

Exact Rate–Distortion Theory for Complex-Projective Born Prediction: Finite-Dimensional Bingham Frontiers, Thermodynamic Coexistence, and Worst-State Capacity

Lluis Eriksson

August 11, 2026

Abstract

We give a unified exact rate–distortion theory for retaining the complete rank-one Born-probability field of a Haar-random pure state in $\mathbb{CP}^{d-1}$. First, a constrained root count, a sharp centered power-sum theorem, and Newton’s expansion prove an all-degree spectral result: at fixed traceless Frobenius norm the projective Laplace transform is maximized by a one-positive-spike spectrum. This reduces the unrestricted Shannon rate–distortion function to an exact scalar complex-Bingham envelope in every dimension, with covariant channels and independently flagged coexistence mixtures attaining every distortion. For $d \geq 3$ directional information turns on discontinuously; the qubit curve and the all-dimensional high-fidelity constant are explicit.

We then solve the full fixed-normalized-distortion limit $d \rightarrow \infty$. If $y_* > 2$ solves $y_* - 1 = 2 \log y_*$, then the limiting free energy has a unique coexistence point with $\alpha_* = y_*^2/[2(y_* - 1)] = 2.455407482\dots$ and $b_* = 1 - y_*^{-1} = 0.715331863\dots$. The information cost per dimension is

$$r_\infty(\delta) = \begin{cases} -\log(1 - \sqrt{1 - \delta}), & 0 < \delta \leq 1 - b_*^2, \\ \alpha_*(1 - \delta), & 1 - b_*^2 \leq \delta \leq 1, \end{cases}$$

and the finite-dimensional onset multiplier obeys $\lambda_{c,d} = \alpha_* d - [\alpha_*/(2 \log y_*)] \log d + O(1)$.

Finally, a pointwise Hilbert projection converts every measurable scalar reporter into a physical density-matrix reporter without increasing risk for any pure input. Hence the exact worst-state channel capacity equals the Haar rate–distortion function, arbitrary joint n -state memories cost exactly $n\mathcal{R}_d(D)$, and transitivity forces zero rate dispersion with an exponential finite-blocklength strong converse. These results concern reusable calibrated probability fields, not click-only simulation, state update, or intrinsic randomness.

1 Introduction

Let $x \in \mathbb{CP}^{d-1}$ label a pure state $\rho_x = |x\rangle\langle x|$. A classical representation of x is asked for the Born probability

$$p_x(y) = \text{tr}(\rho_x \rho_y) \tag{1}$$

of every rank-one query y . The question is not how many outcomes a single measurement can generate, but how much classical information is required to retain an entire calibrated probability field to prescribed accuracy.

For a Haar source, integrated squared calibration loss reduces exactly to squared Frobenius reconstruction of a Haar rank-one projector. The resulting complex-projective Shannon rate–distortion problem has two linked layers. At finite d , the reproduction alphabet is the full

density-matrix body and must first be reduced globally to a scalar complex-Bingham envelope. We prove that reduction in all dimensions and show that the first directional phase is discontinuous for every $d \geq 3$. At large d , the active Bingham field is of order d , so small-parameter expansions do not apply directly [1]. The second layer is the global thermodynamic optimization of that exact finite-dimensional envelope.

Large-parameter Kummer asymptotics themselves are classical [2, 3]; our contribution begins with their global radius optimization, convex conjugate, coexistence law, and finite-size shift. A uniform Laplace principle produces the limiting free energy with coexistence root $y_* - 1 = 2 \log y_*$. General metric-space and complex-Grassmann results provide bounds or high-resolution dimensions, while direct complex-projective and Grassmann quantization treats finite codebooks or metric-ball asymptotics [4, 5, 6, 7]; these works do not evaluate the stochastic all-distortion radial envelope. The exact real-sphere analysis of Dytso and Cardone is the closest symmetry-reduction template, but its Bessel/von-Mises–Fisher partition and full-sphere orbit do not reduce to the Beta/Kummer projective orbit [8].

The centered-projector map embeds $\mathbb{CP}^{d-1}$ as a proper $U(d)$ orbit of real dimension $2d - 2$ in the $(d^2 - 1)$ -dimensional traceless-Hermitian space. Its invariant overlap has law Beta(1, $d - 1$), not the symmetric coordinate law of a uniform real sphere. Thus the real-sphere RDF cannot be imported by setting the sphere dimension to $2d - 2$; in particular its limiting convex geometry has no counterpart of the projective coexistence face.

We also strengthen the operational scope. We call the resulting worst-input channel-capacity frontier “source-universal” only in this defined sense; it is not a claim of a single finite-block universal code for an unknown source distribution. Compound-source, invariant, and minimax rate–distortion theories have a substantial prior literature [9, 10, 11]. The new content here is the pointwise arbitrary-reporter projection and exact complex-projective evaluation. The same argument gives an exact equality for arbitrary joint memories. Finally, the invariant optimal channel has constant D -tilted information, so general nonasymptotic lossy-source converses specialize to a zero-dispersion exponential bound [12].

Our contributions are:

1. an all-degree one-spike projective partition theorem obtained from an exactly identified centered power-sum input, a global constrained root count, parity closure, and Newton’s expansion;
2. the exact unrestricted finite-dimensional RDF, attained by covariant complex-Bingham channels and flagged coexistence mixtures;
3. a discontinuous finite-dimensional onset for every $d \geq 3$, together with the qubit specialization and exact high-fidelity constant;
4. an exact thermodynamic free energy, a closed limiting RDF at every fixed normalized distortion, and dimension-independent coexistence constants α_* , b_* , δ_* ;
5. convergence of first contact with the finite-size correction $\lambda_{c,d} = \alpha_* d - c_* \log d + O(1)$;
6. a pointwise projection of arbitrary probability reporters onto physical density reports, valid for Haar and exact projective 2-design queries;
7. an exact worst-state channel-capacity equality and its finite-block tensorization; and
8. zero rate dispersion and an exponential strong converse.

We do not claim a unique positive branch at every finite d , a matching third-order coding theorem, or a rank- r Grassmann extension. The last would require a new global matrix-Bingham/HCIZ spectral extremum.

This manuscript supersedes the earlier unpublished drafts entitled *Exact Information–Accuracy Frontiers for Born-Probability Prediction: Complex-Bingham Rate–Distortion and a Discontinuous Directional-Information Onset* and *Thermodynamic Limit of Complex-Projective Born-Prediction Rate–Distortion: Exact Coexistence and a Logarithmic Onset Shift*. Their finite-dimensional and thermodynamic results are incorporated here into one self-contained article; they are not intended as separate overlapping publications.

2 Finite-dimensional starting point

Let X be Haar on $\mathbb{CP}^{d-1}$ and set

$$S_X = \rho_X - I/d, \quad R_{0,d}^2 = \|S_X\|_F^2 = 1 - \frac{1}{d}. \quad (2)$$

For a density-matrix reproduction $\hat{\rho}$, use distortion $\|\rho_X - \hat{\rho}\|_F^2$. Denote the unrestricted classical rate–distortion function by

$$\mathcal{R}_d(D) = \inf I(X; \hat{\rho}), \quad \mathbb{E}\|\rho_X - \hat{\rho}\|_F^2 \leq D. \quad (3)$$

The infimum ranges over arbitrary standard-Borel stochastic reproductions.

Let

$$M_d(s) = {}_1F_1(1; d; s) = (d-1) \int_0^1 e^{st} (1-t)^{d-2} dt, \quad K_d(s) = \log M_d(s) - \frac{s}{d}. \quad (4)$$

The complex-Bingham and complex-Watson families and their normalizers are established directional-statistics objects [13, 14]. Here K_d is the cumulant generating function of $T_d - 1/d$ for $T_d \sim \text{Beta}(1, d-1)$. Define

$$G_{d,\lambda}(b) = K_d(2\lambda b) - \lambda R_{0,d}^2 b^2, \quad 0 \leq b \leq 1. \quad (5)$$

It is convenient to denote the optimized Gibbs log partition by

$$c_d(\lambda) = -\lambda R_{0,d}^2 + \max_{0 \leq b \leq 1} G_{d,\lambda}(b). \quad (6)$$

We use the following exact finite-dimensional theorem as the starting point. Its nontrivial spectral input is an all-degree Dirichlet partition extremum: among traceless Hermitian matrices of fixed Frobenius norm, the projective Laplace transform is maximized by one positive eigenvalue and $d-1$ equal negative eigenvalues [15].

