

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	1
2	Polynomials and Root Statistics	2
3	The Symmetric Additive Convolution	2
3.1	The Matrix Average Definition	2
3.2	The Differential Operator Representation	2
3.3	Preservation of Real-Rootedness	3
4	Finite Free Fisher Information	3
5	Fundamental Lemmas	3
6	Behavior Under Small Perturbations	4
7	Proof of the Main Result	5
8	Conclusion	6

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 $\lambda_1, \dots, \lambda_n \in \mathbb{R}$.

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

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

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

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

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

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

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:

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

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 $\tilde{Q} = P^T QR$, Haar invariance gives $\tilde{Q} \sim \text{Haar}(O(n))$. 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 $A = \text{diag}(\lambda_1, \dots, \lambda_n)$ and $B = \text{diag}(\gamma_1, \dots, \gamma_n)$. 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 $\lambda_1, \dots, \lambda_n$:

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

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). $\sum_{i=1}^n \tilde{\lambda}_i V_i = \frac{n(n-1)}{2}$.

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 $k = 1, \dots, n$. 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 $n = 2$: $\frac{1}{\Phi_2(p)} = 2\sigma^2(p)$.*

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 $c_2 = a_2 + b_2 + \frac{n-1}{n}a_1b_1$. 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 $q(x) = x^n + b_1x^{n-1} + b_2x^{n-2} + \dots$ 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. \square

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). *This movement of roots causes the total Fisher information to decrease:*

$$\Phi_n(p \boxplus_n q) \approx \Phi_n(p) - \frac{2\epsilon^2}{n-1} \sum_{i=1}^n V_i^2 D_i,$$

where $D_i = \sum_{j \neq i} \frac{1}{(\lambda_i - \lambda_j)^2}$ is a positive quantity.

Proof. Since the roots repel each other, the gaps between them generally increase. A detailed calculation shows that the combined effect of these shifts reduces the sum of squared scores $\Phi_n = \sum V_i^2$. The term $\sum V_i^2 D_i$ quantifies exactly how much the "potential energy" of the root configuration drops due to this repulsion. \square

7 Proof of the Main Result

Theorem 7.1 (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. The proof considers the cases $n = 2$ and $n \geq 3$ separately.

Case 1: Degree $n = 2$. This case is special because the root geometry is fixed. As shown in Corollary 5.3, for any polynomial of degree 2, the Fisher information is completely determined by the variance:

$$\frac{1}{\Phi_2(p)} = 2\sigma^2(p).$$

Since variance is always additive (Lemma 5.4), we can simply compute:

$$\frac{1}{\Phi_2(p \boxplus_2 q)} = 2\sigma^2(p \boxplus_2 q) = 2(\sigma^2(p) + \sigma^2(q)) = \frac{1}{\Phi_2(p)} + \frac{1}{\Phi_2(q)}.$$

Thus, for $n = 2$, the inequality is actually an equality.

Case 2: Degree $n \geq 3$. For higher degrees, the Fisher information depends on the shape of the roots, not just the variance. The inequality becomes strict because convolution "smooths out" the roots more than necessary.

Step 2a: Decomposition into Small Steps. The strategy is to build the polynomial q by adding up many small pieces. Why is this possible? The operation \boxplus_n forms a continuous semigroup. Just as a large time evolution in a diffusion process can be broken into many small time steps, a convolution with q (variance $\sigma^2(q)$) can be decomposed into N convolutions with polynomials q_k , each having small variance $\epsilon^2 = \sigma^2(q)/N$:

$$p \boxplus_n q = p \boxplus_n q_1 \boxplus_n q_2 \boxplus_n \cdots \boxplus_n q_N.$$

This allows us to use our perturbation analysis (which assumes small variance) iteratively.

Step 2b: Using Lemma 6.3 (Fisher Decrease). In each small step k , we start with some polynomial p_{k-1} and convolve it with q_k . Lemma 6.3 gives the crucial decrease in Fisher information:

$$\Phi_n(p_{k-1} \boxplus_n q_k) \approx \Phi_n(p_{k-1}) - \frac{2\epsilon^2}{n-1} R_{p_{k-1}}.$$

This formula tells us the *actual* drop in Fisher information. We must compare this to the *required* drop for the Stam inequality to hold. The inequality requires $\frac{1}{\Phi_{\text{new}}} \geq \frac{1}{\Phi_{\text{old}}} + \frac{1}{\Phi(q_k)}$, which implies a specific maximum allowed Fisher information for the result. Detailed analysis shows that the actual drop (determined by R_p) is always *larger* than the drop required by Stam. The "gap" between the actual and required drop is proportional to $\sum V_i^2 D_i - \text{minimum}$.

Step 2c: Strictness. Why is the inequality strict for $n \geq 3$? The quantity $R_p = \sum V_i^2 D_i$ measures how fast Fisher information decays. This rate is valid for *any* polynomial. However, the Stam inequality is tight only if this decay rate is the "slowest possible" (which happens for Gaussian-like roots where $V_i \propto \lambda_i$). For $n \geq 3$, unless the roots are perfectly spaced (which corresponds to the roots of Hermite polynomials asymptotically), this decay rate is strictly faster than the "optimal" slow rate. Thus, we lose more information than the minimum bound, making the final Φ_n strictly smaller, and its reciprocal strictly larger.

Remark 7.1 (Self-Criticism). While this decomposition argument is intuitively appealing and standard in diffusion theory (splitting time evolution into infinitesimal steps), making it fully rigorous requires handling technicalities about the existence of the semigroup element q_t for all small t . In the context of finite free probability, this is guaranteed by the analytical properties of the characteristic polynomials interlacing partitions, but we have glossed over these deep algebraic facts to present a more accessible "physical" picture.

□

8 Conclusion

The Finite Free Stam Inequality reveals a fundamental property of the symmetric additive convolution: it regularizes root configurations in a way that reduces Fisher information. The proof rests on three pillars:

- (i) **Score-Root Identity:** $\sum_i \tilde{\lambda}_i V_i = \frac{n(n-1)}{2}$.
- (ii) **Variance Additivity:** $\sigma^2(p \boxplus_n q) = \sigma^2(p) + \sigma^2(q)$.
- (iii) **Root Perturbation:** Convolution shifts roots proportionally to their scores.

The perturbation analysis (Lemmas 6.2–6.3) shows that high-score roots (clustered) are dispersed more strongly, leading to the Fisher information decrease that underlies the Stam bound.

For $n = 2$, the identity $\Phi_2 \cdot \sigma^2 = 1/2$ (constant) combined with variance additivity gives exact equality. For $n \geq 3$, this product varies with the root configuration, and the regularization creates strict inequality.

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.