ON THE FINITE FREE STAM INEQUALITY

A COMPENDIUM OF PROVED RESULTS, DEAD ENDS, STRUCTURAL IDENTITIES, AND OPEN DIRECTIONS

ABSTRACT. Let $\mathcal{P}_n^{\mathbb{R}}$ denote the set of monic, degree-n, real-rooted polynomials and let $\boxplus_n$ denote the Marcus–Spielman–Srivastava finite free additive convolution. For $r \in \mathcal{P}_n^{\mathbb{R}}$ with simple roots $\lambda_1 < \cdots < \lambda_n$, define the *Fisher information* $\Phi_n(r) = \sum_{i=1}^n \big(\sum_{j\neq i}(\lambda_i - \lambda_j)^{-1}\big)^2$.

This document is a working compendium for the *finite free Stam inequality*

$$\frac{1}{\Phi_n(p \boxplus_n q)} \geq \frac{1}{\Phi_n(p)} + \frac{1}{\Phi_n(q)}, \qquad p, q \in \mathcal{P}_n^{\mathbb{R}}, \tag{1}$$

the polynomial analogue of classical Stam. The purpose of this version is to give a compact, accurate handoff for future iterations: what is proved, what is only computer-verified, what failed, and what should be attempted next.

Proof-status conventions. Results are classified as follows: [**Proved**] = fully rigorous proof; [**Conditional**] = result whose proof relies on one or more explicitly identified unproved intermediate statements; [**Computer-verified**] = statement verified by exhaustive symbolic or numerical computation (typically $\geq 10^5$ random feasible trials, zero violations) but lacking a closed-form algebraic certificate; [**Proof sketch**] = outline whose key steps are justified but whose full details are deferred; [**Open**] = statement without proof or reduction. All conditional dependencies and computer-verified steps are explicitly flagged in the text.

Current high-level status: $n = 2, 3$ are fully proved; Gaussian-input Stam for all n is proved *conditional* on the root ODE $\dot{\lambda}_i = V_i/(n-1)$ under Hermite flow (which has not yet been derived from first principles); the $n = 4$ argument is rigorously reduced to computer-verified polynomial inequalities that await closed-form algebraic certificates; a quantitative local Stam inequality near the Hermite manifold is proved for all n (as a proof sketch with an identified uniform-in-w gap); global Stam for $n \geq 5$ remains open. See Table 1 for a complete classification.

CONTENTS

Global status table 4
Global status table 4

Part 1. Foundations 5
1. Setup and definitions 5
1.1. MSS finite free additive convolution 5
1.2. Scores and Fisher information 5
1.3. Variance and additive parameters 5
2. Proved structural identities 6
2.1. Fisher–repulsion identity 6
2.2. Fisher–trace–Laplacian identity 6
2.3. Score identities 6
2.4. Fisher–variance and score-gradient inequalities 7
2.5. Variance additivity and derivative compatibility 7
2.6. Bezoutian representation 8
2.7. Harmonicity of log disc in matrix coordinates 8
2.8. Isoperimetric inequality 8
3. Proved special cases 9
3.1. The $n = 2$ case (equality) 9
3.2. The $n = 3$ case: SOS proof via log-cumulants 9
3.3. Full Stam when one input is Gaussian 9

Date: February 13, 2026.

4. The Stieltjes/Herglotz framework 10
4.1. Transforms 10
4.2. Pick matrix positivity 10
4.3. Contour integral for Φ_n 10
4.4. The Stieltjes PDE under dilation 10
4.5. De Bruijn identity 10
4.6. Cumulant-ratio defect positivity 11
5. The spectral efficiency reformulation 11

Part 2. Dead Ends: Compact Post-Mortem 11
6. Route A 12
7. Route B 12
8. Route C 13
9. Route D 13
9.1. Hessian of $1/\Phi_n$ in log-cumulant coordinates 13
10. Route E 13
11. Route F 13
12. Route G 13
13. Route H 13
14. Route I 13
15. PF sequences / total positivity 13
16. Interlacing / Jensen route 13
17. Induction on degree 13

Part 3. Numerical Landscape 13
18. Master numerical summary 13
18.1. Inequalities and identities that hold universally 13
18.2. Inequalities that fail 14
18.3. Defect scaling law 15

Part 4. Open Conjectures 15

Part 5. Future Directions: Compact Route Maps 15
19. Option C: Marginal / hypergraph decomposition 16
20. Option D: Optimal transport / entropy dissipation 16
21. Additional plausible ideas 16

Part 6. Route J: Cauchy–Schwarz Mixing and the Spectral Efficiency Defect 16
22. The normalised Fisher information G_n 16
23. The Cauchy–Schwarz mixing inequality 17
24. Stam via the spectral efficiency defect 17
25. Hessian of G_n at the Gaussian point 18
26. Quadratic Stam lower bound 18
27. Third proof of Stam for $n = 3$ 19
28. Structure of R_4 and the kurtosis axis 19
29. Complete Stam inequality for $n = 4$ 20
30. General n: the sub-averaging decomposition 22

Part 7. Three Viable Proof Strategies Using MSS Interlacing and Real Stability 23
31. Background: MSS interlacing and real stability 23
32. Approach K: Induction on degree via Score–Cauchy identities 24
32.1. Overview 24
32.2. K-cumulant preservation 24

32.3. The Cauchy interlacing matrix 24
32.4. The deficit telescoping theorem 25
32.5. Numerical evidence for chain dominance 26
32.6. The consecutive deficit ratio 26
32.7. Strengthening Stam: super-additivity margin 27
32.8. Proof of the Stam inequality for all n (modulo chain dominance) 27
32.9. Proof architecture: rigorous path to chain dominance 27
33. Approach L: Real stable symbol and Newton's wall 28
33.1. Overview 28
33.2. The K-transform is real stable 28
33.3. Newton wall: compact feasibility of $\boldsymbol{\tau}$ 28
33.4. The spectral efficiency R_n 29
33.5. Quadratic sub-averaging theorem 30
33.6. Remainder control 30
33.7. Strengthened sub-averaging 31
33.8. Reduction to finite optimisation 31
33.9. Boundary behaviour 31
33.10. Proof architecture: closing the gap 31
34. Approach M: Haar averaging and the mixed characteristic polynomial 32
34.1. Overview 32
34.2. Fisher Jensen inequality 32
34.3. Failure of the harmonic-mean bound 32
34.4. The concavity pathway 33
34.5. Variance additivity and Fisher–variance bound 33
34.6. Score alignment 34
34.7. Eigenvalue fluctuations of M_Q 34
34.8. Revised proof strategy 34
34.9. Monte-Carlo precision limitations 34
35. Comparative assessment and recommended priorities 35

Part 8. Deep Gate Analysis: Summary of Findings 35
36. Gate 2 is a dead end: ladder monotonicity fails 36
37. Gate 1 is a dead end: remainder positivity fails 36
38. New structural facts 36
38.1. Frobenius reduction bound 36
38.2. Universal level-wise Stam 37
38.3. R_n convexity landscape 37
39. Proved corollary: quantitative local Stam for all n 37
40. Proof strategy: proportional Hermite flow 38

Part 9. Conclusion 38
41. Comprehensive handoff for the next agent 38
References 40

Global status table

Table 1 classifies every major result in this document. The status categories are defined in the abstract. References point to the theorem or observation where each result is stated.

Table 1. Proof-status classification of all major results.

Result	**Status**	**Reference**
Structural identities (Part 1)		
Fisher–repulsion $\Phi_n = 2\mathcal{R}$	[**Proved**]	Thm 2.1
Fisher–trace–Laplacian	[**Proved**]	Thm 2.3
Score identities (i)–(iv)	[**Proved**]	Lem 2.4
Fisher–variance inequality	[**Proved**]	Thm 2.5
Score-gradient inequality	[**Proved**]	Thm 2.6
Variance additivity	[**Proved**]	Lem 2.7
Derivative compatibility	[**Proved**]	Lem 2.8
Bezoutian representation	[**Proved**]	Thm 2.9
Harmonicity of $\log \operatorname{disc}$	[**Proof sketch**] + [**Computer-verified**]	Thm 2.10
Contour integral for Φ_n	[**Proved**]	Thm 4.3
Isoperimetric inequality	[**Proved**]	Prop 2.12
Cumulant-ratio defect $\Delta_k \geq 0$	[**Proved**]	Lem 4.9
CS mixing inequality	[**Proved**]	Lem 23.1
Special cases (Part 1)		
Stam for $n = 2$ (equality)	[**Proved**]	§3
Stam for $n = 3$ (SOS proof)	[**Proved**]	Thm 3.1
Stam for $n = 3$ (CS mixing)	[**Proved**]	Thm 27.1
De Bruijn identity	[**Conditional**]	Thm 4.6
Gaussian-input Stam, all n	[**Conditional**]	Thm 3.3
CS mixing framework (Part 6)		
Hessian of G_n, $n = 3, 4$	[**Proved**]	Thm 25.1
Hessian of G_n, $n \geq 5$	[**Computer-verified**]	Thm 25.1
$R_3 = \frac{9}{8}\tau_3^2$ (exact)	[**Proved**]	Thm 33.6
Quadratic Stam, $n = 3$	[**Proved**]	Thm 26.1
Quadratic Stam, $n \geq 4$	[**Conditional**]on Hessian	Thm 26.1
KStam kurtosis axis, $n = 4$	[**Proved**] + [**Computer-verified**]	Thm 28.2
Stam for $n = 4$	[**Computer-verified**] (3 ineqs)	Thm 29.1
MSS interlacing framework (Part 7)		
K-cumulant preservation	[**Proved**]	Thm 32.1
Score–Cauchy identity	[**Proved**]	Thm 32.4
Column-sum vanishing	[**Proved**]	Thm 32.5
Frobenius norm $\|C\|_F^2 = 4\Phi_n$	[**Proved**]	Thm 32.6
Deficit telescoping	[**Proved**]	Thm 32.10
Chain dominance $D_n \geq \delta_n D_3$	[**Computer-verified**]	Conj 32.11
Gate analysis and local results (Part 8)		
Local Stam for all n	[**Proof sketch**]	Thm 39.1
$\mathcal{E}_n \geq 0$	[**Dead End**]	Obs 37.1
Ladder monotonicity $D_k \geq D_{k-1}$	[**Dead End**]	Obs 36.1
Universal level-wise Stam	[**Computer-verified**]	Obs 38.2
Frobenius reduction bound	[**Computer-verified**]	Obs 38.1
Open conjectures		
Stam for all $n \geq 4$	[**Open**]	Conj 18.2
R_n sub-averaging	[**Open**]	Conj 33.11
Flow monotonicity	[**Open**]	Conj 40.1
Real stability of K_r	[**Open**]	Conj 33.1
Newton wall compactness	[**Open**]	Conj 33.3

Part 1. Foundations

1. Setup and definitions

1.1. MSS finite free additive convolution.

Definition 1.1 (MSS convolution)**.** Write $p(x) = \sum_{k=0}^{n} a_k x^{n-k}$, $q(x) = \sum_{k=0}^{n} b_k x^{n-k}$ with $a_0 = b_0 = 1$. The *finite free additive convolution* $r = p \boxplus_n q$ is defined by

$$r(x) = \sum_{k=0}^{n} c_k x^{n-k}, \qquad c_k = \sum_{i+j=k} \frac{(n-i)!\,(n-j)!}{n!\,(n-k)!}\, a_i b_j.$$

By the Marcus–Spielman–Srivastava theorem [1], $\boxplus_n$ preserves $\mathcal{P}_n^{\mathbb{R}}$: if $p, q \in \mathcal{P}_n^{\mathbb{R}}$ then $p \boxplus_n q \in \mathcal{P}_n^{\mathbb{R}}$.

Definition 1.2 (K-transform and log-cumulants)**.** Define $\kappa_k(r) := (n-k)!\, c_k(r)/n!$ and $K_r(z) := \sum_{k=0}^{n} \kappa_k(r)\, z^k$. Then $\boxplus_n$ becomes multiplicative:

$$K_{p \boxplus_n q}(z) = K_p(z) \cdot K_q(z) \pmod{z^{n+1}}. \tag{2}$$

The *log-cumulants* $\ell_k(r) := [z^k] \log K_r(z)$ are computed by $\ell_1 = \kappa_1$, $\ell_k = \kappa_k - \frac{1}{k} \sum_{j=1}^{k-1} j\, \ell_j \kappa_{k-j}$ for $k \geq 2$. They are **additive**: $\ell_k(p \boxplus_n q) = \ell_k(p) + \ell_k(q)$ for all k.

1.2. Scores and Fisher information.

Definition 1.3 (Scores, Fisher information, repulsion)**.** For $r \in \mathcal{P}_n^{\mathbb{R}}$ with simple roots $\lambda_1 < \cdots < \lambda_n$:

$$V_i(r) := \sum_{j \neq i} \frac{1}{\lambda_i - \lambda_j}, \qquad V = (V_1, \ldots, V_n) \quad \text{(score vector)}, \tag{3}$$

$$\Phi_n(r) := \sum_{i=1}^{n} V_i^2 \quad \text{(Fisher information)}, \tag{4}$$

$$\mathcal{R}(r) := \sum_{1 \leq i < j \leq n} \frac{1}{(\lambda_i - \lambda_j)^2} \quad \text{(repulsion energy)}, \tag{5}$$

$$\mathcal{S}(r) := \sum_{1 \leq i < j \leq n} \frac{(V_i - V_j)^2}{(\lambda_i - \lambda_j)^2} \quad \text{(score-gradient energy)}. \tag{6}$$

If r has a repeated root, we set $\Phi_n(r) = \infty$.

Definition 1.4 (Graph Laplacian)**.** The *graph Laplacian* of r is $L \in \mathbb{R}^{n \times n}$ with

$$L_{ij} = \begin{cases} -(\lambda_i - \lambda_j)^{-2} & i \neq j, \\ \sum_{k \neq i} (\lambda_i - \lambda_k)^{-2} & i = j. \end{cases}$$

This is the complete-graph Laplacian with edge weights $w_{ij} = (\lambda_i - \lambda_j)^{-2}$. We have $L\mathbf{1} = 0$, $L \succeq 0$, $\ker L = \operatorname{span}\{\mathbf{1}\}$, $\operatorname{rank} L = n-1$. Equivalently, $L = -\frac{1}{2} \operatorname{Hess}_\lambda(\log \operatorname{disc}(r))$.

1.3. Variance and additive parameters.

Definition 1.5 (Additive mean and variance)**.** For $r \in \mathcal{P}_n^{\mathbb{R}}$ with roots $\lambda_1 < \cdots < \lambda_n$:

$$\mu(r) := -\frac{a_1(r)}{n} = \frac{1}{n} \sum_{i=1}^{n} \lambda_i \qquad \text{(centroid)}, \tag{7}$$

$$\sigma^2(r) := \frac{1}{n} \sum_{i=1}^{n} (\lambda_i - \mu)^2 = \frac{(n-1)\, a_1(r)^2 - 2n\, a_2(r)}{n^2} \qquad \text{(empirical variance)}. \tag{8}$$

Both are **additive**: $\mu(p \boxplus_n q) = \mu(p) + \mu(q)$ and $\sigma^2(p \boxplus_n q) = \sigma^2(p) + \sigma^2(q)$.

Definition 1.6 (Finite Gaussian)**.** The *finite Gaussian* $g_t \in \mathcal{P}_n^{\mathbb{R}}$ of variance $t > 0$ is the unique centred polynomial whose K-transform has $\kappa_k(g_t) = 0$ for all $k \neq 0, 2$. Equivalently, all log-cumulants $\ell_k(g_t) = 0$ for $k \geq 3$, so g_t is characterised by $\sigma^2(g_t) = t$. The Hermite semigroup satisfies $g_s \boxplus_n g_t = g_{s+t}$. For example: $g_t = x^2 - t$ at $n = 2$; $g_t = x^3 - \frac{3}{2} t\, x$ at $n = 3$.

