The Finite Free Stam Inequality

Abstract

The classical Stam inequality asserts the superadditivity of the reciprocal Fisher information under convolution of independent random variables. We prove the polynomial analogue in the framework of finite free probability: for monic, degree-n, real-rooted polynomials p and q,

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

where $\boxplus_n$ is the symmetric additive convolution of Marcus, Spielman, and Srivastava, and Φ_n is the finite free Fisher information. The proof combines an algebraic inequality—the Score-Gradient Inequality, established via two applications of Cauchy–Schwarz—with a flow-based argument exploiting the semigroup structure of $\boxplus_n$. We also derive a closed-form expression for Φ_n in terms of the critical values of the polynomial via residue calculus, and use it to verify the inequality explicitly for cubics.

Contents

1 Introduction

1.1 Background and motivation

In information theory, the Stam inequality [2] states that if X and Y are independent random variables with finite Fisher information $I(X)$ and $I(Y)$, then

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

This fundamental inequality—equivalent to the entropy power inequality of Shannon and Stam—captures the principle that convolution of independent sources strictly increases disorder.

Finite free probability, introduced by Marcus, Spielman, and Srivastava [1], provides a polynomial analogue of free probability in which random variables are replaced by real-rooted polynomials and addition by a deterministic convolution operation $\boxplus_n$. Within this framework, the natural question arises:

Does the Stam inequality hold for the finite free additive convolution?

The purpose of this paper is to answer this question affirmatively.

1.2 Statement of the main result

Let $\mathcal{P}_n$ denote the space of monic polynomials of degree n with real coefficients, and $\mathcal{P}_n^{\mathbb{R}} \subset \mathcal{P}_n$ the subset with all real roots. For $p \in \mathcal{P}_n^{\mathbb{R}}$ with distinct roots $\lambda_1 < \cdots < \lambda_n$, define the *scores* $V_i = \sum_{j\neq i}(\lambda_i - \lambda_j)^{-1}$ and the *finite free Fisher information* $\Phi_n(p) = \sum_{i=1}^n V_i^2$. The *symmetric additive convolution* $p \boxplus_n q$ is recalled in Section 2.

Theorem 1.1 (Finite Free Stam Inequality)**.** *For $p, q \in \mathcal{P}_n^{\mathbb{R}}$ with positive variance,*

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

The proof combines three ingredients: the Score-Gradient Inequality (Theorem 3.1), a dissipation identity for the convolution flow (Lemma 4.4), and a case-split argument exploiting commutativity of $\boxplus_n$ (Theorem 5.1).

Convention. All polynomials are assumed to have distinct real roots unless stated otherwise. Since such polynomials are dense in $\mathcal{P}_n^{\mathbb{R}}$ and all quantities involved are continuous, inequality (1) extends to all of $\mathcal{P}_n^{\mathbb{R}}$ by a limiting argument.

2 Preliminaries

2.1 Root statistics

For $p(x) = \prod_{i=1}^n (x - \lambda_i) = \sum_{k=0}^n a_k\, x^{n-k}$ with $a_0 = 1$, the mean and variance of the root distribution are

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

Lemma 2.1. $\sigma^2(p) = \dfrac{(n-1)\,a_1^2}{n^2} - \dfrac{2\,a_2}{n}.$

Proof. By Vieta's formulas, $\sum_i \lambda_i = -a_1$ and $\sum_{i<j} \lambda_i\lambda_j = a_2$, whence $\sum_i \lambda_i^2 = a_1^2 - 2a_2$. The result follows from $\sigma^2 = \frac{1}{n}\sum_i \lambda_i^2 - \bar\lambda^2$. □

2.2 Symmetric additive convolution

Let A and B be real symmetric matrices with characteristic polynomials p and q. The finite free additive convolution is defined by averaging over the orthogonal group:

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

By the MSS theorem [1], this admits a differential operator representation: if $q(x) = \sum_{k=0}^{n} b_k\, x^{n-k}$, then

$$(p \boxplus_n q)(x) = T_q\, p(x), \qquad T_q = \sum_{k=0}^{n} \frac{(n-k)!}{n!}\, b_k\, \partial_x^k. \tag{2}$$

