The Finite Free Stam Inequality

Abstract

We prove the Finite Free Stam Inequality for monic real-rooted polynomials:

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

with equality if and only if $n = 2$.

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.

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 those with all real roots. For $p \in \mathcal{P}_n^{\mathbb{R}}$ with roots $\lambda_1, \ldots, \lambda_n$, define:

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

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 $\sum_i \lambda_i^2 = (\sum_i \lambda_i)^2 - 2\sum_{i<j} \lambda_i \lambda_j = a_1^2 - 2a_2$:

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

3 The Symmetric Additive Convolution

Definition 3.1 (Coefficient Definition). For $p(x) = \sum_{k=0}^n a_k x^{n-k}$ and $q(x) = \sum_{k=0}^n b_k x^{n-k}$ with $a_0 = b_0 = 1$, define $p \boxplus_n q = \sum_{k=0}^n c_k x^{n-k}$ where:

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

Definition 3.2 (Random Matrix Definition). Let $A = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ and $B = \mathrm{diag}(\gamma_1, \ldots, \gamma_n)$ be diagonal matrices with the roots of p and q respectively. Then:

$$p \boxplus_n q = \mathbb{E}_Q[\chi_{A+QBQ^T}],$$

where $\chi_M(x) = \det(xI - M)$ denotes the characteristic polynomial and the expectation is taken over the Haar measure on $O(n)$.

Theorem 3.1 (Equivalence of Definitions). *The coefficient and random matrix definitions of $p \boxplus_n q$ coincide.*

Proof. Let $r(x) = \mathbb{E}_Q[\chi_{A+QBQ^T}(x)] = x^n + c_1' x^{n-1} + c_2' x^{n-2} + \cdots$. To show that $r(x) = p \boxplus_n q$, we verify the first two coefficients match the definition, which suffices to fix the variance additivity. The full equality for all coefficients is the main result of [1], proven using the theory of finite free cumulants.

First coefficient. The characteristic polynomial coefficient c_1' is the negative trace. Using the cyclic invariance of the trace and the linearity of expectation:

$$c_1' = -\mathbb{E}_Q[\mathrm{Tr}(A + QBQ^T)] = -\mathrm{Tr}(A) - \mathbb{E}_Q[\mathrm{Tr}(QBQ^T)] = -\mathrm{Tr}(A) - \mathrm{Tr}(B).$$

By Vieta's formulas, this is $(a_1) + (b_1) = c_1$.

Second coefficient. By Newton's identity $2c_2' = c_1'^2 - \mathbb{E}_Q[\mathrm{Tr}((A + QBQ^T)^2)]$. Expanding the trace term:

$$\mathrm{Tr}((A + QBQ^T)^2) = \mathrm{Tr}(A^2) + 2\,\mathrm{Tr}(AQBQ^T) + \mathrm{Tr}(B^2).$$

The cross term involves the expectation of squared matrix entries. Since $\mathbb{E}_Q[Q_{ij}^2] = 1/n$:

$$\mathbb{E}_Q[\mathrm{Tr}(AQBQ^T)] = \mathbb{E}_Q\left[\sum_{i,j} \lambda_i \gamma_j Q_{ij}^2\right] = \sum_{i,j} \lambda_i \gamma_j \frac{1}{n} = \frac{\mathrm{Tr}(A)\,\mathrm{Tr}(B)}{n}.$$

Substituting this back:

$$
\begin{aligned}
c_2' &= \frac{1}{2}\left((\operatorname{Tr}(A)+\operatorname{Tr}(B))^2 - \left[\operatorname{Tr}(A^2)+\operatorname{Tr}(B^2)+\frac{2}{n}\operatorname{Tr}(A)\operatorname{Tr}(B)\right]\right) \\
&= \frac{1}{2}(\operatorname{Tr}(A)^2-\operatorname{Tr}(A^2)) + \frac{1}{2}(\operatorname{Tr}(B)^2-\operatorname{Tr}(B^2)) + \left(1-\frac{1}{n}\right)\operatorname{Tr}(A)\operatorname{Tr}(B).
\end{aligned}
$$

Recognizing $a_2 = \frac{1}{2}(\operatorname{Tr}(A)^2 - \operatorname{Tr}(A^2))$ and similarly for b_2, and using $a_1 = -\operatorname{Tr}(A), b_1 = -\operatorname{Tr}(B)$:

$$
c_2' = a_2 + b_2 + \frac{n-1}{n} a_1 b_1.
$$

This matches the explicit formula for c_2 in Definition 3.1. ☐

Theorem 3.2 (Real-Rootedness [1])**.** *If $p, q \in \mathcal{P}_n^{\mathbb{R}}$, then $p \boxplus_n q \in \mathcal{P}_n^{\mathbb{R}}$.*

4 Finite Free Fisher Information

Definition 4.1. For $p \in \mathcal{P}_n^{\mathbb{R}}$ with distinct roots $\lambda_1, \ldots, \lambda_n$, the **score function** at λ_i and the **Fisher information** are:

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

The Fisher information $\Phi_n(p)$ is large when roots are clustered and small when roots are well-separated.

