

Exponential parallel repetition for entangled games

Abstract

We prove that for every two-player one-round game G with entangled value $\text{val}^*(G) < 1$ there are constants $c = c(G) > 0$ and $C = C(G) < \infty$ such that $\text{val}^*(G^{\otimes n}) \leq Ce^{-cn}$ for all $n \geq 1$. The proof routes the coordinates of an arbitrary strategy for $G^{\otimes N}$ into the anchored game of Bavarian, Vidick and Yuen, producing a one-parameter family of anchored strategies whose values are the evaluations of a single real polynomial — the occupancy profile of the strategy — on a segment. A spectral (Hirschfeld–Gebelein–Rényi) expansion of the strategy’s win kernel bounds the same polynomial on a large complex disk, and a two-constants (harmonic majorization) lemma transfers the anchored bound on the segment to the point 1, where the profile dominates the value of the original strategy.

Theorem (Main). *For every two-player one-round game $G = (\mu, V)$ with entangled value $\text{val}^*(G) < 1$ there are constants $c = c(G) > 0$ and $C = C(G) < \infty$ such that*

$$\text{val}^*(G^{\otimes n}) \leq Ce^{-cn} \quad \text{for all } n \geq 1.$$

Overview of the proof

Given a strategy S for $G^{\otimes N}$, we build from it a one-parameter family of strategies for the M -fold repetition of the *anchored* game $G_{\perp\beta}$ of Bavarian–Vidick–Yuen (BVY), by routing the M anchored coordinates to random slots of S and filling the slots that received an anchor symbol $\perp$ with junk questions; a tunable “flag” rate φ lets the $\perp$ -player replace the correlated junk by fresh independent junk. The acceptance probability of this family is shown to equal $\mathcal{G}(B)$, where $\mathcal{G}$ is a real polynomial of degree $\leq M$ — the *occupancy profile* of S — and $B = B(\varphi)$ sweeps the interval $[0, 1 - q]$ (q a fixed constant of the construction). BVY’s anchored parallel repetition theorem caps the entire realizable segment: $\mathcal{G} \leq e^{-\delta_B M}$ on $[0, 1 - q]$. On the other hand $\mathcal{G}(1) \geq \text{val}(S)$, because at $B = 1$ no slot is decorrelated and winning all coordinates of $G^{\otimes N}$ forces winning the routed ones.

The bridge between the segment and the point $B = 1$ is analytic: expanding the strategy’s win kernel in the singular (Hirschfeld–Gebelein–Rényi) basis of μ — whose correlation spectrum is $\leq \rho = \rho_{\max}(\mu) < 1$ on connected supports — yields an operator Parseval ledger for the profile’s coefficients, and a hypergeometric moment bound converts it into $|\mathcal{G}| \leq b^M$ on a complex disk of radius $R = 1 + \ln(1/\rho) \cdot \frac{N-M+1}{M}$, which is large when $N \gg M$. A two-constants (harmonic majorization) lemma then forces

$$\text{val}(S) \leq \mathcal{G}(1) \leq e^{-\delta_B M} \cdot (b^M e^{\delta_B M})^{4\sqrt{q}/\ln R},$$

and choosing $M = \Theta(N)$ with the implied constant small enough that $\ln R$ is large makes the correction factor a fractional power: $\text{val}(S) \leq e^{-\frac{3}{4}\delta_B M}$. This holds for *every* strategy, giving

exponential decay. Disconnected question distributions are handled first by a component reduction. The only results imported from the literature are standard (Naimark dilation; Chernoff; Perron–Frobenius together with Witsenhausen’s characterization of maximal correlation; the conformal map of the exterior of a segment) plus one substantial citation: the anchored parallel repetition theorem of Bavarian, Vidick and Yuen.

Notation. $\mathcal{X}, \mathcal{Y}, \mathcal{A}, \mathcal{B}$ are the question/answer sets, μ_A, μ_B the marginals of μ , $\text{supp } \mu \subseteq \mathcal{X} \times \mathcal{Y}$ the support, $\varepsilon_0 := 1 - \text{val}^*(G)$, $v_n := \text{val}^*(G^{\otimes n})$, $m := \min(|\mathcal{A}|, |\mathcal{B}|)$.

1 Preliminaries

A *strategy* S for a game $H = (\eta, W)$ is a tuple $(\psi, \{A_a^x\}_a, \{B_b^y\}_b)$: finite-dimensional Hilbert spaces $\mathcal{H}_A, \mathcal{H}_B$, a unit vector $\psi \in \mathcal{H}_A \otimes \mathcal{H}_B$, and for each question x (resp. y) a POVM $\{A_a^x\}_a$ on $\mathcal{H}_A$ (resp. $\{B_b^y\}_b$ on $\mathcal{H}_B$); its value is

$$\text{val}(S) = \mathbb{E}_{(x,y) \sim \eta} \sum_{a,b} W(x, y, a, b) \langle \psi | A_a^x \otimes B_b^y | \psi \rangle.$$

Allowing POVMs does not change val^* (Naimark dilation, which preserves finite-dimensionality). Two standard facts are used freely.

(1a) Shared randomness. For strategies $(S_r)_{r \in R}$ (R finite) and a distribution p on R there is a strategy S with $\text{val}(S) = \sum_r p(r) \text{val}(S_r)$: take direct sums $\mathcal{H}_A = \bigoplus_r \mathcal{H}_{A,r}$, $\mathcal{H}_B = \bigoplus_r \mathcal{H}_{B,r}$, state $\psi = \sum_r \sqrt{p(r)} \psi_r$ in the diagonal blocks, and blockwise measurements; block-orthogonality on a single side already kills all cross terms, so every outcome probability is the p -mixture. Consequently the players may use arbitrary finite shared classical randomness, and local (possibly randomness-dependent) pre/post-processing of questions and answers.

Scope of (1a). The index r must be a *common* random variable, drawn independently of the questions and available to both players before they see their questions; the block structure is what makes the cross terms vanish. In particular (1a) does *not* apply when the index is a function of the questions (each player would then select a different block on different question pairs, and the cross terms are no longer controlled). This distinction is used in the proof of Lemma 2.1, where the relevant index is question-determined and a tensor product, not a direct sum, is the correct construction.

(1b) Advice lemma. Let (λ_A, λ_B) be random variables with an arbitrary joint law, independent of $(x, y) \sim \eta$, let ψ be any unit vector and let $\{A_a^{x, \lambda_A}\}$, $\{B_b^{y, \lambda_B}\}$ be POVM families as above. Then

$$\mathbb{E}_{(\lambda_A, \lambda_B)} \mathbb{E}_{(x,y) \sim \eta} \sum_{a,b} W(x, y, a, b) \langle \psi | A_a^{x, \lambda_A} \otimes B_b^{y, \lambda_B} | \psi \rangle \leq \text{val}^*(H).$$

Proof: conditioning on (λ_A, λ_B) gives a bona fide strategy for H (the advice is independent of the questions), of value $\leq \text{val}^*(H)$; now average. $\square$

For a strategy S for $G^{\otimes N}$ and fixed question strings $(\bar{x}, \bar{y})$, the answers are produced by the single product measurement $\{A_a^{\bar{x}} \otimes B_b^{\bar{y}}\}$; for $T \subseteq [N]$ the probability of the event “the answers win all coordinates of T ” is $\sum_{(\bar{a}, \bar{b}) \text{ wins } T} \langle \psi | A_{\bar{a}}^{\bar{x}} \otimes B_{\bar{b}}^{\bar{y}} | \psi \rangle$.

2 Reduction to connected supports

View $\text{supp } \mu$ as a bipartite graph on $\mathcal{X} \sqcup \mathcal{Y}$ (deleting questions of zero marginal). Its connected components induce partitions $\mathcal{X} = \bigsqcup_{\ell} \mathcal{X}_{\ell}$, $\mathcal{Y} = \bigsqcup_{\ell} \mathcal{Y}_{\ell}$ with $\text{supp } \mu \subseteq \bigsqcup_{\ell} (\mathcal{X}_{\ell} \times \mathcal{Y}_{\ell})$; let $p_{\ell} = \mu(\mathcal{X}_{\ell} \times \mathcal{Y}_{\ell})$ and let G_{ℓ} be the component game (μ conditioned on the component, V restricted). The number of components is finite. Each player computes the component label of each coordinate from their own question: write $\ell_A(x)$ for the index with $x \in \mathcal{X}_{\ell_A(x)}$ and $\ell_B(y)$ for the index with $y \in \mathcal{Y}_{\ell_B(y)}$; since $\text{supp } \mu \subseteq \bigsqcup_{\ell} (\mathcal{X}_{\ell} \times \mathcal{Y}_{\ell})$, we have $\ell_A(x) = \ell_B(y)$ almost surely, and we denote this common value by $\ell(x, y)$.

We record explicitly the two features of this labelling that the proof of Lemma 2.1 uses.

(L1) *The label is **locally computable**.* Alice's label map $\ell_A : \mathcal{X} \rightarrow \{\text{components}\}$ depends on x only; Bob's map $\ell_B : \mathcal{Y} \rightarrow \{\text{components}\}$ depends on y only. Neither player needs the other's question, and no communication is involved: the partitions $\mathcal{X} = \bigsqcup_{\ell} \mathcal{X}_{\ell}$ and $\mathcal{Y} = \bigsqcup_{\ell} \mathcal{Y}_{\ell}$ are fixed data of the game G , known to both players before the questions are sent.

(L2) *The two labels **agree** μ -almost surely.* If $\mu(x, y) > 0$ then $(x, y) \in \text{supp } \mu \subseteq \bigsqcup_{\ell} (\mathcal{X}_{\ell} \times \mathcal{Y}_{\ell})$, so $x \in \mathcal{X}_{\ell}$ and $y \in \mathcal{Y}_{\ell}$ for one and the same ℓ ; that is, $\ell_A(x) = \ell_B(y) = \ell(x, y)$. Consequently the events $E_{\ell} := \{(x, y) : \ell(x, y) = \ell\} = (\mathcal{X}_{\ell} \times \mathcal{Y}_{\ell}) \cap \text{supp } \mu$ partition $\text{supp } \mu$, $\mu(E_{\ell}) = p_{\ell}$, and for $p_{\ell} > 0$ the conditional law $\mu_{\ell} := \mu(\cdot | E_{\ell})$ is exactly the question distribution of the component game G_{ℓ} , whose question sets are $\mathcal{X}_{\ell}, \mathcal{Y}_{\ell}$ and whose predicate is $V_{\ell} := V|_{\mathcal{X}_{\ell} \times \mathcal{Y}_{\ell} \times A \times B}$.

Lemma 2.1. $\text{val}^*(G) = \sum_{\ell} p_{\ell} \text{val}^*(G_{\ell})$. In particular, if $\text{val}^*(G) < 1$ there is ℓ^* with $p_{\ell^*} > 0$ and $\text{val}^*(G_{\ell^*}) < 1$.

