The critical alphabet exponent for exact classical simulation of quantum measurements

Kashogi

AI-assisted manuscript. The proof has not been independently verified or peer reviewed.

Kashogi.com · Research announcement · Complete source and raw results

Abstract. Alice receives a quantum state and Bob chooses how to measure it. Can a single classical message and shared randomness reproduce the outcome probabilities exactly? At fixed dimension, a finite average bit cost is possible. The question is how often the simulator must make an exceptionally large menu of messages available. Write L for the menu size after the shared randomness has been fixed. For every complex dimension $d \geq 4$, we propose a proof that the optimal exponent is

$$\gamma_d = \frac{1}{d-2}.$$

The proposed construction attains $\Pr(L > m) = O_d(m^{-\gamma_d})$ by mixing cap scales and grouping their labels. For $d \geq 5$, a strong L^{d-2} estimate implies that every exact simulator has $\mathbb{E}L^{\gamma_d} = \infty$. For $d = 4$, the obstruction already follows from the ququart analysis of de Gois, Santos and Vieira [4, Apps. B.5–B.6] once the sum retains the masses of the input cells. Both obstructions use only rank-one yes/no tests. A reduction to arbitrary measurements increases the declared alphabet by at most a fixed dimension-dependent factor.

What changes in dimension four. Castellano, Sekatski and Brunner [2, Eq. (12), Sec. V and App. A] obtained $1/6 \leq \beta_4 \leq 1$. Their cellwise argument sums to $8D$ over D labels. Keeping the Haar mass $\mu(A_i)$ of each cell gives $8\sum_i \sqrt{\mu(A_i)} \leq 8\sqrt{D}$ (Lemma 13). This yields the obstruction at $1/2$. The matching construction is a separate claim. It gives one protocol with $\mathbb{E}L^{0.49} < \infty$ and $\Pr(L > 2^k) \leq C_4\, 2^{-k/2}$, while every exact protocol has $\mathbb{E}\sqrt{L} = \infty$.

Why the exponent has this value. At zero overlap, the Born probability vanishes. On almost every tape with finitely many positive-measure input cells, the mean acceptance rises more slowly than the required linear law as overlap tends to zero. Section 4 shows that a finite endpoint moment would force this same behavior after averaging over tapes, contradicting the Born law. In the matching construction, the probability of a scale below s is asymptotically proportional to s. At fine scales, grouping candidate messages makes the alphabet grow like $s^{-(d-2)}$ up to a random round factor with finite moments of every order. Sections 5–7 make this calculation exact.

Reading paths. Theorems 3 and 4 specify the quantifiers. Sections 5–7 give the matching protocol. Section 8 treats general measurements. Section 11 and Appendix A describe the tests and their reproducibility.

1 Model, definitions and main results

For the binary experiment defined below, the critical alphabet moment is

$$\beta^{\mathrm{mom}}_{d,\mathrm{bin}} = \sup\{q > 0 : \text{there is an exact protocol } \Pi \text{ with } \mathbb{E}_\lambda D_\Pi(\lambda)^q < \infty\}. \tag{1}$$

The quantifier is existential in the protocol at each order q. The expectation is over its input-independent randomness. All input possibilities are included when forming $D_\Pi(\lambda)$, before the expectation.

Definition 1 (Input pairs and target)**.** *Let $d \geq 2$ be an integer and let*

$$\mathcal{I}_d = \{(\psi, \varphi) \in \mathbb{C}^d \times \mathbb{C}^d : \|\psi\| = \|\varphi\| = 1\}.$$

Alice receives the pure preparation $P_\psi = \psi\psi^*$ and Bob receives the projective test $\{P_\varphi, I - P_\varphi\}$. The target is a single bit B with $\Pr(B = 1) = |\langle\varphi, \psi\rangle|^2$. The input rays $x = [\psi]$, $y = [\varphi]$ range over $X = \mathbb{CP}^{d-1}$. The normalized invariant measure on X is denoted by μ. Exactness is required for every fixed pair $(x, y) \in X^2$.

The accepting effect is a rank-one projector. The rejecting effect has rank $d - 1$. A scalar multiple ηP_φ with $0 \leq \eta \leq 1$ can also be simulated by independent thinning, but the theorem is stated for the projective test. The parameter t below denotes an exact input overlap, $t(x, y) = |\langle\varphi, \psi\rangle|^2$.

Definition 2 (One-way protocol and alphabet)**.** *A protocol consists of a standard Borel probability space $(\Lambda, \mathcal{L}, \rho)$ of public tapes, a Borel encoder $e : \Lambda \times X \to \mathbb{N}$, and Borel decoder functions $b_i : \Lambda \times X \to [0, 1]$. The tape is independent of both inputs. Alice sends the single label $e(\lambda, x)$. Bob accepts with probability $b_{e(\lambda,x)}(\lambda, y)$. Exactness means*

$$\int_\Lambda b_{e(\lambda,x)}(\lambda, y)\, \rho(\,\mathrm{d}\lambda) = t(x, y) \quad \textit{for every } x, y \in X. \tag{2}$$

Local random choices of the encoder can be placed on the public tape. Define the input cells and the two costs by

$$\begin{aligned} A_{\lambda i} &= \{x : e(\lambda, x) = i\}, \\ D(\lambda) = \#\{i : A_{\lambda i} \neq \varnothing\}, &\qquad D_+(\lambda) = \#\{i : \mu(A_{\lambda i}) > 0\}. \end{aligned} \tag{3}$$

Both counts can be infinite. We complete the tape measure when interpreting D. In D_+ the measure μ is applied to preparation cells in X.

The event $\{A_{\lambda i} \neq \varnothing\}$ is the projection of a Borel set and is analytic, hence measurable in the completion of every Borel probability on Λ. The integral $\mu(A_{\lambda i})$ is measurable by Tonelli. Summing the associated indicators defines D and D_+, and $D_+ \leq D$. A countably infinite alphabet covers all finite binary strings as well as finite protocols. Local computation and random-tape size are unrestricted. The resource in (1) is the size of the used alphabet at a fixed tape, allowing the encoder input to range over all preparations.

Theorem 3 (Critical alphabet moment)**.** *For the model in Definitions 1–2, and each fixed $d \geq 4$, put $\gamma_d = 1/(d - 2)$. There is an exact protocol and finite positive constants c_d, C_d, m_d, depending only on d and its fixed construction choices, such that*

$$c_d m^{-\gamma_d} \leq \Pr(D > m) \leq C_d m^{-\gamma_d} \quad (m \geq m_d).$$

Every exact protocol satisfies $\mathbb{E}D_+^{\gamma_d} = \infty$, and therefore $\mathbb{E}D^{\gamma_d} = \infty$. In particular $\beta^{\mathrm{mom}}_{d,\mathrm{bin}} = \gamma_d$.

General measurements. Alice now receives a density matrix ρ. Bob receives a positive operator-valued measurement (POVM) E on a standard Borel outcome space. Positive effects satisfy $E(\Omega) = I$ and are countably additive. Bob must return one outcome with law $\mathrm{tr}(\rho E(\cdot))$ for every input pair. For a discrete POVM this is the probability vector $(\mathrm{tr}(\rho E_i))_i$. The tape is independent of both inputs. The encoder is independent of Bob's measurement. Write $L(\lambda)$ for a finite declared alphabet covering all inputs on tape λ. In particular $D_+ \le D \le L$ on the binary subproblem. For either measurement class let β^{tail} be the supremum of $\beta > 0$ for which some exact protocol has $\Pr(L > m) = O_d(m^{-\beta})$. Equivalently, $\Pr(\log_2 L > k) = O_d(2^{-\beta k})$. This is the tail problem posed in [2, Sec. V].

Theorem 4 (Optimal exponent for all measurements)**.** *For every fixed $d \ge 4$,*

$$\beta^{\mathrm{tail}}_{d,\mathrm{all}} = \beta^{\mathrm{tail}}_{d,\mathrm{bin}} = \frac{1}{d-2}.$$

In each class the upper-tail bound is achieved at this exponent, and the endpoint moment is infinite for every exact protocol. The corresponding optimal moment suprema have the same value.

Lemma 5 (Moments and tails)**.** *For every nonnegative random variable V and $q > 0$, in the extended nonnegative reals,*

$$\mathbb{E}V^q = q \int_0^\infty u^{q-1} \Pr(V > u)\, \mathrm{d}u.$$

A two-sided tail of order $u^{-\gamma}$ with $\gamma > 0$ gives finite qth moment exactly for $0 < q < \gamma$.

Proof. For each $v \ge 0$ write $v^q = \int_0^v q u^{q-1}\, \mathrm{d}u$ and interchange the nonnegative integrands by Tonelli. Above any fixed positive threshold the comparison integral is $\int u^{q-\gamma-1}\, \mathrm{d}u$, which converges precisely for $q < \gamma$. The integral near zero is bounded by the integral of qu^{q-1}. □

Theorem 3 thus gives both directions of (1). Endpoint divergence for every protocol alone would not specify a pointwise tail for every protocol. The two-sided tail is asserted for the constructed simulator.

What a fixed bit budget buys. On a given tape, $\lceil \log_2 L \rceil$ bits suffice to index the declared menu. A k-bit version of the constructed protocol can agree with it on every tape with $L \le 2^k$. On the remaining tapes, identified by the shared randomness before the inputs arrive, Bob substitutes a fixed output law for his measurement. Coupling the two procedures gives, for every input pair, total variation error at most

$$\Pr(L > 2^k) \le C_d\, 2^{-k/(d-2)}.$$

This is a fixed-dimension guarantee for truncating this exact simulator. A converse for separately optimized approximate protocols would require a different argument.

Why the endpoint is still infinite. For a variable with a two-sided tail of order $m^{-\gamma_d}$, the endpoint integrand in Lemma 5 is comparable to $1/m$. Its integral diverges logarithmically. More generally, splitting that integral at the dyadic points gives a necessary condition for *every* exact simulator:

$$\sum_{k=0}^{\infty} 2^{k\gamma_d} \Pr(L > 2^k) = \infty.$$

Indeed, monotonicity bounds the integral on $[2^k, 2^{k+1}]$ by $(2^{\gamma_d} - 1)2^{k\gamma_d} \Pr(L > 2^k)$. In particular, no exact protocol has a uniform tail bound $O_d(m^{-\gamma_d}/(\log m)^{1+\varepsilon})$ for any $\varepsilon > 0$.

1.1 Notation and proof structure

Symbol	Definition
d, a, γ_d	Complex input dimension, $a = d - 2$, critical order $1/a$
X, μ	Ray space $\mathbb{CP}^{d-1}$ and invariant probability on preparations
t, η_t	Fidelity and invariant law of pairs with that fidelity
R_t, T_t	Fidelity average and $(R_t - R_0)/t$
$\mathcal{H}_n, f_n$	Projective harmonic level and orthogonal component of f
$m_t(n), q_t(n)$	Multipliers of R_t and T_t
S_N, h, K_d	Smooth spectral filter, fixed scalar filter, uniform filter norm bound
$A_{\lambda i}, D, D_+$	Encoder cell, number of used labels, number of positive-mass cells
$a_{A,b}, F_\lambda$	Single-cell averaged response and total response at fixed fidelity
U_d, V_d, Q_d	Mixing variable, independent dimension factor, Mellin transform
ν_d, C_s, K_s	Scale law, encoder cap, conditional cap response
δ, c, σ	Fine cutoff, attenuation 3/4, removed inner mass $4^{-(d-1)}$
$L, n, J, \mathcal{A}_{s,\lambda}$	Number of centres, packet capacity, common round bound, packet alphabet

First we identify R_t geometrically and spectrally. A polynomial derivative estimate and a bounded spectral filter yield the standalone endpoint theorem. Orthogonality, Sobolev locality and Hölder give the communication obstruction. We then construct the positive scale mixture, specify the encoding and decoding rules, and prove their exactness and alphabet tail. Section 9 determines the sharp output exponent, and Section 8 transfers the communication result to general measurements.

