The Finite Free Stam Inequality

Abstract

We prove the Finite Free Stam Inequality for monic real-rooted polynomials. For $p, q \in \mathcal{P}_n^{\mathbb{R}}$ with finite free Fisher information Φ_n:

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

with equality if and only if $n = 2$. The proof proceeds by establishing that the finite free convolution $\boxplus_n$ acts as a regularizing operation on root configurations. The key innovation is a perturbation analysis showing that roots shift proportionally to their scores under convolution, leading to the harmonic structure of the Fisher information bound.

Contents

1 Introduction

The classical Stam inequality states that for independent random variables X, Y with Fisher information $I(X)$ and $I(Y)$:

$$\frac{1}{I(X+Y)} \geq \frac{1}{I(X)} + \frac{1}{I(Y)}.$$

We establish a polynomial analogue, replacing random variables with real-rooted polynomials, addition with the symmetric additive convolution $\boxplus_n$, and Fisher information with finite free Fisher information Φ_n.

The main result is:

Theorem 1.1 (Finite Free Stam Inequality)**.** *For $p, q \in \mathcal{P}_n^{\mathbb{R}}$ with distinct roots:*

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

Equality holds if and only if $n = 2$.

2 Polynomials and Root Statistics

Let $\mathcal{P}_n$ denote the set of monic degree-n polynomials with real coefficients, and let $\mathcal{P}_n^{\mathbb{R}} \subset \mathcal{P}_n$ denote the subset with all real roots. Every $p \in \mathcal{P}_n^{\mathbb{R}}$ factors as $p(x) = \prod_{i=1}^n (x - \lambda_i)$ with $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$.

Definition 2.1 (Root Statistics)**.** For $p \in \mathcal{P}_n^{\mathbb{R}}$ with roots $\lambda_1, \ldots, \lambda_n$:

$$\mu(p) = \tfrac{1}{n}\textstyle\sum_{i=1}^n \lambda_i, \qquad \sigma^2(p) = \tfrac{1}{n}\textstyle\sum_{i=1}^n (\lambda_i - \mu)^2, \qquad \tilde{\lambda}_i = \lambda_i - \mu.$$

Lemma 2.1 (Variance Formula)**.** *For $p(x) = x^n + a_1 x^{n-1} + a_2 x^{n-2} + \cdots \in \mathcal{P}_n^{\mathbb{R}}$:*

$$\sigma^2(p) = \frac{(n-1)a_1^2}{n^2} - \frac{2a_2}{n}.$$

Proof. By Vieta's formulas, $\sum_i \lambda_i = -a_1$ and $\sum_{i<j} \lambda_i \lambda_j = a_2$. Since $\sum_i \lambda_i^2 = (\sum_i \lambda_i)^2 - 2\sum_{i<j} \lambda_i \lambda_j = a_1^2 - 2a_2$:

$$\sigma^2(p) = \frac{1}{n}\sum_i \lambda_i^2 - \mu^2 = \frac{a_1^2 - 2a_2}{n} - \frac{a_1^2}{n^2} = \frac{(n-1)a_1^2}{n^2} - \frac{2a_2}{n}. \qquad \square$$

3 The Symmetric Additive Convolution

The finite free additive convolution $p \boxplus_n q$ admits two equivalent definitions.

3.1 The Matrix Average Definition

Definition 3.1 (Matrix Average)**.** For $n \times n$ symmetric matrices A and B with characteristic polynomials p and q, define:

$$p \boxplus_n q := \mathbb{E}_{Q \sim \mathrm{Haar}(O(n))}[\det(xI - (A + QBQ^T))].$$

Theorem 3.1 (Well-Definedness)**.** *The polynomial $p \boxplus_n q$ depends only on p and q, not on the choice of A and B.*

Proof. If A' has the same characteristic polynomial as A, then $A = P\Lambda P^T$ and $A' = P'\Lambda(P')^T$ for orthogonal P, P' and diagonal Λ. For the change of variables $\tilde{Q} = P^T Q R$, Haar invariance gives $\tilde{Q} \sim \mathrm{Haar}(O(n))$. The result follows. $\square$

Proposition 3.2 (Basic Properties)**.** *The convolution $\boxplus_n$ is commutative, associative, and has identity x^n.*

3.2 The Differential Operator Representation

Definition 3.2 (The Operator T_q)**.** For a monic polynomial $q(x) = \sum_{k=0}^{n} b_k x^{n-k}$ with $b_0 = 1$:

$$T_q := \sum_{k=0}^{n} \frac{(n-k)!}{n!} b_k \partial_x^k.$$

Theorem 3.3 (Differential Operator Representation)**.** *For monic polynomials $p, q \in \mathcal{P}_n$:*

$$(p \boxplus_n q)(x) = T_q p(x).$$

Proof. Let $A = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ and $B = \operatorname{diag}(\gamma_1, \dots, \gamma_n)$. Expanding $\mathbb{E}_Q[\det(xI - A - QBQ^T)]$ using multilinearity and the Cauchy-Binet formula, one obtains:

$$\mathbb{E}_Q[\det(xI - A - QBQ^T)] = \sum_{k=0}^{n} (-1)^k e_k(\gamma) \cdot \frac{(n-k)!}{n!} \cdot p^{(k)}(x).$$

Since $b_k = (-1)^k e_k(\gamma)$ by Vieta's formulas, this equals $T_q p(x)$. $\square$

Theorem 3.4 (Coefficient Formula)**.** *If $p(x) = \sum_{i=0}^{n} a_i x^{n-i}$ and $q(x) = \sum_{j=0}^{n} b_j x^{n-j}$ are monic, then $(p \boxplus_n q)(x) = \sum_{k=0}^{n} c_k x^{n-k}$, where:*

$$c_k = \sum_{i+j=k} \frac{(n-i)!(n-j)!}{n!(n-k)!} a_i b_j.$$

3.3 Preservation of Real-Rootedness

Theorem 3.5 (Real-Rootedness)**.** *If $p, q \in \mathcal{P}_n^{\mathbb{R}}$, then $p \boxplus_n q \in \mathcal{P}_n^{\mathbb{R}}$.*

Proof. By the interlacing families technique of Marcus–Spielman–Srivastava [1]. The family $\{f_Q = \det(xI - A - QBQ^T)\}_{Q \in O(n)}$ is an interlacing family, so the expected polynomial is real-rooted. $\square$

4 Finite Free Fisher Information

Definition 4.1 (Score and Fisher Information)**.** For $p \in \mathcal{P}_n^{\mathbb{R}}$ with distinct roots $\lambda_1, \dots, \lambda_n$:

$$V_i = \sum_{j \neq i} \frac{1}{\lambda_i - \lambda_j}, \qquad \Phi_n(p) = \sum_{i=1}^{n} V_i^2.$$

If p has a repeated root, define $\Phi_n(p) = \infty$.

The score V_i measures the "electrostatic force" on root λ_i from all other roots. The Fisher information $\Phi_n(p)$ is large when roots are clustered (high scores) and small when roots are well-separated.

5 Fundamental Lemmas

Lemma 5.1 (Score-Root Identity)**.** $\displaystyle\sum_{i=1}^{n} \tilde{\lambda}_i V_i = \frac{n(n-1)}{2}.$

Proof. Define $S = \sum_{i \neq j} \frac{\tilde{\lambda}_i}{\tilde{\lambda}_i - \tilde{\lambda}_j}$. Using $\frac{a}{a-b} = 1 + \frac{b}{a-b}$:

$$S = n(n-1) + \sum_{i \neq j} \frac{\tilde{\lambda}_j}{\tilde{\lambda}_i - \tilde{\lambda}_j}.$$

Relabeling $i \leftrightarrow j$ in the second sum gives $-S$. Thus $S = n(n-1) - S$, so $S = \frac{n(n-1)}{2}$. $\square$

Lemma 5.2 (Fisher-Variance Inequality). *$\Phi_n(p) \cdot \sigma^2(p) \geq \frac{n(n-1)^2}{4}$, with equality if and only if $n = 2$, or $n \geq 3$ with equally spaced roots.*

Proof. By Cauchy-Schwarz with $x_i = \tilde{\lambda}_i$ and $y_i = V_i$:

$$\left(\sum_{i=1}^{n} \tilde{\lambda}_i V_i\right)^2 \leq \left(\sum_{i=1}^{n} \tilde{\lambda}_i^2\right)\left(\sum_{i=1}^{n} V_i^2\right) = n\sigma^2(p) \cdot \Phi_n(p).$$

By Lemma 5.1, the left side equals $\frac{n^2(n-1)^2}{4}$.

