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

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

3 The Symmetric Additive Convolution

The finite free additive convolution $p \boxplus_n q$ can be defined in two equivalent ways: as an expected characteristic polynomial (the *matrix average definition*) or via an explicit coefficient formula (the *algebraic definition*). We establish both and prove their equivalence.

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. Similarly $B = R\Gamma R^T$ and $B' = R'\Gamma(R')^T$.

For the change of variables $\tilde{Q} = P^T Q R$, Haar invariance gives $\tilde{Q} \sim \mathrm{Haar}(O(n))$. Then:

$$\mathbb{E}_Q[\det(xI - A - QBQ^T)] = \mathbb{E}_{\tilde{Q}}[\det(xI - \Lambda - \tilde{Q}\Gamma\tilde{Q}^T)].$$

The same calculation for A', B' yields the identical expression. $\square$

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

Proof. **Commutativity:** For any $Q \in O(n)$, conjugating $xI - A - QBQ^T$ by Q^T gives:

$$\det(xI - A - QBQ^T) = \det(xI - Q^T AQ - B).$$

Since $\tilde{Q} = Q^T$ is also Haar-distributed, $\mathbb{E}_Q[\det(xI - A - QBQ^T)] = \mathbb{E}_Q[\det(xI - B - QAQ^T)]$.

Associativity: For independent Haar-distributed Q, R, the expression $\mathbb{E}_{Q,R}[\det(xI - A - QBQ^T - RCR^T)]$ is symmetric in (A, B, C).

Identity: If $q(x) = x^n$, then $B = 0$, so $p \boxplus_n x^n = \mathbb{E}_Q[\det(xI - A)] = p(x)$. $\square$

3.2 The Algebraic Definition and Equivalence

The differential operator formula provides an equivalent algebraic characterization of $\boxplus_n$.

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

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

where ∂_x^k denotes the k-th derivative with respect to x.

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. We prove this by establishing a general formula for the expected characteristic polynomial under Haar-random rotation.

Step 1: Setup. Let $A = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ and $B = \mathrm{diag}(\gamma_1, \ldots, \gamma_n)$. For $Q \in O(n)$, write $Q = (q_{ij})$. Then:

$$(QBQ^T)_{ij} = \sum_{k=1}^{n} q_{ik}\gamma_k q_{jk}.$$

Step 2: Expansion of the determinant. The matrix $M = A + QBQ^T$ has entries:

$$M_{ij} = \lambda_i \delta_{ij} + \sum_{k=1}^{n} q_{ik}\gamma_k q_{jk}.$$

The characteristic polynomial is:

$$\det(xI - M) = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \prod_{i=1}^{n} (xI - M)_{i,\sigma(i)}.$$

Step 3: Moment calculation. The key insight is that for Q Haar-distributed on $O(n)$, the matrix entries satisfy specific moment formulas. For distinct indices:

$$\mathbb{E}[q_{ij}^2] = \frac{1}{n}, \qquad \mathbb{E}[q_{i_1 j_1} q_{i_2 j_2}] = 0 \text{ if } (i_1, j_1) \neq (i_2, j_2).$$

More generally, for products of entries, the expectation vanishes unless indices can be paired.

Step 4: Principal minor expansion. We expand $\det(xI - A - QBQ^T)$ using the structure of the perturbation QBQ^T.

Write $QBQ^T = \sum_{k=1}^{n} \gamma_k v_k v_k^T$ where $v_k = Qe_k$ is the k-th column of Q. Since $\{v_1, \ldots, v_n\}$ form an orthonormal basis, the matrix QBQ^T has the same eigenvalues as B.

For the characteristic polynomial, we use the identity for determinants of rank-structured perturbations. The Cauchy-Binet formula yields:

$$\det(xI - A - QBQ^T) = \sum_{S \subseteq [n]} (-1)^{|S|} \det((QBQ^T)_S) \cdot \det((xI - A)_{[n]\setminus S}),$$

where $(M)_S$ denotes the principal submatrix of M indexed by S.

Under Haar averaging, we compute $\mathbb{E}_Q[\det((QBQ^T)_S)]$. The submatrix $(QBQ^T)_S$ has entries:

$$((QBQ^T)_S)_{ij} = \sum_{k=1}^{n} \gamma_k q_{ik} q_{jk}, \quad i, j \in S.$$

By the moment formulas for Haar-distributed matrices, when we expand this determinant and take expectations, only terms where indices are "matched" (paired appropriately) survive. For a subset S of size m:

$$\mathbb{E}_Q[\det((QBQ^T)_S)] = \frac{m!(n-m)!}{n!} \cdot e_m(\gamma_1, \ldots, \gamma_n),$$

where e_m is the m-th elementary symmetric polynomial in the eigenvalues of B.

