The Finite Free Stam Inequality
An Introduction to Polynomial Convolution and Root Dynamics

Abstract

The classical Stam inequality formulates the superadditivity of entropy power, or equivalently, the decrease of Fisher information upon the addition of independent random variables. This paper proves the analogue of this result for real-rooted polynomials, where random variables are replaced by polynomials, addition by the "finite free additive convolution," and Fisher information by a quantity derived from the electrostatic repulsion of roots. We establish a key algebraic inequality—the Score-Gradient Inequality—via two applications of Cauchy–Schwarz, and combine it with a flow-based framework to obtain the full Finite Free Stam Inequality: $1/\Phi_n(p \boxplus_n q) \geq 1/\Phi_n(p) + 1/\Phi_n(q)$.

Contents

1   Introduction

1.1   Analogy with Probability Theory

In probability theory and information theory, the addition of independent random sources strictly increases uncertainty. The convolution of two distributions results in a distribution that is smoother than the originals.

Mathematically, this is captured by the **Stam Inequality**. Let X and Y be independent random variables with Fisher information $I(X)$ and $I(Y)$. Then:

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

Since Fisher information I measures the structural sharpness of a distribution, its reciprocal $1/I$ acts as a measure of smoothness or disorder. This inequality asserts that the total disorder of the sum is at least the sum of the individual disorders.

1.2   The Polynomial Framework

We replace random variables with **polynomials** according to the following correspondence:

Probability	$\longleftrightarrow$	**Polynomials**
Random Variable X	$\longleftrightarrow$	Polynomial $p(x)$
Distribution of X	$\longleftrightarrow$	Roots $\lambda_1, \ldots, \lambda_n$ of $p(x)$
Addition $X + Y$	$\longleftrightarrow$	Symmetric Additive Convolution $p \boxplus_n q$
Fisher Information $I(X)$	$\longleftrightarrow$	Finite Free Fisher Information $\Phi_n(p)$

Finite Free Probability provides a bridge between random matrix theory and the geometry of polynomials. In this framework, we investigate the validity of the Stam inequality for the finite free additive convolution.

This paper presents a self-contained proof of these phenomena. We treat the roots of a polynomial as charged particles on a line and define their "energy" (Fisher information) based on their mutual repulsion. We then show that convolution acts as a diffusion process, lowering the potential energy of the system.

2   Polynomials and Root Statistics

We consider the space of monic polynomials of degree n with real coefficients, denoted $\mathcal{P}_n$. We focus on the subset $\mathcal{P}_n^{\mathbb{R}}$ of polynomials with **all real roots**.

$$p(x) = \prod_{i=1}^{n}(x - \lambda_i), \quad \lambda_i \in \mathbb{R}.$$

We assume the roots are ordered $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$.

Definition 2.1 (Root Statistics). The first and second moments of the root distribution are defined as:

$$\mu(p) = \frac{1}{n}\sum_{i=1}^{n}\lambda_i,$$

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

The variance can be computed directly from the coefficients without solving for the roots.

Lemma 2.1 (The Variance Formula). *Let $p(x) = x^n + a_1 x^{n-1} + a_2 x^{n-2} + \cdots$. Then:*

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

Proof. Recall Vieta's formulas: $\sum \lambda_i = -a_1$ and $\sum_{i<j}\lambda_i\lambda_j = a_2$. The sum of squared roots is $\sum \lambda_i^2 = (\sum \lambda_i)^2 - 2\sum_{i<j}\lambda_i\lambda_j = a_1^2 - 2a_2$. Substituting into the variance definition $\sigma^2(p) = \frac{1}{n}\sum \lambda_i^2 - \mu^2$ yields the result. $\qquad\square$

3 The Symmetric Additive Convolution

The addition of polynomials in this context is not pointwise addition but a convolution operation rooted in random matrix theory.

3.1 Matrix Models and Haar Integration

Let $p(x)$ be the characteristic polynomial of a symmetric matrix A, and $q(x)$ be that of B. The sum $A + B$ has a characteristic polynomial that depends on the relative eigenbasis of A and B. To obtain a basis-independent operation, we average over all possible relative orientations using the Haar measure.