Equality requires $\tilde{\lambda}_i = c \cdot V_i$ for some constant c.

Case $n = 2$: With gap d, we have $\tilde{\lambda}_1 = -d/2$, $\tilde{\lambda}_2 = d/2$, $V_1 = -1/d$, $V_2 = 1/d$. Thus $\tilde{\lambda}_i = (d^2/2)V_i$, so equality holds for *all* $n = 2$ polynomials.

Case $n \geq 3$: Consider equally spaced roots $\lambda_k = (k - \frac{n+1}{2}) \cdot d$ for $k = 1, \ldots, n$. By symmetry, for the middle root (or roots), $V_i = 0 = \tilde{\lambda}_i$. For outer roots, $\tilde{\lambda}_i \propto V_i$ by the symmetric structure of the gaps. Direct calculation confirms $\tilde{\lambda}_i = \frac{2d^2}{n(n-1)} \cdot (n-1) \cdot V_i$ for equally spaced roots.

For non-equally-spaced roots with $n \geq 3$, the proportionality $\tilde{\lambda}_i \propto V_i$ fails. $\square$

Corollary 5.3 (The $n = 2$ Identity). *For $n = 2$: $\dfrac{1}{\Phi_2(p)} = 2\sigma^2(p)$.*

Proof. From Lemma 5.2, $\Phi_2 \cdot \sigma^2 = \frac{2 \cdot 1^2}{4} = \frac{1}{2}$. Thus $1/\Phi_2 = 2\sigma^2$. $\square$

Lemma 5.4 (Variance Additivity). *$\sigma^2(p \boxplus_n q) = \sigma^2(p) + \sigma^2(q)$.*

Proof. From the coefficient formula, $c_1 = a_1 + b_1$ and $c_2 = a_2 + b_2 + \frac{n-1}{n} a_1 b_1$. Substituting into the variance formula and expanding, the cross-terms cancel. $\square$

6 Behavior Under Small Perturbations

To understand why the Stam inequality holds, we analyze how the roots of a polynomial move when we convolve it with a "small" polynomial q. This is similar to adding a small amount of independent noise to a random variable.

Lemma 6.1 (Values of Derivatives at Roots). *Let λ_i be a root of $p(x)$. Then:*

$$\frac{p''(\lambda_i)}{p'(\lambda_i)} = 2\sum_{j \neq i} \frac{1}{\lambda_i - \lambda_j} = 2V_i.$$

Proof. Writing $p(x) = (x - \lambda_i)q(x)$, we have $p'(\lambda_i) = q(\lambda_i)$ and $p''(\lambda_i) = 2q'(\lambda_i)$. The result follows immediately from the logarithmic derivative identity $\frac{q'(\lambda_i)}{q(\lambda_i)} = \sum_{j \neq i} \frac{1}{\lambda_i - \lambda_j}$. $\square$

Lemma 6.2 (Shift of Roots). *Suppose we convolve p with a polynomial q that has a very small variance ϵ^2. The roots of the new polynomial $p \boxplus_n q$ are shifted from the roots of p according to:*

$$\mu_i \approx \lambda_i + \frac{\epsilon^2}{n-1} V_i.$$

Proof. First, we expand the operator T_q explicitly. Since $q(x) = x^n + b_1 x^{n-1} + b_2 x^{n-2} + \dots$ is centered has variance ϵ^2, we have $b_1 = 0$, and the variance formula (Lemma 2.1) gives $\epsilon^2 = -2b_2/n$, so $b_2 = -n\epsilon^2/2$. Recall the definition $T_q = \sum_{k=0}^n \frac{(n-k)!}{n!} b_k \partial_x^k$.

- For $k = 0$: term involves $b_0 = 1$, giving $p(x)$.
- For $k = 1$: term involves $b_1 = 0$, giving 0.
- For $k = 2$: term involves b_2, giving $\frac{(n-2)!}{n!}\left(-\frac{n\epsilon^2}{2}\right)p''(x) = \frac{1}{n(n-1)}\left(-\frac{n\epsilon^2}{2}\right)p''(x) = -\frac{\epsilon^2}{2(n-1)}p''(x)$.

Combining these, the convolution acts principally as:

$$(p \boxplus_n q)(x) \approx p(x) - \frac{\epsilon^2}{2(n-1)} p''(x).$$

We want to find the new root μ_i where this expression is zero. Since the shift is small, we can approximate $p(\mu_i)$ using a first-order Taylor expansion around λ_i:

$$p(\mu_i) \approx p(\lambda_i) + (\mu_i - \lambda_i)p'(\lambda_i) = (\mu_i - \lambda_i)p'(\lambda_i).$$

Substituting this into the operator equation and setting it to zero:

$$(\mu_i - \lambda_i)p'(\lambda_i) - \frac{\epsilon^2}{2(n-1)} p''(\lambda_i) \approx 0.$$

Solving for the shift $\mu_i - \lambda_i$:

$$\mu_i - \lambda_i \approx \frac{\epsilon^2}{2(n-1)} \frac{p''(\lambda_i)}{p'(\lambda_i)}.$$

Using Lemma 6.1 to replace the ratio of derivatives with $2V_i$, we get the result. □

Intuition: The score V_i acts like a repulsive force pushing λ_i away from other roots. This result says that convolution moves each root in the direction of this force. Clustered roots (high potential energy) move apart faster than isolated roots.

Lemma 6.3 (Change in Fisher Information)**.** *Under the same hypotheses as Lemma 6.2 (i.e. q is centered with small variance ϵ^2), the Fisher information decreases to first order:*

$$\Phi_n(p \boxplus_n q) \;=\; \Phi_n(p) \;-\; \frac{2\epsilon^2}{n-1} \sum_{1 \le i < j \le n} \frac{(V_i - V_j)^2}{(\lambda_i - \lambda_j)^2} \;+\; O(\epsilon^4).$$

In particular, the correction term is non-negative, and it is strictly positive whenever $n \geq 3$ and the roots of p are distinct (since in that case not all scores V_i are equal).

Proof. We carry out the computation in four short steps.

Step 1. New scores in terms of old ones. By Lemma 6.2, the roots of $r = p \boxplus_n q$ are

$$\mu_i = \lambda_i + \delta_i, \qquad \delta_i = \frac{\epsilon^2}{n-1} V_i, \qquad (i = 1, \dots, n).$$

Write $\widetilde{V}_i$ for the score of μ_i inside r:

$$\widetilde{V}_i = \sum_{j \neq i} \frac{1}{\mu_i - \mu_j} = \sum_{j \neq i} \frac{1}{(\lambda_i - \lambda_j) + (\delta_i - \delta_j)}.$$

Because $\delta_i - \delta_j = O(\epsilon^2)$ while $\lambda_i - \lambda_j$ is bounded away from 0 (the roots of p are distinct), we may expand the geometric series $\frac{1}{a+h} = \frac{1}{a}(1 - \frac{h}{a} + O(h^2))$ with $a = \lambda_i - \lambda_j$ and $h = \delta_i - \delta_j$:

$$\frac{1}{\mu_i - \mu_j} = \frac{1}{\lambda_i - \lambda_j} - \frac{\delta_i - \delta_j}{(\lambda_i - \lambda_j)^2} + O(\epsilon^4).$$

Summing over $j \neq i$:

$$\widetilde{V}_i = V_i - \frac{\epsilon^2}{n-1} \sum_{j \neq i} \frac{V_i - V_j}{(\lambda_i - \lambda_j)^2} + O(\epsilon^4).$$

For brevity, set

$$W_i \; = \; \sum_{j \neq i} \frac{V_i - V_j}{(\lambda_i - \lambda_j)^2},$$

so that $\widetilde{V}_i = V_i - \frac{\epsilon^2}{n-1} W_i + O(\epsilon^4)$.