Theorem 2.1 (Finite-dimensional projective envelope). *For $0 \leq D \leq R_{0,d}^2$,*

$$\mathcal{R}_d(D) = \sup_{\lambda \geq 0} \left\{ \lambda(R_{0,d}^2 - D) - \max_{0 \leq b \leq 1} G_{d,\lambda}(b) \right\}. \quad (7)$$

Every positive active radius is attained by a covariant complex-Bingham channel whose posterior mean is $I/d + b(\rho_u - I/d)$. Independent revealed mixtures of active radii attain every nonexposed distortion.

Appendix A proves the spectral reduction and the complete Gibbs equality mechanism needed by Theorem 2.1, including the normalization interface to the external centered power-sum theorem.

3 Finite-dimensional phase structure and endpoint regimes

The exact envelope already implies a qualitative transition that is absent in the qubit case. The active radii are precisely the maximizers of $G_{d,\lambda}$.

Theorem 3.1 (Discontinuous finite-dimensional onset). *For every $d \geq 3$, there exists*

$$0 < \lambda_{c,d} < \lambda_{0,d} := \frac{d(d+1)}{2} \quad (8)$$

such that $b = 0$ and at least one $b_{c,d} > 0$ are global maximizers of $G_{d,\lambda_{c,d}}$. The positive maximizer cannot approach zero at first contact. Consequently the exact frontier has a nontrivial linear face adjacent to $(D, R) = (R_{0,d}^2, 0)$, and optimal directional information turns on discontinuously.

Proof. Put $W = T_d - 1/d$. It is centered and lies in an interval of length one. Hoeffding's lemma gives $K_d(s) \leq s^2/8$ for every real s . Hence, uniformly in $b \in [0, 1]$,

$$G_{d,\lambda}(b) \leq \lambda b^2 \left(\frac{\lambda}{2} - R_{0,d}^2 \right) < 0 \quad (b > 0, 0 < \lambda < 2R_{0,d}^2). \quad (9)$$

Thus $b = 0$ is initially the unique global maximizer and $c_d(\lambda) = -\lambda R_{0,d}^2$ on a nonzero interval.

The variance and third centered moment of $T_d \sim \text{Beta}(1, d-1)$ are

$$v_d = \frac{d-1}{d^2(d+1)}, \quad \mu_{3,d} = \frac{2(d-1)(d-2)}{d^3(d+1)(d+2)}. \quad (10)$$

Expansion at $b = 0$ gives

$$G_{d,\lambda}(b) = (-\lambda R_{0,d}^2 + 2\lambda^2 v_d) b^2 + \frac{4}{3} \lambda^3 \mu_{3,d} b^3 + O(b^4). \quad (11)$$

The quadratic coefficient vanishes at $\lambda_{0,d} = d(d+1)/2$. For $d \geq 3$, $\mu_{3,d} > 0$, so a sufficiently small positive b has positive gain at $\lambda_{0,d}$, and the same remains true just below it.

Let $\lambda_{c,d}$ be the infimum of multipliers whose global argmax contains a positive radius. Estimate (9) makes it positive, and the preceding expansion makes it strictly smaller than $\lambda_{0,d}$. Choose positive active radii b_n at $\lambda_n \downarrow \lambda_{c,d}$. Since $K_d(s)/s^2 \rightarrow v_d/2$ as $s \rightarrow 0$ and $\lambda_{c,d} < \lambda_{0,d}$, there exist $\varepsilon, \eta > 0$ such that $G_{d,\lambda}(b) \leq -\eta b^2$ for $0 < b \leq \varepsilon$ and λ near $\lambda_{c,d}$. Thus every b_n is at least ε . Compactness and continuity yield a limiting $b_{c,d} \geq \varepsilon$ coexisting globally with 0. Finally, Danskin's relation (112) and an independently revealed mixture of the two active channels give the linear face. $\square$

Remark 3.2. The theorem proves first contact and its nonzero jump. It does not assert uniqueness of the positive maximizer for every larger multiplier or exclude a later branch exchange. The exact envelope and flagged mixtures already cover such possibilities.

For $d = 2$, the Bloch representation maps centered projectors to the real unit two-sphere up to a factor $1/2$. The exact result is therefore the $n = 3$ real-sphere theorem of Dytso and Cardone [8]. Put $r = \sqrt{1 - 2D}$ and choose $\eta \geq 0$ from

$$r = \coth \eta - \frac{1}{\eta}. \quad (12)$$

For $0 < D < 1/2$, such an η exists uniquely and

$$\mathcal{R}_2(D) = \eta r - \log \frac{\sinh \eta}{\eta}. \quad (13)$$

Moreover $\mathcal{R}_2(0) = +\infty$ by the limit $\eta \rightarrow \infty$, while $\mathcal{R}_2(D) = 0$ for $D \geq 1/2$. This is a consistency check and operational specialization, not a new real-sphere theorem.

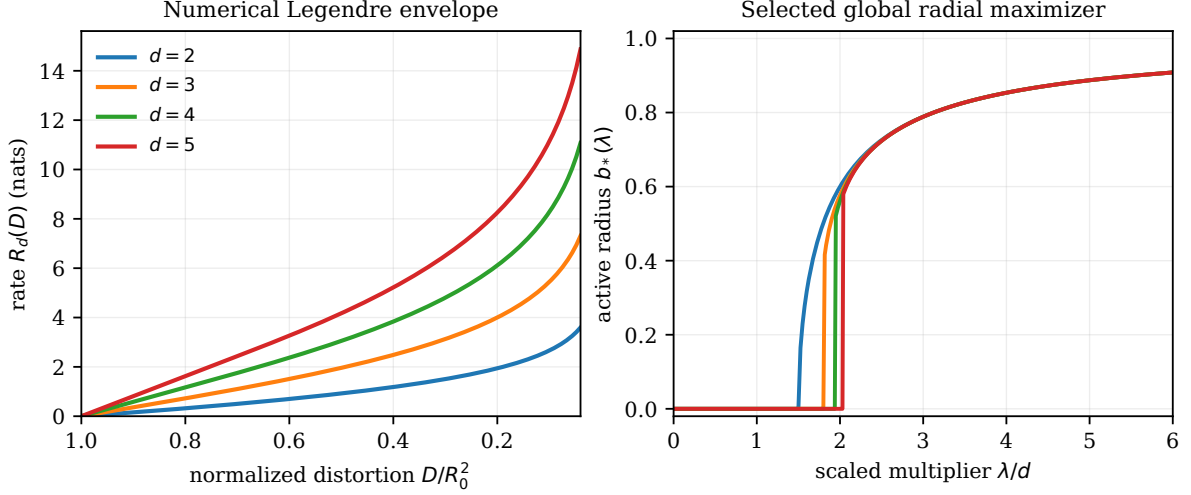


Figure 1: Numerical evaluation of the exact finite-dimensional scalar envelope for $d = 2, 3, 4, 5$. Left: $\mathcal{R}_d(D)$ versus normalized distortion. Right: one selected globally active radius versus scaled multiplier. The qubit radius emerges continuously, while the higher-dimensional curves show the finite jump proved in Theorem 3.1. The bounded scalar computation is a reproducibility diagnostic, not a proof input.

