeneous frequency tasks make  $\kappa$  vanish. Parity-domain coverage makes both residual records vanish. Equation (1) then gives the sharp quantum bound, with no quantum state space in the proof.

The result also gives a way to test the missing domain condition. T2 supplies four distinguishable aligned parity sectors. If there is no room for an additional distinguishable label, these sectors cover the substrate. The same argument must remain valid when another system—a spectator—is attached. A reversible task interchanging the two tested settings then supplies crossed coverage. We prove this capacity criterion, but do not derive its full-substrate capacity bound or setting interchange from T2.

A pair of qubits provides a running example. Fix the second qubit to  $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$  and leave the first unspecified. The sign of the second qubit is a perfectly readable and copyable label, yet the frequencies of the total computational-bit measurement depend on the first qubit. Specifying the preparation resolves that ambiguity. On two completely specified qubits, the four Bell labels exhaust the distinguishable information and jointly record two parities. These are different requirements: a frequency domain must be homogeneous, while a parity domain must be large enough.

We present three results. The first turns an operational frequency condition into calibrated weight neutrality and consistent exclusivity (CE). The second proves Eq. (1) by keeping the residual records in Cabello’s two-copy construction [5]. The third expresses parity coverage as a consequence of information-capacity saturation and a reversible setting-interchange task. The proofs

allow arbitrary state-space dimension and do not use an inner product, local tomography, or a rule assigning Born probabilities.

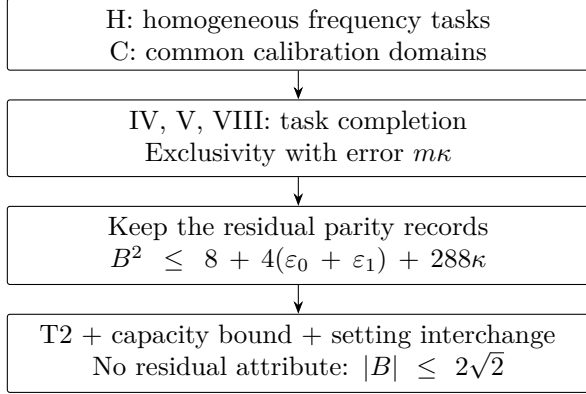


Figure 1: The three steps, under the record-weight and composition conventions of Sec. 2. H and C connect task calibration to weights. Completing a joint parity discriminator retains a possible residual record, whose weight appears in the Bell bound. Capacity saturation and reversible setting interchange remove it.

The relation between nonlocal states and joint measurements is established prior work. Short and Barrett identify the trade-off in boxworld [12]. Czekaj, Horodecki, and Tylec constrain correlations by adding a Bell measurement [13]. Cabello’s extra two-copy contexts supply our final counting argument [5]. The contribution here is a finite-resolution connection from task calibration to weights, a bound retaining unresolved-domain records, and a distinguishable-label certificate for those domains.

The last certificate has a precise relation to no-hypersignaling [14]. Recording distinct labels is a narrower task than simulating every prepare-and-read distribution. We show how no-hypersignaling implies the needed capacity bound when local signaling dimensions match the available label counts and the witnessing attributes are preparable. An exact comparison with the published hypersignaling model shows why those qualifications matter. Section 7 places these results against sharpness, exclusivity, and parity-reading approaches.

## 2 Tasks, preparations, and decision weights

Locators for Refs. [1, 2] refer to arXiv versions v2 and v3, respectively, using their printed page

numbers and named conditions. In particular, a condition on a selected observable is not a condition on every laboratory setting.

### 2.1 The task principles used

An attribute is a set of states. A task specifies an output attribute for every legitimate input attribute, and must work throughout that input set [2, Sec. 2, pp. 8–11]. We write  $a \perp b$  when a possible discrimination task distinguishes them. Possible means arbitrarily accurately achievable, with no finite bound on the required accuracy [1, Sec. 2, p. 8]. Distinguishability is symmetric, is inherited by subsets, and no nonempty attribute is distinguishable from itself.

The information principles are imposed on every finite composite. Interoperability allows independent information variables and their records to be combined; generic resources and task composition allow finite recording and counting networks [1, Sec. 6, III, VI, VII, pp. 12–14]. Principle IV supplies a common discriminator for any finite pairwise distinguishable variable. Principle V upgrades statewise distinguishability to distinguishability of the whole attribute [1, Sec. 6, p. 13]. Thus, with  $\Omega$  the state set of the substrate,

$$\bar{a} = \{s \in \Omega : \{s\} \perp a\}, \quad \text{cl}(a) := \bar{\bar{a}} \quad (2)$$

is an orthogonality closure. An observable attribute is closed under this operation. We use the finite-task portion of the information principles. Principle IX concerns ensemble discrimination and is not needed in these finite constructions [1, Sec. 6, p. 19]. The local-description condition II is not identified with operational local tomography [2, Sec. 2, p. 12].

For a finite measurable variable  $X = \{x_i\}$ , let

$$D_X = \text{cl}\left(\bigcup_i x_i\right). \quad (3)$$

A measurer is calibrated by the task  $x_i \mapsto$  record  $i$ . Nonperturbation, when required, means retention of the input attribute, not preservation of every compatible readout [1, Sec. 5, Eqs. (4)–(5), pp. 11–12]. Principle VIII says that all measurers of  $X$  agree on every sharp output on  $D_X$  [1, Sec. 6, p. 16]. Its observable-closure consequence is that a coarse detector for an outcome

$x_i$  is sharp on  $\text{cl}(x_i)$  [1, Sec. 6, pp. 16–18]. Appendix A states the precise completion properties we use.

These premises do not require a stochastic dynamics. We use ideal readouts in the task-accuracy closure: their record weights are limits of attainable readouts, and finite compositions and postprocessings respect these limits. This is an explicit accuracy convention, not an assertion that perfect constructors physically exist. The source asks subsidiary theories to provide suitable approximation structure [2, Sec. 2, principle I, p. 12]; the operational choice below is additional. When transporting attributes through a physical implementation, we represent it by a total transformation; its action outside the task’s calibration inputs is otherwise unconstrained. This is also part of our operational convention. A finite operational representation of arbitrary dimension is sufficient, but the proofs do not require one.

## 2.2 The record-weight bridge

For a readout  $M$ , its record  $i$ , and a completely specified preparation  $\omega$ , write  $w(M, i \mid \omega)$  for the derived decision weight. Weights are nonnegative and normalized. A sharp record has weight one, and grouping records adds their weights. Parallel repetition on independent instances has the same single-instance marginal weights. In the two-copy Bell argument, we also require

$$w(e, e' \mid \omega \otimes \omega) = w(e \mid \omega)w(e' \mid \omega) \quad (4)$$

for independently prepared experiments. This is Cabello’s product premise [5, step I, Eq. (4)]. These finite-record rules are explicit bridges. Nonsignalling alone does not supply the product rule.

We interpret these numbers as the values of zero-one games played on the final records [2, Secs. 6–7]. In particular, measurement neutrality P4 identifies the values of different implementations of the same game [2, Sec. 7.2, p. 37]. Its scope is important: T1/T2 select observables, rather than asserting a decision-weight extension for every possible readout. We assume the finite-record calculus on the calibrated networks used below. Theorem 1 gives a second, quantitative route to their implementation independence; it does not rederive the whole decision-theoretic representation theorem.

Define an operational pseudometric on preparations of a fixed substrate by

$$d(\omega, \sigma) = \sup_{M \text{ binary}} |w(M, 1 \mid \omega) - w(M, 1 \mid \sigma)|. \quad (5)$$

The supremum includes finite ancillary networks and their ideal readout limits. Distance to an attribute is the infimum over its members; for an empty attribute we set it to one. We make no use of the converse assertion that a weight-one outcome must be sharp.

## 2.3 A finite-resolution repair of T1

Condition E2 asks for one frequency partition on each qualifying attribute; T1 applies it to a selected observable and its two-copy sum [2, Sec. 5, E2, p. 28; Sec. 6, T1, p. 30]. Our domain consists instead of complete preparation instructions, specified independently of a measurement’s weights. A coarser attribute is included only when the same limit works uniformly for all its preparations.

Here is the operational version of the limit. For a binary measurable variable  $X = \{x_0, x_1\}$  and  $L$  copies, form all sharp strings  $x_{s_1} \otimes \cdots \otimes x_{s_L}$ . For a candidate frequency  $f \in [0, 1]$  and resolution  $h > 0$ , define

$$A_{X,L}(f, h) = \text{cl} \bigcup_{\substack{s \in \{0,1\}^L \\ |L^{-1} \sum_j s_j - f| \leq h}} x_{s_1} \otimes \cdots \otimes x_{s_L}. \quad (6)$$

The union is a calibration attribute for the task that counts the ones and reports whether their frequency lies in the window. By VIII, every implementation of that counting task reports yes sharply on its closure. Thus Eq. (6) is defined by attributes and tasks, not by choosing an apparatus’s output distribution.

**Operational homogeneity H.** For every complete preparation  $\omega \in D_X$  and every binary comparison variable used in the proof, including those generated by discrimination completion, there is one  $f$  such that, for each fixed  $h > 0$ ,

$$\delta_{X,L}(\omega; f, h) := d(\omega^{\otimes L}, A_{X,L}(f, h)) \longrightarrow 0. \quad (7)$$

H makes two additions to a mere restriction of T1’s domain. It chooses the operational topology for the frequency limit, and extends the frequency condition to the finite family of tested binary

coarse-grainings. Neither addition is attributed to the printed T1. A law of large numbers for one chosen apparatus is weaker than H.