Definition 1.7 (Normalised cumulant ratios)**.** For centred $r \in \mathcal{P}_n^{\mathbb{R}}$ with $u := -\ell_2(r) > 0$, define $\tau_k(r) := \ell_k(r)/u(r)^{k/2}$ for $k \geq 3$. The *additive variance parameter* $u = \sigma^2/(2(n-1))$ satisfies $u(p \boxplus_n q) = u(p) + u(q)$.

Lemma 1.8 (Normalisation identities)**.** *For centred $r \in \mathcal{P}_n^{\mathbb{R}}$ (i.e., $\mu(r) = 0$), the parameters κ_2, ℓ_2, u, and σ^2 are related by:*

$$\ell_2 = \kappa_2 = \frac{(n-2)!\,a_2}{n!} = \frac{a_2}{n(n-1)}, \qquad u := -\ell_2 > 0, \qquad \sigma^2 = 2(n-1)\,u. \tag{9}$$

Here a_2 is the coefficient of x^{n-2} in r (so $a_2 < 0$ for centred real-rooted r with $n \geq 2$).

Proof. From Definition 1.2: $\kappa_2 = (n-2)!\,a_2/n!$. The log-cumulant recurrence (Definition 1.2) gives $\ell_2 = \kappa_2 - \frac{1}{2}\kappa_1^2 = \kappa_2$ when r is centred ($\kappa_1 = \ell_1 = 0$). From (8) with $a_1 = 0$: $\sigma^2 = -2a_2/n = -2n(n-1)\ell_2/n = 2(n-1)(-\ell_2) = 2(n-1)u$. All three parameters are additive under $\boxplus_n$ because ℓ_2 is additive (Definition 1.2). □

2. Proved structural identities

We collect all rigorously established identities. These are the "library components" on which any future proof of (1) can draw.

2.1. Fisher–repulsion identity.

Theorem 2.1 (Fisher–repulsion identity)**.** *For any $r \in \mathcal{P}_n^{\mathbb{R}}$ with simple roots,*

$$\Phi_n(r) = 2\,\mathcal{R}(r). \tag{10}$$

Proof. Expand $V_i^2 = \sum_{j\neq i}\sum_{k\neq i}(\lambda_i - \lambda_j)^{-1}(\lambda_i - \lambda_k)^{-1}$ and sum over i. The diagonal terms ($j = k$) contribute $\sum_i \sum_{j\neq i}(\lambda_i - \lambda_j)^{-2} = 2\sum_{i<j}(\lambda_i - \lambda_j)^{-2} = 2\mathcal{R}$. The cross terms ($j \neq k$, both $\neq i$) group into triples $\{a, b, c\}$, each contributing

$$\frac{1}{(a-b)(a-c)} + \frac{1}{(b-a)(b-c)} + \frac{1}{(c-a)(c-b)} = 0$$

(partial-fraction identity for the residues of $1/((x-a)(x-b)(x-c))$). Hence $\Phi_n = 2\mathcal{R} + 0 = 2\mathcal{R}$. □

Corollary 2.2 (Stam as harmonic-mean repulsion)**.** *Inequality (1) is equivalent to $1/\mathcal{R}(p \boxplus_n q) \geq 1/\mathcal{R}(p) + 1/\mathcal{R}(q)$.*

2.2. Fisher–trace–Laplacian identity.

Theorem 2.3 (Fisher $= \operatorname{tr}(L) = \lambda^T L^2 \lambda$)**.** *For $r \in \mathcal{P}_n^{\mathbb{R}}$:*

(a) $\Phi_n = \operatorname{tr}(L) = 2\mathcal{R}$.
(b) $V = L\lambda$ (Euler identity).
(c) $\lambda^T L \lambda = \binom{n}{2}$ (universal constant).
(d) $\Phi_n = \|V\|^2 = \|L\lambda\|^2 = \lambda^T L^2 \lambda$.

Proof. (a) follows from $\Phi_n = 2\mathcal{R}$ and $\operatorname{tr}(L) = 2\sum_{i<j}(\lambda_i - \lambda_j)^{-2}$.

(b) We compute $(L\lambda)_i = \sum_{j\neq i} \frac{\lambda_i - \lambda_j}{(\lambda_i - \lambda_j)^2} = \sum_{j\neq i} \frac{1}{\lambda_i - \lambda_j} = V_i$.

(c) $\lambda^T L \lambda = V \cdot \lambda = \frac{1}{2}\sum_i \partial_{\lambda_i} \log \operatorname{disc} \cdot \lambda_i = \frac{n(n-1)}{2}$ by the Euler identity for homogeneity of disc (degree $n(n-1)$).

(d) is immediate from $V = L\lambda$. □

2.3. Score identities.

Lemma 2.4 (Score identities)**.** *For $r \in \mathcal{P}_n^{\mathbb{R}}$ with simple roots:*

(i) $\sum_i V_i = 0$.
(ii) $\sum_i (\lambda_i - \mu)V_i = \binom{n}{2}$ *for any $\mu \in \mathbb{R}$.*
(iii) $\Phi_n = \sum_{i<j} \frac{V_i - V_j}{\lambda_i - \lambda_j}$.
(iv) $V_i = r''(\lambda_i)/(2r'(\lambda_i))$ (score-of-derivative identity).

Proof. (i) By symmetry: $\sum_i V_i = \sum_{i\neq j}(\lambda_i - \lambda_j)^{-1} = 0$ (antisymmetric sum).

(ii) $\sum_i \lambda_i V_i = \sum_{i\neq j} \lambda_i/(\lambda_i - \lambda_j) = \sum_{i\neq j}[1 + \lambda_j/(\lambda_i - \lambda_j)] = n(n-1) + \sum_{i\neq j} \lambda_j/(\lambda_i - \lambda_j)$. Using $\sum_{i\neq j} \lambda_j/(\lambda_i - \lambda_j) = -\sum_{i\neq j} \lambda_i/(\lambda_j - \lambda_i) = -\sum_i \lambda_i V_i$, we get $2\sum_i \lambda_i V_i = n(n-1)$, so $\sum_i \lambda_i V_i = \binom{n}{2}$. By (i), subtracting $\mu \sum V_i = 0$ gives (ii).

(iii) $\sum_{i<j}(V_i - V_j)/(\lambda_i - \lambda_j) = \sum_i V_i \sum_{j\neq i}(\lambda_i - \lambda_j)^{-1} - \sum_{i<j}[V_j/(\lambda_i - \lambda_j) + V_i/(\lambda_j - \lambda_i)]$. Rewriting the double sum: for each pair $i < j$, the term $V_i/(\lambda_i - \lambda_j) + V_j/(\lambda_j - \lambda_i)$ contributes to $\sum_k V_k \cdot$ (sum of $1/(\lambda_k - \lambda_m)$ for $m \neq k$). In fact,

$$\begin{aligned}\sum_{i<j} \frac{V_i - V_j}{\lambda_i - \lambda_j} &= \sum_{i<j} \frac{1}{\lambda_i - \lambda_j}\Big(\sum_{k\neq i} \frac{1}{\lambda_i - \lambda_k} - \sum_{k\neq j} \frac{1}{\lambda_j - \lambda_k}\Big) \\ &= \sum_{i<j}\sum_{k\neq i} \frac{1}{(\lambda_i - \lambda_j)(\lambda_i - \lambda_k)} - \sum_{i<j}\sum_{k\neq j} \frac{1}{(\lambda_i - \lambda_j)(\lambda_j - \lambda_k)}.\end{aligned}$$

Relabelling the second sum by interchanging $i \leftrightarrow j$ (which reverses both the sign of $(\lambda_i - \lambda_j)$ and the ordering to $j < i$, i.e. $i > j$) yields

$$\sum_{j<i}\sum_{k\neq i} \frac{1}{(\lambda_j - \lambda_i)(\lambda_i - \lambda_k)} = -\sum_{i<j}\sum_{k\neq j} \frac{1}{(\lambda_i - \lambda_j)(\lambda_j - \lambda_k)},$$

so the two double sums combine to give

$$2\sum_{i<j}\sum_{k\neq i} \frac{1}{(\lambda_i - \lambda_j)(\lambda_i - \lambda_k)}.$$

Separating the diagonal terms ($k = j$) from the cross terms ($k \neq j$, $k \neq i$): the diagonal terms give $\sum_{i<j}(\lambda_i - \lambda_j)^{-2}$ summed over the appropriate range, yielding $\sum_i \sum_{j\neq i}(\lambda_i - \lambda_j)^{-2} = \Phi_n$. The cross terms, grouped into triples $\{i, j, k\}$, each contribute $1/((\lambda_i - \lambda_j)(\lambda_i - \lambda_k)) + 1/((\lambda_j - \lambda_i)(\lambda_j - \lambda_k)) + 1/((\lambda_k - \lambda_i)(\lambda_k - \lambda_j)) = 0$ by the partial-fraction identity $\sum_{\text{cyc}} 1/((a-b)(a-c)) = 0$ (the same identity as in Theorem 2.1). Hence $\sum_{i<j}(V_i - V_j)/(\lambda_i - \lambda_j) = \Phi_n$.

(iv) Since $r'(\lambda_i) = \prod_{j\neq i}(\lambda_i - \lambda_j)$, $V_i = \sum_{j\neq i}(\lambda_i - \lambda_j)^{-1} = r''(\lambda_i)/(2r'(\lambda_i))$. □

2.4. **Fisher–variance and score-gradient inequalities.**

Theorem 2.5 (Fisher–variance inequality). $\Phi_n(r) \cdot \sigma^2(r) \geq n(n-1)^2/4$.

Proof. Cauchy–Schwarz on $\sum_i(\lambda_i - \mu)V_i = \binom{n}{2}$ with $\sum V_i = 0$: $\left|\sum(\lambda_i - \mu)V_i\right|^2 \leq \left(\sum(\lambda_i - \mu)^2\right)\left(\sum V_i^2\right) = n\sigma^2 \cdot \Phi_n$. □

Theorem 2.6 (Score-Gradient Inequality). *For simple-root $r \in \mathcal{P}_n^{\mathbb{R}}$, $n \geq 2$: $\mathcal{S}(r) \cdot \sigma^2(r) \geq \frac{n-1}{2}\Phi_n(r)$.*

Proof. Two Cauchy–Schwarz applications. From Lemma 2.4(ii): $n\sigma^2 \cdot \Phi_n \geq n^2(n-1)^2/4$. From Lemma 2.4(iii): $\Phi_n^2 \leq \mathcal{S} \cdot n(n-1)/2$. Combining: $\mathcal{S}\sigma^2 \geq (n-1)\Phi_n/2$. □

2.5. **Variance additivity and derivative compatibility.**

Lemma 2.7 (Variance additivity). $\sigma^2(p \boxplus_n q) = \sigma^2(p) + \sigma^2(q)$.

Proof. From the MSS coefficient formula (Definition 1.1): $c_1 = a_1 + b_1$, $c_2 = a_2 + b_2 + \frac{n-1}{n}a_1b_1$. Using (8):

$$\begin{aligned}\sigma^2(p \boxplus_n q) &= \frac{(n-1)(a_1+b_1)^2 - 2n(a_2+b_2+\frac{n-1}{n}a_1b_1)}{n^2} \\ &= \frac{(n-1)a_1^2 - 2na_2}{n^2} + \frac{(n-1)b_1^2 - 2nb_2}{n^2} + \frac{2(n-1)a_1b_1 - 2(n-1)a_1b_1}{n^2} \\ &= \sigma^2(p) + \sigma^2(q).\end{aligned}$$

□

Lemma 2.8 (Derivative compatibility). $(p \boxplus_n q)'/n = (p'/n) \boxplus_{n-1} (q'/n)$.

Proof. Write $p = \sum a_k x^{n-k}$, $q = \sum b_k x^{n-k}$. Then $(p \boxplus_n q)(x) = \sum c_k x^{n-k}$ with $c_k = \sum_{i+j=k} \frac{(n-i)!(n-j)!}{n!(n-k)!} a_i b_j$. The derivative of p is $p'(x)/n = \sum_{k=0}^{n-1} \frac{n-k}{n} a_k x^{n-1-k}$, which is a monic degree-$(n-1)$ polynomial with coefficients $\tilde{a}_k = (n-k)a_k/n$. A direct calculation confirms that the $\boxplus_{n-1}$ convolution of p'/n and q'/n reproduces $(p \boxplus_n q)'/n$, using the identity $\frac{(n-i)(n-j)}{n^2} \cdot \frac{(n-1-i)!(n-1-j)!}{(n-1)!(n-1-k)!} = \frac{(n-k)}{n} \cdot \frac{(n-i)!(n-j)!}{n!(n-k)!}$ for $i + j = k$. □

2.6. **Bezoutian representation.**

Theorem 2.9 (Bezoutian representation of Φ_n)**.** *Let* $\mathrm{Bez}(r, r')$ *denote the Bezoutian of* r *and* r' [9]*. Then*

$$\Phi_n(r) = \left\| \frac{r''}{2} \right\|^2_{\mathrm{Bez}(r,r')} = \sum_{i=1}^{n} \frac{r''(\lambda_i)^2}{4\, r'(\lambda_i)^2}.$$

Proof. The Bezoutian matrix $\mathrm{Bez}(r, r')$ is the unique symmetric matrix $B \in \mathbb{R}^{n\times n}$ satisfying $\sum_{i,j} B_{ij} x^{n-1-i} y^{n-1-j} = (r(x)r'(y) - r'(x)r(y))/(x - y)$. The associated inner product is diagonal in the Lagrange basis $\{L_i(x) = \prod_{j\neq i}(x - \lambda_j)/\prod_{j\neq i}(\lambda_i - \lambda_j)\}$: $\langle f, g\rangle_{\mathrm{Bez}} = \sum_i f(\lambda_i)g(\lambda_i)/r'(\lambda_i)^2$ (see [9, §3] for the diagonalisation in the Lagrange basis). Since $V_i = r''(\lambda_i)/(2r'(\lambda_i))$ (Lemma 2.4(iv)), we get $\Phi_n = \sum V_i^2 = \|r''/2\|^2_{\mathrm{Bez}}$. □

2.7. **Harmonicity of** $\log \operatorname{disc}$ **in matrix coordinates.**

Theorem 2.10 (Harmonicity of $\log \operatorname{disc}$)**.** ***[Proof sketch]** + **[Computer-verified]***
Let $A \in \mathrm{Sym}(n)$ *have simple eigenvalues. Then*

$$\Delta_A \log \operatorname{disc}(\det(xI - A)) = 0,$$

where Δ_A *is the Laplace–Beltrami operator on* $\mathrm{Sym}(n)$*. The eigenvalue Laplacian contribution* $-2\Phi_n$ *is exactly cancelled by the rotation Laplacian* $+2\Phi_n$ *from off-diagonal perturbations.*

Proof status. *The argument below is a* proof sketch*: the diagonal perturbation computation is complete, but the off-diagonal perturbation requires second-order eigenvalue perturbation theory whose algebra is outlined but not expanded line-by-line. The identity has been verified symbolically at* $n = 3$–8 *(error* $< 5 \times 10^{-16}$*).*

Proof sketch. Second-order perturbation theory [10] on $A = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.

Diagonal perturbations. For $H = E_{kk}$, $\partial_H \lambda_k = 1$ and $\partial_H \lambda_j = 0$ for $j \neq k$, so $\partial^2_H \log \operatorname{disc} = -2\sum_{j\neq k}(\lambda_k - \lambda_j)^{-2}$. Summing over k: diagonal Laplacian $= -2\sum_k \sum_{j\neq k}(\lambda_k - \lambda_j)^{-2} = -2\Phi_n$.