Proof. Throughout, sums over ℓ are over the (finitely many) components with $p_{\ell} > 0$; components with $p_{\ell} = 0$ contribute nothing to either side.

“ $\leq$ ”: splitting one strategy for G into one strategy per component, by conditioning on the label. Let $S = (\psi, \{A_a^x\}_{x \in \mathcal{X}}, \{B_b^y\}_{y \in \mathcal{Y}})$ be an arbitrary strategy for G , on finite-dimensional spaces $\mathcal{H}_A, \mathcal{H}_B$ with $\psi \in \mathcal{H}_A \otimes \mathcal{H}_B$.

Step 1 (the restricted data is a legal strategy for G_{ℓ}). Fix ℓ with $p_{\ell} > 0$ and define

$$S_{\ell} := (\psi, \{A_a^x\}_{x \in \mathcal{X}_{\ell}}, \{B_b^y\}_{y \in \mathcal{Y}_{\ell}}),$$

i.e. keep the same Hilbert spaces and the same shared state, and simply forget the measurements attached to questions outside the component. This is a legal strategy for G_{ℓ} : the question sets of G_{ℓ} are $\mathcal{X}_{\ell}$ and $\mathcal{Y}_{\ell}$, for each $x \in \mathcal{X}_{\ell}$ the family $\{A_a^x\}_a$ is a POVM on $\mathcal{H}_A$, for each $y \in \mathcal{Y}_{\ell}$ the family $\{B_b^y\}_b$ is a POVM on $\mathcal{H}_B$, and ψ is a unit vector of $\mathcal{H}_A \otimes \mathcal{H}_B$. It is here that (L1) is used: upon receiving x , Alice knows $\ell_A(x) = \ell$ and hence knows that the measurement she is about to perform is “the S_{ℓ} measurement”; she never needs Bob's label, and symmetrically for Bob. (No new randomness, no communication and no modification of ψ occur in this step — the splitting is purely a relabelling of the data of S according to a quantity each player computes from their own question.)

Step 2 (exact decomposition of the value). By (L2) the events $(E_{\ell})_{\ell}$ partition $\text{supp } \mu$, $\mu(E_{\ell}) = p_{\ell}$, and conditionally on E_{ℓ} the pair (x, y) is distributed as μ_{ℓ} . Splitting the expectation defining $\text{val}(S)$ over this partition,

$$\text{val}(S) = \sum_{\ell} \mu(E_{\ell}) \cdot \mathbb{E}_{(x,y) \sim \mu_{\ell}} \sum_{a,b} V(x, y, a, b) \langle \psi | A_a^x \otimes B_b^y | \psi \rangle = \sum_{\ell} p_{\ell} \text{val}(S_{\ell}),$$

the last equality being the definition of $\text{val}(S_\ell)$ together with the fact that on $\mathcal{X}_\ell \times \mathcal{Y}_\ell$ the predicate V coincides with the predicate V_ℓ of G_ℓ (again (L2)). Note that this is an *identity*, not an inequality: conditioning on the label splits the value of S exactly into the p_ℓ -weighted values of the component strategies S_ℓ .

Step 3 (cap each piece and take a supremum). Each S_ℓ is a legal (finite-dimensional) strategy for G_ℓ by Step 1, so $\text{val}(S_\ell) \leq \text{val}^*(G_\ell)$. Hence $\text{val}(S) \leq \sum_\ell p_\ell \text{val}^*(G_\ell)$ for every strategy S of G , and taking the supremum over S gives $\text{val}^*(G) \leq \sum_\ell p_\ell \text{val}^*(G_\ell)$.

Step 4 (why this does not already give “ $\geq$ ”). The identity of Step 2 produces, from a single S , a *tuple* $(S_\ell)_\ell$ of component strategies sharing one pair of Hilbert spaces and one state ψ . Conversely, an arbitrary tuple of component strategies — in particular a tuple of near-optimal ones, each living on its own spaces with its own entangled state — need not arise in this way from any single S . To obtain “ $\geq$ ” we must therefore *build* a single strategy for G that realizes a prescribed tuple simultaneously; this is the content of the next paragraph.

“ $\geq$ ”: assembling one strategy for G out of one near-optimal strategy per component.

We use a *tensor product* of near-optimal component strategies, one factor per component; note that (1a) is not applicable here, because the component label is determined by the questions rather than by question-independent shared randomness (see the scope remark after (1a)).

Fix $\varepsilon > 0$ and, for each component ℓ with $p_\ell > 0$, a strategy $S_\ell = (\psi_\ell, \{A_{a,\ell}^x\}_{x \in \mathcal{X}_\ell}, \{B_{b,\ell}^y\}_{y \in \mathcal{Y}_\ell})$ for G_ℓ on spaces $\mathcal{H}_{A,\ell} \otimes \mathcal{H}_{B,\ell}$ with $\text{val}(S_\ell) \geq \text{val}^*(G_\ell) - \varepsilon$. Put

$$\mathcal{H}_A := \bigotimes_{\ell} \mathcal{H}_{A,\ell}, \quad \mathcal{H}_B := \bigotimes_{\ell} \mathcal{H}_{B,\ell}, \quad \psi := \bigotimes_{\ell} \psi_\ell,$$

where in the last expression the factors are regrouped so that all $\mathcal{H}_{A,\ell}$ sit in $\mathcal{H}_A$ and all $\mathcal{H}_{B,\ell}$ in $\mathcal{H}_B$ (this regrouping is a fixed unitary identification and is performed once, before any questions are received; it is legal because each ψ_ℓ is a fixed state shared in advance). For $x \in \mathcal{X}$ define the POVM on $\mathcal{H}_A$

$$A_a^x := A_{a,\ell_A(x)}^x \otimes \bigotimes_{\ell' \neq \ell_A(x)} I_{\mathcal{H}_{A,\ell'}},$$

i.e. Alice measures $S_{\ell_A(x)}$'s POVM on the factor labelled by her own question's component and does nothing elsewhere; $\sum_a A_a^x = I$, so this is a POVM. (Local computability (L1) is what makes this definition legal: the factor on which Alice acts is selected by $\ell_A(x)$, a function of her own question alone.) Define B_b^y symmetrically using $\ell_B(y)$.

For $(x, y) \in \text{supp } \mu$ we have, by (L2), $\ell_A(x) = \ell_B(y) =: \ell$, so both players act on the *same* factor ℓ and trivially on the others; since $\psi = \bigotimes_{\ell'} \psi_{\ell'}$ is a product state across the factors and $\langle \psi_{\ell'} | I \otimes I | \psi_{\ell'} \rangle = 1$,

$$\langle \psi | A_a^x \otimes B_b^y | \psi \rangle = \langle \psi_\ell | A_{a,\ell}^x \otimes B_{b,\ell}^y | \psi_\ell \rangle.$$

Therefore, grouping the expectation over $(x, y) \sim \mu$ according to the component (i.e. over the partition $(E_\ell)_\ell$ of (L2), exactly as in Step 2 above),

$$\text{val}(S) = \sum_{\ell} p_{\ell} \mathbb{E}_{(x,y) \sim \mu_{\ell}} \sum_{a,b} V(x, y, a, b) \langle \psi_{\ell} | A_{a,\ell}^x \otimes B_{b,\ell}^y | \psi_{\ell} \rangle = \sum_{\ell} p_{\ell} \text{val}(S_{\ell}) \geq \sum_{\ell} p_{\ell} \text{val}^*(G_{\ell}) - \varepsilon,$$

using $\sum_{\ell} p_{\ell} \leq 1$ in the last step. Since S is a legal finite-dimensional strategy for G , $\text{val}^*(G) \geq \text{val}(S) \geq \sum_{\ell} p_{\ell} \text{val}^*(G_{\ell}) - \varepsilon$; letting $\varepsilon \downarrow 0$ gives “ $\geq$ ”.

The “in particular”. If every component with $p_\ell > 0$ had $\text{val}^*(G_\ell) = 1$, the “ $\geq$ ” direction would give $\text{val}^*(G) \geq \sum_\ell p_\ell = 1$, contradicting $\text{val}^*(G) < 1$. Hence some ℓ^* with $p_{\ell^*} > 0$ has $\text{val}^*(G_{\ell^*}) < 1$. $\square$

Remark 2.2. Only the “in particular” clause of Lemma 2.1 (existence of a component ℓ^* with $p_{\ell^*} > 0$ and $\text{val}^*(G_{\ell^*}) < 1$) is used downstream, namely in Proposition 2.3; that clause rests on the “ $\geq$ ” direction, proved above by the tensor-product construction. The “ $\leq$ ” direction — the per-component splitting of an arbitrary strategy by conditioning on the locally computable label — is recorded only for completeness, and its Step 2 identity $\text{val}(S) = \sum_\ell p_\ell \text{val}(S_\ell)$ is also what makes transparent why the two directions require different constructions (Step 4).

Proposition 2.3. *If the Main Theorem holds for every connected game with entangled value < 1 , it holds for every game with $\text{val}^*(G) < 1$.*

Proof. Pick ℓ^* as in Lemma 2.1; let C^*, c^* be the constants for the connected game G_{ℓ^*} . Fix a strategy S for $G^{\otimes n}$ and condition on the label vector $\bar{\ell}$; let $T(\bar{\ell}) = \{i : \ell_i = \ell^*\}$. Given $\bar{\ell}$, the questions are independent across coordinates, the $T(\bar{\ell})$ -questions are i.i.d. μ_{ℓ^*} and independent of the others, and winning all coordinates implies winning all of $T(\bar{\ell})$. Giving the off- T questions to the players as advice ($\lambda_A = \bar{x}_{T^c}$, $\lambda_B = \bar{y}_{T^c}$, jointly distributed, independent of the T -questions) and reading off the T -answers defines an advised strategy for $G_{\ell^*}^{\otimes |T|}$ whose winning probability equals $\Pr[\text{win all of } T \mid \bar{\ell}]$; by (1b), this is $\leq v_{|T|}(G_{\ell^*}) \leq C^* e^{-c^*|T|}$. Averaging over $|T(\bar{\ell})| \sim \text{Bin}(n, p_{\ell^*})$ and using the Chernoff bound $\Pr[|T| < np_{\ell^*}/2] \leq e^{-np_{\ell^*}/8}$ gives

$$\text{val}(S) \leq C^* e^{-c^* p_{\ell^*} n/2} + e^{-p_{\ell^*} n/8}. \quad \square$$

From now on G is connected, with full marginals.