Quantum preparations satisfy H for projective variables: projection onto a frequency window gives a nearby input on which that window is sharp. Appendix B proves

$$\delta_{X,L}(\rho; f, h) \leq \min \left\{ 1, \frac{1}{2h\sqrt{L}} \right\}. \quad (8)$$

This is a consistency example only. The main theorems use Eq. (7), not quantum projection.

## 2.4 Compositional calibration coverage

A record-producing circuit is not automatically a measurer of an input information variable. We therefore state its required coverage separately.

**Coverage C.** For the finite experimental scenario, there is one enlarged substrate obtained by appending fixed independent finite ancillas. Each tested context is a calibrated measurer of a finite information observable on that substrate, possibly followed by classical coarse-graining. The common enlarged preparation belongs to every corresponding domain  $D_X$ . The identifications of local calibration attributes persist when spectators are appended, so that a shared local setting with opposite records defines distinguishable global calibration attributes. The same condition holds for the finite independent copies used in the experiment.

In the capacity argument we also use independent calibration resources: the calibration attributes in its finite frames admit nonempty independent products, and their local discriminators retain their calibration on those products. Appending any fixed member is an available total preparation map. This specifies the product-input convention used with III and VI, without imposing a numerical factorization on correlated inputs.

This is a task statement about calibration, domains, and composition. It does not require a postmeasurement instrument to preserve compatible statistics. It is therefore not Chiribella–Yuan fundamental sharpness, although their dilation premise is a relevant stronger operational motivation [4, Principle 1 and the reduction to sharp measurements]. The distinction between allowed

effects and specified measurement tasks is also central to restricted-effect theories [6, Sec. IV].

## 3 From frequency limits to exclusivity

Joint distinguishability alone supplies a new measurement, but says nothing numerical about its relation to existing readouts. The next theorem supplies that relation. Its proof concerns attributes and readout weights, without a state-space representation.

**Theorem 1** (Frequency-neutral task completion). *Assume the task principles and finite-record bridge of Sec. 2. Let  $M, N$  be any two measurers of the same binary variable  $X$ , and let  $\omega \in D_X$ . For every  $f, h, L$ ,*

$$\begin{aligned} |w(M, 1 | \omega) - f| &\leq h + \delta_{X,L}, \\ |w(M, 1 | \omega) - w(N, 1 | \omega)| &\leq 2h + 2\delta_{X,L}. \end{aligned} \quad (9)$$

*Under H, all such readouts have the same weights.*

*Consider finitely many covered events on one preparation, each represented by a closed outcome attribute of its context. If these attributes are pairwise distinguishable, their weights sum to at most one. More generally, if every comparison of calibrated binary measurers required in the completion differs by at most  $\kappa$ , a family of  $m$  events obeys*

$$\sum_{i=1}^m w(e_i | \omega) \leq 1 + m\kappa. \quad (10)$$

*Consequently H and C imply CE for covered events and local orthogonality (LO) at every finite-copy level. Product weights are needed only to identify independent-copy statistics with the usual hierarchy.*

*Proof.* Repeat  $M$  in parallel  $L$  times, and call the record frequency  $F$ . Its mean weight is  $w(M, 1 | \omega)$ , by the marginal assumption. On a preparation in  $A_{X,L}(f, h)$ , the window detector is sharp. Equation (5) therefore bounds its failure weight on  $\omega^{\otimes L}$  by  $\delta_{X,L}$ . Since  $0 \leq F, f \leq 1$ , splitting the finite weighted sum into the window and its complement gives

$$|\mathbb{E}_w F - f| \leq \mathbb{E}_w |F - f| \leq h + \delta_{X,L}. \quad (11)$$

The two estimates in Eq. (9) follow. Taking  $L$  to infinity and then  $h$  to zero proves neutrality. Independence among the  $L$  records was not used.

For the completion, let  $x_1, \dots, x_m$  be the closed event attributes and put  $r = \bigcup_i x_i$ . The nonempty members of

$$Z = \{x_1, \dots, x_m, r\} \quad (12)$$

are pairwise distinguishable. Their union has closure  $\Omega$ : its bar is  $\overline{\bigcup_i x_i} \cap \text{cl}(\bigcup_i x_i)$ , which is empty. Principle IV supplies a common discriminator  $D$ .

Coarse-grain  $D$  to “ $i$  or not  $i$ ”. It is a measurer of the same binary calibration variable as a Boolean discriminator for  $\{x_i, \bar{x}_i\}$ . Applying VIII on the full domain shows that this coarse-graining is itself calibrated on that Boolean variable. It is therefore also a measurer of the binary coarse-graining of the original context that defines event  $i$ . The original preparation is in that context’s domain by C. Neutrality identifies the two weights, or bounds their difference by  $\kappa$ . Summing the  $m$  outcomes of  $D$  proves Eq. (10).

For LO, two locally exclusive events differ in the outcome of the same local setting. That local readout distinguishes their global calibration attributes under C. The preceding completion applies to each finite set and to every finite composite. The independent-copy product rule gives the usual hierarchy. The closure and calibration details are proved in Appendix A.  $\square$

The physical content is that possible discrimination must preserve the weights already assigned to its calibration labels. The theorem does not construct a minimally disturbing instrument. Its qualitative neutrality conclusion agrees with the game-level P4 of Ref. [2]; Eq. (9) gives a direct frequency estimate and identifies the input topology needed for that conclusion.

There is a second sufficient reading of the qualitative step. Suppose the zero-one games for every comparison variable in Appendix A fall within the decision-supporting domain, and suppose P4 applies to their implementations. Appending the same reward adder to either calibrated detector implements the same game, so P4 identifies their values. Those values are the two event weights by the record-weight bridge. The completion proof then gives CE without invoking H. Extending the game domain to all these variables is an explicit premise; it is not supplied by citing P4 for one selected observable. The finite estimate in Eq. (9) remains an additional conclusion of H.

There is a useful converse certificate. If a family of  $m$  pairwise distinguishable covered events

has weight sum  $1 + \Delta$ , with  $\Delta > 0$ , at least one binary comparison in the completion differs by at least  $\Delta/m$ . For any proposed common frequency center,

$$\delta_{X,L} \geq \frac{\Delta}{2m} - h \quad \text{for every } L. \quad (13)$$

This obstruction is independent of local dimension and state-space geometry. It identifies a finite-frequency task that fails.

## 4 The Bell bound with a residual record

Exclusivity becomes useful for CHSH only when the additional parity events are represented by calibrated readouts. Instead of assuming that a joint parity variable covers the input, we complete it and retain the extra outcome. This makes the domain issue quantitative.

Let Alice and Bob have settings  $x, y \in \{0, 1\}$  and outcome bits  $a, b \in \{0, 1\}$ , corresponding to signs  $(-1)^a, (-1)^b$ . Define

$$E_{xy} = \sum_{a,b} (-1)^{a+b} w(a, b \mid x, y), \quad (14)$$

$$B = E_{00} + E_{01} + E_{10} - E_{11}.$$

The doubled preparation is  $\omega \otimes \omega$ , on laboratories  $(A, B, A', B')$ . Fixed ancillas needed by C are included in the laboratory substrates.

For  $t = 0, 1$ , compare Alice’s setting  $x$  with Alice’s setting  $x \oplus t$ . The choices  $t = 0$  and  $t = 1$  are called aligned and crossed. Let  $U_{t,x}^c$  be the union of product calibration attributes with  $a \oplus a' = c$ , including arbitrary independent spectator preparations. On the full substrate define

$$p_{t,x}^c = \text{cl}(U_{t,x}^c),$$

$$h_t^{cd} = p_{t,0}^c \cap p_{t,1}^d, \quad V_t = \bigcup_{c,d \in \{0,1\}} h_t^{cd}. \quad (15)$$

The local readouts followed by a parity computation distinguish  $U_{t,x}^0$  and  $U_{t,x}^1$ . VIII extends their sharp records to the closed sectors. Different nonempty  $h_t^{cd}$  are therefore pairwise distinguishable. IV supplies a discriminator for the variable  $H_t$  formed by these attributes. As throughout the paper, empty outcomes are omitted.

A discriminator need not have the required domain. Define its residual attribute and completed variable by

$$r_t = \bar{V}_t, \quad \hat{H}_t = \{h_t^{cd} : h_t^{cd} \neq \emptyset\} \cup \{r_t\}, \quad (16)$$

again omitting  $r_t$  if empty. The completion argument of Theorem 1 gives  $D_{\hat{H}_t} = \Omega$ . Choose one calibrated discriminator  $\hat{J}_t$  of this full-domain variable and put

$$\varepsilon_t = w(\hat{J}_t, r_t \mid \omega \otimes \omega). \quad (17)$$

Set  $\varepsilon_t = 0$  if  $r_t$  is empty. Under H this number is independent of the calibrated implementation. It is an ordinary record weight, not a probability that the constructor malfunctions and not an assumed mixture fraction outside a subspace.

**Parity-domain coverage P.** The domain condition used in the zero-residual theorem is

$$\omega \otimes \omega \in D_{H_t} = \text{cl}(V_t), \quad t = 0, 1. \quad (18)$$

It is the task-domain counterpart of Cabello's joint-parity normalization assumption [5, step III, Eq. (5), and the extension to Eq. (10)]. Its formulation does not assign simultaneous values to the original local settings. P implies  $\varepsilon_0 = \varepsilon_1 = 0$ : the residual detector is sharp no on  $\text{cl}(V_t)$ . We will not use the converse assertion that a zero weight implies a sharp prediction.

