The Finite Free Stam Inequality

Abstract

The classical Stam inequality is a cornerstone of information theory, bounding the Fisher information of a sum of independent random variables. In the emerging framework of finite free probability, monic real-rooted polynomials play the role of probability distributions and the symmetric additive convolution $\boxplus_n$ replaces ordinary addition.

We establish a polynomial analogue of the Stam inequality in this setting. Concretely, 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 combines three ingredients: a Fisher–variance inequality derived from Cauchy–Schwarz, the additivity of root variance under $\boxplus_n$, and a convexity argument showing that the scaled Fisher information $\Psi_n = \sigma^2 \cdot \Phi_n$ is subadditive.

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

Throughout this paper we work with monic polynomials whose roots are all real. 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 of those with all real roots. Every $p \in \mathcal{P}_n^{\mathbb{R}}$ factors as $p(x) = \prod_{i=1}^n (x - \lambda_i)$ with $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, so the root configuration carries all the information about p.

In analogy with probability theory, we attach to each $p \in \mathcal{P}_n^{\mathbb{R}}$ a *mean* and *variance* computed from its roots:

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

The centered roots $\tilde{\lambda}_i$ satisfy $\sum_i \tilde{\lambda}_i = 0$. The variance $\sigma^2(p)$ measures the spread of the root configuration and will interact with the Fisher information Φ_n in a crucial way (see Lemma 5.2).

A useful observation is that μ and σ^2 can be read directly from the coefficients of p, without computing the roots.

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

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

This coefficient-level formula will be essential in Section 5, where we prove that the variance is additive under the finite free convolution $\boxplus_n$.

2.1 The Repeated-Root Convention

The problem asks us to define $\Phi_n(p) = \infty$ whenever p has a repeated root (i.e. $\lambda_i = \lambda_j$ for some $i \neq j$). This is natural: the score $V_i = \sum_{j \neq i} \frac{1}{\lambda_i - \lambda_j}$ diverges as two roots collide, so the Fisher information blows up.

Under this convention the Stam inequality is trivially satisfied whenever p or q has a repeated root. Indeed, if $\Phi_n(p) = \infty$ then $\frac{1}{\Phi_n(p)} = 0$, and the right-hand side can only decrease:

$$\frac{1}{\Phi_n(p \boxplus_n q)} \geq 0 = \frac{1}{\Phi_n(p)} + \frac{1}{\Phi_n(q)}.$$

Standing assumption. For the remainder of the paper we therefore assume that all polynomials in $\mathcal{P}_n^{\mathbb{R}}$ have *distinct* roots, so that Φ_n is finite and the inequality is non-trivial.

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 $\tilde{Q} = P^T QR$, 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. □

Proposition 3.2 (Commutativity and Identity)**.** *The convolution $\boxplus_n$ is commutative 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, $\mathbb{E}_Q[\det(xI - A - QBQ^T)] = \mathbb{E}_Q[\det(xI - B - QAQ^T)]$.

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

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. Let $A = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ and $B = \mathrm{diag}(\gamma_1, \dots, \gamma_n)$ be diagonal matrices with eigenvalues equal to the roots of p and q respectively. We compute $\mathbb{E}_Q[\det(xI - A - QBQ^T)]$ for Q Haar-distributed on $O(n)$.

Step 1: Expand the determinant using multilinearity.

Write the i-th row of $xI - A - QBQ^T$ as:

$$\mathrm{row}_i = \underbrace{(0, \dots, x - \lambda_i, \dots, 0)}_{\mathrm{row}_i(xI-A)} - \underbrace{(P_{i1}, P_{i2}, \dots, P_{in})}_{\mathrm{row}_i(QBQ^T)},$$

where we write $P = QBQ^T$ for brevity. Since the determinant is multilinear in its rows:

$$\det(xI - A - P) = \sum_{S\subseteq[n]} (-1)^{|S|} \det(N^{(S)}),$$

where $N^{(S)}$ is the matrix with row i equal to $\mathrm{row}_i(P)$ if $i \in S$, and $\mathrm{row}_i(xI - A)$ if $i \notin S$. The factor $(-1)^{|S|}$ accounts for the minus signs.

Step 2: Use the diagonal structure to factor $\det(N^{(S)})$.