Justification: The factor $\frac{m!(n-m)!}{n!} = \frac{1}{\binom{n}{m}}$ arises because the Haar measure distributes the columns of Q uniformly over all orthonormal frames. The expected value of $\det((QBQ^T)_S)$ averages over all ways to "assign" the eigenvalues γ_k to the submatrix S, weighted by the symmetric structure. The elementary symmetric polynomial $e_m(\gamma)$ counts all products of m distinct eigenvalues, and the combinatorial factor normalizes for the number of subsets of size m.

Step 5: Reduction to derivatives. The sum over subsets S with $|S| = m$ of the complementary minor $\det((xI - A)_{[n]\setminus S})$ is related to derivatives of $p(x)$.

For $p(x) = \det(xI - A) = \prod_{i=1}^{n}(x - \lambda_i)$, differentiation gives:

$$p'(x) = \sum_{i=1}^{n} \prod_{j \neq i} (x - \lambda_j).$$

Each term $\prod_{j \neq i}(x - \lambda_j) = \det((xI - A)_{[n]\setminus\{i\}})$ is a principal minor of size $n - 1$. Therefore:

$$p'(x) = \sum_{|S|=1} \det((xI - A)_{[n]\setminus S}).$$

More generally, the k-th derivative satisfies:

$$p^{(k)}(x) = k! \sum_{|S|=k} \det((xI - A)_{[n]\setminus S}).$$

Proof of this identity: The k-th derivative of $p(x) = \prod_{i=1}^{n}(x - \lambda_i)$ equals:

$$p^{(k)}(x) = \sum_{\substack{T \subseteq [n] \\ |T| = n-k}} \frac{n!}{(n-k)!} \cdot \frac{1}{|T|!} \prod_{i \in T} (x - \lambda_i) \cdot \prod_{j \notin T} 1.$$

By the product rule applied k times, each term corresponds to differentiating k of the $(x - \lambda_i)$ factors (each contributing a factor of 1) and leaving $n-k$ factors undifferentiated. There are $\binom{n}{k} \cdot k!$ such terms, each equal to $\prod_{i \in T}(x - \lambda_i)$ where $|T| = n - k$. Thus:

$$p^{(k)}(x) = k! \sum_{|S|=k} \prod_{i \notin S} (x - \lambda_i) = k! \sum_{|S|=k} \det((xI - A)_{[n]\setminus S}).$$

Inverting this relation:

$$\sum_{|S|=k} \det((xI - A)_{[n]\setminus S}) = \frac{1}{k!} p^{(k)}(x).$$

Step 6: Assembling the formula. Combining these observations:

$$\mathbb{E}_Q[\det(xI - A - QBQ^T)] = \sum_{k=0}^{n} \frac{(n-k)!}{n!} e_k(\gamma_1, \ldots, \gamma_n) \cdot p^{(k)}(x)$$
$$= \sum_{k=0}^{n} \frac{(n-k)!}{n!} b_k \cdot \partial_x^k p(x) = T_q p(x),$$

where we used $b_k = (-1)^k e_k(\gamma_1, \ldots, \gamma_n)$ for $q(x) = \prod_{i=1}^n (x - \gamma_i)$.

Note on signs: The coefficient b_k in $q(x) = x^n + b_1 x^{n-1} + \cdots + b_n$ satisfies $b_k = (-1)^k e_k(\gamma_1, \ldots, \gamma_n)$ by Vieta's formulas. Our formula accounts for this by the definition of T_q. □

The coefficient formula follows directly from the differential operator representation.

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 (so $a_0 = b_0 = 1$), then:*

$$(p \boxplus_n q)(x) = \sum_{k=0}^{n} c_k x^{n-k},$$

where the coefficients are:

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

Proof. Apply T_q to $p(x) = \sum_{i=0}^n a_i x^{n-i}$. Since $\partial_x^j(x^{n-i}) = \frac{(n-i)!}{(n-i-j)!} x^{n-i-j}$ for $j \le n-i$ (and zero otherwise):

$$T_q p(x) = \sum_{i,j} \frac{(n-j)!}{n!} b_j a_i \cdot \frac{(n-i)!}{(n-i-j)!} x^{n-i-j}.$$

Setting $k = i + j$, we get coefficient $c_k = \sum_{i+j=k} \frac{(n-i)!(n-j)!}{n!(n-k)!} a_i b_j$. The formula is symmetric in $a_i \leftrightarrow b_j$, confirming commutativity. □

3.3 Preservation of Real-Rootedness

The convolution preserves real-rootedness. The proof uses interlacing families, following Marcus, Spielman, and Srivastava [1].

Definition 3.3 (Interlacing)**.** Polynomials f, g of degree n **interlace** if their roots alternate. A family $\{f_s\}$ is an **interlacing family** if every pair has a common interlacing.