Off-diagonal perturbations. For $H = (E_{ab} + E_{ba})/\sqrt{2}$ with $a < b$: $A + \epsilon H$ has eigenvalues $\lambda_i + O(\epsilon^2)$ for $i \neq a, b$ and $\frac{\lambda_a+\lambda_b}{2} \pm \sqrt{(\frac{\lambda_a-\lambda_b}{2})^2 + \epsilon^2/2}$ for the (a, b) pair. At second order: $\partial^2_{ab}\lambda_a = -\partial^2_{ab}\lambda_b = (\lambda_a - \lambda_b)^{-1}$, and for all other eigenvalues the second derivatives involve $(\lambda_i - \lambda_a)^{-1}(\lambda_i - \lambda_b)^{-1}$ terms. Summing the resulting $\partial^2_{ab} \log \operatorname{disc}$ over all $\binom{n}{2}$ pairs (a, b) and using the identity $\sum_{k\neq a,b}\big[(\lambda_k - \lambda_a)^{-1} - (\lambda_k - \lambda_b)^{-1}\big](\lambda_a - \lambda_b)^{-1} = \sum_{k\neq a,b}\big[(\lambda_k - \lambda_a)^{-1}(\lambda_a - \lambda_b)^{-1} + (\lambda_b - \lambda_k)^{-1}(\lambda_a - \lambda_b)^{-1}\big]$ gives off-diagonal Laplacian $= +2\Phi_n$.

Combined: $-2\Phi_n + 2\Phi_n = 0$. (Full perturbation algebra verified symbolically at $n = 3$–8.) □

Remark 2.11. This result is a **fundamental structural obstruction**: Φ_n *cannot* be captured by a matrix-level convexity argument (such as Alexandrov–Fenchel or Loewner ordering). The eigenvalue directions and the rotation directions exactly cancel, so any proof must work in eigenvalue coordinates alone.

2.8. **Isoperimetric inequality.**

Proposition 2.12 (AM–GM isoperimetric)**.** *With* $M = \binom{n}{2}$*:* $\Phi_n(r)\cdot \operatorname{disc}(r)^{1/M} \geq 2M = n(n-1)$*.*

Proof. AM–GM on the M positive terms $(\lambda_i - \lambda_j)^{-2}$: $\frac{\Phi_n}{2M} \geq \left(\prod_{i<j}(\lambda_i - \lambda_j)^{-2}\right)^{1/M} = \operatorname{disc}(r)^{-1/M}$. □

3. Proved special cases

3.1. The $n = 2$ case (equality). $\Phi_2(r) = 2/(\lambda_1 - \lambda_2)^2 = 1/(2\sigma^2)$. Hence $1/\Phi_2 = 2\sigma^2$, and Stam reduces to $2\sigma^2(p \boxplus_n q) \geq 2\sigma^2(p) + 2\sigma^2(q)$, which is variance additivity (equality always holds).

3.2. The $n = 3$ case: SOS proof via log-cumulants.

Theorem 3.1 (Stam for $n = 3$). *For centred $p, q \in \mathcal{P}_3^{\mathbb{R}}$ with $u_p = -\ell_2(p) > 0$, $u_q = -\ell_2(q) > 0$:*

$$D_3 := \frac{1}{\Phi_3(r)} - \frac{1}{\Phi_3(p)} - \frac{1}{\Phi_3(q)} = \frac{3}{2}\left[(1-w)\alpha^2 + w(1-w)(\alpha-\beta)^2 + w\beta^2\right] \geq 0, \tag{11}$$

where $r = p \boxplus_3 q$, $\alpha = \ell_3(p)/u_p$, $\beta = \ell_3(q)/u_q$, $w = u_p/(u_p + u_q)$. Equality holds iff $\ell_3(p) = \ell_3(q) = 0$ (both polynomials have symmetric roots).

Proof. Step 1: deriving $1/\Phi_3$. For centred $r \in \mathcal{P}_3^{\mathbb{R}}$, write $r(x) = x^3 + e_2 x + e_3$ (with $e_1 = 0$). The log-cumulants (Definition 1.2) satisfy $u := -\ell_2 > 0$ and $v := \ell_3$, related to the coefficients by the moment–cumulant inversion for $K_r(z) = 1 + \kappa_1 z + \kappa_2 z^2 + \kappa_3 z^3$ with $\kappa_k = (n-k)!\, e_k/n!$: specifically $\kappa_2 = e_2/6$, $\kappa_3 = e_3/6$, $\ell_2 = \kappa_2 = e_2/6$, and (because $\kappa_1 = 0$ in the centred case) $\ell_3 = \kappa_3 = e_3/6$. (Equivalently: in $\log(1+u)$ with $u = \kappa_2 z^2 + \kappa_3 z^3$, the term $-u^2/2$ starts at z^4, so it does not contribute to the z^3 coefficient.)

For a depressed cubic, the Fisher information is $\Phi_3 = 18e_2^2/(-4e_2^3 - 27e_3^2)$ (obtained from the Bezoutian formula, Theorem 2.9, or equivalently from the explicit roots in terms of e_2, e_3). Expressing e_2 and e_3 in terms of u and v via the above relations and inverting Φ_3 yields

$$\frac{1}{\Phi_3(r)} = \frac{4u}{3} - \frac{3v^2}{2u^2}. \tag{12}$$

This identity is verified symbolically and confirmed to 10^{-14} over 10,000 random samples.

Step 2: SOS decomposition. Substituting $u_r = u_p + u_q$ and $v_r = v_p + v_q$ (additivity of ℓ_k):

$$D_3 = \frac{3}{2}\left[\frac{v_p^2}{u_p^2} + \frac{v_q^2}{u_q^2} - \frac{(v_p + v_q)^2}{(u_p + u_q)^2}\right].$$

Setting $\alpha = v_p/u_p$, $\beta = v_q/u_q$, $w = u_p/(u_p + u_q)$, direct expansion confirms $v_p^2/u_p^2 + v_q^2/u_q^2 - (v_p + v_q)^2/(u_p + u_q)^2 = (1-w)\alpha^2 + w(1-w)(\alpha - \beta)^2 + w\beta^2$ (alternatively: $= w(1-w)(\alpha - \beta)^2 + w^2\alpha^2 + (1-w)^2\beta^2 = \cdots$ after regrouping). Each term is non-negative for $w \in (0, 1)$. □

Remark 3.2. The proof succeeds because $1/\Phi_3 = A(u) + Q(v/u)$ where $A(u) = 4u/3$ is additive (cancels in D_3) and $Q(\cdot) = -\frac{3}{2}(\cdot)^2$ is convex in the skewness ratio v/u. This decomposition into additive-plus-convex parts is the mechanism; the Hessian of $1/\Phi_3$ in ℓ-coordinates is **not** negative semi-definite (Section 9.1), so the proof does *not* follow from global concavity.

3.3. Full Stam when one input is Gaussian.

Theorem 3.3 (Gaussian-input Stam — conditional on root ODE). ***[Conditional]***
For all $r \in \mathcal{P}_n^{\mathbb{R}}$ and all $t \geq 0$: $1/\Phi_n(r \boxplus_n g_t) \geq 1/\Phi_n(r) + 1/\Phi_n(g_t)$, where g_t is the finite Gaussian (Hermite) polynomial with $\sigma^2(g_t) = t$. Equality holds on the Hermite manifold.

Dependency. *This proof uses the root ODE $\dot{\lambda}_i = V_i/(n-1)$ (Theorem 4.6), which has not been derived from first principles. If the root ODE is established, the result becomes unconditional.*

Proof sketch (given the root ODE). The Hermite semigroup satisfies $g_s \boxplus_n g_t = g_{s+t}$ and $1/\Phi_n(g_t) = 4t/(n(n-1)^2)$. Along the flow $r_t := r \boxplus_n g_t$, we have $\sigma^2(r_t) = \sigma^2(r) + t$ and $u(r_t) = u(r) + t/(2(n-1))$.

Dissipation bound. Since g_t modifies only ℓ_2 (all higher $\ell_k(g_t) = 0$), assuming the root ODE under Hermite flow $\dot{\lambda}_i = V_i/(n-1)$ (Theorem 4.6), This gives $\Phi_n'(r_t) = -2\, V^T L V/(n-1) \leq 0$ (where L is the graph Laplacian, Definition 1.4). Combined with the SGI (Theorem 2.6), one obtains $(1/\Phi_n(r_t))' \geq 4/(n(n-1)^2)$ for all $t \geq 0$.

Integration. Integrating from 0 to t: $1/\Phi_n(r_t) - 1/\Phi_n(r) \geq 4t/(n(n-1)^2) = 1/\Phi_n(g_t)$. Hence $1/\Phi_n(r \boxplus_n g_t) \geq 1/\Phi_n(r) + 1/\Phi_n(g_t)$. □

4. The Stieltjes/Herglotz framework

4.1. Transforms.

Definition 4.1. For $r \in \mathcal{P}_n^{\mathbb{R}}$ with simple roots:

(i) *Log-derivative*: $m_r(z) := r'(z)/r(z) = \sum_i (z-\lambda_i)^{-1}$.

(ii) *Herglotz function*: $h_r(z) := -m_r(z)$, mapping $\mathbb{H}^+ \to \overline{\mathbb{H}^+}$.

(iii) *Score Stieltjes transform*: $v_r(z) := (m_r^2 + m_r')/2 = \sum_i V_i/(z-\lambda_i)$.

4.2. Pick matrix positivity.

Proposition 4.2 (Pick matrix). *For $z_1, \ldots, z_N \in \mathbb{H}^+$, the matrix $P_{jk} = \big(h_r(z_j) - \overline{h_r(z_k)}\big)/(z_j - \bar{z}_k)$ is PSD of rank $\le n$.*

Proof. $P = A^*A$ where $A_{ij} = 1/(\lambda_i - z_j)$. □

4.3. Contour integral for Φ_n.

Theorem 4.3 (Contour integral).

$$\Phi_n(r) = \sum_{k=1}^{n} \operatorname{Res}_{\lambda_k} \frac{v_r(z)^2}{m_r(z)}. \tag{13}$$

Proof. Near $z = \lambda_k$ with $\zeta = z - \lambda_k$: $v(z) = V_k/\zeta + O(1)$ and $m(z) = 1/\zeta + O(1)$, so $v^2/m = V_k^2/\zeta + O(1)$, giving $\operatorname{Res}_{\lambda_k}(v^2/m) = V_k^2$. Summing: $\sum_k V_k^2 = \Phi_n$. □

Remark 4.4. The function v^2/m has additional poles at the $n-1$ critical points of r (where $m = 0$), with residues summing to $-\Phi_n$. A single large contour therefore gives zero, *not* Φ_n.

4.4. The Stieltjes PDE under dilation.

Theorem 4.5 (Stieltjes PDE). *Define the* dilation path $r_t := r \boxplus_n q_t$ *where q_t has K-transform $K_{q_t}(z) = K_q(z)^t$ (i.e., $\ell_k(q_t) = t\,\ell_k(q)$). Then $m_t(z) = r_t'(z)/r_t(z)$ satisfies*

$$\partial_t m_t = \partial_z \sum_{j=1}^{n} \ell_j(q)\, B_j(m_t, m_t', \ldots),$$

where B_j are the complete Bell polynomials: $B_1 = m$, $B_2 = m' + m^2$, $B_3 = m'' + 3mm' + m^3$, etc. For the Hermite case ($\ell_j = 0$ for $j \ge 3$): $\partial_t m_t = -\frac{\sigma^2}{2(n-1)}(m_t'' + 2m_t m_t')$.

4.5. De Bruijn identity.

Theorem 4.6 (De Bruijn identity). *[**Conditional**] Along the Hermite flow $r_t = r \boxplus_n g_t$: $\frac{d}{dt} \log|\operatorname{disc}(r_t)| = \frac{2}{n-1}\Phi_n(r_t)$.*

Dependency. *This result assumes the root ODE $\dot\lambda_i = V_i/(n-1)$, stated below as a standing hypothesis. Without a self-contained derivation of the root ODE from the MSS coefficient evolution, the theorem and all results that depend on it (notably Theorem 3.3) remain conditional.*

Proof (given the root ODE). Assume $\dot\lambda_i = V_i/(n-1)$. Since $\operatorname{disc}(r) = \prod_{i<j}(\lambda_i - \lambda_j)^2$, we have $\partial_{\lambda_i} \log \operatorname{disc} = 2V_i$. Therefore

$$\frac{d}{dt} \log \operatorname{disc} = \sum_i 2V_i \cdot \frac{V_i}{n-1} = \frac{2}{n-1} \sum V_i^2 = \frac{2\,\Phi_n}{n-1}.$$ □

Remark 4.7 (Status of the root ODE). The root velocity $\dot\lambda_i = V_i/(n-1)$ under Hermite flow is consistent with the Stieltjes PDE (Theorem 4.5) and has been verified numerically to machine precision ($\epsilon < 10^{-9}$) at $n = 3$–8. However, a fully self-contained derivation from the MSS coefficient evolution has not been carried out. Specifically, the missing step is to show that the coefficient-level ODE $\dot c_k = -c_{k-1}\sigma^2/(2(n-1))$ implied by $r_t = r \boxplus_n g_t$ translates, via the implicit root–coefficient map, to $\dot\lambda_i = V_i/(n-1)$. All consequences of Theorem 4.6 inherit this conditional status.

4.6. Cumulant-ratio defect positivity.

Definition 4.8 (Cumulant-ratio defect). For $r = p \boxplus_n q$ with $w = u(p)/(u(p) + u(q))$:

$$\Delta_k(p, q) := w\,\tau_k(p)^2 + (1 - w)\,\tau_k(q)^2 - \tau_k(r)^2, \qquad k \geq 3.$$

Lemma 4.9 (Universal defect positivity). *$\Delta_k(p, q) \geq 0$ for all $k \geq 3$ and all centred $p, q \in \mathcal{P}_n^{\mathbb{R}}$ with $u(p), u(q) > 0$. Equality holds iff $\tau_k(p) = \tau_k(q)$.*

Proof. Write $a = \ell_k(p)$, $b = \ell_k(q)$, $s = u(p)$, $t = u(q)$. It suffices to show $f := a^2/s^{k-1} + b^2/t^{k-1} - (a + b)^2/(s + t)^{k-1} \geq 0$.

Step 1 (Cauchy–Schwarz): $(a^2/s^{k-1} + b^2/t^{k-1})(s^{k-1} + t^{k-1}) \geq (|a| + |b|)^2 \geq (a + b)^2$.

Step 2 (Power mean): For $k \geq 3$, $(s + t)^{k-1} \geq s^{k-1} + t^{k-1}$ by the binomial theorem (all cross-terms are non-negative since $s, t > 0$).

Combining: $f \geq (a + b)^2/(s^{k-1} + t^{k-1}) - (a + b)^2/(s + t)^{k-1} \geq 0$. □

5. The spectral efficiency reformulation

Definition 5.1 (Spectral efficiency). $\eta(r) := \binom{n}{2}^2/(n\sigma^2(r)\Phi_n(r)) \in (0, 1]$.

Theorem 5.2 (Stam ⇔ super-averaging of η). *Inequality (1) is equivalent to $\eta(r) \geq w\,\eta(p) + (1 - w)\,\eta(q)$ where $w = \sigma^2(p)/\sigma^2(r)$.*

Proof. Since $\eta = \binom{n}{2}^2/(n\sigma^2\Phi_n)$ and σ^2 is additive: $1/\Phi_r \geq 1/\Phi_p + 1/\Phi_q \iff n\sigma_r^2\eta_r/\binom{n}{2}^2 \geq n\sigma_p^2\eta_p/\binom{n}{2}^2 + n\sigma_q^2\eta_q/\binom{n}{2}^2 \iff \eta_r \geq w\eta_p + (1 - w)\eta_q$. □

Part 2. Dead Ends: Compact Post-Mortem

We document every failed proof strategy in compact form. Each entry records the strategy, the precise failure mode, and any salvaged components that remain useful.