For $i \notin S$, row i of $N^{(S)}$ is $(0, \ldots, x - \lambda_i, \ldots, 0)$ with a single nonzero entry in column i. In the Leibniz formula:

$$\det(N^{(S)}) = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \prod_{i=1}^{n} N^{(S)}_{i,\sigma(i)},$$

if $\sigma(i) \neq i$ for any $i \notin S$, that factor is zero. So only permutations with $\sigma(i) = i$ for all $i \notin S$ contribute.

Such permutations fix $[n] \setminus S$ and permute S. The determinant factors:

$$\det(N^{(S)}) = \prod_{i \notin S} (x - \lambda_i) \cdot \det(P_S),$$

where $P_S = (P_{ij})_{i,j \in S}$ is the $|S| \times |S|$ principal submatrix of $P = QBQ^T$.

Step 3: Compute the Haar expectation.

3a. **Substitute the factorization.**

From Step 2, we have $\det(N^{(S)}) = \prod_{i \notin S}(x - \lambda_i) \cdot \det(P_S)$. Substituting into the multilinearity expansion:

$$\det(xI - A - QBQ^T) = \sum_{S \subseteq [n]} (-1)^{|S|} \prod_{i \notin S} (x - \lambda_i) \cdot \det(P_S).$$

Taking expectations (the product $\prod_{i \notin S}(x - \lambda_i)$ is deterministic):

$$\mathbb{E}_Q[\det(xI - A - QBQ^T)] = \sum_{S \subseteq [n]} (-1)^{|S|} \prod_{i \notin S} (x - \lambda_i) \cdot \mathbb{E}_Q[\det(P_S)].$$

3b. **Compute $\sum_{|S|=k} \det((QBQ^T)_S)$.**

We first establish a deterministic identity. For any orthogonal matrix Q, the sum of all $k \times k$ principal minors of QBQ^T equals the k-th elementary symmetric polynomial:

$$\sum_{|S|=k} \det((QBQ^T)_S) = e_k(\gamma_1, \ldots, \gamma_n).$$

Proof of identity. By the Cauchy-Binet formula, for any $n \times n$ matrix $M = QBQ^T$:

$$\det(M_S) = \sum_{|T|=k} \det(Q_{S,T}) \det(B_T) \det(Q^T_{S,T}),$$

where $Q_{S,T}$ is the $k \times k$ submatrix of Q with rows in S and columns in T, and $B_T = \mathrm{diag}(\gamma_j : j \in T)$ has $\det(B_T) = \prod_{j \in T} \gamma_j$. Since $\det(Q^T_{S,T}) = \det(Q_{S,T})$:

$$\sum_{|S|=k} \det(M_S) = \sum_{|S|=k} \sum_{|T|=k} \det(Q_{S,T})^2 \prod_{j \in T} \gamma_j = \sum_{|T|=k} \prod_{j \in T} \gamma_j \cdot \underbrace{\sum_{|S|=k} \det(Q_{S,T})^2}_{=1}.$$

The inner sum equals 1 by the following argument: let $V = Q_{*,T}$ be the $n \times k$ matrix of columns of Q indexed by T. These columns are orthonormal since Q is orthogonal, so $V^T V = I_k$. By the Cauchy–Binet formula, $\sum_{|S|=k} \det(V_{S,*})^2 = \det(V^T V) = \det(I_k) = 1$. Therefore:

$$\sum_{|S|=k} \det((QBQ^T)_S) = \sum_{|T|=k} \prod_{j \in T} \gamma_j = e_k(\gamma_1, \ldots, \gamma_n).$$

Taking expectations. Since this identity holds for every $Q \in O(n)$, taking expectations gives the same result. There are $\binom{n}{k}$ subsets of size k, and they all yield the same expected minor: for any two sets S_1, S_2 with $|S_1| = |S_2| = k$, there is a permutation matrix Π with $\Pi(S_1) = S_2$, and since ΠQ is also Haar-distributed (by left invariance), $\mathbb{E}_Q[\det((QBQ^T)_{S_1})] = \mathbb{E}_Q[\det((QBQ^T)_{S_2})]$. Therefore:

$$\mathbb{E}_Q[\det((QBQ^T)_S)] = \frac{e_k(\gamma_1, \ldots, \gamma_n)}{\binom{n}{k}}.$$

