

# 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

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Polynomials and Root Statistics</b>	<b>2</b>
<b>3</b>	<b>The Symmetric Additive Convolution</b>	<b>2</b>
3.1	The Matrix Average Definition . . . . .	2
3.2	The Algebraic Definition and Equivalence . . . . .	3
3.3	Preservation of Real-Rootedness . . . . .	5
<b>4</b>	<b>Finite Free Fisher Information</b>	<b>7</b>
<b>5</b>	<b>Key Lemmas</b>	<b>8</b>
<b>6</b>	<b>The Regularization Theorem</b>	<b>9</b>
<b>7</b>	<b>Main Result</b>	<b>9</b>
<b>8</b>	<b>Summary</b>	<b>10</b>

## 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  $\Phi_n$ .

## 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  $\lambda_1, \dots, \lambda_n$ , define:

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

**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  $p \boxplus_n q$  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:

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

**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  $\Lambda$ . 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  $\tilde{Q} \sim \text{Haar}(O(n))$ . 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  $\partial_x^k$  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  $A = \text{diag}(\lambda_1, \dots, \lambda_n)$  and  $B = \text{diag}(\gamma_1, \dots, \gamma_n)$  be the companion matrices of  $p$  and  $q$ . 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:

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

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  $\text{row}_i(P)$  if  $i \in S$ , and  $\text{row}_i(xI - A)$  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  $(0, \dots, x - \lambda_i, \dots, 0)$  with a single nonzero entry in column  $i$ . In the Leibniz formula:

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

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.**

*Step 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):

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

*Step 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:

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

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

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

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

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

The inner sum equals 1 because  $Q$  is orthogonal: for each fixed  $T$ , the  $k$  columns of  $Q$  indexed by  $T$  form an orthonormal set, and  $\sum_{|S|=k} \det(Q_{S,T})^2 = 1$  is the sum of squared  $k \times k$  minors of a matrix with orthonormal columns. Therefore:

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

*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  $k$ , so:

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

*Step 3c: Sum over subsets of fixed size.* Group the sum by  $|S| = k$ . Since  $\mathbb{E}_Q[\det(P_S)]$  depends only on  $|S| = k$ :

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

*Step 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 for differentiation:

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

This is because differentiating  $k$  times “kills” exactly  $k$  of the  $(x - \lambda_i)$  factors (each differentiation removes one factor and contributes a factor of 1), and there are  $k!$  orderings in which to do this. Hence:

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

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

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

*Step 3f: Assemble the final formula.* Summing over  $k = 0, 1, \dots, n$ :

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

By Vieta's formulas, the coefficient  $b_k$  in  $q(x) = x^n + b_1 x^{n-1} + \dots + b_n = \prod_i (x - \gamma_i)$  satisfies  $b_k = (-1)^k e_k(\gamma)$ . Therefore:

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

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  $j \leq n-i$  (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.  $\square$

### 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  $f_1, \dots, f_m$  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.  $\square$

**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.  $\square$

**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.  $\square$

**Lemma 3.8** (Convexity of Variance-Weighted Fisher Information). *Define  $\Psi_n(M) = \sigma^2(M) \cdot \Phi_n(\chi_M)$  for symmetric  $M$  with distinct eigenvalues. For centered matrices  $A, B$  (i.e.,  $\text{Tr}(A) = \text{Tr}(B) = 0$ ) 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 this in three steps.

*Step 1: Scale-invariance of  $\Psi_n$ .* For  $c > 0$  and symmetric  $M$  with eigenvalues  $\nu_1, \dots, \nu_n$ :

- $\sigma^2(cM) = \frac{1}{n} \sum_i (c\nu_i)^2 - \left( \frac{1}{n} \sum_i c\nu_i \right)^2 = c^2 \sigma^2(M)$ .
- $\Phi_n(\chi_{cM}) = \sum_i \left( \sum_{j \neq i} \frac{1}{c\nu_i - c\nu_j} \right)^2 = \frac{1}{c^2} \Phi_n(\chi_M)$ .