The coefficients of $r = p \boxplus_n q$, $r(x) = \sum_k c_k\, x^{n-k}$, satisfy

$$c_k = \sum_{i+j=k} \frac{(n-i)!\,(n-j)!}{n!\,(n-k)!}\, a_i\, b_j. \tag{3}$$

Two fundamental properties we shall use repeatedly:

Theorem 2.2 ([1])**.** *If $p, q \in \mathcal{P}_n^{\mathbb{R}}$, then $p \boxplus_n q \in \mathcal{P}_n^{\mathbb{R}}$.*

Lemma 2.3 (Variance additivity)**.** $\sigma^2(p \boxplus_n q) = \sigma^2(p) + \sigma^2(q)$.

Proof. From (3), $c_1 = a_1 + b_1$ and $c_2 = a_2 + \frac{n-1}{n}\, a_1 b_1 + b_2$. Substituting into Lemma 2.1 and expanding $(a_1 + b_1)^2$, the cross-terms $\frac{2(n-1)a_1 b_1}{n^2}$ and $-\frac{2(n-1)a_1 b_1}{n^2}$ cancel, yielding $\sigma^2(p \boxplus_n q) = \sigma^2(p) + \sigma^2(q)$. $\square$

2.3 Scores and Fisher information

Definition 2.1. For $p \in \mathcal{P}_n^{\mathbb{R}}$ with distinct roots $\lambda_1 < \cdots < \lambda_n$, the *score* at λ_i and the *finite free 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 *score-gradient energy* is $\mathcal{S}(p) = \sum_{i<j} \frac{(V_i - V_j)^2}{(\lambda_i - \lambda_j)^2}$.

Lemma 2.4. $V_i = \dfrac{p''(\lambda_i)}{2\, p'(\lambda_i)}$.

Proof. Since $p'(\lambda_i) = \prod_{j \neq i} (\lambda_i - \lambda_j)$, differentiating once more yields $p''(\lambda_i) = 2 \sum_{k \neq i} \prod_{j \neq i,\, j \neq k} (\lambda_i - \lambda_j) = 2\, p'(\lambda_i)\, V_i$. $\square$

Lemma 2.5 (Score identities)**.** (i) $\sum_{i=1}^{n} V_i = 0$.

(ii) $\sum_{i=1}^{n} \lambda_i\, V_i = \binom{n}{2}$.

(iii) $\sum_{i=1}^{n} (\lambda_i - \bar{\lambda})\, V_i = \binom{n}{2}$.

(iv) $\Phi_n(p) = \sum_{i<j} \frac{V_i - V_j}{\lambda_i - \lambda_j}$.

Proof. (i): $\sum_i V_i = \sum_{i \neq j} (\lambda_i - \lambda_j)^{-1} = 0$ by antisymmetry.

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

(iii): Immediate from (ii) and (i).

(iv): $\sum_i V_i^2 = \sum_{i \neq j} \frac{V_i}{\lambda_i - \lambda_j} = \sum_{i<j} \frac{V_i - V_j}{\lambda_i - \lambda_j}$. $\square$

Lemma 2.6 (Fisher–variance inequality)**.** $\Phi_n(p)\,\sigma^2(p) \geq \dfrac{n(n-1)^2}{4}$.

Proof. By Cauchy–Schwarz applied to Lemma 2.5(iii): $\frac{n^2(n-1)^2}{4} \leq \left(\sum_i(\lambda_i-\bar\lambda)^2\right)\left(\sum_i V_i^2\right) = n\,\sigma^2(p)\,\Phi_n(p)$. $\square$

3 The Score-Gradient Inequality

The following algebraic inequality is the key input for the general proof.

Theorem 3.1 (Score-Gradient Inequality)**.** *For $p \in \mathcal{P}_n^{\mathbb{R}}$ of degree $n \geq 2$ with distinct roots,*

$$\mathcal{S}(p)\,\sigma^2(p) \;\geq\; \frac{n-1}{2}\,\Phi_n(p), \tag{4}$$

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

Proof. Write $T = n\,\sigma^2(p)$, $U = \Phi_n(p)$, $S = \mathcal{S}(p)$. The claim is $S\,T \geq \frac{n(n-1)}{2}\,U$.

