

# The Finite Free Stam Inequality

## Abstract

The classical Stam inequality states that Fisher information is superadditive in its reciprocal:  $1/I(X+Y) \geq 1/I(X) + 1/I(Y)$  for independent random variables. We prove the analogue for real-rooted polynomials, where addition is replaced by the finite free additive convolution  $\boxplus_n$  and Fisher information by a quantity  $\Phi_n$  measuring the electrostatic repulsion of roots. The proof combines an algebraic inequality—the Score-Gradient Inequality, established via two applications of Cauchy–Schwarz—with a flow-based Grönwall argument.

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Polynomials and Convolution</b>	<b>2</b>
2.1	The Finite Free Additive Convolution . . . . .	2
<b>3</b>	<b>Scores and Fisher Information</b>	<b>3</b>
<b>4</b>	<b>The Score-Gradient Inequality</b>	<b>4</b>
<b>5</b>	<b>The Convolution Flow</b>	<b>5</b>
5.1	Perturbation Analysis . . . . .	5
5.2	The Continuous Flow . . . . .	6
<b>6</b>	<b>The Stam Inequality</b>	<b>6</b>
<b>7</b>	<b>Concluding Remarks</b>	<b>7</b>

## 1 Introduction

In probability theory, the addition of independent random variables increases disorder. The **Stam inequality** makes this precise: if  $X$  and  $Y$  are independent with Fisher information  $I(X)$  and  $I(Y)$ , then

$$\frac{1}{I(X+Y)} \geq \frac{1}{I(X)} + \frac{1}{I(Y)}.$$

This paper proves the polynomial analogue, where each probabilistic concept is replaced by an algebraic one:

Probability	$\longleftrightarrow$	Polynomials
Random variable $X$	$\longleftrightarrow$	Polynomial $p(x)$
Distribution of $X$	$\longleftrightarrow$	Roots $\lambda_1, \dots, \lambda_n$
Addition $X + Y$	$\longleftrightarrow$	Finite free convolution $p \boxplus_n q$
Fisher information $I(X)$	$\longleftrightarrow$	$\Phi_n(p)$

**Theorem** (Finite Free Stam Inequality). *For monic, degree- $n$  polynomials  $p, q$  with all real roots and positive variance,*

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

The proof proceeds in four stages. We first set up the algebraic framework: root statistics, the convolution  $\boxplus_n$ , and the Fisher information  $\Phi_n$  (§2–3). We then establish the key algebraic input, the Score-Gradient Inequality (§4). Next, we embed  $\boxplus_n$  into a continuous flow and derive an integral identity (§5). Finally, we close the argument via a Grönwall-type integration (§6).

**Convention.** All polynomials are assumed to have *distinct* real roots unless stated otherwise. The final inequality extends to all of  $\mathcal{P}_n^{\mathbb{R}}$  by continuity, since polynomials with distinct roots are dense in  $\mathcal{P}_n^{\mathbb{R}}$ .

## 2 Polynomials and Convolution

Let  $\mathcal{P}_n$  denote the space of monic polynomials of degree  $n$  with real coefficients, and let  $\mathcal{P}_n^{\mathbb{R}} \subset \mathcal{P}_n$  be the subset with all real roots. For  $p \in \mathcal{P}_n^{\mathbb{R}}$  write

$$p(x) = \prod_{i=1}^n (x - \lambda_i), \quad \lambda_1 \leq \dots \leq \lambda_n.$$

**Definition 2.1** (Root Statistics). The mean and variance of the root distribution are

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

**Lemma 2.1** (Variance Formula). *If  $p(x) = x^n + a_1 x^{n-1} + a_2 x^{n-2} + \dots$ , then  $\sigma^2(p) = \frac{(n-1)a_1^2}{n^2} - \frac{2a_2}{n}$ .*

*Proof.* By Vieta's formulas,  $\sum \lambda_i = -a_1$  and  $\sum_{i < j} \lambda_i \lambda_j = a_2$ , so  $\sum \lambda_i^2 = a_1^2 - 2a_2$ . Substituting into  $\sigma^2 = \frac{1}{n} \sum \lambda_i^2 - \mu^2$  gives the result.  $\square$

### 2.1 The Finite Free Additive Convolution

Let  $A$  and  $B$  be real symmetric matrices with characteristic polynomials  $p$  and  $q$ . Averaging the characteristic polynomial of  $A + QBQ^T$  over all orthogonal rotations  $Q$  yields a basis-independent “sum.”