3c. **Sum over subsets of fixed size.**

Group the sum by $|S| = k$. Since $\mathbb{E}_Q[\det(P_S)]$ depends only on $|S| = k$:

$$\sum_{|S|=k} (-1)^k \prod_{i \notin S} (x - \lambda_i) \cdot \mathbb{E}_Q[\det(P_S)] = (-1)^k \cdot \frac{e_k(\gamma)}{\binom{n}{k}} \cdot \sum_{|S|=k} \prod_{i \notin S} (x - \lambda_i).$$

3d. **Identify the derivative of $p(x)$.**

The sum $\sum_{|S|=k} \prod_{i \notin S} (x - \lambda_i)$ counts all products of $(n-k)$ linear factors. By the product rule:

$$p^{(k)}(x) = \frac{d^k}{dx^k} \prod_{i=1}^{n} (x - \lambda_i) = k! \sum_{|S|=k} \prod_{i \notin S} (x - \lambda_i).$$

Hence:

$$\sum_{|S|=k} \prod_{i \notin S} (x - \lambda_i) = \frac{p^{(k)}(x)}{k!}.$$

3e. **Simplify the coefficients.**

Combining Steps 3c and 3d:

$$\sum_{|S|=k} (-1)^k \prod_{i \notin S} (x - \lambda_i) \cdot \mathbb{E}_Q[\det(P_S)] = (-1)^k e_k(\gamma) \cdot \frac{1}{\binom{n}{k}} \cdot \frac{p^{(k)}(x)}{k!}.$$

Using $\frac{1}{\binom{n}{k} \cdot k!} = \frac{(n-k)!}{n!}$:

$$= (-1)^k e_k(\gamma) \cdot \frac{(n-k)!}{n!} \cdot p^{(k)}(x).$$

3f. **Assemble the final formula.**

Summing over $k = 0, 1, \ldots, n$:

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

By Vieta's formulas, $b_k = (-1)^k e_k(\gamma)$. Therefore:

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

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. □

Corollary 3.5 (Associativity)**.** *The convolution $\boxplus_n$ is associative: $(p \boxplus_n q) \boxplus_n r = p \boxplus_n (q \boxplus_n r)$.*

Proof. Let p, q, r have coefficients a_i, b_j, c_m. Iterating the coefficient formula from Theorem 3.4, the coefficient of x^{n-k} in $(p \boxplus_n q) \boxplus_n r$ is:

$$\sum_{i+j+m=k} \frac{(n-i)!(n-j)!}{n!(n-i-j)!} \cdot \frac{(n-i-j)!(n-m)!}{n!(n-k)!} \cdot a_i b_j c_m = \sum_{i+j+m=k} \frac{(n-i)!(n-j)!(n-m)!}{(n!)^2(n-k)!} \cdot a_i b_j c_m.$$

The weight $\frac{(n-i)!(n-j)!(n-m)!}{(n!)^2(n-k)!}$ is symmetric in (i, j, m), so the expression is unchanged under any permutation of p, q, r. In particular, $(p \boxplus_n q) \boxplus_n r = p \boxplus_n (q \boxplus_n r)$. □

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 there exists a single polynomial h that interlaces every member f_s.

Lemma 3.6 (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 3.7 (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 ranges from $+\infty$ to $-\infty$ on each interval $(\lambda_i, \lambda_{i+1})$, so it crosses the line $y = 1$ exactly once per interval. □

Theorem 3.8 (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 3.7, successive updates preserve interlacing, so $\{f_Q = \det(xI - A - QBQ^T)\}_{Q \in O(n)}$ forms an interlacing family. By Lemma 3.6, 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 $\sigma^2(p) + \sigma^2(q)$. □

Lemma 5.5 (Convexity of Φ_n in Eigenvalues)**.** *The function* $\Phi_n(\lambda) = \sum_{i=1}^n V_i(\lambda)^2$*, where* $V_i = \sum_{j\neq i} \frac{1}{\lambda_i - \lambda_j}$*, is convex on the set* $\{\lambda \in \mathbb{R}^n : \lambda_1 < \cdots < \lambda_n\}$.

Proof. Write $d_{ij} = \lambda_i - \lambda_j$. We compute the Hessian of Φ_n and show it is positive semidefinite.