2 The fidelity averaging family

For $t \in [0, 1]$ let η_t be the pushforward of Haar probability on $U(d)$ by

$$G \longmapsto ([Ge_1], [G(\sqrt{t}e_1 + \sqrt{1-t}e_2)]). \tag{4}$$

Its two marginals are μ and it is supported on $t(x, y) = t$. Equivalently, for a unit representative x, draw a uniform unit vector z in $x^\perp$ and set

$$(R_t f)([x]) = \mathbb{E}_z f([\sqrt{t}x + \sqrt{1-t}z]). \tag{5}$$

The distribution is independent of the representative of $[x]$, using the phase invariance of z. These formulas apply to bounded Borel representatives. For L^p classes they define an almost-everywhere operator independent of representatives, by the Haar second marginal. Jensen's inequality and that marginal give $\|R_t f\|_p \le \|f\|_p$ for $1 \le p \le \infty$. They also give

$$\langle g, R_t f\rangle = \int \overline{g(x)} f(y)\, \mathrm{d}\eta_t(x, y). \tag{6}$$

In particular $R_1 = I$, whereas R_0 averages over the projective hyperplane $x^\perp$. For example, on the first nonconstant harmonic level the multiplier is $(dt - 1)/(d - 1)$, which at zero equals $-1/(d - 1)$. No spectral renormalization is applied in the definition

$$T_t = (R_t - R_0)/t, \qquad 0 < t \le 1/2. \tag{7}$$

The interval is fixed independently of harmonic degree.

Theorem 6 (Endpoint estimate)**.** *Let $d \geq 5$, $a = d-2$, and $f \in L^\infty(X,\mu)$. Write M for the essential supremum of $|f|$, so $M = \|f\|_\infty$. With the finite constant $\mathcal{C}_d$ defined in (13),*

$$\sup_{0<t\leq 1/2} \|T_t f\|_a^a \leq \mathcal{C}_d M^{a-2} \|f\|_2^2.$$

For every such f, $T_t f \to (\Delta/4) R_0 f$ strongly in L^a as $t \downarrow 0$. The Laplacian is that of the unit sphere on phase-invariant lifts to S^{2d-1}, so that its projective level-n eigenvalue is $-4n(n+d-1)$. In dimension four, de Gois, Santos and Vieira [4, App. B.5, especially B.5.3, Eq. (167)] give the uniform bound $\|T_t\|_{2\to 2} \leq 8$ and the strong L^2 limit above. Their zero-set argument gives the vanishing cell slope used in Lemma 13. Appendix B.6 of [4] identifies the higher-dimensional spectral decay. The estimate proved here for $d \geq 5$ controls the strong L^{d-2} output norm.

3 Proof of the endpoint estimate

3.1 Jacobi multipliers and their normalization

Let $\mathcal{H}_n$ be the restrictions to S^{2d-1} of harmonic polynomials homogeneous of degree n in both z and $\bar{z}$. They are invariant under $z \mapsto e^{i\theta} z$ and therefore functions on X. Denote the orthogonal projection by $f \mapsto f_n$.

Lemma 7 (Zonal decomposition)**.** *The spaces $\mathcal{H}_n$ give an orthogonal decomposition of $L^2(X,\mu)$. For $a = d-2$, let $B_{n,a} = (n+a)!/(n!\,a!)$, the number of ways to choose a elements from $n+a$. Then*

$$m_t(n) = \frac{P_n^{(a,0)}(2t-1)}{B_{n,a}}, \qquad R_t f_n = m_t(n) f_n, \tag{8}$$

$$m_0(n) = \frac{(-1)^n}{B_{n,a}}, \qquad q_t(n) = \frac{P_n^{(a,0)}(2t-1) - (-1)^n}{t B_{n,a}}, \tag{9}$$

$$T_t f_n = q_t(n) f_n, \qquad m_0'(n) = -n(n+d-1) m_0(n). \tag{10}$$

Here $m_0'(n)$ denotes differentiation in t at zero. In particular $q_t(0) = 0$. The standard Jacobi endpoint value $P_n^{(a,0)}(1) = B_{n,a}$ gives $m_1(n) = 1$ for every $n \geq 0$ [5, Table 18.6.1].

Proof. Polynomials in the entries of zz^* form a self-adjoint algebra containing constants and separating rays. Stone–Weierstrass gives density on the compact ray space. Let V_n be restrictions of polynomials of bidegree (n,n). These spaces are nested, by multiplication by $|z|^2$. In the formal monomial inner product with squared norms $\alpha!\beta!$, multiplication by $\sum z_i \bar{z}_i$ is adjoint to $\sum \partial_{z_i}\partial_{\bar{z}_i}$. Orthogonal splitting gives a harmonic component plus a multiple of $|z|^2$. Iterating gives the decomposition. The Euclidean radial formula for a homogeneous harmonic of real degree k gives spherical eigenvalue $-k(k+2d-2)$. Integration by parts makes the different degrees sphere-orthogonal. This identifies $V_n \ominus V_{n-1}$ with $\mathcal{H}_n$ and proves completeness.

Averaging over the stabilizer of $x = e_1$ turns a bidegree-(n,n) polynomial into a polynomial in $t = |z_1|^2$ of degree at most n. Phase averaging balances the powers of z_1 and $\bar{z}_1$. Unitary averaging on the remaining coordinates gives powers of $1-t$, by transitivity on their sphere. All powers up to degree n occur. In $\mathcal{H}_n$ the invariant subspace is therefore the one-dimensional space of polynomials orthogonal to all lower degrees. Start with independent standard complex Gaussian coordinates. Their squared moduli are independent exponentials. Dividing by their sum gives proportions uniform on the $(d-1)$-simplex. Thus fidelity has density $(d-1)(1-t)^a$. Orthogonality with this weight identifies the invariant polynomial as $P_n^{(a,0)}(2t-1)$, with value $\binom{n+a}{a}$ at one. Stabilizer averaging of f_n has value $f_n(x)$ at the pole, giving (8), including its endpoint values. The contraction property extends the identity to L^2.

The shifted Jacobi equation is

$$t(1-t)m_t''(n) + (1-dt)m_t'(n) + n(n+d-1)m_t(n) = 0.$$

At zero it gives (10). Subtracting $m_0(n)$ and dividing by t gives (9). These formulas are exact for each degree and every t. □

A degree-two check without a Jacobi routine. At $d = 5$, the Haar fidelity has density $4(1-t)^3$ and its first three moments are $1/5, 1/15, 1/35$. The quadratic $Q(t) = 21t^2 - 12t + 1$ is orthogonal to both 1 and t under this law, as substitution of those three moments shows. It has $Q(1) = 10$. The unique degree-two zonal polynomial with value one at the pole is therefore

$$m_t(2) = \frac{21t^2 - 12t + 1}{10}.$$

It gives $m_0(2) = 1/10$, $m_0'(2) = -6/5$ and $m_1(2) = 1$, in agreement with (8)–(10). Dividing Q by $n + a = 5$ would give $m_1(2) = 2$ and violate $R_1 = I$.

For completeness, the reproducing kernel and its dimension are

$$Z_n(x,y) = \sum_{j=1}^{h_n} e_{nj}(x)\overline{e_{nj}(y)} = h_n \frac{P_n^{(a,0)}(2t(x,y)-1)}{\binom{n+a}{a}},$$
$$h_n = \binom{n+d-1}{d-1}^2 - \binom{n+d-2}{d-1}^2 = \frac{2n+a+1}{a+1}\binom{n+a}{a}^2.$$

Here (e_{nj}) is any orthonormal basis of $\mathcal{H}_n$, and a binomial coefficient with upper argument smaller than its lower argument is zero. The dimension formula follows by subtracting the dimensions in the harmonic splitting above. Unitary invariance and the invariant one-dimensional subspace give the zonal form. The diagonal $Z_n(x,x)$ is constant and integrates to h_n, fixing the prefactor. Thus $\int_X Z_n(x,y) f_n(y)\,\mathrm{d}\mu(y) = f_n(x)$. In particular $h_1 = d^2 - 1$.

3.2 Uniform multiplier control

Lemma 8. *If $a = d - 2 \geq 2$, $0 < t \leq 1/2$, and $n \geq 1$, then*

$$|q_t(n)| = |(m_t(n) - m_0(n))/t| \leq H_d n^{-(a-2)}, \qquad H_d = 4 \cdot 2^a a!.$$

Proof. On $-1 \leq x \leq 0$, the bound $|P_n^{(a,0)}(x)| \leq 2^a$ follows by induction. At $a = 0$, the Laplace integral $P_n(x) = \pi^{-1}\int_0^\pi (x + i\sqrt{1-x^2}\cos v)^n\,\mathrm{d}v$ gives the bound one. For the induction use

$$(1-x)P_n^{(k+1,0)}(x) = \frac{2(n+k+1)P_n^{(k,0)}(x) - 2(n+1)P_{n+1}^{(k,0)}(x)}{2n+k+2}.$$

Orthogonality leaves precisely these two possible degrees. Comparing leading coefficients gives $-2(n+1)/(2n+k+2)$ for the coefficient of $P_{n+1}^{(k,0)}$. Evaluating at $x = 1$ then gives $2(n+k+1)/(2n+k+2)$ for the coefficient of $P_n^{(k,0)}$. The absolute coefficients sum to two and $1 - x \geq 1$. Markov's polynomial inequality on $[0, 1/2]$ multiplies the supremum of a degree-n polynomial by at most $4n^2$ on taking one derivative. Divide by $\binom{n+a}{a} \geq n^a/a!$ and apply the mean value theorem. The constant level has zero difference quotient. □

The exact endpoint slope has magnitude

$$|m_0'(n)| = \frac{n(n+d-1)}{\binom{n+a}{a}} \sim a! n^{2-a}.$$

For $d = 4$ this equals $2 - 4/[(n+1)(n+2)]$. Consequently the limit operator $(\Delta/4)R_0$ has L^2 operator norm exactly two in that dimension. The elementary all-dimension estimate in Lemma 8 gives $H_4 = 32$ for the whole difference-quotient family. The ququart-specific estimate of [4, Eq. (167)] gives the smaller uniform bound 8.

There is also an explicit check on the simultaneous limit $t = n^{-2}$. Reflection and the terminating Jacobi series give

$$P_n^{(a,0)}(-1 + 2/n^2) = (-1)^n \sum_{k=0}^{n} \frac{(-n)_k (n+a+1)_k}{n^{2k}(k!)^2}.$$

The series follows by substituting a power series with constant term one into the reflected Jacobi equation. The successive coefficients have ratio $(-n+k)(n+a+1+k)/(k+1)^2$. For $k \le n$, the absolute value of the displayed summand is at most $(a+2)^k/(k!)^2$, and for each fixed k its limit is $(-1)^k/(k!)^2$. Dominated convergence for this summable majorant therefore gives

$$(-1)^{n+1} n^{a-2} q_{1/n^2}(n) \longrightarrow a!(1 - J_0(2)), \qquad J_0(2) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(k!)^2}.$$

The constant $1 - J_0(2)$ is strictly positive, as follows from the alternating series after its first two terms cancel. Thus the high-degree scaling on this sequence is again n^{2-a}. The fixed-degree derivative approximation is not uniform on the scale $n^2 t = 1$.

3.3 The spectral filter and the polynomial bound

We record the complete approximation assertion used from Wang–Sloan [1, Lemma 2.1], with sphere dimension denoted by m to distinguish it from the complex input dimension. Write $\Pi_L(\mathbb{S}^m)$ for spherical polynomials of ordinary degree at most L, $\mathcal{Y}_k$ for their harmonic projections, and $E_L(f)_p = \inf_{P \in \Pi_L} \|f - P\|_p$.