**Definition 2.2** (Symmetric Additive Convolution).

$$p \boxplus_n q := \int_{O(n)} \det(xI - (A + QBQ^T)) d\mu_{\text{Haar}}(Q),$$

where  $\mu_{\text{Haar}}$  is the unique bi-invariant probability measure on the orthogonal group  $O(n)$ .

A theorem of Marcus, Spielman, and Srivastava converts this matrix integral into a differential operator.

**Theorem 2.2** (MSS [1]). *If  $q(x) = \sum_{k=0}^n b_k x^{n-k}$ , then*

$$(p \boxplus_n q)(x) = T_q p(x), \quad T_q := \sum_{k=0}^n \frac{(n-k)!}{n!} b_k \partial_x^k.$$

Convolution with  $q$  thus acts as a diffusion on  $p$ , with weights determined by the root distribution of  $q$ .

**Theorem 2.3** (Preservation of Real Roots). *If  $p, q \in \mathcal{P}_n^{\mathbb{R}}$ , then  $p \boxplus_n q \in \mathcal{P}_n^{\mathbb{R}}$ .*

**Lemma 2.4** (Variance Additivity).  $\sigma^2(p \boxplus_n q) = \sigma^2(p) + \sigma^2(q)$ .

*Proof.* The convolution formula for the coefficients of  $r = p \boxplus_n q$  reads  $c_k = \sum_{i+j=k} \frac{(n-i)!(n-j)!}{n!(n-k)!} a_i b_j$ . Computing  $c_1$  and  $c_2$ :

$$c_1 = a_1 + b_1, \quad c_2 = a_2 + \frac{n-1}{n} a_1 b_1 + b_2.$$

Substituting into the variance formula (Lemma 2.1) and expanding  $(a_1 + b_1)^2$ :

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

The cross-terms cancel exactly. □

### 3 Scores and Fisher Information

We treat the roots  $\lambda_1, \dots, \lambda_n$  as charged particles on a line that repel with force inversely proportional to distance.

**Definition 3.1** (Score and Fisher Information). The *score* at root  $\lambda_i$  is the total repulsive force it experiences:

$$V_i = \sum_{j \neq i} \frac{1}{\lambda_i - \lambda_j}.$$

The *finite free Fisher information* is the total squared force:

$$\Phi_n(p) = \sum_{i=1}^n V_i^2.$$

High  $\Phi_n$  corresponds to tightly clustered roots; low  $\Phi_n$  to well-separated roots.

**Lemma 3.1** (Score via Derivatives).  $V_i = \frac{p''(\lambda_i)}{2p'(\lambda_i)}$ .

*Proof.* Since  $p(x) = \prod_j (x - \lambda_j)$ , we have  $p'(\lambda_i) = \prod_{j \neq i} (\lambda_i - \lambda_j)$ . Differentiating  $p'(x) = \sum_i \prod_{j \neq i} (x - \lambda_j)$  once more and evaluating at  $\lambda_i$ :

$$p''(\lambda_i) = 2 \sum_{k \neq i} \prod_{\substack{j \neq i \\ j \neq k}} (\lambda_i - \lambda_j) = 2p'(\lambda_i) \sum_{k \neq i} \frac{1}{\lambda_i - \lambda_k} = 2p'(\lambda_i) V_i. \quad \square$$