Definition 3.1 (Haar Measure on $O(n)$). The group of orthogonal matrices $O(n)$ admits a unique probability measure μ_{Haar} that is invariant under left and right multiplication: $\mu(S) = \mu(gS) = \mu(Sg)$ for any $g \in O(n)$ and measurable set S. This is the uniform distribution on the rotation group.

Definition 3.2 (Symmetric Additive Convolution). The finite free additive convolution of p and q is defined as the expected characteristic polynomial of the sum of randomly rotated matrices:

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

In this expression, QBQ^T represents the matrix B rotated by a random orthogonal matrix Q. The term $\det(xI - (A + QBQ^T))$ is a random polynomial. Its expectation is a deterministic polynomial whose roots effectively represent the "sum" of the root sets of p and q.

3.1.1 Relation between Matrix Eigenvalues and Differential Operators

A fundamental result by Marcus, Spielman, and Srivastava connects this matrix integral to a deterministic differential operator. This connection relies on the relationship between the eigenvalues of B and the coefficients of the operator.

Let the eigenvalues of B be the roots of $q(x)$, denoted $\beta_1, \ldots, \beta_n$. The characteristic polynomial is $q(x) = \prod_{i=1}^{n} (x - \beta_i) = \sum_{k=0}^{n} b_k x^{n-k}$. The coefficients b_k are the elementary symmetric polynomials of the eigenvalues:

$$b_k = (-1)^k \sum_{1 \le i_1 < \cdots < i_k \le n} \beta_{i_1} \ldots \beta_{i_k}.$$

The expected characteristic polynomial expansion can be written in terms of these coefficients.

Theorem 3.1 (MSS Theorem). *The expectation over the orthogonal group factorizes into a differential operator acting on p:*

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

This theorem establishes that the eigenvalues of the matrix B (encapsulated in the coefficients b_k) determine the weights of the differentiation in the operator T_q.

3.2 The Differential Operator Representation

Based on the MSS theorem, we define the operator formally.

Definition 3.3 (Convolution Operator T_q). For a monic polynomial $q(x) = \sum_{k=0}^{n} b_k x^{n-k}$, define:

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

Then the convolution satisfies:

$$(p \boxplus_n q)(x) = T_q p(x).$$

This operator formalism reveals that convolution with q acts as a diffusion process on p, smoothing its features solely based on the root distribution of q.

Theorem 3.2 (Preservation of Real Roots). *If p and q have all real roots, then $p \boxplus_n q$ also has all real roots.*

Lemma 3.3 (Variance Additivity). *The variance is additive under convolution:*

$$\sigma^2(p \boxplus_n q) = \sigma^2(p) + \sigma^2(q).$$

Proof. By the convolution coefficient formula $c_k = \sum_{i+j=k} \frac{(n-i)!\,(n-j)!}{n!\,(n-k)!} \, a_i \, b_j$, we compute the first two coefficients of $p \boxplus_n q$.

For $k = 1$: only the terms $(i, j) = (1, 0)$ and $(0, 1)$ contribute (since $a_0 = b_0 = 1$):

$$c_1 = \frac{(n-1)!\,n!}{n!\,(n-1)!} \, a_1 + \frac{n!\,(n-1)!}{n!\,(n-1)!} \, b_1 = a_1 + b_1.$$

For $k = 2$: the contributing terms are $(i, j) \in \{(2, 0), (1, 1), (0, 2)\}$:

$$c_2 = a_2 + \frac{(n - 1)!\,(n - 1)!}{n!\,(n - 2)!}\,a_1 b_1 + b_2 = a_2 + \tfrac{n-1}{n}\,a_1 b_1 + b_2.$$