Lemma 9 (Wang–Sloan filter statement)**.** *Let $m \ge 2$, $r = \lfloor (m+1)/2 \rfloor$, and let a filter h equal one on $[0,1]$ and zero on $[2,\infty)$. Assume $h \in W^r BV$, meaning that $h^{(r-1)}$ is absolutely continuous and the one-sided rth derivatives exist and have bounded variation, with the lower derivatives absolutely continuous. For $V_L f = \sum_{k \ge 0} h(k/L)\mathcal{Y}_k f$, with integer $L \ge 1$, the following hold. The output belongs to Π_{2L-1}, and $V_L P = P$ for $P \in \Pi_L$. Its uniform operator norm is bounded independently of L. For $1 \le p \le \infty$, $\|V_L f\|_p \le C_{m,h,p}\|f\|_p$ and $\|f - V_L f\|_p \le C'_{m,h,p} E_L(f)_p$. For every $s > 0$ and $f \in W_p^s(\mathbb{S}^m)$ the error is at most $C''_{m,h,p,s} L^{-s} \|f\|_{W_p^s}$. All displayed constants are independent of L and f.*

The proof of this cited lemma uses summation by parts and uniformly bounded Cesàro means. Only its polynomial reproduction and uniform norm bound are needed here. We fix the filter explicitly. Put $\chi(v) = e^{-1/v}$ for $v > 0$ and $\chi(v) = 0$ for $v \le 0$, and set

$$h(v) = \frac{\chi(2-v)}{\chi(2-v) + \chi(v-1)} \quad (v \ge 0).$$

The denominator is strictly positive everywhere, $0 \le h \le 1$, and this fixed smooth filter has the required plateau and support. Every derivative is smooth with compact transition support, so all the hypotheses in Lemma 9 hold for $m = 2d-1$ and $r = d$. Define the finite constant

$$K_d = \max\left\{1, \sup_{L \in \mathbb{N},\, L \ge 1} \|V_L\|_{L^\infty(\mathbb{S}^{2d-1}) \to L^\infty(\mathbb{S}^{2d-1})}\right\}, \qquad L_d = 8K_d. \tag{11}$$

These constants are determined by d and the displayed filter. Theorem 6 needs their finiteness, with no numerical estimate assumed.

Lemma 10 (Low-frequency difference quotient)**.** *For the fixed filter above, let S_N be V_{2N} restricted to phase-invariant functions on $\mathbb{S}^{2d-1}$, where $N \geq 1$ is an integer. Then $(S_N f)_n = h(n/N) f_n$, S_N reproduces projective levels through N, and*

$$\|T_t S_N f\|_\infty \leq L_d N^2 \|f\|_\infty \quad (0 < t \leq 1/2).$$

Proof. The full-sphere operator commutes with rotations and in particular with global complex phase. A projective level n has ordinary spherical degree $2n$, so the factor is exactly $h(2n/(2N)) = h(n/N)$. The surviving projective degrees are $n < 2N$. For a fixed pole x,

$$R_u S_N f(x) = \sum_{n=0}^{2N-1} h(n/N) \frac{P_n^{(a,0)}(2u-1)}{\binom{n+a}{a}} f_n(x). \tag{12}$$

This finite expression is a polynomial in u of degree at most $2N-1$, bounded on $[0,1]$ by $\|S_N f\|_\infty \leq K_d \|f\|_\infty$ using (5). Markov's inequality on $[0,1]$ bounds its derivative by $2(2N)^2 K_d \|f\|_\infty$. Apply the mean value theorem between zero and t. Complex values are handled by a rotated real part at the point where the derivative is being estimated. The finite sum is a spectral cutoff of f, and every retained multiplier is the exact Jacobi polynomial. □

3.4 Summation at the endpoint

For later use define

$$\mathcal{C}_d = H_d^2 (2L_d)^{a-2} \left[\frac{a}{a-2} + \frac{4(4^a - 1)}{1 - 2^{-2(a-2)}} \right] \quad (d \geq 5), \qquad B_d^{\mathrm{op}} = \mathcal{C}_d^{1/a}. \tag{13}$$

In dimension four use $B_4^{\mathrm{op}} = 8$ from [4, App. B.5.3, Eq. (167)].

Proof of Theorem 6. First suppose $d \geq 5$. The zero function satisfies the assertion directly. Normalize $\|f\|_\infty = 1$ and set $g = T_t f$. Lemma 8 gives $\|g\|_2 \leq H_d \|f\|_2$. For a level $u_j = 2L_d 4^j$ put $N = 2^j$ $(j = 0, 1, \ldots)$. Lemma 10 bounds the low-frequency part by half the level. Chebyshev's inequality, orthogonality and the exact low-frequency plateau give

$$\mu(|g| > u_j) \leq \frac{4H_d^2}{u_j^2} \sum_{n > 2^j} n^{-2(a-2)} \|f_n\|_2^2. \tag{14}$$

For the contribution of $u \in (0, u_0)$ to the layer-cake integral $a \int_0^\infty u^{a-1} \mu(|g| > u)\, \mathrm{d}u$, the bound $\mu(|g| > u) \leq \|g\|_2^2 / u^2$ gives at most $a u_0^{a-2} H_d^2 \|f\|_2^2 / (a-2)$. The contribution from $[u_j, 4u_j]$ is at most $(4^a - 1) u_j^a \mu(|g| > u_j)$. Exchange the order of the nonnegative sums in (14). Each fixed harmonic degree contributes its energy times a constant bounded by

$$4H_d^2 (4^a - 1)(2L_d)^{a-2} n^{-2(a-2)} \sum_{2^j < n} 2^{2j(a-2)} \leq \frac{4H_d^2 (4^a - 1)(2L_d)^{a-2}}{1 - 2^{-2(a-2)}}.$$

The sum of the two displayed bounds is exactly $\mathcal{C}_d \|f\|_2^2$ in (13). Parseval finishes the normalized estimate. For a general nonzero f, apply it to $h = f/M$ with $M = \|f\|_\infty$. Linearity gives, uniformly in t,

$$\|T_t f\|_a^a = M^a \|T_t h\|_a^a \leq M^a \mathcal{C}_d \|h\|_2^2 = \mathcal{C}_d M^{a-2} \|f\|_2^2.$$

For the strong limit, $S_N f \to f$ in L^2 and $\|(I - S_N) f\|_\infty \leq (1 + K_d) \|f\|_\infty$. Thus the bound just proved makes $T_t(I - S_N) f$ small in L^a uniformly in t. On a fixed finite spectral sum termwise

differentiation gives convergence even in the uniform norm. The Cauchy criterion yields strong L^a convergence on bounded inputs. The Jacobi equation at zero gives $m_0'(n) = -n(n+d-1)m_0(n)$. Since L^a embeds in L^2 under probability measure, the limit is identified by its coefficients as $(\Delta/4)R_0 f$.

To recover the known strong limit at $d = 4$ from our general estimates, Lemma 8 gives $|q_t(n)| \le H_4 = 32$ without decay. Parseval proves the uniform L^2 bound. For any fixed $f \in L^2$, termwise convergence of each multiplier to $m_0'(n)$ and domination by $4H_4^2\|f_n\|_2^2$ prove strong convergence by summing the squares. This argument avoids the factor $1/(a-2)$ in (13). □

4 Why orthogonality forces an infinite moment

For a Borel cell $A \subset X$ and Borel decoder $b : X \to [0,1]$ define the rectangle response $r_{A,b}(x,y) = \mathbf{1}_A(x)b(y)$ and its averaged response

$$a_{A,b}(t) = \int r_{A,b}(x,y)\,\mathrm{d}\eta_t(x,y) = \langle \mathbf{1}_A, R_t b\rangle. \tag{15}$$

The rectangle is bounded and measurable. The function to which Sobolev locality will apply is $R_0 b$ on the preparation space.

Lemma 11 (Orthogonality for almost every tape)**.** *Assume Definitions 1–2 and exactness (2). There exists a measurable set Λ_0 of tape probability one such that for every $\lambda \in \Lambda_0$ and every label i,*

$$\int \mathbf{1}_{A_{\lambda i}}(x) b_i(\lambda, y)\,\mathrm{d}\eta_0(x,y) = 0. \tag{16}$$

The integrand vanishes η_0-almost everywhere for these tapes and labels. For each fixed $t \in [0,1]$, the measurable response $F_\lambda(t) = \sum_i a_{A_{\lambda i}, b_i(\lambda,\cdot)}(t)$ obeys $\mathbb{E}F_\lambda(t) = t$.

Proof. Use the common parametrization (4) to express every integral over $U(d)$. Joint Borel measurability of the encoder and decoder makes the integrands measurable in the tape and group element. Countable summation and Tonelli are applicable because all terms are nonnegative. The cells partition X, so summing the integrands over i gives the total conditional acceptance. Exactness at each fixed pair now gives

$$\int_\Lambda \sum_i a_{A_{\lambda i}, b_i(\lambda,\cdot)}(t)\rho(\,\mathrm{d}\lambda) = \int_{X^2} t(x,y)\,\mathrm{d}\eta_t(x,y) = t.$$

At $t = 0$ the integral of the nonnegative sum is zero. Outside a single null set its every nonnegative summand is zero. Each zero integral has a nonnegative integrand, proving the almost-everywhere statement. This uses a countable intersection over labels after integrating the orthogonal pairs, with no uncountable intersection over input pairs. □

For $d = 4$, the following zero-set conclusion is the one used in the ququart zero-slope proof of de Gois, Santos and Vieira [4, App. B.5]. We give the rotation argument here also for $d > 4$.

Lemma 12 (Regularity and locality of the excluded average)**.** *Let $d \ge 4$, $a = d-2$, $A \subset X$ measurable and $b \in L^\infty(X)$ with $0 \le b \le 1$. If $a_{A,b}(0) = 0$, then $w = R_0 b$ vanishes almost everywhere on A, belongs to $W^{a,2}(X)$, and satisfies $\Delta w = 0$ almost everywhere on A. Consequently*

$$\lim_{t \downarrow 0} a_{A,b}(t)/t = 0.$$

The Sobolev normalization is inherited from the phase-invariant lift to the unit sphere.

Proof. Positivity of the conditional average gives $w \geq 0$. Its integral on A is zero, so it vanishes there almost everywhere. Define the spectral $W^{a,2}$ norm using weights $(1 + 4n(n + d - 1))^a$ on the squared harmonic coefficients. For $n \geq 1$,

$$(1 + 4n(n + d - 1))^{a/2}|m_0(n)| \leq (1 + 4d)^{a/2}a!.$$

Here $n + d - 1 \leq dn$ and $\binom{n+a}{a} \geq n^a/a!$ were used. The constant level has multiplier one. This proves the spectral Sobolev assertion. In integer order it also controls weak derivatives. The following rotation argument gives directly the derivatives needed for locality.

Let $V_{ij} = x_i\partial_j - x_j\partial_i$ on the real sphere with $2d$ coordinates. Expanding on radially constant extensions gives $\Delta = \sum_{i<j} V_{ij}^2$. The rotations are skew-adjoint and preserve the full space of harmonic polynomials of ordinary degree k. Integration by parts yields

$$\sum_{i<j} \|V_{ij}v\|_2^2 = k(k + 2d - 2)\|v\|_2^2.$$

Apply the individual derivative bound again to $V_{ij}v$ in the same space to get $\|V_{ij}^2 v\|_2 \leq k(k+2d-2)\|v\|_2$. Products of up to a such derivatives are bounded by the corresponding powers of the square root of the eigenvalue. Distinct full harmonic degrees remain orthogonal. Thus the partial harmonic sums of w converge with these weak derivatives in L^2. The vector fields span each tangent space, and on each coordinate patch ordinary tangent derivatives are smooth linear combinations of them, identifying the usual local Sobolev class.

For completeness the level-set locality argument needs only two derivatives. Outside the fixed locus $x_i = x_j = 0$, a set of sphere measure zero, use local orbit coordinates for the V_{ij} rotation. In these coordinates V_{ij} is angle differentiation. Fubini for weak derivatives gives one-dimensional $W^{2,2}$ slices for almost every transverse coordinate. One-dimensional $W^{1,2}$ functions have absolutely continuous representatives. At almost every point of a measurable zero set both differentiation and the density property hold. Taking difference quotients along zeros proves that the first derivative vanishes there. Apply this to w on the slice and then to its first derivative, whose zero set contains almost all of the first zero set. This gives $V_{ij}^2 w = 0$ almost everywhere on the lifted A. A finite sum proves the Laplacian assertion. No interior of A is assumed.

Since $a_{A,b}(0) = 0$, the quotient in the statement is $\langle \mathbf{1}_A, T_t b\rangle$. The strong convergence in Theorem 6, also for $d = 4$, gives its limit as $\langle \mathbf{1}_A, (\Delta/4)w\rangle = 0$. □

Lemma 13 (Bridge inequality)**.** *Let $d \geq 4$, $a = d - 2$ and let $\lambda \in \Lambda_0$ be a tape from Lemma 11 with $D_+(\lambda) < \infty$. Then for every $0 < t \leq 1/2$,*

$$0 \leq F_\lambda(t)/t \leq B_d^{\mathrm{op}} D_+(\lambda)^{1/a}, \qquad \lim_{t\downarrow 0} F_\lambda(t)/t = 0,$$

where B_d^{op} is defined after (13). In particular $B_4^{\mathrm{op}} = 8$, using the ququart difference-quotient and zero-slope estimates of de Gois, Santos and Vieira [4, Apps. B.5–B.6].