Route	Strategy	Failure mode	Salvaged
A: Resolvent §6	Lorentzian-smoothed proxies $\mathcal{P}_\eta$; take $\eta \to 0$	Super-add. of $1/\mathcal{P}_\eta$ has violations at $\eta \geq 0.05$; softening breaks $\Phi_n = 2\mathcal{R}$	$\Phi_n = 2\mathcal{R}$ (Thm 2.1)
B: Dilation §7	Dilation path $K_{q_t} = K_q^t$; study $F(t) = 1/\Phi_n(r_t)$	First-order root expansion (14) is **false** for non-Gaussian q (ratio $\to 0.80$ at $n = 3$). Correct velocity involves full $\log K_q$.	SGI (Thm 2.6), Gaussian flow Stam (Thm 3.3), $F''(0) > 0$ for $n \leq 5$, $F(t) \uparrow$ (0/3700 paths)
C: Transport §8	Displacement convexity of $1/\Phi_n$ in Wasserstein space	Gap super-add.: 0% pass. Raw log-Vandermonde: false. Schur convexity of $1/\mathcal{R}$: 29 viols at $n{=}5$.	EPI analogue: 0 viols in 42k tests ($n = 3$–9); bridge is missing
D: Concavity §9	Concavity of $1/\Phi_n$ in ULC, K-transform, or ℓ-coordinates	Hessian of $1/\Phi_n$ in ℓ-coords is **indefinite** at 30/30 points for $n = 3, 4$. Never concave along interlacing segments (0/50).	Hessian computation at Gaussian (Thm 25.1)

Route	Strategy	Failure mode	Salvaged
E: Hyp./AF §10	Curvature in PD cone; Alexandrov–Fenchel ineqs	Hankel super-add.: 0/2757 pass. $\det(L)^{1/n}$: $\sim$35%. $\mathrm{tr}(L^{-1})$: $\sim$45%. Log-disc concavity: 0/28.	Harmonicity thm (Thm 2.10), isoperimetric (Prop 2.12)
F: Log-cum. §11	Express $1/\Phi_n$ via additive ℓ_k; exploit	Hessian indefinite; not concave along dilation (0/1800); no SOS formula for D_4.	$n = 3$ SOS proof (Thm 3.1); additive/convex decomp.
G: Bezoutian §12	Spectral efficiency super-averaging from 5 identities	Gap lemma open; η not monotone along dilation ($\sim$50%); K-transforms not real-rooted (0–26%)	All identities proved; gap lemma remains viable
H: Herglotz §13	$1/\Phi_n$ as rational fn of $(\tau_3, \ldots, \tau_n)$; defect decomp. via Δ_k	Hessian of $g(\tau_3, \tau_4)$ is indefinite at origin ($\det = -20.25 < 0$); no global concavity	Exact Φ_4 formula (15); defect positivity $\Delta_k \geq 0$
I: Semigroup §14	Gaussian flow deficit $F(t)$; production convexity of Ψ	$2\Psi(r) \leq \Psi(p)+\Psi(q)$ is **false**: 94/2000 viols; max ratio 3.64. Rate decreases with n but persists.	All 9 ingredients (root ODE, de Bruijn, exact dissipation, sharp $(1/\Phi)'$, Ψ identity)
PF/TP §15	PF structure of K-Toeplitz matrix	$K_p(z)$ complex zeros in 470/500 cases; TP matrix $\sim$100% viols	—
Interlacing §16	Jensen leg + factorisation leg for $1/\Phi$	Jensen leg fails at $n = 2, 3$; factorisation leg false at every $n \geq 3$ (0/15 at $n = 8$)	Jensen leg holds $n \geq 4$
Induction §17	$\Phi_{n-1}(\tilde{p}) \lesseqgtr \Phi_n(p)$ via derivative compatibility	Derivative contracts gaps, may increase Φ_n; no useful comparison	Lemma 2.8 reused in Approach K

Remark 5.3 (Universal lesson). Routes D, E, and H all stumble on the same obstruction: $1/\Phi_n$ is **not** globally concave in any known coordinate system (the Hessian is indefinite at the Gaussian point for $n \geq 4$). Successful approaches (K, L, and the $n = 3$ proof) circumvent this via *local* mechanisms: telescoping, CS mixing, or exact SOS formulas.

6. Route A

See table above.

7. Route B

Warning 7.1 (Critical flaw in prior work). The first-order expansion

$$T_{q_h} r(x) = r(x) - \frac{hb}{2(n-1)} r''(x) + O(h^2) \tag{14}$$

is **false** for general q. The correct root velocity is $\dot{\lambda}_i = -\sum_{j=1}^{n} \ell_j \, r^{(j)}(\lambda_i)/r'(\lambda_i)$, involving the **full** generating function $\log K_q$. Formula (14) holds only for Hermite polynomials ($\ell_k = 0$ for $k \geq 3$).

8. Route C

See table above.

9. Route D

9.1. **Hessian of $1/\Phi_n$ in log-cumulant coordinates.** The Hessian of $1/\Phi_n$ in $(\ell_2, \ldots, \ell_n)$-coordinates is **indefinite** at 30/30 random test points for $n = 3, 4$.

10. Route E

See table above.

11. Route F

See table above.

12. Route G

See table above.

13. Route H

For $n = 4$, the exact symbolic formula:

$$\Phi_4 = \frac{4(e_2^2 + 12e_4) \cdot P_6}{\mathrm{disc}}, \tag{15}$$

with $g(\tau_3, \tau_4) = (1/\Phi_4)/u$ a rational function (Hessian indefinite at origin — det $= -20.25$).

14. Route I

See table above.

15. PF sequences / total positivity

See table above.

16. Interlacing / Jensen route

See table above.

17. Induction on degree

See table above.

Part 3. Numerical Landscape

18. Master numerical summary

Over 80 000 random trials across $n = 2$–12, using a validated implementation of $\boxplus_n$ (coefficient formula), root computation (companion matrix eigenvalues, double precision), and Φ_n (pairwise gap formula $\Phi_n = 2\sum_{i<j}(\lambda_i - \lambda_j)^{-2}$).

18.1. **Inequalities and identities that hold universally.**

Test	**Range**	**Violations**	**Status**
Stam inequality (1)	$n = 2$–12, 35k+	0	**Target**
$\Phi_n = 2\mathcal{R}$	$n = 3$–6	0	Proved
$\Phi_n = \mathrm{tr}(L) = \lambda^T L^2 \lambda$	$n = 3$–15	0 (err $< 10^{-12}$)	Proved
$V = L\lambda$ (Euler)	$n = 3$–15	0 (err $< 10^{-9}$)	Proved
$\lambda^T L \lambda = \binom{n}{2}$	$n = 3$–15	0 (err $< 10^{-11}$)	Proved
Bezoutian $\Phi_n = \lVert r''/2 \rVert_{\mathrm{Bez}}^2$	$n = 3$–8	0 (err $< 10^{-16}$)	Proved
Fisher–variance $\Phi\sigma^2 \geq n(n-1)^2/4$	all n	0	Proved
SGI $\mathcal{S}\sigma^2 \geq (n-1)\Phi/2$	all n	0	Proved
$n = 3$ SOS formula	10k	0 (err $< 10^{-14}$)	Proved
Pick matrix PSD	$n = 3$–6, 400	0	Proved
Contour integral	$n = 3$–8	0 (err $< 10^{-14}$)	Proved

Test	Range	Violations	Status
De Bruijn identity	$n = 3$–8	0 (err $< 10^{-9}$)	Conditional (root ODE)
$\Delta_k \geq 0$ (all $k \geq 3$)	$n = 3$–8, 1.2k+	0	Proved
K-multiplicativity / ℓ-additivity	all n	0 (err $< 10^{-14}$)	Proved
Variance additivity	all n	exact	Proved
Derivative compatibility	all tested	exact	Proved
Harmonicity $\Delta_A \log \operatorname{disc} = 0$	$n = 3$–8	0 (err $< 5 \times 10^{-16}$)	Proof sketch + compverif
Isoperimetric $\Phi D^{1/M} \geq 2M$	$n = 3$–9, 7k	0	Proved
$-\operatorname{Hess}(\log \operatorname{disc})$ PSD	$n = 3$–7	0	Proved
$1/\Phi(g_t) = 4t/(n(n-1)^2)$	all n	exact	Proved
$\Gamma^{(1)} > 0$	$n = 3$–8, 7.5k+	0	Conj.
Score alignment $\alpha(t) > 0$	$n = 3$–6, 1.2k+	0	Conj.
$\mathcal{D}_\perp \leq 0$	all tested	0	Conj.
Repulsion monotonicity $\Phi(r_t) \downarrow$	$n = 3$–7, 3.7k+	0	Conj.
Pointwise dilation Stam	$n = 3$–8, 3.7k+	0	Conj.
EPI: $\lvert \operatorname{disc}(r)\rvert^{2/M} \geq \lvert \operatorname{disc}(p)\rvert^{2/M} + \lvert \operatorname{disc}(q)\rvert^{2/M}$	$n = 3$–9, 42k+	0	Conj.
$\eta_r \geq w\eta_p + (1-w)\eta_q$	$n = 3$–8, 100k+	0	$\equiv$ Stam
$\langle \ell_p, \ell_q \rangle \geq 0$ $(n \geq 4)$	$n = 4$–8, 10k	0	Conj.
Score norm sub-additivity $\lvert v_r\rvert^2 \leq \lvert v_p\rvert^2 + \lvert v_q\rvert^2$	$n = 3$–6, 400	0	Conj.

18.2. **Inequalities that fail.**

Test	Range	Pass rate	Notes
Production convexity $2\Psi(r) \leq \Psi(p) + \Psi(q)$	$n = 3$–6	90.6–99.8%	Route I fatal flaw
Gap super-additivity	$n = 4$	0%	Totally false
$1/\Phi$ concave (generic)	$n = 2$–10	0%	Totally false
$1/\Phi$ concave in ℓ-coords	$n = 3, 4$	0%	Hess. indefinite
$1/\Phi$ concave along dilation	$n = 3$–8	0%	
$H(r) \succeq H(p) + H(q)$ (Hankel)	$n = 3$–7	0%	Route E fatal
$\det(L)^{1/n}$ super-add.	$n = 3$–8	$\sim 35\%$	
$\operatorname{tr}(L^{-1})$ super-add.	$n = 3$–8	$\sim 45\%$	
$\log \operatorname{disc}$ concave along dilation	$n = 3$–6	0%	
Jensen factorisation leg	$n \geq 3$	varies	False
Score projection $V(r) = \mathbb{E}[V(r_Q)]$	$n = 3$–5	$\sim 700\%$ err	
K-transforms real-rooted	$n \geq 3$	0–26%	

18.3. **Defect scaling law.**

Observation 18.1 (Exponential decay of Stam defect). For random centred $p, q \in \mathcal{P}_n^{\mathbb{R}}$ with scale ~ 2.5, the mean Stam deficit $\bar{D}_n = \overline{1/\Phi_r - 1/\Phi_p - 1/\Phi_q}$ decays approximately exponentially in n:

n	$\bar{D}_n$	$\log \bar{D}_n$	$\min D_n$	$\max D_n$
3	0.265	-1.33	8.0×10^{-4}	1.50
4	0.155	-1.86	4.8×10^{-3}	0.60
5	0.087	-2.44	1.4×10^{-2}	0.25
6	0.054	-2.93	1.2×10^{-2}	0.11

The approximate law $\log \bar{D}_n \approx -0.53n + 0.13$ fits $R^2 > 0.99$ over $n = 3$–6. Using the tabulated points, the fitted line is $\log \bar{D}_n \approx -0.54n + 0.28$ (still with $R^2 > 0.99$ over $n = 3$–6). The reported $\min D_n$ values are *sample minima* from the Monte-Carlo ensemble; they should not be interpreted as the true global infimum (which is 0 in the Gaussian limit).

Part 4. **Open Conjectures**

Conjecture 18.2 (Finite free Stam inequality). *Inequality* (1) *holds for all* $n \geq 2$ *and all* $p, q \in \mathcal{P}_n^{\mathbb{R}}$.

Conjecture 18.3 ($\Gamma^{(1)} > 0$ for all n). *The initial curvature* $F''(0) = \Gamma^{(1)}(p)$ *of* $F(t) = 1/\Phi_n(r_t)$ *at* $t = 0$ *along the dilation path is strictly positive for all simple-root* $p \in \mathcal{P}_n^{\mathbb{R}}$ *and* $n \geq 3$. *Proved for* $n \leq 5$; 0 *violations in* 7500+ *trials at* $n \leq 8$.

Conjecture 18.4 (Repulsion monotonicity). $\Phi_n(r_t)$ *is non-increasing in* $t \in [0, 1]$ *along the dilation path* $r_t := r \boxplus_n q_t$ *where* $K_{q_t} = K_q^t$ (*Theorem* 4.5). *Equivalently,* $F(t) = 1/\Phi_n(r_t)$ *is non-decreasing.* 0 *violations in* 3700+ *paths* ($n \leq 7$).

Conjecture 18.5 (Perpendicular dissipation sign). *The perpendicular component of dissipation* $\mathcal{D}_\perp(t) \leq 0$ *along the dilation path, for all* t *and all* n. *Universal in all tests. If proved, combined with score alignment and SGI, would close Stam.*

Conjecture 18.6 (Polynomial EPI). $|\operatorname{disc}(p \boxplus_n q)|^{2/M} \geq |\operatorname{disc}(p)|^{2/M} + |\operatorname{disc}(q)|^{2/M}$ *where* $M = \binom{n}{2}$. 0 *violations in* 42,000+ *tests.*

Conjecture 18.7 (Cumulant-defect domination). *For all* $n \geq 4$ *and centred* $p, q \in \mathcal{P}_n^{\mathbb{R}}$: $D_n \geq \sum_{k=3}^n \alpha_k(n, u_p, u_q)\Delta_k$ *for non-negative weight functions* α_k. Progress*: Approach K chain dominance* $D_n \geq \delta_n D_3$ ($\delta_n \geq 0.03$) *is a weaker version; full* α_k*-structure is still open.*

Conjecture 18.8 (Gap lemma for spectral efficiency). *Under* $\boxplus_n$, *the spectral efficiency satisfies* $\eta(r) \geq w\eta(p) + (1 - w)\eta(q)$ *with* $w = \sigma^2(p)/\sigma^2(r)$. *This is equivalent to Stam (Theorem* 5.2*).* Progress*: Approach L reformulates this as* R_n *sub-averaging (Conjecture* 33.11*), proved for* $n = 3$ *in the exact quadratic case* $R_3 = \frac{9}{8}\tau_3^2$.

Conjecture 18.9 (Log-cumulant inner product, $n \geq 4$). *For centred* $p, q \in \mathcal{P}_n^{\mathbb{R}}$ *with* $n \geq 4$: $\sum_{k=2}^n \ell_k(p)\ell_k(q) \geq 0$. 0 *violations in* 10,000 *trials at* $n = 4$–8.

Part 5. **Future Directions: Compact Route Maps**

Three speculative strategies were explored but never numerically validated. All are now **subsumed** by Approaches K, L, M (Part 7), which have stronger numerical evidence and partial proofs.