Proposition 3.3 (Exact fixed-dimensional high-fidelity asymptotic). *For every fixed $d \geq 2$,*

$$\mathcal{R}_d(D) = (d-1) \log \frac{2(d-1)}{D} - (d-1) - \log \Gamma(d) + o(1), \quad D \downarrow 0. \quad (14)$$

Proof. Put $m = d - 1$. For integer d ,

$$M_d(s) = \Gamma(d) s^{-m} \left(e^s - \sum_{j=0}^{m-1} \frac{s^j}{j!} \right). \quad (15)$$

For $s > 0$ let $\ell_d(s) = \log(1 - e^{-s} \sum_{j=0}^{m-1} s^j/j!) \leq 0$. For a traceless Hermitian report A , define

$$Q_\lambda(A) = \int \exp\{-\lambda \|S_x - A\|_F^2\} d\mu_d(x). \quad (16)$$

For the axial reproduction $A_{b,u} = b(\rho_u - I/d)$,

$$\log Q_\lambda(A_{b,u}) = -m \log(2\lambda) + \log \Gamma(d) - \lambda R_{0,d}^2 (1-b)^2 - m \log b + \ell_d(2\lambda b). \quad (17)$$

Taking $b = 1$ gives a lower bound with vanishing error. For the upper bound, $K_d(s) \leq R_{0,d}^2 s$ for $s \geq 0$, so radii $b \leq 1/2$ are eventually excluded because their radial gain is at most $-\lambda R_{0,d}^2/4$. On $b \geq 1/2$, use $-\log b \leq 2(1-b)$ and $\ell_d \leq 0$ to bound the remaining correction by

$$-\lambda R_{0,d}^2 (1-b)^2 + 2m(1-b) \leq \frac{m^2}{\lambda R_{0,d}^2} = o(1). \quad (18)$$

Consequently

$$c_d(\lambda) = -m \log(2\lambda) + \log \Gamma(d) + o(1). \quad (19)$$

Secant inequalities for the convex function c_d imply $-\lambda g_\lambda \rightarrow m$ for every $g_\lambda \in \partial c_d(\lambda)$. If λ_D maximizes the exact conjugate, then $-D \in \partial c_d(\lambda_D)$ and $\lambda_D D \rightarrow m$. Therefore

$$\mathcal{R}_d(D) = -\lambda_D D - c_d(\lambda_D) \quad (20)$$

$$= -m + m \log(2\lambda_D) - \log \Gamma(d) + o(1) \quad (21)$$

$$= m \log \frac{2m}{D} - m - \log \Gamma(d) + o(1), \quad (22)$$

which proves (14). $\square$

The limits $d \rightarrow \infty$ at fixed normalized distortion and $D \downarrow 0$ at fixed d are distinct. The following sections resolve the first regime; a uniform double-scaling theory remains open.

4 Uniform projective Laplace principle

The transition occurs at λ of order d . We first isolate the corresponding asymptotic partition.

Lemma 4.1 (Uniform beta Laplace principle). *For $y \geq 0$, put*

$$F_d(y) = \frac{1}{d} \log M_d(dy). \quad (23)$$

Then $F_d \rightarrow \Phi$ locally uniformly on $[0, \infty)$, where

$$\Phi(y) = \begin{cases} 0, & 0 \leq y \leq 1, \\ y - 1 - \log y, & y > 1. \end{cases} \quad (24)$$

Proof. Rewrite the beta integral as

$$M_d(dy) = (d-1) \int_0^1 \exp\{dyt + (d-2)\log(1-t)\} dt. \quad (25)$$

For fixed $\varepsilon > 0$, the usual compact-interval Laplace bounds apply on $[0, 1 - \varepsilon]$. On the remaining interval the original factor $(1-t)^{d-2}$ supplies exponential control; letting ε decrease after $d \rightarrow \infty$ gives the pointwise limit

$$\sup_{0 \leq t < 1} \{yt + \log(1-t)\}. \quad (26)$$

The maximizer is $t = 0$ for $y \leq 1$ and $t = 1 - y^{-1}$ for $y > 1$, which yields (24). Moreover,

$$F'_d(y) = \mathbb{E}_{d,y} T_d \in [0, 1], \quad (27)$$

where $\mathbb{E}_{d,y}$ is expectation under exponential tilt dy . Hence the family is equi-Lipschitz. A finite-net argument upgrades pointwise convergence to local-uniform convergence because Φ is continuous. $\square$

For $\alpha \geq 0$, define the normalized optimized gain

$$g_d(\alpha) = \frac{1}{d} \max_{0 \leq b \leq 1} G_{d,\alpha d}(b). \quad (28)$$

Corollary 4.2 (Thermodynamic free-energy convergence). *The functions g_d converge locally uniformly to*

$$g(\alpha) = \max_{0 \leq b \leq 1} \{\Phi(2\alpha b) - \alpha b^2\}. \quad (29)$$

Proof. Lemma 4.1, $R_{0,d}^2 \rightarrow 1$, and $d^{-1}K_d(2\alpha db) = F_d(2\alpha b) - 2\alpha b/d$ give uniform convergence of the objective on compact α intervals and $b \in [0, 1]$. Taking a maximum preserves it. $\square$

5 Global coexistence and the dimension-independent constants

The limiting scalar problem is globally solvable.

Theorem 5.1 (Exact limiting phase diagram). *Let $y_* > 2$ be the unique nontrivial root of*

$$y_* - 1 = 2 \log y_*. \quad (30)$$

Define

$$\alpha_* = \frac{y_*^2}{2(y_* - 1)}, \quad b_* = 1 - \frac{1}{y_*}. \quad (31)$$

Then $g(\alpha) = 0$ for $0 \leq \alpha \leq \alpha_$. At α_* the global maximizers in (29) are exactly 0 and b_* . For $\alpha > \alpha_*$ the positive global maximizer is unique and equals*

$$b_+(\alpha) = 1 - \frac{1}{y_+(\alpha)}, \quad y_+(\alpha) = \alpha + \sqrt{\alpha(\alpha - 2)}. \quad (32)$$

Proof. Set $y = 2\alpha b$. The optimization becomes

$$\max_{0 \leq y \leq 2\alpha} \left\{ \Phi(y) - \frac{y^2}{4\alpha} \right\}. \quad (33)$$

For $y \leq 1$ the objective is nonpositive and vanishes only at $y = 0$. For $y > 1$, an interior stationary point obeys

$$1 - \frac{1}{y} - \frac{y}{2\alpha} = 0, \quad \alpha = \frac{y^2}{2(y - 1)}. \quad (34)$$

The two positive stationary points appear at $\alpha = 2$; the larger root is $y_+(\alpha)$ and is the local maximum. At any stationary point the gain is

$$\Phi(y) - \frac{y^2}{4\alpha} = \frac{y - 1}{2} - \log y. \quad (35)$$

The function $y - 1 - 2 \log y$ decreases on $(1, 2)$, increases on $(2, \infty)$, and has exactly the roots 1 and y_* . Thus the larger stationary point first ties the zero-radius solution at y_* . More explicitly, for $y > 1$,

$$\frac{d}{dy} \left(\Phi(y) - \frac{y^2}{4\alpha} \right) = -\frac{(y - y_-)(y - y_+)}{2\alpha y}. \quad (36)$$

For $\alpha < 2$ the derivative is strictly negative. For $\alpha \geq 2$, y_- is a local minimum and y_+ is the only local maximum; the objective decreases after y_+ , so the endpoint $y = 2\alpha$ cannot improve it. Equation (35) then decides exactly when that maximum beats zero and proves the complete statement. $\square$

Numerically,

$$\begin{aligned} y_* &= 3.512862417252\dots, \\ \alpha_* &= 2.455407482284\dots, \\ b_* &= 0.715331862959\dots, \\ \delta_* &:= 1 - b_*^2 = 0.488300325835\dots \end{aligned} \quad (37)$$

6 Closed thermodynamic information–accuracy frontier

We normalize distortion by the zero-information value $R_{0,d}^2$.

Theorem 6.1 (Thermodynamic Born-prediction frontier). *For each fixed $0 < \delta \leq 1$,*

$$\lim_{d \rightarrow \infty} \frac{1}{d} \mathcal{R}_d(\delta R_{0,d}^2) = r_\infty(\delta), \quad (38)$$

where

$$r_\infty(\delta) = \begin{cases} -\log(1 - \sqrt{1 - \delta}), & 0 < \delta \leq \delta_*, \\ \alpha_*(1 - \delta), & \delta_* \leq \delta \leq 1. \end{cases} \quad (39)$$

The branches agree at δ_* , where $r_\infty(\delta_*) = \log y_*$.