**Theorem 2** (Residual-weight Bell bound). *Assume the task principles, the finite-record bridge, and coverage C. Let  $\kappa \geq 0$  bound every calibrated binary comparison in the sixteen discrimination completions specified below. With independent-copy factorization,*

$$B^2 \leq 8 + 4(\varepsilon_0 + \varepsilon_1) + 288\kappa. \quad (19)$$

*Under H, one may set  $\kappa = 0$ . Under P as well, the consequence is*

$$|B| \leq 2\sqrt{2}, \quad (20)$$

*and this constant is sharp. With P and finite calibration discrepancies,*

$$B^2 \leq 8 + 288\kappa. \quad (21)$$

*More generally,*

$$\varepsilon_t \leq d(\omega \otimes \omega, D_{H_t}), \quad (22)$$

*so either residual weights or operational distances to the parity domains give valid error bounds.*

*Proof.* Let  $S$  be the sum of the eight weights satisfying the CHSH winning relation  $a \oplus b = xy$ . Normalization gives

$$S = 2 + \frac{B}{2}, \quad \sum_{a \oplus b \neq xy} w(a, b \mid x, y) = 4 - S. \quad (23)$$

For each  $t, c, d, k \in \{0, 1\}$ , form eight four-laboratory events indexed by  $x, a, b \in \{0, 1\}$ :

$$\begin{aligned} x' &= x \oplus t, & y &= k, & y' &= k \oplus 1 \oplus c \oplus d, \\ a' &= a \oplus 1 \oplus \gamma_x, & (\gamma_0, \gamma_1) &= (c, d), \\ b' &= b \oplus 1 \oplus \gamma_x \oplus xy \oplus x'y'. \end{aligned} \quad (24)$$

Adjoin the outcome  $h_t^{cd}$  of  $\hat{J}_t$ , unless it is empty. These are at most nine pairwise distinguishable closed event attributes. Within one  $x$  block, different  $(a, b)$  are locally exclusive. Between the blocks,  $b \oplus b'$  changes while Bob's settings stay fixed, so at least one Bob record differs. Each fine event has Alice parity opposite to the  $x$  component of  $h_t^{cd}$ .

Theorem 1 bounds each event sum by  $1 + 9\kappa$ . Across the sixteen choices, every ordered winning–winning or losing–losing pair occurs exactly once, and each joint parity outcome occurs twice. Appendix C proves this using an explicit inverse to Eq. (24). Product factorization and normalization of the completed readouts now give

$$S^2 + (4 - S)^2 + 2 \sum_{t=0}^1 (1 - \varepsilon_t) \leq 16 + 144\kappa. \quad (25)$$

Substitution of Eq. (23) proves Eq. (19). H removes  $\kappa$  by Theorem 1. P removes the two residual weights by their sharp-no calibration. The qubit construction in Appendix E attains equality.

Finally, the residual detector is zero on every preparation in  $D_{H_t}$ . Comparison with any such preparation in Eq. (5), followed by an infimum, proves Eq. (22). The empty-domain convention covers that case too.  $\square$

The only change to Cabello's numerical argument is retaining the two residual weights. Their presence identifies what is needed from the task domains. Figure 2 plots the resulting bound. A common frequency-window radius  $h$  and distance bound  $\delta$  in the binary comparisons replace the last term in Eq. (19) by  $576(h + \delta)$ .

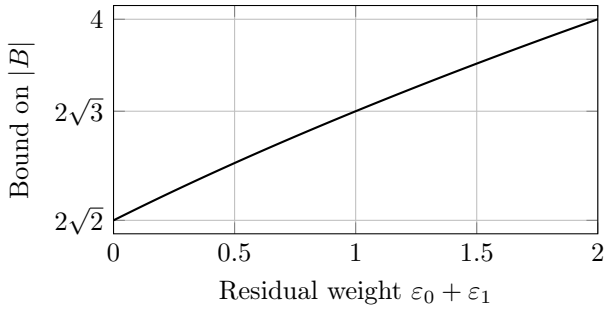


Figure 2: The residual-weight bound at  $\kappa = 0$ :  $|B| \leq \sqrt{8 + 4(\varepsilon_0 + \varepsilon_1)}$ . Zero residual weight gives  $2\sqrt{2}$ ; no restriction on the residual records gives the algebraic bound four. Attainability is established at zero residual weight, not along the whole curve.

## 5 What T2 contributes: sectors and capacity

The previous theorem reduces the Bell question to two calibrated residual records. To connect it to T2, we first construct the aligned sectors and then ask what would make their residual attribute empty. The relevant task notion is the number of distinguishable information labels, not the number of state coordinates.

### 5.1 The four aligned sectors

Let  $S_A$  denote Alice’s active substrate and let  $X = \{x_0, x_1\}$  and  $Y = \{y_0, y_1\}$  be the binary sector selected by T2. The source’s labels  $x_1, x_2, y_+, y_-$  have only been relabelled. R1 places these attributes in the required generalized-mixture relations. R2 supplies label flips that preserve the other variable’s attributes. R3 identifies specified intrinsic measurement attributes. R4 supplies a two-copy generalized mixture  $q$  of the equal- $X$  products and of the equal- $Y$  products [2, Sec. 6, T2, R1–R4, pp. 30–32].

On  $S_A \otimes S_A$ , use the same parity definition as Eq. (15), without spectators for now. R2 and R4 imply that all four aligned sectors are nonempty. Indeed, write  $T_X$  for the R2 operation on the first copy that flips  $x_0, x_1$  while preserving both  $y$  attributes, and  $T_Y$  for its counterpart. Then

$$\begin{aligned} q &\subseteq h_0^{00}, & T_X q &\subseteq h_0^{10}, \\ T_Y q &\subseteq h_0^{01}, & T_X T_Y q &\subseteq h_0^{11}. \end{aligned} \quad (26)$$

Appendix D proves that possible total transformations transport distinguishability closure, which justifies these inclusions without assuming

globally reversible label flips. All four attributes are pairwise distinguishable. R1 and R3 remain conditions on a T2 theory, but are not used in this sector-construction lemma.

### 5.2 A task certificate for an exhaustive domain

For a substrate  $S$ , let  $\text{cap}(S)$  be the greatest number of nonempty pairwise distinguishable attributes it admits, when finite. IV turns every such family into a common discrimination task; call the family a frame. Thus an upper bound on  $\text{cap}(S)$  is an impossibility statement about recording too many mutually distinguishable input labels. It is not a bound on the dimension of an operational state space.

Deutsch and Marletto call a variable maximal when the bar of its union is empty [1, Sec. 6, p. 14]. The elementary connection needed here is that a frame saturating the information capacity is maximal: a nonempty residual attribute would add one more distinguishable label. We apply that observation while keeping the spectators explicit.

**Theorem 3** (Capacity obstruction to a missing parity domain). *Assume the task principles of Sec. 2, the total-transformation convention, and the independent-calibration convention stated with C. Suppose  $S_A$  has the T2 sector described above. Let  $R$  be any finite spectator substrate with  $\text{cap}(R) = n < \infty$ , and form the parity attributes on  $S_A \otimes S_A \otimes R$ .*

*If the aligned parity domain is not the whole substrate, then*

$$\text{cap}(S_A \otimes S_A \otimes R) \geq 4n + 1. \quad (27)$$

*In particular, the additional capacity bound*

$$\text{cap}(S_A \otimes S_A \otimes R) \leq 4\text{cap}(R) \quad (28)$$

*forces full aligned coverage, including preparations correlated with  $R$ .*

*Suppose, in addition, that an available reversible transformation  $G$  of  $S_A$ , with available inverse, interchanges the tested attributes:*

$$G(x_a) = y_a, \quad G(y_a) = x_a, \quad a \in \{0, 1\}. \quad (29)$$

*Then full aligned coverage implies full crossed coverage. Under this interchange task, failure of either domain implies Eq. (27).*

*Proof.* Choose one nonempty attribute  $q_{cd}$  in each sector of Eq. (26), and an  $n$ -element frame  $\{z_j\}$  on  $R$ . Their  $4n$  independent products  $q_{cd} \otimes z_j$  are nonempty and pairwise distinguishable. A parity readout separates different  $(c, d)$ ; a discriminator on  $R$  separates different  $j$ . Each product lies in the corresponding full-substrate sector  $h_0^{cd}$ . To see the latter inclusion, append a fixed member of  $z_j$  and apply the closure-transport lemma to each of the two parity calibration attributes. Taking the union over members of  $z_j$  gives the claim.

If the aligned domain is not all of  $\Omega$ , then  $r_0 = \overline{V_0}$  is nonempty: otherwise  $\text{cl}(V_0) = \overline{\emptyset} = \Omega$ . This attribute is distinguishable from every product in the  $4n$ -element frame. IV then gives a  $4n + 1$ -element discrimination task, proving Eq. (27). This argument includes a possible residual state correlated with  $R$ ; no factorization of that state was used.

Apply  $G$  to the second Alice copy and the identity elsewhere. An available reversible map preserves both distinguishability and its closure, by applying closure transport to the map and its inverse. Equation (29) therefore sends  $h_0^{cd}$  onto  $h_1^{cd}$  and their closed union onto the crossed domain. It maps  $\Omega$  onto itself, proving the remaining statements.  $\square$

The theorem supplies a sufficient task extension of T2. If Eq. (28) holds for the spectators in the doubled Bell experiment and the setting-interchange task is available, P follows and Theorem 2 applies. One sufficient form of the capacity bound is a binary information capacity  $\text{cap}(S_A) = 2$  together with multiplicativity on the relevant finite composites.

