

A finite free-entropy gap for two self-adjoint variables

Myeongjin Kang

October 18, 2026

Abstract

Starting from the stipulated finite-entropy separation in the pinned OpenAI Math preprint, we construct exactly two bounded self-adjoint elements with $-\infty < \chi < \chi^* < \infty$. The microstates entropy is the original operator-norm-cutoff entropy with an ordinary limsup over all matrix sizes. For the source tuple X of length $n = 2^{64} + 1$, put $k = 2^{32}$. A normal-slice encoding $E_\varepsilon(X)$ satisfies $\liminf_{\varepsilon \downarrow 0} (\chi^*(E_\varepsilon(X)) - \chi(E_\varepsilon(X))) \geq k^{-2}(\chi^*(X) - \chi(X))$. The proof couples an unconditional Fisher-information estimate with a matched microstates-volume estimate. The latter recovers matrix units from arbitrary pair microstates after removing a vanishing fraction of the dimension, so the upper bound includes every output size. All analytic estimates needed for this transfer are proved below; the source separation and explicitly named classical free-entropy interfaces are inputs. No assertion about generation of the ambient von Neumann algebra is made.

1 Statement, source, and contribution

All tracial von Neumann algebras in this paper have a faithful normal tracial state. All tuples are finite and consist of bounded self-adjoint elements. The two free entropies have the normalizations of Section 2. In particular, χ never denotes a relative entropy or an ultrafilter entropy unless explicitly decorated.

Theorem 1.1. *There are a tracial von Neumann algebra $(\mathcal{N}, \tau_{\mathcal{N}})$ and two bounded self-adjoint elements $Z_1, Z_2 \in \mathcal{N}$ such that*

$$-\infty < \chi(Z_1, Z_2) < \chi^*(Z_1, Z_2) < \infty.$$

Here χ uses the original operator-norm cutoff and ordinary limsup over matrix sizes.

The source input is the main theorem and the specified construction of [1], at the immutable revision recorded in Appendix A. Its theorems are stipulated inputs to this work; we do not reprove or audit them. More precisely, that construction gives

$$n = 2^{64} + 1, \quad X = C + \sqrt{h}S, \quad h > 0, \quad -\infty < \chi(X) \leq \chi^*(X) - \frac{1}{2} < \infty. \quad (1)$$

Here $C = (B_1, \dots, B_p, Y)$, $p = 2^{64}$, the B_i are standard free semicircular variables, and Y is uniform on p equally spaced points in $[-1, 1]$ in a separate tensor factor. The family S is standard semicircular and free from C . The source constructs bounded matrix models. Thus $W^*(X)$ is embeddable into a tracial ultrapower of the hyperfinite factor, and

$$\Phi^*(X) \leq n/h < \infty. \quad (2)$$

These properties, the length n , and the positive finite gap are the only information imported from the source construction.

Set $k = 2^{32}$, so $n = k^2 + 1$. Let $D_1 = \text{diag}(1, \dots, k)$ and let D_2 be the adjacency matrix of the path on $\{1, \dots, k\}$. On Hermitian pairs use $\langle A, B \rangle = \sum_{s=1}^2 \text{tr}_k(A_s B_s)$. Let $V_1, \dots, V_n$ be any real orthonormal basis of the orthogonal complement of

$$\{(i[H, D_1], i[H, D_2]) : H \in M_k(\mathbb{C})^{\text{sa}}\}.$$

Such a complement has dimension $k^2 + 1$, as proved below. Define

$$E_\varepsilon(X)_s = D_s \otimes 1 + \varepsilon \sum_{a=1}^n V_{as} \otimes X_a, \quad s = 1, 2. \quad (3)$$

The ambient algebra is $M_k(W^*(X))$ with trace $\text{tr}_k \otimes \tau$.

Theorem 1.2 (Transfer for the specified source). *For X in (1), both entropies of $E_\varepsilon(X)$ are finite for every sufficiently small $\varepsilon > 0$, and*

$$\liminf_{\varepsilon \downarrow 0} \mathcal{G}(E_\varepsilon(X)) \geq k^{-2} \mathcal{G}(X), \quad \mathcal{G}(A) = \chi^*(A) - \chi(A). \quad (4)$$

Consequently every sufficiently small positive encoding parameter gives a strictly positive gap. No effective useful value of this parameter is asserted.

Although several lemmas below hold for arbitrary fixed $k \geq 2$ and bounded input tuples, the stated transfer target is the source just specified. No uniformity as $k \rightarrow \infty$ is claimed. The parameter k is held fixed in every matrix-size and small-parameter limit.

The contribution is confined to three coupled operations: the unconditional Fisher bound of Proposition 5.1; the unconditional, all-output-size upper volume bound of Proposition 4.1, with the same constant as its lower bound; and their coupling in Section 6. The underlying matrix algebra, orbit volumes, lower packing pattern, orthogonal semicircular rotation, relative change of variables, and final entropy-increment comparison are classical or supplied by the source. Section 8 gives the precise comparison and its literature-screen limits. In particular, we make no worldwide priority claim.

2 Entropy conventions and classical interfaces

Write $\text{tr}_d = d^{-1} \text{Tr}_d$ and equip $M_d(\mathbb{C})^{\text{sa}}$ with real inner product $\text{Tr}_d(AB)$. All matrix Lebesgue volumes refer to this *unnormalized* Hilbert–Schmidt metric. For a q -tuple A , let $\Gamma_R(A; m, d, \gamma)$ be the tuples $a \in (M_d^{\text{sa}})^q$ with $\max_j \|a_j\| \leq R$ and

$$|\text{tr}_d(a_{j_1} \cdots a_{j_\ell}) - \tau(A_{j_1} \cdots A_{j_\ell})| < \gamma \quad (1 \leq \ell \leq m).$$

With $\log 0 = -\infty$, set

$$\chi_R(A) = \inf_{m \in \mathbb{N}, \gamma > 0} \limsup_{d \rightarrow \infty} \left[d^{-2} \log \text{Vol } \Gamma_R(A; m, d, \gamma) + \frac{q}{2} \log d \right], \quad (5)$$

$$\chi(A) = \sup_{R > 0} \chi_R(A). \quad (6)$$

We use cutoff independence: $\chi_R(A) = \chi(A)$ when $R > \max_j \|A_j\|$ [7].

The conjugate system $\xi_j \in L^2(W^*(A), \tau)$ is characterized by

$$\tau(\xi_j P(A)) = (\tau \otimes \tau)(\partial_j P(A)), \quad \partial_j x_i = \delta_{ij} 1 \otimes 1.$$

The free difference quotient has its usual outer bimodule Leibniz rule. Set $\Phi^*(A) = \sum_j \|\xi_j\|_2^2$, or $+\infty$ if no conjugate system exists. With standard free semicircular noise S free from A ,

$$\chi^*(A) = \frac{q}{2} \log(2\pi e) + \frac{1}{2} \int_0^\infty \left(\frac{q}{1+s} - \Phi^*(A + \sqrt{s}S) \right) ds. \quad (7)$$

The standard interfaces used here are the comparison $\chi \leq \chi^*$ [10, 11], the free heat identity

$$\chi^*(A + \sqrt{t}S) - \chi^*(A) = \frac{1}{2} \int_0^t \Phi^*(A + \sqrt{s}S) ds, \quad (8)$$

when the quantities are finite, and the projection bound $\Phi^*(A + \sqrt{s}S) \leq \Phi^*(A)$ for finite initial Fisher information [8]. The normalization (7) and the finite heat identity agree with equations `eq:chistar` and `eq:freeflow` of [1]. The heat identity follows by shifting the integral in (7); the logarithmic reference terms cancel. Finite initial Fisher information makes its lower endpoint integrable in our applications.

We also use the covariance upper bound

$$\chi^*(A) \leq \frac{q}{2} \log(2\pi e c_A), \quad c_A = q^{-1} \sum_j \tau((A_j - \tau(A_j))^2) > 0. \quad (9)$$

For clarity, the classical derivation is short. Conjugate identities and Cauchy–Schwarz give $\Phi^*(A + \sqrt{s}S) \geq q/(c_A + s)$; if conjugates do not exist the inequality is automatic. Substitution in (7), using $\int_0^\infty ((1+s)^{-1} - (c_A + s)^{-1}) ds = \log c_A$, proves (9). A scalar tuple has entropy $-\infty$ and presents no exceptional finite-entropy case.