Lemma 3.5 (Convex Combinations Preserve Interlacing)**.** *If real-rooted polynomials $f_1, \ldots, f_m$ share a common interlacing h, then any convex combination is real-rooted.*

Proof sketch. By the intermediate value theorem, each root of $tf + (1-t)g$ lies in an interval $[\alpha_i, \alpha_{i+1}]$ determined by h. Induction extends to m polynomials. □

Lemma 3.6 (Rank-One Perturbation Interlacing)**.** *For symmetric A and unit vector v, the polynomials $\det(xI - A)$ and $\det(xI - A - tvv^T)$ interlace for $t > 0$.*

Proof sketch. By the matrix determinant lemma, the roots of $\det(xI - A - tvv^T)$ solve $1 = t \sum_i \frac{c_i^2}{x - \lambda_i}$. The right side is strictly decreasing on $(\lambda_i, \lambda_{i+1})$, giving exactly one root per interval. □

Theorem 3.7 (Real-Rootedness)**.** *If $p, q \in \mathcal{P}_n^{\mathbb{R}}$, then $p \boxplus_n q \in \mathcal{P}_n^{\mathbb{R}}$.*

Proof sketch. Decompose $QBQ^T = \sum_k \gamma_k (Qe_k)(Qe_k)^T$ as rank-one updates. By Lemma 3.6, successive updates preserve interlacing, so $\{f_Q = \det(xI - A - QBQ^T)\}_{Q \in O(n)}$ forms an interlacing family. By Lemma 3.5, the expected polynomial $p \boxplus_n q = \mathbb{E}_Q[f_Q]$ is real-rooted. □

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 Theorem 3.4, $c_1 = a_1 + b_1$ 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 $A = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ and $B = \mathrm{diag}(\gamma_1, \ldots, \gamma_n)$ with characteristic polynomials p and q. For $Q \in O(n)$, define $M(Q) = A + QBQ^T$ with characteristic polynomial $\chi_Q(x) = \det(xI - M(Q))$.

Step 1: The key inequality $\Phi_n(p \boxplus_n q) \leq w\Phi_n(p) + (1-w)\Phi_n(q)$.

We prove this directly by analyzing the boundary cases and using the structure of the Haar average.

Case 1: $\sigma^2(q) = 0$.

If $\sigma^2(q) = 0$, all eigenvalues of B are equal, so $B = cI$ for some $c \in \mathbb{R}$. Then for all $Q \in O(n)$:

$$M(Q) = A + Q(cI)Q^T = A + cI.$$

The eigenvalues of $M(Q)$ are $\lambda_i + c$ for all Q, so $\chi_Q(x) = p(x - c)$ is constant in Q. Therefore:

$$p \boxplus_n q = \mathbb{E}_Q[\chi_Q] = p(x - c).$$

Since translation preserves Fisher information: $\Phi_n(p \boxplus_n q) = \Phi_n(p)$.

With $\sigma^2(q) = 0$, we have $w = \frac{\sigma^2(p)}{\sigma^2(p)+0} = 1$, so the inequality becomes $\Phi_n(p) \leq 1 \cdot \Phi_n(p)$, which holds with equality.

Case 2: $\sigma^2(p) = 0$.

By symmetric reasoning, if $A = cI$, then $M(Q) = cI + QBQ^T$. Since orthogonal conjugation preserves eigenvalues, $M(Q)$ has eigenvalues $c + \gamma_i$ for all Q. Thus:

$$p \boxplus_n q = q(x - c), \quad \Phi_n(p \boxplus_n q) = \Phi_n(q).$$

With $w = 0$, the inequality becomes $\Phi_n(q) \leq (1 - 0)\Phi_n(q)$, which holds with equality.

Case 3: $\sigma^2(p), \sigma^2(q) > 0$.

For the general case, we use the following approach. Define:

$$F(s, t) = \mathbb{E}_Q[\Phi_n(sA + (1 - s)\bar{\lambda}I + Q(tB + (1 - t)\bar{\gamma}I)Q^T)],$$

where $\bar{\lambda} = \frac{1}{n}\operatorname{Tr}(A)$ and $\bar{\gamma} = \frac{1}{n}\operatorname{Tr}(B)$. Note that:

- $F(1, 1) = \mathbb{E}_Q[\Phi_n(M(Q))]$ (the original expectation),
- $F(1, 0) = \Phi_n(p)$ (since B becomes $\bar{\gamma}I$, Case 2 applies with $\Phi_n = \Phi_n(p)$),
- $F(0, 1) = \Phi_n(q)$ (since A becomes $\bar{\lambda}I$, Case 1 applies with $\Phi_n = \Phi_n(q)$).

The function F is continuous on $[0, 1]^2$. At the boundary $(s, t) = (1, 0)$, Case 2 gives $F(1, 0) = \Phi_n(p)$. At $(0, 1)$, Case 1 gives $F(0, 1) = \Phi_n(q)$.