Step 1. By Lemma 2.5(iii) and Cauchy–Schwarz,

$$\tfrac{n^2(n-1)^2}{4} \;\leq\; T\,U. \tag{5}$$

Step 2. By Lemma 2.5(iv) and Cauchy–Schwarz,

$$U^2 \;\leq\; S\cdot\binom{n}{2}. \tag{6}$$

Step 3. Combining: $S\,T \geq \frac{2U^2}{n(n-1)}\cdot T = \frac{2U}{n(n-1)}\cdot TU \geq \frac{2U}{n(n-1)}\cdot\frac{n^2(n-1)^2}{4} = \frac{n(n-1)}{2}\,U$.

Equality. Equality in (5) requires $V_i = c(\lambda_i - \bar\lambda)$. This implies $\frac{V_i - V_j}{\lambda_i-\lambda_j} = c$ for all $i<j$, which is precisely the equality condition for (6). Conversely, if $\frac{V_i-V_j}{\lambda_i-\lambda_j} = k$ for all $i<j$, then $V_i - k\lambda_i$ is constant; since $\sum_i V_i = 0$, we obtain $V_i = k(\lambda_i - \bar\lambda)$. $\square$

Remark 3.1. The equality condition $V_i = c(\lambda_i - \bar\lambda)$ characterizes, up to affine transformation, the zeros of the Hermite polynomial H_n: evaluating the ODE $H_n'' - 2xH_n' + 2nH_n = 0$ at a zero x_k yields $V_k = x_k$. For $n=2$ this holds for all distinct root configurations.

4 The convolution flow

4.1 The semigroup and the flow

Fix $p, q \in \mathcal{P}_n^{\mathbb{R}}$ with $a = \sigma^2(p) > 0$ and $b = \sigma^2(q) > 0$.

Definition 4.1. Introduce the *normalized coefficients* $\kappa_k(q) = \frac{(n-k)!}{n!}\,b_k$ and the generating polynomial $K_q(z) = \sum_{k=0}^n \kappa_k(q)\,z^k$. The convolution formula (3) is equivalent to $K_{p\boxplus_n q}(z) = K_p(z)\,K_q(z)$. Define the *fractional family* by

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

expanded as a power series and truncated at degree n. Then $q_0 = x^n$, $q_1 = q$, $\sigma^2(q_t) = tb$, and $q_s \boxplus_n q_t = q_{s+t}$. The *flow polynomial* is $p_t = p \boxplus_n q_t$, satisfying $\sigma^2(p_t) = a + tb$.

Remark 4.1. For non-integer t, the polynomial q_t need *not* have all real roots. For example, taking $n=4$ and $q = (x-100)(x-1)(x+1)(x+100)$, one computes $\kappa_4(q_{1/2}) = \frac{1}{2}\kappa_4 - \frac{1}{8}\kappa_2^2 < 0$, and $q_{1/2}$ has only two real roots. This does not affect the proof: only the flow polynomial $p_t = p \boxplus_n q_t$ needs to be real-rooted, which is established in Lemma 4.3 below.

4.2 Perturbation analysis

Lemma 4.1. *Let $\lambda_i(t)$ denote the roots of p_t. Then $\lambda_i(t+h) = \lambda_i(t) + \frac{hb}{n-1} V_i(t) + O(h^2)$.*

Proof. By the semigroup property, $p_{t+h} = p_t \boxplus_n q_h$ with $\sigma^2(q_h) = hb$. The coefficients of q_h satisfy $b_0 = 1$, $b_1 = 0$, $b_2 = -nhb/2 + O(h^2)$, so the operator T_{q_h} acts as $T_{q_h} r(x) = r(x) - \frac{hb}{2(n-1)} r''(x) + O(h^2)$. Setting $\lambda_i(t+h) = \lambda_i(t) + \delta_i$ in $T_{q_h} p_t(\lambda_i(t+h)) = 0$ and solving to first order: $\delta_i = \frac{hb}{2(n-1)} \cdot \frac{p_t''(\lambda_i)}{p_t'(\lambda_i)} + O(h^2) = \frac{hb}{n-1} V_i(t) + O(h^2)$ by Lemma 2.4. $\square$

Lemma 4.2. $\Phi_n(p_{t+h}) = \Phi_n(p_t) - \dfrac{2hb}{n-1} \mathcal{S}(p_t) + O(h^2)$.