3 The singular basis of μ

Lemma 3.1. *There are real orthonormal bases $\{\phi_i\}_{i \geq 0}$ of $L^2(\mathcal{X}, \mu_A)$ and $\{\psi_{i'}\}_{i' \geq 0}$ of $L^2(\mathcal{Y}, \mu_B)$ with $\phi_0 = \psi_0 \equiv 1$ and*

$$\mathbb{E}_{(x,y) \sim \mu} [\phi_i(x) \psi_{i'}(y)] = \varrho_i \delta_{ii'}, \quad 1 = \varrho_0 > \rho := \varrho_1 \geq \varrho_2 \geq \dots \geq 0$$

(rectangular convention: $\delta_{ii'}$ fires for in-range equal indices; $\varrho_i := 0$ beyond the rank). Here $\rho = \rho_{\max}(\mu) < 1$, the maximal correlation of μ ; $\rho < 1$ precisely because $\text{supp } \mu$ is connected.

Proof. Let $Q_{xy} = \mu(x, y) / \sqrt{\mu_A(x) \mu_B(y)}$ and take a real singular value decomposition $Q = U \Sigma V^T$; set $\phi_i(x) = U_{xi} / \sqrt{\mu_A(x)}$ and $\psi_{i'}(y) = V_{yi'} / \sqrt{\mu_B(y)}$. Then $\{\phi_i\}$, $\{\psi_{i'}\}$ are orthonormal in the stated spaces and $\mathbb{E}_\mu[\phi_i \psi_{i'}] = (U^T Q V)_{ii'} = \varrho_i \delta_{ii'}$. The vectors $u_0(x) = \sqrt{\mu_A(x)}$, $v_0(y) = \sqrt{\mu_B(y)}$ satisfy $Q v_0 = u_0$, $Q^T u_0 = v_0$, and $\|Q\| \leq 1$ (conditional expectation is an L^2 -contraction), so (u_0, v_0) may be taken as the top singular pair, giving $\phi_0 = \psi_0 = 1$. Multiplicity of the singular value 1: $Q^T Q$ is conjugate to the two-step Markov chain $y \rightarrow x \rightarrow y'$ with stationary law μ_B ; its eigenvalue-1 multiplicity equals its number of communicating classes (Perron–Frobenius), which equals the number of connected components of $\text{supp } \mu$ (Witsenhausen). Connectedness gives multiplicity 1, i.e. $\varrho_1 < 1$; and ϱ_1 equals

$$\rho_{\max}(\mu) = \sup\{\mathbb{E}_\mu[fg] : \mathbb{E}f = \mathbb{E}g = 0, \|f\|_{L^2(\mu_A)} = \|g\|_{L^2(\mu_B)} = 1\}$$

by the variational principle for singular values. $\square$

For a finite index set U and $S \in \{0, 1, 2, \dots\}^U$ put $\Phi_S(\bar{x}_U) = \prod_{j \in U} \phi_{S_j}(x_j)$, $\Psi_S(\bar{y}_U) = \prod_{j \in U} \psi_{S_j}(y_j)$, $\text{supp } S = \{j : S_j \neq 0\}$, $w(S) = |\text{supp } S|$, and $\varrho^{(S)} = \prod_{j \in \text{supp } S} \varrho_{S_j} \leq \rho^{w(S)}$. Then $\{\Phi_S\}_S$ is an orthonormal basis of $L^2(\mathcal{X}^U, \mu_A^{\otimes U})$, and similarly $\{\Psi_S\}_S$ of $L^2(\mathcal{Y}^U, \mu_B^{\otimes U})$.

4 Anchoring

For $\beta \in (0, 1)$, the *anchored game* $G_{\perp\beta} = (\mu_{\perp}, V_{\perp})$ of BVY has question sets $\mathcal{X} \cup \{\perp\}$ and $\mathcal{Y} \cup \{\perp\}$; a question pair is sampled by drawing $(x, y) \sim \mu$ and independently replacing x by $\perp$ with probability β and y by $\perp$ with probability β ; the predicate accepts whenever either question is $\perp$, and equals V otherwise.

Lemma 4.1. $\text{val}^*(G_{\perp\beta}) = 1 - (1 - \beta)^2 \varepsilon_0 < 1$.

Proof. “ $\geq$ ”: play a near-optimal strategy for G , answering arbitrarily at $\perp$; rejection requires both questions real and V failing. “ $\leq$ ”: conditioned on both questions being real the questions are $\sim \mu$ and the restricted data is a strategy for G ; the value is $1 - (1 - \beta)^2(1 - \text{that conditional value})$. $\square$

Theorem 4.2 (BVY, cited). *For every anchored game H (two-sided anchoring probability $\beta \in (0, 1)$) with $\text{val}^*(H) < 1$ there are $\delta_B = \delta_B(H) > 0$ and $M_1 = M_1(H)$ such that $\text{val}^*(H^{\otimes M}) \leq e^{-\delta_B M}$ for all $M \geq M_1$.*

This is the anchored parallel repetition theorem of Bavarian, Vidick and Yuen (“Hardness amplification for entangled games via anchoring”, STOC 2017), with explicit $\delta_B = \text{poly}(\beta, 1 - \text{val}^*(H)) / \log(|\mathcal{A}||\mathcal{B}|)$.

Form of the imported statement, and what we use from it. Since Theorem 4.2 is the unique non-elementary input of this paper, we record precisely the hypotheses we verify and the form of the conclusion we consume.

Hypotheses verified at the point of use.

- (B1) *Two-player one-round game with finite alphabets.* $H = G_{\perp\beta}$ has question sets $\mathcal{X} \cup \{\perp\}$, $\mathcal{Y} \cup \{\perp\}$ and answer sets $\mathcal{A}, \mathcal{B}$, all finite.
- (B2) *Two-sided β -anchoring.* The question distribution $\mu_{\perp}$ is exactly of BVY’s anchored form: it is obtained from μ by replacing each player’s question independently by the dedicated anchor symbol $\perp$ with probability β , and the predicate accepts automatically whenever at least one question is $\perp$. This is the construction of §4 verbatim, with $\beta = \frac{3}{4}$ fixed below.
- (B3) *Strictly sub-unit entangled value.* $\text{val}^*(G_{\perp\beta}) = 1 - (1 - \beta)^2 \varepsilon_0 < 1$ by Lemma 4.1, using $\varepsilon_0 > 0$.
- (B4) *Finite-dimensional (tensor-product) strategies.* val^* is the supremum over finite-dimensional strategies as defined in §1; this is the class in which BVY’s theorem is stated, and the class in which all our strategies live (Naimark dilation preserves finite-dimensionality; Lemma 5.1 and (1a) produce finite-dimensional strategies).

Form of the conclusion consumed. We use only: there exist $\delta_B > 0$ and $M_1 < \infty$, depending on $H = G_{\perp\beta}$ alone (through $\beta, 1 - \text{val}^*(H)$ and $|\mathcal{A}|, |\mathcal{B}|$), such that

$$\text{val}^*(H^{\otimes M}) \leq e^{-\delta_B M} \quad \text{for every integer } M \geq M_1.$$

Two points deserve emphasis, as they are what the later argument needs.

- *Constants independent of the repetition number.* δ_B and M_1 do not depend on M ; they are fixed once and for all in the present section, *before* N , M , the flag rate φ , and the strategy S are introduced. In particular the bound is uniform in $M \geq M_1$, which is essential since in §8 we let $M = \lfloor \theta_0(N+1) \rfloor \rightarrow \infty$.
- *Strategy-independence.* $\text{val}^*(H^{\otimes M})$ is a supremum over *all* finite-dimensional strategies for $H^{\otimes M}$; we apply it to the particular strategies $\mathcal{R}_\varphi[S]$ built in §5, whose dimensions grow with N . No property of $\mathcal{R}_\varphi[S]$ beyond legality is needed.
- *Prefactors are harmless.* If the cited bound is quoted in the form $\text{val}^*(H^{\otimes M}) \leq C_B e^{-\delta M}$ with $C_B < \infty$, Theorem 4.2 as stated follows with $\delta_B := \delta/2$ and $M_1 := \max(M_1, \lceil 2 \ln \max(C_B, 1)/\delta \rceil)$, since then $C_B e^{-\delta M} \leq e^{-\delta M/2}$. Likewise, any bound of the form $\text{val}^*(H^{\otimes M}) \leq (1 - \delta')^M$ gives Theorem 4.2 with $\delta_B = \ln \frac{1}{1-\delta'}$. Thus the exponential-in- M shape, with constants independent of M , is the only feature used.

We fix from now on

$$\beta := \frac{3}{4}, \quad \ell := (1 - \beta)^2 = \frac{1}{16}, \quad q := \Pr[\text{one-sided} \mid \text{not live}] = \frac{2\beta(1 - \beta)}{1 - \ell} = \frac{2}{5},$$

and let $\delta_B = \delta_B(G, \beta) \in (0, 1]$ and M_1 be as in Theorem 4.2 for $H = G_{\perp\beta}$ (shrinking δ_B to ensure $\delta_B \leq 1$ is harmless).

5 The flagged routed family

Fix a strategy $S = (\psi, \{A_a^{\bar{x}}\}, \{B_b^{\bar{y}}\})$ for $G^{\otimes N}$, a number $M \leq N$ of anchored coordinates, and a flag rate $\varphi \in [0, 1] \cap \mathbb{Q}$. Define a strategy $\mathcal{R}_\varphi[S]$ for $(G_{\perp\beta})^{\otimes M}$ as follows.

Shared randomness: a uniformly random injective map $\iota : [M] \rightarrow [N]$; i.i.d. junk pairs $(x'_i, y'_i) \sim \mu$ for $i \in [N]$; i.i.d. fresh letters $y''_i \sim \mu_B$ for $i \in [N]$; i.i.d. flags $F \subseteq [M]$ with $\Pr[c \in F] = \varphi$. All four ingredients are mutually independent, and independent of the anchored questions; since φ is rational all of them may be generated from finitely many uniform bits, so (1a) applies.

Play: Alice, holding anchored questions $(\tilde{x}_1, \dots, \tilde{x}_M)$, assembles $\bar{x} \in \mathcal{X}^N$ by: $\bar{x}_{\iota(c)} := \tilde{x}_c$ if $\tilde{x}_c \neq \perp$, else $\bar{x}_{\iota(c)} := x'_{\iota(c)}$; and $\bar{x}_i := x'_i$ for $i \notin \text{im } \iota$. Bob assembles $\bar{y}$ by: $\bar{y}_{\iota(c)} := \tilde{y}_c$ if $\tilde{y}_c \neq \perp$; if $\tilde{y}_c = \perp$ then $\bar{y}_{\iota(c)} := y''_{\iota(c)}$ if $c \in F$ and $\bar{y}_{\iota(c)} := y'_{\iota(c)}$ otherwise; and $\bar{y}_i := y'_i$ off the image. Each player measures S 's measurement for their assembled string and answers coordinate c with their $\iota(c)$ -th answer. By (1a) this is a legal finite-dimensional strategy for $(G_{\perp\beta})^{\otimes M}$: each player's map from (own anchored questions, shared randomness) to a question string of S and back from S 's answers to anchored answers is local pre/post-processing.