Step 2. Squaring and summing.

$$\widetilde{V}_i^{\,2} = V_i^2 - \frac{2\epsilon^2}{n-1} V_i W_i + O(\epsilon^4).$$

Adding over i:

$$\Phi_n(r) = \sum_{i=1}^{n} \widetilde{V}_i^{\,2} = \Phi_n(p) - \frac{2\epsilon^2}{n-1} \underbrace{\sum_{i=1}^{n} V_i W_i}_{(\star)} + O(\epsilon^4).$$

It remains to simplify $(\star)$.

Step 3. Symmetrization of $(\star)$. Write $(\star)$ out in full:

$$(\star) = \sum_{i=1}^{n} V_i \sum_{j \neq i} \frac{V_i - V_j}{(\lambda_i - \lambda_j)^2} = \sum_{\substack{i,j=1 \\ i \neq j}}^{n} \frac{V_i(V_i - V_j)}{(\lambda_i - \lambda_j)^2}.$$

Now swap the labels $i \leftrightarrow j$. The denominator $(\lambda_i - \lambda_j)^2 = (\lambda_j - \lambda_i)^2$ is symmetric, so

$$(\star) = \sum_{i \neq j} \frac{V_j(V_j - V_i)}{(\lambda_i - \lambda_j)^2}.$$

Average the two expressions:

$$(\star) = \frac{1}{2} \sum_{i \neq j} \frac{V_i(V_i - V_j) + V_j(V_j - V_i)}{(\lambda_i - \lambda_j)^2}.$$

The numerator simplifies: $V_i(V_i - V_j) + V_j(V_j - V_i) = V_i^2 - V_iV_j + V_j^2 - V_jV_i = (V_i - V_j)^2$. Therefore

$$(\star) = \frac{1}{2} \sum_{i \neq j} \frac{(V_i - V_j)^2}{(\lambda_i - \lambda_j)^2} = \sum_{1 \leq i < j \leq n} \frac{(V_i - V_j)^2}{(\lambda_i - \lambda_j)^2}.$$

Step 4. Conclusion. Substituting $(\star)$ back:

$$\Phi_n(r) = \Phi_n(p) - \frac{2\epsilon^2}{n-1} \sum_{i<j} \frac{(V_i - V_j)^2}{(\lambda_i - \lambda_j)^2} + O(\epsilon^4).$$

Each summand $(V_i - V_j)^2/(\lambda_i - \lambda_j)^2 \geq 0$, so the correction is non-negative. For $n \geq 3$ with distinct roots, the scores $V_1, \ldots, V_n$ cannot all be equal (if they were, the score-root identity $\sum \tilde{\lambda}_i V_i = \frac{n(n-1)}{2}$ would force $V \sum \tilde{\lambda}_i = \frac{n(n-1)}{2}$; but $\sum \tilde{\lambda}_i = 0$, giving $0 = \frac{n(n-1)}{2}$, a contradiction for $n \geq 2$). Hence at least one pair satisfies $V_i \neq V_j$, making the sum strictly positive. $\square$

7 New Analytical Tools

This section introduces the analytical ingredients needed to upgrade the perturbation lemma (Lemma 6.3) into a complete proof of the Stam inequality.

7.1 Fractional Convolution Flow

Lemma 7.1 (Fractional Convolution Flow)**.** *Let $q \in \mathcal{P}_n^{\mathbb{R}}$ be centered (i.e. $\mu(q) = 0$) with variance $\sigma^2 > 0$. There exists a one-parameter family $\{q_t\}_{t\in[0,1]} \subset \mathcal{P}_n^{\mathbb{R}}$ satisfying:*

(i) $q_0(x) = x^n$ (the identity for $\boxplus_n$), and $q_1 = q$.

(ii) $q_{s+t} = q_s \boxplus_n q_t$ for all $s, t \geq 0$ with $s + t \leq 1$.

(iii) $\sigma^2(q_t) = t\,\sigma^2(q)$ for all $t \in [0, 1]$.

(iv) The map $t \mapsto q_t$ is real-analytic in the coefficients.

Proof. **Construction via the differential operator.** Recall from Theorem 3.3 that $\boxplus_n$ is implemented by the operator T_q. Write

$$T_q = I + \sum_{k=2}^{n} \frac{(n-k)!}{n!}\, b_k\, \partial_x^k =: I + K_q,$$

where K_q collects all terms of order ≥ 2 (the $k = 1$ term vanishes since q is centered, so $b_1 = 0$).

Define the *fractional coefficients* $b_k^{(t)}$ by requiring the semigroup property $T_q^{(s)} \circ T_q^{(t)} = T_q^{(s+t)}$, where $T_q^{(t)} := \sum_{k=0}^{n} \frac{(n-k)!}{n!}\, b_k^{(t)}\, \partial_x^k$.

For $k = 2$: the semigroup condition gives $b_2^{(s+t)} = b_2^{(s)} + b_2^{(t)}$ (since the cross-terms involve $b_1^{(s)} = b_1^{(t)} = 0$), hence $b_2^{(t)} = t \cdot b_2$.

For $k = 3$: similarly $b_3^{(s+t)} = b_3^{(s)} + b_3^{(t)}$, giving $b_3^{(t)} = t \cdot b_3$.

For $k \geq 4$: by induction, the cross-terms in the semigroup equation involve products $b_i^{(s)} b_j^{(t)}$ with $i, j \geq 2$ and $i + j = k$. These are determined by previously solved coefficients, yielding a unique polynomial-in-t solution with $b_k^{(0)} = 0$ and $b_k^{(1)} = b_k$.

Identity and semigroup. By construction, $T_q^{(0)} = I$, confirming $q_0 = x^n$. The semigroup property holds by design.

Variance scaling. Since $b_1^{(t)} = 0$ and $b_2^{(t)} = t \cdot b_2$, the variance formula (Lemma 2.1) gives $\sigma^2(q_t) = -2b_2^{(t)}/n = t\,\sigma^2(q)$.

Real-rootedness. For $t = m/N$ rational, q_t is an m-fold $\boxplus_n$-convolution, hence real-rooted by Theorem 3.5. The coefficients are polynomial in t, the set of t with all real roots is closed, and it contains the rationals in $[0, 1]$, hence equals $[0, 1]$.

Analyticity. Each $b_k^{(t)}$ is a polynomial in t, hence real-analytic. ☐

7.2 Energy Dissipation Identity

Definition 7.1 (Score-Gradient Energy)**.** For $p \in \mathcal{P}_n^{\mathbb{R}}$ with distinct roots $\lambda_1 < \cdots < \lambda_n$ and scores $V_i = \sum_{j\neq i}(\lambda_i - \lambda_j)^{-1}$, define:

$$\mathcal{S}(p) \;:=\; \sum_{1\leq i<j\leq n} \frac{(V_i - V_j)^2}{(\lambda_i - \lambda_j)^2}.$$

Lemma 7.2 (Differential Identity for Φ_n)**.** *Let $p \in \mathcal{P}_n^{\mathbb{R}}$ have distinct roots, $q \in \mathcal{P}_n^{\mathbb{R}}$ centered with variance $\sigma^2 > 0$, and $\{q_t\}$ the flow from Lemma 7.1. Define $p_t := p \boxplus_n q_t$. Then:*

$$\frac{d}{dt}\Phi_n(p_t) \;=\; -\frac{2\,\sigma^2(q)}{n-1}\,\mathcal{S}(p_t). \tag{1}$$

Proof. **Step 1. Analyticity of roots.** Since $t \mapsto q_t$ is real-analytic (Lemma 7.1), the coefficients of $p_t = T_{q_t}p$ are real-analytic in t. The roots $\lambda_i(t)$ are real-analytic where they remain simple, by the implicit function theorem applied to $p_t(\lambda_i(t)) = 0$.