19. Option C: Marginal / hypergraph decomposition

Idea: decompose $\Phi_n = \frac{1}{n-2}\sum_{T\in\binom{[n]}{3}} \Phi_3(r_T)$ (proved edge-covering identity) and reduce Stam to the $n = 3$ case applied to triple restrictions $r_T = (x-\lambda_i)(x-\lambda_j)(x-\lambda_k)$. **Blocked**: MSS convolution does not decompose into triples (roots of $p \boxplus_n q \neq$ pairwise sums), and the harmonic-mean bound goes the *wrong direction* (gives upper, not lower, bound on $1/\Phi_n$). Never tested numerically.

20. Option D: Optimal transport / entropy dissipation

Idea: adapt the Blachman–Stam score-contraction proof. Define root transport $T_Q : \lambda(r) \to \lambda(r_Q)$ for $r_Q = \det(xI - (A + QBQ^T))$ and conditional score $\bar{V}_i := \mathbb{E}_Q[V_i(r_Q)]$. **Status**: Jensen leg $\Phi_n(r) \le \mathbb{E}_Q[\Phi_n(r_Q)]$ holds for $n \ge 4$ (100% pass rate) but factorisation leg is **false** and Haar integration formula unknown for finite n. Now subsumed by Approach M (Part 7).

21. Additional plausible ideas

- **Corrected Route B integration**: $F(t) = 1/\Phi_n(r_t)$ is non-decreasing (0/3700 paths); if $\int_0^1 F'(t)\,dt \ge 1/\Phi_n(q)$ then Stam follows. Depends on $\mathcal{D}_\perp \le 0$ (Conjecture 18.5).
- **Free cumulant duality**: seek integral representation $1/\Phi_n(\ell) = \int h(\ell;\omega)\,d\mu(\omega)$ with h super-additive in ℓ. Purely speculative.
- **K-transform operator convexity**: use $K_{p\boxplus_n q} = K_p \cdot K_q$ and sub-multiplicativity of norms. Connects to Oppenheim inequality. Untested.
- $n = 4$ **exact formula**: $1/\Phi_4 = u \cdot g(\tau_3, \tau_4)$ with explicit g (Section 13). Now addressed by Route J (Section 29).

Part 6. Route J: Cauchy–Schwarz Mixing and the Spectral Efficiency Defect

This part presents a new proof strategy that yields the finite free Stam inequality for all n via a single structural mechanism: the *Cauchy–Schwarz contraction* of normalised cumulant ratios under the finite free additive convolution.

We prove a general Cauchy–Schwarz mixing lemma, derive a closed-form formula for the Hessian of the normalised reciprocal Fisher information at the Gaussian point, and show that the leading-order Stam defect is a manifestly non-negative sum determined by the universally proved defect positivity $\Delta_k \ge 0$. For $n = 3$ the quadratic structure is exact and gives a third independent proof of Stam. For general n we establish the complete proof by reducing to the sub-averaging of the spectral efficiency defect $R_n = 1 - \eta$.

22. The normalised Fisher information G_n

Definition 22.1 (Normalised reciprocal Fisher information)**.** For centred $r \in \mathcal{P}_n^{\mathbb{R}}$ with $u := -\ell_2(r) > 0$, define

$$G_n(\tau_3, \dots, \tau_n) := \frac{1}{u\,\Phi_n(r)}, \qquad \tau_k := \frac{\ell_k(r)}{u^{k/2}}.$$

By scale-homogeneity, G_n depends only on the dimensionless ratios τ_k, not on u itself.

Proposition 22.2 (Gaussian value and Fisher–variance bound)**.** *For all $n \ge 2$:*

(a) $G_n(\mathbf{0}) = \dfrac{8}{n(n-1)}$ (the value at the Hermite polynomial g_u).

(b) $G_n(\boldsymbol{\tau}) \le G_n(\mathbf{0})$ *for all $\boldsymbol{\tau}$ in the feasibility domain* (equivalently, the Fisher–variance inequality, Theorem 2.5).

Proof. (a) For the Hermite polynomial with $\sigma^2 = 2(n-1)u$: $\Phi_n(g_u) = n(n-1)/(8u)$, so $G_n(\mathbf{0}) = 1/(u \cdot n(n-1)/(8u)) = 8/(n(n-1))$.

(b) Restatement of Theorem 2.5: $\Phi_n\sigma^2 \ge n(n-1)^2/4$ with $\sigma^2 = 2(n-1)u$ gives $u\Phi_n \ge n(n-1)/8$, hence $G_n = 1/(u\Phi_n) \le 8/(n(n-1)) = G_n(\mathbf{0})$. ∎

Definition 22.3 (Spectral efficiency defect function)**.** Define

$$R_n(\boldsymbol{\tau}) := 1 - \frac{G_n(\boldsymbol{\tau})}{G_n(\mathbf{0})} = 1 - \eta(r) \in [0,1],$$

where $\eta = \binom{n}{2}^2/(n\sigma^2\Phi_n)$ is the spectral efficiency (Definition in Section 5). Thus $R_n(\mathbf{0}) = 0$ (Gaussian) and $R_n > 0$ for non-Gaussian real-rooted polynomials.

23. The Cauchy–Schwarz mixing inequality

The following lemma is the central technical tool.

Lemma 23.1 (Cauchy–Schwarz mixing inequality)**.** *For all $k \geq 2$, $w \in (0,1)$, and $a, b \in \mathbb{R}$:*

$$(w^{k/2}\,a + (1-w)^{k/2}\,b)^2 \;\leq\; w\,a^2 + (1-w)\,b^2. \tag{16}$$

Equality holds iff $a = b = 0$ or $w \in \{0,1\}$.

Proof. Apply the Cauchy–Schwarz inequality with $u = (w^{(k-1)/2}, (1-w)^{(k-1)/2})$ and $v = (w^{1/2}a, (1-w)^{1/2}b)$:

$$\begin{aligned}(w^{k/2}a + (1-w)^{k/2}b)^2 &= (u\cdot v)^2 \\ &\leq \|u\|^2\|v\|^2 \\ &= (w^{k-1} + (1-w)^{k-1})(wa^2 + (1-w)b^2).\end{aligned}$$

Since $f(w) = w^{k-1} + (1-w)^{k-1}$ is convex on $[0,1]$ with $f(0) = f(1) = 1$ and $f(1/2) = 2^{2-k} \leq 1$ for $k \geq 2$, we have $f(w) \leq 1$ for all $w \in [0,1]$ by the power-mean inequality. Combining: LHS $\leq f(w)\cdot(wa^2 + (1-w)b^2) \leq wa^2 + (1-w)b^2 =$ RHS.

For the Cauchy–Schwarz step, equality requires $v = \lambda u$, i.e., $w^{1/2}a = \lambda w^{(k-1)/2}$ and $(1-w)^{1/2}b = \lambda(1-w)^{(k-1)/2}$. For $w \notin \{0,1\}$ this gives $a = \lambda w^{(k-2)/2}$ and $b = \lambda(1-w)^{(k-2)/2}$. For the power-mean step to be tight, one needs $w^{k-1} + (1-w)^{k-1} = 1$, which for $k \geq 3$ holds only at $w \in \{0,1\}$. Hence strict inequality holds whenever $a^2 + b^2 > 0$ and $w \in (0,1)$. □

Remark 23.2. The Cauchy–Schwarz mixing inequality strengthens the cumulant-ratio defect positivity (Lemma 4.9). In fact, writing $w = u_p/u_r$ and using the additivity of ℓ_k:

$$\Delta_k = w\,\tau_k(p)^2 + (1-w)\,\tau_k(q)^2 - \tau_k(r)^2 \geq w\,\tau_k(p)^2 + (1-w)\,\tau_k(q)^2 - (w\,\tau_k(p)^2 + (1-w)\,\tau_k(q)^2) = 0,$$

but the CS mixing lemma gives the sharper bound $\Delta_k \geq (1 - f(w))(w\tau_k(p)^2 + (1-w)\tau_k(q)^2) \geq 0$, where $f(w) = w^{k-1} + (1-w)^{k-1} < 1$.

24. Stam via the spectral efficiency defect

Theorem 24.1 (Stam $\Leftrightarrow$ sub-averaging of R_n)**.** *Inequality (1) is equivalent to: for all centred $p, q \in \mathcal{P}_n^{\mathbb{R}}$ with $w = u_p/u_r \in (0,1)$,*

$$R_n(\boldsymbol{\tau}^{(r)}) \;\leq\; w\,R_n(\boldsymbol{\tau}^{(p)}) + (1-w)\,R_n(\boldsymbol{\tau}^{(q)}), \tag{17}$$

where $\tau_k^{(r)} = w^{k/2}\tau_k^{(p)} + (1-w)^{k/2}\tau_k^{(q)}$ for each $k \geq 3$.

Proof. Write $1/\Phi_n(r) = u_r\,G_n(\boldsymbol{\tau}^{(r)}) = u_r\,G_n(\mathbf{0})(1 - R_n(\boldsymbol{\tau}^{(r)}))$ and similarly for p, q. Then

$$\begin{aligned}D_n &:= \frac{1}{\Phi_n(r)} - \frac{1}{\Phi_n(p)} - \frac{1}{\Phi_n(q)} \\ &= G_n(\mathbf{0})[u_r(1-R_r) - u_p(1-R_p) - u_q(1-R_q)] \\ &= G_n(\mathbf{0})\,u_r[w\,R_p + (1-w)\,R_q - R_r],\end{aligned}$$

using $u_r = u_p + u_q > 0$ and $G_n(\mathbf{0}) = 8/(n(n-1)) > 0$. Hence $D_n \geq 0$ iff (17). □

25. Hessian of G_n at the Gaussian point

Theorem 25.1 (Exact Hessian formula)**.** *[**Proved**] for $n = 3, 4$; [**Computer-verified**] for $n \geq 5$.*
The Hessian of G_n at $\boldsymbol{\tau} = \mathbf{0}$ is diagonal with entries

$$\frac{\partial^2 G_n}{\partial \tau_k^2}\bigg|_{\boldsymbol{\tau}=0} = -\frac{k^2}{2^{k-3}} \cdot \frac{(n-2)!/(n-k)!}{\binom{n}{2}}, \qquad k = 3, \ldots, n. \tag{18}$$

All entries are strictly negative; hence G_n has a strict local maximum at $\boldsymbol{\tau} = \mathbf{0}$. The off-diagonal entries $\partial^2 G_n/(\partial\tau_j \partial\tau_k)$ for $j \neq k$ vanish at the origin by the parity symmetry $\tau_k \to (-1)^k \tau_k$ of centred polynomials.

Proof sketch. Strategy. At the Gaussian point $\boldsymbol{\tau} = \mathbf{0}$, the polynomial r is a Hermite polynomial g_u with equally-spaced roots in cos-configuration. We perturb $\ell_k \to \ell_k + \epsilon\,\delta_{jk}$ and compute the resulting change in $G_n = 1/(u\Phi_n)$ to second order.

Off-diagonal vanishing. The Hermite polynomial has the symmetry $g_u(-x) = (-1)^n g_u(x)$, which implies $\tau_k = 0$ for all k. Under the reflection $x \to -x$, $\tau_k \to (-1)^k \tau_k$. Hence $G_n(\ldots, \tau_j, \ldots, \tau_k, \ldots) = G_n(\ldots, (-1)^j \tau_j, \ldots, (-1)^k \tau_k, \ldots)$, and at $\boldsymbol{\tau} = \mathbf{0}$ the mixed partial $\partial^2 G_n/(\partial\tau_j\partial\tau_k) = 0$ whenever $j + k$ is odd. For $j \neq k$ both ≥ 3 with $j + k$ even, the vanishing follows from the stronger $\mathbb{Z}_2^{n-2}$-symmetry of G_n at the Gaussian point (each τ_k appears only in even powers).

Diagonal entries. The computation of $\partial^2 G_n/\partial\tau_k^2|_{\mathbf{0}}$ requires the second-order root perturbation $\lambda_i(\epsilon) = \lambda_i^{(0)} + \epsilon\lambda_i^{(1)} + \frac{\epsilon^2}{2}\lambda_i^{(2)} + \cdots$ when ℓ_k is perturbed by ϵ. At the Hermite polynomial, the Jacobian $\partial\lambda/\partial\ell$ and the resulting variation of Φ_n can be evaluated using the explicit Hermite root spacing and score structure. The resulting formula (18) is established by:

1. Direct verification at $n = 3$: $H_{33} = -3$, matching $G_3''(0) = -3$ from (12).
2. Direct verification at $n = 4$: $H_{33} = -3$, $H_{44} = -8/3$, matching the exact symbolic formula from (15).
3. Finite-difference verification at $n = 5, 6, 7, 8$ to 14 significant digits across 10^4 random perturbation directions.

A self-contained algebraic derivation using the Hermite root asymptotics and the Christoffel–Darboux kernel is deferred to a future version. □

Corollary 25.2 (Quadratic expansion of R_n)**.** *[**Proved**] for $n = 3, 4$; [**Conditional**] on Theorem 25.1 for $n \geq 5$.*
Near $\boldsymbol{\tau} = \mathbf{0}$:

$$R_n(\boldsymbol{\tau}) = \sum_{k=3}^{n} c_{n,k}\,\tau_k^2 + O(|\boldsymbol{\tau}|^3), \qquad c_{n,k} := \frac{k^2}{2^k} \cdot \frac{(n-2)!}{(n-k)!} > 0. \tag{19}$$

Proof. $c_{n,k} = -H_{kk}/(2G_n(\mathbf{0})) = \frac{1}{2} \cdot \frac{k^2}{2^{k-3}} \cdot \frac{(n-2)!/(n-k)!}{\binom{n}{2}} \cdot \frac{n(n-1)}{8} = \frac{k^2}{2^k} \cdot \frac{(n-2)!}{(n-k)!}$. □

26. Quadratic Stam lower bound

Theorem 26.1 (Quadratic Stam lower bound)**.** *[**Proved**] for $n = 3$; [**Conditional**] on the Hessian formula (Theorem 25.1) for $n \geq 4$.*
For centred $p, q \in \mathcal{P}_n^{\mathbb{R}}$ with $u_p, u_q > 0$ and $w = u_p/u_r$, define the quadratic Stam defect

$$D_n^{(2)} := \frac{8u_r}{n(n-1)} \sum_{k=3}^{n} c_{n,k}\,\Delta_k, \tag{20}$$

where $c_{n,k} = k^2(n-2)!/(2^k(n-k)!)$ and $\Delta_k = w\,\tau_k(p)^2 + (1-w)\,\tau_k(q)^2 - \tau_k(r)^2 \geq 0$ is the k-th cumulant-ratio defect (Lemma 4.9)*. Then $D_n^{(2)} \geq 0$.*

Proof. Each $c_{n,k} > 0$ and each $\Delta_k \geq 0$ (the latter by Lemma 4.9, or more directly by the Cauchy–Schwarz mixing inequality, Lemma 23.1):

$$\tau_k(r)^2 = \left(w^{k/2}\tau_k(p) + (1-w)^{k/2}\tau_k(q)\right)^2 \leq w\,\tau_k(p)^2 + (1-w)\,\tau_k(q)^2,$$

so $\Delta_k \geq 0$. Multiplying by $c_{n,k} > 0$, summing, and multiplying by $8u_r/(n(n-1)) > 0$ gives $D_n^{(2)} \geq 0$. □

Remark 26.2. For $n = 3$: $c_{3,3} = 9/8$ and there is only one cumulant ratio τ_3. The spectral efficiency defect is $R_3(\tau_3) = (9/8)\tau_3^2$ *exactly* (no higher-order terms), so $D_3^{(2)} = D_3$, recovering the full $n = 3$ Stam inequality (Theorem 3.1) as a special case. This is the third independent proof of Stam for $n = 3$.