Classify each coordinate $c \in [M]$ by its anchor pattern: *live* (both real; probability ℓ), *A-sev* ($\tilde{x}_c = \perp$, $\tilde{y}_c$ real), *B-sev* (mirror), *null* (both $\perp$); the types are i.i.d. across c , independent of the underlying real pairs $(s_c, t_c) \sim \mu$ (i.i.d.), of ι , of F and of all junk. Set

$$T := \iota(\text{Live}), \quad C' := \iota([M] \setminus \text{Live}), \quad \Sigma := \iota(\text{ASev} \cup \text{BSev} \cup (\text{Null} \cap F)), \quad B := (1 - q)(1 - \varphi).$$

Lemma 5.1 (conditional law). *Conditionally on (types, F , ι), the assembled pair $(\bar{x}, \bar{y})$ is distributed as the product measure $\nu_\Sigma := \bigotimes_{i=1}^N \nu_i$ with $\nu_i = \mu_A \otimes \mu_B$ for $i \in \Sigma$ and $\nu_i = \mu$*

otherwise; the questions at the live slots are the live coordinates' real pairs; and the acceptance event of $\mathcal{R}_\varphi[S]$ is the event that the answers win all coordinates of T (a function of the T -components of the answers and of the T -questions only). Moreover, conditionally on $t := |\text{Live}|$:

- (i) T is a uniform t -subset of $[N]$;
- (ii) given T , C' is a uniform $(M - t)$ -subset of $[N] \setminus T$;
- (iii) given (T, C') , the set Σ contains each slot of C' independently with probability $1 - B$, and no other slot; moreover the family of indicators $(\mathbf{1}[i \in \Sigma])_{i \in C'}$ is independent of (T, C') itself, i.e. its conditional law given (T, C') is the i.i.d. Bernoulli($1 - B$) law regardless of which sets (T, C') are;
- (iv) the live questions $\bar{z}_T := (s_c, t_c)_{c \text{ live}} \sim \mu^{\otimes t}$, independent of (T, C', Σ) — more precisely, the conditional law of $\bar{z}_T$ given (t, T, C', Σ) is $\mu^{\otimes t}$ (the coordinates of $\bar{z}_T$ being indexed by T through ι), so that $\bar{z}_T$ may be integrated out independently of the combinatorial data.

Proof. The randomness, listed. The experiment is driven by the following mutually independent ingredients:

$$\underbrace{(\tau_c)_{c \in [M]}}_{\text{types, i.i.d.}}, \quad \underbrace{(s_c, t_c)_{c \in [M]}}_{\text{i.i.d. } \mu}, \quad \underbrace{\iota}_{\text{uniform injective}}, \quad \underbrace{F}_{\text{i.i.d. Bernoulli}(\varphi)}, \quad \underbrace{(x'_i, y'_i)_{i \in [N]}}_{\text{i.i.d. } \mu}, \quad \underbrace{(y''_i)_{i \in [N]}}_{\text{i.i.d. } \mu_B},$$

where $\tau_c \in \{\text{live}, \text{A-sev}, \text{B-sev}, \text{null}\}$ with probabilities $(1 - \beta)^2, \beta(1 - \beta), \beta(1 - \beta), \beta^2$, and the anchored question pair of coordinate c is obtained from (s_c, t_c) by replacing the halves dictated by τ_c with $\perp$. (This is exactly the BVY sampling of $\mu_{\perp}^{\otimes M}$, rewritten as “type $\otimes$ real pair”; the independence of τ_c from (s_c, t_c) is the definition of anchoring.)

Slot laws given (types, F , ι). A live slot carries $(s_c, t_c) \sim \mu$. An A-sev slot carries $(x'_{\iota(c)}, t_c)$: the junk half is $\sim \mu_A$ and independent of $t_c \sim \mu_B$, so the law is $\mu_A \otimes \mu_B$ (the y -half $y'_{\iota(c)}$ is unused). A B-sev slot carries $(s_c, y'_{\iota(c)})$ if $c \notin F$ and $(s_c, y''_{\iota(c)})$ if $c \in F$ — in both cases $\mu_A \otimes \mu_B$ (independent halves): the flag is a law-level no-op at B-sev coordinates. A null slot carries $(x'_{\iota(c)}, y'_{\iota(c)}) \sim \mu$ if $c \notin F$ (the two halves of one shared pair: exactly μ -correlated), and $(x'_{\iota(c)}, y''_{\iota(c)}) \sim \mu_A \otimes \mu_B$ if $c \in F$. Off-image slots carry $(x'_i, y'_i) \sim \mu$. Distinct slots use disjoint ingredients, so the joint law is the stated product ν_Σ . A coordinate with a $\perp$ is accepted automatically; a live coordinate c is accepted iff $V(s_c, t_c, a_{\iota(c)}, b_{\iota(c)}) = 1$ — so acceptance reads only the T -questions and the T -answers.

Conditional structure (i)–(iv). Define the *severance indicator* of a non-live coordinate c ,

$$\sigma_c := \mathbf{1}[\tau_c \in \{\text{A-sev}, \text{B-sev}\}] \vee \mathbf{1}[\tau_c = \text{null}, c \in F],$$

so that, by the slot laws just computed, $\Sigma = \iota(\{c : \tau_c \neq \text{live}, \sigma_c = 1\})$ is exactly the set of slots carrying the product law $\mu_A \otimes \mu_B$.

Let $\mathcal{L} := \{c : \tau_c = \text{live}\}$ and $t = |\mathcal{L}|$. Since the τ_c are i.i.d., conditionally on t the set $\mathcal{L}$ is a uniform t -subset of $[M]$, and conditionally on $\mathcal{L}$ the types $(\tau_c)_{c \notin \mathcal{L}}$ are i.i.d. with $\Pr[\text{one-sided}] = q$ and $\Pr[\text{null}] = 1 - q$ (these are the conditional probabilities given “not live”). The flags are i.i.d. (φ) and independent of the types, hence conditionally on $(t, \mathcal{L})$ the variables $(\sigma_c)_{c \notin \mathcal{L}}$ are i.i.d. Bernoulli with

$$\Pr[\sigma_c = 1] = q + (1 - q)\varphi = 1 - (1 - q)(1 - \varphi) = 1 - B,$$

and they are independent of $\mathcal{L}$ (their conditional law does not depend on which t -set $\mathcal{L}$ is) and of $(s_c, t_c)_{c \in [M]}$.

The map ι is uniform injective and independent of $((\tau_c), F, (s_c, t_c))$. Hence, conditionally on t : $(T, C') = (\iota(\mathcal{L}), \iota(\mathcal{L}^c))$ is a uniformly random ordered pair of disjoint subsets of $[N]$ of sizes $(t, M - t)$, which is exactly (i) and (ii). Transporting the indicators through the bijection $\iota|_{\mathcal{L}^c} : \mathcal{L}^c \rightarrow C'$ and using that $(\sigma_c)_{c \notin \mathcal{L}}$ is i.i.d. and independent of $(\mathcal{L}, \iota)$, we get: conditionally on (t, T, C') , the indicators $(\mathbf{1}[i \in \Sigma])_{i \in C'} = (\sigma_{\iota^{-1}(i)})_{i \in C'}$ are i.i.d. Bernoulli($1 - B$), with conditional law not depending on (T, C') , and $\Sigma \subseteq C'$; this is (iii). Finally the real pairs (s_c, t_c) are independent of $((\tau_c), \iota, F, \text{junk})$ and i.i.d. μ , so conditionally on (t, T, C', Σ) — all of which are functions of $((\tau_c), \iota, F)$ — the live pairs $\bar{z}_T = (s_c, t_c)_{c \in \mathcal{L}}$, re-indexed by T via ι , are i.i.d. μ ; this is (iv).

Implementability. Bob can carry out his rule: he knows his own $\perp$ -set, ι , F and all junk; note that he need not know *which* of his $\perp$ -coordinates are null — the flag rule is applied to all of them, and by the slot-law computation above it only has an effect (severing) at nulls. Likewise Alice needs only her own $\perp$ -set, ι and the junk (x'_i) . $\square$

Remark 5.2 (the joint independence structure, made explicit). It is worth isolating the probabilistic skeleton of Lemma 5.1, since every one of the assertions (i)–(iv) is a purely measure-theoretic consequence of it.

The underlying product space. All randomness of the experiment analysed above lives on the product probability space

$$\Omega := \underbrace{\mathcal{T}^{[M]}}_{\text{types } (\tau_c)} \times \underbrace{(\mathcal{X} \times \mathcal{Y})^{[M]}}_{\text{real pairs } (s_c, t_c)} \times \underbrace{\text{Inj}([M], [N])}_{\iota} \times \underbrace{\{0, 1\}^{[M]}}_{\text{flags } F} \times \underbrace{(\mathcal{X} \times \mathcal{Y})^{[N]}}_{\text{junk pairs } (x'_i, y'_i)} \times \underbrace{\mathcal{Y}^{[N]}}_{\text{fresh letters } (y''_i)}$$

equipped with the *product* of the six laws: i.i.d. type law on $\mathcal{T} = \{\text{live}, \text{A-sev}, \text{B-sev}, \text{null}\}$ with weights $((1 - \beta)^2, \beta(1 - \beta), \beta(1 - \beta), \beta^2)$; i.i.d. μ ; uniform on injections $[M] \rightarrow [N]$; i.i.d. Bernoulli(φ); i.i.d. μ ; i.i.d. μ_B . Thus:

(types, flags, ι , junk pairs, fresh letters, underlying real pairs) are jointly independent by construction,

and every assertion (i)–(iv) is a deterministic function of this independence. By “by construction” we mean that this joint independence is not derived from anything: the six blocks are the *inputs* of the experiment, declared independent in the “Shared randomness” paragraph of §5 (the four shared-randomness blocks ι , junk pairs, fresh letters, flags are mutually independent and independent of the anchored questions) together with the BVY sampling of $\mu_{\perp}^{\otimes M}$ (each anchored question pair is “type $\otimes$ real pair” with the type independent of the real pair, i.i.d. across c). Everything else in §5 — the anchored questions $(\tilde{x}_c, \tilde{y}_c)$, the classification into live/A-sev/B-sev/null, the sets $\mathcal{L}, T, C', \Sigma$, the severance indicators σ_c , the assembled strings $\bar{x}, \bar{y}$, the measurement outcomes’ conditional distribution — is a *deterministic* (measurable) function of the six blocks; in particular no feedback or circular dependence is created by the players’ actions, since each player’s assembled string is a deterministic function of their own anchored questions and of the shared randomness.

