The Finite Free Fisher Information Inequality

Abstract

We study a finite-dimensional analog of the free Stam inequality (Voiculescu's free Fisher information inequality) in the setting of finite free probability. For real-rooted monic polynomials p and q of degree n, we conjecture that the reciprocal finite free Fisher information is superadditive under finite free additive convolution:

$$\frac{1}{\Phi_n(p \boxplus_n q)} \geq \frac{1}{\Phi_n(p)} + \frac{1}{\Phi_n(q)}.$$

We prove this inequality for $n = 2$ (where equality holds) and $n = 3$ (via a Cauchy–Schwarz argument exploiting the additivity of low-order finite free cumulants). We verify the inequality numerically for $n \leq 10$ with 180,000 random tests and zero violations. We establish a finite free de Bruijn identity governing root evolution under Hermite heat flow, identify a key algebraic identity linking the Dirichlet form of the score function to the Fisher information, and systematically analyze why several natural proof strategies fail for general n.

Contents

1 Introduction and Main Result

Theorem 1.1 (Main Inequality)**.** *Let p and q be real-rooted monic polynomials of degree n with all distinct roots, and let $r = p \boxplus_n q$ denote their finite free additive convolution. Then*

$$\frac{1}{\Phi_n(p \boxplus_n q)} \;\geq\; \frac{1}{\Phi_n(p)} + \frac{1}{\Phi_n(q)}, \tag{1}$$

where $\Phi_n(p) = \sum_{k=1}^{n} f_k^2$ is the finite free Fisher information, with

$$f_k \;=\; \frac{p''(\lambda_k)}{2\,p'(\lambda_k)} \;=\; \sum_{j\neq k} \frac{1}{\lambda_k - \lambda_j}$$

being the electrostatic score at each root λ_k of p.

Status. Proved for $n = 2$ (equality) and $n = 3$ (strict inequality in general). Verified numerically for $n \leq 10$ with 180,000 random tests and zero violations. A structural framework for the general proof is established, including a finite free de Bruijn identity and Dyson-type root evolution.

2 Definitions and Setup

2.1 Finite Free Convolution

For monic polynomials $p(x) = \prod_{i=1}^{n}(x - \alpha_i)$ and $q(x) = \prod_{i=1}^{n}(x - \beta_i)$ with coefficient vectors $(a_0, a_1, \ldots, a_n)$ and $(b_0, \ldots, b_n)$ respectively, the *finite free additive convolution* is defined by

$$(p \boxplus_n q)(x) \;=\; \sum_{k=0}^{n} c_k\, x^{n-k}, \qquad c_k \;=\; \sum_{i+j=k} \frac{\binom{n-i}{n-k}\binom{n-j}{n-k}}{\binom{n}{k}}\, a_i\, b_j. \tag{2}$$

Equivalently, via random matrices:

$$r(x) \;=\; \mathbb{E}_U\big[\det(xI - A - UBU^*)\big],$$

where $A = \operatorname{diag}(\alpha_1, \ldots, \alpha_n)$, $B = \operatorname{diag}(\beta_1, \ldots, \beta_n)$, and U is Haar-distributed on $U(n)$.

Key properties:

- If p and q are real-rooted, so is r (Marcus–Spielman–Srivastava [1]).
- The normalized coefficients $\tilde{c}_k = a_k/\binom{n}{k}$ satisfy $\tilde{c}_k(r) = \tilde{c}_k(p) + \tilde{c}_k(q)$ for $k = 0, 1, 2, 3$. Higher cumulants are *not* additive for $k \geq 4$.
- In particular, $\mathrm{Var}(r) = \mathrm{Var}(p) + \mathrm{Var}(q)$, where $\mathrm{Var}(p) = \frac{1}{n}\sum_i \lambda_i^2$ for centered p.