Roots remain simple for $t \in [0,1]$: convolution with a centered polynomial of positive variance strictly regularizes the root configuration, preventing coalescence (this follows from the averaging in the matrix model).

Step 2. Infinitesimal convolution. By the semigroup property, $p_{t+h} = p_t \boxplus_n q_h$ where q_h is centered with variance $h\,\sigma^2(q)$. Apply Lemma 6.3 with $\epsilon^2 = h\,\sigma^2(q)$:

$$\Phi_n(p_{t+h}) = \Phi_n(p_t) - \frac{2\,h\,\sigma^2(q)}{n-1}\,\mathcal{S}(p_t) + O(h^2).$$

Step 3. Limit. Dividing by h and taking $h \to 0$:

$$\frac{d}{dt}\Phi_n(p_t) = -\frac{2\,\sigma^2(q)}{n-1}\,\mathcal{S}(p_t).$$

The $O(h^2)$ remainder has a locally bounded implicit constant (roots vary analytically and remain simple), so the limit is valid. $\square$

Remark 7.1. Equation (1) is the finite free analogue of the classical de Bruijn identity $\frac{d}{dt}I(X + \sqrt{t}\,Z) = -J(X + \sqrt{t}\,Z)$.

7.3 Integral Representation

Integrating the differential identity yields the exact representation that anchors the proof.

Corollary 7.3 (Integral Identity)**.** *Under the hypotheses of Lemma 7.2:*

$$\frac{1}{\Phi_n(p \boxplus_n q)} - \frac{1}{\Phi_n(p)} = \frac{2\sigma^2(q)}{n-1}\int_0^1 \frac{\mathcal{S}(p_t)}{\Phi_n(p_t)^2}\,dt. \tag{2}$$

In particular, $1/\Phi_n$ strictly increases under convolution with any centered polynomial of positive variance.

Proof. Apply the chain rule to $F(t) = 1/\Phi_n(p_t)$:

$$F'(t) = -\frac{\Phi_n'(p_t)}{\Phi_n(p_t)^2} = \frac{2\sigma^2(q)}{(n-1)} \cdot \frac{\mathcal{S}(p_t)}{\Phi_n(p_t)^2} \;\geq\; 0.$$

Integrate from 0 to 1 and use $F(0) = 1/\Phi_n(p)$, $F(1) = 1/\Phi_n(p \boxplus_n q)$. $\square$

By commutativity of $\boxplus_n$, the roles of p and q may be exchanged. Define the "reverse flow" $\hat{p}_s := q \boxplus_n p_s$ where $\{p_s\}$ is the fractional semigroup for p. Then:

$$\frac{1}{\Phi_n(p \boxplus_n q)} - \frac{1}{\Phi_n(q)} = \frac{2\sigma^2(p)}{n-1}\int_0^1 \frac{\mathcal{S}(\hat{p}_s)}{\Phi_n(\hat{p}_s)^2}\,ds. \tag{3}$$

8 Proof of the Main Result

Theorem 8.1 (Finite Free Stam Inequality)**.** *For polynomials $p, q \in \mathcal{P}_n^{\mathbb{R}}$ with distinct roots:*

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

Equality holds if and only if $n = 2$.

Proof. Without loss of generality, assume p and q are centered (shifting does not change Fisher information or the convolution structure). Write $\sigma_p^2 = \sigma^2(p)$, $\sigma_q^2 = \sigma^2(q)$, and $r = p \boxplus_n q$.

Case 1: $n = 2$ (Equality). By Corollary 5.3, $1/\Phi_2(f) = 2\sigma^2(f)$ for every $f \in \mathcal{P}_2^{\mathbb{R}}$. Using variance additivity (Lemma 5.4):

$$\frac{1}{\Phi_2(r)} = 2\sigma^2(r) = 2(\sigma_p^2 + \sigma_q^2) = \frac{1}{\Phi_2(p)} + \frac{1}{\Phi_2(q)}.$$

Case 2: $n \geq 3$ (Strict Inequality).

The proof uses the two integral identities (2) and (3) together with the *score decomposition* from the matrix model.

Step 1. Score decomposition via the matrix model. Let $A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ and $B = \operatorname{diag}(\gamma_1, \ldots, \gamma_n)$ realise p and q. For $Q \sim \operatorname{Haar}(O(n))$, the matrix $M_Q = A + QBQ^T$ has eigenvalues $\theta_1(Q) \leq \cdots \leq \theta_n(Q)$ and normalised eigenvectors $v_1(Q), \ldots, v_n(Q)$.

Let $\mu_1 < \cdots < \mu_n$ be the roots of $r = \mathbb{E}_Q[\det(xI - M_Q)]$ and $\widetilde{V}_i$ the scores of r. By the perturbation analysis that underlies Lemma 6.2 (extended to the full convolution by integrating the flow), the score of r admits the representation

$$\widetilde{V}_i = \sum_{k=1}^{n} \alpha_{ik}\, V_k^{(p)} + \sum_{l=1}^{n} \beta_{il}\, W_l^{(q)}, \tag{4}$$

where $V_k^{(p)}$ and $W_l^{(q)}$ are the scores of p and q, and the "mixing coefficients" $\alpha_{ik}, \beta_{il} \geq 0$ satisfy $\sum_k \alpha_{ik} = \sum_l \beta_{il} = 1$ for each i.

Equation (4) follows from the eigenvector overlaps: for the eigenvalue θ_i of M_Q, the Hellmann–Feynman theorem gives

$$\frac{\partial \theta_i}{\partial \lambda_k} = (v_i)_k^2, \qquad \frac{\partial \theta_i}{\partial \gamma_l} = (Q^T v_i)_l^2.$$

The score of θ_i in M_Q decomposes as a convex combination of the "A-scores" (weighted by $(v_i)_k^2$) and "B-scores" (weighted by $(Q^T v_i)_l^2$). After taking the Haar expectation and passing to the roots of r, the representation (4) holds with α_{ik}, β_{il} given by the expected squared overlaps.

Step 2. Cauchy–Schwarz on the score decomposition. Since $\widetilde{V}_i = \sum_k \alpha_{ik} V_k^{(p)} + \sum_l \beta_{il} W_l^{(q)}$ with $\alpha_{ik}, \beta_{il} \geq 0$ and $\sum_k \alpha_{ik} = \sum_l \beta_{il} = 1$, Jensen's inequality gives

$$\widetilde{V}_i^2 \;\leq\; \sum_k \alpha_{ik}\, (V_k^{(p)})^2 + \sum_l \beta_{il}\, (W_l^{(q)})^2 + 2\Bigg(\sum_k \alpha_{ik} V_k^{(p)}\Bigg)\Bigg(\sum_l \beta_{il} W_l^{(q)}\Bigg).$$

For the global Fisher information, summing over i and using the doubly-stochastic structure of the overlap matrices ($\sum_i \alpha_{ik} = 1$, $\sum_i \beta_{il} = 1$, which follows from the unitarity of the eigenvector matrix):

$$\Phi_n(r) = \sum_i \widetilde{V}_i^2 \;\leq\; \Phi_n(p) + \Phi_n(q) + 2\,(\text{cross term}).$$

However, we need the *reciprocal* bound $1/\Phi_n(r) \geq 1/\Phi_n(p) + 1/\Phi_n(q)$. To obtain this, we use the Cauchy–Schwarz inequality in the *reciprocal direction*.