How each assertion follows. Write $\mathcal{F}_{\text{comb}} := \sigma((\tau_c)_c, \iota, F)$ for the “combinatorial” σ -algebra.

- *Slot laws and ν_{Σ} :* distinct slots consume disjoint coordinates of the junk/fresh/real-pair blocks, so conditionally on $\mathcal{F}_{\text{comb}}$ the N assembled pairs are independent, each with the law computed in the proof; this is exactly the product form ν_{Σ} .

- (i), (ii): these use only that $\iota \perp (\tau_c)_c$. Independence implies that conditioning on $t = |\mathcal{L}|$ leaves ι uniform on injections and leaves $\mathcal{L}$ a uniform t -subset of $[M]$; the pair $(\iota(\mathcal{L}), \iota(\mathcal{L}^c))$ is then a uniform ordered pair of disjoint subsets of $[N]$ of sizes $(t, M - t)$.
- (iii): σ_c is a deterministic function of $(\tau_c, \mathbf{1}[c \in F])$ alone. Since the pairs $(\tau_c, \mathbf{1}[c \in F])$ are i.i.d. across c and independent of ι and of the real pairs, conditionally on $(t, \mathcal{L})$ the variables $(\sigma_c)_{c \notin \mathcal{L}}$ are i.i.d. Bernoulli($1 - B$) with a law not depending on which t -set $\mathcal{L}$ is, and independent of ι ; pushing forward through the bijection $\iota|_{\mathcal{L}^c}$ gives (iii) verbatim, including the clause that the conditional law of the indicator family given (T, C') does not depend on (T, C') .
- (iv): the block $(s_c, t_c)_c$ is independent of the remaining five blocks, and (t, T, C', Σ) is $\mathcal{F}_{\text{comb}}$ -measurable; independence of a random vector from a σ -algebra is preserved under conditioning on that σ -algebra, so the conditional law of $(s_c, t_c)_{c \in \mathcal{L}}$ given (t, T, C', Σ) is still i.i.d. μ , i.e. $\mu^{\otimes t}$ after re-indexing by T through ι . This is what licenses integrating $\bar{z}_T$ out separately from the combinatorial data in Lemma 7.1(i).

No other independence is used anywhere downstream: Lemma 7.1(i) consumes exactly (i)–(iv) and the reduction of the acceptance event to “win all of T ”.

6 The spectral representation and the ledger

Fix t , a t -subset T , live questions $\bar{z}_T = (\bar{x}_T, \bar{y}_T)$, and write $U := [N] \setminus T$ and $n' := |U| = N - t$. Define the coarse POVMs on $\mathcal{H}_A$ (suppressing $\bar{z}_T$), for $\bar{a} \in \mathcal{A}^T$:

$$A_{\bar{a}}(\bar{x}_U) := \sum_{\bar{a}_U} A_{(\bar{a}, \bar{a}_U)}^{(\bar{x}_T, \bar{x}_U)}, \quad \text{and on } \mathcal{H}_B : \quad B_{\bar{b}}(\bar{y}_U) := \sum_{\bar{b}_U} B_{(\bar{b}, \bar{b}_U)}^{(\bar{y}_T, \bar{y}_U)},$$

so that $0 \preceq A_{\bar{a}}(\cdot) \preceq I$, $\sum_{\bar{a}} A_{\bar{a}} \equiv I$, and likewise for B . Let $E = E(\bar{z}_T) \subseteq \mathcal{A}^T \times \mathcal{B}^T$ be the set of answer tuples winning all of T , and put

$$B_{E_{\bar{a}}} := \sum_{\bar{b}: (\bar{a}, \bar{b}) \in E} B_{\bar{b}}, \quad A_{E_{\bar{b}}} := \sum_{\bar{a}: (\bar{a}, \bar{b}) \in E} A_{\bar{a}}$$

(so $0 \preceq B_{E_{\bar{a}}} \preceq I$ and $0 \preceq A_{E_{\bar{b}}} \preceq I$).

The win kernel. Define

$$\mathcal{W}_{T, \bar{z}_T}(\bar{x}_U, \bar{y}_U) := \sum_{(\bar{a}, \bar{b}) \in E} \langle \psi | A_{(\bar{a}, \cdot)}^{(\bar{x}_T, \bar{x}_U)} \otimes B_{(\bar{b}, \cdot)}^{(\bar{y}_T, \bar{y}_U)} | \psi \rangle \quad (6.1)$$

where the inner sum over the U -answers is the one implicit in the coarse POVMs; concretely,

$$\mathcal{W}_{T, \bar{z}_T}(\bar{x}_U, \bar{y}_U) = \sum_{\bar{a} \in \mathcal{A}^T} \langle \psi | A_{\bar{a}}(\bar{x}_U) \otimes B_{E_{\bar{a}}}(\bar{y}_U) | \psi \rangle = \sum_{\bar{b} \in \mathcal{B}^T} \langle \psi | A_{E_{\bar{b}}}(\bar{x}_U) \otimes B_{\bar{b}}(\bar{y}_U) | \psi \rangle. \quad (6.2)$$

Thus $\mathcal{W}_{T, \bar{z}_T}(\bar{x}_U, \bar{y}_U) \in [0, 1]$ is simply the probability that S 's answers win all coordinates of T when the assembled question strings are $(\bar{x}_T, \bar{x}_U)$ and $(\bar{y}_T, \bar{y}_U)$. It is an intrinsic object: it is defined from the strategy S and the data $(T, \bar{z}_T)$ alone, with no reference to whether one groups

the answers by $\bar{a}$ or by $\bar{b}$; the two expressions in (6.2) are two groupings of the same finite sum (6.1).

For $\Sigma \subseteq U$ the win- T probability under ν_Σ is

$$F_{T, \bar{z}_T}(\Sigma) := \mathbb{E}_{(\bar{x}_U, \bar{y}_U) \sim \nu_\Sigma} [\mathcal{W}_{T, \bar{z}_T}(\bar{x}_U, \bar{y}_U)] = \mathbb{E}_{(\bar{x}_U, \bar{y}_U) \sim \nu_\Sigma} \sum_{\bar{a} \in \mathcal{A}^T} \langle \psi | A_{\bar{a}}(\bar{x}_U) \otimes B_{E_{\bar{a}}}(\bar{y}_U) | \psi \rangle \in [0, 1].$$

Define the operator Fourier coefficients

$$\hat{A}_{\bar{a}, S} := \mathbb{E}_{\bar{x}_U \sim \mu_A^{\otimes U}} [A_{\bar{a}}(\bar{x}_U) \Phi_S(\bar{x}_U)], \quad \hat{B}_{\bar{a}, S} := \mathbb{E}_{\bar{y}_U \sim \mu_B^{\otimes U}} [B_{E_{\bar{a}}}(\bar{y}_U) \Psi_S(\bar{y}_U)]$$

(self-adjoint; finite sums), and the scalar Fourier coefficients of the win kernel with respect to the *product* reference measure $\mu_A^{\otimes U} \otimes \mu_B^{\otimes U}$:

$$\widehat{\mathcal{W}}_{S, S'} := \mathbb{E}_{\bar{x}_U \sim \mu_A^{\otimes U}} \mathbb{E}_{\bar{y}_U \sim \mu_B^{\otimes U}} [\mathcal{W}_{T, \bar{z}_T}(\bar{x}_U, \bar{y}_U) \Phi_S(\bar{x}_U) \Psi_{S'}(\bar{y}_U)] \in \mathbb{R}, \quad (6.3)$$

so that $\mathcal{W}_{T, \bar{z}_T} = \sum_{S, S'} \widehat{\mathcal{W}}_{S, S'} \Phi_S \Psi_{S'}$ pointwise (finite expansion in an orthonormal product basis). Since the two reference measures in (6.3) are independent, substituting (6.2) and expanding by bilinearity gives the *diagonal identity*

$$\widehat{\mathcal{W}}_{S, S} = \sum_{\bar{a} \in \mathcal{A}^T} \langle \psi | \hat{A}_{\bar{a}, S} \otimes \hat{B}_{\bar{a}, S} | \psi \rangle = \sum_{\bar{b} \in \mathcal{B}^T} \langle \psi | \hat{A}_{E_{\bar{b}}, S} \otimes \hat{B}_{\bar{b}, S} | \psi \rangle, \quad (6.4)$$

where on the right the hats denote the analogous Fourier coefficients of $A_{E_{\bar{b}}}(\cdot)$ and $B_{\bar{b}}(\cdot)$. In words: $\widehat{\mathcal{W}}_{S, S}$ is a function of the win kernel alone, and both the $\bar{a}$ -decomposition and the $\bar{b}$ -decomposition compute it. This observation is used in Lemma 6.2 to justify taking the minimum of the two alphabet bounds.

Lemma 6.1 (representation). *For every $\Sigma \subseteq U$,*

$$F_{T, \bar{z}_T}(\Sigma) = \sum_{S: \text{supp } S \cap \Sigma = \emptyset} \varrho^{(S)} \widehat{\mathcal{W}}_{S, S} = \sum_{\bar{a}} \sum_{S: \text{supp } S \cap \Sigma = \emptyset} \varrho^{(S)} \langle \psi | \hat{A}_{\bar{a}, S} \otimes \hat{B}_{\bar{a}, S} | \psi \rangle.$$

Proof. Expand $\mathcal{W}_{T, \bar{z}_T} = \sum_{S, S'} \widehat{\mathcal{W}}_{S, S'} \Phi_S \Psi_{S'}$ (equivalently, Fourier inversion in each matrix entry gives $A_{\bar{a}}(\bar{x}_U) = \sum_S \hat{A}_{\bar{a}, S} \Phi_S(\bar{x}_U)$ pointwise, and similarly for B). By independence across slots, $\mathbb{E}_{\nu_\Sigma}[\Phi_S \Psi_{S'}] = \prod_{j \in U} c_j(S_j, S'_j)$ with $c_j(s, s') = \varrho_s \delta_{ss'}$ at slots $j \notin \Sigma$ (Lemma 3.1) and $c_j(s, s') = \mathbb{E}_{\mu_A}[\phi_s] \mathbb{E}_{\mu_B}[\psi_{s'}] = \delta_{s0} \delta_{s'0}$ at $j \in \Sigma$. Thus only matched pairs $S = S'$ with $\text{supp } S \cap \Sigma = \emptyset$ survive, with coefficient $\varrho^{(S)}$. The second expression is (6.4). $\square$

Lemma 6.2 (ledger). $\sum_S |\widehat{\mathcal{W}}_{S, S}| \leq m^{t/2}$, where $m = \min(|\mathcal{A}|, |\mathcal{B}|)$.