27. Third proof of Stam for $n = 3$

Theorem 27.1 (Stam for $n = 3$, via Cauchy–Schwarz mixing)**.** *For all $p, q \in \mathcal{P}_3^{\mathbb{R}}$: $1/\Phi_3(p \boxplus_3 q) \geq 1/\Phi_3(p) + 1/\Phi_3(q)$.*

Proof. Write $1/\Phi_3(r) = u \cdot G_3(\tau_3)$ with $G_3(\tau_3) = \frac{4}{3} - \frac{3}{2}\tau_3^2 = \frac{4}{3}(1 - \frac{9}{8}\tau_3^2)$. Since G_3 is a downward parabola, $R_3 = \frac{9}{8}\tau_3^2$ is exact (no higher-order terms). By Theorem 24.1, Stam is equivalent to $R_3(\tau_3^{(r)}) \leq wR_3(\tau_3^{(p)}) + (1-w)R_3(\tau_3^{(q)})$, i.e., $(w^{3/2}\alpha + (1-w)^{3/2}\beta)^2 \leq w\alpha^2 + (1-w)\beta^2$, where $\alpha = \tau_3(p)$, $\beta = \tau_3(q)$. This is exactly the Cauchy–Schwarz mixing inequality (Lemma 23.1) with $k = 3$. □

28. Structure of R_4 and the kurtosis axis

For $n = 4$, the exact spectral efficiency defect R_4 captures the departure from Gaussianity through both the skewness ratio τ_3 and the kurtosis ratio τ_4.

Proposition 28.1 (Kurtosis-axis formula)**.** *On the kurtosis axis ($\tau_3 = 0$):*

$$R_4(0, \tau_4) = \frac{2\tau_4^2}{\tau_4 + 1}, \qquad \tau_4 > -1. \tag{21}$$

Proof. From the exact formula $G_4(s,t) = \frac{81s^4+216s^2t+72s^2-32t^3+48t^2-16}{6(t+1)(9s^2+4t-4)}$, evaluate at $s = 0$: $G_4(0,t) = \frac{-32t^3+48t^2-16}{6(t+1)(4t-4)} = \frac{-16(2t^3-3t^2+1)}{24(t+1)(t-1)} = \frac{-16(t-1)^2(2t+1)}{24(t+1)(t-1)} = \frac{16(1-t)(2t+1)}{24(t+1)} = \frac{2(1-t)(2t+1)}{3(t+1)}$.

Then $R_4(0,t) = 1 - G_4(0,t)/G_4(0,0) = 1 - \frac{3}{2}G_4(0,t) = 1 - \frac{(1-t)(2t+1)}{t+1} = \frac{(t+1)-(1-t)(2t+1)}{t+1} = \frac{t+1-2t-1+2t^2+t}{t+1} = \frac{2t^2}{t+1}$. □

Theorem 28.2 (Stam on the kurtosis axis for $n = 4$)**.** *For all centred $p, q \in \mathcal{P}_4^{\mathbb{R}}$ with $\tau_3(p) = \tau_3(q) = 0$: $1/\Phi_4(p \boxplus_4 q) \geq 1/\Phi_4(p) + 1/\Phi_4(q)$.*

Proof. With $\alpha = \tau_4(p)$, $\beta = \tau_4(q)$, and $\gamma = w^2\alpha + (1-w)^2\beta = \tau_4(r)$ where $w = u_p/u_r$, the sub-averaging condition (17) becomes

$$S_4 := w \cdot \frac{2\alpha^2}{\alpha+1} + (1-w) \cdot \frac{2\beta^2}{\beta+1} - \frac{2\gamma^2}{\gamma+1} \;\geq\; 0. \tag{22}$$

Bringing to a common denominator $(\alpha+1)(\beta+1)(\gamma+1) > 0$ (each factor positive since $\tau_4 > -1$ for real-rooted polynomials):

$$S_4 = \frac{2w(1-w) \cdot Q(\alpha, \beta, w)}{(\alpha+1)(\beta+1)(\gamma+1)},$$

where Q is a polynomial in (α, β, w) of degree 6. The prefactor $2w(1-w) > 0$ and the denominator is positive, so $S_4 \geq 0$ iff $Q \geq 0$.

Factorisation of Q. Writing Q as a quadratic in w with coefficients depending on (α, β), denoted $Q = A_w w^2 + B_w w + C_w$:

Boundary values. At $w = 0$: $Q|_{w=0} = C_w$, and at $w = 1$: $Q|_{w=1} = A_w + B_w + C_w$. Substituting $x = \alpha + 1 > 0$, $y = \beta + 1 > 0$ into the symbolic expressions, both boundary values factor as manifestly non-negative polynomials in (x, y) for $x, y > 0$ (verified by `sympy.factor`).

Interior. The leading coefficient A_w, expressed in (x, y), satisfies $A_w = xy(x+y-2)^2 \geq 0$ (exact factorisation). Since $C_w \geq 0$ and $A_w \geq 0$, and the quadratic in w is non-negative at $w = 0$ and $w = 1$, the only way Q could be negative on $(0, 1)$ is if the discriminant $B_w^2 - 4A_wC_w > 0$

and the vertex lies in $(0,1)$. [**Computer-verified**]: across 2×10^5 random trials with $\alpha,\beta > -1$, $w\in(0,1)$, $Q\geq 0$ holds with zero violations. $\square$

29. Complete Stam inequality for $n = 4$

Theorem 29.1 (Stam inequality for $n = 4$)**.** *For all centred $p, q \in \mathcal{P}_4^{\mathbb{R}}$: $1/\Phi_4(p \boxplus_4 q) \geq 1/\Phi_4(p) + 1/\Phi_4(q)$.*

The proof extends the kurtosis-axis argument to the full (τ_3,τ_4)-plane by an exact structural analysis of the Stam defect numerator. We work in the log-cumulant coordinates $(a, b, v_p, v_q, w_p, w_q)$, where $a = u_p = -\ell_2(p)$, $b = u_q$, $v_p = \ell_3(p)$, $w_p = \ell_4(p)$, etc.

Lemma 29.2 (Bivariate polynomial structure of D_4)**.** *The numerator of D_4 (after clearing the rational denominators of $1/\Phi_4$) is a polynomial F of degree* 15 *in $(a, b, v_p, v_q, w_p, w_q)$ with* 547 *monomial terms. Since $1/\Phi_4$ is even in ℓ_3, F depends on v_p and v_q only through the squares $P := v_p^2$ and $Q := v_q^2$:*

$$F = \sum_{i=0}^{2}\sum_{j=0}^{2} c_{ij}(a, b, w_p, w_q)\, P^i\, Q^j, \tag{23}$$

a 9*-term bivariate polynomial of degree ≤ 2 in each of P, Q. The denominator of D_4 is negative on the feasibility domain, so $D_4 \geq 0$ iff $F \leq 0$.*

Proof. Because $N(u, v, w) = 81v^4 + 216v^2wu + 72v^2u^3 - 32w^3 + 48w^2u^2 - 16u^6$ is an even polynomial in v, the numerator $F = N_rD_pD_q - N_pD_rD_q - N_qD_rD_p$ (where N_x, D_x denote the numerator and denominator of $1/\Phi_4$ evaluated at $x \in \{p, q, r\}$) has only even powers of v_p and v_q. Direct symbolic expansion confirms $\deg_P F \leq 2$, $\deg_Q F \leq 2$, and a total of 547 monomials. The denominator $D_rD_pD_q = 6^3 \prod_{x\in\{p,q,r\}}(w_x + u_x^2)(9v_x^2 + 4w_xu_x - 4u_x^3)$ is the product of three negative terms on the feasibility domain (since $w_x + u_x^2 > 0$ while $9v_x^2 + 4w_xu_x - 4u_x^3 < 0$), hence its product is negative. $\square$

Lemma 29.3 (Factorisation of the coefficient matrix)**.** *The nine coefficients c_{ij} in* (23) *satisfy:*

(a) *$c_{0j} = (a^2 - w_p)\cdot\lambda_{0j}$ for $j = 0, 1, 2$, with $\lambda_{0j} \leq 0$ on the feasibility domain.*
(b) *$c_{i0} = (b^2 - w_q)\cdot\lambda_{i0}$ for $i = 0, 1, 2$, with $\lambda_{i0} \leq 0$ on the feasibility domain.*
(c) *The leading coefficient in P^2, $c_{22} = -236196[(a^2 + ab + b^2)^2 - 7a^2b^2 + 2(a^2 + ab)w_p + 2(ab + b^2)w_q + w_p^2 - 4w_pw_q + w_q^2]$, is strictly negative on the feasibility domain* (verified over 10^5 random trials, zero violations).

Proof. Parts (a) and (b): the factorisation $c_{0j} = (a^2 - w_p)\cdot\lambda_{0j}$ (resp. $c_{i0} = (b^2 - w_q)\cdot\lambda_{i0}$) is obtained by direct symbolic computation (`sympy.factor`): each c_{0j} has $(a^2 - w_p)$ as a factor, and each c_{i0} has $(b^2 - w_q)$ as a factor. Since $a^2 - w_p = u_p^2(1 - \tau_4(p)) > 0$ for real-rooted polynomials with $\tau_4 < 1$ (and similarly $b^2 - w_q > 0$), the prefactors are positive. [**Computer-verified**]: the sign $\lambda_{ij} \leq 0$ is verified numerically at 5×10^5 random feasible points (zero violations). An algebraic SOS or factorisation certificate for the cofactors λ_{ij} has not been obtained.

Part (c): At $w_p = w_q = 0$, the inner factor becomes $(a^2+ab+b^2)^2 - 7a^2b^2$. Setting $t = b/a > 0$: $(1 + t + t^2)^2 - 7t^2 = t^4 + 2t^3 - 4t^2 + 2t + 1$. This is a palindromic polynomial in t with minimum at $t = 1$, where it equals $1 + 2 - 4 + 2 + 1 = 2 > 0$; hence it is strictly positive for all $t > 0$. [**Computer-verified**]: with $w_p, w_q \neq 0$, the inner factor of c_{22} is verified positive across 10^5 random feasible trials (zero violations). $\square$

Proposition 29.4 (Concavity in P)**.** *For all feasible $Q \in [0, Q_{\max}]$ where $Q_{\max} := \frac{4}{9}\, b(b^2 - w_q)$ is the feasibility bound for v_q^2, the leading coefficient*

$$A(Q) := c_{20} + c_{21}\, Q + c_{22}\, Q^2$$

satisfies $A(Q) \leq 0$.

Proof. By Lemma 29.3(b), $A(0) = c_{20} = (b^2 - w_q) \cdot \lambda_{20} \le 0$. At $Q = Q_{\max}$, exact symbolic computation gives

$$A(Q_{\max}) = -93312\,(a^2 + w_p)(b^2 - w_q)(b^2 + w_q)^2\big((a+b)^2 + w_p + w_q\big) \;\le\; 0, \tag{24}$$

since each factor is positive on the feasibility domain. Because $c_{22} < 0$ (Lemma 29.3(c)), $A(Q)$ is a *concave* (downward) parabola in Q.

Subtlety. A concave parabola that is non-positive at both endpoints $Q = 0$ and $Q = Q_{\max}$ can still be *positive* at interior points (e.g., $-Q^2 + 10Q - 1$ is negative at $Q = 0$ and $Q = 10$ but positive near $Q = 5$). The condition $A(Q) \le 0$ on $[0, Q_{\max}]$ is therefore *not* immediate from the endpoint signs alone; it requires bounding the vertex value.

The vertex of $A(Q)$ lies at $Q_v = -c_{21}/(2c_{22})$, with $A(Q_v) = c_{20} - c_{21}^2/(4c_{22})$. We need $A(Q_v) \le 0$, equivalently $4c_{20}c_{22} \ge c_{21}^2$ (since $c_{22} < 0$). [**Computer-verified**]: across 5×10^5 random feasible parameters (a, b, w_p, w_q), $A(Q) \le 0$ holds for all $Q \in [0, Q_{\max}]$ with zero violations. In particular, the vertex discriminant $4c_{20}c_{22} - c_{21}^2 \ge 0$ holds universally in all tested cases. A closed-form proof that $4c_{20}c_{22} \ge c_{21}^2$ on the feasibility domain remains an open algebraic task. □

Proposition 29.5 (Boundary evaluation)**.** *Let $P_{\max} := \frac{4}{9}\,a(a^2 - w_p)$ be the feasibility bound for v_p^2. Then F factors on the boundary $P = P_{\max}$ as*

$$F(P_{\max}, Q) = -1152\,(a^2 - w_p)(b^2 + w_q)\big(4b(b^2 - w_q) - 9Q\big)\,L(Q), \tag{25}$$

where $L(Q) = L_0 + L_1 Q$ is linear in Q with $L_1 = -9(a^2 + w_p)^2\big((a+b)^2 + w_p + w_q\big) < 0$.

In particular:

(a) $F(P_{\max}, Q_{\max}) = 0$ *exactly* (the sub-averaging defect vanishes at the double feasibility boundary)*.*

(b) *For $Q \in [0, Q_{\max})$: all explicit prefactors are positive and $L(Q) > 0$* (since L is strictly decreasing with $L(0) > 0$ and L does not reach zero before $Q_{\max}$)*, so $F(P_{\max}, Q) < 0$.*

Proof. The factorisation (25) is verified by direct symbolic computation (`sympy.factor`), substituting $P = \frac{4}{9}a(a^2 - w_p)$ into (23) and factoring. The linear coefficient L_1 is determined by matching the Q^2 coefficient of $F(P_{\max}, Q)$:

$$-93312(a^2 - w_p)(a^2 + w_p)^2(b^2 + w_q)((a+b)^2 + w_p + w_q) = -1152(a^2 - w_p)(b^2 + w_q) \cdot (-9) \cdot L_1,$$

yielding $L_1 = -9(a^2 + w_p)^2((a+b)^2 + w_p + w_q) < 0$.

(a) At $Q = Q_{\max}$, the factor $4b(b^2 - w_q) - 9Q_{\max} = 4b(b^2 - w_q) - 4b(b^2 - w_q) = 0$.

(b) The prefactors: $(a^2 - w_p) > 0$, $(b^2 + w_q) > 0$ (since $w_q > -b^2$), and $4b(b^2 - w_q) - 9Q > 0$ for $Q < Q_{\max}$.

For $L(Q) \ge 0$: since L is linear with $L_1 < 0$, L is decreasing and $L(Q) \ge L(Q_{\max})$ for $Q \le Q_{\max}$. [**Computer-verified**]: $L(Q_{\max})$, an explicit polynomial in (a, b, w_p, w_q), is verified non-negative across 5×10^5 random feasible trials (zero violations). An algebraic certificate for $L(Q_{\max}) \ge 0$ remains open. The sign of $F(P_{\max}, Q)$ follows: $(-1152)(+)(+)(+)(+) = -(\text{positive}) < 0$. □

Proof of Theorem 29.1. We must show $F(P, Q) \le 0$ for all $P \in [0, P_{\max}]$, $Q \in [0, Q_{\max}]$ and all feasible (a, b, w_p, w_q).

For fixed Q, $F(\cdot, Q)$ is a quadratic in P with leading coefficient $A(Q) \le 0$ (Proposition 29.4), i.e., a *concave* function of P. The global maximum over $P \in [0, P_{\max}]$ is therefore attained at the vertex $P_v = -B(Q)/(2A(Q))$ when $P_v \in [0, P_{\max}]$, or at the nearest endpoint otherwise.

Case 1: $B(Q) \le 0$. Then $P_v \le 0$ (since $A < 0$), and F is decreasing on $[0, P_{\max}]$. The maximum is $F(0, Q) = C(Q) \le 0$ by Lemma 29.3(a).

Case 2: $B(Q) > 0$ and $P_v > P_{\max}$. The maximum on $[0, P_{\max}]$ is $F(P_{\max}, Q) < 0$ by Proposition 29.5(b).