Proof. Cells of zero mass contribute nothing because the first marginal of η_t is μ. Let $I_+ = \{i : \mu(A_{\lambda i}) > 0\}$ and $\mu_i = \mu(A_{\lambda i})$. Write $a_i(t)$ for the ith cell response. The positive cells partition a set of full measure.

Dimension four. On the full-measure set of tapes from Lemma 11, $a_i(0) = 0$ for every i. Apply Cauchy–Schwarz on X and the multiplier bound of [4, App. B.5.3, Eq. (167)]. Since $\|b_i\|_2 \leq 1$,

$$0 \leq a_i(t)/t = \langle \mathbf{1}_{A_{\lambda i}}, T_t b_i\rangle \leq 8\sqrt{\mu_i}.$$

A second Cauchy–Schwarz inequality, this time on the finite set of cells, gives

$$0 \leq F_\lambda(t)/t \leq 8 \sum_{i\in I_+} \sqrt{\mu_i} \leq 8\sqrt{D_+(\lambda)}.$$

Each summand tends to zero by the zero-slope result in [4, App. B.5], including for cells of empty interior. This proves both assertions at $d = 4$. It also gives the obstruction directly. If $\mathbb{E}\sqrt{D_+} < \infty$, then $D_+ < \infty$ almost surely and the last bound permits dominated convergence. But $\mathbb{E}F_\lambda(t)/t = 1$ for every $t > 0$, a contradiction. In particular $\mathbb{E}\sqrt{L} = \infty$. Castellano, Sekatski and Brunner [2, App. A, Eqs. (92)–(94)] bound each cell by 8 and sum to $8D$ for an alphabet of size D. The square root enters here because the individual estimates retain μ_i.

Dimensions $d \geq 5$. For $i \in I_+$ subtract (16) and apply Hölder with conjugate exponents $a/(a-1)$ and a.

$$0 \leq a_i(t)/t = \langle \mathbf{1}_{A_{\lambda i}}, T_t b_i \rangle \leq \mu_i^{1-1/a} \|T_t b_i\|_a \leq B_d^{\mathrm{op}} \mu_i^{1-1/a}.$$

The last inequality uses $\|b_i\|_\infty, \|b_i\|_2 \leq 1$ in the strong L^a estimate of Theorem 6. Sum the finite set of terms. A second use of Hölder, on counting measure over I_+, gives

$$\sum_{i \in I_+} \mu_i^{1-1/a} \leq \left(\sum_{i \in I_+} \mu_i \right)^{1-1/a} |I_+|^{1/a} = D_+(\lambda)^{1/a}.$$

The total positive-cell mass is one because the partition is countable. Each cell quotient tends to zero by Lemma 12, so their finite sum does too. □

Universal part of Theorem 3. Assume $\mathbb{E}D_+^{1/a} < \infty$. Then $D_+ < \infty$ almost surely. Intersect this set with Λ_0 and choose the deterministic sequence $t_k = 1/(k+2)$. Lemma 13 supplies one integrable bound independent of k and a pointwise limit zero. Dominated convergence gives $\mathbb{E}F_\lambda(t_k)/t_k \to 0$. Lemma 11 gives $\mathbb{E}F_\lambda(t_k)/t_k = 1$ for every k, a contradiction. Since $D_+ \leq D$, the asserted obstruction holds also for the full used alphabet. □

The declared-alphabet converse also holds on a general probability space for shared randomness under the joint measurability assumed in [2]. Only D_+ enters the contradiction. Each cell mass is measurable by integration on the product space, and $D_+ \leq L$ pointwise. Thus the argument for a measurable declared L never needs to project a nonempty input cell onto that more general tape space.

5 An exact mixture of cap responses

The scale distribution must be positive, reproduce the exact target at every fidelity, and have a specified mass near zero. The following lemmas establish these properties. Let $C_s(x) = \{z : t(x,z) > 1 - s\}$ for $0 < s \leq 1/2$. Let $K_s(t)$ be the probability that a uniform point of $C_s(x)$ has fidelity greater than s with y, given $t(x,y) = t$. Write $(v)_k$ for the rising factorial.

Lemma 14. *For every $d \geq 4$ there is a positive probability law of $U \in (0,1]$ with moments, for integers $k \geq 0$,*

$$Q_d(k) = \frac{(2)_k (d)_k (d+1)_k}{(d+2)_k ((d+1)/2)_k ((d+2)/2)_k}.$$

It has $\mathbb{E}U^{-1} = (d+1)/4$ and $\mathbb{P}(U \leq u) \sim r_d u^2/2$ with $r_d = d(d+1)(d-3)/[16(d-1)]$.

Proof. At $d = 4$ take the density

$$f(u) = \tfrac{5}{4} \left[u(1-u)^{-1/2} - 2u^4 \int_u^1 v^{-4}(1-v)^{-1/2} \, \mathrm{d}v \right].$$

The Mellin integral at $x > -2$ is $\frac{5}{4}(x+3)B(x+2, 1/2)/(x+5)$ by Fubini on each nonnegative term. Its values at zero and minus one are 1 and $5/4$. For positivity set $h = 4f/(5u^4)$. Differentiation gives $h' = (-1 + 3u/2)/[u^4(1-u)^{3/2}]$. At its minimum $u = 2/3$ write $z = 1/\sqrt{3}$. The integral

is $2\int_0^z (1-w^2)^{-4}\,\mathrm{d}w \le 113z/24$ by the convex chord in w^2 between 0 and $1/3$. Thus $h(2/3) \ge (81/8 - 113/12)z > 0$. At zero $u^3\int_u^1 v^{-4}(1-v)^{-1/2}\,\mathrm{d}v \to 1/3$, so $f(u) \sim 5u/12$. This gives $r_4 = 5/12$.

Gamma duplication expresses the proposed transform as

$$Q_d(x) = 4^x\Gamma(x+2)\frac{d(d+1)\Gamma(x+d)}{(x+d+1)\Gamma(2x+d+1)}.$$

The ratio $Q_{d+1}(x)/Q_d(x)$ is

$$B_d(x) = \frac{d+2}{d}\frac{(x+d)(x+d+1)}{(x+d+2)(2x+d+1)}.$$

This is the Mellin transform of V_d having mass $(d+2)/(2d)$ at one and the following density on $(0,1)$:

$$\frac{d+2}{4d(d+3)}\left[(d^2-1)v^{(d-1)/2} - 8v^{d+1}\right].$$

For $0 < v < 1$, the bracket equals $v^{(d-1)/2}[(d^2-1) - 8v^{(d+3)/2}]$, which is positive since $d^2 - 1 \ge 15 > 8$. Its integral against v^x, including the atom, is

$$\frac{d+2}{2d} + \frac{d+2}{4d(d+3)}\left[\frac{d^2-1}{x+(d+1)/2} - \frac{8}{x+d+2}\right] = B_d(x).$$

Multiplying both sides by $(x+d+2)(2x+d+1)$ reduces this identity to equality of three quadratic coefficients. At $x = 0$ the right side is one, so the atom and density define a probability law. This is an identity in the real variable $x > -(d+1)/2$, not a finite moment interpolation. The integrals also exist at $x = -2$. Taking $U_{d+1} = U_d V_d$ independently now yields the moments and the claimed negative first moment by induction. If $H_d(u) = \mathbb{P}(U_d \le u) \sim r_d u^2/2$, then $H_d(v) \le c_d v^2$ for some finite dimension-dependent constant for all $v > 0$, extending the CDF by one above one. Apply dominated convergence to $u^{-2}\mathbb{E}H_d(u/V_d)$, dominated by $c_d V_d^{-2}$. This gives $r_{d+1} = r_d B_d(-2)$. Here

$$B_d(-2) = \frac{(d+2)(d-2)(d-1)}{d^2(d-3)}.$$

Multiplying this by $r_d = d(d+1)(d-3)/[16(d-1)]$ gives $(d+1)(d+2)(d-2)/(16d) = r_{d+1}$. Also $B_d(-1) = (d+2)/(d+1)$, which propagates the negative first moment. The negative second moment of the factor V_d is finite. No such moment is required of U_d. □

Lemma 15. *Let η be the law of $S = (1-\sqrt{1-U})/2$ from Lemma 14, and set*

$$\mathrm{d}\nu_d(s) = \frac{\mathrm{d}\eta(s)}{d(d+1)s(1-s)^d}.$$

Then ν_d is a probability measure, $\int K_s(t)\,\mathrm{d}\nu_d(s) = t$ for all $t \in [0,1]$, and $\nu_d((0,s]) \sim [(d-3)/(d-1)]s$.

Proof. Finiteness follows from $s^{-1}(1-s)^{-d} \le 2^{d+1}/U$. For $\ell \ge 1$ write $j = \ell - 1$, $B_{\ell,d-2} = \binom{\ell+d-2}{d-2}$ and $c_{\ell,d} = (d-1)/(\ell B_{\ell,d-2})$. A cap of Haar mass v^{d-1} has unnormalized level-ℓ multiplier

$$F_\ell(v) = \frac{d-1}{B_{\ell,d-2}}\int_0^v w^{d-2}P_\ell^{(d-2,0)}(1-2w)\,\mathrm{d}w = c_{\ell,d}v^{d-1}(1-v)P_j^{(d-1,1)}(1-2v).$$

For the second equality, shifted Rodrigues gives $[v^{d-1}(1-v)P_j^{(d-1,1)}(1-2v)]' = \ell v^{d-2}P_\ell^{(d-2,0)}(1-2v)$, and the primitive vanishes at zero. Thus the two multipliers before normalizing the encoder are

$$F_\ell(s) = c_{\ell,d}s^{d-1}(1-s)P_j^{(d-1,1)}(1-2s),$$
$$F_\ell(1-s) = c_{\ell,d}(1-s)^{d-1}sP_j^{(d-1,1)}(2s-1).$$

The encoder draws uniformly inside its cap, whose mass is s^{d-1}, while the decoder uses an indicator. Composition of the radial operators therefore gives the level-ℓ multiplier of K_s.

$$\frac{F_\ell(s)F_\ell(1-s)}{s^{d-1}} = c_{\ell,d}^2 s(1-s)^d P_j^{(d-1,1)}(1-2s)P_j^{(d-1,1)}(2s-1).$$

Here is the coefficient computation. Set $z = 1-2s$, $u = 1-z^2$, $p = P_j^{(d-1,1)}(z)$, $q = p(-z)$, $Y = pq$, $Z = p'q - pq'$, $W = p'q'$, $\alpha = j(j+d+1)$. Primes on p, q are z-derivatives. All three products are even. The Jacobi equations give $up'' = [(d+2)z + d - 2]p' - \alpha p$ and the reflected equation for q, changing $+d-2$ to $-(d-2)$. Substitution and $\partial_z = -2z\partial_u$ give

$$4u(1-u)Y_{uu} + [2(d+2)(1-u) - 2u]Y_u + 2\alpha Y - (d-2)Z - 2uW = 0,$$
$$uZ_u = -(d+2)Z/2 + (d-2)Y_u, \qquad uW_u = -(d+2)W - \alpha Y_u.$$

Write $Y = \sum b_k u^k$. The last two identities give $Z_k = (d-2)(k+1)b_{k+1}/[k + (d+2)/2]$ and $W_k = -\alpha(k+1)b_{k+1}/(k+d+2)$. In the first equation the coefficient multiplying b_{k+1} becomes $4(k+1)(k+2)(k+d)/[k+(d+2)/2]$, and that multiplying b_k becomes $2(2k+d+1)[\alpha - k(k+d+1)]/(k+d+1)$. Multiplying the resulting ratio by $Q_d(k+1)/Q_d(k)$ cancels all but $(k-j)(k+j+d+1)/[(k+1)(k+d+2)]$. For $j \geq 1$ the mixed coefficient is zero by the finite-difference identity

$$\frac{\mathbb{E}Y}{b_0} = \sum_{k=0}^{j}(-1)^k\binom{j}{k}\frac{(j+d+1)_k}{(d+2)_k} = \sum_{k=0}^{j}(-1)^k\binom{j}{k}\frac{(k+d+2)_{j-1}}{(d+2)_{j-1}} = 0.$$

Here $b_0 = P_j^{(d-1,1)}(1)P_j^{(d-1,1)}(-1) \neq 0$. The last ratio is a polynomial in k of degree $j-1$. Its jth forward difference vanishes. The constant $c_{\ell,d}^2$ is independent of k, so every mixed multiplier of degree $\ell \geq 2$ vanishes.

