THE FINITE FREE STAM INEQUALITY

JORGE GARZA VARGAS, NIKHIL SRIVASTAVA, AND ZACK STIER

Let $\boxplus_n$ and $\Phi_n(\cdot)$ be defined as in the problem statement. In this note we prove the following result, which was conjectured by D. Shlyakhtenko.

Theorem 0.1. *Let $p(x)$ and $q(x)$ be any two monic real-rooted polynomials of degree n. Then*

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

1. Notation and preliminaries

1.1. **Polynomials and the finite free convolution.** Given a polynomial $p(x)$ of degree n we say that $\alpha = (\alpha_1, \ldots, \alpha_n)$ is a vector of roots for $p(x)$ if the α_i are the roots of $p(x)$. We will say that α is ordered if $\alpha_1 \geq \cdots \geq \alpha_n$. Recall that for monic polynomials $p(x)$ and $q(x)$, $p(x) \boxplus_n q(x)$ may be expressed as:

$$(1.1) \qquad p(x) \boxplus_n q(x) = \sum_{\pi \in S_n} \prod_{i=1}^{n} (x - \alpha_i - \beta_{\pi(i)}),$$

where α and β are vectors of roots for $p(x)$ and $q(x)$, respectively, and S_n is the symmetric group on n elements (see Theorem 2.11 of [MSS22] for a proof). Walsh [Wal22] proved that if $p(x)$ and $q(x)$ are real-rooted, then so is $p(x) \boxplus_n q(x)$. Therefore, the finite free convolution induces a map

$$\Omega_{\boxplus_n} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n,$$

where if α and β are vectors of roots for $p(x)$ and $q(x)$, then $\Omega_{\boxplus_n}(\alpha, \beta)$ is defined to be the ordered vector of roots for $p(x) \boxplus_n q(x)$.

Other than the fact that $\boxplus_n$ preserves real-rootedness, our proof will crucially exploit each of the following well-known properties of the finite free convolution. It was shown to us by D. Shlakyhenko. In what follows we will use $\mathbb{1}_n$ to denote the all-ones vector of dimension n. We will use the notation

$$m_k(\alpha) := \frac{1}{n} \sum_{i=1}^{n} \alpha_i^k \quad \text{and} \quad \mathrm{Var}(\alpha) := m_2(\alpha) - m_1(\alpha)^2.$$

Proposition 1.1 (Properties of $\boxplus_n$). *If $\alpha, \beta \in \mathbb{R}^n$ and $\gamma = \Omega_{\boxplus_n}(\alpha, \beta)$, then:*
 i) (Additivity) $m_1(\gamma) = m_1(\alpha) + m_1(\beta)$ and $\mathrm{Var}(\gamma) = \mathrm{Var}(\alpha) + \mathrm{Var}(\beta)$.
 ii) (Commutation with translation) For all $t \in \mathbb{R}$, $\Omega_{\boxplus_n}(\alpha + t\mathbb{1}_n, \beta) = \gamma + t\mathbb{1}_n$ and $\Omega_{\boxplus_n}(\alpha, \beta + t\mathbb{1}_n) = \gamma + t\mathbb{1}_n$.

Proof. (i) Follows from the definition of $p \boxplus_n q$ in terms of the coefficients of p and q and the Newton identities. (ii) Follows from (1.1). $\qquad\square$

1.2. **The heat flow and the finite free Fisher information.** Given a vector of roots $\alpha \in \mathbb{R}^n$ we will define the its finite free score vector $\mathscr{J}_n(\alpha) \in (\mathbb{R} \cup \{\infty\})^n$ as

$$\mathscr{J}_n(\alpha) := \left(\sum_{j:j \neq i} \frac{1}{\alpha_i - \alpha_j} \right)^n_{i=1}.$$

Given a real-rooted polynomial $p(x)$ with vector of roots α, define its finite free Fisher information as

$$\Phi_n(p) := \| \mathscr{J}_n(\alpha) \|^2.$$

The following fact will allow us to write the finite free Fisher information of the polynomial $p(x)$ in terms of the dynamics of its roots under the reverse heat flow.

Lemma 1.1 (Score vectors as derivatives). *Assume $p(x)$ has simple roots. Let $p_t(x) := \exp\left(-\frac{t}{2}\partial_x^2\right) p(x)$ and let $\alpha(t) = (\alpha_1(t), \ldots, \alpha_n(t))$ be the ordered vector of roots of $p_t(x)$. Then*

$$\alpha_i'(0) = \sum_{j:j \neq i} \frac{1}{\alpha_i - \alpha_j},$$

and in particular $\alpha'(0) = \mathscr{J}_n(\alpha)$.