Interoperability supplies the product lower bound, not an upper bound. Deutsch and Marletto do state that locality II makes their classical information capacity additive [1, Sec. 4, p. 10]. Their capacity is the logarithm of the largest clonable computation variable;  $\text{cap}$  here counts distinguishable frames. Under preparation access, the record-and-prepare construction of their Sec. 8.1, pp. 20–21, turns a finite frame into an information variable: read its label, then prepare a permuted label or two copies. On that domain, their additivity statement gives multiplicativity of  $\text{cap}$ . Applying it here requires preparation access for the frames counted on the relevant composites, as well as the global condition  $\text{cap}(S_A) = 2$ .

T2 selects a binary sector; it does not impose

that global capacity. For example, the same qubit sector can occupy two levels of a qutrit. It satisfies T2, while the carrier and its square have capacities three and nine. Equation (28) then fails for a trivial spectator. Appendix E gives this embedding explicitly. Hardy’s composition axiom also has a product rule for information capacity [11, Axiom 4]; we do not use his separate rule for the number of state parameters. Equation (29) requires a further reversible task. R2 flips labels within a variable, rather than interchanging  $X$  and  $Y$ .

More generally, the same proof works whenever a frame of  $M$  attributes inside the aligned sectors saturates  $\text{cap}(S_A \otimes S_A) = M$  and this saturation persists as the upper bound  $M \text{cap}(R)$  after adjoining  $R$ . Appendix D gives the statement. This general form applies to higher-capacity carriers as well.

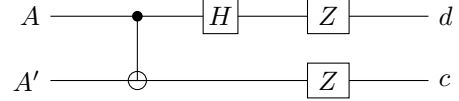


Figure 3: The running qubit example. A controlled-NOT followed by a Hadamard gate and computational readout measures the aligned Bell labels:  $c$  is the  $Z \otimes Z$  parity and  $d$  the  $X \otimes X$  parity. Four orthogonal Bell states fill the four-dimensional space, so no residual label remains. A Hadamard on the second input converts this to the crossed context.

For the running qubit example, the four Bell states constitute a saturating frame. The parity readout in Fig. 3 has no residual record, even with a correlated finite-dimensional quantum spectator. A Hadamard gate interchanges the tested  $Z$  and  $X$  attributes and produces the crossed context. Appendix E verifies the capacity and circuit claims directly.

### 5.3 What no-hypersignaling would supply

The capacity premise concerns the size of a perfectly readable alphabet. No-hypersignaling concerns a different quantity: the classical cost of reproducing every finite prepare-and-read channel, with shared randomness allowed [14, Definitions 2–3, Eq. (2)]. The connection is useful, but it requires stating which attributes can be prepared.

In this subsection, assume that each nonempty attribute in a finite discrimination frame contains an available complete preparation. In particular,

this applies to a proposed residual attribute. This preparation-access premise extends VI, whose wording guarantees preparation of information-instantiating attributes [1, Sec. 6, VI, p. 13]. It is an explicit premise when applied to arbitrary residual attributes, not a consequence of their being nonempty.

Write  $s(S)$  for the signaling dimension, with value infinity if no finite simulator exists. It is the smallest integer  $m$  such that every finite channel  $p(j | i)$  obtained by preparing  $S$  and making one fixed readout can be simulated by sending one of  $m$  classical symbols, with free shared randomness. There is no independently chosen input at the receiver. Equivalently, each such channel belongs to the convex hull of deterministic maps whose image has at most  $m$  elements. The classical simulator is only a benchmark; no hidden stochastic dynamics is assumed for  $S$ .

Any  $m$ -symbol channel with the same  $N$  input and output labels satisfies

$$\sum_{i=1}^N p(i | i) \leq m. \quad (30)$$

Indeed, a deterministic map with at most  $m$  output values has at most  $m$  fixed points. Convex mixtures preserve the inequality. An  $N$ -label discriminator with worst-case preparation-and-readout error at most  $\xi$  gives the opposite bound  $N(1 - \xi)$  on this sum. Letting  $\xi$  vanish proves

$$\text{cap}(S) \leq s(S) \quad (31)$$

under preparation access. The closed classical polytope includes the ideal task limit.

**Corollary 5.1** (A no-hypersignaling certificate for parity coverage). *Assume the hypotheses of Theorem 3 and preparation access for its frames. Let  $\text{cap}(R) = n$ . Suppose the local prepare-and-read simulations obey*

$$s(S_A) = 2, \quad s(R) = n, \quad (32)$$

*and no-hypersignaling holds on the relevant composite:*

$$s(S_A \otimes S_A \otimes R) \leq s(S_A)^2 s(R). \quad (33)$$

*Then the aligned parity domain is exhaustive. The reversible setting interchange of Eq. (29) makes the crossed domain exhaustive too. With  $H$ ,  $C$ ,*

*the finite-record bridge, and independent-copy factorization as in Theorem 2, the consequence is  $|B| \leq 2\sqrt{2}$ .*

*More explicitly, a missing aligned domain with a preparable residual would permit a channel on  $N = 4n + 1$  labels for which*

$$\begin{aligned} \frac{1}{4n+1} \sum_i p(i | i) &\geq 1 - \xi, \\ \frac{1}{4n+1} \sum_i p(i | i) &\leq \frac{4n}{4n+1} \quad (\text{classical}). \end{aligned} \quad (34)$$

*The second line applies to every  $4n$ -symbol simulator. Thus any realization with  $\xi < 1/(4n+1)$  violates that simulation bound.*

*Proof.* Equations (31)–(33) imply Eq. (28); Theorem 3 then applies. Alternatively, its construction supplies  $4n$  calibration labels and one residual label. Preparation access and IV turn them into a nearly noiseless  $4n + 1$ -label channel. Equation (30) proves Eq. (34), contradicting Eq. (33) at sufficiently small  $\xi$ .  $\square$

This is the distinguishable-message instance of the communication-game witness used to formulate no-hypersignaling [14, Eq. (3)]. The task result supplies its messages from a failed parity domain. Without Eq. (32), no-hypersignaling gives only  $\text{cap}(S_A^{\otimes 2} R) \leq s(S_A)^2 s(R)$ , not the required  $4n$ . It is therefore stronger than the capacity upper bound only after these local simulation budgets and preparation access are fixed.

The distinction is strict in general operational theories. For the elementary square and product-state composite in the hypersignaling model of Ref. [14], the largest jointly readable alphabets have sizes two and four. Nevertheless, a seven-preparation channel defeats every four-symbol simulation by the exact payoff gap  $1/2 - 10/21 = 1/42$ . Appendix F proves the alphabet counts and checks that published witness. The square violates principle IV, so this comparison separates two capacity notions; it is not a countermodel to the constructor premises.

## 6 Two finite tests of the premises

The running qubit example isolates preparation homogeneity. A five-event certificate then excludes the joint application of  $H$  and  $C$  to specified Bell weights, without assuming any state-space geometry.

Return to

$$Z_+ = \{\rho_A \otimes |+\rangle\langle+| : \rho_A \text{ arbitrary}\}. \quad (35)$$

This is a closed support attribute. Together with  $Z_-$  it has a unitarily readable and copyable sign label. For the total computational-bit observable  $N_A + N_B$ , where  $N = |1\rangle\langle 1|$ , its frequency vector is

$$\left(\frac{r}{2}, \frac{1}{2}, \frac{1-r}{2}\right), \quad r = \langle 0 | \rho_A | 0 \rangle. \quad (36)$$

No state in  $Z_+$  lies in a sum eigenspace, and no sum eigenspace is orthogonal to all of  $Z_+$ . These are the relevant generalized-mixture conditions [2, Sec. 3, “Generalised mixtures”, pp. 17–18]. Yet the members of  $Z_+$  have different frequency partitions.

Let  $F_{0,L}$  be the frequency operator for outcome zero on  $L$  copies. The exact optimized error about a single common label is

$$\inf_f \sup_{\rho_1, \dots, \rho_L \in Z_+} \text{Tr} \left[ \left( \bigotimes_j \rho_j \right) (F_{0,L} - fI)^2 \right] \quad (37)$$

$$= \left( \frac{L+1}{4L} \right)^2.$$

It approaches  $1/16$ . For one completely specified member, the mean-square error about its own frequency is instead at most  $1/(4L)$ . Appendix B proves both statements. H therefore uses complete preparation instructions, not every copyable information label.

For a second, geometry-independent test, consider the positive nonsignalling weights

$$w(a, b | x, y) = \frac{1 + (-1)^{a+b+xy}t}{4}, \quad t = \frac{81}{100}. \quad (38)$$

They give  $B = 81/25$ . The following five two-copy events, in party order  $(A, B, A', B')$ , are pairwise locally exclusive:

$$\begin{aligned} (0000 | 0000), & (0011 | 0100), (0111 | 1101), \\ (1100 | 1010), & (1110 | 0011). \end{aligned} \quad (39)$$

Each wins both CHSH games. Their exact total weight is

$$5 \left( \frac{181}{400} \right)^2 = 1 + \frac{761}{32000}. \quad (40)$$

This is an application of local orthogonality [8]. Under C and the discrimination principles, Theorem 1 forces some binary comparison to differ by

at least  $761/160000$ . Its proposed homogeneous frequency limit then obeys

$$\delta_{X,L} \geq \frac{761}{320000} - h \quad \text{for every } L. \quad (41)$$

Thus no theory with these covered weights can satisfy H, regardless of its geometry. The residual Bell theorem gives the additional necessary relation

$$\varepsilon_0 + \varepsilon_1 + 72\kappa \geq \frac{1561}{2500}. \quad (42)$$

Equation (42) is a necessary certificate, not an assertion that these weights are compatible with H and C.