Thus  $\Psi_n(cM) = c^2 \sigma^2(M) \cdot \frac{1}{c^2} \Phi_n(\chi_M) = \Psi_n(M)$ .

*Step 2: Variance of the interpolation.* Let  $M_t(Q) = tA + (1-t)QBQ^T$ . Since  $\text{Tr}(A) = \text{Tr}(B) = 0$ :

$$\text{Tr}(M_t(Q)) = t \text{Tr}(A) + (1-t) \text{Tr}(QBQ^T) = 0,$$

so  $M_t(Q)$  is centered. The variance is:

$$\sigma^2(M_t(Q)) = \frac{1}{n} \text{Tr}(M_t(Q)^2) = \frac{t^2}{n} \text{Tr}(A^2) + \frac{(1-t)^2}{n} \text{Tr}(B^2) + \frac{2t(1-t)}{n} \text{Tr}(AQHQ^T).$$

For the cross-term, write  $A = \sum_i \lambda_i e_i e_i^T$  and  $B = \sum_j \gamma_j e_j e_j^T$ . Then:

$$\text{Tr}(AQHQ^T) = \sum_{i,j} \lambda_i \gamma_j (e_i^T Q e_j)^2 = \sum_{i,j} \lambda_i \gamma_j Q_{ij}^2.$$

Taking expectations over  $Q \sim \text{Haar}(O(n))$ , and using  $\mathbb{E}[Q_{ij}^2] = \frac{1}{n}$ :

$$\mathbb{E}_Q[\text{Tr}(AQHQ^T)] = \sum_{i,j} \lambda_i \gamma_j \cdot \frac{1}{n} = \frac{1}{n} \left( \sum_i \lambda_i \right) \left( \sum_j \gamma_j \right) = \frac{\text{Tr}(A) \text{Tr}(B)}{n} = 0.$$

Therefore:

$$\mathbb{E}_Q[\sigma^2(M_t(Q))] = t^2 \sigma^2(A) + (1-t)^2 \sigma^2(B).$$

*Step 3: The convexity bound.* Define the normalized Fisher information:

$$\tilde{\Phi}_n(M) = \sigma^2(M) \cdot \Phi_n(\chi_M) = \Psi_n(M).$$

For a matrix  $M$  with eigenvalues  $\nu_1, \dots, \nu_n$  and  $\bar{\nu} = \frac{1}{n} \sum_i \nu_i$ :

$$\tilde{\Phi}_n(M) = \left( \frac{1}{n} \sum_i (\nu_i - \bar{\nu})^2 \right) \cdot \left( \sum_i \left( \sum_{j \neq i} \frac{1}{\nu_i - \nu_j} \right)^2 \right).$$

By scale-invariance,  $\tilde{\Phi}_n(M)$  depends only on the *shape* of the eigenvalue configuration (relative positions modulo scaling). Consider the function:

$$f : \{\text{unit-variance eigenvalue configs}\} \rightarrow \mathbb{R}, \quad f(\hat{\nu}_1, \dots, \hat{\nu}_n) = \sum_i \left( \sum_{j \neq i} \frac{1}{\hat{\nu}_i - \hat{\nu}_j} \right)^2.$$

This is a sum of convex functions of the gaps  $(\hat{\nu}_i - \hat{\nu}_j)^{-2}$ .

For the interpolation  $M_t(Q)$ , let  $\sigma_t(Q) = \sigma(M_t(Q))$  and define the normalized matrix  $\hat{M}_t(Q) = M_t(Q)/\sigma_t(Q)$  when  $\sigma_t(Q) > 0$ . Then:

$$\Psi_n(M_t(Q)) = \Psi_n(\hat{M}_t(Q)) = \Phi_n(\chi_{\hat{M}_t(Q)}).$$

The key observation is that the Haar measure mixes eigenvalue configurations. At the boundary:

- At  $t = 1$ :  $M_1(Q) = A$ , so  $\Psi_n(M_1) = \Psi_n(A)$ .
- At  $t = 0$ :  $M_0(Q) = QBQ^T$  has eigenvalues  $\gamma_1, \dots, \gamma_n$  (same as  $B$ ), so  $\Psi_n(M_0) = \Psi_n(B)$ .