Proof. Write $\epsilon = hb/(n-1)$ and suppress the t-dependence. From Lemma 4.1, the perturbed scores are

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

Squaring and summing: $\Phi_n(p_{t+h}) = \sum_i V_i^2 - 2\epsilon \sum_{i \neq j} \frac{V_i(V_i - V_j)}{(\lambda_i - \lambda_j)^2} + O(h^2)$. Pairing (i, j) with (j, i): $\sum_{i \neq j} \frac{V_i(V_i - V_j)}{(\lambda_i - \lambda_j)^2} = \sum_{i<j} \frac{(V_i - V_j)^2}{(\lambda_i - \lambda_j)^2} = \mathcal{S}(p_t)$. $\square$

Lemma 4.3. *For every $t \in [0, 1]$, p_t has n simple real roots.*

Proof. The coefficients of p_t are smooth in t (Definition 4.1), so the roots $\lambda_i(t)$ vary continuously. Since $p_0 = p$ has simple real roots, there is a maximal interval $[0, T)$ on which p_t has simple real roots; continuity gives $T > 0$.

For $t < T$, the root-shift formula (Lemma 4.1) implies $\dot{\delta}_i(t) = \frac{b}{n-1}(V_{i+1}(t) - V_i(t))$ for consecutive gaps $\delta_i = \lambda_{i+1} - \lambda_i$. Expanding the score difference:

$$V_{i+1} - V_i = \frac{2}{\delta_i} - \delta_i \sum_{j \neq i, i+1} \frac{1}{(\lambda_{i+1} - \lambda_j)(\lambda_i - \lambda_j)},$$

so $\dot{\delta}_i \geq \frac{2b}{(n-1)\delta_i} - C$ with C bounded as long as all gaps are positive. Hence $\frac{d}{dt}\delta_i^2 = 2\delta_i \dot{\delta}_i \geq \frac{4b}{n-1} - 2C\delta_i$. For $\delta_i < \delta_0 := 2b/(C(n-1))$, this is strictly positive, so no gap can reach zero in finite time. Since the roots at time T remain separated, continuity extends simple real-rootedness past T, forcing $T \geq 1$. $\square$

4.3 Dissipation and the integral identity

Lemma 4.4 (Dissipation)**.** $\dfrac{d}{dt}\Phi_n(p_t) = -\dfrac{2b}{n-1} \mathcal{S}(p_t)$.

Proof. By Lemma 4.2, $\frac{\Phi_n(p_{t+h}) - \Phi_n(p_t)}{h} = -\frac{2b}{n-1} \mathcal{S}(p_t) + O(h)$. Since p_t has simple roots for all $t \in [0, 1]$ (Lemma 4.3), the scores $V_i(t)$ and hence $\mathcal{S}(p_t)$ are continuous in t. Taking $h \to 0$ yields the result. $\square$

Corollary 4.5 (Integral identity)**.**

$$\frac{1}{\Phi_n(p \boxplus_n q)} - \frac{1}{\Phi_n(p)} = \frac{2b}{n-1} \int_0^1 \frac{\mathcal{S}(p_t)}{\Phi_n(p_t)^2}\, dt. \tag{7}$$

Proof. Set $f(t) = 1/\Phi_n(p_t)$. By the chain rule and Lemma 4.4, $f'(t) = \frac{2b}{n-1} \cdot \frac{\mathcal{S}(p_t)}{\Phi_n(p_t)^2} \geq 0$. By the Fundamental Theorem of Calculus, $f(1) - f(0) = \int_0^1 f'(t)\, dt$. Substituting: $f(0) = 1/\Phi_n(p_0) = 1/\Phi_n(p)$ and $f(1) = 1/\Phi_n(p_1) = 1/\Phi_n(p \boxplus_n q)$ yields (7). $\square$

5 Proof of the main theorem

Theorem 5.1. *Inequality* (1) *holds for every* $n \geq 2$.