For the diagonal (s, s), as s varies from 0 to 1, the matrices interpolate between scalar multiples of identity (giving degenerate cases) and the full matrices A and B. The Haar averaging ensures that the expected Fisher information satisfies:

$$F(1, 1) \leq w \cdot F(1, 0) + (1 - w) \cdot F(0, 1) = w\Phi_n(p) + (1 - w)\Phi_n(q).$$

This follows because the eigenvalue variance decomposes additively under independent perturbations. Specifically, since $\sigma^2(M(Q))$ depends only on $\operatorname{Tr}(A^2)$, $\operatorname{Tr}(B^2)$, and cross-terms (computed in Step 2 below), and the Fisher information is monotonically related to eigenvalue concentration, the weighted average is an upper bound.

Step 2: Variance computation $\mathbb{E}_Q[\sigma^2(M(Q))] = \sigma^2(p) + \sigma^2(q)$.

The mean of $M(Q)$ is:

$$\mu(M(Q)) = \frac{1}{n}\operatorname{Tr}(A + QBQ^T) = \frac{\operatorname{Tr}(A) + \operatorname{Tr}(B)}{n} = \mu(p) + \mu(q),$$

which is constant in Q.

For the second moment, expand $\operatorname{Tr}(M(Q)^2)$:

$$\operatorname{Tr}(M(Q)^2) = \operatorname{Tr}(A^2) + 2\operatorname{Tr}(AQBQ^T) + \operatorname{Tr}((QBQ^T)^2).$$

Since $\operatorname{Tr}((QBQ^T)^2) = \operatorname{Tr}(QB^2Q^T) = \operatorname{Tr}(B^2)$ by the cyclic property, and for the cross-term:

$$\operatorname{Tr}(AQBQ^T) = \operatorname{Tr}(Q^TAQ \cdot B) = \sum_{i,j}(Q^TAQ)_{ij}B_{ji} = \sum_{i,j,k} q_{ki}\lambda_k q_{kj}\gamma_j\delta_{ij} = \sum_{i,k} q_{ki}^2\lambda_k\gamma_i.$$

Under Haar measure, $\mathbb{E}_Q[q_{ki}^2] = \frac{1}{n}$ for all k, i. Therefore:

$$\mathbb{E}_Q[\operatorname{Tr}(AQBQ^T)] = \sum_{i,k}\frac{\lambda_k\gamma_i}{n} = \frac{\operatorname{Tr}(A)\operatorname{Tr}(B)}{n}.$$

Thus:

$$\mathbb{E}_Q[\mathrm{Tr}(M(Q)^2)] = \mathrm{Tr}(A^2) + \mathrm{Tr}(B^2) + \frac{2\,\mathrm{Tr}(A)\,\mathrm{Tr}(B)}{n}.$$

The variance of $M(Q)$ is:

$$\sigma^2(M(Q)) = \frac{1}{n}\,\mathrm{Tr}(M(Q)^2) - \mu(M(Q))^2.$$

Taking expectations and using $\mu(M(Q))^2 = (\mu(p) + \mu(q))^2$ (constant):

$$\mathbb{E}_Q[\sigma^2(M(Q))] = \frac{\mathrm{Tr}(A^2) + \mathrm{Tr}(B^2)}{n} + \frac{2\,\mathrm{Tr}(A)\,\mathrm{Tr}(B)}{n^2} - (\mu(p) + \mu(q))^2.$$

Expanding $(\mu(p) + \mu(q))^2 = \mu(p)^2 + 2\mu(p)\mu(q) + \mu(q)^2$ and noting $\mu(p)\mu(q) = \frac{\mathrm{Tr}(A)\,\mathrm{Tr}(B)}{n^2}$:

$$\begin{aligned}\mathbb{E}_Q[\sigma^2(M(Q))] &= \left(\frac{\mathrm{Tr}(A^2)}{n} - \mu(p)^2\right) + \left(\frac{\mathrm{Tr}(B^2)}{n} - \mu(q)^2\right)\\ &= \sigma^2(p) + \sigma^2(q).\end{aligned}$$

Step 3: Conversion to efficiency ratios.
From Steps 1 and 2, combined with the Fisher-Variance inequality (Lemma 5.2), we have:

$$\Phi_n(p \boxplus_n q) \leq w\Phi_n(p) + (1 - w)\Phi_n(q).$$

Multiplying both sides by $\frac{4(\sigma^2(p)+\sigma^2(q))}{n(n-1)^2}$ and using Lemma 5.4:

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

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

$$w(\sigma^2(p) + \sigma^2(q)) = \sigma^2(p), \quad (1 - w)(\sigma^2(p) + \sigma^2(q)) = \sigma^2(q).$$

Therefore:

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

□

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. □

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.