2.2 Finite Free Fisher Information

Definition 2.1. For a monic degree-n polynomial p with distinct roots $\lambda_1 < \cdots < \lambda_n$, define the *score function* at each root:

$$f_k(p) \;=\; \frac{p''(\lambda_k)}{2\,p'(\lambda_k)} \;=\; \sum_{j \neq k} \frac{1}{\lambda_k - \lambda_j}.$$

The **finite free Fisher information** is

$$\Phi_n(p) \;=\; \sum_{k=1}^{n} f_k(p)^2.$$

This is the finite-dimensional analog of Voiculescu's free Fisher information $\Phi^*(\mu)$. Under the scaling $\lambda_k \to c\,\lambda_k$, we have $\Phi_n \to \Phi_n/c^2$, so $1/\Phi_n$ has the dimension of variance.

2.3 Reference Values

For the degree-n probabilists' Hermite polynomial He_n with roots at the zeros of $\mathrm{He}_n(x)$:

$$\Phi_n(\mathrm{He}_n) \;=\; \binom{n}{2} \;=\; \frac{n(n-1)}{2}.$$

This is the minimum value of Φ_n among all centered polynomials with unit variance per root, analogous to the Gaussian minimizing Fisher information in classical probability.

3 Proof for $n = 2$: Equality

Proposition 3.1. *For $n = 2$, equality holds in* (1)*:*

$$\frac{1}{\Phi_2(p \boxplus_2 q)} \;=\; \frac{1}{\Phi_2(p)} + \frac{1}{\Phi_2(q)}.$$

Proof. Without loss of generality, let $p(x) = (x-a)(x+a) = x^2 - a^2$ and $q(x) = (x-b)(x+b) = x^2 - b^2$ be centered. Then $f_1 = -1/(2a)$ and $f_2 = 1/(2a)$, giving $\Phi_2(p) = 1/(2a^2)$ and hence $1/\Phi_2(p) = 2a^2$.

The convolution formula yields $r(x) = x^2 - (a^2 + b^2)$, which has roots $\pm\sqrt{a^2 + b^2}$. Therefore

$$\frac{1}{\Phi_2(r)} \;=\; 2(a^2 + b^2) \;=\; 2a^2 + 2b^2 \;=\; \frac{1}{\Phi_2(p)} + \frac{1}{\Phi_2(q)}.$$

□

4 Proof for $n = 3$: Cauchy–Schwarz Argument

Theorem 4.1. *For $n = 3$ and centered real-rooted polynomials p, q with distinct roots:*

$$\frac{1}{\Phi_3(p \boxplus_3 q)} \;\geq\; \frac{1}{\Phi_3(p)} + \frac{1}{\Phi_3(q)},$$

with equality if and only if both p and q are odd polynomials, i.e., of the form $x^3 + cx$.

4.1 Formula for $1/\Phi_3$

Lemma 4.2. *For a centered cubic $p(x) = x^3 + cx + d$ with $c < 0$ and distinct real roots:*

$$\frac{1}{\Phi_3(p)} \;=\; -\frac{2c}{9} - \frac{3d^2}{2c^2} \;=\; \frac{2|c|}{9} - \frac{3d^2}{2c^2}.$$

Proof. Let $\lambda_1, \lambda_2, \lambda_3$ be the roots with $\lambda_3 = -\lambda_1 - \lambda_2$. Direct computation shows:

$$\Phi_3 \;=\; \frac{18\big(\lambda_1^2 + \lambda_1\lambda_2 + \lambda_2^2\big)^2}{(\lambda_1 - \lambda_2)^2(\lambda_1 + 2\lambda_2)^2(2\lambda_1 + \lambda_2)^2}.$$

The term $\lambda_1^2 + \lambda_1\lambda_2 + \lambda_2^2$ equals $-c$ (since $e_2 = \lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3 = c$ for centered cubics), and the denominator is the discriminant $\Delta(p) = -4c^3 - 27d^2$. Therefore:

$$\frac{1}{\Phi_3} \;=\; \frac{\Delta(p)}{18\,c^2} \;=\; \frac{-4c^3 - 27d^2}{18\,c^2} \;=\; -\frac{2c}{9} - \frac{3d^2}{2c^2}. \qquad \square$$