Proof. Theorem 2.1, with $\lambda = \alpha d$, gives

$$\frac{1}{d} \mathcal{R}_d(\delta R_{0,d}^2) = \sup_{\alpha \geq 0} \{ \alpha R_{0,d}^2(1 - \delta) - g_d(\alpha) \}. \quad (40)$$

We first justify passage of the supremum to the limit. Fix $\delta > 0$ and put $\eta = \delta/8$. The beta tail is exact: $\mathbb{P}\{T_d \geq 1 - \eta\} = \eta^{d-1}$. Evaluating the partition on this event and choosing $b = 1$ gives

$$g_d(\alpha) \geq \alpha \left(1 - 2\eta - \frac{1}{d} \right) + \left(1 - \frac{1}{d} \right) \log \eta. \quad (41)$$

Consequently, for every $d \geq 2$ and $\alpha \geq 0$,

$$\alpha R_{0,d}^2(1 - \delta) - g_d(\alpha) \leq -\frac{\delta}{4} \alpha - \left(1 - \frac{1}{d} \right) \log \eta. \quad (42)$$

Thus all maximizing multipliers lie in a common compact interval. Local uniform convergence from Corollary 4.2 now yields

$$r_\infty(\delta) = \sup_{\alpha \geq 0} \{ \alpha(1 - \delta) - g(\alpha) \}. \quad (43)$$

The right derivative of g jumps at α_* from 0 to b_*^2 . Therefore gains $1 - \delta \leq b_*^2$ lie on the coexistence face and give $\alpha_*(1 - \delta)$. If $1 - \delta > b_*^2$, envelope differentiation on the unique positive branch gives

$$1 - \delta = g'(\alpha) = b_+(\alpha)^2. \quad (44)$$

Put $b = \sqrt{1 - \delta}$ and $y = (1 - b)^{-1}$. Substitution into the conjugate is explicit: at a stationary point

$$\alpha b^2 - g(\alpha) = 2\alpha b^2 - \Phi(y) = y - 1 - \Phi(y) = \log y. \quad (45)$$

Thus it reduces to $\log y = -\log(1 - \sqrt{1 - \delta})$. At coexistence, (30) makes this value $\log y_* = \alpha_* b_*^2$. $\square$

The fixed- δ order of limits is essential. The theorem does not assert uniformity as $\delta \downarrow 0$; the fixed-dimensional high-resolution limit is a different asymptotic regime.

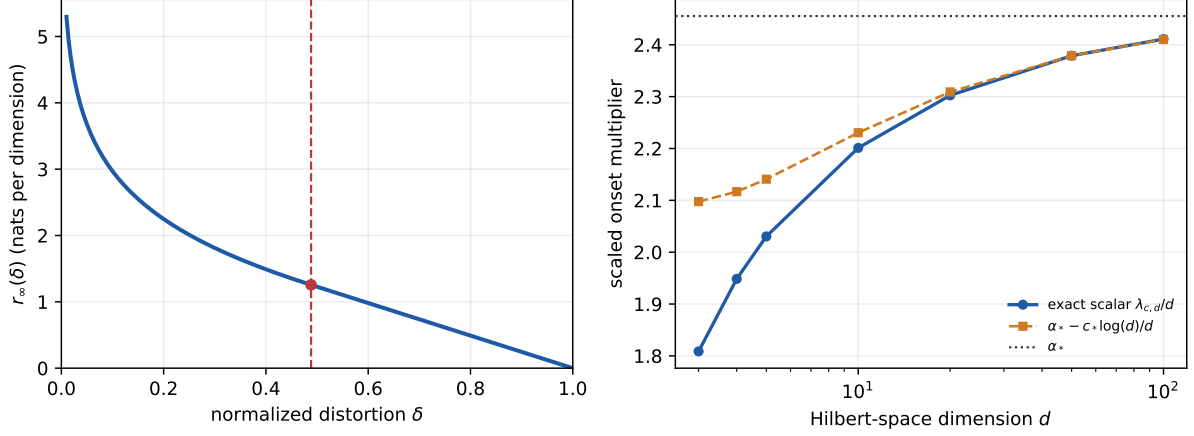


Figure 2: Left: the exact limiting information–accuracy curve and its coexistence point δ_* . Right: finite-dimensional first-contact multipliers from the exact scalar problem, the leading logarithmic approximation, and the limit α_* . The plotted finite-dimensional values are reproducibility diagnostics, not proof inputs.

7 Thermodynamic first contact and its logarithmic finite-size correction

Retain the first multiplier $\lambda_{c,d}$ at which a positive radius is a global maximizer of (5), and choose any positive active radius $b_{c,d}$ there.

Lemma 7.1 (Beta-mgf exclusion bound). *For $0 \leq s < d$,*

$$K_d(s) \leq -\log(1 - s/d) - s/d. \quad (46)$$

Consequently a minimizing first-contact tilt cannot satisfy $s/d \rightarrow 0$.

Proof. The hypergeometric series and $(d)_k \geq d^k$ give

$$M_d(s) = \sum_{k \geq 0} \frac{s^k}{(d)_k} \leq \sum_{k \geq 0} (s/d)^k = (1 - s/d)^{-1}. \quad (47)$$

After centering, this is (46). If $u = s/d \rightarrow 0$, then the right side is $u^2/2 + O(u^3)$, so the ratio term is at least $(1 + o(1))R_{0,d}^2 d^2/2$. On the other hand, evaluating (60) at $s = dy_*$ gives $\lambda_{c,d} = O(d)$. Therefore a minimizing tilt cannot have $s/d \rightarrow 0$. $\square$

Lemma 7.2 (Uniform interior Kummer–Laplace expansion). *Let $J \Subset (1, \infty)$ be compact. Uniformly for $y \in J$,*

$$M_d(dy) = e^{d\Phi(y)} y \sqrt{2\pi d} [1 + O_J(d^{-1})], \quad (48)$$

and hence

$$\log M_d(dy) = d\Phi(y) + \frac{1}{2} \log d + \log(y\sqrt{2\pi}) + O_J(d^{-1}). \quad (49)$$

Proof. Write

$$M_d(dy) = (d-1) \int_0^1 e^{df_y(t)} g(t) dt, \quad f_y(t) = yt + \log(1-t), \quad g(t) = (1-t)^{-2}. \quad (50)$$

The unique maximizer is $t_y = 1 - y^{-1}$, and as y ranges over J the points t_y remain in a fixed compact subinterval of $(0, 1)$. Choose $\rho > 0$ so that $[t_y - \rho, t_y + \rho] \subset (0, 1)$ for every $y \in J$. Compactness and strict concavity give a uniform $c_J > 0$ such that

$$f_y(t) \leq \Phi(y) - c_J \quad \text{whenever } |t - t_y| \geq \rho, y \in J. \quad (51)$$

Thus the exterior integral is exponentially smaller than the main term.

On the interior neighborhood,

$$f_y''(t_y) = -y^2, \quad f_y'''(t_y) = -2y^3, \quad f_y^{(4)}(t_y) = -6y^4, \quad g(t_y) = y^2. \quad (52)$$

All derivatives of f_y through order five and of g through order three are uniformly bounded there. With $z = \sqrt{d}(t - t_y)$, Taylor's theorem through fourth order gives a Gaussian integrand $e^{-y^2 z^2/2}$ times a factor whose even contribution is $1 + O_J(d^{-1})(1 + |z|^6)$; the order- $d^{-1/2}$ term is odd and integrates to zero on the symmetric neighborhood. On $|z| \leq d^{1/10}$ the remainder is uniformly dominated by an integrable Gaussian times $O_J(d^{-1})(1 + |z|^{10})$; on the remainder of the interior neighborhood strict concavity gives an exponentially small bound. Therefore

$$M_d(dy) = (d-1)e^{d\Phi(y)}g(t_y)\sqrt{\frac{2\pi}{d|f_y''(t_y)|}}[1 + O_J(d^{-1})] \quad (53)$$

$$= (d-1)e^{d\Phi(y)}y^2\sqrt{\frac{2\pi}{dy^2}}[1 + O_J(d^{-1})] \quad (54)$$

$$= e^{d\Phi(y)}y\sqrt{2\pi d}[1 + O_J(d^{-1})]. \quad (55)$$

Taking logarithms proves (49). $\square$

Theorem 7.3 (First-contact asymptotics). *As $d \rightarrow \infty$,*

$$\frac{\lambda_{c,d}}{d} \rightarrow \alpha_*, \quad b_{c,d} \rightarrow b_*. \quad (56)$$

The coexistence distortion and rate satisfy

$$\frac{D_{c,d}}{R_{0,d}^2} \rightarrow \delta_*, \quad \frac{R_{c,d}}{d} \rightarrow \log y_*. \quad (57)$$

More sharply,

$$\boxed{\lambda_{c,d} = \alpha_* d - c_* \log d + O(1)}, \quad c_* = \frac{\alpha_*}{2 \log y_*} = \frac{y_*^2}{2(y_* - 1)^2}. \quad (58)$$