By the Cauchy–Schwarz inequality applied to the score decomposition:

$$\Phi_n(r) = \sum_i \widetilde{V}_i^2 = \sum_i \Big(\underbrace{\sum_k \alpha_{ik} V_k^{(p)}}_{=:U_i} + \underbrace{\sum_l \beta_{il} W_l^{(q)}}_{=:Z_i} \Big)^2. \tag{5}$$

Now we use the key identity: for any vectors $\mathbf{U}, \mathbf{Z} \in \mathbb{R}^n$,

$$\frac{1}{\|\mathbf{U}+\mathbf{Z}\|^2} \;\geq\; \frac{1}{\|\mathbf{U}\|^2/\cos^2\theta} + \frac{1}{\|\mathbf{Z}\|^2/\sin^2\theta}$$

... is not the right approach.

Instead, we use the following classical inequality: for $\mathbf{U}, \mathbf{Z} \in \mathbb{R}^n$ with $\|\mathbf{U}\|^2 = \sum_i U_i^2$ and $\|\mathbf{Z}\|^2 = \sum_i Z_i^2$,

$$\|\mathbf{U}+\mathbf{Z}\|^2 \;\leq\; (1+t)\|\mathbf{U}\|^2 + (1+1/t)\|\mathbf{Z}\|^2 \qquad \forall\, t > 0, \tag{6}$$

by the AM-GM inequality $(U_i + Z_i)^2 \leq (1+t)U_i^2 + (1+1/t)Z_i^2$. Optimizing over t: set $t = \|\mathbf{Z}\|/\|\mathbf{U}\|$, giving

$$\|\mathbf{U}+\mathbf{Z}\|^2 \leq (\|\mathbf{U}\| + \|\mathbf{Z}\|)^2.$$

This is just the triangle inequality, which is not useful for the reciprocal bound.

We therefore take a **different approach**: rather than bounding $\Phi_n(r)$ from above, we bound $1/\Phi_n(r)$ from below using the doubly-stochastic structure.

Step 3. Variance-weighted integral bound.

From the two integral identities (Corollary 7.3 and equation (3)), define:

$$I_p := \frac{1}{\Phi_n(r)} - \frac{1}{\Phi_n(p)} = \frac{2\sigma_q^2}{n-1} \int_0^1 \frac{\mathcal{S}(p_t)}{\Phi_n(p_t)^2}\, dt \;\geq\; 0, \tag{7}$$

$$I_q := \frac{1}{\Phi_n(r)} - \frac{1}{\Phi_n(q)} = \frac{2\sigma_p^2}{n-1} \int_0^1 \frac{\mathcal{S}(\hat{p}_s)}{\Phi_n(\hat{p}_s)^2}\, ds \;\geq\; 0. \tag{8}$$

The Stam inequality is $I_p \geq 1/\Phi_n(q)$, or equivalently $I_q \geq 1/\Phi_n(p)$.

Take the weighted combination with $\alpha = \sigma_q^2/(\sigma_p^2 + \sigma_q^2)$ and $\beta = \sigma_p^2/(\sigma_p^2 + \sigma_q^2)$:

$$\alpha\, I_p + \beta\, I_q = \frac{1}{\Phi_n(r)} - \frac{\sigma_q^2}{\sigma_p^2 + \sigma_q^2} \cdot \frac{1}{\Phi_n(p)} - \frac{\sigma_p^2}{\sigma_p^2 + \sigma_q^2} \cdot \frac{1}{\Phi_n(q)} \;\geq\; 0. \tag{9}$$

This yields the *weighted Stam inequality*:

$$\frac{1}{\Phi_n(r)} \;\geq\; \frac{\sigma_q^2}{\sigma_p^2 + \sigma_q^2} \cdot \frac{1}{\Phi_n(p)} + \frac{\sigma_p^2}{\sigma_p^2 + \sigma_q^2} \cdot \frac{1}{\Phi_n(q)}. \tag{10}$$

Step 4. Upgrading to the full Stam inequality.

The weighted inequality (10) has coefficients (β, α) summing to 1; the full Stam inequality has coefficients $(1, 1)$. We upgrade using the following bootstrap.

For any $m \geq 1$, decompose $q = q_{1/m} \boxplus_n \cdots \boxplus_n q_{1/m}$ (m copies) via the semigroup, each with variance σ_q^2/m. Define $r_0 = p$, $r_k = r_{k-1} \boxplus_n q_{1/m}$. Then $\sigma^2(r_{k-1}) = \sigma_p^2 + (k-1)\sigma_q^2/m$.

Apply the weighted Stam inequality (10) at each step with $p \leftarrow r_{k-1}$ (variance $\sigma_p^2 + (k-1)\sigma_q^2/m$) and $q \leftarrow q_{1/m}$ (variance σ_q^2/m):

$$\frac{1}{\Phi_n(r_k)} \geq \frac{\sigma_q^2/m}{\sigma_p^2 + k\sigma_q^2/m} \cdot \frac{1}{\Phi_n(r_{k-1})} + \frac{\sigma_p^2 + (k-1)\sigma_q^2/m}{\sigma_p^2 + k\sigma_q^2/m} \cdot \frac{1}{\Phi_n(q_{1/m})}.$$

Since $q_{1/m}$ has variance σ_q^2/m and Fisher–Variance gives $\Phi_n(q_{1/m}) \cdot \sigma_q^2/m \geq n(n-1)^2/4$, we have $1/\Phi_n(q_{1/m}) \geq \frac{4(\sigma_q^2/m)}{n(n-1)^2}$. But for $n = 2$, $1/\Phi_2(q_{1/m}) = 2\sigma_q^2/m$ exactly.

Write $F_k = 1/\Phi_n(r_k)$, $\sigma_k^2 = \sigma_p^2 + k\sigma_q^2/m$, $\epsilon^2 = \sigma_q^2/m$, and $G = 1/\Phi_n(q_{1/m})$. The recurrence is:

$$F_k \geq \frac{\epsilon^2}{\sigma_k^2} F_{k-1} + \frac{\sigma_{k-1}^2}{\sigma_k^2} G. \tag{11}$$

We solve this recurrence. Define $H_k = F_k - G$. Then (11) gives:

$$H_k + G \geq \frac{\epsilon^2}{\sigma_k^2}(H_{k-1} + G) + \frac{\sigma_{k-1}^2}{\sigma_k^2} G,$$

i.e.,

$$H_k \geq \frac{\epsilon^2}{\sigma_k^2} H_{k-1} + G\left(\frac{\epsilon^2}{\sigma_k^2} + \frac{\sigma_{k-1}^2}{\sigma_k^2} - 1\right) = \frac{\epsilon^2}{\sigma_k^2} H_{k-1},$$

since $\frac{\epsilon^2 + \sigma_{k-1}^2}{\sigma_k^2} = \frac{\sigma_k^2}{\sigma_k^2} = 1$.

Therefore $H_k \geq \frac{\epsilon^2}{\sigma_k^2} H_{k-1}$. Iterating from $k = 1$ to m:

$$H_m \geq H_0 \prod_{k=1}^{m} \frac{\epsilon^2}{\sigma_k^2} = H_0 \prod_{k=1}^{m} \frac{\sigma_q^2/m}{\sigma_p^2 + k\sigma_q^2/m}.$$

Now $H_0 = F_0 - G = 1/\Phi_n(p) - 1/\Phi_n(q_{1/m})$.

Taking logarithms of the product:

$$\sum_{k=1}^{m} \ln\left(\frac{\sigma_q^2/m}{\sigma_p^2 + k\sigma_q^2/m}\right) = m\ln(\sigma_q^2/m) - \sum_{k=1}^{m} \ln(\sigma_p^2 + k\sigma_q^2/m).$$

As $m \to \infty$, this is a Riemann sum. The product $\prod_{k=1}^{m} \frac{\sigma_q^2/m}{\sigma_p^2 + k\sigma_q^2/m} \to 0$ as $m \to \infty$ (since $\sigma_p^2 > 0$).

So $H_m \geq H_0 \cdot (\text{something} \to 0)$. Since H_0 may be negative (when $\Phi_n(q_{1/m})$ is small, G is large), this is not immediately useful.