The relative entropy $\chi(A | B)$ used once below is Shlyakhtenko’s supremum-over-conditioning-models entropy [4, Definition 2.3]. For a finite conditioning tuple, take the supremum of the fiber volume over its bounded matrix models, then the ordinary limsup, infimum over joint moment accuracies, and supremum over cutoffs; for an algebra take the infimum over finite conditioning families. It is not defined by subtraction of joint entropies. The particular classical statements needed are:

- (i) $\chi(A | T) \leq \chi(A : T) \leq \chi(A)$, by [4, Proposition 2.8] and projection of joint microstate sets.
- (ii) If A is free from an embeddable algebra B , then $\chi(A | B) = \chi(A)$ [4, Theorem 2.19].
- (iii) An invertible self-adjoint noncommutative power-series map F with coefficients in B , with both inverse series converging on multiradii strictly exceeding the evaluated norms, satisfies

$$\chi(F(A) | B) = \chi(A | B) + (\text{Tr}_q \otimes \tau \otimes \tau)(\log |D_B F(A)|). \quad (10)$$

This is [4, Theorem 4.1]. The proof is omitted in the read arXiv v1, by reference to Voiculescu’s argument; we use its stated theorem. Every printed analytic hypothesis is verified in Section 6.2.

- (iv) For embeddable $W^*(A)$ and $t > 0$,

$$\chi(A + \sqrt{t}S) \geq \frac{q}{2} \log(2\pi e t). \quad (11)$$

This is [4, Proposition 4.2], with the $q/2$ constant in its proof. The displayed proposition in arXiv v1 prints $q \log$ instead; that is a normalization typo. The embeddability hypothesis is essential to this use.

3 Uniform normal geometry

For this and the next two sections, fix $k \geq 2$ and $n = k^2 + 1$. Constants may depend on k, D and the indicated operator-norm bounds, but never on a matrix approximation size. For tuples, $\|\cdot\|$ denotes the maximum coordinate norm; changing to the sum norm changes only such constants.

Commuting with D_1 forces a scalar matrix to be diagonal, and commuting with every path edge of D_2 equates its diagonal entries. Thus their common commutant is $\mathbb{C}1$. The real map

$$L(H) = (i[H, D_1], i[H, D_2])$$

has kernel $\mathbb{R}1$ and rank $k^2 - 1$. Write $\mathcal{O} = \text{ran } L$ and $\mathcal{V} = \mathcal{O}^\perp$, so $\dim_{\mathbb{R}} \mathcal{V} = n$. The basis V_a in (3) is orthonormal for normalized tr_k . A pair V is normal precisely when $\sum_s [D_s, V_s] = 0$.

At size kd let $\mathcal{K}_d$ be the orthogonal complement of $I_k \otimes M_d^{\text{sa}}$ in M_{kd}^{sa} . The map

$$\mathcal{L}_d(K, Y) = i[K, D \otimes I_d] + Y, \quad K \in \mathcal{K}_d, \quad Y \in \mathcal{V} \otimes_{\mathbb{R}} M_d^{\text{sa}}, \quad (12)$$

is an isomorphism onto the pair space. In Hermitian coefficient bases it is the fixed scalar map tensored with the identity on M_d^{sa} . Its Hilbert–Schmidt inverse norm is therefore independent of d . Its operator-norm inverse is uniformly bounded as well: extract the finitely many $k \times k$ coefficients by block maps, apply the fixed scalar inverse matrix, and reconstruct. Each block map is bounded in operator norm independently of d . This same argument works with entries in any unital C^* -algebra.

Lemma 3.1 (Uniform analytic normal tube). *There is $r_0 > 0$, depending only on k, D , such that the associated bundle map*

$$[U, Y] \mapsto U(D \otimes I_d + Y)U^*, \quad U \in U(kd), \quad Y \in \mathcal{V} \otimes_{\mathbb{R}} M_d^{\text{sa}}, \quad \|Y\| < r_0, \quad (13)$$

over $U(kd)/(I_k \otimes U(d))$ is injective, for every d . Locally the normal coordinates and the transverse unitary have absolutely convergent noncommutative power-series descriptions, with uniform radii and derivative bounds. These descriptions hold over any coefficient C^* -algebra. On $\|Y\| \leq C_R \varepsilon$, the Jacobian relative to orbit volume times Euclidean normal-fiber volume lies between $\exp(-C\varepsilon d^2)$ and $\exp(C\varepsilon d^2)$.

Proof. Let P_o, P_v be the fixed coefficient projections onto the orbit and normal spaces. Choose the sign of L as above and put

$$G(K, W) = P_o(e^{-iK}(D + W)e^{iK} - D) = P_o W - LK + \mathcal{R}(K, W).$$

The exponential series gives, for small K, W ,

$$\begin{aligned} \|\mathcal{R}(K, W)\| &\leq C(\|K\|^2 + \|K\|\|W\|), \\ \|\mathcal{R}(K, W) - \mathcal{R}(K', W)\| &\leq C(\|K\| + \|K'\| + \|W\|)\|K - K'\|. \end{aligned}$$

All coefficient projections and L^{-1} are the fixed finite maps just discussed. Hence

$$K = L^{-1}P_o W + L^{-1}\mathcal{R}(K, W) \quad (14)$$

is a strict contraction on $\|K\| \leq 2\|L^{-1}P_o\|\|W\|$ for a sufficiently small uniform radius.

For the analytic assertion, give a series in finitely many noncommuting letters the absolute majorant norm obtained by summing the norms of its coefficients times the products of assigned positive letter radii. This is a Banach algebra norm and is submultiplicative. The exponential series

and the two estimates above hold in this norm. Iteration of (14) starting at zero gives an absolutely convergent series. Choose the radius of actual evaluation strictly smaller than its majorant radius. Differentiation is then bounded, since a degree- ℓ term acquires at most ℓ factors and the radius margin dominates this factor. The same observation gives a second derivative bound. The resulting normal displacement is $P_v(e^{-iK}(D+W)e^{iK}-D)$.

The local map $(K, Y) \mapsto e^{iK}(D+Y)e^{-iK}$ is injective on a uniform convex small ball. Indeed its derivative differs from (12) by $O(r)$ in operator norm and in Hilbert–Schmidt operator norm on a ball of radius r . The derivative formula for the exponential and $\|AB\|_{\text{HS}} \leq \|A\|\|B\|_{\text{HS}}$ prove both bounds. Integrating $\mathcal{L}_d^{-1}D\Psi - I$ along a segment gives a Lipschitz error less than $1/2$ after decreasing r .

For global uniqueness, if two normal representatives in (13) coincide, their transition unitary u satisfies $\|[u, D_s]\| \leq 2r_0$. From D_1 , $\|u_{ij}\| \leq 2r_0/|i-j|$ for $i \neq j$. In the tree-edge entries of $[u, D_2]$, the other terms are off-diagonal blocks of this size; thus $\|u_{ii} - u_{jj}\| = O_k(r_0)$ on each edge and then along every path. The average diagonal block is close to a unitary, and its polar part v satisfies $\|u - I_k \otimes v\| = O_k(r_0)$. Near this subgroup the map $(K, v) \mapsto e^{iK}(I_k \otimes v)$ has a uniform analytic inverse. To see this without a size-dependent inverse function theorem, use the bounded coefficient projections onto $\mathcal{K}_d$ and its complement, the convergent logarithm near the identity, and the same contraction argument as (14). Removing v reduces to the local injective chart. The representatives therefore agree in the associated bundle. This argument is valid over any coefficient algebra, a fact needed when identifying recovered laws.