**Lemma 3.2** (Score-Root Identity).  $\sum_{i=1}^n (\lambda_i - \mu) V_i = \frac{n(n-1)}{2}$ .

*Proof.* First,  $\sum_i V_i = \sum_{i \neq j} \frac{1}{\lambda_i - \lambda_j} = 0$  by antisymmetry. Next,

$$\sum_{i=1}^n \lambda_i V_i = \sum_{i \neq j} \frac{\lambda_i}{\lambda_i - \lambda_j} = \sum_{i < j} \left( \frac{\lambda_i}{\lambda_i - \lambda_j} + \frac{\lambda_j}{\lambda_j - \lambda_i} \right) = \sum_{i < j} 1 = \binom{n}{2}.$$

Since  $\sum V_i = 0$ , we conclude  $\sum (\lambda_i - \mu) V_i = \sum \lambda_i V_i = \frac{n(n-1)}{2}$ .  $\square$

**Lemma 3.3** (Fisher–Variance Inequality).  $\Phi_n(p) \cdot \sigma^2(p) \geq \frac{n(n-1)^2}{4}$ .

*Proof.* Apply Cauchy–Schwarz to Lemma 3.2:

$$\frac{n^2(n-1)^2}{4} = \left( \sum_i (\lambda_i - \mu) V_i \right)^2 \leq \left( \sum_i (\lambda_i - \mu)^2 \right) \left( \sum_i V_i^2 \right) = n \sigma^2(p) \cdot \Phi_n(p). \quad \square$$

## 4 The Score-Gradient Inequality

We now establish the key algebraic estimate that powers the full Stam inequality.

**Lemma 4.1** (Score Decomposition).  $\Phi_n(p) = \sum_{i < j} \frac{V_i - V_j}{\lambda_i - \lambda_j}$ .

*Proof.*

$$\sum_i V_i^2 = \sum_i V_i \sum_{j \neq i} \frac{1}{\lambda_i - \lambda_j} = \sum_{i \neq j} \frac{V_i}{\lambda_i - \lambda_j} = \sum_{i < j} \frac{V_i - V_j}{\lambda_i - \lambda_j}. \quad \square$$

**Definition 4.1** (Score-Gradient Energy).  $\mathcal{S}(p) = \sum_{i < j} \frac{(V_i - V_j)^2}{(\lambda_i - \lambda_j)^2}$ .

**Theorem 4.2** (Score-Gradient Inequality). For  $p \in \mathcal{P}_n^{\mathbb{R}}$  of degree  $n \geq 2$  with distinct roots,

$$\mathcal{S}(p) \cdot \sigma^2(p) \geq \frac{n-1}{2} \Phi_n(p),$$

with equality if and only if  $V_i = c(\lambda_i - \mu)$  for some constant  $c$ .

*Proof.* Write  $T = n \sigma^2(p)$ ,  $U = \Phi_n(p)$ ,  $S = \mathcal{S}(p)$ . We must show  $ST \geq \frac{n(n-1)}{2} U$ .

**Step 1 (Cauchy–Schwarz on the Score-Root Identity).** By Lemma 3.2:

$$\frac{n^2(n-1)^2}{4} = \left( \sum_i (\lambda_i - \mu) V_i \right)^2 \leq T U. \quad (1)$$

**Step 2 (Cauchy–Schwarz on the Score Decomposition).** By Lemma 4.1:

$$U^2 = \left( \sum_{i < j} \frac{V_i - V_j}{\lambda_i - \lambda_j} \right)^2 \leq S \cdot \binom{n}{2} = \frac{n(n-1)}{2} S. \quad (2)$$

**Step 3 (Combination).** From (2):  $S \geq \frac{2U^2}{n(n-1)}$ . Multiply by  $T$  and apply (1):

$$ST \geq \frac{2U^2 T}{n(n-1)} = \frac{2U}{n(n-1)} \cdot TU \geq \frac{2U}{n(n-1)} \cdot \frac{n^2(n-1)^2}{4} = \frac{n(n-1)}{2} U.$$