For level one the constant in each primitive is one, giving mixed multiplier $1/[d(d+1)]$. For level zero the mixed multiplier is $\int(1-s)^{d-1}\,\mathrm{d}\nu_d = 4Q_d(-1)/[d(d+1)] = 1/d$. The kernel t has precisely these multipliers because $t = [1 + (d-1)\phi_1]/d$, and the projection kernel on level one is $(d^2-1)\phi_1$, since this space consists of traceless forms. In general the projection kernel is the dimension times the normalized zonal function by invariance and its diagonal trace. Completeness gives equality of the kernels almost everywhere. Both are continuous, since rotating either cap changes its indicator continuously in L^1 and ν_d is finite. Hence equality holds everywhere. At $t = 1$ all the kernels equal one, proving normalization.

The small-s CDF of η is asymptotic to $8r_d s^2$ by the change of variables. For $H(s) = \eta((0,s])$, integration by parts gives $\int_0^s v^{-1}\,\mathrm{d}H(v) = H(s)/s + \int_0^s H(v)v^{-2}\,\mathrm{d}v \sim 16r_d s$. Multiplication by $(1-s)^{-d}$ changes the leading term by a factor tending to one. Substitution of r_d gives the assertion. □

5.1 Explicit form and sampling of the scale law

The multiplicative recursion also gives a density construction. Let f_4 be the density in Lemma 14, let $\alpha_d = (d+2)/(2d)$, and let g_d be the displayed density of the continuous part of V_d. Then recursively

$$f_{d+1}(u) = \alpha_d f_d(u) + \int_u^1 f_d(u/v)g_d(v)\frac{\mathrm{d}v}{v}, \qquad 0 < u < 1. \tag{17}$$

This follows by conditioning on V_d and changing variables in the density of $U_d V_d$. It has unit mass and is nonnegative by the same conditioning. The factors have positive atoms only at one, while U_4 has a density, so every U_d has a density. The measure ν_d on the interval $(0, 1/2)$ consequently has density

$$v_d(s) = \frac{4(1-2s)f_d(4s(1-s))}{d(d+1)s(1-s)^d}. \tag{18}$$

The normalization and exact mixture identity have been proved in Lemma 15. Shared uniforms sample any probability distribution used below by its generalized inverse CDF. Thus (17) and (18) specify the scale distribution without reference to any private data or stored certificate.

6 An exact protocol for binary tests

Fix $d \geq 4$ and retain $a = d - 2$. Let

$$b = 2(d-1), \quad c = 3/4, \quad \sigma = 4^{-(d-1)}, \quad C = 2^b(d-1)(5/4)^2, \quad H = 9^b. \tag{19}$$

These are numerical constants once d is fixed. Choose $\delta = 2^{-k_d}$, where $k_d \geq 5$ is the first integer such that

$$m_\delta := \nu_d((0,\delta)) < \frac{\sigma}{2(1/3+\sigma)}. \tag{20}$$

Lemma 15 implies $m_\delta \to 0$ as $\delta \to 0$, so k_d exists. The strict factor two leaves positive slack in the budget. A scale is called fine if $s < \delta$.

Lemma 16 (Metric, cap volume and endpoint response)**.** *On X, $d_c(x,z) = \|P_x - P_z\|_{\mathrm{op}} = \sqrt{1-t(x,z)}$ is a metric. A radius-r open ball has Haar mass r^b for $0 \leq r \leq 1$. The cap $C_s(x)$ is its radius-$\sqrt{s}$ ball. For $0 < s \leq 1/2$, $K_s(0) = 0$, $K_s(1) = 1$, and $K_s(t) = 1$ if $t > 4s(1-s)$.*

Proof. For distinct rays, $P_x - P_z$ on their span has trace zero and $\mathrm{tr}((P_x - P_z)^2) = 2(1-t)$, hence eigenvalues $\pm\sqrt{1-t}$. For coincident rays the difference is zero. The operator norm triangle inequality gives the metric. Integrating the Haar fidelity density above $1 - r^2$ gives $r^{2(d-1)}$. At orthogonality the projection of an encoder cap point onto $x^\perp$ has squared norm less than s. At unit fidelity the encoder threshold $1 - s$ is at least the decoder threshold s. For the remaining assertion decompose z along x. The smallest possible overlap amplitude with y for z in the cap is bounded below by $\sqrt{(1-s)t} - \sqrt{s(1-t)}$. For $s < 1/2$ it exceeds $\sqrt{s}$ when $\sqrt{1-t} < 1 - 2s$, equivalent to $t > 4s(1-s)$. This also follows by squaring the two nonnegative sides of $\sqrt{(1-s)t} > \sqrt{s} + \sqrt{s(1-t)}$. At $s = 1/2$ the sufficient interval is empty and the endpoint $t = 1$ was already treated. □

6.1 Public lists and the exact anchor law

Lemma 17 (Rotated net sampler)**.** *For each fine scale s there is a Borel public-tape sampler selecting a point uniformly in $C_s(x)$ for every fixed input x. Its successful round is at most a geometric random variable J of parameter H^{-1} simultaneously for all x, almost surely. J is independent of s.*

Proof. Fix a dense countable list of rays by normalizing nonzero Gaussian-rational vectors and identifying phases by their projectors. For each dyadic h, scan the list and retain a point whenever its distance to every earlier retained point is at least h. The retained points have disjoint open $h/2$-balls, so there are finitely many. If some point had distance strictly greater than h from the retained list, its neighborhood would contain a dense-list point that should have been retained. Thus the covering radius is at most h.

Set $r = \sqrt{s}$ and choose the unique dyadic h with $r/4 < h \leq r/2$, with this half-open convention. Let $z_1, \ldots, z_L$ be the resulting deterministic list. Covering and packing imply $h^{-b} \leq L \leq (2/h)^b$. Every radius-$r$ open ball contains at least one centre. The disjoint $h/2$-balls of centres in such a ball are contained in a ball of radius $r + h/2$. Their number is at most $(2r/h + 1)^b \leq 9^b = H$.

Each round uses a fresh Haar rotation G_j and independent uniforms U_j, V_j. Let $k_j(x)$ count the hits $G_j z_i \in C_s(x)$. Alice accepts the round when $U_j \leq k_j(x)/H$ and uses V_j to select a uniform hit, ordered by the original list index. On a Borel set E, one-round success with selected point in E has probability

$$\frac{1}{H}\mathbb{E}\#\{i : G_j z_i \in E \cap C_s(x)\} = \frac{L}{H}\mu(E \cap C_s(x)).$$

Success has probability $Ls^{d-1}/H > 0$. Independence of rounds and geometric summation give selected law $\mu(E \cap C_s(x))/s^{d-1}$. Define $J = \min\{j : U_j \leq 1/H\}$. Every input still searching succeeds by round J, since it always has at least one hit. The uniforms defining J are independent of the scale. □

Lemma 18 (Uniform unsafe fraction)**.** *For a fine scale s, every rotated list of Lemma 17 and every test y, the number B_y of centres z with $t(y,z) \le s$ satisfies $B_y/L \le Cs$.*

Proof. For $d_c(w,z) < h/2$, decompose a unit representative of w along z. The perpendicular component has norm $d_c(w,z)$ and the parallel coefficient has modulus at most one. Thus $|\langle y,w\rangle| \le |\langle y,z\rangle| + h/2 \le r + h/2$ for an unsafe centre. The band with amplitude at most $v \le 1$ has mass $1-(1-v^2)^{d-1} \le (d-1)v^2$. The disjoint balls therefore give $B_y(h/2)^b \le (d-1)(r+h/2)^2$. Since $L(h/2)^b \ge 2^{-b}$ and $h/2 \le r/4$, division gives exactly $B_y/L \le Cs$. The band bound applies since $s < \delta \le 1/32$. □

6.2 Packet alphabet and decoder

Lemma 19 (Grouped response and actual range)**.** *At each fine scale set*

$$n(s) = \max\{1, \lfloor (8Cs)^{-1} \rfloor\}, \qquad M_d = 16C\,8^b + 1.$$

There is a protocol with response $cK_s(t)$ for every input pair. Its used alphabet has size at most $M_d J s^{-a}$ and at least $1/\lceil n(s)s^{d-1}\rceil$ almost surely. For $s < 1/(8C)$ the latter is at least $8Cs^{-a}$.

Proof. Independently of the round data used to choose the anchor, uniformly permute each round's L labels. Divide the permuted list into consecutive packets of size $n(s)$, allowing a shorter last packet. The public declared alphabet is

$$\mathcal{A}_{s,\lambda} = \{(j,g) : 1 \le j \le J(\lambda),\ 1 \le g \le \lceil L/n(s)\rceil\}.$$

Alice sends the pair consisting of the first successful round and the packet containing its selected anchor. Neither selection uses the permutation. Bob reconstructs the indicated round and packet and counts B_y for the entire round. If any packet member is unsafe, he rejects. If every member is safe and its actual size is g, he computes

$$p_g = \frac{\binom{L-1-B_y}{g-1}}{\binom{L-1}{g-1}}, \qquad p_1 = 1,$$

then accepts with probability c/p_g, using a fresh independent coin.

To justify this rule, condition on the round list, the selected anchor and the size of its packet. The anchor selection is independent of the permutation. For a safe anchor the remaining $g-1$ members are a uniform subset of the other $L-1$ labels. Thus p_g is the conditional all-safe probability. Lemma 18, $L \ge 2$ and a union bound give

$$p_g \ge 1 - \frac{(g-1)B_y}{L-1} \ge 1 - 2Cs(g-1) \ge 1 - 2Cs(n(s)-1) \ge 3/4 = c.$$

The last inequality holds also when $n = 1$. Conditional on a safe anchor, averaging over the packet and the additional coin gives acceptance $p_g(c/p_g) = c$. An unsafe anchor forces rejection. Averaging the exactly uniform anchor proves response $cK_s(t)$. Conditioning on the size treats the short final packet explicitly.

The full declared alphabet has size $J\lceil L/n\rceil$. Since $n \ge 1/(16Cs)$, $L \le 8^b s^{-(d-1)}$, and $s < 1$, this is at most $M_d J s^{-a}$. A used pair (j,g) can encode only inputs whose anchor lies in its packet. These inputs belong to a union of at most n cap balls, of total mass at most ns^{d-1}. The union over the actually used labels covers X, proving the lower bound. If $s < 1/(8C)$ then $n \le (8Cs)^{-1}$, giving its stated form. These inequalities concern the range over all inputs on one fixed tape. The geometric round factor remains random at fixed scale. □

6.3 Coarse sampling and a common finite alphabet

Lemma 20 (Uniform cap and annulus samples)**.** *For every fixed $s > 0$, a first-hit iid Haar sampler in a cap or in the annulus $C_s(x) \setminus C_{s/4}(x)$ has the uniform conditional distribution there. There is a common random integer Z dominating all first-hit indices for every x and every $\delta/128 \le s \le 1/2$, with*

$$\Pr(Z > j) \le N_d(1-\epsilon^b)^j, \qquad \epsilon = \sqrt{\delta/128}/16, \quad N_d \le (2/\epsilon)^b.$$

Proof. Let E be the chosen region of mass $p > 0$. For any Borel set G, geometric summation gives the selected distribution $\sum_{j\ge1}(1-p)^{j-1}\mu(G\cap E) = \mu(G\cap E)/p$. Choose a fixed ϵ-net by the packing construction, of size at most $(2/\epsilon)^b$. Define Z as the first prefix of the Haar tape hitting each net point's open ϵ-ball. A union bound gives the displayed tail. This prefix is a 2ϵ-net. For a given x and radius $r = \sqrt{s} \ge \sqrt{\delta/128}$ choose a ray at distance $3r/4$ from x on a projective line. A prefix ray within $2\epsilon \le r/8$ has distance strictly between $r/2$ and r from x, and qualifies for both sampling rules. Thus their first-hit indices never exceed Z on that tape, simultaneously for all inputs and scales. In particular $Z \ge 1$ and all its positive moments are finite. □

Lemma 21 (Two local exact responses)**.** *There are two protocols with responses $P_{\mathrm{near}}(t) = t$ for $0 \le t \le \delta/4$ and $P_{\mathrm{out}}(t) = t$ for $\delta/16 \le t \le 1$. Only the near protocol uses the fine packets of Lemma 19. All its other branches, and the outer protocol, have range bounded by a variable Z of Lemma 20.*

Proof. For $s \ge \delta$, the inner cap $C_{s/4}(x)$ has conditional mass $\sigma = 4^{-(d-1)}$ inside $C_s(x)$. If $t(x,y) \le \delta/4$, an inner point satisfies $|\langle y,z\rangle| \le \sqrt{t} + d_c(x,z) < \sqrt{s}$. The removed part always rejects. If $A_s(t)$ denotes the annular response, then $K_s(t) = (1-\sigma)A_s(t)$ on this interval.