Substituting into the variance formula (Lemma 2.1):

$$\sigma^2(p \boxplus_n q) = \frac{(n - 1)c_1^2}{n^2} - \frac{2c_2}{n} = \frac{(n - 1)(a_1 + b_1)^2}{n^2} - \frac{2}{n}\left(a_2 + \tfrac{n-1}{n}a_1 b_1 + b_2\right).$$

Expanding $(a_1 + b_1)^2 = a_1^2 + 2a_1 b_1 + b_1^2$ and collecting terms:

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

The cross-terms cancel exactly, yielding $\sigma^2(p \boxplus_n q) = \sigma^2(p) + \sigma^2(q)$. $\qquad\square$

4   Fisher Information: The Energy of Roots

We define the finite free Fisher information based on the electrostatic interaction of roots.

4.1   Scores and Forces

Roots repel each other with a force proportional to the inverse distance. The total force on a root λ_i is its **score**.

Definition 4.1 (The Score V_i)**.**

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

4.2   Finite Free Fisher Information

The Fisher information is defined as the sum of the squared scores, representing the total stress in the system.

Definition 4.2 (Fisher Information Φ_n)**.**

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

- **High Information:** Clustered roots produce large scores V_i, hence large Φ_n.

- **Low Information:** Well-separated roots produce small scores, hence small Φ_n.

Lemma 4.1 (Score via Derivatives)**.** *For any monic polynomial p of degree n with distinct roots $\lambda_1, \ldots, \lambda_n$,*

$$V_i = \frac{p''(\lambda_i)}{2\,p'(\lambda_i)}.$$

Proof. Since $p(x) = \prod_{j=1}^{n}(x - \lambda_j)$, its derivative at a root λ_i is $p'(\lambda_i) = \prod_{j \neq i}(\lambda_i - \lambda_j)$. To find $p''(\lambda_i)$, differentiate $p'(x) = \sum_{i=1}^{n} \prod_{j \neq i}(x - \lambda_j)$ once more. At $x = \lambda_i$, only the terms where x differentiates one factor of $\prod_{j \neq i}(x - \lambda_j)$ survive:

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

Dividing both sides by $2\, p'(\lambda_i)$ gives $V_i = p''(\lambda_i)/(2\, p'(\lambda_i))$. $\qquad\square$

Remark 4.1. This lemma says that the score at λ_i is simply the logarithmic derivative of p' evaluated at λ_i, divided by 2. It provides a coordinate-free way to compute scores directly from p and its derivatives.

4.3 Key Identities

Lemma 4.2 (Score-Root Identity)**.**

$$\sum_{i=1}^{n}(\lambda_i - \mu)V_i = \frac{n(n-1)}{2}.$$

Proof. First, $\sum_{i=1}^{n} V_i = \sum_{i \neq j} \frac{1}{\lambda_i - \lambda_j} = 0$, since pairing (i, j) with (j, i) gives cancellation. Next,

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

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

Lemma 4.3 (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 Cauchy–Schwarz applied to the Score-Root Identity:

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

Equality in Cauchy–Schwarz requires proportionality $V_i = c(\lambda_i - \mu)$ for a constant c. For $n = 2$ this holds automatically (see Lemma 4.4). $\qquad\square$

Lemma 4.4 (The $n = 2$ Identity)**.** *For quadratic polynomials ($n = 2$), information is inversely proportional to variance:*

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

Proof. Let roots be $-d/2$ and $d/2$. $V_1 = -1/d$, $V_2 = 1/d$. $\Phi_2 = 2/d^2$. $\sigma^2 = d^2/4$. Thus $1/\Phi_2 = d^2/2 = 2\sigma^2$. $\qquad\square$

5 Local Analysis: Perturbation Theory

We examine the behavior of roots under convolution with a noise polynomial q_ϵ of small variance ϵ^2.

5.1   Root Dynamics

The convolution operator induces a shift in the roots of p.