Proof. Write $s = 2\lambda b$. Nonnegative radial gain and $b \leq 1$ are equivalent to

$$\lambda \geq \frac{R_{0,d}^2 s^2}{4K_d(s)}, \quad \lambda \geq \frac{s}{2}. \quad (59)$$

Hence

$$\lambda_{c,d} = \inf_{s>0} \max \left\{ \frac{s}{2}, \frac{R_{0,d}^2 s^2}{4K_d(s)} \right\}. \quad (60)$$

Evaluating this expression at $s = dy_*$ and using local-uniform Laplace convergence gives $\limsup_d \lambda_{c,d}/d \leq \alpha_*$. This comparison will be used below to localize every minimizing sequence. For $s = dy$ with $y > 1$, Lemma 4.1 makes the second term divided by d converge locally uniformly to

$$A(y) = \frac{y^2}{4\Phi(y)}. \quad (61)$$

Its unique global minimizer is y_* : the equation $A'(y) = 0$ is exactly (30), $A(y_*) = \alpha_*$, and $y_*/2 < \alpha_*$. Lemma 7.1 excludes $s/d \rightarrow 0$. If instead $s/d \rightarrow y \in (0, 1]$, local-uniform Laplace convergence gives $K_d(s)/d \rightarrow \Phi(y) = 0$, so the ratio term in (60), after division by d , diverges. The term $s/2$ excludes $s/d \rightarrow \infty$. Thus every minimizing sequence lies in a compact subset of

$(1, \infty)$, where uniform convergence and uniqueness imply $s_{c,d}/d \rightarrow y_*$ and $\lambda_{c,d}/d \rightarrow \alpha_*$. For all large d the second term in (60) is active, so

$$b_{c,d} = \frac{s_{c,d}}{2\lambda_{c,d}} = \frac{2K_d(s_{c,d})}{R_{0,d}^2 s_{c,d}} \rightarrow \frac{2\Phi(y_*)}{y_*} = b_*. \quad (62)$$

The distortion and rate limits follow from $D = R_{0,d}^2(1 - b^2)$ and $R = \lambda R_{0,d}^2 b^2$ at coexistence.

It remains to refine the multiplier. Apply Lemma 7.2 on a compact neighborhood of y_* . Using $K_d(dy) = \log M_d(dy) - y$ and $R_{0,d}^2 = 1 - d^{-1}$ gives

$$\frac{R_{0,d}^2 d^2 y^2}{4K_d(dy)} = dA(y) - \frac{A(y)}{2\Phi(y)} \log d + O(1) \quad (63)$$

uniformly there. Put $q(y) = A(y)/(2\Phi(y))$. On a fixed neighborhood U of y_* there are constants $c, L, C > 0$ such that $A(y) \geq A(y_*) + c|y - y_*|^2$, $|q(y) - q(y_*)| \leq L|y - y_*|$, and the remainder in (63) has absolute value at most C . Thus, with $z = y - y_*$, the right side is bounded below by

$$dA(y_*) - q(y_*) \log d + dc z^2 - L(\log d)|z| - C. \quad (64)$$

The variable terms are at least $-L^2(\log d)^2/(4cd) = o(1)$. Evaluation at $y = y_*$ gives the matching upper bound up to $O(1)$. Since the global minimizer has already been shown to enter U , we obtain $\lambda_{c,d} = dA(y_*) - q(y_*) \log d + O(1)$. Finally $A(y_*) = \alpha_*$ and $\Phi(y_*) = \log y_*$ prove (58). $\square$

8 Pointwise physical projection and source universality

So far the source is Haar. We now allow every pure input state and charge a representation channel by its Shannon capacity over all input priors.

Let ν be Haar measure on rank-one queries, or a weighted exact complex projective 2-design. For a scalar report $q \in L^2(\nu)$, define its risk against x by

$$\ell_x(q) = \int |q(y) - \text{tr}(\rho_x \rho_y)|^2 d\nu(y). \quad (65)$$

Theorem 8.1 (Pointwise projection to a physical Born reporter). *For every $q \in L^2(\nu)$ there exists a density matrix $\sigma(q)$ such that*

$$\ell_x(y \mapsto \text{tr}(\sigma(q) \rho_y)) \leq \ell_x(q) \quad (66)$$

simultaneously for every pure x . The map $q \mapsto \sigma(q)$ may be chosen Borel measurable. Moreover,

$$\int |\text{tr}[(\rho_x - \sigma) \rho_y]|^2 d\nu(y) = \frac{\|\rho_x - \sigma\|_F^2}{d(d+1)}. \quad (67)$$

Proof. Let

$$\mathcal{B} = \{y \mapsto \text{tr}(A \rho_y) : A = A^*\} \subset L^2(\nu). \quad (68)$$

The 2-design identity makes this a finite-dimensional closed subspace with matrix norm

$$\|A\|_*^2 = \frac{\text{tr}(A^2) + (\text{tr } A)^2}{d(d+1)}. \quad (69)$$

Let f_A be the orthogonal projection of q onto $\mathcal{B}$. Because $f_{\rho_x} \in \mathcal{B}$,

$$\|f_{\rho_x} - q\|_2^2 = \|f_{\rho_x} - f_A\|_2^2 + \|q - f_A\|_2^2. \quad (70)$$

Next let σ be the metric projection of A in the norm (69) onto the compact convex density-matrix body. Since ρ_x belongs to that body, the Hilbert projection inequality gives

$$\|f_{\rho_x} - f_\sigma\|_2 \leq \|f_{\rho_x} - f_A\|_2. \quad (71)$$

This proves (66). Orthogonal projection onto a finite-dimensional subspace is continuous in L^2 , and metric projection onto a closed convex set in finite dimension is continuous, establishing the measurable choice. Finally, both matrices in (67) have trace one, so the trace term in (69) vanishes. $\square$

Consider a channel W from pure states to an arbitrary standard-Borel classical memory Z , followed by jointly measurable probability reports $q_Z(y) \in [0, 1]$; bounded joint measurability makes $z \mapsto q_z$ a Borel $L^2(\nu)$ -valued map. Its worst-state Born risk is $\sup_x \mathbb{E}[\ell_x(q_Z)|X = x]$. Its information cost is its channel capacity

$$C(W) = \sup_P I_P(X; Z), \quad (72)$$

where the supremum ranges over all pure-state priors. For $0 \leq D \leq R_{0,d}^2$, let

$$\mathfrak{W}_d(D) = \left\{ (W, q) : \sup_x \mathbb{E}[\ell_x(q_Z)|X = x] \leq \frac{D}{d(d+1)} \right\}. \quad (73)$$

Define the source-universal frontier by

$$\mathcal{C}_d^{\text{univ}}(D) = \inf_{(W,q) \in \mathfrak{W}_d(D)} \sup_P I_P(X; Z). \quad (74)$$

Theorem 8.2 (Exact source-universal minimax capacity). *For every $0 \leq D \leq R_{0,d}^2$, the minimum channel capacity in (74) is exactly*

$$\boxed{\mathcal{C}_d^{\text{univ}}(D) = \mathcal{R}_d(D)}. \quad (75)$$

The Haar prior is least favorable for an optimal covariant channel.

Proof. Theorem 8.1 lets us replace every report by a density report without increasing any statewise risk or channel capacity. Every uniformly feasible channel is therefore feasible under Haar input, so

$$C(W) \geq I_{\text{Haar}}(X; Z) \geq \mathcal{R}_d(D). \quad (76)$$

Conversely, Theorem 2.1 supplies a covariant complex-Bingham channel W_D , or an independently flagged mixture of active channels on a coexistence face, with constant statewise distortion D . Let Q_D be its output law under Haar input. For one active orbit, covariance makes the divergence below independent of x , and Haar averaging identifies its value with the attained mutual information. On a coexistence face, write the revealed-flag channel and output law as $W_D = \sum_j a_j \delta_j \otimes W_j$ and $Q_D = \sum_j a_j \delta_j \otimes Q_j$. The flag is independent of x , so the same conclusion follows from $D_{\text{KL}}(W_D(\cdot|x) \| Q_D) = \sum_j a_j D_{\text{KL}}(W_j(\cdot|x) \| Q_j)$. Thus

$$D_{\text{KL}}(W_D(\cdot|x) \| Q_D) = \mathcal{R}_d(D) \quad (77)$$

independent of x . For every input prior P , the classical information-radius identity [16] gives

$$I_P(X; Z) = \int D_{\text{KL}}(W_D(\cdot|x) \| Q_D) dP(x) - D_{\text{KL}}(PW_D \| Q_D) \quad (78)$$

$$\leq \mathcal{R}_d(D). \quad (79)$$

Haar input attains equality, proving both capacity and minimax equality for $0 < D < R_{0,d}^2$. At $D = R_{0,d}^2$ the constant report has zero capacity. At $D = 0$, exact reconstruction of the non-atomic projective source has infinite capacity, matching $\mathcal{R}_d(0) = \infty$. $\square$

This specialization is not a novelty claim for invariant, compound-source, or minimax rate-distortion theory itself; such formulations go back at least to Sakrison, with compact-group and modern universal-coding developments in [9, 10, 11]. The result here is the exact evaluation of the stated projective worst-state frontier.