Finally fix $[U, Y]$ and use coordinates $(\kappa, y) \mapsto Ue^{i\kappa}(D+Y+y)e^{-i\kappa}U^*$ at that point. At the coordinate origin the bundle measure has density $|\det \mathcal{L}_d|$ relative to the product Lebesgue measure: its horizontal density is the orbit differential and its vertical density is Euclidean. The differential of the map is $\mathcal{L}_d + E$, where $E(\kappa, y) = i[\kappa, Y]$. Thus its relative Jacobian is

$$\frac{|\det(\mathcal{L}_d + E)|}{|\det \mathcal{L}_d|} = |\det(I + \mathcal{L}_d^{-1}E)|.$$

Here $\|\mathcal{L}_d^{-1}E\|_{\text{HS} \rightarrow \text{HS}} = O(\varepsilon)$. All singular values lie between $1 - C\varepsilon$ and $1 + C\varepsilon$ in real dimension $2k^2d^2$, giving the stated exponential bounds. This is a pointwise local-coordinate calculation, so invariance makes it valid throughout the bundle. $\square$

3.1 The common volume constant

Let c_D be the product of the $k^2 - 1$ positive singular values of L on the scalar traceless Hermitian space, with normalized trace metrics. Each is repeated d^2 times at size kd . The base-orbit volume is

$$A_d = c_D^{d^2} \frac{\text{Vol } U(kd)}{k^{d^2/2} \text{Vol } U(d)}. \quad (15)$$

The factor $k^{d^2/2}$ comes from the metric induced by $v \mapsto I_k \otimes v$. The normal coefficient map has Jacobian

$$(\varepsilon\sqrt{k})^{nd^2}. \quad (16)$$

Indeed a Hermitian coefficient basis orthonormal for Tr_d is multiplied in pair norm by $\sqrt{k}$ on tensoring with V_a .

The classical unitary volume formula in this metric is

$$\text{Vol } U(l) = \frac{(2\pi)^{l(l+1)/2}}{\prod_{j=1}^{l-1} j!}, \quad c = \frac{3}{4} + \frac{1}{2} \log(2\pi). \quad (17)$$

For example, projecting to the first column gives a quotient sphere whose $2l - 2$ horizontal metric directions have squared length multiplied by 2, and whose phase direction is unchanged. Its volume is $(2\pi)^l/(l-1)!$; induction proves the formula. Stirling summation then yields

$$\frac{\log A_d}{d^2} = -\frac{k^2 - 1}{2} \log d - \frac{k^2 + 1}{2} \log k + (k^2 - 1)c + \log c_D + o(1). \quad (18)$$

The orbit/fiber factorization and the constant c are established antecedents: see Nica–Shlyakhtenko–Speicher [5, Lemmas 2.6 and 2.8], including their reference to Marinov, and the upper/lower orbit patterns in their Lemmas 2.9 and 2.10. We claim no novelty for these constants or the bookkeeping (15)–(18).

4 Matched microstates volumes at every output size

Proposition 4.1. *Let A be a bounded self-adjoint $n = k^2 + 1$ tuple with $\chi(A) > -\infty$ and $\max_a \|A_a\| \leq R$. For all sufficiently small $\varepsilon > 0$,*

$$\left| \chi(E_\varepsilon(A)) - k^{-2} \chi(A) - \frac{n}{k^2} \log \varepsilon - C_D \right| \leq C_{k,R} \varepsilon, \quad (19)$$

where

$$C_D = \log k + \frac{k^2 - 1}{k^2} c + \frac{\log c_D}{k^2}. \quad (20)$$

The constants and smallness threshold can be chosen uniformly for a fixed coordinate bound R . Both sides use (5).

4.1 Lower bound and preservation of source subsequences

Fix a source cutoff $R_1 > \max_a \|A_a\|$. A source microstate set $\Gamma_{R_1}(A; m, d, \gamma)$ is invariant under simultaneous $U(d)$ conjugation. It therefore defines an associated-bundle fiber. For any prescribed output moments and a sufficiently large fixed output cutoff, every point of the bundle over a sufficiently accurate source microstate set is an output microstate: substitution in (3) is polynomial, and conjugation preserves every trace word. Lemma 3.1 and (16) give Here $\Gamma_{\text{out},kd}$ denotes the prescribed output microstate neighborhood in size kd .

$$\text{Vol } \Gamma_{\text{out},kd} \geq A_d(\varepsilon\sqrt{k})^{nd^2} \text{Vol } \Gamma_{R_1}(A; m, d, \gamma) e^{-C\varepsilon d^2}. \quad (21)$$

Every subsequence realizing a source limsup gives output sizes kd . This suffices for an output lower limsup bound and loses no source subsequence. Substituting (18) cancels the $\log d$ terms, since $n + k^2 - 1 = 2k^2$. The orbit contribution $-n \log k/2$ cancels the fiber contribution $n \log k/2$. After division by $(kd)^2$, the output normalization $\log(kd)$ leaves (20). Taking infima over moments proves the lower half of (19). In particular $\chi(E_\varepsilon(A)) > -\infty$.

4.2 Rank perturbations and intrinsic recovery

The upper bound requires arbitrary output sizes N , not just multiples of k . We first give the rank estimate used in recovery.

Lemma 4.2 (Rank perturbation of moments). *Suppose two tuples a, b in M_N have coordinate norms at most $R_0 \geq 1$ and $\text{rank}(a_j - b_j) \leq m$ for every j . A word w of length L satisfies*

$$|\text{tr}_N(w(a) - w(b))| \leq 2LR_0^L m/N.$$

Compression by a projection of corank m changes any fixed trace word by $O_{L,R_0}(m/N)$, including the normalization change to the retained corner when $m/N \rightarrow 0$.

Proof. Telescope $w(a) - w(b)$ into L terms, each containing one difference. Each term has rank at most m and norm at most $2R_0^L$. The estimate $|\text{Tr } T| \leq \text{rank}(T)\|T\|$ proves the assertion. For compression, regard Pa_jP as a matrix on the original space; $a_j - Pa_jP$ has rank at most $2m$. Apply the first assertion and then $\text{tr}_{\text{rank } P} = (N/\text{rank } P)\text{tr}_N$ on the corner. $\square$

Lemma 4.3 (Uniform recovery from pair microstates). *Fix A, R as in Proposition 4.1 and a sufficiently small $\varepsilon > 0$. There is $R' > \max_a \|A_a\|$, depending only on k, R , with the following property. Given a finite output cutoff R_{out} , source parameters $m_0 \in \mathbb{N}, \gamma_0 > 0$, and $\theta > 0$, there are output parameters m, γ and N_0 such that every*

$$z \in \Gamma_{R_{\text{out}}}(E_\varepsilon(A); m, N, \gamma), \quad N \geq N_0,$$

has a projection P of rank $r = kd \geq (1 - \theta)N$ and exact matrix units on $P\mathbb{C}^N$ for which PzP is in the normal tube and its recovered coefficients lie in $\Gamma_{R'}(A; m_0, d, \gamma_0)$. The cutoff R' is independent of $N, \theta, R_{\text{out}}$.

Proof. We prove the sequential assertion first. Let $z^{(N)}$ be any sequence at a common output cutoff whose joint moments tend to those of $Z = E_\varepsilon(A)$. All $o(N)$ terms below are along this sequence.

Spectral bands and their traces. The actual Z_1 has k spectral bands around $1, \dots, k$. Along $D_1 + s\varepsilon \sum_a V_{a1} \otimes A_a$, $0 \leq s \leq 1$, these bands remain separated for small ε . Their Riesz projections vary continuously in norm. Projections at distance less than one are unitarily equivalent, by the polar decomposition of their intertwiner; subdivision of the path shows that the actual band projections p_i are equivalent to $E_{ii} \otimes 1$. Thus $\tau_{\mathcal{N}}(p_i) = 1/k$ exactly, even when the coefficient algebra is not a factor.

Choose slightly enlarged band intervals with boundaries in the actual gaps. The corresponding matrix spectral projections $p_i^{(N)}$ have ranks $N/k + o(N)$, and their complement has rank $o(N)$. Indeed bounded moment convergence implies weak convergence of spectral measures; closed gap neighborhoods have zero limiting mass. Smooth functions agreeing with the indicators off these neighborhoods differ from the spectral projections only on rank $o(N)$. Uniform polynomial approximation then transfers joint moments.

Tree-polar synchronization. In the actual algebra, the projections can be simultaneously identified with $E_{ii} \otimes 1$ by a unitary within $O_{k,R}(\varepsilon)$ of 1. For example, take the polar part of $\sum_i p_i(E_{ii} \otimes 1)$. In these coordinates every tree-edge block of Z_2 is within $O_{k,R}(\varepsilon)$ of 1, so its singular values lie in $[1/2, 2]$ after shrinking ε . Its polar unitary is also close to 1.