## 7 Discussion

A possible discriminator, an exhaustive measurement domain, and implementation-independent weights are distinct requirements. The information principles construct discriminators. H compares their weights, while the residual record measures the remaining domain issue. Under H, the result is  $B^2 \leq 8 + 4(\varepsilon_0 + \varepsilon_1)$ . Theorem 3 converts failure of a full domain into a surplus distinguishable label; Corollary 5.1 turns a preparable surplus into a communication witness.

**States versus measurements.** Short and Barrett show that maximal nonsignalling state spaces have no entangled effects. For one and two elementary systems their measurements are mixtures of basic sequential measurements, whereas this stronger description already fails for three systems [12, Sec. IV, Corollary 1, Theorems 2–4]. Their result is an antecedent of the mechanism here: admitting more states can prevent a useful joint measurement. Our addition is narrower. Principles IV and V already supply a discriminator for the calibrated parity sectors, so the unresolved point becomes the domain on which that discriminator can function as the required joint measurement.

Czekaj, Horodecki, and Tylec start from three locally complete binary measurements and add a four-outcome Bell measurement with quantum-inherited parity identities. Positivity on products gives the local Bloch-ball constraint and hence quantum bipartite correlations [13, Secs. II–III, Eqs. (4), (9)]. This is a stronger correlation conclusion under a different measurement and state-description premise. We assume neither those

three coordinates nor local tomography, and do not reconstruct a ball. We retain a possible extra record and derive only the CHSH consequence supported by its weight. Their noisy Bell effects also differ from our residual record, which is a calibrated outcome, not an unsharpness parameter.

Selby and collaborators study parity-reading effects through compositional positivity constraints [10, Secs. 3, 5, 7–8]. Their Sec. 7.1 already recovers the CHSH quantum bound numerically in specified locally tomographic two- and three-fiducial-measurement cases. Our contribution is not that parity readout can restrict CHSH. It is an analytic implication from the stated task and calibration premises, with an explicit residual-domain term. Their general parity-reading conjecture remains distinct from this conditional result. Restricted-effect models likewise motivate keeping the available transformations explicit [6]; the present theorems do not infer the availability of every positive functional.

**Exclusivity and sharpness.** Cabello’s extra two-copy contexts are the direct antecedent of P [5, step III and Eq. (10)]. Our sixteen-set generator certifies his counting construction for arbitrary biased weights. The new quantitative term is the retained residual weight, not a new route from exclusivity alone to Tsirelson’s bound. Chiribella and Yuan instead obtain exclusivity from repeatability, minimal disturbance, and closure of sharp measurements [4, Principles 1–3 and Methods]. Our frequency-neutral completion bypasses that instrument-level assumption. The reset example in Appendix E shows why retaining an input attribute is weaker than minimal disturbance.

The finite-copy LO conclusion uses the principle of Fritz and collaborators [8]. It is not identified with the almost-quantum representation of Navascués and collaborators [9, Definition 1], whose vector structure already bounds CHSH. Nor is the present proof an uncertainty-and-steering derivation of the kind given by Oppenheim and Wehner [3]. These distinctions delimit the contribution: it is a task formulation with a finite-resolution estimate and a residual-domain certificate, not a quantum reconstruction.

**Capacity and the remaining implication.** The source’s classical-capacity additivity statement and Hardy’s capacity composi-

tion rule precede the sufficient bound used here [1, Sec. 4, p. 10][11, Axiom 4]. Section 5.2 explains the preparation-access condition that connects the source’s information variables to the frames counted in Theorem 3. Dall’Arno and collaborators constrain the larger signaling-simulation problem [14, Definitions 2–3]. Our no-hypersignaling corollary needs both access to the residual preparations and the stated local simulation budgets; under those conditions a missing domain has the communication test in Eq. (34).

H, C, the finite-record rules, and independent-copy factorization are added premises. T2 yields the aligned sectors, not the global binary capacity in Eq. (28); the qutrit embedding already separates those assertions. We have not derived the reversible setting interchange, or preparation-specific coverage P, from T2 and the information principles. Nor have we constructed a superquantum theory satisfying all those source premises. The remaining Bell question concerns coverage of the tested preparation, not the already-false unrestricted inference from a binary sector to global binary capacity.

For a universal Bell bound, the selected T2 settings must also cover the tested settings or experiments approaching the theory’s supremum. Requiring P for every setting pair would be too strong: the tilted-qubit example in Appendix E lacks an exhaustive aligned domain. Failure of one sufficient certificate does not establish independence of P.

**Reproducibility.** The source package contains this manuscript, its TikZ and pgfplots figures, a build script, and exact certificate code. The code checks the finite enumeration, rational identities, and quantum examples; it is not a formalization of the universal task theorems.

## References

- [1] David Deutsch and Chiara Marletto. Constructor theory of information. *Proceedings of the Royal Society A* **471**, 20140540 (2015). [arXiv:1405.5563v2](#).
- [2] Chiara Marletto. Constructor theory of probability. *Proceedings of the Royal Society A* **472**, 20150883 (2016). [arXiv:1507.03287v3](#).
- [3] Jonathan Oppenheim and Stephanie Wehner. The uncertainty principle de-

- termines the non-locality of quantum mechanics. *Science* **330**, 1072–1074 (2010). [arXiv:1004.2507v2](#).
- [4] Giulio Chiribella and Xiao Yuan. Measurement sharpness cuts nonlocality and contextuality in every physical theory. [arXiv:1404.3348v2](#) (2014).
- [5] Adán Cabello. Exclusivity principle and the quantum bound of the Bell inequality. *Physical Review A* **90**, 062125 (2014). [arXiv:1406.5656v3](#).
- [6] Peter Janotta and Raymond Lal. Generalized probabilistic theories without the no-restriction hypothesis. *Physical Review A* **87**, 052131 (2013). [arXiv:1302.2632v1](#).
- [7] Emanuel Knill and Raymond Laflamme. Theory of quantum error-correcting codes. *Physical Review A* **55**, 900–911 (1997). [arXiv:quant-ph/9604034](#).
- [8] Tobias Fritz, Ana Belén Sainz, Remigiusz Augusiak, Jonatan Bohr Brask, Rafael Chaves, Anthony Leverrier, and Antonio Acín. Local orthogonality as a multipartite principle for quantum correlations. *Nature Communications* **4**, 2263 (2013). [arXiv:1210.3018v3](#).
- [9] Miguel Navascués, Yelena Guryanova, Matty J. Hoban, and Antonio Acín. Almost quantum correlations. *Nature Communications* **6**, 6288 (2015). [arXiv:1403.4621v1](#).
- [10] John H. Selby, Ana Belén Sainz, Victor Magron, Łukasz Czekaj, and Michał Horodecki. Correlations constrained by composite measurements. *Quantum* **7**, 1080 (2023). [arXiv:2009.04994v4](#).
- [11] Lucien Hardy. Quantum theory from five reasonable axioms. [arXiv:quant-ph/0101012v4](#) (2001).
- [12] Anthony J. Short and Jonathan Barrett. Strong nonlocality: A trade-off between states and measurements. *New Journal of Physics* **12**, 033034 (2010). [arXiv:0909.2601v2](#).
- [13] Łukasz Czekaj, Michał Horodecki, and Tomasz Tylec. Bell measurement ruling out supraquantum correlations. *Physical Review A* **98**, 032117 (2018). [arXiv:1802.09510v2](#).
- [14] Michele Dall’Arno, Sarah Brandsen, Alessandro Tosini, Francesco Buscemi, and Vlatko Vedral. No-hypersignaling principle. *Phys-*

*ical Review Letters* **119**, 020401 (2017). [arXiv:1609.09237v3](#).

## A Task closure and the completion argument

The proofs use ideal calibration tasks and their limiting readout weights, as specified in Sec. 2. Empty outcomes are omitted; a constant detector handles the associated trivial variables [1, Sec. 6, footnote on p. 15].

### A.1 Orthogonality closure

Principle V and symmetry imply

$$a \perp b \iff \forall s \in a, t \in b : \{s\} \perp \{t\}. \quad (43)$$

For the reverse direction, first use V to infer  $\{s\} \perp b$  for each  $s$ , then use it again for  $a$ . Hence the bar operation is the polarity of a symmetric, irreflexive relation on states. It reverses inclusions,  $a \subseteq \bar{a}$ , and  $\bar{\bar{a}} = a$ . Double bar is therefore extensive, monotone, and idempotent. Its closed sets are closed under intersections. Moreover,

$$\overline{a \cup b} = \bar{a} \cap \bar{b}, \quad \bar{a} \cap \text{cl}(a) = \emptyset. \quad (44)$$

The second equality follows because a state in the intersection would be orthogonal to itself. These facts prove the closure claims used in Eq. (12).

### A.2 Calibration on the closure of an attribute

For any attribute  $a$ , the pair  $\{\text{cl}(a), \bar{a}\}$  is distinguishable by IV and V and has full domain. A measurer of it also measures  $\{a, \bar{a}\}$ . If a variable  $X$  contains  $a$ , the Boolean measurer is a measurer of the binary coarse-graining that asks whether the input label is  $a$ . By VIII, all measurers of that coarse-graining are sharp on  $\text{cl}(a)$ , since  $\text{cl}(a) \subseteq D_X$ . This proves the closure property used for frequency windows and parity sectors.