The near protocol uses the following unconditional branch measures, all sampled publicly. On $(0,\delta)$ give scale measure $\mathrm{d}\nu_d/c$ and use the packet rule of response cK_s. On $[\delta, 1/2)$ give scale measure $(1-\sigma)\,\mathrm{d}\nu_d$ and use the annular rule. The remaining branch always rejects with one fixed label. The total nondefault mass is

$$W = c^{-1}m_\delta + (1-\sigma)(1-m_\delta), \qquad 1 - W = \sigma - (1/3+\sigma)m_\delta > \sigma/2$$

by (20). Hence this is a probability distribution. On the near interval its response is exactly the full mixture in Lemma 15.

For the outer protocol sample a scale from ν_d. Always accept at $s < \delta/128$, with one label. Otherwise use the first-hit cap sampler. If $t \ge \delta/16$, the discarded scales already have $K_s(t) = 1$, since $4s(1-s) < \delta/32 < t$. Thus the outer response is exact there. Lemma 20 supplies all claimed range bounds. The branch and scale are public and require no transmitted branch label. □

Lemma 22 (Exact deterministic routing)**.** *There is a Borel one-way protocol exact for every pair $(x,y) \in X^2$, with actual range bounded by $N_0 D_{\mathrm{near}} D_{\mathrm{out}}$, where $N_0 \le (64/\delta)^b$ is fixed and the two subprotocol tapes are independent.*

Proof. Fix a $\delta/32$-net. Alice sends a nearest point index x_0, breaking ties by its fixed order, and both near and outer messages. Encode this finite tuple as one natural-number label by a fixed injective coding. Bob uses the near message if $t(x_0,y) \le \delta/8$ and uses the outer message otherwise. The projector norm bound gives $|t(x_0,y) - t(x,y)| \le d_c(x_0,x) \le \delta/32$. The first case implies $t \le 5\delta/32 < \delta/4$, and the second implies $t > 3\delta/32 > \delta/16$. The routing decision depends only on the fixed inputs and deterministic net. Conditional selection therefore preserves the exact mean of the selected random subprotocol. Multiplying the three range bounds gives the assertion, regardless of whether every tuple is realizable.

To verify measurability, all base lists belong to a countable dyadic family, with a Borel scale-to-list choice. Each round uses finitely many strict fidelity comparisons. The uniformly chosen hit uses

$\lfloor kV \rfloor$ in fixed list order. A uniform permutation of a finite list is obtained by sorting iid continuous priorities, resolving ties by index. First successful rounds and first-hit indices are countable stopping rules. The common bounds J and Z are measurable finite random variables almost surely. On their null failure sets one can assign a default label. All threshold conventions are included above, and the extra acceptance coin is independent. This defines a joint Borel encoder and decoder satisfying Definition 2. □

7 The two-sided alphabet tail

Lemma 23 (Explicit tail constants)**.** *The protocol of Lemma 22 has the tail in Theorem 3. Its fine-scale input-independent law, conditional kernels, and common round bound determine constants depending only on d.*

Proof. Put $\gamma = 1/a$ and $k_d^\nu = (d-3)/(d-1) > 0$. Define the finite constant

$$A_d = \sup_{0<r\le\delta} \frac{\nu_d((0,r))}{cr}.$$

Its finiteness follows from Lemma 15 at zero and from bounded total mass away from zero. Extend the fine-branch cumulative mass by its total mass above δ. It remains bounded by $A_d r$. Conditional on J, the event $M_d J s^{-a} > m$ restricts s to an interval of length $(M_d J/m)^\gamma$. Integrating the scale first gives a near fine-tail bound $A_d M_d^\gamma m^{-\gamma} \mathbb{E} J^\gamma$. Markov's inequality for the other branches gives at most $m^{-\gamma}\mathbb{E}Z^\gamma$. Consequently, for all $m > 0$,

$$\Pr(D_{\mathrm{near}} > m) \le B_d m^{-\gamma}, \quad B_d = A_d M_d^\gamma \mathbb{E} J^\gamma + \mathbb{E} Z^\gamma.$$

All expectations are finite by their explicit geometric tail bounds. The independent outer tape has range bounded by a copy Z' of Z. Condition on Z' in $D \le N_0 D_{\mathrm{near}} Z'$ to obtain

$$\Pr(D > m) \le B_d N_0^\gamma \mathbb{E} Z^\gamma \, m^{-\gamma}.$$

This supplies a finite upper constant specified entirely by d and the fixed net choices.

Choose $s_0 < \min(\delta, 1/(8C))$ such that $\nu_d((0,r)) \ge k_d^\nu r/2$ for all $0 < r \le s_0$. At a fine tape with $s < s_0$, projecting joint transcripts onto their near packet label shows $D \ge 8Cs^{-a}$. For $m \ge 8Cs_0^{-a}$, every fine scale $s < (8C/m)^\gamma$ therefore gives $D > m$. Its unconditional branch mass is $\nu_d((0,(8C/m)^\gamma))/c$. Hence

$$\Pr(D > m) \ge \frac{k_d^\nu}{2c}(8C)^\gamma m^{-\gamma}.$$

Together the bounds give the assertion. Lemma 5 yields finite moments for every $q < \gamma$ for this one protocol. Combined with Section 4, this completes Theorem 3 for every $d \ge 4$. □

8 From binary tests to general measurements

Theorem 24 (Finite-overhead reduction)**.** *In fixed dimension $d \ge 2$, an exact one-way binary rank-one simulator with declared alphabet $L(\lambda)$ gives an exact one-way simulator for arbitrary states and POVMs with declared alphabet at most $C_d L(\lambda)$ on every augmented tape, up to the harmless default on null tapes. The constant depends only on d.*

The proof constructs a decoder whose outputs form a probability distribution for each tape and input. Its expectation matches the Born law. Finite and countable outcomes are handled first.

8.1 A finite local description and a filtered query

Constants s, r, δ, U in this section are local to the reduction and independent of the random scale in the binary construction. Write $P_x = xx^*$ for a unit vector x. Choose

$$s = \frac{1}{32d}, \quad r = \frac{s}{4}, \quad m = d^2 - 1, \quad \delta = \frac{r}{\sqrt{2m}}, \quad U = 2s.$$

Take a finite δ-net of pure projectors in operator distance. Alice sends a nearest centre P_0, breaking ties by list order. Thus $\|P_x - P_0\|_{\mathrm{op}} \le \delta$. The net is fixed and does not depend on Bob's measurement.

A formal matrix with the right mean. Fix a real Hilbert–Schmidt orthonormal basis $H_1, \dots, H_m$ of the traceless Hermitian matrices. Write $P_x - P_0 = \sum_k \xi_k H_k$. Two rank-one projectors satisfy $\|P_x - P_0\|_{\mathrm{HS}} = \sqrt{2}\|P_x - P_0\|_{\mathrm{op}}$. Therefore $\sum_k |\xi_k| \le r$. With probability $|\xi_k|/r$ Alice chooses $Q = P_0 + r\,\mathrm{sign}(\xi_k) H_k$. She chooses P_0 with the remaining probability. She sends the vertex index. The $1 + 2m$ possible matrices satisfy

$$\operatorname{tr} Q = 1, \qquad \|Q - P_0\|_{\mathrm{op}} \le r, \qquad \mathbb{E}Q = P_x.$$

Some vertices have negative eigenvalues. The decoder below uses only probabilities proved nonnegative for each such vertex.

A filtered binary query. Define $A = \sqrt{s}P_0 + (I - P_0)$, and let

$$u = \|Ax\|^2 = s + (1 - s)\|P_x - P_0\|_{\mathrm{op}}^2 \le U, \qquad x' = Ax/\sqrt{u}.$$

Alice sends the original binary encoder's single message for x'. Using a fresh independent uniform she sends a bit G with probability u/U of equaling one. All additional sender randomness may be placed on the public tape. Alice needs no information about the measurement.

Write $b(y') \in [0, 1]$ for the original decoder on this tape and message queried at a test ray y'. For each y, put $v_y = \|A^{-1}y\|^2$ and $y' = A^{-1}y/\sqrt{v_y}$. Then

$$|\langle x', y'\rangle|^2 = \frac{|\langle x, y\rangle|^2}{u v_y}, \qquad \mathbb{E}[U v_y G b(y')] = |\langle x, y\rangle|^2. \tag{21}$$

The first identity uses self-adjointness and invertibility of A. The second uses independence of the gate and the binary tape, and exactness for x', y'. Only rays with $t_0(y) := \operatorname{tr}(P_0 P_y) \le s$ will use this score. For them $v_y = 1 + (s^{-1} - 1)t_0(y) \le 2$. Its maximum contribution per unit of effect weight is at most $4s$, uniformly on every tape.

8.2 One normalized distribution on every tape

First take a rank-one POVM $E_i = w_i P_{y_i}$, omitting zero effects. Its weights obey $\sum_i w_i P_{y_i} = I$ and $\sum_i w_i = d$. For a Hermitian H, put $f_i(H) = w_i \operatorname{tr}(H P_{y_i})$. Each is linear, and $\sum_i f_i(H) = \operatorname{tr} H$.

Partition the indices into $N = \{i : t_0(y_i) \le s\}$ and its complement F. The far reference mass $B = \sum_{i \in F} f_i(P_0)$ is at least $1 - ds = 31/32$. For $j \in F$ set $R_j = f_j(P_0)/B$, giving a probability distribution on F depending only on the centre and the measurement.

For a fixed tape and received message, Bob defines

$$\begin{aligned} q_i &= w_i U v_{y_i} G b(y_i'), && i \in N, \\ S &= \sum_{i \in N} q_i, \quad T = \sum_{i \in N} f_i(Q), \\ q_j &= f_j(Q) + R_j(T - S), && j \in F. \end{aligned} \tag{22}$$

He samples one index from this distribution. Evaluations of b at different rays use one and the same binary message. Joint independence of those scores is unnecessary.

Normalization. On each tape $\sum_i q_i = S + (1 - T) + (T - S) = 1$.

Positivity. Near entries are nonnegative and $S \le 4sd = 1/8$. Since $f_i(P_0) \ge 0$ and $\|Q - P_0\|_{\mathrm{op}} \le r$, we have $T \ge -dr = -1/128$. For a far index $t_0(y_j) > s$ implies $f_j(Q) \ge (1 - r/s) f_j(P_0) = 3f_j(P_0)/4$. Consequently

$$q_j \ge f_j(P_0) \left[\frac{3}{4} - \frac{1/8 + 1/128}{31/32} \right] = \frac{19}{31} f_j(P_0) > 0.$$

The estimate holds for any values of the binary scores in $[0, 1]$, even under adversarial correlation. The gate and the same finite message suffice for every Bob input.

Exactness. For fixed preparation and measurement, (21) gives $\mathbb{E}q_i = f_i(P_x)$ on N. Linearity and $\mathbb{E}Q = P_x$ give $\mathbb{E}T = \sum_{i \in N} f_i(P_x) = \mathbb{E}S$. Since R_j is fixed by the inputs, $\mathbb{E}q_j = f_j(P_x)$ on F. A near outcome with zero Born probability has zero mean and is nonnegative on each tape, so is zero almost surely for the fixed input. The proof includes orthogonal inputs and the near-set boundary.