Proof. Since the $\alpha_i(t)$ are continuous in t, the roots remain simple in a neighborhood of $t = 0$. Implicitly differentiating the expression

$$p(\alpha_i(t)) - t p''(\alpha_i(t))/2 + t^2 R(\alpha_i(t), t) = 0$$

(where $R(x,t)$ is a polynomial) at $t = 0$ one obtains

$$\alpha_i'(0) = \frac{1}{2} \frac{p''(\alpha_i)}{p'(\alpha_i)},$$

which is equal to the advertised expression. $\qquad\square$

2. Proof of Stam's inequality

We now prove Theorem 0.1. The following Lemma allows us to restrict attention to the case when p, q, and $p \boxplus_n q$ all have simple roots.

Lemma 2.1 (Approximation by Simple Rooted Polynomials). *Let $\epsilon > 0$ and define the differential operator $T_\epsilon := (1 - \epsilon \cdot d/dx)^n$. If $p(x)$ is a monic real-rooted polynomial of degree n, then*

  i) $(T_\epsilon p)(x)$ is monic and real-rooted of degree n with simple roots.
  ii) $\Phi_n(T_\epsilon p) \to \Phi_n(p)$ as $\epsilon \to 0$.
  iii) $(T_\epsilon p) \boxplus_n (T_\epsilon q) = T_\epsilon^2 (p \boxplus_n q)$.

Proof. (i) was shown in [Nui68]. (ii) is because Φ_n is continuous in the roots of p, which are continuous in ϵ. (iii) follows because $\boxplus_n$ commutes with differential operators (see e.g. [MSS22].) $\qquad\square$

Thus, establishing Theorem 0.1 for the simple case implies the general case by using (iii) above and taking $\epsilon \to 0$. In what follows, $p(x)$ and $q(x)$ are monic real-rooted polynomials, α and β are vectors of roots for $p(x)$ and $q(x)$, $\gamma := \Omega_{\boxplus_n}(\alpha, \beta)$, and α, β, γ all have distinct entries, implying that they are smooth functions of the coefficients of the corresponding polynomials. Let $J_{\boxplus_n}$ denote the Jacobian of $\Omega_{\boxplus_n}$ at the point (α, β).

Our proof can be separated into three steps. The second step is the most substantial one and we will defer its detailed discussion to Section 2.1.

Step 1 (Jacobians and score vectors). We first note that the following relation between score vectors holds.

Observation 2.1 (Relating score vectors)**.** *Using the above notation, for any $a, b \geq 0$*

$$J_{\boxplus_n}(a\, \mathscr{J}_n(\alpha), b\, \mathscr{J}_n(\beta)) = (a + b)\, \mathscr{J}_n(\gamma).$$

Proof. For every $t \geq 0$ let $p_t(x) = \exp(-\frac{t}{2}\partial_x^2)p(x)$, let $\alpha(t)$ be the ordered vector of roots of p_t, and define q_t, r_t and $\beta(t), \gamma(t)$ in an analogous way. Since the finite free convolution commutes with any differential operator, it follows that

$$r_{(a+b)t} = p_{at} \boxplus_n q_{bt}.$$

Hence $\gamma((a+b)t) = \Omega_{\boxplus_n}(\alpha_{at}, \beta_{bt})$ for every t. So, if we differentiate this relation with respect to t, using the chain rule for the right-hand side, we get

$$(a + b)\gamma'(0) = J_{\boxplus_n}\left(\begin{array}{c} a \cdot \alpha'(0) \\ b \cdot \beta'(0) \end{array} \right).$$

A direct application of Lemma 1.1 concludes the proof. $\square$

Step 2 (Understanding the Jacobian). The substance of our proof lies in understanding $J_{\boxplus_n}$. In particular, we will show the following.

Proposition 2.1. *If $u, v \in \mathbb{R}^n$ are orthogonal to $\mathbb{1}_n$ then*

$$\|J_{\boxplus_n}(u, v)\|^2 \leq \|u\|^2 + \|v\|^2.$$

This proposition will be proven in Section 2.1, for now we show how it is used.

Step 3 (Proof of Theorem 0.1 à la Blachman). With Observation 2.1 and Proposition 2.1 in hand we can conclude the proof using the same argument that Blachman used in [Bla65].