A Boolean measurer can have sharp yes only on its closed yes attribute. Indeed, on a sharp-yes input  $s$ , that same measurer distinguishes  $\{s\}$  from the no attribute  $\bar{a}$ . Thus  $s \in \text{cl}(a)$ . Conversely, sharp yes on  $\text{cl}(a)$  is its calibration. No inference from the numerical condition  $w = 1$  to sharpness is involved.

### A.3 Full completion and comparison with the original contexts

Let  $x_i$  be the finite family of nonempty closed, pairwise distinguishable attributes in Theorem 1, and  $r = \bigcup_i x_i$ . By Eq. (44), the closure of their union with  $r$  is  $\Omega$ . All members are distinguishable, including  $r$  when nonempty. Let  $D$  be the common discriminator from IV.

For fixed  $i$ , put  $b_i = r \cup \bigcup_{j \neq i} x_j$ . The coarse-graining  $D_i$  is calibrated on  $Y_i = \{x_i, b_i\}$ , and  $D_{Y_i} = \Omega$ . Since  $b_i \subseteq \overline{x_i}$ , a Boolean measurer  $T_i$  of  $\{x_i, \overline{x_i}\}$  is also a measurer of  $Y_i$ . It is sharp no on  $\overline{x_i}$ . VIII, applied on  $D_{Y_i}$ , forces  $D_i$  to be sharp no there too. It is already sharp yes on  $x_i$ . Thus  $D_i$  is a Boolean measurer on all of  $\Omega$ .

Now let the original event arise as outcome  $x_i$  of context  $X_i$ . Both  $D_i$  and the original coarse detector measure

$$V_i = \left\{ x_i, \bigcup_{x \in X_i, x \neq x_i} x \right\}, \quad (45)$$

whose domain equals  $D_{X_i}$ . Coverage puts the tested preparation in that domain. Equation (9) then compares the original event weight with the corresponding  $D$  outcome. If a covered event is a coarse union rather than a closed outcome, replace that union by its closure; the preceding calibration property makes the replacement a closed event with the same detector on its legitimate domain.

When  $\Delta > 0$  in Eq. (13),  $\sum_i w(D, i | \omega) \leq 1$  gives

$$\sum_i [w(e_i | \omega) - w(D, i | \omega)] \geq \Delta. \quad (46)$$

At least one summand is at least  $\Delta/m$ . Combining it with Eq. (9) proves the stated lower bound for every  $L$  and  $f$ .

### A.4 Approximation and task possibility

For each  $L$ , the completion uses only finitely many discriminators, relabellings, and parallel networks. Under the task-accuracy convention, successively more accurate realizations retain all terminal records, including failure records, while uniform calibration errors vanish. Equations (11), (10), and (25) are finite inequalities in bounded record weights and persist in that limit. The frequency limit  $L \rightarrow \infty$  is taken after fixing

the finite-resolution task and its ideal readout weights; no unbounded constructor is claimed to be a finite apparatus.

## B Frequency homogeneity and its quantum consistency

### B.1 A uniform partition is a homogeneous attribute

Let  $P_1, \dots, P_k$  be a projective observable in finite dimension and let  $A$  be an attribute of density operators. Define

$$F_{j,L} = \frac{1}{L} \sum_{\ell=1}^L P_j^{(\ell)}. \quad (47)$$

A common frequency vector  $f$  in mean square on all independent products from  $A$  exists if and only if

$$\text{Tr}(\rho P_j) = f_j \quad \text{for every } \rho \in A \text{ and every } j. \quad (48)$$

To prove this, set  $p_{j,\ell} = \text{Tr}(\rho_\ell P_j)$  and  $\bar{p}_j = L^{-1} \sum_{\ell} p_{j,\ell}$ . Direct expansion gives

$$\begin{aligned} & \text{Tr} \left[ \left( \bigotimes_{\ell=1}^L \rho_\ell \right) (F_{j,L} - f_j I)^2 \right] \\ &= (\bar{p}_j - f_j)^2 + \frac{1}{L^2} \sum_{\ell=1}^L p_{j,\ell} (1 - p_{j,\ell}). \end{aligned} \quad (49)$$

Taking every factor equal proves necessity. If Eq. (48) holds, the right side is at most  $1/(4L)$  for every product, which proves sufficiency. For a closed support attribute consisting of all density operators on a subspace  $S$ , this is equivalent to

$$P_S P_j P_S = f_j P_S. \quad (50)$$

The equivalence follows by evaluating all pure-state quadratic forms on  $S$  and using polarization. This scalar-compression calculation is elementary; the analogous structure for a general error family appears in quantum error correction [7].

### B.2 The exact one-sixteenth obstruction

For  $Z_+$ , the probability or frequency parameter of outcome zero is  $p \in [0, 1/2]$ . The quantum frequency identification here is the one explicitly made in Ref. [2, Sec. 5, E2, p. 28]. For fixed

mean  $q = L^{-1} \sum_{\ell} p_{\ell}$ , concavity of  $p(1-p)$  bounds Eq. (49) by

$$(q - f)^2 + \frac{q(1-q)}{L}, \quad (51)$$

with equality when all  $p_{\ell} = q$ . This expression is convex in  $q \in [0, 1/2]$ , including the affine case  $L = 1$ . Its maximum is therefore

$$\max \left\{ f^2, (1/2 - f)^2 + \frac{1}{4L} \right\}. \quad (52)$$

The minimum occurs when the two entries agree, at  $f = (L + 1)/(4L)$ , proving Eq. (37).

The sign attributes  $Z_{\pm}$  are closed because they are complete density-operator attributes on orthogonal subspaces. Applying  $Z$  to the second qubit exchanges the signs. A controlled sign flip on a blank target copies the label; equivalently, conjugate a controlled-NOT by Hadamard gates on its source and target sign qubits. The unknown first qubit is untouched. Neither of the subspaces  $\mathbb{C}^2 \otimes \text{span}\{|\pm\rangle\}$  has a nonzero intersection with any eigenspace of  $N_A + N_B$ , but each has nonzero overlap with each eigenspace. This proves the generalized-mixture and information-variable assertions used in Sec. 6.

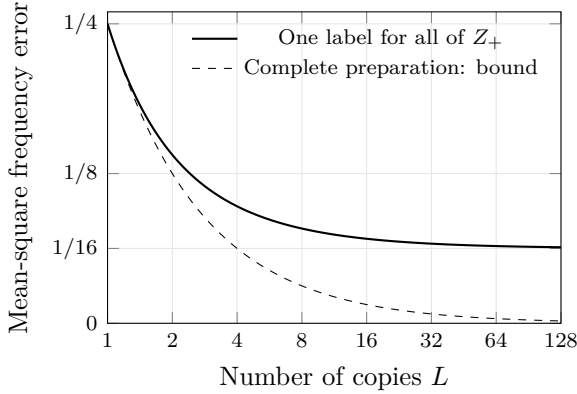


Figure 4: A common frequency center for the unspecified first qubit leaves the exact error  $((L + 1)/(4L))^2$ , approaching  $1/16$ . For one specified preparation, the variance about its own frequency is at most  $1/(4L)$ .

### B.3 Complete quantum preparations satisfy H

For a binary projective observable  $\{P_0, P_1\}$ , let  $f = \text{Tr}(\rho P_1)$  and let  $Q$  be the spectral projector of  $L^{-1} \sum_j P_1^{(j)}$  onto the window  $[f - h, f + h]$ . The input attribute in Eq. (6) is exactly the density operators supported on  $Q$ . From Eq. (49) and

the spectral inequality outside the window,

$$1 - \text{Tr}(Q\rho^{\otimes L}) \leq \frac{1}{4Lh^2}. \quad (53)$$

When  $p = \text{Tr}(Q\rho^{\otimes L}) > 0$ , choose

$$\sigma = Q\rho^{\otimes L}Q/p. \quad (54)$$

Purify  $\rho^{\otimes L}$  to  $|\psi\rangle$ . Its normalized projection  $|\phi\rangle = (Q \otimes I)|\psi\rangle/\sqrt{p}$  has squared overlap  $p$ . Diagonalizing the difference of their rank-one projectors in their two-dimensional span gives trace distance  $\sqrt{1-p}$ . Trace distance cannot increase under partial trace: in its variational expression as a supremum over effects, every effect on the reduced system is also an effect on the purification. Hence

$$\frac{1}{2} \|\rho^{\otimes L} - \sigma\|_1 \leq \sqrt{1-p}. \quad (55)$$

The quantum version of Eq. (5) is bounded by trace distance, including finite ancillary readouts, again by the effect variational expression. Combining with Eq. (53) proves Eq. (8); the bound one covers the remaining cases.

## C The sixteen-set and five-set certificates

### C.1 Pairwise exclusivity and exact coverage

Fix  $t, c, d, k$  in Eq. (24), and write  $k' = k \oplus 1 \oplus c \oplus d$ . The eight fine events have Bob parity

$$b \oplus b' = 1 \oplus \gamma_x \oplus xk \oplus (x \oplus t)k'. \quad (56)$$

Its change between  $x = 0$  and  $x = 1$  is

$$c \oplus d \oplus k \oplus k' = 1. \quad (57)$$

Bob's two local settings are the same in both blocks. Two events from different blocks therefore disagree in at least one Bob output at the same setting. Events within a block differ in  $a$  or  $b$ , also at the same setting. This proves all twenty-eight fine-event comparisons in each set. The remaining eight comparisons are supplied by the Alice parity opposite to  $h_t^{cd}$ . If that sector is empty, omit it and retain the same upper bound for at most nine events.