**Equality.** Both Cauchy–Schwarz applications must be tight. Inequality (1) requires  $V_i = c(\lambda_i - \mu)$ . Inequality (2) requires  $\frac{V_i - V_j}{\lambda_i - \lambda_j}$  to be constant over all pairs. These conditions are equivalent:  $V_i = c(\lambda_i - \mu)$  implies  $\frac{V_i - V_j}{\lambda_i - \lambda_j} = c$ . Conversely,  $\frac{V_i - V_j}{\lambda_i - \lambda_j} = k$  for all  $i < j$  forces  $V_i - k\lambda_i$  to be constant; since  $\sum V_i = 0$ , this gives  $V_i = k(\lambda_i - \mu)$ .  $\square$

*Remark 4.1.* The equality condition  $V_i = c(\lambda_i - \mu)$  characterizes (up to affine transformation) the zeros of the Hermite polynomials: if  $x_1, \dots, x_n$  are the zeros of the physicist's Hermite polynomial  $H_n$ , the ODE  $H_n'' - 2xH_n' + 2nH_n = 0$  evaluated at a zero  $x_k$  gives  $V_k = x_k$ .

## 5 The Convolution Flow

We study how  $\Phi_n$  evolves under convolution, first infinitesimally, then along a continuous flow.

### 5.1 Perturbation Analysis

Let  $q_\epsilon$  be a centered polynomial with  $\sigma^2(q_\epsilon) = \epsilon^2$ . We analyze  $p \boxplus_n q_\epsilon$  for small  $\epsilon$ .

**Lemma 5.1** (Shift of Roots). *The roots  $\mu_i$  of  $p \boxplus_n q_\epsilon$  satisfy  $\mu_i = \lambda_i + \frac{\epsilon^2}{n-1} V_i + O(\epsilon^3)$ .*

*Proof.* Since  $q_\epsilon$  is centered with variance  $\epsilon^2$ , its coefficients satisfy  $b_0 = 1$ ,  $b_1 = 0$ ,  $b_2 = -n\epsilon^2/2$  (by Lemma 2.1). The operator  $T_{q_\epsilon}$  acts as

$$T_{q_\epsilon} p(x) = p(x) - \frac{\epsilon^2}{2(n-1)} p''(x) + O(\epsilon^3).$$

Setting  $\mu_i = \lambda_i + \delta_i$  with  $\delta_i = O(\epsilon^2)$  and expanding  $T_{q_\epsilon} p(\mu_i) = 0$ :

$$0 = \underbrace{p(\lambda_i)}_{=0} + \delta_i p'(\lambda_i) - \frac{\epsilon^2}{2(n-1)} p''(\lambda_i) + O(\epsilon^3).$$

Solving and applying Lemma 3.1:  $\delta_i = \frac{\epsilon^2}{2(n-1)} \cdot \frac{p''(\lambda_i)}{p'(\lambda_i)} + O(\epsilon^3) = \frac{\epsilon^2}{n-1} V_i + O(\epsilon^3)$ .  $\square$

Roots move in the direction of the repulsive force  $V_i$ : isolated roots barely shift, while clustered roots are pushed apart.

**Lemma 5.2** (Change in Fisher Information).  $\Phi_n(p \boxplus_n q_\epsilon) = \Phi_n(p) - \frac{2\epsilon^2}{n-1} \mathcal{S}(p) + O(\epsilon^3)$ .