Proof of Theorem 0.1. First note that

$$\sum_{i=1}^{n} \sum_{j:j\neq i} \frac{1}{\alpha_i - \alpha_j} = 0,$$

since each term in the sum appears once with a plus and once with a minus. Therefore $\mathscr{J}_n(\alpha)$ is orthogonal to $\mathbb{1}_n$ and, arguing analogously, $\mathscr{J}_n(\beta)$ is orthogonal to $\mathbb{1}_n$. So, Proposition 2.1 implies

$$\|J_{\boxplus_n}(a\, \mathscr{J}_n(\alpha), b\, \mathscr{J}_n(\beta))\|^2 \leq a^2 \|\mathscr{J}_n(\alpha)\|^2 + b^2 \|\mathscr{J}_n(\beta)\|^2.$$

3

Combining this with Observation 2.1 yields

$$(a+b)^2 \|\mathscr{J}_n(\gamma)\|^2 \le a^2 \|\mathscr{J}_n(\alpha)\|^2 + b^2 \|\mathscr{J}_n(\beta)\|^2.$$

Now, by choosing $a = \frac{1}{\|\mathscr{J}_n(\alpha)\|^2}$ and $b = \frac{1}{\|\mathscr{J}_n(\beta)\|^2}$, the above inequality turns into

$$\left(\frac{1}{\|\mathscr{J}_n(\alpha)\|^2} + \frac{1}{\|\mathscr{J}_n(\beta)\|^2}\right)^2 \|\mathscr{J}_n(\gamma)\|^2 \le \frac{1}{\|\mathscr{J}_n(\alpha)\|^2} + \frac{1}{\|\mathscr{J}_n(\beta)\|^2},$$

which after simple algebraic manipulations can be turned into the inequality claimed in Theorem 0.1. $\qquad\square$

2.1. Understanding $J_{\boxplus_n}$.

Let $(\Omega_{\boxplus_n,1}, \dots, \Omega_{\boxplus_n,n})$ be the coordinate functions of $\Omega_{\boxplus_n}$, that is $\gamma_i = \Omega_{\boxplus_n,i}(\alpha, \beta)$. The starting point of our approach to proving Proposition 2.1 is the observation that the matrix $J_{\boxplus_n} J_{\boxplus_n}^*$ is related to the Hessians of the functions $\Omega_{\boxplus_n,i}$. It will be helpful to introduce the notation

$$H_{\boxplus_n}^{(i)} := \mathrm{Hess}_{\Omega_{\boxplus_n,i}}.$$

For this discussion it will prove useful to define the $(2n-2)$-dimensional subspace

$$\mathcal{V} = \left\{ (u,v) \in \mathbb{R}^n \times \mathbb{R}^n : u^* \mathbb{1}_n = v^* \mathbb{1}_n = 0 \right\}.$$

And, given $w \in \mathbb{R}^n \times \mathbb{R}^n$ and $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ we will use $D_w f$ to denote the directional derivative of f in the direction of w, that is $D_w = \sum_i w_i \partial_i$.

Lemma 2.2 (The Hessian of $\Omega_{\boxplus_n}$). *Using the above notation*

$$(2.1) \qquad w^* J_{\boxplus_n} J_{\boxplus_n}^* w = w^* \left(I_n \oplus I_n - \sum_{i=1}^n \gamma_i H_{\boxplus_n}^{(i)} \right) w, \quad \forall w \in \mathcal{V}.$$

Proof. Fix $w = (u,v) \in \mathcal{V}$ and define

$$\alpha(t) := \alpha + tu, \quad \beta(t) := \beta + tv, \quad \text{and} \quad \gamma(t) := \Omega_{\boxplus_n}(\alpha(t), \beta(t)),$$

and note that the variance additivity from Proposition 1.1 i) implies that

$$m_2(\gamma(t)) - m_1(\gamma(t))^2 = m_2(\alpha(t)) + m_2(\beta(t)) - (m_1(\alpha(t))^2 + m_1(\beta(t))^2).$$

Now, the fact that $(u,v) \in \mathcal{V}$ implies that the means $m_1(\alpha(t))$ and $m_1(\beta(t))$ are a constant function of t and therefore, again by Proposition 1.1 i), the mean $m_1(\gamma(t))$ is also a constant function of t. So, differentiating the above equation twice with respect to t we get

$$(2.2) \qquad \partial_t^2 m_2(\gamma(t))\big|_{t=0} = \partial_t^2 \big(m_2(\alpha(t)) + m_2(\beta(t))\big)\big|_{t=0}.$$

Now we inspect both sides of the above equation. First

$$n \, \partial_t^2 m_2(\gamma(t))\big|_{t=0} = \sum_{i=1}^n D_w^2(\gamma_i^2)$$

$$= 2 \sum_{i=1}^n \left((D_w \gamma_i)^2 + \gamma_i D_w^2 \gamma_i \right)$$

4

$$= 2\left(w^* J_{\boxplus_n} J_{\boxplus_n}^* w + \sum_{i=1}^{n} \gamma_i w^* H_{\boxplus_n}^{(i)} w\right).$$

Second

$$n\, \partial_t^2 (m_2(\alpha(t)) + m_2(\beta(t))) = \partial_t^2 ((\alpha + tu)^*(\alpha + tu) + (\beta + tv)^*(\beta + tv))$$
$$= 2(u^*u + v^*v)$$
$$= 2w^*w.$$

Finally, plugging (2.3) and (2.4) back into (2.2) yields

$$w^* J_{\boxplus_n} J_{\boxplus_n}^* w + \sum_{i=1}^{n} \gamma_i w^* H_{\boxplus_n}^{(i)} w = w^*w,$$

which is equivalent to the advertised result. $\square$

We now apply a result of Bauschke et al. [BGLS01, Corollary 3.3].