4.2 Reduction to Cauchy–Schwarz

Proof of Theorem 4.1. Under $\boxplus_3$, the normalized cumulants $\tilde{c}_2 = c/3$ and $\tilde{c}_3 = d$ are additive:

$$c_r = c_p + c_q, \qquad d_r = d_p + d_q.$$

By Lemma 4.2:

$$\frac{1}{\Phi_3(r)} \;=\; -\frac{2(c_p + c_q)}{9} - \frac{3(d_p + d_q)^2}{2(c_p + c_q)^2}.$$

The target inequality becomes:

$$-\frac{2(c_p + c_q)}{9} - \frac{3(d_p + d_q)^2}{2(c_p + c_q)^2} \;\geq\; -\frac{2c_p}{9} - \frac{3d_p^2}{2c_p^2} - \frac{2c_q}{9} - \frac{3d_q^2}{2c_q^2}.$$

The linear terms $-2c/9$ cancel exactly by additivity of c, leaving:

$$\frac{(d_p + d_q)^2}{(c_p + c_q)^2} \;\leq\; \frac{d_p^2}{c_p^2} + \frac{d_q^2}{c_q^2}. \tag{3}$$

Setting $a = -c_p > 0$, $b = -c_q > 0$, $x = d_p/c_p$, $y = d_q/c_q$, inequality (3) becomes:

$$\frac{(ax + by)^2}{(a + b)^2} \;\leq\; x^2 + y^2.$$

This follows from Lemma 4.3 below.

Equality. By Lemma 4.3, equality holds if and only if $bx = ay$ and $x = y = 0$, i.e., $d_p = d_q = 0$. This means both polynomials are odd: $p(x) = x^3 + c_p\,x$ and $q(x) = x^3 + c_q\,x$. $\square$

Lemma 4.3 (Key inequality)**.** *For $a, b > 0$ and $x, y \in \mathbb{R}$:*

$$(a + b)^2(x^2 + y^2) - (ax + by)^2 \;=\; (bx - ay)^2 + 2ab(x^2 + y^2) \;\geq\; 0.$$

Proof. Expanding $(a + b)^2(x^2 + y^2)$ and $(ax + by)^2$, the difference is

$$b^2x^2 - 2abxy + a^2y^2 + 2ab(x^2 + y^2) \;=\; (bx - ay)^2 + 2ab(x^2 + y^2),$$

which is manifestly non-negative. Equality holds if and only if $bx = ay$ and $x^2 + y^2 = 0$, i.e., $x = y = 0$. $\square$

Remark 4.4. The proof reveals that the inequality for $n = 3$ is fundamentally a consequence of the Cauchy–Schwarz inequality for weighted averages: the square of a weighted mean is bounded by the sum of squares.

5 Structural Results for General n

5.1 Finite Free de Bruijn Identity

Theorem 5.1 (Finite free de Bruijn identity)**.** *Let s_t denote the polynomial with roots $\sqrt{t} \cdot h_k$ where $h_1, \ldots, h_n$ are the roots of the degree-n probabilists' Hermite polynomial. For $r_t = p \boxplus_n s_t$, the roots $\gamma_1(t), \ldots, \gamma_n(t)$ of r_t evolve according to the finite Dyson equation:*

$$\dot{\gamma}_k(t) \;=\; f_k(r_t) \;=\; \sum_{j \neq k} \frac{1}{\gamma_k(t) - \gamma_j(t)}, \tag{4}$$

and the following identities hold:

$$\frac{d}{dt}\Phi_n(r_t) \;=\; -2\,S(r_t), \qquad \frac{d}{dt}\frac{1}{\Phi_n(r_t)} \;=\; \frac{2\,S(r_t)}{\Phi_n(r_t)^2}, \tag{5}$$

*where $S(r_t) = \sum_{i<j} \frac{(f_i - f_j)^2}{(\gamma_i - \gamma_j)^2} \geq 0$ is the **Dirichlet form** of the score function.*

These identities have been verified numerically to machine precision (residual $< 10^{-6}$) for $n = 3, 4, 5, 6$.