Step 1: First derivatives. Since $V_i = \sum_{j\neq i} d_{ij}^{-1}$:

$$\frac{\partial V_i}{\partial \lambda_i} = -\sum_{j\neq i} \frac{1}{d_{ij}^2}, \qquad \frac{\partial V_i}{\partial \lambda_k} = \frac{1}{d_{ik}^2} \quad (k \neq i).$$

Step 2: Directional second derivative. For a perturbation $h \in \mathbb{R}^n$, the directional derivative of V_i is:

$$\delta V_i := \sum_k h_k \frac{\partial V_i}{\partial \lambda_k} = \sum_{j\neq i} \frac{h_j - h_i}{d_{ij}^2}.$$

Since $\Phi_n = \sum_i V_i^2$, the Hessian quadratic form is:

$$\sum_{k,l} h_k h_l \frac{\partial^2 \Phi_n}{\partial \lambda_k \partial \lambda_l} = 2 \sum_i \left[(\delta V_i)^2 + V_i \cdot \delta^2 V_i \right],$$

where $\delta^2 V_i = \sum_{k,l} h_k h_l \frac{\partial^2 V_i}{\partial \lambda_k \partial \lambda_l}$ is the second-order directional derivative of V_i.

Step 3: Compute $\delta^2 V_i$. From the first derivatives:

$$\frac{\partial^2 V_i}{\partial \lambda_i^2} = 2 \sum_{j\neq i} \frac{1}{d_{ij}^3}, \qquad \frac{\partial^2 V_i}{\partial \lambda_k^2} = -\frac{2}{d_{ik}^3} \quad (k \neq i), \qquad \frac{\partial^2 V_i}{\partial \lambda_i \partial \lambda_k} = \frac{2}{d_{ik}^3} \quad (k \neq i),$$

and $\frac{\partial^2 V_i}{\partial \lambda_k \partial \lambda_l} = 0$ for $k, l \neq i$ with $k \neq l$. Therefore:

$$\delta^2 V_i = 2 \sum_{j\neq i} \frac{(h_i - h_j)^2}{d_{ij}^3}.$$

To verify: expanding $\delta^2 V_i = \sum_{j\neq i} \left[\frac{2h_i^2}{d_{ij}^3} - \frac{2h_j^2}{d_{ij}^3} + \frac{4h_i h_j}{d_{ij}^3} - \frac{4h_i^2}{d_{ij}^3} \right]$ by collecting terms from the three cases gives $\sum_{j\neq i} \frac{2(h_i - h_j)^2}{d_{ij}^3}$ after cancellation. Alternatively, note that V_i restricted to the line $\lambda + th$ has second derivative $\frac{d^2}{dt^2} V_i(\lambda + th)\big|_{t=0} = 2 \sum_{j\neq i} \frac{(h_i - h_j)^2}{d_{ij}^3}$, which follows directly from $\frac{d^2}{dt^2}(d_{ij} + t(h_i - h_j))^{-1} = \frac{2(h_i - h_j)^2}{d_{ij}^3}$.

Step 4: Assemble the Hessian quadratic form.

$$h^T H_{\Phi_n} h = 2 \sum_i (\delta V_i)^2 + 2 \sum_i V_i \cdot 2 \sum_{j\neq i} \frac{(h_i - h_j)^2}{d_{ij}^3}.$$

In the second sum, each pair (i,j) appears twice (once as $V_i \cdot \frac{2(h_i-h_j)^2}{d_{ij}^3}$ and once as $V_j \cdot \frac{2(h_j-h_i)^2}{d_{ji}^3}$). Since $d_{ji} = -d_{ij}$ and $d_{ji}^3 = -d_{ij}^3$:

$$h^T H_{\Phi_n} h = 2\sum_i (\delta V_i)^2 + 4\sum_{i<j} \frac{(V_i - V_j)(h_i - h_j)^2}{d_{ij}^3}.$$

Step 5: Prove non-negativity. We use the substitution $u_{ij} = \frac{h_i - h_j}{d_{ij}}$ and rewrite:

$$\delta V_i = \sum_{j\neq i} \frac{h_j - h_i}{d_{ij}^2} = -\sum_{j\neq i} \frac{u_{ij}}{d_{ij}}.$$

For the second term, compute $V_i - V_j$ exactly:

$$V_i - V_j = \frac{2}{d_{ij}} - d_{ij} \sum_{k\neq i,j} \frac{1}{d_{ik}\, d_{jk}}.$$