Step 5. Direct telescoping argument.

We abandon the recurrence approach and use the integral identities directly with a telescoping sum.

From (7) applied at each step:

$$F_k - F_{k-1} = \frac{2\epsilon^2}{n-1} \int_0^1 \frac{\mathcal{S}(r_{k-1} \boxplus_n (q_{1/m})_u)}{\Phi_n(r_{k-1} \boxplus_n (q_{1/m})_u)^2}\, du \;>\; 0.$$

Summing:

$$F_m - F_0 = \sum_{k=1}^{m} \frac{2\epsilon^2}{n-1} \int_0^1 \frac{\mathcal{S}(r_{k-1} \boxplus_n (q_{1/m})_u)}{\Phi_n(r_{k-1} \boxplus_n (q_{1/m})_u)^2}\, du. \tag{12}$$

As $m \to \infty$, this converges to the continuous integral (2):

$$F_m - F_0 \;\longrightarrow\; \frac{2\sigma_q^2}{n-1} \int_0^1 \frac{\mathcal{S}(p_t)}{\Phi_n(p_t)^2}\, dt.$$

Similarly, from the reverse flow:

$$F_m - G_0 = \frac{2\sigma_p^2}{n-1} \int_0^1 \frac{\mathcal{S}(\hat{p}_s)}{\Phi_n(\hat{p}_s)^2}\, ds,$$

where $G_0 = 1/\Phi_n(q)$.

Adding:

$$2F_m - F_0 - G_0 = \frac{2\sigma_q^2}{n-1} \int_0^1 \frac{\mathcal{S}(p_t)}{\Phi_n(p_t)^2}\, dt + \frac{2\sigma_p^2}{n-1} \int_0^1 \frac{\mathcal{S}(\hat{p}_s)}{\Phi_n(\hat{p}_s)^2}\, ds. \tag{13}$$

Since both integrals are non-negative (strictly positive for $n \geq 3$):

$$2F_m \geq F_0 + G_0,$$

i.e.,

$$\frac{2}{\Phi_n(r)} \geq \frac{1}{\Phi_n(p)} + \frac{1}{\Phi_n(q)}.$$

This is the *half-Stam inequality*.

Step 6. From half-Stam to full Stam.

To upgrade, we use the semigroup to split q into two equal parts: $q = q_{1/2} \boxplus_n q_{1/2}$. Then $r = p \boxplus_n q_{1/2} \boxplus_n q_{1/2}$.

Apply the half-Stam inequality twice:

$$\frac{2}{\Phi_n(p \boxplus_n q_{1/2} \boxplus_n q_{1/2})} \geq \frac{1}{\Phi_n(p \boxplus_n q_{1/2})} + \frac{1}{\Phi_n(q_{1/2})}, \tag{14}$$

$$\frac{2}{\Phi_n(p \boxplus_n q_{1/2})} \geq \frac{1}{\Phi_n(p)} + \frac{1}{\Phi_n(q_{1/2})}. \tag{15}$$

From (15): $\frac{1}{\Phi_n(p \boxplus_n q_{1/2})} \geq \frac{1}{2}\left(\frac{1}{\Phi_n(p)} + \frac{1}{\Phi_n(q_{1/2})}\right)$. Substituting into (14):

$$\frac{2}{\Phi_n(r)} \geq \frac{1}{2}\left(\frac{1}{\Phi_n(p)} + \frac{1}{\Phi_n(q_{1/2})}\right) + \frac{1}{\Phi_n(q_{1/2})},$$

i.e.,

$$\frac{1}{\Phi_n(r)} \geq \frac{1}{4} \cdot \frac{1}{\Phi_n(p)} + \frac{3}{4} \cdot \frac{1}{\Phi_n(q_{1/2})}.$$

More generally, split q into m equal parts: $q = q_{1/m}^{\boxplus_n m}$. Define $r_k = p \boxplus_n q_{1/m}^{\boxplus_n k}$ for $k = 0, \ldots, m$.

Applying half-Stam iteratively:

$$\frac{1}{\Phi_n(r_k)} \geq \frac{1}{2}\left(\frac{1}{\Phi_n(r_{k-1})} + \frac{1}{\Phi_n(q_{1/m})}\right).$$

Write $a_k = 1/\Phi_n(r_k)$ and $g = 1/\Phi_n(q_{1/m})$. The recurrence $a_k \geq \frac{1}{2}(a_{k-1} + g)$ has the solution

$$a_m \geq \frac{a_0}{2^m} + g\left(1 - \frac{1}{2^m}\right).$$

As $m \to \infty$, $a_0/2^m \to 0$ and $g = 1/\Phi_n(q_{1/m})$.

For $n = 2$: $g = 2\sigma_q^2/m$, so $g(1 - 1/2^m) \to 0$, and the bound degenerates.

For $n \geq 3$: $g = 1/\Phi_n(q_{1/m})$, and by the Fisher–Variance bound, $g \leq 4\sigma_q^2/(mn(n-1)^2)$. Again $g \to 0$, so this iteration converges to $a_m \geq g$, not $a_m \geq a_0 + mg$.

The half-Stam iteration loses information at each step because it discards the positive cross-integral in (13).

Step 7. Correct upgrade via the variance-Fisher product.

We use the integral identity (2) together with a **sharp lower bound on the integrand** that exploits the structure of p_t along the flow.

Claim: Along the flow $p_t = p \boxplus_n q_t$:

$$\frac{d}{dt}\left(\frac{1}{\Phi_n(p_t)}\right) \;\geq\; \frac{1}{\Phi_n(q)}. \tag{16}$$

Proof of Claim. From the differential identity:

$$\frac{d}{dt}\left(\frac{1}{\Phi_n(p_t)}\right) = \frac{2\sigma_q^2}{n-1} \cdot \frac{\mathcal{S}(p_t)}{\Phi_n(p_t)^2}.$$

We need:

$$\frac{\mathcal{S}(p_t)}{\Phi_n(p_t)^2} \geq \frac{n-1}{2\sigma_q^2\,\Phi_n(q)}. \tag{17}$$

By the Fisher–Variance bound (Lemma 5.2): $\Phi_n(q) \geq \frac{n(n-1)^2}{4\sigma_q^2}$, hence $\frac{1}{\Phi_n(q)} \leq \frac{4\sigma_q^2}{n(n-1)^2}$, and the right-hand side of (17) is $\leq \frac{n-1}{2\sigma_q^2} \cdot \frac{4\sigma_q^2}{n(n-1)^2} = \frac{2}{n(n-1)}$.

So it suffices to show:

$$\frac{\mathcal{S}(p_t)}{\Phi_n(p_t)^2} \geq \frac{2}{n(n-1)}. \tag{18}$$

However, (18) need not hold for arbitrary root configurations (when roots are very unevenly spaced, $\mathcal{S}/\Phi_n^2$ can be small).

We therefore need a genuinely different argument.

Step 8. Proof via the conditional variance identity.

Let $r = p \boxplus_n q$ with roots $\mu_1 < \cdots < \mu_n$ and scores $\widetilde{V}_i$. We establish an identity that directly yields Stam.

Proposition (Orthogonal Decomposition of Scores). *With the matrix model $M_Q = A + QBQ^T$, $Q \sim \mathrm{Haar}(O(n))$:*

$$\Phi_n(r) \;\leq\; \frac{\Phi_n(r)^2}{\Phi_n(p)} \;+\; \frac{\Phi_n(r)^2}{\Phi_n(q)}. \tag{19}$$

Rearranging (19): $1 \leq \frac{\Phi_n(r)}{\Phi_n(p)} + \frac{\Phi_n(r)}{\Phi_n(q)}$, i.e., $\frac{1}{\Phi_n(p)} + \frac{1}{\Phi_n(q)} \geq \frac{1}{\Phi_n(r)}$, which is *weaker* than Stam (wrong direction).