For the matrix edge block $y = p_i^{(N)} z_2 p_j^{(N)}$, take its polar partial isometry on singular values in a fixed, slightly larger interval, for example $[1/3, 3]$. The excluded initial and final ranks are $o(N)$ by convergence of the spectral distribution of y^*y in its band corner and the preceding rank approximation. The truncated polar map is bounded; on the good interval its formula is $y(y^*y)^{-1/2}$. A continuous extension of this bounded function approximates it in moments, with discrepancies confined to rank $o(N)$. No inversion near zero occurs.

Choose band 1 as root. Along each finite tree path pull the next good initial domain back to the root and intersect with the domains already retained. Each step loses at most the excluded rank, together with the $o(N)$ discrepancy of band ranks. A partial isometry is an isometry on its good domain, so pulling back does not increase this loss. The final root subspace has dimension $d_0 = N/k - o(N)$. Restricted path maps v_i are isometries from that root into mutually orthogonal band subspaces, with v_1 the root inclusion. Then $f_{ij} = v_i v_j^*$ are exact matrix units on their sum. The absence of cycles removes any consistency condition.

Root moments. The root entries $a_{ij}^{(s,N)} = v_i^* z_s v_j$ are uniformly bounded by R_{out} . Replace each spectral or polar operation by its continuous functional-calculus approximation. Every discarded-domain restriction changes the resulting bounded products by rank $o(N)$. The root intersection need not itself be a functional calculus: Lemma 4.2 applies directly to its compression. Consequently the joint d_0 -normalized moments of the root entries tend to the moments of the actual polar-coordinate entries, with the normalized root trace. Polynomial approximation supplies the continuous-function step, and $d_0/N \rightarrow 1/k$ supplies the trace normalization. In the actual algebra, these entries form a pair $D + O_{k,R}(\varepsilon)$.

A second root truncation gives norm control. For each real and imaginary Hermitian part h of each root entry $a_{ij}^{(s,N)} - (D_s)_{ij}1$, fix a positive bound $a_0 = C_{k,R}\varepsilon$ for the norm of its actual limiting counterpart. This bound may be positive even if that counterpart is zero. A continuous nonnegative function vanishing on $[-a_0, a_0]$ and equal to one outside $[-2a_0, 2a_0]$ shows that the bad spectral subspace of h has rank $o(N)$. Intersect all the good root subspaces and call the resulting projection Q . There are at most $2k^2$ Hermitian parts per coordinate, hence at most $4k^2$ in total; this is a fixed number before $N \rightarrow \infty$. Their rank losses therefore sum to $o(N)$.

Restrict *every* path map to this one Q . Exact matrix units remain, now on a corner of rank $kd = N - o(N)$. Since $Q \leq 1_{[-2a_0, 2a_0]}(h)$, $\|QhQ\| \leq 2a_0$; invariance of Q under h is unnecessary. Each complex entry deviation has norm at most $4a_0$, so a conservative bound for the whole pair is $4k^2a_0 = O_{k,R}(\varepsilon)$. Recompression changes each prescribed moment by $o(1)$ by Lemma 4.2.

Normal inversion and source-law identification. Decrease ε so that this last norm bound is strictly inside the analytic domain of Lemma 3.1. The normal inverse gives coefficient norms at most $C_k a_0/\varepsilon \leq C'_{k,R}$. Choose R' strictly larger than that bound and $\max \|A_a\|$. The actual polar-coordinate pair is a unitary conjugate of $E_\varepsilon(A)$ by a unitary close to the scalar stabilizer. The uniqueness in Lemma 3.1, valid over the actual coefficient algebra, identifies its recovered coefficients with a common unitary conjugate of A . Hence their joint law is that of A . The inverse series converges absolutely on a ball strictly containing both actual and truncated matrix entries. Uniform polynomial approximation to this series transfers every desired source moment. This proves the sequential recovery assertion and its fixed cutoff.

Finally, if the quantified statement failed for fixed $R_{\text{out}}, m_0, \gamma_0, \theta$, one could choose bad microstates with output accuracies $m_j \rightarrow \infty$, $\gamma_j \rightarrow 0$ and sizes $N_j \rightarrow \infty$. They form a bounded moment-convergent sequence, contradicting the sequential assertion. This proves uniformity over *all* sufficiently accurate output microstates. $\square$

4.3 The discarded-block volume and the ordinary limsup

We complete the upper bound of Proposition 4.1. Fix ε , an output cutoff, and a desired source neighborhood, leaving strict slack in its moment tolerances and cutoff. Let $r = kd$, $l = N - r$. Rank- l projections admit an operator-norm η -net of cardinality

$$\#\mathcal{P}_{N,l,\eta} \leq (C/\eta)^{2Nl}. \quad (22)$$

To prove this, represent them by N -by- l isometric frames, all in the operator-norm unit ball of a real $2Nl$ -dimensional normed space. A maximal separated set has the usual disjoint-ball volume bound; passing to projections changes only C .

For projections with $\|P - Q\| < \eta < 1/4$, the polar part w of $QP + (1 - Q)(1 - P)$ satisfies $wPw^* = Q$ and $\|w - 1\| \leq C\eta$. Thus QzQ differs in norm by at most $CR_{\text{out}}\eta$ from the transported recovered block $wPzPw^*$. The analytic inverse is locally Lipschitz, with constants allowed to depend on fixed ε ; choose η small enough that the recovered source coefficients of QzQ lie in the

prescribed relaxed source neighborhood. Strict slack keeps it inside the normal tube. Because the tube image is $U(r)$ -invariant, no choice of frame for Q matters. This covers the union of possible recovered corners without choosing a measurable recovery map and without estimating an uncontrolled thickening of a source microstate set.

For each net corner, Lemma 3.1 and the area formula bound its pair-block volume by

$$A_d(\varepsilon\sqrt{k})^{nd^2} \text{Vol} \Gamma_{R'}(A; m_0, d, \gamma_0) e^{C\varepsilon d^2}, \quad (23)$$

where the neighborhood on the right is the relaxed one. The remaining Hermitian pair coordinates have real dimension $M = 2(N^2 - r^2)$ and lie in a Euclidean ball of radius $R_{\text{out}}\sqrt{2N}$. Orthogonal decomposition of the ambient Hilbert–Schmidt space makes Lebesgue measure factor into the corner and these remaining coordinates. Cross blocks carry their induced factor-two metric; that metric is already included in this ball. The Euclidean ball formula and Stirling’s bound give, when $M > 0$,

$$\begin{aligned} \log \text{Vol}_{\text{rem}} &\leq -(N^2 - r^2) \log N \\ &\quad + C(N^2 - r^2) \left[1 + \log^+ R_{\text{out}} + \log \frac{N^2}{N^2 - r^2} \right]. \end{aligned} \quad (24)$$

For $M = 0$ the complementary volume is 1.

Multiply (23) by this bound and (22), then sum over the at most $N + 1$ possible ranks. Let $\theta' = 1 - r/N \leq \theta$. The source logarithm, orbit logarithm, and fiber factor combine into

$$-r^2 \log r + r^2 \left[k^{-2} h_d + \frac{n}{k^2} \log \varepsilon + C_D + O(\varepsilon) + o(1) \right],$$

where $h_d = d^{-2} \log \text{Vol} \Gamma_{R'}(A; m_0, d, \gamma_0) + (n/2) \log d$. The logarithmic size terms, including the output normalization, are

$$-r^2 \log r - (N^2 - r^2) \log N + N^2 \log N = r^2 \log(N/r).$$

After division by N^2 , all complementary, net, rank-sum, and normalization errors are bounded by

$$C [\theta \log(C/\theta) + \theta \log(C/\eta)] + o(1). \quad (25)$$

Constants here may depend on the fixed output cutoff and ε .

For fixed source parameters, the definition of a limsup bounds h_d above for every sufficiently large d by its limsup plus an arbitrary positive error. If that limsup is $-\infty$, use an arbitrary finite upper bound instead and then send it to $-\infty$. Every retained $d \geq (1 - \theta)N/k$ tends to infinity; thus this estimate applies to all output sizes, regardless of their residue modulo k . Lemma 4.3 allows θ to be arbitrarily small by choosing the output moment accuracy sufficiently strong.