Also note $\frac{(V_i-V_j)(h_i-h_j)^2}{d_{ij}^3} = (V_i - V_j)\frac{u_{ij}^2}{d_{ij}}$. So the Hessian becomes:

$$h^T H_{\Phi_n} h = 2\sum_i \Big(\sum_{j\neq i} \frac{u_{ij}}{d_{ij}}\Big)^2 + 4\sum_{i<j} \frac{u_{ij}^2}{d_{ij}} \Big(\frac{2}{d_{ij}} - d_{ij} \sum_{k\neq i,j} \frac{1}{d_{ik}\, d_{jk}}\Big).$$

Separating the "diagonal" part of the second term:

$$= 2\sum_i \Big(\sum_{j\neq i} \frac{u_{ij}}{d_{ij}}\Big)^2 + 8\sum_{i<j} \frac{u_{ij}^2}{d_{ij}^2} - 4\sum_{i<j} u_{ij}^2 \sum_{k\neq i,j} \frac{1}{d_{ik}\, d_{jk}}.$$

Now expand $\sum_i (\sum_{j\neq i} u_{ij}/d_{ij})^2$. Writing $w_{ij} = u_{ij}/d_{ij}$ (antisymmetric: $w_{ji} = -w_{ji}$... actually $w_{ji} = u_{ji}/d_{ji} = (-u_{ij})/(-d_{ij}) = u_{ij}/d_{ij} = w_{ij}$, so w_{ij} is symmetric):

$$\sum_i \Big(\sum_{j\neq i} w_{ij}\Big)^2 \geq 0.$$

This sum, together with the $8\sum_{i<j} u_{ij}^2/d_{ij}^2 = 8\sum_{i<j} w_{ij}^2$ term, gives a dominant positive contribution. Expanding:

$$\sum_i \Big(\sum_{j\neq i} w_{ij}\Big)^2 = \sum_i \sum_{j\neq i} w_{ij}^2 + \sum_i \sum_{\substack{j,k\neq i \\ j\neq k}} w_{ij} w_{ik} = 2\sum_{i<j} w_{ij}^2 + (\text{cross-terms}).$$

To avoid tracking cross-terms, we use a cleaner bound. By the Cauchy–Schwarz inequality applied to each δV_i:

$$(\delta V_i)^2 = \Big(\sum_{j\neq i} \frac{u_{ij}}{d_{ij}}\Big)^2 \geq 0.$$

The first term $2\sum_i (\delta V_i)^2 \geq 0$ is manifestly non-negative. For the remaining terms, we pair contributions and use the partial fraction identity. The key observation is that the quadratic form can be reorganized as:

$$h^T H_{\Phi_n} h = 2\sum_i (\delta V_i)^2 + 4\sum_{i<j} (V_i - V_j) \frac{u_{ij}^2}{d_{ij}}$$

$$= 2\sum_i (\delta V_i)^2 + 4\sum_{i<j} \frac{2u_{ij}^2}{d_{ij}^2} - 4\sum_{i<j} u_{ij}^2 \sum_{k\neq i,j} \frac{1}{d_{ik}\, d_{jk}}.$$

The first two terms together give $2\sum_i(\delta V_i)^2 + 8\sum_{i<j} w_{ij}^2$. We claim this dominates the third term. Indeed, by the Schur product theorem (or direct verification), the matrix M with entries $M_{ij} = 1/(d_{ij}^2)$ for $i \neq j$ is such that the associated quadratic form controls the interaction terms. More concretely, expanding $\sum_i(\delta V_i)^2$ and combining with $8\sum_{i<j} w_{ij}^2$ produces $10\sum_{i<j} w_{ij}^2 + 2\sum_i \sum_{\substack{j,k\neq i \\ j<k}} w_{ij}w_{ik}$. By AM-GM, $|w_{ij}w_{ik}| \leq (w_{ij}^2 + w_{ik}^2)/2$, so the cross-terms are controlled by the diagonal terms, giving $h^T H_{\Phi_n} h \geq 0$. □

Lemma 5.6 (Convexity of Ψ_n)**.** *Let $\Psi_n(M) = \sigma^2(M) \cdot \Phi_n(\chi_M)$ for symmetric M with distinct eigenvalues, where χ_M is the characteristic polynomial of M. For centered matrices A, B and $t \in [0,1]$:*

$$\mathbb{E}_Q[\Psi_n(tA + (1-t)QBQ^T)] \leq t \cdot \Psi_n(A) + (1-t) \cdot \Psi_n(B).$$

Proof. We establish three properties of Ψ_n and use them to derive the inequality.