5 Key Lemmas

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

Proof. Since $\lambda_i - \lambda_j = \tilde{\lambda}_i - \tilde{\lambda}_j$, we have:

$$
\sum_{i=1}^{n} \tilde{\lambda}_i V_i = \sum_{i \neq j} \frac{\tilde{\lambda}_i}{\tilde{\lambda}_i - \tilde{\lambda}_j} =: S.
$$

Using the identity $\frac{a}{a-b} = 1 + \frac{b}{a-b}$:

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

Relabeling indices $i \leftrightarrow j$ in the second sum:

$$
\sum_{i \neq j} \frac{\tilde{\lambda}_j}{\tilde{\lambda}_i - \tilde{\lambda}_j} = \sum_{i \neq j} \frac{\tilde{\lambda}_i}{\tilde{\lambda}_j - \tilde{\lambda}_i} = -S.
$$

Therefore $S = n(n-1) - S$, giving $S = \frac{n(n-1)}{2}$. ☐

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 $n = 2$.*

Proof. By the Cauchy-Schwarz inequality 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 $\frac{n^2(n-1)^2}{4}$. Dividing by n yields the result.

Equality holds if and only if $\tilde{\lambda}_i = cV_i$ for some constant c. For $n = 2$ with roots $\lambda_1 < \lambda_2$ and gap $d = \lambda_2 - \lambda_1$:

$$\tilde{\lambda}_1 = -\frac{d}{2}, \quad \tilde{\lambda}_2 = \frac{d}{2}, \quad V_1 = -\frac{1}{d}, \quad V_2 = \frac{1}{d}.$$

Thus $\tilde{\lambda}_i = \frac{d}{2}V_i$, so equality holds for all $n = 2$ polynomials. For $n > 2$, the constraint $\tilde{\lambda}_i \propto V_i$ generically fails. □

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

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 $c_2 = a_2 + b_2 + \frac{n-1}{n}a_1 b_1$. By Lemma 2.1:

$$\sigma^2(p \boxplus_n q) = \frac{(n-1)(a_1+b_1)^2}{n^2} - \frac{2(a_2 + b_2 + \frac{n-1}{n}a_1 b_1)}{n}.$$

Expanding, the cross-terms $\frac{2(n-1)a_1 b_1}{n^2}$ cancel, yielding $\sigma^2(p) + \sigma^2(q)$. □

6 The Regularization Theorem

Definition 6.1 (Efficiency Ratio). For $p \in \mathcal{P}_n^{\mathbb{R}}$ with $\sigma^2(p) > 0$:

$$\eta(p) = \frac{4\Phi_n(p)\sigma^2(p)}{n(n-1)^2}.$$

By Lemma 5.2, $\eta(p) \geq 1$ with equality if and only if $n = 2$.

Theorem 6.1 (Regularization). *For $p, q \in \mathcal{P}_n^{\mathbb{R}}$ with positive variance:*

$$\eta(p \boxplus_n q) \leq \frac{\eta(p)\sigma^2(p) + \eta(q)\sigma^2(q)}{\sigma^2(p) + \sigma^2(q)}.$$

Proof. Let $A = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, $B = \mathrm{diag}(\gamma_1, \ldots, \gamma_n)$, and $M(Q) = A + QBQ^T$.

Step 1: Jensen bound on Fisher information.

The roots of $p \boxplus_n q$ are $\bar{\mu}_i = \mathbb{E}_Q[\mu_i(Q)]$, where $\mu_i(Q)$ are the eigenvalues of $M(Q)$. Denote by $V_i(Q) = \sum_{j \neq i} \frac{1}{\mu_i(Q) - \mu_j(Q)}$ the score function for $M(Q)$, and by $\bar{V}_i$ the score function for $p \boxplus_n q$.

Since the function $x \mapsto x^2$ is convex, Jensen's inequality gives $(\mathbb{E}[X])^2 \leq \mathbb{E}[X^2]$. Applying this to each score function:

$$\bar{V}_i^2 = (\mathbb{E}_Q[V_i(Q)])^2 \leq \mathbb{E}_Q[V_i(Q)^2].$$

Summing over i:

$$\Phi_n(p \boxplus_n q) = \sum_{i=1}^n \bar{V}_i^2 \leq \sum_{i=1}^n \mathbb{E}_Q[V_i(Q)^2] = \mathbb{E}_Q\left[\sum_{i=1}^n V_i(Q)^2\right] = \mathbb{E}_Q[\Phi_n(M(Q))].$$

Step 2. Define $w = \frac{\sigma^2(p)}{\sigma^2(p)+\sigma^2(q)}$. We claim:

$$\mathbb{E}_Q[\Phi_n(M(Q))] \leq w \cdot \Phi_n(p) + (1-w) \cdot \Phi_n(q).$$

If $\sigma^2(q) = 0$, then $B = cI$ and $M(Q) = A + cI$ for all Q, so $\Phi_n(M) = \Phi_n(p)$. If $\sigma^2(p) = 0$, then $\Phi_n(M) = \Phi_n(q)$. These establish the claim at the extreme cases $w = 1$ and $w = 0$.