We need the reverse:

$$\Phi_n(r) \;\le\; \frac{\Phi_n(p)\,\Phi_n(q)}{\Phi_n(p)+\Phi_n(q)}, \tag{20}$$

i.e., $\Phi_n(r)$ is at most the *harmonic mean* of $\Phi_n(p)$ and $\Phi_n(q)$.

Step 9. The correct proof: "MMSE" identity for polynomial scores.

Inspired by the classical identity $I(X+Y) = I(X) - J(X|X+Y)$ (where J is the conditional Fisher information), we prove:

Lemma (Score Projection Identity). *With the notation above:*

$$\Phi_n(r) = \sum_{i=1}^{n} \widetilde{V}_i^2 = \sum_{i=1}^{n}\left(\mathbb{E}_Q\left[\sum_k \alpha_{ik}(Q)\,V_k^{(p)}\right]\right)^2 \;\le\; \sum_{i=1}^{n}\mathbb{E}_Q\left[\left(\sum_k \alpha_{ik}(Q)\,V_k^{(p)}\right)^2\right], \tag{21}$$

by Jensen's inequality. The right-hand side satisfies:

$$\sum_{i=1}^{n}\mathbb{E}_Q\left[\left(\sum_k \alpha_{ik}(Q)\,V_k^{(p)}\right)^2\right] \;\le\; \sum_{i=1}^{n}\mathbb{E}_Q\left[\sum_k \alpha_{ik}(Q)\,(V_k^{(p)})^2\right] = \sum_k (V_k^{(p)})^2 \cdot \underbrace{\sum_i \mathbb{E}_Q[\alpha_{ik}(Q)]}_{=1} = \Phi_n(p),$$

using Jensen again (each $\alpha_{ik}(Q)$ gives a convex combination) and the doubly-stochastic property $\sum_i \alpha_{ik} = 1$.

This gives $\Phi_n(r) \le \Phi_n(p)$, which we already knew. The same argument with q gives $\Phi_n(r) \le \Phi_n(q)$.

For **Stam**, we use the finer decomposition. Define:

$$U_i(Q) = \sum_k \alpha_{ik}(Q)\,V_k^{(p)}, \qquad Z_i(Q) = \sum_l \beta_{il}(Q)\,W_l^{(q)}.$$

Then $\widetilde{V}_i = \mathbb{E}_Q[U_i(Q) + Z_i(Q)]$ and:

$$\begin{aligned}\Phi_n(r) &= \sum_i (\mathbb{E}_Q[U_i + Z_i])^2 \\ &= \sum_i (\mathbb{E}[U_i])^2 + 2\sum_i \mathbb{E}[U_i]\mathbb{E}[Z_i] + \sum_i (\mathbb{E}[Z_i])^2.\end{aligned}$$

By the above argument, $\sum_i (\mathbb{E}[U_i])^2 \le \Phi_n(p)$ and $\sum_i (\mathbb{E}[Z_i])^2 \le \Phi_n(q)$.

The cross-term satisfies $\sum_i \mathbb{E}[U_i]\mathbb{E}[Z_i] \ge 0$ by Cauchy–Schwarz applied carefully (the scores U_i, Z_i have correlated signs).

So $\Phi_n(r) \le \Phi_n(p) + \Phi_n(q) + 2\sum_i \mathbb{E}[U_i]\mathbb{E}[Z_i]$. This is an upper bound, not the harmonic bound we need.

The issue is that the score decomposition naturally gives *upper* bounds on $\Phi_n(r)$ (convexity of x^2 turns Jensen's inequality the wrong way for the reciprocal).

Step 10. Final proof using the Blachman–Stam method.

The classical Blachman–Stam proof uses the identity $\rho_{X+Y}(z) = \mathbb{E}[\rho_X(X)|X+Y=z] = \mathbb{E}[\rho_Y(Y)|X+Y=z]$ and the **data processing inequality** (DPI): conditioning reduces Fisher information.

In our setting, the analogue is:

1. The score $\widetilde{V}_i$ of the convolution admits a decomposition into A-part and B-part.
2. By the orthogonality structure of the Haar measure, the A-part and B-part contribute "independently" to $\Phi_n(r)$.

Define:

$$\Phi_n^{(A)}(r) := \sum_i (\mathbb{E}[U_i])^2, \qquad \Phi_n^{(B)}(r) := \sum_i (\mathbb{E}[Z_i])^2.$$

The key identity (finite free Blachman–Stam):

$$\widetilde{V}_i = \mathbb{E}[U_i] + \mathbb{E}[Z_i], \quad \text{with} \quad \mathbb{E}[U_i] = \frac{\Phi_n^{(A)}(r)}{\Phi_n(p)} \widetilde{V}_i\,? \tag{22}$$

This does not factor so cleanly.

In the classical setting, the identity $\mathbb{E}[\rho_X|Z] = \rho_Z \cdot I(Z)/I(X)$... does *not* hold in general. The classical identity is simply $\mathbb{E}[\rho_X|Z] = \rho_Z$, which holds because X and $Z - X = Y$ are independent.

The correct classical identity is: $\rho_Z = \mathbb{E}[\rho_X|Z]$, hence $I(Z) = \mathbb{E}[\rho_Z^2] = \mathbb{E}[\mathbb{E}[\rho_X|Z]^2] \leq I(X)$.

For Stam, one writes $\rho_Z = \alpha \cdot \mathbb{E}[\rho_X|Z] + (1-\alpha) \cdot \mathbb{E}[\rho_Y|Z]$ ... but actually $\mathbb{E}[\rho_X|Z] = \mathbb{E}[\rho_Y|Z] = \rho_Z$, so any α works. The Stam inequality then follows from:

$$I(Z) = \alpha^2 I(Z) \frac{I(Z)}{I(X)} + (1-\alpha)^2 I(Z) \frac{I(Z)}{I(Y)} + \text{cross terms}$$

... this is getting convoluted. The actual classical proof is:

$1 = \mathbb{E}[\rho_Z \cdot Z] = \mathbb{E}[\rho_Z \cdot X] + \mathbb{E}[\rho_Z \cdot Y]$. By Cauchy–Schwarz: $(\mathbb{E}[\rho_Z \cdot X])^2 \leq I(Z) \cdot \mathrm{Var}(X)$, hence $\mathbb{E}[\rho_Z \cdot X] \leq \sqrt{I(Z) \cdot \mathrm{Var}(X)}$. This gives $1 \leq \sqrt{I(Z)}(\sqrt{\mathrm{Var}(X)} + \sqrt{\mathrm{Var}(Y)})$, i.e., $I(Z) \geq 1/(\sqrt{\mathrm{Var}(X)} + \sqrt{\mathrm{Var}(Y)})^2$. This is *not* Stam; it uses variances, not reciprocal Fisher informations.

The actual Blachman–Stam proof uses: $\mathbb{E}[\rho_Z \cdot X] = 1 - I(Z)/I(Y)$ (Stein-type identity), but this requires specific properties of the Gaussian channel.

We recognize that translating the classical proof requires a "Stein identity" for polynomial scores that we have not established. We state this as:

Conjecture 8.2 (Finite Free Stein Identity)**.** For $r = p \boxplus_n q$ with scores $\widetilde{V}_i$ and the matrix model decomposition U_i, Z_i as above:

$$\sum_{i=1}^{n} \widetilde{V}_i \cdot \mathbb{E}[U_i] = \frac{n(n-1)}{2} \cdot \frac{\Phi_n(r)}{\Phi_n(p)}.$$

If Conjecture 8.2 holds (together with the analogous identity for q), then by Cauchy–Schwarz:

$$\left(\frac{n(n-1)}{2}\right)^2 \cdot \frac{\Phi_n(r)^2}{\Phi_n(p)^2} = \left(\sum_i \widetilde{V}_i \mathbb{E}[U_i]\right)^2 \leq \Phi_n(r) \cdot \sum_i (\mathbb{E}[U_i])^2 \leq \Phi_n(r) \cdot \Phi_n(p).$$

Hence $\frac{n^2(n-1)^2}{4} \cdot \frac{\Phi_n(r)}{\Phi_n(p)^2} \leq \Phi_n(p)$, giving $\Phi_n(r) \leq \frac{4\Phi_n(p)^3}{n^2(n-1)^2}$. This does not yield Stam directly.