Lemma 5.1 (Shift of Roots). *Let λ_i be the roots of p. The roots μ_i of $p \boxplus_n q_\epsilon$ satisfy:*

$$\mu_i = \lambda_i + \frac{\epsilon^2}{n-1}V_i + O(\epsilon^4).$$

Proof. Consider the operator representation $p \boxplus_n q_\epsilon = T_{q_\epsilon} p$. Since q_ϵ is centered with variance ϵ^2 and close to x^n, its coefficients satisfy $b_0 = 1$, $b_1 = 0$, and $b_2 = -n\epsilon^2/2$ (from the variance formula applied to q_ϵ). Higher coefficients are $O(\epsilon^3)$ or smaller. The operator $T_{q_\epsilon} = \sum_k \frac{(n-k)!}{n!} b_k \partial_x^k$ therefore acts as:

$$T_{q_\epsilon} p(x) = p(x) + \frac{(n-2)!}{n!} b_2 p''(x) + O(\epsilon^3) = p(x) - \frac{\epsilon^2}{2(n-1)} p''(x) + O(\epsilon^3).$$

Now let $\mu_i = \lambda_i + \delta_i$ with $\delta_i = O(\epsilon^2)$. Expanding $T_{q_\epsilon} p(\mu_i) = 0$:

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

Since $p(\lambda_i) = 0$ and $\delta_i^2 = O(\epsilon^4)$, we solve:

$$\delta_i = \frac{\epsilon^2}{2(n-1)} \cdot \frac{p''(\lambda_i)}{p'(\lambda_i)} + O(\epsilon^4).$$

By the Score via Derivatives lemma (Lemma 4.1), $p''(\lambda_i)/(2\,p'(\lambda_i)) = V_i$, so $\delta_i = \frac{\epsilon^2}{n-1} V_i + O(\epsilon^4)$. $\square$

Interpretation: Roots move in the direction of the repulsive force V_i. Isolated roots (small $|V_i|$) barely move, while roots in a cluster (large $|V_i|$) are pushed apart. The convolution acts like a physical relaxation.

5.2   Monotonicity of Fisher Information

This relaxation leads to a decrease in the total energy.

Lemma 5.2 (Change in Fisher Information).

$$\Phi_n(p \boxplus_n q_\epsilon) = \Phi_n(p) - \frac{2\epsilon^2}{n-1}\mathcal{S}(p) + O(\epsilon^4),$$

*where $\mathcal{S}(p)$ is the **Score-Gradient Energy**:*

$$\mathcal{S}(p) = \sum_{1 \leq i < j \leq n} \frac{(V_i - V_j)^2}{(\lambda_i - \lambda_j)^2}.$$

Proof. From Lemma 5.1, the perturbed roots are $\mu_i = \lambda_i + \frac{\epsilon^2}{n-1}V_i + O(\epsilon^4)$. The new scores are:

$$V_i^{(\epsilon)} = \sum_{j \neq i} \frac{1}{\mu_i - \mu_j} = \sum_{j \neq i} \frac{1}{(\lambda_i - \lambda_j) + \frac{\epsilon^2}{n-1}(V_i - V_j) + O(\epsilon^4)}$$

$$= \sum_{j \neq i} \frac{1}{\lambda_i - \lambda_j} \cdot \frac{1}{1 + \frac{\epsilon^2}{n-1}\frac{V_i - V_j}{\lambda_i - \lambda_j} + O(\epsilon^4)}.$$

Using $\frac{1}{1+u} = 1 - u + O(u^2)$ for small u:

$$V_i^{(\epsilon)} = V_i - \frac{\epsilon^2}{n-1} \sum_{j \neq i} \frac{V_i - V_j}{(\lambda_i - \lambda_j)^2} + O(\epsilon^4).$$

Squaring and summing:

$$\Phi_n(p_\epsilon) = \sum_i (V_i^{(\epsilon)})^2 = \sum_i V_i^2 - \frac{2\epsilon^2}{n-1} \sum_i V_i \sum_{j \neq i} \frac{V_i - V_j}{(\lambda_i - \lambda_j)^2} + O(\epsilon^4).$$

The key observation is that the cross-sum simplifies:

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

where in the second step we used $V_i(V_i - V_j) + V_j(V_j - V_i) = (V_i - V_j)^2$. This yields the stated formula. $\square$

Since $\mathcal{S}(p) \geq 0$, the Fisher information strictly decreases under convolution (it decreases strictly unless all scores are equal, i.e. $V_i = 0$ for all i, which only happens for $n = 1$).