Explicitly, hold ε , source accuracy, and output cutoff fixed; choose η for the source slack; choose output accuracy for a given θ ; take the limsup as $N \rightarrow \infty$; let $\theta \downarrow 0$; then let the net slack and source tolerance tend to zero and the source moment order to infinity. Lastly take the supremum over output cutoffs. The surviving constant C_D and $O_{k,R}(\varepsilon)$ error are independent of the output cutoff, since the recovered cutoff R' is. Cutoff independence gives $\chi_{R'}(A) = \chi(A)$. This proves the upper half of (19), and completes its proof.

5 Unconditional scores in the coarse bimodule

This section supplies the estimate that cannot be obtained by simply conditioning on scalar matrix units. Work in $\mathcal{N} = M_k(\mathcal{A})$ with normalized trace $\text{tr}_k \otimes \tau$. The coarse bimodule has outer actions

$x \cdot (a \otimes b) \cdot y = xa \otimes by$. Inner products use normalized traces. Tensor notation may equivalently use the opposite algebra on the second factor; the displayed outer-action convention fixes all matrix-unit orientations.

Let $O_1, \dots, O_{k^2-1}$ be a real orthonormal basis of $\mathcal{O}$. For bounded self-adjoint tuples $U = (U_a)_{a=1}^n$ and $T = (T_b)_{b=1}^{k^2-1}$, put

$$Z_s = D_s + \varepsilon \left(\sum_a V_{as} \otimes U_a + \sum_b O_{bs} \otimes T_b \right). \quad (26)$$

Proposition 5.1. *Suppose U has finite conjugate system ξ_a and is free from T . For a common coordinate norm bound R , and sufficiently small $\varepsilon > 0$, the pair (26) has unconditional conjugates and*

$$\Phi^*(Z) \leq \frac{\Phi^*(U)}{\varepsilon^2 k^4} + C_{k,R} \varepsilon^{-1} (1 + \Phi^*(U)). \quad (27)$$

5.1 The finite coarse decomposition

In $M_k \otimes M_k$ set

$$w_{bc} = \sum_{i=1}^k E_{ib} \otimes E_{ci}, \quad W_{a,bc} = (V_{a1} w_{bc}, V_{a2} w_{bc}).$$

For every scalar matrix H , $H w_{bc} = w_{bc} H$, by direct matrix-unit multiplication. Conversely, solving these equations for all E_{ij} shows that the w_{bc} span the scalar-central subspace. The same calculation, with normalized tensor traces, gives

$$\langle W_{a,bc}, W_{e,fg} \rangle = k^{-1} \delta_{ae} \delta_{bf} \delta_{cg}. \quad (28)$$

Let $J\eta = ([D_1, \eta], [D_2, \eta])$ in the coarse pair space. Since $\sum_s [D_s, V_{as}] = 0$, $J^* W_{a,bc} = 0$. The kernel of J is the scalar-central subspace, of complex dimension k^2 , because D_1, D_2 generate M_k . Hence

$$\text{rank } J = k^4 - k^2, \quad \dim \text{span}\{W_{a,bc}\} = k^2(k^2 + 1).$$

These dimensions add to $2k^4$. Thus J on its kernel complement and the span of the W give an orthogonal direct-sum isomorphism of the entire coarse pair space. This is a calculation in the *coarse* module, rather than an inference from the ordinary tangent count. Tensoring the finite coefficient identity with $\mathcal{A} \otimes \mathcal{A}$ preserves it.

5.2 Derivations, coefficient placement, and bilinear scores

Because U is free from T , its conjugates ξ_a are also conjugates relative to $\text{alg}(T)$. Existence of these conjugates makes the free difference quotients well defined on the represented algebra: the adjoint formula against polynomial elementary tensors makes the derivative of a zero polynomial orthogonal to a dense set, hence zero. This standard domain implication is part of the conjugate-variable calculus [8].

Define entry derivations $\Delta_{a,bc}$, vanishing on M_k and on T , with $\Delta_{a,bc}(U_e) = \delta_{ae} w_{bc}$. Compatibility with scalar-matrix relations follows from centrality of w_{bc} . On a matrix monomial $E_{ij} \otimes P(U, T)$, trace contraction of this derivative is $\delta_{ib} \delta_{jc} / k^2$ times the entry difference-quotient contraction. Throughout this section a *bilinear score* $\mathbf{b}(\delta)$ means a vector satisfying

$$(\text{tr}_k \otimes \tau)(\mathbf{b}(\delta)P) = ((\text{tr}_k \otimes \tau) \otimes (\text{tr}_k \otimes \tau))(\delta(P))$$

for every represented polynomial P . This is not Hilbert-adjoint notation. For the Hilbert inner product $\langle x, y \rangle = \tau(x^*y)$, a Hilbert adjoint at $1 \otimes 1$ gives $\mathfrak{b}(\delta) = (\delta_H^*(1 \otimes 1))^*$. Trace contraction of the entry derivation gives

$$\mathfrak{b}(\Delta_{a,bc}) = k^{-1} E_{cb} \otimes \xi_a. \quad (29)$$

The transpose and k^{-1} in this formula are essential. It is the same entry-adjoint method underlying [6, Proposition 4.1]; the normalization here uses real Hermitian coefficients orthonormal for tr_k .

For a coefficient $u \otimes v \in \mathcal{A} \otimes \mathcal{A}$, insert it on the *inner* legs of the derivative: the contribution at an occurrence PU_aQ is

$$\sum_i P(E_{ib} \otimes u) \otimes (E_{ci} \otimes v)Q. \quad (30)$$

This insertion commutes with both outer actions and therefore preserves the Leibniz rule. Arbitrary outer multiplication of an already formed derivative would not do so and is not used. Include also inner derivations $P \mapsto [P, \eta]$.

For precision, for polynomial coefficients $u, v \in \mathcal{A}$ define

$$\mathcal{C}_a(u, v) = v\xi_a u - v(\tau \otimes \text{id})(\partial_a u) - (\text{id} \otimes \tau)(\partial_a v)u. \quad (31)$$

If $\delta_{uv}(P) = \sum_{P=P_1U_aP_2} P_1u \otimes vP_2$, then

$$(\tau \otimes \tau)(\delta_{uv}(P)) = \tau(\mathcal{C}_a(u, v)P). \quad (32)$$

Indeed apply the conjugate relation to uPv , expand

$$\partial_a(uPv) = (\partial_a u)(1 \otimes Pv) + (u \otimes 1)(\partial_a P)(1 \otimes v) + (uP \otimes 1)(\partial_a v),$$

and move factors cyclically inside each scalar trace. The middle term has trace $\sum \tau(P_1u)\tau(vP_2)$. The other two terms give the two subtractions in (31), while $\tau(\xi_a uPv) = \tau(v\xi_a uP)$. This proves the identity without mixing bilinear and sesquilinear conventions. For the matrix insertion (30), the bilinear score is $k^{-1} E_{cb} \otimes \mathcal{C}_a(u, v)$, by the same entry contraction as in (29).

The resulting bound is

$$\|\mathcal{C}_a(u, v)\|_2 \leq \|u\| \|v\| \|\xi_a\|_2 + \|v\| \|\partial_a u\|_\pi + \|\partial_a v\|_\pi \|u\|.$$

Thus a sum of inserted coefficients has the same projective coefficient and differentiated-coefficient bounds used below. The reversed order of factors changes no norm bound. For an inner derivation, trace contraction gives the difference of the two partial traces of η ; its bilinear score has norm at most $2\|\eta\|_\pi$. Equivalently these are Hilbert-adjoint bounds after the explicit adjoint conversion above.

5.3 Exact inverse and the leading energy

Let $H = VU + OT$ denote the bounded perturbation in (26). For a sum of coefficient-inserted normal derivatives and an inner derivation, prescribe its two values on Z . Use unknowns $(\varepsilon c, \eta)$, where c collects all normal coefficients and η lies in the fixed complement of $\ker J$. The map to generator values is

$$\mathcal{L}_0 + \varepsilon \mathcal{B}_H. \quad (33)$$

Here $\mathcal{L}_0$ is exactly the finite direct-sum isomorphism from the preceding subsection. The normal derivations vanish on T and D , and their values on VU are constant, so the only perturbation is $\varepsilon[H, \eta]$.