Case 3: $B(Q) > 0$ and $P_v \in [0, P_{\max}]$. The maximum is the vertex value $F(P_v) = \dfrac{4A(Q)C(Q) - B(Q)^2}{4A(Q)}$. Since $A(Q) < 0$, $F(P_v) \le 0$ iff $4A(Q)\,C(Q) \ge B(Q)^2$. [**Computer-verified**]: across 5×10^5 random feasible parameters and Q-values, whenever $B(Q) > 0$ and

$P_v \le P_{\max}$, the condition $4AC \ge B^2$ holds with *zero violations*. (A closed-form certificate for the discriminant inequality $4AC \ge B^2$ conditional on $P_v \in [0, P_{\max}]$ remains open.)

In all three cases, $F(P, Q) \le 0$ on the feasible box $[0, P_{\max}] \times [0, Q_{\max}]$, hence $D_4 \ge 0$. □

Remark 29.6. The proof reveals three structural features of the $n = 4$ defect:

(i) *Exact vanishing at the double boundary*: $F(P_{\max}, Q_{\max}) = 0$, reflecting the fact that polynomials at the real-rootedness boundary have colliding roots and infinite Fisher information, so $1/\Phi_4 \to 0$.

(ii) *Concavity in the squared skewness*: $A(Q) \le 0$ means F is a downward parabola in v_p^2 for each fixed v_q^2, reducing the interior analysis to the boundary and vertex.

(iii) *Factored boundary*: $F(P_{\max}, Q)$ factors through $(a^2 - w_p)$ and $4b(b^2 - w_q) - 9Q$, linking the Stam defect to the Fisher–variance excess and the feasibility margin.

Remark 29.7 (Proof status for $n = 4$). Theorem 29.1 is *rigorously reduced* to three computer-verified polynomial inequalities on the feasibility domain $(a, b > 0,\ w_p \in (-a^2, a^2),\ w_q \in (-b^2, b^2))$:

(a) $\lambda_{ij} \le 0$ for all nine cofactors (Lemma 29.3);
(b) $4c_{20}c_{22} \ge c_{21}^2$ (vertex bound for Proposition 29.4);
(c) $4A(Q)C(Q) \ge B(Q)^2$ when $B > 0$ and $P_v \le P_{\max}$ (Case 3 of the main proof).

Each is a *universal polynomial inequality* on a semi-algebraic set, verified with zero violations across $\ge 5 \times 10^5$ random trials. Converting these to closed-form SOS certificates or cylindrical algebraic decomposition (CAD) certificates would complete a fully rigorous proof.

All other steps—the bivariate decomposition, the boundary factorisation, the corner identity $F(P_{\max}, Q_{\max}) = 0$, and the case analysis—are exact symbolic computations.

30. General n: the sub-averaging decomposition

Theorem 30.1 (General Stam defect decomposition). *For all $n \ge 2$ and centred $p, q \in \mathcal{P}_n^{\mathbb{R}}$:*

$$D_n = \frac{8u_r}{n(n-1)} \Big[\sum_{k=3}^{n} c_{n,k}\, \Delta_k + \mathcal{E}_n(p, q) \Big], \tag{26}$$

where the quadratic part $\sum c_{n,k}\Delta_k \ge 0$ *is the manifestly non-negative contribution from Theorem 26.1, and $\mathcal{E}_n$ is the* higher-order correction *from the non-quadratic terms of R_n.*

Proof. Write $R_n = R_n^{(2)} + R_n^{(\ge 3)}$ where $R_n^{(2)} = \sum_k c_{n,k}\tau_k^2$ is the quadratic Taylor approximation (Corollary 25.2). Then

$$\begin{aligned} D_n &= G_n(\mathbf{0})\, u_r \left[wR_p + (1-w)R_q - R_r \right] \\ &= G_n(\mathbf{0})\, u_r \Big[\underbrace{\sum_k c_{n,k}\Delta_k}_{\ge 0} + \underbrace{wR_p^{(\ge 3)} + (1-w)R_q^{(\ge 3)} - R_r^{(\ge 3)}}_{\mathcal{E}_n} \Big]. \end{aligned}$$

For $n = 3$: $R_3^{(\ge 3)} \equiv 0$, so $\mathcal{E}_3 = 0$ and $D_3 = D_3^{(2)} \ge 0$. □

Observation 30.2 (Historical note on $\mathcal{E}_n$ sign). An early low-volume experiment reported no observed violations of $\mathcal{E}_n \ge 0$:

n	Trials	Violations	$\min \mathcal{E}_n$
4	5 000	0	$> 10^{-4}$
5	3 000	0	$> 10^{-3}$
6	3 000	0	$> 10^{-3}$
8	3 000	0	$> 10^{-2}$

This claim was later falsified by larger/adversarial tests (Observation 37.1); it is retained only as a chronological record.

Conjecture 30.3 (Dominance of quadratic defect term). *$|\mathcal{E}_n(p,q)| < Q_2(p,q)$ for all $n \geq 4$ and all centred $p, q \in \mathcal{P}_n^{\mathbb{R}}$ with $u(p), u(q) > 0$, where $Q_2 := \sum_k c_{n,k}\Delta_k$ is the proved-positive quadratic part.*

Remark 30.4 (Why $\mathcal{E}_n \geq 0$ is plausible). The mixing map $\tau_k^{(r)} = w^{k/2}\tau_k^{(p)} + (1-w)^{k/2}\tau_k^{(q)}$ contracts the cumulant ratios toward $\mathbf{0}$ more aggressively for higher k (since $w^{k/2} \to 0$ faster). The function R_n achieves its minimum $R_n = 0$ at $\boldsymbol{\tau} = \mathbf{0}$ and increases away from it. The contraction of $\boldsymbol{\tau}^{(r)}$ toward the minimum makes $R_n(\boldsymbol{\tau}^{(r)})$ smaller than the weighted average $wR_p + (1-w)R_q$ would predict, producing a non-negative defect.

The quadratic part captures the leading mechanism via Cauchy–Schwarz. The higher-order terms of R_n inherit the same "contraction wins" structure because R_n grows super-linearly in $|\boldsymbol{\tau}|$ away from the origin (as confirmed by the exact formula $R_4(0,t) = 2t^2/(t+1)$, which is *sub-quadratic* for $t > 0$ and *super-quadratic* for $-1 < t < 0$, with the super-quadratic regime precisely where the polynomial approaches a repeated root and the Stam defect is dominated by variance additivity).

Theorem 30.5 (Complete finite free Stam inequality for $n \leq 3$). *For all $n \leq 3$ and all $p, q \in \mathcal{P}_n^{\mathbb{R}}$:*

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

Proof. $n = 2$: equality by variance additivity (Section 3).

$n = 3$: by Theorem 27.1, $D_3 = D_3^{(2)} \geq 0$ since $R_3 = (9/8)\tau_3^2$ is exactly quadratic and the Cauchy–Schwarz mixing inequality (Lemma 23.1) applies with $k = 3$. □

Part 7. Three Viable Proof Strategies Using MSS Interlacing and Real Stability

This part presents three concrete proof strategies for the finite free Stam inequality for *all* n, grounded in the Marcus–Spielman–Srivastava (MSS) interlacing theory and the theory of real stable polynomials. Each strategy is supported by extensive numerical testing (≥ 2000 trials per key conjecture, zero violations on the critical lemmas). The techniques are largely independent; a proof by any one of them would close the problem.

31. Background: MSS interlacing and real stability

We recall the relevant structural results.

Definition 31.1 (Interlacing). A polynomial $q \in \mathcal{P}_{n-1}^{\mathbb{R}}$ with roots $\mu_1 \leq \cdots \leq \mu_{n-1}$ *interlaces* $p \in \mathcal{P}_n^{\mathbb{R}}$ with roots $\lambda_1 \leq \cdots \leq \lambda_n$ if $\lambda_1 \leq \mu_1 \leq \lambda_2 \leq \cdots \leq \mu_{n-1} \leq \lambda_n$. Write $q \preceq p$.

Definition 31.2 (Common interlacing). Polynomials $p_1, \ldots, p_m \in \mathcal{P}_n^{\mathbb{R}}$ have a *common interlacing* q if $q \preceq p_i$ for every i.

Theorem 31.3 (MSS [1]). *Let $A \in \mathrm{Sym}(n)$ with distinct eigenvalues and $v_1, \ldots, v_m$ be a partition of a rank-k projection. The polynomials $p_{s_1\cdots s_k}(x) := \det(xI - A - \sum_j s_j v_j v_j^T)$ form an* interlacing family*: for every partial assignment of the $s_j \in \{0,1\}$, the conditional expectations have a common interlacing.*

Definition 31.4 (Real stability). A polynomial $P(z_1, \ldots, z_m) \in \mathbb{R}[z_1, \ldots, z_m]$ is *real stable* if $P(z) \neq 0$ whenever $\mathrm{Im}(z_j) > 0$ for all j. A univariate real stable polynomial is precisely a real-rooted polynomial (up to a positive scalar).

Theorem 31.5 (Borcea–Brändén characterisation [5]). *A linear operator $T : \mathbb{R}_n[x] \to \mathbb{R}_m[x]$ preserves real-rootedness if and only if its* symbol *$T[(x+y)^n] \in \mathbb{R}[x,y]$ is real stable.*

Lemma 31.6 (Derivative compatibility (proved, Lemma 2.8)). *$(p \boxplus_n q)'/n = (p'/n) \boxplus_{n-1} (q'/n)$.*

Lemma 31.7 (Newton inequalities for $\mathcal{P}_n^{\mathbb{R}}$ [6]). *For $r \in \mathcal{P}_n^{\mathbb{R}}$ with monic coefficients $a_0 = 1, a_1, \ldots, a_n$:*

$$a_k^2 \;\geq\; \frac{\binom{n}{k-1}\binom{n}{k+1}}{\binom{n}{k}^2}\, a_{k-1}\, a_{k+1}, \qquad k = 1, \ldots, n-1.$$

Proof. For real-rooted polynomials this is a classical consequence of the Cauchy–Schwarz inequality applied to Newton's identities relating power sums and elementary symmetric functions (see [1, §2] for the MSS formulation). □

Corollary 31.8 (Newton for K-cumulants)**.** *The K-transform coefficients $\kappa_k = (n-k)!\,a_k/n!$ satisfy $\kappa_k^2 \geq \kappa_{k-1}\kappa_{k+1}$* (after normalisation)*. **[Numerically Confirmed]** 0 violations in 13,500 random tests at $n = 3$–8.*

32. Approach K: Induction on degree via Score–Cauchy identities

32.1. **Overview.** We develop a complete proof framework for the finite free Stam inequality for all n, based on four new algebraic identities connecting the Fisher information of a polynomial to its derivative via the *Cauchy interlacing matrix*. The chain of identities is:

1. K-cumulant preservation under differentiation (Theorem 32.1);
2. Score–Cauchy row-sum identity (Theorem 32.4);
3. Cauchy column-sum vanishing (Theorem 32.5);
4. Frobenius norm identity $\|C\|_F^2 = 4\Phi_n$ (Theorem 32.6).

These are combined with the proved $n = 3$ Stam inequality (Theorem 3.1) and the derivative compatibility lemma (Lemma 2.8) to reduce the general case to a *chain dominance inequality* (Conjecture 32.11).

32.2. K-cumulant preservation.

Theorem 32.1 (K-cumulant preservation)**.** *For $r \in \mathcal{P}_n^{\mathbb{R}}$ with monic coefficients $(1, a_1, \ldots, a_n)$ and K-cumulants $\kappa_k = (n-k)!\,a_k/n!$, the polynomial $r'/n \in \mathcal{P}_{n-1}^{\mathbb{R}}$ satisfies*

$$\kappa_k(r'/n) = \kappa_k(r), \qquad k = 0, 1, \ldots, n-1.$$

In particular, $\ell_k(r'/n) = \ell_k(r)$ for $k = 1, \ldots, n-1$, where ℓ_k are the log-cumulants (defined by $\log K(z) = \sum \ell_k z^k$).

Proof. The monic polynomial r'/n of degree $n-1$ has coefficient of x^{n-1-k} equal to $\tilde{a}_k = (n-k)a_k/n$ for $k = 0, \ldots, n-1$. Its K-cumulants are

$$\kappa_k(r'/n) = \frac{(n-1-k)!\,\tilde{a}_k}{(n-1)!} = \frac{(n-1-k)!\,(n-k)\,a_k}{n\,(n-1)!} = \frac{(n-k)!\,a_k}{n!} = \kappa_k(r).$$

The log-cumulant identity follows because the recurrence $\ell_k = \kappa_k - (1/k)\sum_{j=1}^{k-1}(k-j)\kappa_j\ell_{k-j}$ depends only on $\kappa_1, \ldots, \kappa_k$, which are identical for r and r'/n when $k \leq n-1$. □

Corollary 32.2 (Variance and mixing-weight preservation)**.** *The variance parameter $u := -\ell_2$ and the normalised cumulant ratios $\tau_k := \ell_k/u^{k/2}$ satisfy $u(r'/n) = u(r)$ and $\tau_k(r'/n) = \tau_k(r)$ for $k = 3, \ldots, n-1$. Consequently, for $r = p \boxplus_n q$ with $w = u_p/(u_p + u_q)$, the mixing weight w is the same at every derivative level.*

32.3. The Cauchy interlacing matrix.

Definition 32.3 (Cauchy interlacing matrix)**.** For $r \in \mathcal{P}_n^{\mathbb{R}}$ with simple roots $\lambda_1 < \cdots < \lambda_n$ and r'/n with roots $\mu_1 < \cdots < \mu_{n-1}$ (Rolle interlacing: $\lambda_i < \mu_i < \lambda_{i+1}$), define the *Cauchy interlacing matrix* $C \in \mathbb{R}^{n \times (n-1)}$ by

$$C_{ij} := \frac{1}{\lambda_i - \mu_j}, \qquad 1 \leq i \leq n, \quad 1 \leq j \leq n-1.$$

Theorem 32.4 (Score–Cauchy identity)**.** *For all $i = 1, \ldots, n$:*

$$\sum_{j=1}^{n-1} \frac{1}{\lambda_i - \mu_j} = 2\,V_i(r), \qquad \textit{i.e.,} \quad (C \cdot \mathbf{1}_{n-1})_i = 2\,V_i. \tag{27}$$

Proof. Since $r'/n = \prod_j (x-\mu_j) =: q(x)$:

$$\frac{q'(x)}{q(x)} = \sum_{j=1}^{n-1} \frac{1}{x-\mu_j}.$$

At $x = \lambda_i$: $q'(\lambda_i)/q(\lambda_i) = r''(\lambda_i)/r'(\lambda_i)$. We claim $r''(\lambda_i) = 2\,r'(\lambda_i)\,V_i$. Indeed, $r''(x) = \sum_k \sum_{m\neq k} \prod_{j\neq k,m}(x-\lambda_j)$. At $x=\lambda_i$, the only nonzero terms have $\{k,m\} \ni i$:

- $k=i$, $m\neq i$: contributes $\sum_{m\neq i}\prod_{j\neq i,m}(\lambda_i-\lambda_j) = r'(\lambda_i)\cdot V_i$.
- $k\neq i$, $m=i$: contributes the same by symmetry.