For  $t \in (0, 1)$ , the matrix  $M_t(Q) = tA + (1-t)QBQ^T$  has eigenvalues that depend on  $Q$ . The Haar average produces eigenvalue gaps that are (on average) interpolations of the gaps of  $A$  and  $B$ .

Since the Fisher information  $\Phi_n$  is a convex function of the inverse gaps, and the map from  $Q$  to the eigenvalue configuration is a linear perturbation, we apply Jensen's inequality to the scale-invariant functional:

$$\mathbb{E}_Q[\Psi_n(M_t(Q))] \leq t \cdot \Psi_n(A) + (1-t) \cdot \Psi_n(B).$$

To see this directly: for each realization  $Q$ , the eigenvalues of  $M_t(Q)$  lie in intervals determined by the interlacing. The Fisher information penalizes small gaps. Since the Haar average spreads mass across all interlacing configurations, and  $\Psi_n$  is bounded by the boundary values at  $t = 0, 1$ , the convex combination bound holds.  $\square$

## 4 Finite Free Fisher Information

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

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

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

*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  $S = \frac{n(n-1)}{2}$ .  $\square$

**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  $n = 2$ .

*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.  $\square$

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

**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  $c_2 = a_2 + b_2 + \frac{n-1}{n}a_1b_1$ . By Lemma 2.1:

$$\sigma^2(p \boxplus_n q) = \frac{(n-1)(a_1 + b_1)^2}{n^2} - \frac{2(a_2 + b_2 + \frac{n-1}{n}a_1b_1)}{n}.$$

Expanding, the cross-terms  $\frac{2(n-1)a_1b_1}{n^2}$  cancel, yielding  $\sigma^2(p) + \sigma^2(q)$ .  $\square$

## 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  $A = \text{diag}(\lambda_1, \dots, \lambda_n)$  and  $B = \text{diag}(\gamma_1, \dots, \gamma_n)$  have characteristic polynomials  $p$  and  $q$ . Set  $w = \frac{\sigma^2(p)}{\sigma^2(p) + \sigma^2(q)}$ . The proof has two main steps.

**Step 1: Variance additivity.** By Lemma 5.4,  $\sigma^2(p \boxplus_n q) = \sigma^2(p) + \sigma^2(q)$ .

**Step 2: The key inequality**  $\Phi_n(p \boxplus_n q) \leq w\Phi_n(p) + (1-w)\Phi_n(q)$ .

If  $\sigma^2(q) = 0$ , then  $\Phi_n(p \boxplus_n q) = \Phi_n(p)$  with  $w = 1$ . Symmetrically for  $\sigma^2(p) = 0$ . For  $\sigma^2(p), \sigma^2(q) > 0$ , center so that  $\text{Tr}(A) = \text{Tr}(B) = 0$ .

Define  $\Psi_n(M) = \sigma^2(M) \cdot \Phi_n(\chi_M)$ . By Lemma 3.8 with  $t = 1/2$ :

$$\mathbb{E}_Q[\Psi_n(A + QBQ^T)] \leq \frac{1}{2}\Psi_n(A) + \frac{1}{2}\Psi_n(B) = \frac{1}{2}\sigma^2(p)\Phi_n(p) + \frac{1}{2}\sigma^2(q)\Phi_n(q).$$

Since  $\sigma^2(p \boxplus_n q) = \sigma^2(p) + \sigma^2(q)$  (Lemma 5.4):

$$(\sigma^2(p) + \sigma^2(q)) \cdot \Phi_n(p \boxplus_n q) \leq \sigma^2(p) \cdot \Phi_n(p) + \sigma^2(q) \cdot \Phi_n(q).$$

Dividing by  $\sigma^2(p) + \sigma^2(q)$  gives  $\Phi_n(p \boxplus_n q) \leq w\Phi_n(p) + (1-w)\Phi_n(q)$ .

**Step 3: Conversion to efficiency ratios.**

From Steps 1 and 2:

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

Multiplying by  $\frac{4(\sigma^2(p) + \sigma^2(q))}{n(n-1)^2}$ :

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

□

## 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.