Proof. Step 1 (the $\bar{a}$ -decomposition gives $|\mathcal{A}|^{t/2}$). Completeness of the basis gives $\sum_S \Phi_S(\bar{x}) \Phi_S(\bar{x}') = \delta_{\bar{x}\bar{x}'} / \mu_A^{\otimes U}(\bar{x})$, whence

$$\sum_S \hat{A}_{\bar{a}, S} \hat{A}_{\bar{a}, S}^\dagger = \mathbb{E}_{\bar{x}_U} [A_{\bar{a}}(\bar{x}_U)^2] \preceq \mathbb{E}_{\bar{x}_U} [A_{\bar{a}}(\bar{x}_U)]$$

(using $0 \preceq A \preceq I \Rightarrow A - A^2 = A^{1/2}(I - A)A^{1/2} \succeq 0$), and the same on Bob's side. Sandwiching with ψ :

$$\sum_S \|(\hat{A}_{\bar{a}, S}^\dagger \otimes I)\psi\|^2 \leq \alpha_{\bar{a}} := \langle \psi | \mathbb{E}[A_{\bar{a}}] \otimes I | \psi \rangle, \quad \sum_S \|(I \otimes \hat{B}_{\bar{a}, S})\psi\|^2 \leq \beta_{\bar{a}} := \langle \psi | I \otimes \mathbb{E}[B_{E_{\bar{a}}}] | \psi \rangle.$$

By Cauchy–Schwarz (twice), $\sum_S |\langle \psi | \hat{A}_{\bar{a},S} \otimes \hat{B}_{\bar{a},S} | \psi \rangle| \leq \sqrt{\alpha_{\bar{a}} \beta_{\bar{a}}}$, and $\sum_{\bar{a}} \sqrt{\alpha_{\bar{a}} \beta_{\bar{a}}} \leq \sqrt{\sum_{\bar{a}} \alpha_{\bar{a}}} \sqrt{\sum_{\bar{a}} \beta_{\bar{a}}}$ with $\sum_{\bar{a}} \alpha_{\bar{a}} = 1$ and $\sum_{\bar{a}} B_{E_{\bar{a}}} = \sum_{\bar{b}} \#\{\bar{a} : (\bar{a}, \bar{b}) \in E\} B_{\bar{b}} \preceq |\mathcal{A}|^t I$, so $\sum_{\bar{a}} \beta_{\bar{a}} \leq |\mathcal{A}|^t$. Hence, using (6.4) and the triangle inequality,

$$\sum_S |\widehat{\mathcal{W}}_{S,S}| \leq \sum_{\bar{a}} \sum_S |\langle \psi | \hat{A}_{\bar{a},S} \otimes \hat{B}_{\bar{a},S} | \psi \rangle| \leq |\mathcal{A}|^{t/2}.$$

Step 2 (the $\bar{b}$ -decomposition gives $|\mathcal{B}|^{t/2}$). Repeat Step 1 with the roles of the players exchanged, i.e. using the second grouping in (6.2): $\sum_{\bar{b}} \langle \psi | A_{E_{\bar{b}}} \otimes B_{\bar{b}} | \psi \rangle$, with $\sum_{\bar{b}} \langle \psi | I \otimes \mathbb{E}[B_{\bar{b}}] | \psi \rangle = 1$ and $\sum_{\bar{b}} A_{E_{\bar{b}}} \preceq |\mathcal{B}|^t I$. This yields $\sum_S |\widehat{\mathcal{W}}_{S,S}| \leq |\mathcal{B}|^{t/2}$.

Step 3 (taking the minimum is legitimate). The left-hand side $\sum_S |\widehat{\mathcal{W}}_{S,S}|$ is defined by (6.3) in terms of the win kernel only — it does not refer to any grouping of the answers. Steps 1 and 2 bound this *same* quantity by $|\mathcal{A}|^{t/2}$ and by $|\mathcal{B}|^{t/2}$ respectively, so both bounds hold simultaneously and $\sum_S |\widehat{\mathcal{W}}_{S,S}| \leq \min(|\mathcal{A}|, |\mathcal{B}|)^{t/2} = m^{t/2}$. $\square$

Since $\varrho^{(S)} \leq \rho^{w(S)}$, Lemma 6.2 yields, after averaging over a uniform t -subset T and $\bar{z}_T \sim \mu^{\otimes T}$, the **coefficient ledger** for the averaged coefficients below. (Convention: since $\varrho^{(S)} \leq \rho^{w(S)} \leq \rho'^{w(S)}$ for any $\rho' \in [\rho, 1)$, the ledger and everything downstream hold with any such ρ' in place of ρ ; we therefore assume without loss of generality that $\rho \geq \frac{1}{2}$, replacing ρ by $\max(\rho, \frac{1}{2})$ throughout. This also covers the degenerate case $\rho = 0$, e.g. product μ .) The ledger reads: with

$$c_w^{(t)} := \mathbb{E}_{T, \bar{z}_T} \left[\sum_{S: w(S)=w} \varrho^{(S)} \widehat{\mathcal{W}}_{S,S} \right] \stackrel{(6.4)}{=} \mathbb{E}_{T, \bar{z}_T} \left[\sum_{\bar{a}} \sum_{S: w(S)=w} \varrho^{(S)} \langle \psi | \hat{A}_{\bar{a},S} \otimes \hat{B}_{\bar{a},S} | \psi \rangle \right] \in \mathbb{R}, \quad (6.5)$$

one has

$$\sum_{w \geq 0} \rho^{-w} |c_w^{(t)}| \leq \mathbb{E}_{T, \bar{z}_T} \left[\sum_S \rho^{-w(S)} \varrho^{(S)} |\widehat{\mathcal{W}}_{S,S}| \right] \leq \mathbb{E}_{T, \bar{z}_T} \left[\sum_S |\widehat{\mathcal{W}}_{S,S}| \right] \leq m^{t/2}. \quad (6.6)$$

Remark 6.3 ($c_w^{(t)}$ is decomposition-independent). By (6.5) the coefficient $c_w^{(t)}$ is an average of the diagonal Fourier coefficients $\widehat{\mathcal{W}}_{S,S}$ of the win kernel, weighted by $\varrho^{(S)}$ and restricted to $w(S) = w$. Since $\widehat{\mathcal{W}}_{S,S}$ is determined by $\mathcal{W}_{T, \bar{z}_T}$ — an object built from S , T and $\bar{z}_T$ only — the quantities $c_w^{(t)}$ do not depend on whether one decomposes the win kernel by Alice’s answers or by Bob’s. Equivalently, $\Sigma \mapsto F_{T, \bar{z}_T}(\Sigma)$ is a fixed real function on subsets of U , and by Lemma 6.1 its multilinear (Fourier) expansion in the variables $(\mathbf{1}[j \in \Sigma])_{j \in U}$ has coefficients determined by the numbers $\varrho^{(S)} \widehat{\mathcal{W}}_{S,S}$; $c_w^{(t)}$ is the average of the total mass of those coefficients at Hamming weight w . This is what makes the “min” in Lemma 6.2 (Step 3) legitimate, and it is the only place where the two decompositions are compared.

7 The occupancy profile

Let $H_w^{(t)}$ denote a hypergeometric variable: the size of the intersection of a fixed w -subset of an $(N - t)$ -set with a uniformly random $(M - t)$ -subset. Define the real polynomial (of degree $\leq M$)

$$\mathcal{G}(z) := \sum_{t=0}^M \pi_t \tilde{Q}_t(z), \quad \tilde{Q}_t(z) := \sum_{w=0}^{N-t} c_w^{(t)} \mathbb{E}[z^{H_w^{(t)}}], \quad \pi_t = \binom{M}{t} \ell^t (1 - \ell)^{M-t}.$$

Lemma 7.1. (i) For every $\varphi \in [0, 1] \cap \mathbb{Q}$: $\text{val}(\mathcal{R}_\varphi[S]) = \mathcal{G}((1 - q)(1 - \varphi))$.

(ii) $0 \leq \mathcal{G}(B) \leq 1$ for all $B \in [0, 1]$.

(iii) $\mathcal{G}(1) \geq \text{val}(S)$.

(iv) If $M \geq M_1$ then $0 \leq \mathcal{G}(B) \leq e^{-\delta_B M}$ for all $B \in [0, 1 - q]$.

Proof. (i) Condition on t , then on $(T, C', \bar{z}_T)$ (Lemma 5.1(i),(ii),(iv)). For a fixed multi-index S with $W := \text{supp } S$, $|W| = w$, average the indicator $\mathbf{1}[W \cap \Sigma = \emptyset]$ appearing in Lemma 6.1 first over the severance indicators and then over C' . By Lemma 5.1(iii) the severance indicators within C' are i.i.d. Bernoulli($1 - B$) with conditional law not depending on (T, C') , so

$$\mathbb{E}[\mathbf{1}[W \cap \Sigma = \emptyset] \mid T, C'] = B^{|W \cap C'|},$$

and by Lemma 5.1(ii) C' is a uniform $(M - t)$ -subset of $U = [N] \setminus T$, so $|W \cap C'| \sim H_w^{(t)}$ for any fixed w -subset $W \subseteq U$ (exchangeability); hence

$$\mathbb{E}[\mathbf{1}[W \cap \Sigma = \emptyset] \mid T] = \mathbb{E}_{C'}[B^{|W \cap C'|}] = \mathbb{E}[B^{H_w^{(t)}}],$$

which depends only on (t, w) and is independent of $\bar{z}_T$ and of the spectral data (Lemma 5.1(iv) is what allows $\bar{z}_T$ to be averaged separately). All sums are finite, so by Lemma 6.1 and (6.5), $\mathbb{E}[\text{accept} \mid t] = \tilde{Q}_t(B)$; average over $t \sim \text{Bin}(M, \ell)$, which is the law of $|\text{Live}|$, with $\ell = (1 - \beta)^2$.

(ii) For any $B \in [0, 1]$, $\tilde{Q}_t(B)$ is the expectation of $F_{T, \bar{z}_T}(\Sigma) \in [0, 1]$ under the well-defined law “sever each slot of C' i.i.d. with probability $1 - B$ ” (no strategy-realizability is needed for this), by the same computation as in (i).

(iii) For every t ,

$$\tilde{Q}_t(1) = \sum_w c_w^{(t)} = \mathbb{E}_{T, \bar{z}_T}[F_{T, \bar{z}_T}(\emptyset)] = \mathbb{E}_T \Pr_{\mu^{\otimes N}}[\text{win all of } T] \geq \Pr_{\mu^{\otimes N}}[\text{win all}] = \text{val}(S),$$

since winning all coordinates implies winning T , pointwise.

(iv) We spell out the chain, as this is the step that combines the construction of §5 with the imported Theorem 4.2.