9 Exact finite-block cost and zero dispersion

For convenience extend $\mathcal{R}_d(D)$ by zero for $D \geq R_{0,d}^2$; this preserves convexity and monotonicity.

Allow an arbitrary joint channel from n pure input states to a shared classical memory with a standard-Borel output alphabet and n joint reports, each Borel as a map from the memory into $L^2(\nu)$. Define

$$\mathcal{C}_{d,n}^{\text{univ}}(D) = \inf_{(W^n, q^n)} \sup_{P_{X^n}} I_P(X^n; Z), \quad (80)$$

where the infimum is subject to

$$\sup_{x^n} \mathbb{E} \left[\frac{1}{n} \sum_{t=1}^n \ell_{x_t}(q_t, Z) \middle| X^n = x^n \right] \leq \frac{D}{d(d+1)}. \quad (81)$$

Theorem 9.1 (Exact arbitrary-joint-memory finite-block law). *For every integer $n \geq 1$ and $0 \leq D \leq R_{0,d}^2$,*

$$\boxed{\mathcal{C}_{d,n}^{\text{univ}}(D) = n\mathcal{R}_d(D)}. \quad (82)$$

Arbitrary correlations in the representation do not lower the cost; a memoryless causal product of covariant complex-Bingham channels attains it.

Proof. For the lower bound, choose the iid Haar prior and apply the pointwise projection coordinatewise. If D_t is the marginal mean projector distortion, independence and the mutual-information chain rule give

$$I(X^n; Z) - \sum_t I(X_t; Z) = \sum_t I(X_t; X^{t-1} | Z) \geq 0. \quad (83)$$

Data processing and the one-letter converse therefore give

$$I(X^n; Z) \geq \sum_{t=1}^n I(X_t; \hat{\rho}_t) \geq \sum_{t=1}^n \mathcal{R}_d(D_t) \geq n\mathcal{R}_d\left(\frac{1}{n} \sum_t D_t\right), \quad (84)$$

where the last step is convexity. The worst-sequence constraint implies $n^{-1} \sum_t D_t \leq D$ under iid Haar input, so monotonicity yields the claimed lower bound. For the upper bound use $W_D^{\otimes n}$. Relative to $Q_D^{\otimes n}$, every input sequence has divergence $n\mathcal{R}_d(D)$ by (77); the information-radius identity upper-bounds capacity by that value, and iid Haar input attains it. $\square$

We next turn from stochastic representations to fixed-length lossy codes. An (n, M, D, ϵ) code maps an iid Haar source X^n to one of M messages and then to physical density reports $\hat{\rho}^n$, with

$$\mathbb{P} \left\{ \frac{1}{n} \sum_{t=1}^n \|\rho_{X_t} - \hat{\rho}_t\|_F^2 > D \right\} \leq \epsilon. \quad (85)$$

Theorem 9.2 (Zero dispersion and exponential strong converse). *Let $0 < D < R_{0,d}^2$ be a differentiability point of $\mathcal{R}_d$; this includes the relative interior of every linear face. Then the corresponding D -tilted information is constant:*

$$j_X(x, D) = \mathcal{R}_d(D) \quad \text{for Haar-almost every } x. \quad (86)$$

Hence the rate dispersion is zero. Every fixed-length n -state code obeys

$$\boxed{\epsilon \geq [1 - \exp\{\log M - n\mathcal{R}_d(D)\}]_+}. \quad (87)$$

In particular, if $\log M \leq n(\mathcal{R}_d(D) - \eta)$, then $\epsilon \geq 1 - e^{-n\eta}$.

Proof. Compactness of the source and physical reproduction spaces and continuity of the bounded distortion give an attaining invariant KKT reproduction law $P_{\hat{\rho}}^*$ at every stated D . Set $\lambda = -\mathcal{R}'_d(D) > 0$ and define

$$J_X(x, D) = -\lambda D - \log \int e^{-\lambda \|\rho_x - \hat{\rho}\|_F^2} dP_{\hat{\rho}}^*(\hat{\rho}). \quad (88)$$

On a linear face take the invariant mixture of the supporting active physical orbit laws. By transitivity, (88) is independent of x ; its Haar mean is $\mathcal{R}_d(D)$ by dual equality. This proves (86) and zero variance. Additivity makes the n -letter tilted information identically $n\mathcal{R}_d(D)$. Theorem 7 of Kostina and Verdú [12] states that, for every $\gamma \geq 0$,

$$\epsilon \geq \mathbb{P}\{J_{X^n}(X^n, nD) \geq \log M + \gamma\} - e^{-\gamma}. \quad (89)$$

Taking $\gamma = n\mathcal{R}_d(D) - \log M$ when positive gives (87). $\square$

Zero dispersion is a symmetry consequence, not a uniquely quantum effect. The theorem is a converse; it does not assert a matching constant or $\frac{1}{2} \log n$ third-order achievability term.

10 Causal and finite-query consequences

Exact weighted complex projective 2-designs replace Haar query integration in Theorem 8.1 without changing any constants. This makes the probability-field risk finitely witnessable. Sequential extensions require an explicit isotropy condition: a policy that always asks one fixed ray tests only one overlap and cannot force full-state information.

Corollary 10.1 (History-dependent conditional-design tensorization). *Let $X_1, \dots, X_n$ be fresh independent Haar preparations. At time t , a causal classical encoder produces Z_t from (X^t, Z^{t-1}) , and the reporter may use Z^t ; any externally randomized menu or seed visible to the reporter is included in this transcript. For every allowed history on which the menu may depend, suppose its conditional scoring measure is a weighted exact projective 2-design and the random query draw has no additional access to the fresh X_t . Let $\hat{\rho}_t(Z^t)$ be the measurable physical projection of the current report and define*

$$D_t = \mathbb{E} \|\rho_{X_t} - \hat{\rho}_t(Z^t)\|_F^2. \quad (90)$$

Then

$$I(X^n; Z^n) \geq \sum_{t=1}^n \mathcal{R}_d(D_t) \geq n\mathcal{R}_d\left(\frac{1}{n} \sum_t D_t\right). \quad (91)$$

For the causal factorization $W(dz^n \| x^n) = \prod_t W_t(dz_t | z^{t-1}, x^t)$, define

$$I(X^n \rightarrow Z^n) = \sum_{t=1}^n I(X^t; Z_t | Z^{t-1}). \quad (92)$$

For this exogenous iid source, Massey's conservation identity [17] gives $I(X^n \rightarrow Z^n) = I(X^n; Z^n)$.

Proof. Apply the conditional 2-design projection and the one-letter converse at each history. Independence gives

$$I(X^n; Z^n) = \sum_t I(X_t; Z^n | X^{t-1}) \quad (93)$$

$$\geq \sum_t I(X_t; Z^t | X^{t-1}) \quad (94)$$

$$= \sum_t I(X_t; Z^t, X^{t-1}) \quad (95)$$

$$\geq \sum_t I(X_t; Z^t), \quad (96)$$

and the one-letter bounds plus convexity complete the proof. Because the source law has no feedback dependence on past Z , reverse directed information vanishes. $\square$

At fixed $0 < \delta < 1$, Theorem 6.1 also implies a classical predictive information cost of $d r_\infty(\delta) / \log 2 + o(d)$ bits per state. A direct d -level quantum system has logarithmic Hilbert-space size, but this comparison must not be misread: a single quantum specimen is not an exact reusable classical probability oracle, and remote-state preparation or shared entanglement is a different resource model [18, 19].