The last equation in Eq. (24) gives

$$a \oplus b \oplus xy = a' \oplus b' \oplus x'y'. \quad (58)$$

Thus both component events win, or both lose. Conversely, consider any ordered pair satisfying Eq. (58). Recover  $t = x \oplus x'$ ,  $k = y$ , and  $\gamma_x = 1 \oplus a \oplus a'$ . The equation  $c \oplus d = 1 \oplus y \oplus y'$  uniquely fixes the other  $\gamma$  component. Equation (58) then fixes the displayed value of  $b'$ . The event pair occurs in exactly one generated set.

There are eight winning events and eight losing events in a single CHSH experiment. The sixteen sets therefore cover all  $8^2 + 8^2 = 128$  ordered same-status pairs exactly once. Each parity outcome is included once for each  $k$ , hence twice. This proves Eq. (25) for arbitrary, possibly biased, Bell weights and certifies the construction underlying Ref. [5, Eq. (10)].

### C.2 The residual polynomial

Write  $e = \varepsilon_0 + \varepsilon_1$ . The left side of Eq. (25) minus its right side is exactly

$$\left(2 + \frac{B}{2}\right)^2 + \left(2 - \frac{B}{2}\right)^2 + 4 - 2e - 16 - 144\kappa = \frac{B^2 - 8 - 4e - 288\kappa}{2}. \quad (59)$$

For each  $t$ , the four displayed parity weights sum to  $1 - \varepsilon_t$  by normalization of  $\hat{J}_t$ , including the case of empty sectors. The eight parity weights occur twice in the sixteen event sums. This proves the residual coefficient four. The bound can also be combined with the algebraic inequality  $|B| \leq 4$ , which follows from  $|E_{xy}| \leq 1$ .

### C.3 The noisy-box rational certificates

For Eq. (38), each winning event has weight  $181/400$  and each losing event  $19/400$ . Direct summation gives normalized weights, uniform marginals independent of the remote setting, and

$$B = \frac{81}{25}, \quad B^2 - 8 = \frac{1561}{625}. \quad (60)$$

For the five events of Eq. (39), list their ten pairs in lexicographic order. A party witnessing equal setting and opposite output in each pair is respectively

$$3, 3, 2, 1, 2, 4, 1, 1, 4, 3. \quad (61)$$

Each event wins both games. Equation (40) follows by squaring its single-game weight and adding five times. Equations (13) and (21) then give, without numerical approximation,

$$\kappa \geq \frac{761}{160000} \quad \text{or, with P,} \quad \kappa \geq \frac{1561}{180000}. \quad (62)$$

The first inequality concerns the five-event completion; the second concerns the sixteen completions with P. Retaining the residual records instead gives Eq. (42) from  $B^2 - 8 = 1561/625$ .

## D Transporting closure and the T2 parity sectors

The following lemma avoids assuming a reversible symmetry on the whole state space.

**Lemma D.1** (Closure transport). *Let  $T$  be an available total transformation, including a fixed independent resource, and suppose  $T(a) \subseteq b$ . Then*

$$T(\text{cl}(a)) \subseteq \text{cl}(b). \quad (63)$$

*Proof.* Follow  $T$  by a Boolean measurer of  $\{\text{cl}(b), \bar{b}\}$ . This circuit is sharp yes on  $a$ , so it is a measurer of the one-label calibration variable  $\{a\}$ . A constant-yes circuit measures the same variable and is sharp on  $\text{cl}(a)$ . VIII therefore makes the composed circuit sharp yes throughout  $\text{cl}(a)$ . A Boolean measurer has sharp yes only on  $\text{cl}(b)$ , by Appendix A. This proves Eq. (63). The one-label case uses the trivial constant-task convention.  $\square$

For the selected T2 sector, let  $U_X^0 = (x_0, x_0) \cup (x_1, x_1)$  and  $U_X^1 = (x_0, x_1) \cup (x_1, x_0)$ ; define  $U_Y^0, U_Y^1$  analogously. R4 gives

$$q \subseteq \text{cl}(U_X^0) \cap \text{cl}(U_Y^0). \quad (64)$$

The R2 task  $T_X$  on the first copy maps each  $U_X$  parity attribute into the opposite one while preserving both  $U_Y$  parity labels. Similarly  $T_Y$  exchanges the  $Y$  parity and preserves the  $X$  parity. These are simultaneous calibration requirements on the same implementations, as specified by R2. Lemma D.1 extends these actions to the closed parity sectors.

It follows that  $q, T_X q, T_Y q, T_X T_Y q$  lie in the four sectors of Eq. (26). Images of a nonempty physical attribute under a total available transformation are nonempty. Any two of these four attributes differ in an  $X$  or  $Y$  parity, and the corresponding parity measurer distinguishes them. Their enclosing intersections of closed sectors are also pairwise distinguishable and closed. IV supplies a common four-valued discriminator.

## D.1 Reversible transport and spectator-stable saturation

If  $G$  and  $G^{-1}$  are available total transformations, applying Lemma D.1 in both directions gives

$$G(\text{cl}(a)) = \text{cl}(G(a)). \quad (65)$$

The same maps preserve distinguishability: a discriminator can be preceded by the inverse transformation. Hence  $G(\bar{a}) = \overline{G(a)}$ . In particular, Eq. (29) on the second Alice copy maps the aligned product calibration unions to the crossed unions. Equation (65) then maps each closed parity sector, its intersections, and its union closure as claimed in Theorem 3.

Here is the general saturation criterion. Let  $K$  be an active composite with pairwise distinguishable closed attributes  $h_1, \dots, h_k$ . Suppose their union contains a frame  $\{f_1, \dots, f_M\}$ , each  $f_i$  lying in one of the  $h_j$ . For a spectator  $R$  with  $\text{cap}(R) = n$ , assume the extended sectors contain every independent product  $f_i \otimes z$ , with  $z$  a spectator attribute. If

$$\text{cap}(K \otimes R) \leq Mn, \quad (66)$$

then the extended sectors have full domain.

To prove the claim, choose an  $n$ -element frame  $\{z_j\}$ . The  $Mn$  product attributes  $f_i \otimes z_j$  are pairwise distinguishable: a discriminator on  $K$  separates different  $i$ , and one on  $R$  separates different  $j$ . All products belong to the extended union. A nonempty bar of this union would supply a further distinguishable attribute, contradicting Eq. (66). Its bar is therefore empty, which is exactly full double-bar coverage. If  $M = \text{cap}(K)$ , the premise is the spectator-stable upper half of multiplicativity. The upper bound concerns the full composite, including correlated residual attributes, not just the independent calibration products.

## E Quantum examples and comparisons

### E.1 Saturation and the parity domains

Use  $|\Phi\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$  in each Bell experiment. Alice measures  $A_0 = Z$ ,  $A_1 = X$ , and Bob measures

$$B_0 = (Z + X)/\sqrt{2}, \quad B_1 = (Z - X)/\sqrt{2}. \quad (67)$$

The correlators are  $(1, 1, 1, -1)/\sqrt{2}$ , obtained directly from  $\langle Z \otimes Z \rangle = \langle X \otimes X \rangle = 1$  and the vanishing cross correlators. Hence  $B = 2\sqrt{2}$ .

On Alice's doubled subsystem the aligned parity operators are  $Z \otimes Z$  and  $X \otimes X$ . The crossed operators are  $Z \otimes X$  and  $X \otimes Z$ . In each pair the operators commute and square to the identity. Their four joint projectors are

$$\Pi^{cd} = \frac{1}{4}(I + (-1)^c C_0)(I + (-1)^d C_1). \quad (68)$$

They are orthogonal, have trace one, and sum to the identity. Thus their support attributes cover every two-qubit preparation, including arbitrary spectators. They realize P without imposing a two-coordinate local state space. The decoder in Fig. 3 is  $U = (H \otimes I) \text{CNOT}$ , where  $H = (X + Z)/\sqrt{2}$ . Direct conjugation gives

$$\begin{aligned} U^\dagger(Z \otimes I)U &= X \otimes X, \\ U^\dagger(I \otimes Z)U &= Z \otimes Z. \end{aligned} \quad (69)$$

Also  $HZH = X$  and  $HXH = Z$ , proving the setting-interchange property and the crossed decoder.

A quantum system of Hilbert-space dimension  $D$  has information capacity  $D$ . Orthogonal pure states attain this value. Conversely, an effect that accepts every state of an attribute and rejects every state of another forces their support subspaces to be orthogonal. Indeed,  $0 \leq E \leq I$  and  $\text{Tr}(E\rho) = 1$  imply that  $E$  is the identity on  $\text{supp } \rho$ , whereas  $\text{Tr}(E\sigma) = 0$  implies it vanishes on  $\text{supp } \sigma$ . Arbitrarily accurate discrimination gives the same conclusion by compactness of the effect interval. A finite frame thus occupies pairwise orthogonal nonzero subspaces, so its cardinality cannot exceed  $D$ . A spectator of dimension  $D_R$  gives capacity  $4D_R$  for Alice's doubled qubit and the spectator. This proves Eq. (28) for the full quantum composite, not just its product states.