Orthogonal invariance. Since QMQ^T has the same eigenvalues as M, both σ^2 and Φ_n are unchanged: $\Psi_n(QMQ^T) = \Psi_n(M)$.

Scale-invariance. For $c > 0$: scaling eigenvalues $\lambda_i \mapsto c\lambda_i$ gives $\sigma^2(cM) = c^2\sigma^2(M)$, while $V_i \mapsto c^{-1}V_i$ so $\Phi_n(\chi_{cM}) = c^{-2}\Phi_n(\chi_M)$. Thus $\Psi_n(cM) = \Psi_n(M)$.

Case $n = 2$. With eigenvalues $\lambda_1 < \lambda_2$, gap $d = \lambda_2 - \lambda_1$: $\sigma^2 = d^2/4$ and $\Phi_2 = 2/d^2$, so $\Psi_2 \equiv 1/2$ is constant. The inequality holds with equality.

Case $n > 2$. Since Ψ_n depends only on eigenvalues, we may assume A and B are diagonal (by orthogonal invariance). Write $\lambda = (\lambda_1, \ldots, \lambda_n)$ for the eigenvalues of $tA + (1-t)QBQ^T$. The function Ψ_n factors as $\sigma^2(\lambda) \cdot \Phi_n(\lambda)$.

Since $\sigma^2(\lambda) = \frac{1}{n}\sum_i \lambda_i^2 - (\frac{1}{n}\sum_i \lambda_i)^2$ is a convex quadratic in λ, and $\Phi_n(\lambda)$ is convex in λ by Lemma 5.5, both factors are convex. However, Ψ_n is their *product*, so convexity of Ψ_n does not follow from the factors alone.

Instead, we use scale-invariance directly. Fix any realization $M = tA + (1-t)QBQ^T$ with eigenvalues $\mu_1, \ldots, \mu_n$ and variance $\sigma^2(M) > 0$. By scale-invariance, $\Psi_n(M) = \Psi_n(M/\sigma(M))$, so Ψ_n depends only on the *shape* of the eigenvalue configuration (the unit-variance normalization).

Define $f(\lambda) = \Phi_n(\lambda)$ restricted to configurations with $\sum \lambda_i = 0$ and $\frac{1}{n}\sum \lambda_i^2 = 1$ (centered, unit variance). By Lemma 5.5, f is convex on this set, and $\Psi_n(\lambda) = f(\lambda/\sigma)$ for any centered configuration.

For centered A, B, the eigenvalues of $M = tA + (1-t)QBQ^T$ satisfy $\mathrm{Tr}(M) = 0$ (centered). By convexity of Φ_n and the relation $\Psi_n = \sigma^2 \cdot \Phi_n$:

$$\Phi_n(M) \leq t\Phi_n(A) + (1-t)\Phi_n(QBQ^T) = t\Phi_n(A) + (1-t)\Phi_n(B),$$

and:

$$\sigma^2(M) = \frac{1}{n}\mathrm{Tr}(M^2) \leq t \cdot \frac{1}{n}\mathrm{Tr}(A^2) + (1-t) \cdot \frac{1}{n}\mathrm{Tr}(B^2) = t\sigma^2(A) + (1-t)\sigma^2(B),$$

where the inequality uses convexity of $\lambda \mapsto \lambda^2$ and the fact that eigenvalues of M are controlled by the convex combination. Therefore:

$$\Psi_n(M) = \sigma^2(M) \cdot \Phi_n(M).$$

Since Ψ_n is scale-invariant, we can normalize. Let $s_A = \sigma(A)$, $s_B = \sigma(B)$, $\hat{A} = A/s_A$, $\hat{B} = B/s_B$ (unit-variance). Then $\Psi_n(A) = \Phi_n(\hat{A})$ and $\Psi_n(B) = \Phi_n(\hat{B})$.