5.3   The Score-Gradient Inequality

The score-gradient energy $\mathcal{S}(p)$ satisfies a fundamental lower bound relating it to the Fisher information and the variance. This inequality is the key algebraic input for the full Stam inequality.

Lemma 5.3 (Score Decomposition)**.**

$$\Phi_n(p) = \sum_{i=1}^n V_i^2 = \sum_{1 \leq i < j \leq n} \frac{V_i - V_j}{\lambda_i - \lambda_j}.$$

Proof. Expanding the left-hand side using the definition of V_i:

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

Theorem 5.4 (Score-Gradient Inequality)**.** *For any monic polynomial p of degree $n \geq 2$ with distinct real roots,*

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

that is,

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

Equality holds if and only if there exists a constant c such that $V_i = c(\lambda_i - \mu)$ for all i.

Proof. Write $T = \sum_{i=1}^n (\lambda_i - \mu)^2 = n\sigma^2$, $U = \Phi_n(p) = \sum V_i^2$, and $S = \mathcal{S}(p)$. The inequality to establish is

$$S\,T \geq \frac{n(n-1)}{2} U. \tag{1}$$

Step 1 (Cauchy–Schwarz on the Score-Root Identity). By Lemma 4.2 and the Cauchy–Schwarz inequality,

$$\frac{n^2(n-1)^2}{4} = \left(\sum_{i=1}^n (\lambda_i - \mu)\, V_i\right)^2 \le T\, U. \tag{2}$$

Step 2 (Cauchy–Schwarz on the Score Decomposition). By Lemma 5.3 and the Cauchy–Schwarz inequality,

$$U^2 = \left(\sum_{i<j} \frac{V_i - V_j}{\lambda_i - \lambda_j}\right)^2 \le \left(\sum_{i<j} \frac{(V_i - V_j)^2}{(\lambda_i - \lambda_j)^2}\right)\left(\sum_{i<j} 1\right) = S \cdot \frac{n(n-1)}{2}. \tag{3}$$

Step 3 (Combination). From (3), $S \ge \frac{2U^2}{n(n-1)}$. Multiplying by T:

$$ST \ge \frac{2\, U^2\, T}{n(n-1)} = \frac{2\, U}{n(n-1)} \cdot T\, U \ge \frac{2\, U}{n(n-1)} \cdot \frac{n^2(n-1)^2}{4} = \frac{n(n-1)}{2}\, U,$$

where the second inequality uses (2). This establishes (1).

Equality analysis. Equality requires both Cauchy–Schwarz applications to be tight.

- Inequality (2) is an equality if and only if $V_i = c(\lambda_i - \mu)$ for some constant c.
- Inequality (3) is an equality if and only if $\frac{V_i - V_j}{\lambda_i - \lambda_j}$ is constant over all pairs $i < j$.

If $V_i = c(\lambda_i - \mu)$, then $\frac{V_i - V_j}{\lambda_i - \lambda_j} = c$, so both conditions hold simultaneously. Conversely, if $\frac{V_i - V_j}{\lambda_i - \lambda_j} = k$ for all $i < j$, then $V_i - k\lambda_i$ is constant across i; since $\sum V_i = 0$, this forces $V_i = k(\lambda_i - \mu)$. Therefore, equality holds if and only if $V_i = c(\lambda_i - \mu)$ for all i.

This condition characterizes (up to affine transformation) the zeros of the classical Hermite polynomials: if $x_1, \ldots, x_n$ are the zeros of the physicist's Hermite polynomial H_n, then the ODE $H_n''(x) - 2x\, H_n'(x) + 2n\, H_n(x) = 0$ evaluated at a zero x_k yields $V_k = x_k$, confirming the proportionality $V_k = 1 \cdot (x_k - 0)$ since the zeros are symmetric about the origin. $\square$

6 Global Analysis: The Flow Approach

We construct a continuous flow interpolating between p and $p \boxplus_n q$.