A four-outcome example. Take $d = 4$, a net containing $P_0 = P_{e_1}$, and a preparation $x = \sqrt{1 - \eta}\, e_1 + \sqrt{\eta}\, e_2$ so close to e_1 that P_0 is the nearest centre. If Bob measures in the coordinate basis, only outcome 1 is far. Its redistribution weight is $R_1 = 1$. Since $\sum_i f_i(Q) = 1$, the rule gives $q_1 = 1 - S$ on every tape. For $i = 2, 3, 4$ one has $v_{e_i} = 1$ and $q_i = UGb(e_i)$. Equation (21) gives $\mathbb{E}q_2 = \eta$ and $\mathbb{E}q_3 = \mathbb{E}q_4 = 0$. Thus Bob's average law is $(1 - \eta, \eta, 0, 0)$ even though the total near mass S may change from tape to tape. Removing the subtraction of S in (22) would give total mass $1 + S$ whenever $S > 0$.

8.3 General effects, mixed states and alphabet range

Arbitrary effects. Fix an orthonormal basis (e_k). For an outcome with effect E, refine it into its at most d rank-one terms

$$E = \sum_{k=1}^{d} |E^{1/2} e_k\rangle\langle E^{1/2} e_k|.$$

Discard zero columns, apply (22) to the refinement, then forget the refinement index. This avoids choices of eigenvectors at degeneracies. Finite and countable sums converge absolutely because on the near set $q_i \le 4sw_i$ and $|f_i(Q)| \le (1 + r)w_i$ everywhere.

Standard Borel outcomes. Fix the outcome space $(\Omega, \mathcal{B})$ and equip its POVMs with the evaluation σ-algebra generated by the matrix entries of $E(B)$ for $B \in \mathcal{B}$. For a POVM $E(B)$ take the finite measure $\sigma(B) = \operatorname{tr} E(B)$. Its total mass is d. In finite dimension there is a measurable positive matrix density $H(z) = dE/d\sigma$, with trace one σ-almost everywhere. Refine its square-root columns as above. On pairs (z, k) let $d\omega = \|H(z)^{1/2} e_k\|^2 d\sigma(z)$ and y be the normalized nonzero column. Then $\int P_y \, d\omega = I$ and $\omega(\text{all}) = d$. Formula (22) applies to densities relative to ω, replacing sums over N, F by integrals. Boundedness allows integration over any outcome event and interchange with tape expectation. The resulting kernel has total mass one and its projection onto z has the prescribed law. On a fixed standard Borel output space, a jointly measurable density can also be chosen as the limit of ratios $E(C)/\sigma(C)$ on refining finite generating partitions, taking a default value on zero-mass or nonconvergent points. This gives the usual kernel interpretation for varying measurement inputs.

Mixed states. Sample a pure state from the normalized columns of $\rho^{1/2}$ with probabilities equal to their squared norms. Their mean projector is ρ. One independent uniform implements this choice, and the pure-state alphabet envelope covers all selections.

Cost. With N_d net centres the tuple has at most $C_d L(\lambda)$ slots, where $C_d = 2N_d(1 + 2m)$. Extra uniforms never increase this count. A maximal δ-separated net can be chosen with $N_d \le (2/\delta)^{2(d-1)}$ since disjoint projector-distance balls of radius $\delta/2$ have Haar mass $(\delta/2)^{2(d-1)}$. The only use of this estimate is finiteness. The overhead in log alphabet size is additive at fixed d.

This proves Theorem 24. In particular, restriction transfers any attainable upper-tail exponent from general measurements to binary tests, and the reduction transfers it in the opposite direction. This equivalence holds for every $d \geq 2$, even before the binary exponent is evaluated.

Proof of Theorem 4. For the forward construction, the bounds in Lemmas 19, 20 and 22 apply to the declared menus. In a fine branch declare all $J\lceil L/n \rceil$ slots, bounded by $M_d J s^{-(d-2)}$. At a coarse branch declare all indices up to Z (or one label at a default branch). The independent outer tape has its own bound Z'. These common finite envelopes are tape-measurable. The tail integration in Lemma 23 applies to these declared bounds and gives $\Pr(L_{\mathrm{bin}} > m) = O_d(m^{-1/(d-2)})$. Theorem 24 preserves this rate under multiplication by C_d. It also gives finite moments below that exponent.

Conversely, restrict any all-measurement protocol to pure states and binary rank-one tests. Its declared alphabet dominates the used range on the restricted experiment. Theorem 3 makes the endpoint moment infinite. If an exponent strictly above $1/(d-2)$ were attained, Lemma 5 would give a finite endpoint moment, a contradiction. The binary construction has a two-sided tail for its used alphabet. For general measurements the proof gives an attainable upper tail and a universal moment obstruction. □

9 Sharpness of the analytic exponent

Corollary 25 (Maximum strong output exponent)**.** *Fix $d \geq 4$ and $a = d - 2$. There is no finite constant M and exponent $a < p \leq \infty$ for which*

$$\|T_t f\|_p \leq M \|f\|_\infty \quad \text{for every bounded } f \text{ and every } 0 < t \leq 1/2.$$

The same assertion holds on any smaller fixed interval $0 < t \leq t_0$ with $t_0 > 0$. For finite $p > a$, such a bound would force $\mathbb{E}D_+^{1/p} = \infty$ for every exact protocol, contradicting the protocol of Theorem 3.

Proof. Assume the bound with finite $p > a$. On tapes from Lemma 11, each positive-mass cell satisfies

$$0 \leq a_i(t)/t = \langle \mathbf{1}_{A_i}, T_t b_i \rangle \leq M \mu(A_i)^{1-1/p}.$$

If $D_+ < \infty$, summing gives $F_\lambda(t)/t \leq M D_+^{1/p}$. The excluded-cell limit remains available from Lemma 12. If its dominating moment were finite, dominated convergence would contradict exactness as before. On the other hand $1/p < 1/a$, and the constructed tail gives

$$\mathbb{E}D^{1/p} = \frac{1}{p} \int_0^\infty u^{1/p-1} \Pr(D > u)\, \mathrm{d}u < \infty$$

by Lemma 5. This is a contradiction since $D_+ \leq D$. A uniform L^∞ output bound implies every finite L^p bound on a probability space, so it too is excluded. For a shorter interval choose the sequence approaching zero inside that interval. □

9.1 An explicit family above the endpoint

The following independent proof of divergence for finite $p > d - 2$ also specifies a measurable event for the numerical tests. Fix n centres $y_1, \ldots, y_n$. Let f be the indicator of the union of bands $\{z : t(y_j, z) \leq s\}$, let $t = 5s \leq 1/2$, and put $h = s^{d-1}$. Sample a uniform label j, then X uniformly in $C_s(y_j)$, and finally Z uniformly on the fidelity-t sphere about X. Define m_* as the probability that X lies in exactly one centre cap and Z misses every band. The tagged sampling density is the number of covering caps divided by nh. On the uniquely covered region A, $R_0 f = 1$ and $e = 1 - R_t f \geq 0$. Hence $\int_A e \, d\mu = nhm_*$, while $\mu(A) \leq nh$. For $p \geq 1$, Hölder yields the fixed-decoder inequality

$$\|T_t f\|_p^p \geq nh(m_*/t)^p. \tag{23}$$

For independent Haar centres, the tagged band is escaped because $t > 4s(1-s)$ (Lemma 16). Each remaining centre must avoid both the cap about X, of mass h, and the band about Z, of mass $b = 1-(1-s)^{d-1}$. These sets are disjoint by the same escape inequality. Conditional on the tagged centre, X, Z, the other centres remain independent. Thus $\mathbb{E}_{\text{centres}} m_* = (1-h-b)^{n-1}$. Jensen applied to (23) gives an expected lower bound $nht^{-p}(1-h-b)^{p(n-1)}$. For $s = \kappa/[(d-1)n]$ with $\kappa > 0$ fixed, the bound is asymptotic to

$$\frac{\kappa e^{-p\kappa}}{(d-1)5^p}\, s^{d-2-p}.$$

At least one realized union at each sufficiently large n attains the expected lower bound. This supplies explicit finite families that rule out a uniform estimate above the endpoint. The endpoint upper estimate itself comes from Theorem 6.

10 Comparison and dimension dependence

Castellano, Sekatski and Brunner [2, Definition 1, Eq. (12) and Sec. V] give an all-measurement construction with tail exponent $1/(2d-2)$ and exclude finite expected alphabet size for $d \geq 4$. Their binary obstruction already applies to pure preparations and rank-one projective tests. Their bounds for the optimal all-measurement tail exponent were $1/(2d-2) \leq \beta^{\text{tail}}_{d,\text{all}} \leq 1$. Theorems 3 and 24 supply matching bounds under the same input-independent shared-randomness convention. In dimension four the earlier interval was $[1/6, 1]$ and the exponent obtained here is $1/2$. These results concern tail decay at fixed d. The optimal dependence of the expected bit cost on d remains open.

Comparison with the qutrit cap protocol. De Gois, Santos and Vieira [4, App. D] already realize qutrit Born probabilities by coupling cap thresholds to a shared Haar ray and conditioning on Alice's acceptance. Their finite binary simulator truncates the search and uses a covering-ball fallback. In Appendix D.3 they simulate general qutrit measurements using 26 independent binary copies, rejection and a multi-outcome fallback. Here Lemmas 14–15 construct an exact scale mixture in $d \geq 4$ with $\nu_d((0,s]) \sim (d-3)s/(d-1)$. Grouping the labels at scale s then gives the matching tail in Lemma 23. Theorem 24 uses one binary message on each tape. Its normalized decoder gives a declared alphabet at most $C_d L(\lambda)$ for *any* exact binary simulator, preserving its tail exponent at fixed d.

d	Optimal tail exponent	Basis
2	∞	Bounded alphabet for arbitrary measurements [3].
3	∞	Bounded alphabet for arbitrary measurements [4].
4	$1/2$	The ququart estimates of [4], summed with cell masses in Lemma 13, and the matching construction with Theorem 24.
≥ 5	$1/(d-2)$	The strong L^{d-2} estimate, the binary construction and the finite-overhead reduction.

There is a direct reason the positive cap recursion cannot be extended down to $d = 3$ using the same proposed moments. Its first two moments would be $Q_3(1) = 24/25$ and $Q_3(2) = 32/35$, so $1 - 2Q_3(1) + Q_3(2) = -1/175$. A positive law on $[0,1]$ must have $\mathbb{E}(1-U)^2 \geq 0$. The low-dimensional bounded protocols therefore use a different construction.

The sphere kernel $(u \cdot v)^2$ for real unit vectors $u, v \in \mathbb{R}^5$ is obtained by restricting the complex dimension-five experiment to real inputs. The upper construction restricts to those inputs, but the Haar-cell converse integrates over the full complex projective space. The real projective locus has zero complex Haar measure, as follows from its smaller real dimension in coordinate charts. Consequently the present lower bound does not decide the finite-alphabet question for the real five-dimensional square or its exceptional polynomial-contact case. The common squared-overlap formula alone does not transfer the complex lower bound to that restricted domain.

11 Computational checks

The experiments have three distinct targets. Exact arithmetic detects a wrong multiplier normalization and a missing term in the general-measurement decoder. Quadrature and sampling test the cap response against the Born probability at specified fidelities. Realized finite decoders give lower confidence bounds on their own operator norms. The lower bound over all exact protocols follows from Section 4. None of the sampling routines enumerates protocols.

11.1 Exact arithmetic and normalized multipliers

In `check_dimension.py`, exact arithmetic confirms the dimension-factor transform for $d = 4, \dots, 30$ and Mellin arguments $-2, \dots, 24$. It also checks the Jacobi-product cancellation for $d = 4, \dots, 16$ and degrees $1, \dots, 12$, the dyadic sum and the degree-one quotient. The floating-point multiplier grid covers $d = 4, \dots, 12$, degrees $1, \dots, 400$, and 110 positive fidelities including a repeated join at 0.01. It also checks the endpoint slopes. The largest ratio of an observed quotient to the bound in Lemma 8 was 0.062499224575.