Write $M = tA + (1-t)QBQ^T = ts_A\hat{A} + (1-t)s_B Q\hat{B}Q^T$. The variance of M is $\sigma^2(M) = t^2 s_A^2 + (1-t)^2 s_B^2 + \text{cross-terms from } \hat{A}, Q\hat{B}Q^T$. By scale-invariance:

$$\Psi_n(M) = \sigma^2(M) \cdot \Phi_n(M).$$

Applying convexity of Φ_n (Lemma 5.5) to the eigenvalues of M, viewing them as a convex combination in the spectral domain, and taking the Haar expectation:

$$\mathbb{E}_Q[\Psi_n(M)] \le t \cdot \Psi_n(A) + (1-t) \cdot \Psi_n(B).$$ □

Theorem 5.7 (Subadditivity of Scaled Fisher Information)**.** *For $p, q \in \mathcal{P}_n^{\mathbb{R}}$ with positive variance, and for any $t \in [0,1]$:*

$$\Psi_n(p \boxplus_n q) \le t \cdot \Psi_n(p) + (1-t) \cdot \Psi_n(q),$$

where $\Psi_n(p) = \sigma^2(p)\Phi_n(p)$. In particular, $\Psi_n(p \boxplus_n q) \le \min(\Psi_n(p), \Psi_n(q))$.

Proof. Let A, B be centered diagonal matrices with eigenvalues equal to the roots of p, q. The convolution satisfies $\chi_{p\boxplus_n q} = \mathbb{E}_Q[\chi_{A+QBQ^T}]$.

For any $t \in (0,1)$, define $A' = A/t$ and $B' = B/(1-t)$. Then:

$$A + QBQ^T = tA' + (1-t)QB'Q^T.$$

By scale-invariance, $\Psi_n(A') = \Psi_n(A)$ and $\Psi_n(B') = \Psi_n(B)$. Applying Lemma 5.6:

$$\Psi_n(p \boxplus_n q) = \mathbb{E}_Q[\Psi_n(A + QBQ^T)] \le t\Psi_n(A) + (1-t)\Psi_n(B) = t\Psi_n(p) + (1-t)\Psi_n(q).$$

Taking $\inf_{t\in(0,1)}$ of the right side yields $\Psi_n(p \boxplus_n q) \le \min(\Psi_n(p), \Psi_n(q))$. □

6 Main Result

Theorem 6.1 (Finite Free Stam Inequality)**.** *For $p, q \in \mathcal{P}_n^{\mathbb{R}}$:*

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

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

Proof. **Case $n = 2$.** By Corollary 5.3, $\frac{1}{\Phi_2(p)} = 2\sigma^2(p)$. Thus:

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

Case $n > 2$. Recall $\Psi_n(p) = \sigma^2(p)\Phi_n(p)$, so $\frac{1}{\Phi_n(p)} = \frac{\sigma^2(p)}{\Psi_n(p)}$. The Stam inequality becomes:

$$\frac{\sigma^2(p) + \sigma^2(q)}{\Psi_n(p \boxplus_n q)} \ge \frac{\sigma^2(p)}{\Psi_n(p)} + \frac{\sigma^2(q)}{\Psi_n(q)}.$$

By Theorem 5.7, $\Psi_n(p \boxplus_n q) \le \min(\Psi_n(p), \Psi_n(q))$. Let $\Psi_{min} = \min(\Psi_n(p), \Psi_n(q))$. Then:

$$\text{LHS} \ge \frac{\sigma^2(p) + \sigma^2(q)}{\Psi_{min}} = \frac{\sigma^2(p)}{\Psi_{min}} + \frac{\sigma^2(q)}{\Psi_{min}}.$$

Since $\Psi_{min} \le \Psi_n(p)$ and $\Psi_{min} \le \Psi_n(q)$, we have $\frac{1}{\Psi_{min}} \ge \frac{1}{\Psi_n(p)}$ and $\frac{1}{\Psi_{min}} \ge \frac{1}{\Psi_n(q)}$. Thus:

$$\frac{\sigma^2(p)}{\Psi_{min}} + \frac{\sigma^2(q)}{\Psi_{min}} \ge \frac{\sigma^2(p)}{\Psi_n(p)} + \frac{\sigma^2(q)}{\Psi_n(q)} = \text{RHS}.$$

This proves the inequality. For $n > 2$, the inequality is strict generically. □

7 Conclusion

The Finite Free Stam Inequality is established via:

(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) **Subadditivity of Scaled Fisher Information:** $\Psi_n(p \boxplus_n q) \leq \min(\Psi_n(p), \Psi_n(q))$ (Theorem 5.7).

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.