Equip the finitely many coefficient arrays with sum projective tensor norm. If $K = \|\mathcal{L}_0^{-1}\|$ and $\|\mathcal{B}_H\| \leq H_R$, choose $\varepsilon K H_R \leq 1/2$. Then the Neumann series for (33) converges absolutely, has norm at most $2K$, and differs from $\mathcal{L}_0^{-1}$ by at most $2\varepsilon K^2 H_R$. Its entries are series of words in U, T on the two legs. Differentiating a degree- ℓ summand with respect to an entry U_a inserts a derivative at at most ℓ positions. The nonconstant part of the inverse therefore has differentiated norm bounded by a constant times

$$\sum_{\ell \geq 1} \ell \varepsilon (\varepsilon K H_R)^{\ell-1} = O_{k,R}(\varepsilon).$$

After undoing the scaling εc , the correction to the leading normal coefficient and its differentiated norm are $O_{k,R}(1)$; the inner coefficient and its norm are also $O_{k,R}(1)$. Apply (31)–(32) first to every finite Neumann sum, with the inner-derivation score just given. The coefficient and differentiated-coefficient series converge absolutely, so their bilinear scores are Cauchy in L^2 . For a fixed polynomial P , the inserted derivatives converge as well; passing to the limit in the trace pairing gives a bilinear score for the limiting derivation. The corresponding Hilbert adjoint exists by the stated adjoint conversion. This supplies the same remainder norm bound.

Apply the inverse to the target pair $(\delta_{js}1 \otimes 1)_{s=1}^2$ to obtain a derivation δ_j with $\delta_j(Z_s) = \delta_{js}1 \otimes 1$. It agrees on every polynomial in Z with its free difference quotient, by Leibniz. Project its bilinear score onto $L^2(W^*(Z))$. The trace pairing against every polynomial in Z is preserved by this conditional expectation, so the projection is the unconditional conjugate variable of Z_j .

The projection of the target pair onto the normal coarse subspace, computed from (28), is

$$\frac{1}{k} \sum_{a,b,c} (V_{aj})_{cb} W_{a,bc}.$$

Using (29) gives the leading unprojected score

$$\Xi_j^0 = \frac{1}{\varepsilon k^2} \sum_a V_{aj} \otimes \xi_a, \quad \sum_{j=1}^2 \|\Xi_j^0\|_2^2 = \frac{\Phi^*(U)}{\varepsilon^2 k^4}. \quad (34)$$

Indeed the cross terms sum to $\sum_j \text{tr}_k(V_{aj} V_{ej}) = \delta_{ae}$; no orthogonality among the ξ_a is assumed. The remainder estimate just proved is

$$\left(\sum_j \|\Xi_j - \Xi_j^0\|_2^2 \right)^{1/2} \leq C_{k,R} (1 + \sqrt{\Phi^*(U)}).$$

Squaring the triangle inequality and using contraction of orthogonal projection proves (27), absorbing the bounded square remainder into $C\varepsilon^{-1}(1 + \Phi^*(U))$ for $\varepsilon \leq 1$. This completes the proof of Proposition 5.1. The inner derivations supply the orbit component of the target; the normal leading scores alone would fail the generator tests.

6 The exact noisy path and analytic straightening

Return to the source X of (1), and write $Z_\varepsilon = E_\varepsilon(X)$. Let S^{pair} be a standard free semicircular pair free from the full ambient algebra.

6.1 Noise normalization and the nonmicrostates increment

A standard scalar semicircular pair over $M_k(\mathcal{A})$ can be realized by $2k^2$ standard free semicircular entry coefficients divided by k , in an orthonormal Hermitian pair basis for normalized tr_k . This is the classical free semicircular/circular matrix construction [9]. One way to check the normalization is the completeness identity $\sum_\alpha B_\alpha A B_\alpha = k \text{Tr}_k(A) I$ for a tr_k -orthonormal Hermitian basis. After division by k^2 the operator-valued covariance is the scalar normalized trace; higher free cumulants vanish. The two pair-coordinate covariances are orthogonal, yielding a standard free pair free from $M_k(\mathcal{A})$.

Orthogonal rotation to the bases V and O preserves a standard free semicircular family. Thus, exactly in distribution at $t = \varepsilon^2 k^2 u$,

$$Z_\varepsilon + \sqrt{t} S^{\text{pair}} = D + \varepsilon(VU(u) + OT(u)), \quad U(u) = X + \sqrt{u} S^n, \quad T(u) = \sqrt{u} S^o. \quad (35)$$

All coefficient semicirculars are free from X , and the normal and orbital families are free from one another. In particular $U(u)$ is free from $T(u)$. This is an exact linear identity of laws, so there is no use of upper semicontinuity of Fisher information or of a separate hidden copy of X .

Fix $U_0 < \infty$. The coefficient norms in (35) are uniformly bounded for $0 \leq u \leq U_0$, and $\Phi^*(U(u)) \leq \Phi^*(X) < \infty$. Proposition 4.1 and $\chi \leq \chi^*$, together with (9), show that both entropies of Z_ε are finite. For its noisy endpoints the Fisher bound is integrable on every such interval, and the standard heat identity applies. Proposition 5.1 and $dt = \varepsilon^2 k^2 du$ give

$$\begin{aligned} \chi^*(Z_\varepsilon + \varepsilon k \sqrt{U_0} S^{\text{pair}}) - \chi^*(Z_\varepsilon) &\leq \frac{1}{2k^2} \int_0^{U_0} \Phi^*(U(u)) du + C_{U_0} \varepsilon (1 + \Phi^*(X)) \\ &= k^{-2} (\chi^*(X + \sqrt{U_0} S^n) - \chi^*(X)) + O_{U_0}(\varepsilon). \end{aligned} \quad (36)$$

In particular the leading k^{-4} energy in (34) becomes k^{-2} after changing heat time. The error constant can depend on U_0 , which is fixed before $\varepsilon \downarrow 0$.

6.2 The precise power-series hypotheses

Fix a positive U_0 and abbreviate $U = U(U_0)$, $T = T(U_0)$. Apply the normal solver (14) to $W = \varepsilon(VU + OT)$. Its first-order term $K_1 = L^{-1} P_o W$ depends only on T , since $P_o VU = 0$. Consequently

$$e^{-iK}(D + \varepsilon(VU + OT))e^{iK} = D + \varepsilon V H_\varepsilon(U, T), \quad H_\varepsilon(U, T) = U + \varepsilon R_\varepsilon(U, T). \quad (37)$$

The remainder and its derivative in U have bounded absolute power-series majorant norms on fixed radii strictly larger than the coefficient norms. The cancellation of the first-order orbital part leaves the normal part $\varepsilon VU + O(\varepsilon^2)$, so division by ε causes no singular remainder.

Here are the hypotheses of [4, Theorem 4.1] in its own terms. Let $B = \text{alg}(T)$ be the unital algebra generated by the orbital tuple; its von Neumann closure is embeddable. Group consecutive T -letters in each monomial of the series in (U, T) . This expresses $F = H_\varepsilon$ as a noncommutative power series in the U letters with coefficients in B . The sum of products of coefficient norms is bounded by the original majorant evaluated at a T -radius exceeding $\|T\|$. Thus it has a multiradius strictly exceeding every $\|U_a\|$.

The inverse series is constructed, with a radius margin, by the contraction

$$A = H - \varepsilon R_\varepsilon(A, T). \quad (38)$$

Choose a coefficient bound larger than both the evaluated U and H norms, then a strictly larger majorant radius. Uniform bounds on the remainder and its derivative make (38) a contraction in the associated Banach algebra of convergent series for small ε . Iteration yields G ; uniqueness proves both formal identities $F \circ G = \text{id}$ and $G \circ F = \text{id}$, as well as their analytic identities. The radii for F and G can be chosen with a common strict margin beyond the evaluated norms, as required by the printed condition $\|U_a\| < \min(r_a, R_a)$.