The normalization audit in `normalization_checks` verifies $m_1(n) = 1$, the reproducing-kernel dimension, and the dimension-four endpoint slope using exact rational arithmetic. The routine `diagonal_scaling_checks` independently evaluates the terminating Jacobi series with 80 decimal digits at $t = n^{-2}$ for $n = 16, 64, 256, 1024, 4096$. In dimension five its largest tested degree gives $|q_{1/n^2}(n)| = 0.0011360387157$, consistent with the proved n^{-1} scaling. For a direct norm test, take $f_n([z]) = (2z_1\overline{z}_2)^n$ on unit representatives. Its complex Euclidean Laplacian is zero, so $f_n \in \mathcal{H}_n$, and $\|f_n\|_\infty = 1$. The simplex law of the squared moduli gives, for $r > 0$,

$$\|f_n\|_r^r = \frac{2^{rn}\Gamma(d)\Gamma(rn/2+1)^2}{\Gamma(rn+d)}.$$

The routine `harmonic_endpoint_test` evaluates both these norms and the multiplier. For $d = 5$, the ratio $\|T_{1/n^2} f_n\|_3^3/(\|f_n\|_\infty \|f_n\|_2^2)$ decreases from 0.003733516007646 at $n = 16$ to $3.548397241413 \times 10^{-10}$ at $n = 4096$. These values omit the fixed factor $\mathcal{C}_5$ on the right-hand side. The test uses the norm ratio directly. For every nonzero $f_n \in \mathcal{H}_n$, the elementary inequality $\|f_n\|_3^3 \le \|f_n\|_\infty \|f_n\|_2^2$ bounds this ratio by $|q_t(n)|^3$.

The only specialized approximation input to the analytic proof is the Wang–Sloan filter bound in Lemma 9. Lemma 12 proves the required Sobolev locality on arbitrary measurable cells.

11.2 Finite CPU experiments

The reference panel contains three seeded waves in dimensions 4 through 8. Each wave has five cap-mixture cells, 25 finite-decoder cells and one cell for four packet configurations, giving 93 scheduled cells in total. Seeds, interval allocations and raw counts are included in the embedded source.

Born probabilities. Importance sampling of the multiplicative dimension factors, followed by rejection sampling of the mixing law in Lemma 15, produced 6,943,070 accepted cap anchors from 62,914,560 proposals. For each anchor the code tested nine fidelities, $0, 1/1024, 1/64, 1/8, 0.37, 1/2, 0.81, 63/64, 1$, by directly computing the overlap with the test ray. Thus 62,487,630 bits were evaluated, with nine correlated bits per anchor. The cap law takes $1 - u = sV^{1/(d-1)}$ for a uniform V. The squared projection on a chosen transverse direction is independently $\mathrm{Beta}(1, d-2)$. These and a uniform phase determine the overlap without using the target probability as an output coin.

d	Accepted anchors	Max. $\lvert\widehat{p}-t\rvert$	Finest-scale $p=d-1$ lower bound
4	3,146,183	0.00086744	0.872750
5	1,834,756	0.00095795	0.106409
6	1,046,648	0.00181094	0.0126192
7	588,901	0.00219489	0.00155111
8	326,582	0.00164808	0.000185654

The largest residual was at $d=7$, $t=1/2$ in wave 3. There were 97,638 successes in 196,137 trials, giving $\widehat{p}=0.4978051$ and a simultaneous interval $[0.492764, 0.502846]$. The intervals are two-sided Clopper–Pearson intervals. Zero-indexed wave w receives error budget $0.01/[(w+1)(w+2)]$, split among all its intervals, including normalization and small-scale mass checks. The union bound permits dependence across fidelities. The budgets sum to 0.01 even if later waves are added. The core required each Born interval to contain t and fit within $t \pm 0.012$. A wrong-threshold control had to produce false accepts at $t=0$. The scale normalization and all scheduled checks met their criteria.

As an independent $d=4$ check, direct nested quadrature evaluates $\int K_s(t)\, d\nu_4(s)$ on eight fidelities. Its maximum absolute residual in the reference run was 2.38×10^{-8} (rounded up). All 15 diagnostics in the quadrature validation module passed. These fixed-fidelity computations are separate from the random cap samples.

Finite decoders and packets. At each dimension the core used $n = 2, 8, 32, 128, 512$ bands. Each of 75 saved decoders received 131,072 independent tagged trials conditional on its centres. Across all decoders this required 2,681,733,120 pairwise centre fidelities. Substituting the lower limit for m_* into (23) gives a lower confidence limit for the full-space p-th power of the norm of that particular decoder. The last column reports the smallest such bound across the three 512-centre decoders in each dimension. Separately, 3,145,728 finite population trials checked the hypergeometric packet correction of Lemma 19, including a short packet, a singleton and the boundary $p_g = c$. The source also includes an optional GPU runner. The reported measurements used a CPU.

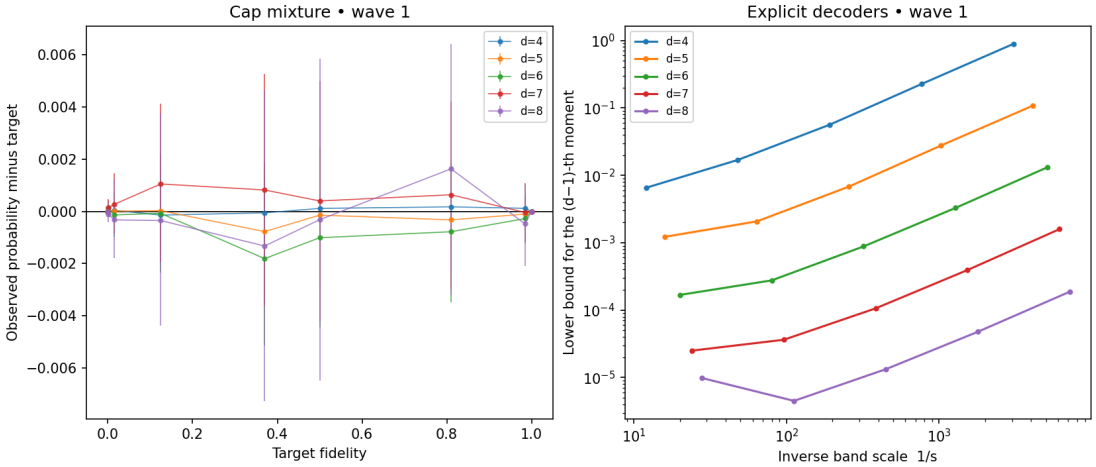


Figure 1: First of the three prespecified CPU waves. The left panel shows deviation of the simulated cap response from its target at each fidelity with simultaneous binomial intervals. The right panel gives lower confidence limits from (23) for archived unions of bands at $p=d-1$. The sequence of finite witnesses illustrates the construction in Section 9.1.

Arbitrary-measurement decoder. The twelve-outcome rational fixture enumerates 18 positive-probability tape blocks at each of two scales, the original $s=1/256$ and the theorem's $s=1/(32d)=$

1/128 at $d = 4$. On every block it forms both full vectors in exact arithmetic. The correct vector uses $f_j(Q) + R_j(T - S)$ on far outcomes. The faulty vector uses $f_j(Q) + R_j T$ there and keeps the near entries unchanged. At each scale the correct vector is nonnegative, has unit mass on all 18 blocks and reproduces all twelve Born probabilities after averaging. The faulty vector fails normalization on six of the 18 blocks at each scale. The released JSON records both measured totals on every block. A second suite tests 720 frames in dimensions $4, 5, 7, 8$ and 6,480 fixed-tape vectors. The maximum normalization error was 1.1103×10^{-15} (rounded up) and the maximum deviation of the computed mean identity was 4.9961×10^{-16} (rounded up). The minimum output component was zero. In 1,071 vertex–measurement pairs a raw score before correction was negative. These checks exercise the algebra of the reduction using an exactly specified binary-response fixture. The optional geometric simulator in `ququart_lab/protocol.py` replaces the deterministic covering net of Lemma 17 with finitely many random centres and a finite round limit. A random list may miss a cap. Its counts refer only to sampled inputs. The 93-cell panel tests components of the construction. Neither experiment measures the distribution of L over complete tapes.

12 Conclusion

For $d \geq 4$, a classical one-way simulator of arbitrary quantum measurements can attain alphabet-tail exponent $1/(d-2)$, and every exact simulator has an infinite endpoint moment. The matching protocol is constructed in Sections 5–7. At $d = 4$ the obstruction follows by summing the ququart estimates of [4] with cell masses. For $d \geq 5$ it uses the strong L^{d-2} bound of Theorem 6. Section 8 transfers the exponent to arbitrary measurements. The optimal constants and the dependence of the optimal mean bit cost on dimension remain open.

References

[1] H. Wang and I. H. Sloan, *On filtered polynomial approximation on the sphere*, arXiv:1509.03792, Lemma 2.1.

[2] R. Castellano, P. Sekatski and N. Brunner, The cost of classically simulating quantum communication is as high as it could be, arXiv:2610.19060v1, Definition 1, Eq. (12), Sec. V and Appendix A.

[3] M. J. Renner, A. Tavakoli and M. T. Quintino, Classical cost of transmitting a qubit, Phys. Rev. Lett. 130, 120801 (2023), doi:10.1103/PhysRevLett.130.120801, arXiv:2207.02244.

[4] C. de Gois, T. S. R. Santos and C. Vieira, Quantum communication and Bell nonlocality require infinite classical communication to simulate, arXiv:2609.04182v1, Apps. B.5–B.6 and D.

[5] NIST Digital Library of Mathematical Functions, *Symmetry, special values, and limits to monomials*, Table 18.6.1.

A Source code and reproduction

The research archive `Kashogi_Source.zip` is embedded in the PDF and available directly from the Kashogi.com research page. The public download address is `https://kashogi.com/research/critical-alphabet/source.zip`. It contains this manuscript's LaTeX source, the complete Python checks, 93 raw result cells, the 75 saved decoder arrays, the sampling plan and a CPU replay log. After installing the pinned dependencies, the checks run locally. The reference environment used Python 3.12.14, NumPy 2.3.5, SciPy 1.17.0 and Matplotlib 3.10.8. The verification script first checks the files listed in `SHA256SUMS`.
With Poppler and `unzip` installed, extract the archive in the directory containing the PDF.

```
pdfdetach -saveall Kashogi_Critical_Alphabet.pdf
unzip Kashogi_Source.zip
cd Critical_Alphabet
```

Implementation	Formula or claim to audit
`check_dimension.py`	`cap_moments` and `cancellation` check dimension recursion and finite Jacobi products. `endpoint_sum`, `degree_one_check` and `multiplier_checks` check the dyadic sum and multiplier bound.
`normalization_checks` in the same file	Exact tests of (8) at $t = 1$, the reproducing-kernel dimension, and the $d = 4$ slope. At $d = 5, n = 2$, dividing $P_2^{(3,0)}(1) = 10$ by $\binom{5}{3}$ gives 1. Dividing by $n + a$ gives 2. The code computes and asserts both results.
`diagonal_scaling_checks` and `harmonic_endpoint_test`	Independent terminating series at $t = n^{-2}$ and Dirichlet-integral harmonic norms.
`check_povm_lift.py`	Exact rational and floating-point fixtures for (21) and (22), including a faulty decoder control.
`ququart_lab/evidence.py` and `campaign.py` in that directory	Scale sampler, direct overlap bits, finite packet trials, (23), deterministic seed mapping and interval allocation.
`reference/cells/` and `verify.py`	The driver reconstructs totals, intervals and lower bounds from raw integer counts and hashed centre matrices.

With Python 3.12, run the checks from the extracted directory.

```
python3 -m venv .venv
. .venv/bin/activate
python -m pip install -r requirements-pinned.txt
python verify.py
```

For a fresh CPU simulation of all 93 cells on Linux, run `python verify.py --full`. It compares each new result against the archived cell, excluding elapsed time. Exact seeded replay targets the stated software environment. Floating-point residuals may vary on other systems. To inspect just the endpoint normalization, run

```
python -c "from check_dimension import normalization_checks as f; f()"
```

Keep Python assertions enabled. A replay with the same code and seeds matched all 93 archived cells, excluding elapsed times. See `reference/replay.log` for the record and `README.md` for the plain-text formulas, build instructions and optional GPU commands.

Preparation note. An AI assistant contributed to the development of the proofs, the manuscript and the numerical code.