Definition 6.1 (Fractional Convolution Flow). Write $b = \sigma^2(q) > 0$. We construct a one-parameter family $\{q_t\}_{t \in [0,1]}$ of polynomials such that $q_0 = x^n$ (the identity for $\boxplus_n$), $q_1 = q$, and the variance scales linearly: $\sigma^2(q_t) = t\, b$.

Concretely, define q_t by scaling the normalized coefficients: if $\kappa_k = \frac{(n-k)!}{n!}\, b_k$ are the operator coefficients of q, set $b_k(t) = \frac{n!}{(n-k)!}\, \kappa_k^t$. One can verify that this family satisfies the semigroup property $q_s \boxplus_n q_t = q_{s+t}$.

Define the *flow polynomial*: $p_t := p \boxplus_n q_t$. By the variance additivity (Lemma 2.1), $\sigma^2(p_t) = \sigma^2(p) + t\, b$.

6.1   The Dissipation Identity

Lemma 6.1 (Dissipation)**.**

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

Proof. By the semigroup property, $p_{t+h} = p_t \boxplus_n q_h$ where $\sigma^2(q_h) = h\,b$. Applying Lemma 5.2 to the polynomial p_t convolved with the small-variance polynomial q_h:

$$\Phi_n(p_{t+h}) = \Phi_n(p_t) - \frac{2\,h\,b}{n-1}\mathcal{S}(p_t) + O(h^2).$$

Dividing by h and sending $h \to 0$ yields the result. $\square$

6.2   The Integral Formula

We now derive a formula for the change in the *reciprocal* Fisher information.
Theorem 6.2 (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. Apply the chain rule to $f(t) = 1/\Phi_n(p_t)$:

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

where the second equality uses the Dissipation Lemma 6.1. Since $\mathcal{S}(p_t) \geq 0$, the integrand is non-negative, so f is non-decreasing. Integrating from $t = 0$ to $t = 1$, and noting $p_0 = p$ and $p_1 = p \boxplus_n q$, gives the result. $\square$

This integral identity is the engine of the proof. It expresses the "information gain" $1/\Phi_n(p \boxplus_n q) - 1/\Phi_n(p)$ as an integral of the normalized score-gradient energy along the flow.

7   Main Results

7.1   The Weak Stam Inequality

Since the integral is positive:

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

7.2   The Half-Stam Inequality

Summing the integral identities for the forward and reverse flows yields:
Theorem 7.1 (Half-Stam Inequality)**.** *For any real-rooted polynomials p, q with distinct roots:*

$$\frac{2}{\Phi_n(p \boxplus_n q)} \geq \frac{1}{\Phi_n(p)} + \frac{1}{\Phi_n(q)}.$$

Proof. From the Integral Identity (Theorem 6.2), the integrand is non-negative, so:

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

By symmetry of the convolution ($p \boxplus_n q = q \boxplus_n p$), running the same flow from q instead of p gives:

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

Adding these two inequalities yields the Half-Stam inequality. $\square$

Remark 7.1. This result is unconditional and independent of n. It uses only the positivity of $\mathcal{S}$, not its quantitative lower bound. The full Stam inequality requires the sharper Score-Gradient Inequality.

7.3 The Full Stam Inequality

The Score-Gradient Inequality (Theorem 5.4) provides sufficient control on the dissipation rate to upgrade the Half-Stam to the full inequality. The argument proceeds via a Grönwall-type differential inequality along the flow, applied in both directions.