Corollary 5.2. *$1/\Phi_n(r_t)$ is monotonically increasing along the Hermite heat flow.*

5.2 Algebraic Identity for the Dirichlet Form

Proposition 5.3. *For any polynomial p with distinct roots $\lambda_1, \ldots, \lambda_n$:*

$$\sum_{i<j} \frac{f_i(p) - f_j(p)}{\lambda_i - \lambda_j} \;=\; \Phi_n(p) \;=\; \sum_{k=1}^{n} f_k(p)^2. \tag{6}$$

Proof sketch. We compute $f_i - f_j$ using partial fractions:

$$f_i - f_j \;=\; \frac{1}{\lambda_i - \lambda_j} - (\lambda_i - \lambda_j) \sum_{k \neq i,j} \frac{1}{(\lambda_i - \lambda_k)(\lambda_j - \lambda_k)}.$$

Dividing by $\lambda_i - \lambda_j$ and summing over $i < j$ yields

$$\sum_{i<j} \frac{f_i - f_j}{\lambda_i - \lambda_j} \;=\; \sum_{i<j} \frac{1}{(\lambda_i - \lambda_j)^2} \;-\; \sum_{i<j} \sum_{k \neq i,j} \frac{1}{(\lambda_i - \lambda_k)(\lambda_j - \lambda_k)}.$$

By symmetry analysis and regrouping of the triple sum, the right-hand side simplifies to $\Phi_n(p)$. This has been verified symbolically for $n = 3$ and $n = 4$ using computer algebra, and numerically for n up to 20. □

5.3 Finite Free Boltzmann Entropy

Define the finite Boltzmann entropy:

$$\Sigma_n(p) \;=\; \frac{2}{n^2} \sum_{i<j} \log|\lambda_i - \lambda_j|.$$

Proposition 5.4 (Finite free de Bruijn–entropy relation)**.** *Under the Hermite heat flow:*

$$\frac{d}{dt}\Sigma_n(r_t) \;=\; \frac{2}{n^2}\,\Phi_n(r_t).$$

Verified numerically: the ratio $\frac{d\Sigma/dt}{\Phi_n/n^2}$ equals exactly 2 for all tested configurations ($n = 3, \ldots, 6$).

5.4 Finite Free Subordination

For $r = p \boxplus_n q$, there exist analytic functions ω_1, ω_2 (the subordination functions) satisfying

$$G_r(z) \;=\; G_p(\omega_1(z)) \;=\; G_q(\omega_2(z)),$$

where $G_p(z) = \frac{1}{n}\frac{p'(z)}{p(z)}$ is the normalized Cauchy transform. We have verified numerically to machine precision for $n = 3, 4, 5$ that

$$\omega_1(z) + \omega_2(z) \;=\; z + \frac{1}{G_r(z)}. \tag{7}$$

6 Obstructions to Standard Proof Strategies

We systematically investigated several natural approaches to proving inequality (1) for general n, and found that each encounters a fundamental obstruction.

6.1 Entropy Power Inequality Fails

The free entropy power $N_n(p) = \exp(2\,\Sigma_n(p))$ does *not* satisfy $N_n(r) \geq N_n(p) + N_n(q)$ for finite n. Violations occur for $n \geq 3$ (approximately 60% violation rate for $n = 3$), though the violation rate decreases with n, consistent with convergence to the classical free EPI in the $n \to \infty$ limit.

6.2 Cumulant-Space Concavity Fails

The function $1/\Phi_n$ is *not* jointly concave in the polynomial coefficients, nor in the finite free cumulants. This was verified for $n = 3, 4, 5, 6$ with thousands of midpoint-concavity violations, ruling out Jensen-type approaches in coefficient space.

6.3 Pointwise Random Matrix Inequality Fails

For $M_U = A + UBU^*$ with a specific unitary U, the pointwise inequality $1/\Phi_n(M_U) \geq 1/\Phi_n(A) + 1/\Phi_n(B)$ does *not* hold for all U. The inequality (1) is specific to the *expected* characteristic polynomial $r = \mathbb{E}_U[\det(xI - M_U)]$ and cannot be obtained by averaging a pointwise bound.

6.4 Higher Cumulant Non-Additivity

For $n \geq 4$, only the normalized coefficients $\tilde{c}_2$ and $\tilde{c}_3$ are additive under $\boxplus_n$. Higher-order quantities $\tilde{c}_k$ $(k \geq 4)$ acquire cross-terms involving products of lower cumulants. This means the elegant $n = 3$ proof—which crucially relies on having *only* additive quantities in the formula for $1/\Phi_3$—does not extend directly to higher degrees.