*Proof.* From Lemma 5.1, the perturbed scores are

$$V_i^{(\epsilon)} = \sum_{j \neq i} \frac{1}{\mu_i - \mu_j} = \sum_{j \neq i} \frac{1}{(\lambda_i - \lambda_j) \left(1 + \frac{\epsilon^2}{n-1} \frac{V_i - V_j}{\lambda_i - \lambda_j} + O(\epsilon^3)\right)} = V_i - \frac{\epsilon^2}{n-1} \sum_{j \neq i} \frac{V_i - V_j}{(\lambda_i - \lambda_j)^2} + O(\epsilon^3).$$

Squaring and summing:

$$\Phi_n(p_\epsilon) = \sum_i V_i^2 - \frac{2\epsilon^2}{n-1} \sum_{i \neq j} \frac{V_i(V_i - V_j)}{(\lambda_i - \lambda_j)^2} + O(\epsilon^3).$$

Pairing  $(i, j)$  with  $(j, i)$ :  $\sum_{i \neq j} \frac{V_i(V_i - V_j)}{(\lambda_i - \lambda_j)^2} = \sum_{i < j} \frac{(V_i - V_j)^2}{(\lambda_i - \lambda_j)^2} = \mathcal{S}(p)$ .  $\square$

Since  $\mathcal{S}(p) \geq 0$ , the Fisher information is non-increasing under convolution.

## 5.2 The Continuous Flow

To prove the full Stam inequality we embed the convolution into a one-parameter flow.

**Definition 5.1** (Fractional Convolution Flow). Let  $b = \sigma^2(q) > 0$ . Using the normalized coefficients  $\kappa_k(p) = \frac{(n-k)!}{n!} a_k$ , the convolution formula becomes a Cauchy product:  $K_{p \boxplus_n q}(z) = K_p(z) K_q(z)$ , where  $K_p(z) = \sum_k \kappa_k(p) z^k$ . Define

$$K_{q_t}(z) := K_q(z)^t, \quad t \in [0, 1],$$

truncated at degree  $n$ . Then  $q_0 = x^n$ ,  $q_1 = q$ ,  $\sigma^2(q_t) = tb$ , and the semigroup property  $q_s \boxplus_n q_t = q_{s+t}$  holds at the operator level. The *flow polynomial* is  $p_t := p \boxplus_n q_t$ , so that

$$\sigma^2(p_t) = \sigma^2(p) + tb. \quad (3)$$

**Lemma 5.3** (Dissipation).  $\frac{d}{dt} \Phi_n(p_t) = -\frac{2b}{n-1} \mathcal{S}(p_t)$ .

*Proof.* By the semigroup property,  $p_{t+h} = p_t \boxplus_n q_h$  with  $\sigma^2(q_h) = hb$ . Lemma 5.2 gives  $\Phi_n(p_{t+h}) = \Phi_n(p_t) - \frac{2hb}{n-1} \mathcal{S}(p_t) + O(h^2)$ . Dividing by  $h$  and letting  $h \rightarrow 0$  yields the result.  $\square$

**Theorem 5.4** (Integral Identity).

$$\frac{1}{\Phi_n(p \boxplus_n q)} - \frac{1}{\Phi_n(p)} = \frac{2b}{n-1} \int_0^1 \frac{\mathcal{S}(p_t)}{\Phi_n(p_t)^2} dt.$$

*Proof.* Set  $f(t) = 1/\Phi_n(p_t)$ . By the chain rule and Lemma 5.3:

$$f'(t) = -\frac{\dot{\Phi}_n(p_t)}{\Phi_n(p_t)^2} = \frac{2b}{n-1} \cdot \frac{\mathcal{S}(p_t)}{\Phi_n(p_t)^2} \geq 0.$$

Integrating from 0 to 1 and noting  $f(0) = 1/\Phi_n(p)$ ,  $f(1) = 1/\Phi_n(p \boxplus_n q)$  gives the result.  $\square$

## 6 The Stam Inequality

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

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

*Proof.* Write  $a = \sigma^2(p)$ ,  $b = \sigma^2(q)$ , and  $f(t) = 1/\Phi_n(p_t)$ .