11 Operational boundary and open Grassmann frontier

The theorems concern probability-field calibration. They do not derive Born's rule, simulate state update, certify intrinsic randomness, or exclude click-only hidden-variable models. Indeed, reproducing sampled clicks can require far less information than retaining calibrated probabilities for all queries. Channel capacity is an operational asymptotic information resource; it is not a literal count of physical memory states in one realization.

The natural mathematical extension replaces a ray by a Haar rank- r projector. Axial overlaps then have beta law $\text{Beta}(r, d - r)$, and the skewness changes sign under $r \leftrightarrow d - r$. However, the full partition is matrix-Jacobi rather than scalar Dirichlet. An axial ansatz is not a proof: exact rank- r rate-distortion requires showing that a specified reconstruction spectrum globally maximizes a matrix-Bingham/HCIZ integral at fixed Frobenius norm. Existing complex-projective and Grassmann quantization results, arbitrary-radius metric-ball asymptotics, and general rate-distortion bounds do not supply that extremum [4, 5, 6, 7]. We therefore leave it as an explicit open problem.

12 Conclusion

We have given a self-contained rate-distortion theory for calibrated Born prediction on complex projective space. The all-degree one-spike partition theorem reduces the unrestricted finite-dimensional problem to an exact scalar complex-Bingham envelope, including flagged mixtures at coexistence. It yields a discontinuous directional-information onset for every $d \geq 3$, the exact qubit specialization, and the sharp fixed- d high-fidelity constant. In the thermodynamic regime, a uniform projective Laplace principle gives the closed two-piece frontier, the nontrivial coexistence root $y - 1 = 2 \log y$, and the finite-size correction $\lambda_{c,d} = \alpha_* d - c_* \log d + O(1)$. Finally, a pointwise physical-report projection upgrades Haar-average optimality to exact worst-state channel capacity, arbitrary-joint-memory tensorization, zero dispersion, and an exponential strong converse. The resulting theorem package connects exact finite-dimensional probability prediction to its high-dimensional phase law without identifying calibrated prediction with click generation or particle dynamics.

Data and code availability

The accompanying archive `repro_v1.zip` is supplied as supplementary material with this manuscript. It has the immutable content identifier SHA-256

54a0fd17a2b68bdfc62a188043be104ef533b1b20581789a9992ba912902821e.

From the extracted directory, the exact replay commands are

```
python verify_frontier.py
python verify_thermodynamic_limit.py --max-d 100 --output-dir .
```

They evaluate the finite-dimensional scalar envelope, the coexistence root, the bounded first-contact problem, and the limiting curve using one process, without branch-and-bound or heavy search. The generated JSON files record the environment, grids, residuals, and runtime. These outputs are diagnostic; all headline results are analytic.

AI assistance and author responsibility

AI tools assisted with literature search, algebraic stress testing, editing, and reproducibility checks. The author is responsible for the definitions, proofs, claims, citations, and final manuscript.

A Finite-dimensional spectral reduction

We give the complete reduction so that every finite-dimensional input used in the main text is verifiable within this manuscript.

Proposition A.1 (Brazitikos–Pandis normalized interface). *Let $d \geq 2$, let $\gamma_1, \gamma_2, \gamma_3$ be nonnegative integer multiplicities with $\gamma_1 + \gamma_2 + \gamma_3 = d$, and let $u, v, w \in \mathbb{R}$ satisfy*

$$\gamma_1 u + \gamma_2 v + \gamma_3 w = 0, \quad \gamma_1 u^2 + \gamma_2 v^2 + \gamma_3 w^2 = 1. \quad (97)$$

For the grouped vector

$$x = (\underbrace{u, \dots, u}_{\gamma_1}, \underbrace{v, \dots, v}_{\gamma_2}, \underbrace{w, \dots, w}_{\gamma_3}), \quad (98)$$

and every positive integer m ,

$$p_m(x) \leq p_m(a_1^*), \quad a_1^* = \left(\sqrt{\frac{d-1}{d}}, -\frac{1}{\sqrt{d(d-1)}}, \dots, -\frac{1}{\sqrt{d(d-1)}} \right). \quad (99)$$

Here zero multiplicities cover the two-value case. This is precisely Proposition 5.5 of Brazitikos and Pandis in their unit-norm normalization [15]. Homogeneity gives radius r by multiplying the right side by r^m .

Remark A.2. Proposition A.1 is the only external extremal inequality used below. It asserts the grouped power-sum maximum. The reduction of an arbitrary spectrum to at most three values, the odd-sign closure, the Newton step, the projective Laplace theorem, and the rate–distortion application are proved here.

Lemma A.3 (All-degree centered power sums). *Let $d \geq 3$, $r > 0$, and let $a \in \mathbb{R}^d$ satisfy $\sum_i a_i = 0$ and $\sum_i a_i^2 = r^2$. Put*

$$a^* = r \left(\sqrt{\frac{d-1}{d}}, -\frac{1}{\sqrt{d(d-1)}}, \dots, -\frac{1}{\sqrt{d(d-1)}} \right). \quad (100)$$

Then, for every integer $m \geq 2$,

$$|p_m(a)| \leq p_m(a^*), \quad p_m(a) = \sum_i a_i^m. \quad (101)$$

Proof. For $m = 2$, the assertion is immediate because $p_2(a) = r^2 = p_2(a^*)$ is fixed by the norm constraint. Hence assume $m \geq 3$. The centered sphere is compact. Its constraint gradients $\mathbf{1}$ and $2a$ are independent because a is nonzero and centered. At a constrained global maximum or minimum of p_m , Lagrange multipliers give

$$mx^{m-1} - 2\xi x - \zeta = 0 \quad (102)$$

for every coordinate value x . The derivative of this sparse polynomial is $m(m-1)x^{m-2} - 2\xi$. It has at most one real zero when m is odd and at most two when m is even. Rolle's theorem therefore leaves at most two, respectively three, distinct coordinate values at every global extremum. Proposition A.1, applied precisely to this grouped centered fixed-norm class, places the maximum of p_m at a^* . For odd m , the map $a \mapsto -a$ converts the minimum to minus the maximum; for even m , $p_m(a) \geq 0$. This proves (101). The external proposition supplies the grouped power-sum extremum, not the all-degree Laplace conclusion below. $\square$

If $P \sim \text{Dirichlet}(1, \dots, 1)$, then

$$\mathbb{E}\langle a, P \rangle^k = \frac{k!(d-1)!}{(d+k-1)!} h_k(a), \quad (103)$$

where h_k is the complete homogeneous symmetric polynomial. Since $p_1(a) = 0$, Newton's expansion becomes

$$h_k(a) = \sum_{\substack{m_1+2m_2+\dots+km_k=k \\ m_1=0}} \prod_{j=2}^k \frac{p_j(a)^{m_j}}{m_j! j^{m_j}}. \quad (104)$$

For $d \geq 3$, every $p_j(a^*)$ with $j \geq 2$ is positive. Hence

$$h_k(a) \leq |h_k(a)| \leq \sum_{\substack{m_1+2m_2+\dots+km_k=k \\ m_1=0}} \prod_{j=2}^k \frac{|p_j(a)|^{m_j}}{m_j! j^{m_j}} \leq h_k(a^*). \quad (105)$$

Equations (103) and (105), summed termwise in the entire exponential series, prove the axial projective partition inequality in every degree. When $d = 2$, the centered fixed-radius sphere consists only of the two permutations of $(r/\sqrt{2}, -r/\sqrt{2})$, so the result is immediate.

It remains to derive and attain the scalar envelope. Write a density reproduction as $I/d + A$, where $\text{tr } A = 0$ and $\|A\|_F \leq R_{0,d}$. For

$$Q_\lambda(A) = \int e^{-\lambda \|S_x - A\|_F^2} d\mu_d(x), \quad (106)$$

conditional Gibbs duality, applied to every posterior given A , gives

$$I(X; A) \geq -\lambda D - \mathbb{E} \log Q_\lambda(A) \geq -\lambda D - \log \sup_A Q_\lambda(A). \quad (107)$$

At fixed $\|A\|_F$, diagonalization and the all-degree partition inequality replace its spectrum by the valid axial density comparator $A = b(\rho_u - I/d)$, $0 \leq b \leq 1$. For this comparator,

$$\log Q_\lambda(A) = -\lambda R_{0,d}^2 + K_d(2\lambda b) - \lambda R_{0,d}^2 b^2 = -\lambda R_{0,d}^2 + G_{d,\lambda}(b). \quad (108)$$

Maximizing first over b and then over λ proves the converse in Theorem 2.1.