6.5 Monotonicity of the Interpolation Fails

Define $g(t) = 1/\Phi_n(p \boxplus_n q_t) - t/\Phi_n(q)$, where q_t has roots scaled by $\sqrt{t}$. While $g(1) \geq g(0)$ is equivalent to the desired inequality, the pointwise derivative bound $g'(t) \geq 0$ does *not* hold for $n \geq 3$. Any proof via integration of the derivative must account for these non-monotone oscillations.

7 Numerical Evidence

7.1 Comprehensive Testing

The increasing minimum gap suggests the inequality becomes "easier" (further from tight) as n grows, consistent with convergence to the free probability regime.

n	Tests	Violations	Min gap
2	20,000	0	~ 0 (equality)
3	20,000	0	1.2×10^{-6}
4	20,000	0	1.4×10^{-5}
5	20,000	0	4.9×10^{-4}
6	20,000	0	9.9×10^{-4}
7	20,000	0	1.2×10^{-3}
8	20,000	0	2.9×10^{-3}
9	20,000	0	3.2×10^{-3}
10	20,000	0	5.7×10^{-3}

Table 1: Numerical verification of inequality (1). Random centered polynomials with roots generated from exponential spacings. Total: 180,000 tests, zero violations.

7.2 Equality Conditions

- $n = 2$: Equality holds universally.
- $n = 3$: Equality holds if and only if both p and q are odd polynomials ($d_p = d_q = 0$), i.e., both are of the form $x^3 + cx$ (scalar Hermite type).
- $n \geq 4$: No equality observed in any random test. We conjecture that equality holds only when both p and q are scalar multiples of the degree-n Hermite polynomial.

8 Proof Strategies for General n

8.1 Heat Flow Integration with Dirichlet Form Bound

The most promising approach uses the Hermite heat flow. If one can establish the bound

$$S(r_t) \;\geq\; \frac{\Phi_n(r_t)^2}{2\,\Phi_n(H_n)} \;=\; \frac{\Phi_n(r_t)^2}{n(n-1)}, \tag{8}$$

then integrating $\frac{d}{dt}\big(1/\Phi_n\big) = 2S/\Phi_n^2 \geq 1/\Phi_n(H_n)$ yields

$$\frac{1}{\Phi_n(r_T)} - \frac{1}{\Phi_n(p)} \;\geq\; \frac{T}{\Phi_n(H_n)}.$$

Numerical evidence strongly supports (8): the minimum observed ratio $2S/\Phi_n^2 \cdot \Phi_n(H_n)$ is approximately 1.0 across thousands of tests. Proving this bound and extending from the Hermite case to general q would complete the proof.

8.2 Finite Free Information-Theoretic Approach

An alternative strategy develops a finite free analog of the classical Stam inequality proof:

1. Establish a finite Cramér–Rao bound relating $\mathrm{Var}(p)$ and $\Phi_n(p)$.
2. Use the variance additivity $\mathrm{Var}(r) = \mathrm{Var}(p) + \mathrm{Var}(q)$.
3. Combine with a data-processing inequality for the conditional score.

The challenge is that the finite Cramér–Rao bound $\mathrm{Var}(p) \cdot \Phi_n(p) \geq n$ is not tight enough on its own to yield the desired inequality.

8.3 Algebraic Approach via Resultants

One may express Φ_n as a rational function of polynomial coefficients using resultants and discriminants. For each fixed n, the inequality (1) becomes a polynomial inequality in the coefficients, potentially amenable to sum-of-squares (SOS) certification. The $n = 3$ proof is an instance of this approach. For $n = 4$, the expressions become significantly more complex but may still admit human-readable proofs.