(1) *Legality.* For each rational $\varphi \in [0, 1]$, the routed family $\mathcal{R}_\varphi[S]$ of §5 is a legal finite-dimensional strategy for $(G_{\perp\beta})^{\otimes M}$: its shared randomness is finite (rationality of φ), the pre-processing (assembling $\bar{x}$, resp. $\bar{y}$) and post-processing (reading coordinate $\iota(c)$) are local, and the underlying state and measurements are those of S ; see (1a) and the paragraph “Play” in §5.

(2) *Its value is a value of the profile.* By part (i), $\text{val}(\mathcal{R}_\varphi[S]) = \mathcal{G}((1 - q)(1 - \varphi))$. Here the identification of the acceptance probability with $\tilde{Q}_t$ -averages rests on Lemma 5.1: the conditional product law ν_Σ (which makes the spectral representation Lemma 6.1 available), the reduction of the acceptance event to “win all of T ”, and the independence structure (iii)–(iv) (see also Remark 5.2).

(3) *The cap.* By definition of val^* , $\text{val}(\mathcal{R}_\varphi[S]) \leq \text{val}^*((G_{\perp\beta})^{\otimes M})$. The game $G_{\perp\beta}$ satisfies hypotheses (B1)–(B4) of §4 — it is a finite two-player one-round game, two-sidedly β -anchored by construction with $\beta = \frac{3}{4}$, and $\text{val}^*(G_{\perp\beta}) < 1$ by Lemma 4.1 — so Theorem 4.2 applies and gives $\text{val}^*((G_{\perp\beta})^{\otimes M}) \leq e^{-\delta_B M}$ for all $M \geq M_1$. Crucially, δ_B and M_1 were fixed in §4 as functions of G and β only: they do not depend on M , on N , on φ , or on the strategy S . Hence

$$\mathcal{G}((1 - q)(1 - \varphi)) \leq e^{-\delta_B M} \quad \text{for every rational } \varphi \in [0, 1], \text{ provided } M \geq M_1. \quad (7.1)$$

(4) *From a dense set of points to the whole segment.* From now on $M \geq M_1$ is fixed; in particular the right-hand side $e^{-\delta_B M}$ of (7.1) is a fixed constant, independent of φ . Let

$$\Theta : [0, 1] \longrightarrow [0, 1 - q], \quad \Theta(\varphi) := (1 - q)(1 - \varphi),$$

and let $\mathcal{B} := \Theta([0, 1] \cap \mathbb{Q}) = \{(1 - q)(1 - \varphi) : \varphi \in [0, 1] \cap \mathbb{Q}\}$ be the set of points of the segment that are realized as values of legal strategies by (7.1).

(4a) *Θ is an affine homeomorphism, so $\mathcal{B}$ is dense in $[0, 1 - q]$.* The map Θ is affine with slope $-(1 - q) \neq 0$ (recall $q = \frac{2}{5}$, so $1 - q = \frac{3}{5} > 0$), hence a decreasing bijection of $[0, 1]$ onto $[0, 1 - q]$, with continuous inverse $\Theta^{-1}(B) = 1 - \frac{B}{1 - q}$; both Θ and Θ^{-1} are Lipschitz. Consequently Θ maps dense subsets of $[0, 1]$ to dense subsets of $[0, 1 - q]$. Explicitly: given $B_0 \in [0, 1 - q]$, put $\varphi_0 := \Theta^{-1}(B_0) = 1 - \frac{B_0}{1 - q} \in [0, 1]$; since $\mathbb{Q}$ is dense in $\mathbb{R}$, choose rationals $\varphi_k \in [0, 1] \cap \mathbb{Q}$ with $\varphi_k \rightarrow \varphi_0$ (for instance $\varphi_k := \min\{\lceil k\varphi_0 \rceil / k, 1\}$, which lies in $[0, 1] \cap \mathbb{Q}$ and satisfies $|\varphi_k - \varphi_0| \leq 1/k$). Then $B_k := \Theta(\varphi_k) \in \mathcal{B}$ and

$$|B_k - B_0| = (1 - q) |\varphi_k - \varphi_0| \leq \frac{1 - q}{k} \xrightarrow{k \rightarrow \infty} 0.$$

(4b) *$\mathcal{G}$ is continuous.* $\mathcal{G}$ is, by its very definition at the head of this section, a finite real linear combination of the polynomials $z \mapsto \mathbb{E}[z^{H_w^{(t)}}] = \sum_h \Pr[H_w^{(t)} = h] z^h$ (finitely many $t \leq M$ and $w \leq N - t$, each expectation a finite sum), hence a polynomial of degree $\leq M$ with real coefficients; polynomials are continuous on $\mathbb{R}$ (indeed on $\mathbb{C}$).

(4c) *Passing to the limit.* Fix $B_0 \in [0, 1 - q]$ and let $(B_k) \subseteq \mathcal{B}$ with $B_k \rightarrow B_0$ as in (4a). By (7.1), $\mathcal{G}(B_k) \leq e^{-\delta_B M}$ for every k . By (4b), $\mathcal{G}(B_k) \rightarrow \mathcal{G}(B_0)$. Since the bound $e^{-\delta_B M}$ is a constant not depending on k , and since non-strict inequalities are preserved under limits,

$$\mathcal{G}(B_0) = \lim_{k \rightarrow \infty} \mathcal{G}(B_k) \leq e^{-\delta_B M}.$$

(4d) *Endpoints.* No limiting argument is in fact needed at the two endpoints of the segment: $\varphi = 0$ and $\varphi = 1$ are rational, so $B = 1 - q = \Theta(0)$ and $B = 0 = \Theta(1)$ already belong to $\mathcal{B}$ and are covered directly by (7.1). The approximation in (4a)–(4c) is therefore only needed at the (irrational) interior points $B_0 \in (0, 1 - q) \setminus \mathcal{B}$; in particular the limit points B_k used there may always be taken inside $[0, 1 - q]$, so no value of $\mathcal{G}$ outside the segment is ever invoked.

(4e) *The lower bound.* Finally, $[0, 1 - q] \subseteq [0, 1]$ (again $q = \frac{2}{5} \in (0, 1)$), so part (ii) — which holds for *every* real $B \in [0, 1]$, with no realizability by a strategy required — gives $\mathcal{G}(B) \geq 0$ there. Combining with (4c)–(4d),

$$0 \leq \mathcal{G}(B) \leq e^{-\delta_B M} \quad \text{for all } B \in [0, 1 - q],$$

which is the assertion of (iv). □

Remark 7.2 (the nexus). Lemma 7.1(iv) is the only point where the literature is used, and it is used exactly once, in the following shape: *an upper bound $e^{-\delta_B M}$, with $\delta_B > 0$ independent of M , on the value of every finite-dimensional strategy for $(G_{\perp\beta})^{\otimes M}$.* The role of the flag sweep of §5 is solely to make the *entire* segment $[0, 1 - q]$ consist of values of legal anchored strategies (up to the density/continuity step (4) above, which costs nothing since $\mathcal{G}$ is a polynomial), so that the single imported inequality becomes a bound on a continuum of points of the polynomial $\mathcal{G}$; this is what the analytic step of §8 requires (a bound at one point would be useless). Conversely, nothing about the internal structure of the BVY proof is used, and no property of $\mathcal{R}_\varphi[S]$ beyond legality enters.

Lemma 7.3 (disk bound). *Let*

$$R := 1 + \ln(1/\rho) \frac{N - M + 1}{M}, \quad b := 1 - \ell + \ell\sqrt{m} = 1 + \frac{\sqrt{m} - 1}{16} \quad (\ell = \frac{1}{16}).$$

Then $|\mathcal{G}(z)| \leq b^M$ for all $|z| \leq R$.

Proof. We claim that $\mathbb{E}[R^{H_w^{(t)}}] \leq \rho^{-w}$ for all $0 \leq t \leq M$ and $0 \leq w \leq N - t$. Write $n = N - t$ and $K = M - t$, and note that $n - K + 1 = N - M + 1$ for every t , so that, with R as displayed,

$$(R - 1) \frac{K}{n - K + 1} = \ln(1/\rho) \frac{N - M + 1}{M} \cdot \frac{M - t}{N - M + 1} = \ln(1/\rho) \frac{M - t}{M} \leq \ln(1/\rho). \quad (7.2)$$

Case $w \leq n - K + 1$: sample the K elements of C' sequentially; at each draw, the conditional probability of hitting the fixed w -set is at most $\frac{w}{n - K + 1}$ (at least $n - K + 1$ slots remain available and at most w of them are in the set). Hence H_w is stochastically dominated by $\text{Bin}(K, \frac{w}{n - K + 1})$, and since $y \mapsto y^h$ is increasing for $y \geq 1$,

$$\mathbb{E}[R^{H_w}] \leq \left(1 + (R - 1) \frac{w}{n - K + 1}\right)^K \leq \exp\left((R - 1) \frac{wK}{n - K + 1}\right) \stackrel{(7.2)}{\leq} e^{w \ln(1/\rho)} = \rho^{-w}.$$

Case $w > n - K + 1$: here $H_w \leq K$ always, and

$$K \ln R \leq K(R - 1) = \ln(1/\rho) K \frac{N - M + 1}{M} \leq \ln(1/\rho)(N - M + 1) < w \ln(1/\rho),$$

so $\mathbb{E}[R^{H_w}] \leq R^K \leq \rho^{-w}$.

Therefore, for $|z| \leq R$: $|\mathbb{E}[z^{H_w}]| \leq \mathbb{E}[R^{H_w}] \leq \rho^{-w}$, so $|\tilde{Q}_t(z)| \leq \sum_w |c_w^{(t)}| \rho^{-w} \leq m^{t/2}$ by (6.6), and

$$|\mathcal{G}(z)| \leq \sum_t \pi_t m^{t/2} = (1 - \ell + \ell\sqrt{m})^M = b^M,$$

using $\sum_t \binom{M}{t} \ell^t (1 - \ell)^{M - t} m^{t/2} = (1 - \ell + \ell\sqrt{m})^M$. $\square$

Remark 7.4 (calibration and robustness of the disk bound). The hypergeometric moment bound $\mathbb{E}[R^{H_w^{(t)}}] \leq \rho^{-w}$ matches the ledger weight ρ^{-w} of (6.6) exactly; this is by construction, since R was defined in Lemma 7.3 as the largest radius for which (7.2) holds for all t (the worst case being $t = 0$). The matching is therefore a choice of parameter, not a coincidence, and there is ample slack downstream, for three independent reasons.