**Step 1: Differential inequality.** From Theorem 5.4,  $f'(t) = \frac{2b}{n-1} \cdot \frac{\mathcal{S}(p_t)}{\Phi_n(p_t)^2}$ . The Score-Gradient Inequality (Theorem 4.2) applied to  $p_t$  gives  $\mathcal{S}(p_t) \geq \frac{(n-1)\Phi_n(p_t)}{2\sigma^2(p_t)}$ . Substituting:

$$f'(t) \geq \frac{b}{\sigma^2(p_t)} f(t) = \frac{b}{a+tb} f(t).$$

**Step 2: Grönwall integration.** Since  $f(t) > 0$ , divide by  $f(t)$ :

$$\frac{d}{dt} \ln f(t) \geq \frac{b}{a+tb} = \frac{d}{dt} \ln(a+tb).$$

Integrate from 0 to 1:  $\ln \frac{f(1)}{f(0)} \geq \ln \frac{a+b}{a}$ , hence

$$\frac{1}{\Phi_n(p \boxplus_n q)} \geq \frac{a+b}{a} \cdot \frac{1}{\Phi_n(p)}. \quad (\text{Forward})$$

**Step 3: Reverse bound by symmetry.** Since  $p \boxplus_n q = q \boxplus_n p$ , the identical argument with  $p$  and  $q$  swapped yields

$$\frac{1}{\Phi_n(p \boxplus_n q)} \geq \frac{a+b}{b} \cdot \frac{1}{\Phi_n(q)}. \quad (\text{Reverse})$$

**Step 4: Weighted combination.** Set  $w_1 = \frac{a}{a+b}$  and  $w_2 = \frac{b}{a+b}$ , so that  $w_1 + w_2 = 1$ . Multiply (Forward) by  $w_1$  and (Reverse) by  $w_2$ :

$$\begin{aligned} w_1 \cdot \frac{1}{\Phi_n(p \boxplus_n q)} &\geq \frac{a}{a+b} \cdot \frac{a+b}{a} \cdot \frac{1}{\Phi_n(p)} = \frac{1}{\Phi_n(p)}, \\ w_2 \cdot \frac{1}{\Phi_n(p \boxplus_n q)} &\geq \frac{b}{a+b} \cdot \frac{a+b}{b} \cdot \frac{1}{\Phi_n(q)} = \frac{1}{\Phi_n(q)}. \end{aligned}$$

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

*Remark 6.1.* The weight  $w_1 = a/(a+b)$  is the unique choice making the factor  $w_1 \cdot (a+b)/a$  equal to 1. This is precisely where variance additivity ( $\sigma^2(p \boxplus_n q) = a+b$ ) is essential: without it, the cancellation in Step 4 would fail.

*Remark 6.2.* The inequality is strict for generic  $p$  and  $q$ . Full equality requires both Grönwall bounds to be simultaneously tight along the entire flow, i.e.  $V_i(p_t) = c(t)(\lambda_i(t) - \mu(t))$  for all  $t \in [0, 1]$ , which forces both  $p$  and  $q$  to have roots at the (affinely rescaled) zeros of the Hermite polynomial  $H_n$ .

## 7 Concluding Remarks

The proof rests on four ingredients, each with a clear role:

Ingredient	Role in the proof
$\mathcal{S}(p) \geq 0$	Monotonicity of $1/\Phi_n$ along the flow
Score-Gradient Inequality	Quantitative ODE: $f' \geq (b/(a+tb))f$
Grönwall integration	Forward and reverse bounds
Commutativity + variance additivity	Weighted combination

Two weaker results also follow immediately. The *weak Stam inequality*,  $1/\Phi_n(p \boxplus_n q) \geq 1/\Phi_n(p)$ , uses only  $\mathcal{S} \geq 0$  (the integral identity has a non-negative integrand) and requires none of the Grönwall machinery. Averaging this bound with its symmetric counterpart  $1/\Phi_n(q \boxplus_n p) \geq 1/\Phi_n(q)$  gives the *half-Stam inequality*:  $2/\Phi_n(p \boxplus_n q) \geq 1/\Phi_n(p) + 1/\Phi_n(q)$ .

## 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.
- [2] A. J. Stam, *Some inequalities satisfied by the quantities of information of Fisher and Shannon*, Inform. Control **2** (1959), 101–112.