Theorem 2.2 (Bauschke et al.). *Let $f \in \mathbb{R}[x_1, \ldots, x_m]$ be a hyperbolic polynomial in the direction $w \in \mathbb{R}^m$ and for every $a \in \mathbb{R}^m$ let $\lambda_1(a) \geq \cdots \geq \lambda_m(a)$ be the roots of $g_a(t) := f(a + tw)$. Then, for every $k = 1, \ldots, m$, the function $\sigma_k(a) := \sum_{i=1}^{k} \lambda_i(a)$ is convex in a.*

In our context this implies the following.

Corollary 2.1. *For any real numbers $c_1 \geq \cdots \geq c_n$, the matrix $\sum_{i=1}^{n} c_i H_{\boxplus_n}^{(i)}$ is PSD.*

Proof. Define the multivariate polynomial

$$f(x, a_1, \ldots, a_n, b_1, \ldots, b_n) := \sum_{\pi \in S_n} \prod_{i=1}^{n} (x - a_i - b_{\pi(i)}).$$

Since the above polynomial is homogeneous and the finite free convolution preserves real rootedness, f is hyperbolic in the direction $e_1 = (1, 0 \cdots, 0)$. Now, by Theorem 2.2 the functions

$$\sigma_k(x, a, b) = \sum_{i=1}^{k} \lambda_i(x, a, b)$$

are convex, where $\lambda_1(x, a, b) \geq \cdots \geq \lambda_n(x, a, b)$ denote the roots of $f((x, a, b) + te_1)$. And, because the c_i are ordered we moreover have that the function

$$L(x, a, b) := \sum_{i=1}^{n} c_i \lambda_i(x, a, b)$$

is convex, as it can be written as a positive linear combination of the σ_k. It follows that $\text{Hess}_L = \sum_{i=1}^{n} c_i \text{Hess}_{\lambda_i}$ at any (x, a, b) is PSD. But, on the other hand, when $x = 0$, $a = \alpha$ and $b = \beta$, we have that $\text{Hess}_{\lambda_i} = H_{\boxplus_n}^{(i)}$, which in turn gives that $\sum_{i=1}^{n} c_i H_{\boxplus_n}^{(i)}$ is PSD. $\square$

We can now complete the proof of Proposition 2.1.

Proof of Proposition 2.1. Let $(u, v) \in \mathcal{V}$. Then

$$\|J_{\boxplus_n}(u, v)\|^2 = (u, v)^* J_{\boxplus_n} J_{\boxplus_n}^*(u, v) = \|u\|^2 + \|v\|^2 - \sum_{i=1}^{n} \gamma_i(u, v)^* H_{\boxplus_n}^{(i)}(u, v),$$

where the last equality follows from Lemma 2.2. Now, applying Corollary 2.1 with $c_i = \gamma_i$ gives that $\sum_{i=1}^{n} \gamma_i H_{\boxplus_n}^{(i)}$ is PSD, and hence

$$\sum_{i=1}^{n} \gamma_i(u, v)^* H_{\boxplus_n}^{(i)}(u, v) \geq 0.$$

The proof follows from putting the two expressions together. □

REFERENCES

[BGLS01] Heinz H Bauschke, Osman Güler, Adrian S Lewis, and Hristo S Sendov. Hyperbolic polynomials and convex analysis. *Canadian Journal of Mathematics*, 53(3):470–488, 2001.

[Bla65] Nelson Blachman. The convolution inequality for entropy powers. *IEEE Transactions on Information theory*, 11(2):267–271, 1965.

[MSS22] Adam W Marcus, Daniel A Spielman, and Nikhil Srivastava. Finite free convolutions of polynomials. *Probability Theory and Related Fields*, 182(3):807–848, 2022.

[Nui68] Wim Nuij. A note on hyperbolic polynomials. *Mathematica Scandinavica*, 23(1):69–72, 1968.

[Wal22] Joseph L Walsh. On the location of the roots of certain types of polynomials. *Transactions of the American Mathematical Society*, 24(3):163–180, 1922.