(a) *The exponent ρ may be relaxed.* As recorded in the convention before (6.6), the ledger holds verbatim with any $\rho' \in [\rho, 1)$ in place of ρ , because $\varrho^{(S)} \leq \rho^{w(S)} \leq \rho'^{w(S)}$. Consequently, suppose the disk step were only able to give a weaker moment bound of the form

$$\mathbb{E}[R'^{H_w^{(t)}}] \leq K^w \rho'^{-w} \quad \text{with some fixed } K > 1$$

(e.g. a loss of a factor $(1 + o(1))^w$, or K^w for a constant K). Running the argument with the ledger at exponent $\rho'' := \rho'/K$ instead of ρ' restores the conclusion $|\tilde{Q}_t| \leq m^{t/2}$, provided only $\rho'' \geq \rho$, i.e. $K \leq \rho'/\rho$. Taking, say, $\rho' = \sqrt{\rho}$ (legitimate since $\rho \geq \frac{1}{2}$ by the convention, so $\sqrt{\rho} < 1$) leaves room for any $K \leq \rho^{-1/2}$, a fixed constant > 1 , at the sole cost of replacing $\ln(1/\rho)$ by $\frac{1}{2} \ln(1/\rho)$ in the definition of R — which changes θ_0 in Theorem 9.1 by a bounded factor and nothing else. Since the Main Theorem asserts only the existence of some $c(G) > 0$, such a change is immaterial.

(b) *The final calibration is strict.* In the proof of Theorem 9.1 we use $\frac{4\sqrt{q}}{\ln R}(\delta_B + \ln b) \leq \frac{\delta_B}{4}$, i.e. the two-constants correction is forced to be at most one quarter of the main term, leaving the factor $e^{-\frac{3}{4}\delta_B M}$ with a gap of $\frac{1}{4}\delta_B M$ to spare. Any multiplicative loss of the form $e^{O(1) \cdot M}$ with a small enough constant, or any replacement of $\Xi = b^M$ by $\Xi = (Cb)^M$ with fixed C , is absorbed by enlarging Λ (equivalently, shrinking θ_0), again changing only the value of $c(G) > 0$.

(c) *No cancellation is relied upon.* The ledger (6.6) bounds the sum of absolute values of the coefficients, and the disk bound bounds $|\mathcal{G}|$ term by term; the argument never uses sign information or near-cancellation, so it is stable under perturbation of any of the coefficients.

8 The two-constants lemma

Lemma 8.1. *Let $0 < q \leq \frac{1}{2}$, $L := 1 - q$, $R > 1$, $0 < \varepsilon \leq \Xi$, and let f be a polynomial (of any degree) with $|f| \leq \varepsilon$ on $E := [0, L]$ and $|f| \leq \Xi$ on $\overline{D}(0, R)$. Then*

$$\ln |f(1)| \leq \ln \varepsilon + \frac{4\sqrt{q}}{\ln R} \ln \frac{\Xi}{\varepsilon}.$$

Proof. Assume $f \not\equiv 0$; then $\ln |f|$ is subharmonic on $\mathbb{C}$. Let $\phi(\zeta) = \zeta + \sqrt{\zeta^2 - 1}$, analytic on $\mathbb{C} \setminus [-1, 1]$ with $\phi(\zeta) \sim 2\zeta$ at infinity, mapping $\mathbb{C} \setminus [-1, 1]$ conformally onto $\{|w| > 1\}$, with $|\phi| \rightarrow 1$ at $[-1, 1]$. Set $g(z) := \ln |\phi(\frac{2z-L}{L})|$: harmonic on $\mathbb{C} \setminus E$, continuous up to E with boundary value 0, and positive off E .

(a) At $z = 1$: $\zeta_0 = \frac{2-L}{L} = \frac{1+q}{1-q}$, and

$$g(1) = \operatorname{arccosh} \frac{1+q}{1-q} = \int_0^q \frac{ds}{\sqrt{s}(1-s)} \leq \frac{1}{1-q} \int_0^q \frac{ds}{\sqrt{s}} = \frac{2\sqrt{q}}{1-q} \leq 4\sqrt{q}$$

for $q \leq \frac{1}{2}$. (The derivative identity: $\frac{d}{dq} \operatorname{arccosh} \frac{1+q}{1-q} = \frac{2/(1-q)^2}{\sqrt{(\frac{1+q}{1-q})^2 - 1}} = \frac{2/(1-q)^2}{2\sqrt{q}/(1-q)} = \frac{1}{\sqrt{q}(1-q)}$.)

(b) On $|z| = R$: from $\phi + \phi^{-1} = 2\zeta$ and $|\phi| \geq 1$ we get $|\phi(\zeta)| \geq 2|\zeta| - 1$; here $|\zeta| = \frac{|2z-L|}{L} \geq \frac{2R-L}{L}$, so $|\phi| \geq \frac{4R-3L}{L} \geq 4R-3 \geq R$ (using $L \leq 1 \leq R$). Hence $m_R := \min_{|z|=R} g \geq \ln R$.

(c) On $\Omega := D(0, R) \setminus E$, the function $u := \ln |f| - [\ln \varepsilon + \frac{g}{m_R}(\ln \Xi - \ln \varepsilon)]$ is subharmonic (g is harmonic on Ω). On $|z| = R$: $g \geq m_R$, so the bracket is $\geq \ln \Xi \geq \ln |f|$ and $u \leq 0$. At each point of E : $g \rightarrow 0$ and $\ln |f| \leq \ln \varepsilon$, so $\limsup u \leq 0$. By the maximum principle, $u \leq 0$ on Ω ; evaluate at $z = 1 \in \Omega$ and use (a), (b). $\square$

9 Proof of the Main Theorem

Theorem 9.1 (connected case). *Let G be connected with $\operatorname{val}^*(G) < 1$; let $\rho < 1$ (Lemma 3.1), m , $b = 1 + \frac{\sqrt{m-1}}{16}$, and $\delta_B \in (0, 1]$, M_1 as in §4. Define*

$$\Lambda := \frac{16\sqrt{q}(\delta_B + \ln b)}{\delta_B} \quad (q = \tfrac{2}{5}), \quad \theta_0 := \frac{\ln(1/\rho)}{\ln(1/\rho) + e^\Lambda}, \quad N_0 := \left\lceil \max\left(\frac{2M_1}{\theta_0}, \frac{4}{\theta_0}\right) \right\rceil.$$

Then for all $N \geq N_0$,

$$v_N \leq \exp\left(-\frac{3}{8}\delta_B\theta_0 N\right).$$

Proof. Fix $N \geq N_0$ and a strategy S with $\text{val}(S) = e^{-u} > 0$. Let $M := \lfloor \theta_0(N+1) \rfloor$; then $M \geq \theta_0 N_0 - 1 \geq 2M_1 - 1 \geq M_1$ and $M \leq N$. Moreover

$$\frac{N - M + 1}{M} \geq \frac{N + 1}{M} - 1 \geq \frac{1}{\theta_0} - 1 = \frac{e^\Lambda}{\ln(1/\rho)} \implies R \geq 1 + e^\Lambda, \quad \ln R \geq \Lambda.$$

Apply Lemma 8.1 to $f = \mathcal{G}$ with $\varepsilon = e^{-\delta_B M}$ (Lemma 7.1(iv); legitimate since $M \geq M_1$) and $\Xi = b^M$ (Lemma 7.3; note $\varepsilon \leq 1 \leq \Xi$). Using Lemma 7.1(iii),

$$-u \leq \ln \mathcal{G}(1) \leq -\delta_B M + \frac{4\sqrt{q}}{\ln R} M(\delta_B + \ln b) \leq -\delta_B M + \frac{4\sqrt{q}(\delta_B + \ln b)}{\Lambda} M = -\delta_B M + \frac{\delta_B}{4} M,$$

so $u \geq \frac{3}{4}\delta_B M \geq \frac{3}{4}\delta_B(\theta_0(N+1) - 1) \geq \frac{3}{8}\delta_B\theta_0 N$, the last step because $\theta_0 N \geq \theta_0 N_0 \geq 4$ gives $\theta_0(N+1) - 1 \geq \theta_0 N - 1 \geq \frac{1}{2}\theta_0 N$. Take the supremum over S . $\square$

Corollary 9.2 (general case = Main Theorem). *Every game G with $\text{val}^*(G) < 1$ satisfies $v_n \leq C(G)e^{-c(G)n}$ for all $n \geq 1$: for connected G take $c(G) = \frac{3}{8}\delta_B\theta_0$ and $C(G) = e^{c(G)N_0}$ (so the bound is trivial for $n < N_0$); for general G combine Theorem 9.1 with Proposition 2.3 (whose hypothesis is supplied by the “in particular” clause of Lemma 2.1). ■*

10 Remarks

1. **Where each hypothesis is consumed.** The hypothesis $\text{val}^*(G) < 1$ enters only through $\delta_B > 0$ (Theorem 4.2 applied to $G_{\perp 3/4}$, via Lemma 4.1) — the unique step that consumes finite-dimensionality of the strategies, as it must (for commuting-operator strategies the anchored repeated value can equal 1, and the argument correctly yields nothing). Connectedness enters only through $\rho < 1$, which makes $R > 1$; for disconnected supports the profile bound genuinely fails (e.g. perfectly correlated questions with parity answers give a profile collapsing at every positive severance rate while $\mathcal{G}(1) = 1$), and Proposition 2.3 handles them. The “board length” N/M enters only through the radius R .
2. **Constants.** All constants are explicit given BVY’s: $c(G) = \frac{3}{8}\delta_B\theta_0$ with $\theta_0 = \ln(1/\rho)/(\ln(1/\rho) + e^\Lambda)$ and $\Lambda = 16\sqrt{2/5}(1 + \ln b/\delta_B) \leq 11 + \frac{11\ln b}{\delta_B}$. Since $\delta_B \leq \ln \frac{1}{\text{val}^*(G_{\perp \beta})} \leq \ln \frac{1}{1-\varepsilon_0/16}$, we have $c(G) < \ln \frac{1}{1-\varepsilon_0}$, consistent with the trivial lower bound $v_N \geq (1-\varepsilon_0)^N$. The rate is tiny (θ_0 is exponentially small in $\ln b/\delta_B$), but the theorem asks only for some $c(G) > 0$; see Remark 7.4 for why constant-factor losses anywhere in the pipeline only degrade $c(G)$.
3. **Monotone consequences.** Since $\mathcal{R}_0[S]$ (no flags) is included, the argument in particular bounds every strategy for $G^{\otimes N}$ against the anchored repetition bound through the profile at the single point $B = 1 - q$; the flags (the φ -sweep) are what turn the single inequality into control on a whole segment, which is what the analytic step needs (cf. Remark 7.2).
4. **Summary of imported versus self-contained content.** Self-contained: §1–§3, §5–§8 (the routing construction and its conditional law, the spectral representation and ledger, the occupancy profile, the hypergeometric moment bound, the two-constants lemma), together with the standard facts listed in the Overview. Imported: Theorem 4.2 only, in the precise form stated in §4 (hypotheses (B1)–(B4); conclusion exponential in M with constants independent of M and of the strategy), consumed exactly once, in Lemma 7.1(iv).