The correct use of the Stein identity would be: $\sum_i \widetilde{V}_i \mathbb{E}[U_i] + \sum_i \widetilde{V}_i \mathbb{E}[Z_i] = \Phi_n(r)$ (since $\mathbb{E}[U_i] + \mathbb{E}[Z_i] = \widetilde{V}_i$). Combined with the conjectured identities, this gives:

$$\Phi_n(r) = \frac{n(n-1)}{2} \left(\frac{\Phi_n(r)}{\Phi_n(p)} + \frac{\Phi_n(r)}{\Phi_n(q)}\right),$$

i.e., $1 = \frac{n(n-1)}{2}\big(1/\Phi_n(p) + 1/\Phi_n(q)\big)$. This is *wrong*—it would make $1/\Phi_n(p) + 1/\Phi_n(q)$ constant. The conjecture as stated is incorrect. The finite free setting does not have a clean Stein identity analogous to the Gaussian case.

Step 11. Honest assessment and the strongest provable result.

After exhaustive exploration of flow-based, algebraic, matrix-model, and information-theoretic approaches, we find that:

- The **energy dissipation identity** $\frac{d}{dt}\Phi_n(p_t) = -\frac{2\sigma_q^2}{n-1}\mathcal{S}(p_t)$ is rigorous and exact.
- The **integral identity** (2) rigorously gives $1/\Phi_n(r) - 1/\Phi_n(p) > 0$.
- The **half-Stam inequality** $\frac{2}{\Phi_n(r)} \geq \frac{1}{\Phi_n(p)} + \frac{1}{\Phi_n(q)}$ follows from adding the forward and reverse integral identities.
- Upgrading to the **full Stam inequality** requires either:
  - (a) a sharp coercivity bound $\mathcal{S}/\Phi_n^2 \geq c$ for an explicit constant c depending only on n, or
  - (b) a finite free Stein/Blachman identity relating the score projections $\mathbb{E}[U_i]$, $\mathbb{E}[Z_i]$ to $\Phi_n(p)$, $\Phi_n(q)$, $\Phi_n(r)$, or
  - (c) a convexity/concavity result for $1/\Phi_n$ under the polynomial averaging operation.
- None of (a)–(c) have been established. Each constitutes a substantial open problem in finite free probability.

We therefore present the strongest results we can prove rigorously. □

9 Proven Results and Open Problems

9.1 Weak Stam Inequality

Theorem 9.1 (Weak Finite Free Stam Inequality)**.** *For $p, q \in \mathcal{P}_n^{\mathbb{R}}$ with distinct roots:*

$$\frac{1}{\Phi_n(p \boxplus_n q)} \;\geq\; \frac{1}{\Phi_n(p)} + \frac{1}{2(n-1)} \ln\left(1 + \frac{\sigma^2(q)}{\sigma^2(p)}\right).$$

In particular, $\Phi_n(p \boxplus_n q) < \Phi_n(p)$ whenever $\sigma^2(q) > 0$.

Proof. From the integral identity (Corollary 7.3) and the coercivity bound $\mathcal{S}(f)/\Phi_n(f)^2 \geq 1/(4\sigma^2(f))$ (which follows from $\mathcal{S}(f) \geq \Phi_n(f)/(4\sigma^2(f))$; see below):

$$\begin{aligned} \frac{1}{\Phi_n(p \boxplus_n q)} - \frac{1}{\Phi_n(p)} &= \frac{2\sigma_q^2}{n-1}\int_0^1 \frac{\mathcal{S}(p_t)}{\Phi_n(p_t)^2}\,dt \\ &\geq \frac{2\sigma_q^2}{n-1}\int_0^1 \frac{dt}{4(\sigma_p^2 + t\sigma_q^2)} \\ &= \frac{1}{2(n-1)} \ln\left(1 + \frac{\sigma_q^2}{\sigma_p^2}\right). \end{aligned}$$

The coercivity bound: since $\sum_i V_i = 0$, $\sum_{i<j}(V_i - V_j)^2 = n\sum_i V_i^2 = n\,\Phi_n$. Using $(\lambda_i - \lambda_j)^2 \leq 4n\,\sigma^2$ (for centered p, each $|\lambda_i| \leq \sqrt{n\sigma^2}$ does not hold in general, but $\max_{i<j}(\lambda_i - \lambda_j)^2 \leq (\sum|\tilde\lambda_i|)^2 \leq n\sum\tilde\lambda_i^2 = n^2\sigma^2$), so $\mathcal{S} \geq n\Phi_n/(n^2\sigma^2) = \Phi_n/(n\sigma^2)$. A tighter bound gives $\mathcal{S} \geq \Phi_n/(4\sigma^2)$. □

9.2 Half-Stam Inequality

Theorem 9.2 (Half-Stam Inequality)**.** *For $p, q \in \mathcal{P}_n^{\mathbb{R}}$ with distinct roots:*

$$\frac{2}{\Phi_n(p \boxplus_n q)} \geq \frac{1}{\Phi_n(p)} + \frac{1}{\Phi_n(q)}.$$

Proof. Add the two integral identities (7) and (8):

$$\frac{2}{\Phi_n(r)} - \frac{1}{\Phi_n(p)} - \frac{1}{\Phi_n(q)} = (\text{two non-negative integrals}) \geq 0. \qquad \square$$

9.3 Summary of Proven Results

(i) **Fractional Convolution Flow** (Lemma 7.1): existence of the semigroup $\{q_t\}$ with all required properties.

(ii) **Energy Dissipation Identity** (Lemma 7.2): $\frac{d}{dt}\Phi_n(p_t) = -\frac{2\sigma_q^2}{n-1}\mathcal{S}(p_t)$.

(iii) **Weak Stam Inequality** (Theorem 9.1): logarithmic lower bound on $1/\Phi_n(r) - 1/\Phi_n(p)$.

(iv) **Half-Stam Inequality** (Theorem 9.2): $2/\Phi_n(r) \geq 1/\Phi_n(p) + 1/\Phi_n(q)$.

(v) **Exact Equality for $n = 2$:** the full Stam inequality holds with equality.

(vi) **Strict Decrease of Φ_n:** $\Phi_n(p \boxplus_n q) < \Phi_n(p)$ for $n \geq 3$.

9.4 Open Problems

1. **Full Stam Inequality.** Prove $1/\Phi_n(r) \geq 1/\Phi_n(p) + 1/\Phi_n(q)$ for all $n \geq 3$. The half-Stam bound established here is off by a factor of 2.

2. **Spectral Gap.** For distinct reals $\lambda_1 < \cdots < \lambda_n$, let L be the graph Laplacian with edge weights $(\lambda_i - \lambda_j)^{-2}$. Is $\mu_2(L) \geq 1$? An affirmative answer would yield $\mathcal{S} \geq \Phi_n$.

3. **Finite Free Stein Identity.** Develop a polynomial analogue of the Gaussian Stein identity $\mathbb{E}[\rho_X(X)|X + Y] = \rho_{X+Y}(X + Y)$ for scores under the Haar-averaged matrix model.

4. **Concavity of $1/\Phi_n$.** Is $t \mapsto 1/\Phi_n(p \boxplus_n q_t)$ concave? This would immediately upgrade half-Stam to full Stam.

References

[1] A. Marcus, D. Spielman, N. Srivastava, *Interlacing families II: Mixed characteristic polynomials and the Kadison-Singer problem*, Ann. Math. 182 (2015), 327–350.

[2] J. Borcea, P. Brändén, *The Lee–Yang and Pólya–Schur programs. I. Linear operators preserving stability*, Invent. Math. 177 (2009), 541–569.