The distinction between a binary sector and a binary-capacity carrier can be seen in a qutrit. Take  $x_0, x_1$  to be the pure-state attributes of  $|0\rangle, |1\rangle$ , and  $y_0, y_1$  those of  $(|0\rangle \pm |1\rangle)/\sqrt{2}$ . The selected two-attribute variables are legitimate information observables without being maximal [1, Sec. 7, p. 19]. Extend the qubit flips and the Hadamard gate by the identity on  $|2\rangle$ . The generalized-mixture relations, simultaneous flip calibrations, intrinsic measurement attributes, and the state  $|\Phi\rangle$  then satisfy R1–R4

within the two-level sector. The four aligned Bell sectors span only the four-dimensional space  $\text{span}\{|0\rangle, |1\rangle\}^{\otimes 2}$ . Their residual support is five-dimensional and contains  $|22\rangle$ . Nevertheless,  $\text{cap}(S_A) = 3$  and  $\text{cap}(S_A \otimes S_A) = 9$ . Thus T2 does not imply the full-carrier bound in Eq. (28), even with a trivial spectator. Preparations covered by the selected binary local observables still lie in the two-level sector and satisfy P in this example.

## E.2 Parity coverage is not automatic for arbitrary quantum settings

Let  $A_0 = Z$  and  $A_1 = (3Z + 4X)/5$ . They are binary projective observables because  $A_1^2 = I$ . Their aligned parity operators satisfy

$$[A_0 \otimes A_0, A_1 \otimes A_1] = \frac{24i}{25}(I \otimes Y + Y \otimes I), \quad (70)$$

whose operator norm is  $48/25$ . The norm follows from the eigenvalues  $2, 0, 0, -2$  of the parenthesized operator. If their common eigenspaces spanned the whole space, a basis of common eigenvectors would make the operators commute. Equation (70) rules that out. Thus their common parity-sector variable does not have full domain, although it can still cover particular preparations.

## E.3 Nonperturbation need not be minimal disturbance

On two qubits, consider the instrument

$$\mathcal{M}_i(\rho) = |i0\rangle\langle i0| \text{Tr}[(|i\rangle\langle i| \otimes I)\rho], \quad i = 0, 1. \quad (71)$$

It is completely positive, and its sum preserves trace. It reads the first-qubit label and leaves that label unchanged, so it is nonperturbing on its calibrated attributes and repeatable. The compatible readout  $Q = I \otimes |1\rangle\langle 1|$  has weight one on  $|01\rangle\langle 01|$  and zero after the instrument. The disturbance is exactly one. It therefore fails the minimal-disturbance condition of Ref. [4, Eq. (4)]. The same first-qubit effects do admit a Lüders implementation; this example separates instrument conditions, not their possible statistics.

## E.4 The almost-quantum comparison

The representation in Ref. [9, Definition 1 and Sec. 6] assigns projectors to local outcomes and a

state on which the relevant products have the required order independence. For binary outcomes put  $A_x = P_0^x - P_1^x$  and  $B_y = Q_0^y - Q_1^y$ . On that state  $|\psi\rangle$ , define unit vectors  $u_x = A_x|\psi\rangle$  and  $v_y = B_y|\psi\rangle$ . Their real inner products are the correlators. Therefore

$$|B| \leq \|v_0 + v_1\| + \|v_0 - v_1\| \leq \sqrt{2(\|v_0 + v_1\|^2 + \|v_0 - v_1\|^2)} = 2\sqrt{2}. \quad (72)$$

This is the standard Gram-vector bound for that representation.

## F Readable labels and the published hypersignaling witness

This appendix separates two dimension notions using the HS model of Ref. [14, Eq. (4) and Supplemental Material]. The model and the seven-input witness are prior work. We give an analytic proof of its readable-alphabet bound and evaluate the witness exactly.

### F.1 Two locally readable labels, four jointly

For this comparison only, let  $d_{\text{rec}}(S)$  be the maximum number of prepared states distinguished by one measurement. Under IV and preparation access it equals  $\text{cap}(S)$ , but we will not impose IV on the comparison model.

The elementary normalized state space is the convex hull of

$$\begin{aligned} v_0 &= (1, 0, 1)^T, & v_1 &= (0, 1, 1)^T, \\ v_2 &= (-1, 0, 1)^T, & v_3 &= (0, -1, 1)^T. \end{aligned} \quad (73)$$

The unit effect is  $u = (0, 0, 1)^T$ , and all positive effects are allowed. The four extreme rays of the effect cone are

$$\begin{aligned} e_0 &= (1, 1, 1)^T, & e_1 &= (-1, 1, 1)^T, \\ e_2 &= (-1, -1, 1)^T, & e_3 &= (1, -1, 1)^T. \end{aligned} \quad (74)$$

These rays follow by imposing nonnegative evaluation on the four displayed vertices. Any finite measurement can be refined into effects proportional to  $e_i$ . Normalization forces their total coefficients to have the form  $(a, b, a, b)$  with  $2a + 2b = 1$ . Thus every measurement is a mixture of the two binary measurements  $\{e_0/2, e_2/2\}$  and  $\{e_1/2, e_3/2\}$ , followed by classical postprocessing. Sending the binary outcome simulates

any such readout with shared randomness. Hence  $s(S) = 2$ , with equality because  $v_0, v_2$  are perfectly distinguishable.

Let  $c = (0, 0, 1)^T$  denote the center state. Since  $2c - v_i$  is a normalized state, any effect  $f$  satisfies  $f(v_i) \leq 2f(c)$ , and convexity extends this to every normalized state. If  $N$  states are read with error at most  $\xi$ , summing over their respective effects gives  $N(1 - \xi) \leq 2$ . Consequently  $d_{\text{rec}}(S) = 2$ .

The bipartite normalized states are all convex mixtures of  $v_i \otimes v_j$ , again with all positive effects. Cone positivity gives

$$4c \otimes c - v_i \otimes v_j = (2c - v_i) \otimes 2c + v_i \otimes (2c - v_j) \geq 0. \quad (75)$$

Applying the same bound to every outcome of a common measurement gives  $N(1 - \xi) \leq 4$ . The products of  $v_0, v_2$  form a four-label readable frame, so

$$d_{\text{rec}}(S \otimes S) = 4 = d_{\text{rec}}(S)^2. \quad (76)$$

This proof includes arbitrary mixed preparations and arbitrarily accurate discrimination limits.

Each pair of different elementary vertices can nevertheless be distinguished by some binary measurement: evaluation of the  $e_i/2$  gives an opposite sharp pair. Since no one measurement distinguishes all four, principle IV fails. The comparison therefore does not decide whether the two dimension notions can differ under every constructor premise.

## F.2 A channel that needs more than four classical symbols

Represent the product state  $v_i \otimes v_j$  by the matrix  $\Omega_{4i+j} = v_i v_j^T$ , and put  $E_{4i+j} = e_i e_j^T$ , with  $i, j \in \{0, 1, 2, 3\}$ . Matrix evaluation is  $\text{Tr}(\Omega^T E)$ . Define

$$E_{23} = \begin{pmatrix} -1 & -1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (77)$$

The subscript follows the source's enumeration. The seven measurement effects are

$$\frac{1}{8}E_0, \frac{1}{8}E_1, \frac{1}{8}E_6, \frac{1}{8}E_8, \frac{1}{8}E_{10}, \frac{1}{8}E_{15}, \frac{1}{4}E_{23}. \quad (78)$$

Their sum is  $uu^T$ . The first six terms are positive products. Since the planar parts of  $v_i, v_j$  are signed coordinate vectors,  $E_{23}$  evaluates to either

zero or two on every product vertex. Convexity proves validity on the entire declared state space.

For preparations  $(\Omega_0, \Omega_2, \Omega_6, \Omega_7, \Omega_{12}, \Omega_{13}, \Omega_{15})$ , respectively, the two nonzero output positions in each row are

$$(1, 6), (2, 7), (2, 3), (3, 7), \\ (4, 6), (4, 7), (5, 6). \quad (79)$$

Each displayed position has weight  $1/2$ , defining a channel  $\Xi$ . This is the channel in Ref. [14, Eq. (4)]. Its payoff matrix is

$$\Gamma = \frac{1}{21} \begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 2 \\ 0 & 2 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 1 & 0 \end{pmatrix}. \quad (80)$$

Direct substitution gives  $\sum_{ij} \Gamma_{ij} \Xi_{ij} = 1/2$ .

A deterministic four-symbol simulation uses a set  $J$  of at most four output columns, so its best payoff is  $\sum_i \max_{j \in J} \Gamma_{ij}$ . Nonnegative entries let us enlarge  $J$  to four columns. Let  $k$  count the selected columns from  $\{2, 3, 7\}$ , let  $l$  count those from  $\{1, 4, 5\}$ , and let  $b$  indicate whether column six is selected. Then  $k + l + b = 4$ . The unscaled row-maximum sum is exactly

$$r(k) + 2l + b(3 - l), \\ (r(0), r(1), r(2), r(3)) = (0, 4, 6, 6). \quad (81)$$

For  $b = 0$ , the pair  $(k, l)$  is  $(1, 3)$ ,  $(2, 2)$ , or  $(3, 1)$ , giving respective scores 10, 10, and 8. For  $b = 1$ , it is  $(0, 3)$ ,  $(1, 2)$ ,  $(2, 1)$ , or  $(3, 0)$ , giving 6, 9, 10, and 9. Thus the maximum is  $10/21$ . Shared randomness cannot increase a linear payoff. We have proved the exact separation

$$\sum_{ij} \Gamma_{ij} \Xi_{ij} - \max_{\text{four symbols}} \sum_{ij} \Gamma_{ij} p(j | i) = \frac{1}{42}. \quad (82)$$

It follows that  $s(S \otimes S) > 4 = s(S)^2$ , despite Eq. (76). The certificate program also evaluates all thirty-five four-column choices and reconstructs the channel directly from Eq. (78).

**Acknowledgments.** AI assistance supported literature review, proof exploration, and preparation of the manuscript and verification code. The author is responsible for the results.