Theorem 7.2 (Full Stam Inequality). *For any real-rooted polynomials $p, q \in \mathcal{P}_n^{\mathbb{R}}$ with positive variances,*

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

Proof. **Step 1: Differential inequality from the flow.** Consider the fractional convolution flow $p_t = p \boxplus_n q_t$ where $\sigma^2(q_t) = t\,\sigma^2(q)$. By the integral formula (Section 6.2):

$$\frac{d}{dt}\left(\frac{1}{\Phi_n(p_t)}\right) = \frac{2\,\sigma^2(q)}{n-1}\,\frac{\mathcal{S}(p_t)}{\Phi_n(p_t)^2}.$$

Applying the Score-Gradient Inequality (Theorem 5.4) to p_t, which has distinct real roots along the flow:

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

Substituting and writing $f(t) = 1/\Phi_n(p_t)$:

$$f'(t) \geq \frac{2\,\sigma^2(q)}{n-1} \cdot \frac{(n-1)}{2\,\sigma^2(p_t)} \cdot \frac{1}{\Phi_n(p_t)} = \frac{\sigma^2(q)}{\sigma^2(p_t)}\,f(t).$$

Since $\sigma^2(p_t) = \sigma^2(p) + t\,\sigma^2(q)$, we recognize this as

$$\frac{f'(t)}{f(t)} \geq \frac{\sigma^2(q)}{\sigma^2(p) + t\,\sigma^2(q)} = \frac{d}{dt}\ln(\sigma^2(p) + t\,\sigma^2(q)).$$

Step 2: Integration. Integrating from $t = 0$ to $t = 1$:

$$\ln\frac{f(1)}{f(0)} \geq \ln\frac{\sigma^2(p) + \sigma^2(q)}{\sigma^2(p)},$$

yielding the *forward bound*:

$$\frac{1}{\Phi_n(p \boxplus_n q)} \geq \frac{\sigma^2(p) + \sigma^2(q)}{\sigma^2(p)} \cdot \frac{1}{\Phi_n(p)}. \tag{4}$$

By symmetry—running the flow from q (setting $q_t = q \boxplus_n r_t$ with $\sigma^2(r_t) = t\,\sigma^2(p)$, so that $q_1 = p \boxplus_n q$)—the identical argument yields the *reverse bound*:

$$\frac{1}{\Phi_n(p \boxplus_n q)} \geq \frac{\sigma^2(p) + \sigma^2(q)}{\sigma^2(q)} \cdot \frac{1}{\Phi_n(q)}. \tag{5}$$

Step 3: Combining via weighted average. Multiply (4) by $w_1 = \frac{\sigma^2(p)}{\sigma^2(p)+\sigma^2(q)}$ and (5) by $w_2 = \frac{\sigma^2(q)}{\sigma^2(p)+\sigma^2(q)}$:

$$\frac{w_1}{\Phi_n(p \boxplus_n q)} \geq \frac{1}{\Phi_n(p)}, \qquad \frac{w_2}{\Phi_n(p \boxplus_n q)} \geq \frac{1}{\Phi_n(q)}.$$

Adding, and using $w_1 + w_2 = 1$:

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

Remark 7.2. The forward and reverse bounds (4)–(5) are individually stronger than the Stam inequality: each asserts that the reciprocal Fisher information grows by at least a factor proportional to the relative variance contribution. In terms of the efficiency ratio $\eta(p) = 4\Phi_n(p)\sigma^2(p)/[n(n-1)^2]$, the forward bound (4) is equivalent to $\eta(p \boxplus_n q) \leq \eta(p)$, confirming that convolution always regularizes.

8 Alternative Characterizations

The Stam inequality can be reformulated in several instructive ways that highlight different aspects of the convolution process.

8.1 The Efficiency Ratio and Regularization

A natural way to measure the "non-Gaussianity" of a polynomial's root distribution is by comparing its Fisher information to the lower bound established in Lemma 4.3.

Definition 8.1 (Efficiency Ratio)**.** For $p \in \mathcal{P}_n^{\mathbb{R}}$ with positive variance, the *efficiency ratio* $\eta(p)$ is defined as:

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

By the Fisher–Variance inequality, $\eta(p) \geq 1$, with equality if and only if $n = 2$ (or in the $n \to \infty$ Gaussian limit for random polynomials).

In this language, the Stam inequality is related to the following regularization property.