For equality, let $b > 0$ be active, put $\kappa = 2\lambda b$, and take

$$\frac{dP_{X|U=u}}{d\mu_d}(x) = \frac{e^{\kappa |\langle x, u \rangle|^2}}{M_d(\kappa)}, \quad U \sim \mu_d. \quad (109)$$

Invariance makes the source marginal Haar. If $m_d(\kappa) = \partial_\kappa \log M_d(\kappa)$, stationarity of the active radius is

$$b = \frac{dm_d(\kappa) - 1}{d - 1}. \quad (110)$$

Thus the conditional mean is exactly $(1 - b)I/d + b\rho_u$, so the channel is its own posterior-mean reproduction. It attains

$$D_b = R_{0,d}^2(1 - b^2), \quad I_b = \kappa m_d(\kappa) - \log M_d(\kappa) = \lambda(R_{0,d}^2 - D_b) - \max_{0 \leq v \leq 1} G_{d,\lambda}(v). \quad (111)$$

The optimized log partition $c_d(\lambda) = -\lambda R_{0,d}^2 + \max_b G_{d,\lambda}(b)$ is convex, and Danskin's theorem gives

$$-\partial c_d(\lambda) = \text{conv}\{D_b : b \text{ is active at } \lambda\}. \quad (112)$$

Independently flagged mixtures of the corresponding invariant channels attain every distortion in this interval. The independent channel gives $D = R_{0,d}^2$, $I = 0$; exact reconstruction is obtained as the limiting endpoint and has infinite information. Thus all $0 \leq D \leq R_{0,d}^2$ are covered without assuming uniqueness of the finite-dimensional radial maximizer.

References

- [1] Armine Bagyan and Donald Richards. “Complete Asymptotic Expansions and the High-Dimensional Bingham Distributions”. In: *TEST* 33.2 (2024), pp. 540–563. DOI: 10.1007/s11749-023-00910-w. eprint: 2303.10290. URL: <https://doi.org/10.1007/s11749-023-00910-w>.
- [2] Nico M. Temme. “Uniform Asymptotic Expansions of Confluent Hypergeometric Functions”. In: *IMA Journal of Applied Mathematics* 22.2 (1978), pp. 215–223. DOI: 10.1093/imamat/22.2.215. URL: <https://doi.org/10.1093/imamat/22.2.215>.
- [3] José L. López and Pedro J. Pagola. “The Confluent Hypergeometric Functions $M(a, b; z)$ and $U(a, b; z)$ for Large b and z ”. In: *Journal of Computational and Applied Mathematics* 233.6 (2010), pp. 1570–1576. DOI: 10.1016/j.cam.2009.02.072. URL: <https://doi.org/10.1016/j.cam.2009.02.072>.
- [4] Bishwarup Mondal, Satyaki Dutta, and Robert W. Heath Jr. “Quantization on the Complex Projective Space”. In: *Proceedings of the Data Compression Conference*. 2006. DOI: 10.1109/DCC.2006.68. URL: <https://doi.org/10.1109/DCC.2006.68>.
- [5] Wei Dai, Youjian Liu, and Brian Rider. “Quantization Bounds on Grassmann Manifolds and Applications to MIMO Communications”. In: *IEEE Transactions on Information Theory* 54.3 (2008), pp. 1108–1123. DOI: 10.1109/TIT.2007.915691. eprint: cs/0603039. URL: <https://doi.org/10.1109/TIT.2007.915691>.
- [6] Renaud-Alexandre Pitaval, Lu Wei, Olav Tirkkonen, and Jukka Corander. “Volume of Metric Balls in High-Dimensional Complex Grassmann Manifolds”. In: *IEEE Transactions on Information Theory* 62.9 (2016), pp. 5105–5116. DOI: 10.1109/TIT.2016.2594289. arXiv: 1508.00256. URL: <https://doi.org/10.1109/TIT.2016.2594289>.
- [7] Erwin Riegler, Günther Koliander, and Helmut Bölcskei. “Lossy Compression of General Random Variables”. In: *Information and Inference: A Journal of the IMA* 12.3 (2023), pp. 1759–1829. DOI: 10.1093/imaiai/iaac035. eprint: 2111.12312. URL: <https://doi.org/10.1093/imaiai/iaac035>.
- [8] Alex Dytso and Martina Cardone. *Uniform Distribution on $(n - 1)$ -Sphere: Rate-Distortion under Squared Error Distortion*. 2024. arXiv: 2401.04248 [cs.IT]. URL: <https://arxiv.org/abs/2401.04248>.

- [9] David J. Sakrison. “The Rate Distortion Function for a Class of Sources”. In: *Information and Control* 15.2 (1969), pp. 165–195. DOI: 10.1016/S0019-9958(69)90403-3. URL: [https://doi.org/10.1016/S0019-9958\(69\)90403-3](https://doi.org/10.1016/S0019-9958(69)90403-3).
- [10] Peter Harremoës. “Maximum Entropy on Compact Groups”. In: *Entropy* 11.2 (2009), pp. 222–237. DOI: 10.3390/e11020222. URL: <https://doi.org/10.3390/e11020222>.
- [11] Adeel Mahmood and Aaron B. Wagner. *Minimax Rate-Distortion*. 2022. arXiv: 2202.04481 [cs.IT]. URL: <https://arxiv.org/abs/2202.04481>.
- [12] Victoria Kostina and Sergio Verdú. “Fixed-Length Lossy Compression in the Finite Blocklength Regime”. In: *IEEE Transactions on Information Theory* 58.6 (2012), pp. 3309–3338. DOI: 10.1109/TIT.2012.2186786. eprint: 1102.3944. URL: <https://doi.org/10.1109/TIT.2012.2186786>.
- [13] John T. Kent. “The Complex Bingham Distribution and Shape Analysis”. In: *Journal of the Royal Statistical Society: Series B* 56.2 (1994), pp. 285–299. DOI: 10.1111/j.2517-6161.1994.tb01978.x. URL: <https://doi.org/10.1111/j.2517-6161.1994.tb01978.x>.
- [14] Kanti V. Mardia and Ian L. Dryden. “The Complex Watson Distribution and Shape Analysis”. In: *Journal of the Royal Statistical Society: Series B* 61.4 (1999), pp. 913–926. DOI: 10.1111/1467-9868.00210. URL: <https://doi.org/10.1111/1467-9868.00210>.
- [15] Silouanos Brazitikos and Christos Pandis. *Sharp Inequalities for Symmetric Polynomials, Hunter’s Conjecture, and Moments of Exponential Random Variables*. 2025. arXiv: 2512.12254 [math.PR]. URL: <https://arxiv.org/abs/2512.12254>.
- [16] Imre Csiszár and Paul C. Shields. “Information Theory and Statistics: A Tutorial”. In: *Foundations and Trends in Communications and Information Theory* 1.4 (2004), pp. 417–528. DOI: 10.1561/01000000004. URL: <https://doi.org/10.1561/01000000004>.
- [17] James L. Massey. “Causality, Feedback and Directed Information”. In: *Proceedings of the International Symposium on Information Theory and Its Applications*. 1990, pp. 303–305. URL: https://www.isiweb.ee.ethz.ch/archive/massey_pub/pdf/BI532.pdf.
- [18] Charles H. Bennett, David P. DiVincenzo, Peter W. Shor, John A. Smolin, Barbara M. Terhal, and William K. Wootters. “Remote State Preparation”. In: *Physical Review Letters* 87.7 (2001), p. 077902. DOI: 10.1103/PhysRevLett.87.077902. eprint: quant-ph/0006044. URL: <https://doi.org/10.1103/PhysRevLett.87.077902>.
- [19] Charles H. Bennett, Patrick Hayden, Debbie W. Leung, Peter W. Shor, and Andreas Winter. “Remote Preparation of Quantum States”. In: *IEEE Transactions on Information Theory* 51.1 (2005), pp. 56–74. DOI: 10.1109/TIT.2004.839476. eprint: quant-ph/0307100. URL: <https://doi.org/10.1109/TIT.2004.839476>.