Both F and G preserve self-adjointness. The normal solver operates in the self-adjoint part, V is a real Hermitian basis, and coefficient extraction is the scalar partial trace $\sum_s (\text{tr}_k \otimes \text{id})(V_{as} Y_s)$. This is self-adjoint for Hermitian inputs because the scalar matrix factor can be moved through the partial trace. The same holds for the fixed point inverse.

Finally the coefficient-algebra Jacobian satisfies

$$D_B F(U) = I + O_{U_0}(\varepsilon) \quad \text{in } M_n(\mathcal{A} \bar{\otimes} \mathcal{A}^{\text{op}}).$$

The projective majorant estimate implies this operator-norm estimate. The singular values lie between $1 - C\varepsilon$ and $1 + C\varepsilon$, hence

$$|(\text{Tr}_n \otimes \tau \otimes \tau) \log |D_B F(U)|| \leq C_{U_0} n \varepsilon. \quad (39)$$

Since n is fixed it is absorbed into C_{U_0} . This verifies the convergent inverse, strict multiradius, self-adjointness, and logarithmic Jacobian requirements; a merely formal inverse would not suffice.

6.3 The microstates increment

Theorem 2.19 of [4] applies because U is free from B and $W^*(B)$ is the embeddable semicircular algebra. In this particular application T has bounded matrix models at every sufficiently large size. Uniform asymptotic freeness of bounded deterministic U models with independent GUE models for T , followed by averaging over T , provides a conditioning model whose good U -fiber has proportion tending to one. This explains directly why free conditioning retains every source limsup subsequence here; it does not require aligning two unrelated maximizing subsequences.

Using Theorems 2.19 and 4.1, then Proposition 2.8 and projection of joint microstate sets, yields

$$\chi(H_\varepsilon(U, T)) \geq \chi(H_\varepsilon(U, T) \mid B) = \chi(U) + (\text{Tr}_n \otimes \tau \otimes \tau) \log |D_B F(U)| \geq \chi(U) - C_{U_0} \varepsilon. \quad (40)$$

This inequality does not use the equality left to the reader in Proposition 2.12 of [4]. Also $\chi(U) > -\infty$ by (11), since X is embeddable. The coefficient norm bound for H_ε is uniform for small ε .

Unitary conjugation in the tracial algebra preserves the joint law, so (37) and the lower half of Proposition 4.1 give a lower bound at the noisy endpoint. Use its upper half at X . Their identical constants C_D and $(n/k^2) \log \varepsilon$ cancel, giving

$$\begin{aligned} \chi(Z_\varepsilon + \varepsilon k \sqrt{U_0} S^{\text{pair}}) - \chi(Z_\varepsilon) \\ \geq k^{-2} (\chi(X + \sqrt{U_0} S^n) - \chi(X)) - O_{U_0}(\varepsilon). \end{aligned} \quad (41)$$

Only a relative *microstates* change-of-variables theorem was used. The nonmicrostates estimate is the separate unconditional construction in Section 5.

7 Gap transfer and finiteness

Let $g = \mathcal{G}(X) \geq 1/2$ and $X_u = X + \sqrt{u} S^n$. Every entropy in (36) and (41) is finite: the lower bounds just proved supply finite microstates entropy, while comparison and covariance bound the

nonmicrostates entropy below and above. For the noisy pair, $\chi \leq \chi^*$; therefore subtraction of the two finite increment inequalities gives

$$\mathcal{G}(Z_\varepsilon) \geq k^{-2}(g - \mathcal{G}(X_{U_0})) - C_{U_0}\varepsilon. \quad (42)$$

This final comparison is the standard two-noise-scale strictness pattern used in [1]. The work here supplies the two unconditional estimates to which that pattern is applied.

Embeddability and (11) give $\chi(X_u) \geq (n/2)\log(2\pi eu)$ for $u > 0$. If $C_X = n^{-1}\sum_a \text{Var}(X_a)$, covariance and comparison give

$$0 \leq \mathcal{G}(X_u) \leq \frac{n}{2} \log \frac{u + C_X}{u} \longrightarrow 0. \quad (43)$$

For each fixed U_0 let $\varepsilon \downarrow 0$ in (42), and only then let $U_0 \rightarrow \infty$. This proves (4). For an explicit order of existential choices, first choose U_0 so that $\mathcal{G}(X_{U_0}) < g/2$; then choose $\varepsilon > 0$ below all geometric and analytic thresholds and so small that $C_{U_0}\varepsilon < g/(4k^2)$. Equation (42) then gives $\mathcal{G}(Z_\varepsilon) > g/(4k^2) > 0$. Proposition 4.1 gives $\chi(Z_\varepsilon) > -\infty$, and (9) gives $\chi^*(Z_\varepsilon) < \infty$. The tensor trace on $M_k(W^*(X))$ is faithful and normal, and both coordinates of (3) are bounded and self-adjoint. This proves Theorems 1.2 and 1.1.

Every matrix-size limit in Section 4 precedes this ε limit. In particular no uniformity of microstate neighborhoods as $\varepsilon \downarrow 0$, no replacement of a limsup by a liminf in (5), and no ultrafilter are part of the argument.

8 Relation to prior operations and scope of comparison

8.1 Orbit volumes and supplied matrix-unit compression

The polar measure factorization of [5, Lemma 2.6], its unitary-volume asymptotic in Lemma 2.8, and its upper/lower volume patterns in Lemmas 2.9 and 2.10 are antecedents of the geometric bookkeeping here. The last lower-packing lemma has free Haar and positive-variable hypotheses. Neither that special-case theorem nor the classical formula (17) by itself proves the intrinsic recovery statement of Lemma 4.3. Uniform normal-tube geometry is supporting analysis, not a separate novelty claim. The specific additional upper-volume operation is recovery from the pair's own microstates with a common-root truncation and vanishing-rank accounting, valid for every N with the ordinary limsup and the same C_D .

Shlyakhtenko [4, Section 5, Theorem 5.1] already uses the corner/compression/packing pattern. Its proof takes a corner $p \leq e_{11}$ with $[e_{11}, p] = 0$ under *supplied* matrix-unit conditioning models, compresses with the map T made from those models, and packs entry microstates back by a measure-preserving map. This pattern is credited here. Its equality is an ultrafilter entropy identity, with its specified dilated ultrafilter. The displayed ordinary-limsup statement is one-sided, and its proof is explicitly left to the reader. The second displayed intermediate claim repeats the first inequality direction; the following packing proof establishes the reverse direction, a printed direction error. We use neither that misprinted direction nor an inferred ordinary-limsup equality. The upper bound here instead recovers its matrix units intrinsically and handles all output ranks.

Nica–Shlyakhtenko–Speicher [6, Proposition 4.1] proves matrix-entry Fisher identities relative to the scalar matrix algebra, and unconditional projection inequalities. Its raw complex-entry normalization has a factor k^3 . Formula (34) uses real Hermitian coefficients normalized for tr_k , yielding k^4 in the score energy. The classical entry-adjoint method and orthogonal Gaussian rotation are used openly. The additional unconditional estimate combines normal entry derivatives with inner derivatives and controls their adjoints; it is not obtained by deleting matrix-unit conditioning from a relative identity.

8.2 Why relative identities do not prove the finite gap

At $Z_\varepsilon = D + \varepsilon V X$, its $2k^2$ real entry coordinates lie in an affine subspace of dimension $k^2 + 1$. Thus there are $k^2 - 1$ independent real-linear relations. Entry microstates entropy is $-\infty$, including each ultrafilter version. Under the ultrafilter identity in [4, Theorem 5.1], the corresponding relative matrix microstates entropy is $-\infty$. The one-sided ordinary-limsup statement, whose lower side is then $-\infty$, does *not* compute the ordinary-limsup relative value.

On the Fisher side, the paid relative matrix-entry identity on the positive noisy path has normal and orbital contribution

$$\Phi^*(Z_\varepsilon + \sqrt{t}S^{\text{pair}} : M_k) = \frac{\Phi^*(U(u)) + (k^2 - 1)/u}{\varepsilon^2 k^4}, \quad t = \varepsilon^2 k^2 u.$$

Its orbital term is $((k^2 - 1)/k^2)t^{-1}$, nonintegrable at zero. The relative nonmicrostates entropy at the unperturbed pair is therefore $-\infty$. Between positive source noise times u_0 and U_0 , this orbital term contributes

$$\frac{k^2 - 1}{2k^2} \log(U_0/u_0).$$

Thus the relative identities at the target are degenerate. Where the relative quantities are $-\infty$ and the unconditional ones finite, conditioning-removal excesses are $+\infty$. Subtracting such excesses is not a proof of a finite unconditional gap. In the ordinary-limsup microstates case the cited one-sided theorem leaves even that relative value undetermined. Our proof never subtracts infinities; it compares direct unconditional finite increments.