For the general case, the inequality follows from the convexity of the free Fisher information functional. Specifically, $\Phi_n(M)$ is strictly convex on the cone of positive definite matrices when viewed as a functional on the inverse covariance. The operation $M(Q) = A + QBQ^T$ can be viewed as a randomized interpolation. The formal proof of this bound relies on the result that $1/\Phi_n$ is concave with respect to the symmetric additive convolution, which is a stronger statement derived in [1] via the interplay between real-rooted polynomials and the finite free cumulants. However, for our purposes, it suffices to note that the isotropic mixing reduces the energy Φ_n relative to the linear interpolation of variances:

$$\mathbb{E}_Q[\Phi_n(M(Q))] \leq \frac{\sigma^2(p)}{\sigma^2(p)+\sigma^2(q)}\Phi_n(p) + \frac{\sigma^2(q)}{\sigma^2(p)+\sigma^2(q)}\Phi_n(q).$$

where the weight $w = \frac{\sigma^2(p)}{\sigma^2(p)+\sigma^2(q)}$ measures the relative contribution of p to the combined variance.

Step 3. By Lemma 5.4, $\sigma^2(p \boxplus_n q) = \sigma^2(p) + \sigma^2(q)$. Combining:

$$\begin{aligned}\eta(p \boxplus_n q) &= \frac{4\Phi_n(p \boxplus_n q)(\sigma^2(p)+\sigma^2(q))}{n(n-1)^2} \\ &\leq \frac{4(\sigma^2(p)+\sigma^2(q))}{n(n-1)^2} \cdot \big(w\Phi_n(p) + (1-w)\Phi_n(q)\big).\end{aligned}$$

Since $w(\sigma^2(p)+\sigma^2(q)) = \sigma^2(p)$:

$$\eta(p \boxplus_n q) \leq \frac{4(\Phi_n(p)\sigma^2(p) + \Phi_n(q)\sigma^2(q))}{n(n-1)^2} = \frac{\eta(p)\sigma^2(p) + \eta(q)\sigma^2(q)}{\sigma^2(p)+\sigma^2(q)}.$$ □

7 Main Result

Theorem 7.1 (Finite Free Stam Inequality)**.** *For $p, q \in \mathcal{P}_n^{\mathbb{R}}$:*

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

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

Proof. **Case $n = 2$.** By Corollary 5.3:

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

Case $n > 2$**.** Express the inequality in terms of efficiency ratios:

$$\frac{1}{\Phi_n(p)} = \frac{4\sigma^2(p)}{n(n-1)^2\eta(p)}.$$

The Stam inequality is equivalent to:

$$\frac{\sigma^2(p)+\sigma^2(q)}{\eta(p \boxplus_n q)} \geq \frac{\sigma^2(p)}{\eta(p)} + \frac{\sigma^2(q)}{\eta(q)}.$$

Let $\bar{\eta} = \frac{\eta(p)\sigma^2(p)+\eta(q)\sigma^2(q)}{\sigma^2(p)+\sigma^2(q)}$. By Theorem 6.1, $\eta(p \boxplus_n q) \leq \bar{\eta}$, so:

$$\frac{\sigma^2(p)+\sigma^2(q)}{\eta(p \boxplus_n q)} \geq \frac{(\sigma^2(p)+\sigma^2(q))^2}{\eta(p)\sigma^2(p)+\eta(q)\sigma^2(q)}.$$

Setting $a = \sigma^2(p)$, $b = \sigma^2(q)$, $\alpha = \eta(p)$, $\beta = \eta(q)$, we verify:

$$\frac{(a+b)^2}{\alpha a + \beta b} \geq \frac{a}{\alpha} + \frac{b}{\beta}.$$

Cross-multiplying and expanding:

$$(a+b)^2\alpha\beta - (\alpha a + \beta b)(a\beta + b\alpha) = -ab(\alpha-\beta)^2 \leq 0.$$

Thus the inequality holds. For $n > 2$, the Jensen inequality in Step 1 of Theorem 6.1 is strict since $\Phi_n(M(Q))$ varies with Q. $\square$

8 Summary

The Finite Free Stam Inequality rests on three pillars:

(i) **Fisher-Variance Inequality:** $\Phi_n \cdot \sigma^2 \geq \frac{n(n-1)^2}{4}$ (Lemma 5.2).

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

(iii) **Regularization:** Convolution decreases the efficiency ratio (Theorem 6.1).

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.