Theorem 8.1 (Regularization Property)**.** *For $p, q \in \mathcal{P}_n^{\mathbb{R}}$ with positive variances, the efficiency ratio of the convolution satisfies:*

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

Proof. Write $a = \sigma^2(p)$ and $b = \sigma^2(q)$. Recall the forward and reverse bounds from the proof of the Full Stam Inequality (Theorem 7.2):

$$\Phi_n(p \boxplus_n q) \leq \frac{a\,\Phi_n(p)}{a + b} \quad \text{and} \quad \Phi_n(p \boxplus_n q) \leq \frac{b\,\Phi_n(q)}{a + b}.$$

Multiply the first by $\frac{a}{a+b}$ and the second by $\frac{b}{a+b}$, then add:

$$\Phi_n(p \boxplus_n q) \leq \frac{a^2\,\Phi_n(p) + b^2\,\Phi_n(q)}{(a+b)^2}.$$

Multiplying both sides by $\frac{4(a+b)}{n(n-1)^2}$:

$$\underbrace{\frac{4\,\Phi_n(p \boxplus_n q)(a+b)}{n(n-1)^2}}_{=\,\eta(p\boxplus_n q)} \leq \frac{a \cdot \frac{4a\,\Phi_n(p)}{n(n-1)^2} + b \cdot \frac{4b\,\Phi_n(q)}{n(n-1)^2}}{a+b} = \frac{a\,\eta(p) + b\,\eta(q)}{a+b}.$$

Note that $\sigma^2(p \boxplus_n q) = a + b$ was used in identifying $\eta(p \boxplus_n q)$. $\qquad\square$

This theorem asserts that the efficiency ratio of the convolution is bounded by the variance-weighted average of the efficiency ratios of the components: convolution always pushes the root configuration toward a more regular (lower η) state.

8.2 Convex Mixing Interpretation

The symmetric additive convolution can be viewed as a form of convex mixing at the level of the matrix model. Let A and B be centered symmetric matrices with characteristic polynomials p and q. The functional $\Psi_n(M) = \sigma^2(M)\,\Phi_n(\chi_M)$, where χ_M denotes the characteristic polynomial of M, is scale-invariant.

Remark 8.1 (Convexity under Haar Averaging). The Regularization Property can be interpreted as a matrix-level Jensen's inequality: the Haar integration over $O(n)$ in the definition of $\boxplus_n$ acts as an averaging operation that always reduces the efficiency ratio η. Heuristically, rotating B by a random orthogonal matrix Q and adding it to A produces a sum whose eigenvalues are more "spread out" than either A or B alone. Since η measures deviation from the most spread-out configuration (the Hermite roots), this averaging cannot increase η.

This interpretation complements the flow-based proof: both perspectives confirm that convolution acts as a regularization process, driving root configurations toward the universally optimal Hermite arrangement.

9 Conclusion

We have established the full Finite Free Stam Inequality, linking the geometry of polynomial roots to information-theoretic inequalities.

1. We defined root interactions via electrostatic forces and their associated Fisher information $\Phi_n(p)$.

2. We proved the Score-Gradient Inequality (Theorem 5.4): $\mathcal{S}(p) \cdot \sigma^2(p) \geq \frac{n-1}{2}\Phi_n(p)$, via two applications of Cauchy–Schwarz to the Score-Root Identity and the Score Decomposition.

3. We derived a precise integral formula for the increase in the reciprocal Fisher information along the convolution flow.

4. We combined the Score-Gradient Inequality with the flow to obtain the full Stam inequality (Theorem 7.2):

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

with equality characterized by the Hermite polynomial condition $V_i = c(\lambda_i - \mu)$.

References

[1] A. Marcus, D. Spielman, N. Srivastava, *Interlacing families II: Mixed characteristic polynomials and the Kadison-Singer problem*, Ann. Math. 182 (2015).

[2] A. J. Stam, *Some inequalities satisfied by the quantities of information of Fisher and Shannon*, Information and Control (1959).