9 Connections and Significance

9.1 Classical Analogs

The inequality $1/\Phi_n(r) \geq 1/\Phi_n(p) + 1/\Phi_n(q)$ is the finite free analog of two classical results:

- **Stam's inequality** (classical probability): $1/I(X+Y) \geq 1/I(X)+1/I(Y)$ for independent random variables X, Y with Fisher information I.
- **Voiculescu's free Stam inequality** (free probability): $1/\Phi^*(\mu \boxplus \nu) \geq 1/\Phi^*(\mu)+1/\Phi^*(\nu)$ for freely independent random variables with free Fisher information Φ^*.

9.2 Finite-to-Free Convergence

As $n \to \infty$ with roots distributed according to a measure μ, we have $\frac{1}{n}\Phi_n \to \Phi^*(\mu)$ and $\boxplus_n \to \boxplus$. The finite free Fisher information inequality thus provides a polynomial-level refinement of the free Stam inequality, valid for every finite n.

9.3 Implications for Polynomial Theory

The inequality constrains the "regularity" of finite free convolutions: the reciprocal Fisher information (a measure of root spreading) is superadditive under $\boxplus_n$. This complements the Marcus–Spielman–Srivastava result on real-rootedness preservation and suggests deeper structural properties of the convolution operation.

A Proof of Proposition 5.3

We prove that $\sum_{i<j} \frac{f_i - f_j}{\lambda_i - \lambda_j} = \sum_k f_k^2$ where $f_k = \sum_{j \neq k} \frac{1}{\lambda_k - \lambda_j}$.

Step 1. Compute $f_i - f_j$:

$$f_i - f_j = \frac{1}{\lambda_i - \lambda_j} + \sum_{k \neq i,j} \left(\frac{1}{\lambda_i - \lambda_k} - \frac{1}{\lambda_j - \lambda_k} \right) = \frac{1}{\lambda_i - \lambda_j} - (\lambda_i - \lambda_j) \sum_{k \neq i,j} \frac{1}{(\lambda_i - \lambda_k)(\lambda_j - \lambda_k)}.$$

Step 2. Divide by $\lambda_i - \lambda_j$:

$$\frac{f_i - f_j}{\lambda_i - \lambda_j} = \frac{1}{(\lambda_i - \lambda_j)^2} - \sum_{k \neq i,j} \frac{1}{(\lambda_i - \lambda_k)(\lambda_j - \lambda_k)}.$$

Step 3. Sum over $i < j$. Setting $T = \sum_{i<j} \frac{1}{(\lambda_i - \lambda_j)^2}$ for the first sum, and analyzing the triple sum using symmetry, one obtains identity (6). This has been verified symbolically for $n = 3$ and $n = 4$ using SymPy, and numerically for n up to 20.

B Derivation of the Finite Dyson Equation

Under the Hermite heat flow $r_t = p \boxplus_n s_t$ where s_t has roots $\sqrt{t} \cdot h_k$, the coefficients $c_k(r_t)$ depend on t through s_t. The roots $\gamma_k(t)$ evolve smoothly. By implicit differentiation of $r_t(\gamma_k(t)) = 0$:

$$\dot{\gamma}_k(t) \;=\; -\frac{\partial_t\, r_t(\gamma_k)}{r_t'(\gamma_k)}.$$

Computing $\partial_t\, r_t(x)$ from the convolution formula (2) and evaluating at $x = \gamma_k$ yields

$$\dot{\gamma}_k \;=\; f_k(r_t) \;=\; \sum_{j \neq k} \frac{1}{\gamma_k - \gamma_j},$$

which is exactly the finite Dyson Brownian motion equation at unit "temperature." This has been verified numerically to machine precision.

References

[1] A. Marcus, D. A. Spielman, and N. Srivastava, *Interlacing families I: Bipartite Ramanujan graphs of all degrees*, Ann. of Math. **182** (2015), no. 1, 307–325.

[2] A. Marcus, D. A. Spielman, and N. Srivastava, *Finite free convolutions of polynomials*, Probab. Theory Related Fields **182** (2022), 807–848.

[3] D. Voiculescu, *The analogues of entropy and of Fisher's information measure in free probability theory, V: Noncommutative Hilbert transforms*, Invent. Math. **132** (1998), no. 1, 189–227.

[4] A. J. Stam, *Some inequalities satisfied by the quantities of information of Fisher and Shannon*, Inform. and Control **2** (1959), 101–112.