8.3 Generation, source scope, and limits of the screen

The pinned preprint [3], Theorem `thm:main`, gives two self-adjoint W^* -generators for each separable-predual II_1 factor. Its introduction credits the classical 2×2 matrix-generation techniques of Douglas–Percy and Wogen, using Dykema–Sinclair–Smith–White Proposition 1.1. Its own Corollary `cor:entropy-generators` shows that the free entropy dimensions δ, δ_0 depend on the generating tuple. That theorem supplies generation only; it supplies no finite χ/χ^* -gap invariance under replacement of generators. The companion [2], in its amplification and entropy-dimension sections, transports tuples coordinatewise through trace-preserving isomorphisms. This preserves their joint law and length. It also does not reduce the number of entropy coordinates. We assert no generation property of Z_ε and use no factor-isomorphism invariance of entropy.

The source [1] supplies the many-variable finite gap, finite initial Fisher information, and bounded matrix models. The Gaussian lower bound, covariance maximum, comparison $\chi \leq \chi^*$, relative analytic change of variables, and two-scale strictness pattern are all credited inputs or classical deductions. The claimed additional mechanism is precisely the unconditional Fisher bound, the matched all-size upper volume bound, and their coupling.

The comparison is bounded. Operative bodies of the three primary papers [4, 5, 6] and the specified pinned source and companion statements were read. Dimension, orbital-free-entropy, and free-mutual-information literature remains a further screen, including Jung’s [math/0410594](#), Shlyakhtenko’s [2201.09503](#), Hiai–Miyamoto–Ueda’s [math/0702745](#), Ueda’s [1210.6421](#) and [1610.04118](#), Izumi–Ueda’s [1306.5372](#), and the Connes–Shlyakhtenko framework of dimensions and inner derivations. Their operative theorem bodies were not screened for this manuscript, so no claim that these papers lack a relevant transfer theorem is made. This limitation concerns literature priority; none is invoked as a mathematical premise above.

A Validation, provenance, and rights

This paper is a self-contained presentation of the transfer proof, with published and stipulated inputs identified in Section 2. The source revision is

fd4aeeb2ee4fc729c18d98444fed42fd0529eeeb

of [OpenAI Math](#). The accompanying local package records source paths, statement labels, and hashes, the frozen author packet, its independent proof and contribution reviews, and both attribution addenda. Reviewer completions are integrated into the rank lemma, the quantified recovery lemma, the local-coordinate Jacobian proof, and the literal power-series hypothesis check. The old exploratory status text in working notes is not part of the theorem or proof.

The frozen proof packet received the independent verdict `PASS_FULL_T298P`; the contribution review classified its mechanism as substantial conditional on correctness and required four attribution clarifications. Section 8 incorporates those clarifications. These records concern the frozen analytic proof; they do not constitute an independent review of this typeset manuscript or a certification of worldwide novelty. This manuscript is an author submission for that separate review.

A companion Lean package has four support declarations and four counterexample controls. It proves a real inner-product coefficient isometry assuming a column Gram identity, its scaled leading-score energy identity, a scalar gap rearrangement assuming both analytic increment bounds and endpoint comparison, and strictness under an explicit strict error bound. Its controls demonstrate failures with k^2 in place of k^4 , missing endpoint comparison, a reversed source gap sign, or a non-strict error threshold. It does not formalize conjugate variables, analytic inverse series, microstate recovery, entropy limits, the source theorem, or the theorem of this paper. No full formalization is claimed. The scalar support statements quantify over finite real numbers, not extended-real entropy functions. Their reported axiom dependencies are `propext`, `Classical.choice`, and `Quot.sound`. Reproduction builds fresh support and control oleans while trusting the installed Lean toolchain and pre-existing pinned dependency oleans. It rebuilds neither the compiler nor mathlib; Git source cleanliness does not authenticate those cached binaries. The independently reviewed supplement checks ten additional concrete examples and five expected rejected mutations, without enlarging the analytic coverage.

Finite symbolic and numerical diagnostics in the earlier proof package check matrix coefficients, the coarse Gram identities, noise covariance, metric multiplicities, and finite recovery examples. Such diagnostics are support for calculations, not proofs of the entropy limit theorem. The local reproducibility package builds this PDF from its complete LaTeX source, records the compiler and PDF hashes, and supplies a read-only externally anchored integrity verifier. Integrity checks and visual inspection verify files and layout, not mathematical truth. The construction is existential at the fixed large value $k = 2^{32}$; no useful quantitative ε or low-dimensional numerical witness has been produced.

The author is Myeongjin Kang, with no institutional affiliation stated. AI assistance was used in research drafting, review workflows, and typesetting; the detailed local provenance records the role boundaries. Original manuscript prose and LaTeX are licensed under [CC BY 4.0](#). Original reproducibility code is licensed under MIT. Third-party papers, theorem statements, software, and earlier packet materials retain their own rights and are credited separately; the licenses here do not relicense them. No third-party full paper is redistributed in this manuscript package.

References

- [1] OpenAI, *A finite-entropy separation of microstates and nonmicrostates free entropy*, September 25, 2026. Pinned OpenAI Math preprint, main theorem, Sections 2, 4–6. [Immutable source](#).
- [2] OpenAI, *An isomorphism of the free group factors*, September 23, 2026. Pinned OpenAI Math preprint, amplification and entropy-dimension sections. [Immutable source](#).
- [3] OpenAI, *Relative generation and the generator problem for finite factors*, September 23, 2026. Pinned OpenAI Math preprint, introduction Theorem `thm:main`; consequences Corollary `cor:entropy-generators`. [Immutable theorem source](#).
- [4] D. Shlyakhtenko, *A microstates approach to relative free entropy*, [arXiv:math/9812079v1](#) (1998). Definition 2.3, Proposition 2.8, Theorems 2.19 and 4.1, Proposition 4.2, and Section 5. Version-specific omissions and typos are discussed in text.
- [5] A. Nica, D. Shlyakhtenko, and R. Speicher, *Maximality of the microstates free entropy for R -diagonal elements*, [arXiv:math/9809081v1](#) (1998). Lemmas 2.6, 2.8–2.10 and Theorem 3.1.
- [6] A. Nica, D. Shlyakhtenko, and R. Speicher, *Some minimization problems for the free analogue of the Fisher information*, [arXiv:math/9809080v1](#) (1998). Proposition 4.1.
- [7] D. Voiculescu, *The analogues of entropy and of Fisher’s information measure in free probability theory. II*, Invent. Math. **118** (1994), 411–440. [doi:10.1007/BF01231539](#).
- [8] D. Voiculescu, *The analogues of entropy and of Fisher’s information measure in free probability theory. V. Noncommutative Hilbert transforms*, Invent. Math. **132** (1998), 189–227. [doi:10.1007/s002220050222](#).
- [9] D. Voiculescu, *Circular and semicircular systems and free product factors*, in *Operator Algebras, Unitary Representations, Enveloping Algebras, and Invariant Theory*, Progress in Mathematics 92, Birkhäuser (1990), 45–60. Classical semicircular matrix construction; also recalled in [5, 6].
- [10] P. Biane, M. Capitaine, and A. Guionnet, *Large deviation bounds for matrix Brownian motion*, Invent. Math. **152** (2003), 433–459. [doi:10.1007/s00222-002-0281-4](#).
- [11] D. Jekel and J. Pi, *An elementary proof of the inequality $\chi \leq \chi^*$ for conditional free entropy*, Doc. Math. **29** (2024), 1085–1124. [doi:10.4171/DM/969](#).
- [12] M. S. Marinov, *Invariant volumes of compact groups*, J. Phys. A **13** (1980), 3357–3366. The normalization used here is the $\text{Re Tr}(ab^*)$ convention of [5, Lemma 2.8].