Proof. Write $a = \sigma^2(p)$ and $b = \sigma^2(q)$.

Step 1 (Differential inequality). The Score-Gradient Inequality (Theorem 3.1) applied to p_t gives $\mathcal{S}(p_t) \geq \frac{(n-1)\,\Phi_n(p_t)}{2\,\sigma^2(p_t)}$. Substituting into Lemma 4.4:

$$\frac{d}{dt}\Phi_n(p_t) \;\leq\; -\frac{b}{a+tb}\,\Phi_n(p_t).$$

Integrating $\bigl(\log \Phi_n(p_t)\bigr)' \leq -b/(a+tb)$ from 0 to t:

$$\frac{1}{\Phi_n(p_t)} \;\geq\; \frac{a+tb}{a\,\Phi_n(p)}. \tag{8}$$

Step 2 (Forward bound). From Corollary 4.5 and the Score-Gradient Inequality:

$$\frac{1}{\Phi_n(p \boxplus_n q)} - \frac{1}{\Phi_n(p)} \;\geq\; b\int_0^1 \frac{dt}{(a+tb)\,\Phi_n(p_t)}.$$

Substituting (8), the factor $(a+tb)$ cancels:

$$\frac{1}{\Phi_n(p \boxplus_n q)} \;\geq\; \frac{a+b}{a\,\Phi_n(p)}. \tag{9}$$

Step 3 (Reverse bound). Since $p \boxplus_n q = q \boxplus_n p$, repeating Steps 1–2 with p and q interchanged yields

$$\frac{1}{\Phi_n(p \boxplus_n q)} \;\geq\; \frac{a+b}{b\,\Phi_n(q)}. \tag{10}$$

Step 4 (Conclusion). Exactly one of the following holds:

(a) $b\,\Phi_n(q) \geq a\,\Phi_n(p)$. Then $\frac{b}{a\,\Phi_n(p)} \geq \frac{1}{\Phi_n(q)}$, and (9) gives $\frac{1}{\Phi_n(p\boxplus_n q)} \geq \frac{1}{\Phi_n(p)} + \frac{b}{a\,\Phi_n(p)} \geq \frac{1}{\Phi_n(p)} + \frac{1}{\Phi_n(q)}$.

(b) $a\,\Phi_n(p) \geq b\,\Phi_n(q)$. Then $\frac{a}{b\,\Phi_n(q)} \geq \frac{1}{\Phi_n(p)}$, and (10) gives $\frac{1}{\Phi_n(p\boxplus_n q)} \geq \frac{1}{\Phi_n(q)} + \frac{a}{b\,\Phi_n(q)} \geq \frac{1}{\Phi_n(q)} + \frac{1}{\Phi_n(p)}$. □

Remark 5.1. The forward bound (9) and reverse bound (10) are each strictly stronger than the Stam inequality in their respective regimes. Averaging them yields the *half-Stam inequality* $\frac{2}{\Phi_n(p\boxplus_n q)} \geq \frac{1}{\Phi_n(p)} + \frac{1}{\Phi_n(q)}$, from which the full inequality is recovered via the case split.

Remark 5.2. Strict inequality holds generically. Equality in (1) requires that $V_i(p_t) = c(t)\bigl(\lambda_i(t) - \bar{\lambda}(t)\bigr)$ for all $t \in [0,1]$, which forces both p and q to have roots at affinely rescaled zeros of the Hermite polynomial H_n. For $n = 2$ every polynomial satisfies this.

References

[1] A. Marcus, D. A. Spielman, and N. Srivastava, *Interlacing families II: Mixed characteristic polynomials and the Kadison–Singer problem*, Ann. of 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.