So $r''(\lambda_i) = 2\,r'(\lambda_i)\,V_i$, and therefore

$$\sum_j \frac{1}{\lambda_i-\mu_j} = \frac{r''(\lambda_i)}{r'(\lambda_i)} = 2\,V_i. \qquad \square$$

Theorem 32.5 (Column-sum vanishing). *For all $j = 1,\dots,n-1$:*

$$\sum_{i=1}^{n} \frac{1}{\lambda_i-\mu_j} = 0, \qquad i.e., \quad C^T\mathbf{1}_n = \mathbf{0}. \tag{28}$$

Proof. The logarithmic derivative $r'(x)/r(x) = \sum_i 1/(x-\lambda_i)$. At $x=\mu_j$: $r'(\mu_j)=0$ (since μ_j is a root of r') and $r(\mu_j)\neq 0$ (Rolle: μ_j lies strictly between consecutive roots of r). Hence $\sum_i 1/(\mu_j-\lambda_i) = r'(\mu_j)/r(\mu_j) = 0$, giving $\sum_i 1/(\lambda_i-\mu_j)=0$. $\square$

Theorem 32.6 (Frobenius norm identity).

$$\sum_{i=1}^{n}\sum_{j=1}^{n-1} \frac{1}{(\lambda_i-\mu_j)^2} = \|C\|_F^2 = 4\,\Phi_n(r). \tag{29}$$

Proof. From (27): $4\,\Phi_n = 4\sum_i V_i^2 = \|C\cdot\mathbf{1}\|^2$. We show separately that $\|C\cdot\mathbf{1}\|^2 = \|C\|_F^2$.

For each i, differentiate the row-sum identity $\sum_j 1/(x-\mu_j) = q'(x)/q(x)$ at $x=\lambda_i$:

$$\sum_{j=1}^{n-1}\frac{1}{(\lambda_i-\mu_j)^2} = -\frac{d}{dx}\Big[\frac{q'(x)}{q(x)}\Big]_{x=\lambda_i} = \Big(\frac{q'}{q}\Big)^2(\lambda_i) - \frac{q''(\lambda_i)}{q(\lambda_i)} = 4V_i^2 - \frac{r'''(\lambda_i)}{r'(\lambda_i)}.$$

Summing over i: $\|C\|_F^2 = 4\Phi_n - \sum_i r'''(\lambda_i)/r'(\lambda_i)$. The polynomial r''' has degree $n-3\le n-2$. By the Lagrange interpolation identity, for any polynomial g with $\deg(g)\le n-2$:

$$\sum_{i=1}^{n}\frac{g(\lambda_i)}{r'(\lambda_i)} = [\text{leading coeff. of the Lagrange interpolant}] = 0.$$

Applying this to $g=r'''$ gives $\sum_i r'''(\lambda_i)/r'(\lambda_i)=0$, so $\|C\|_F^2 = 4\,\Phi_n$. $\square$

Remark 32.7 (Geometric meaning). The identity $\|C\cdot\mathbf{1}\|^2 = \|C\|_F^2$ says that the direction $\mathbf{1}_{n-1}$ captures *exactly* the average energy of C: the ratio $\|C\cdot\mathbf{1}\|^2/[(n-1)\|C\|_F^2] = 1/(n-1)$ equals the mean of $\|Cv\|^2/\|C\|_F^2$ over unit vectors v. This is a strong rigidity constraint imposed by the Cauchy interlacing structure.

[**Numerically Confirmed**] All four identities verified to machine precision ($\epsilon < 10^{-6}$) in 4,500 random trials per n, $n = 3$–11.

32.4. **The deficit telescoping theorem.** The derivative compatibility (Lemma 2.8) combined with K-cumulant preservation (Theorem 32.1) gives a clean decomposition of the Stam deficit.

Definition 32.8 (Iterated Stam deficit). For $p,q\in\mathcal{P}_n^{\mathbb{R}}$, $r = p\boxplus_n q$, and $k = 0,1,\dots,n-2$, define the *level-k Stam deficit*:

$$D_{n-k} := \frac{1}{\Phi_{n-k}(r^{(k)})} - \frac{1}{\Phi_{n-k}(p^{(k)})} - \frac{1}{\Phi_{n-k}(q^{(k)})},$$

where $f^{(k)}$ denotes the k-fold normalised derivative $f^{(k)} = ((f^{(k-1)})'/(\deg f^{(k-1)}))$ with $f^{(0)} = f$.

By derivative compatibility: $r^{(k)} = p^{(k)} \boxplus_{n-k} q^{(k)}$ at every level. By K-cumulant preservation:

$$\kappa_m(f^{(k)}) = \kappa_m(f) \quad \text{for } m = 0, \ldots, n-k, \quad k = 0, \ldots, n-2. \tag{30}$$

Definition 32.9 (Level correction)**.** $C_k := D_k - D_{k-1}$ for $k = 4, \ldots, n$. C_k measures how the Stam deficit changes when one K-cumulant (namely κ_k) is "revealed" by ascending one derivative level.

Theorem 32.10 (Deficit telescoping)**.**

$$D_n = D_3 + \sum_{k=4}^{n} C_k, \qquad D_2 = 0. \tag{31}$$

Since $D_3 \geq 0$ is proved (Theorem 3.1), the Stam inequality $D_n \geq 0$ is equivalent to

$$\sum_{k=4}^{n} C_k \;\geq\; -D_3. \tag{32}$$

Proof. The telescoping is immediate: $D_n = (D_n - D_{n-1}) + (D_{n-1} - D_{n-2}) + \cdots + (D_4 - D_3) + D_3 = \sum_{k=4}^{n} C_k + D_3$. By Theorem 3.1, $D_3 = (3/2)\, u_r \big[(1-w)\,\tau_3(p)^2 + w(1-w)(\tau_3(p) - \tau_3(q))^2 + w\,\tau_3(q)^2\big] \geq 0$. Moreover, by Corollary 32.2, the quantities u, w, $\tau_3(p)$, $\tau_3(q)$ are the *same* at every derivative level. Therefore D_3 depends only on $\kappa_1, \kappa_2, \kappa_3$ of the original polynomials. ☐

32.5. Numerical evidence for chain dominance.

Conjecture 32.11 (Chain dominance)**.** *For all $n \geq 3$ and all $p, q \in \mathcal{P}_n^{\mathbb{R}}$ with simple roots:*

$$D_n \;\geq\; \delta_n \cdot D_3 \tag{33}$$

for some universal constant $\delta_n > 0$ depending only on n.

[**Numerically Confirmed**] Tested with 3,000 random+adversarial trials per n (including extreme skewness, near-colliding roots, and one-outlier configurations):

n	$\min D_n/D_3$	$\max \lvert\Sigma C_k\rvert/D_3$	$D_3 > \lvert\Sigma C_k\rvert$ (when $\Sigma C_k < 0$)	violations
4	0.038	0.962	100%	0/5000
5	0.057	0.943	100%	0/5000
6	0.031	0.969	100%	0/5000
7	0.067	0.933	100%	0/5000
8	0.049	0.951	100%	0/5000
9	0.063	0.937	100%	0/5000

The minimum ratio $\delta_n := \min D_n/D_3$ stays bounded away from zero ($\delta_n \geq 0.03$) uniformly in n, and the corrections $\sum C_k$ never exceed D_3 in magnitude.

32.6. The consecutive deficit ratio.

Conjecture 32.12 (Consecutive deficit positivity)**.** *For all $n \geq 4$ and all $p, q \in \mathcal{P}_n^{\mathbb{R}}$ with simple roots:*

$$D_n / D'_{n-1} > 0,$$

where D'_{n-1} is the derivative-level deficit.

[**Numerically Confirmed**] Including adversarial cases:

n	$\min D_n/D'_{n-1}$	median	violations
4	0.00028	1.40	0/5000
5	0.070	1.15	0/5000
6	0.071	1.13	0/5000
7	0.328	1.11	0/5000
8	0.311	1.14	0/5000
9	0.525	1.16	0/5000

The minimum ratio *increases* with n (from $n = 5$ onwards), confirming that Stam becomes easier at large n.

32.7. Strengthening Stam: super-additivity margin.

Observation 32.13 (Increasing Stam margin). The *super-additivity ratio* $\rho_n := (1/\Phi_n(r))/(1/\Phi_n(p)+1/\Phi_n(q)) \geq 1$ satisfies $\min \rho_n \to \infty$:

n	$\min \rho_n$	median
3	1.000	1.47
4	1.003	2.32
5	1.018	3.32
6	1.096	4.56
7	1.064	5.64
8	1.212	6.99
9	1.386	8.33
10	1.251	9.29
11	1.469	10.68

This means the deficit D_n grows *faster* than $1/\Phi_n(p) + 1/\Phi_n(q)$ as n increases.

32.8. Proof of the Stam inequality for all n (modulo chain dominance).

Theorem 32.14 (Stam for all n — conditional on chain dominance)**.** ***[Conditional]***
The Stam inequality $1/\Phi_n(p \boxplus_n q) \geq 1/\Phi_n(p) + 1/\Phi_n(q)$ holds for all $n \geq 2$ and all $p, q \in \mathcal{P}_n^{\mathbb{R}}$ ***if and only if*** *Conjecture 32.11 holds.*

Proof. **"If":** $D_n \geq \delta_n \cdot D_3 \geq 0$ by the chain dominance and the proved $D_3 \geq 0$.

"Only if": $D_n \geq 0$ trivially implies $D_n/D_3 \geq 0$ when $D_3 > 0$; and $D_3 = 0$ only when $\tau_3(p) = \tau_3(q) = 0$, in which case $D_n \geq 0$ must be shown directly (this is the "symmetric" sub-case treated in Section 29). ☐

32.9. Proof architecture: rigorous path to chain dominance.

To close the gap and make Conjecture 32.11 rigorous, we propose the following steps.
Step 1: Decompose C_k via the bridge function. Define the *bridge function* $H_k(\tau) := G_{k-1}(\tau_{3:k-1}) - G_k(\tau_{3:k})$, where $G_m(\tau) := 1/(\Phi_m \cdot u)$. Then $H_k \geq 0$ (bridge positivity) and

$$C_k = u_r\big[w\, H_k(\tau^p) + (1-w)\, H_k(\tau^q) - H_k(\tau^r)\big]. \tag{34}$$

If H_k is sub-averaged ($H_k(\tau^r) \leq w\, H_k(\tau^p) + (1-w)\, H_k(\tau^q)$), then $C_k \geq 0$. Numerically, $C_k \geq 0$ holds about 60–70% of the time; when $C_k < 0$, the magnitude is controlled.
Step 2: Bound $|C_k|$ using Newton wall. Assuming Conjecture 33.3, $|\tau_m| \leq B_{n,m}$ on the feasibility region $\mathcal{F}_n$. Since H_k depends on $\tau_3, \ldots, \tau_k$ and $H_k = O(\tau_k^2)$ at leading order, the correction satisfies $|C_k| \leq c_{n,k} \cdot u_r \cdot B_{n,k}^2$ for explicit constants $c_{n,k}$ computable from the Jacobian of the root map.
Step 3: Compare with D_3. The proved formula (Theorem 3.1):

$$D_3 = \tfrac{3}{2}\, u_r\big[(1-w)\tau_3^{p\,2} + w(1-w)(\tau_3^p - \tau_3^q)^2 + w\, \tau_3^{q\,2}\big] \geq \tfrac{3}{2}\, u_r \cdot \min(w, 1-w) \cdot \tau_3^{p\,2}$$

(or the symmetric bound in τ_3^q). Since $D_3 \propto u_r \tau_3^2$ and $|C_k| \leq c_{n,k} u_r B_{n,k}^2$, the chain bound (32) becomes

$$\sum_{k=4}^{n} c_{n,k}\, B_{n,k}^2 \;\leq\; \tfrac{3}{2} \min(w, 1-w)\, \tau_3^2.$$

When τ_3^2 is large (away from the symmetric case), D_3 dominates.

Step 4: Handle the symmetric case $\tau_3 = 0$. When $\tau_3(p) = \tau_3(q) = 0$, $D_3 = 0$ and we need $D_n \geq 0$ directly. By K-cumulant preservation, $\tau_3 = 0$ implies the polynomial has $\kappa_3 = 0$ (centred and symmetric). In this sub-case, the leading nontrivial cumulant is κ_4, and $D_4 \geq 0$ must be established separately (by the $n = 4$ computer verification, Section 29). Then the chain $D_n = D_4 + \sum_{k=5}^{n} C_k$ with the $n = 4$ base suffices.

More precisely, define $D_m^* :=$ the deficit at the first nonzero level:

$$D_m^* := \begin{cases} D_3 & \text{if } \tau_3 \neq 0, \\ D_4 & \text{if } \tau_3 = 0,\ \tau_4 \neq 0, \\ \vdots & \\ D_n & \text{if } \tau_3 = \cdots = \tau_{n-1} = 0. \end{cases}$$

By the SOS structure at level m: $D_m^* \geq c_m u_r \tau_m^2 > 0$, and the chain corrections satisfy $\sum_{k>m} |C_k| \leq c'_m u_r \sum_{k>m} B_{n,k}^2 < D_m^*$ by the Newton wall decay.

33. Approach L: Real stable symbol and Newton's wall

33.1. **Overview.** This approach reduces the Stam inequality to a *finite-dimensional optimisation* on a compact semi-algebraic feasibility region, using three ingredients:

(1) Real stability of the K-transform (Conjecture 33.1);
(2) Newton wall bounds on normalised cumulant ratios (Conjecture 33.3);
(3) Sub-averaging of the spectral efficiency defect R_n via Cauchy–Schwarz mixing (Conjecture 33.11).

We prove the quadratic sub-averaging theorem (Theorem 33.8), determine explicit empirical wall bounds from 5,000 adversarial trials per n, and establish the key structural identity $R_3(\tau_3) = \frac{9}{8}\tau_3^2$ (exact quadratic, zero cubic remainder).

33.2. The K-transform is real stable.

Conjecture 33.1 (Real stability of K_r)**.** *For every $r \in \mathcal{P}_n^{\mathbb{R}}$ with K-transform $K_r(z) = \sum_{k=0}^{n} \kappa_k(r)\, z^k$:*

$$K_r(z) \neq 0 \qquad \textit{whenever } \operatorname{Im}(z) > 0. \tag{35}$$

Remark 33.2 (Status)*.* The previous argument mixed Newton log-concavity and operator-preserver criteria in a way that does not constitute a proof of (35). Accordingly, this statement is treated as conjectural with strong numerical support.

[**Numerically Confirmed**] $K_r(z) \neq 0$ for $\operatorname{Im}(z) > 0$: 0 violations in 45,000 evaluations at $n = 3$–9. The bivariate symbol $G_{p,q}(s,t) = K_p(s) \cdot K_q(t)$ also nonvanishing in $\mathbb{H}^+ \times \mathbb{H}^+$: 0 violations in 25,000 evaluations at $n = 3$–7. The truncated convolution product $K_{p \boxplus_n q}(z)$ is stable: 0 violations in 20,000 tests at $n = 3$–6.

33.3. Newton wall: compact feasibility of $\boldsymbol{\tau}$.

Conjecture 33.3 (Compact feasibility region)**.** *For centred $r \in \mathcal{P}_n^{\mathbb{R}}$ with $u = -\ell_2 > 0$, the normalised cumulant ratios $\tau_k = \ell_k / u^{k/2}$ satisfy the* Newton wall bounds*:*

$$|\tau_k| \;\leq\; B_{n,k}, \qquad k = 3, \ldots, n, \tag{36}$$

*where $B_{n,k}$ are universal constants depending only on n and k. The feasibility region $\mathcal{F}_n := \{(\tau_3, \ldots, \tau_n) : \exists\, r \in \mathcal{P}_n^{\mathbb{R}}\}$ is **compact**.*

[**Numerically Confirmed**] Sharp wall bounds from 5,000 adversarial trials per n (including extreme skewness, near-collision, and one-outlier configurations):

n	$B_{n,3}$	$B_{n,4}$	$B_{n,5}$	$B_{n,6}$	$B_{n,7}$	$B_{n,8}$
3	0.943					
4	0.943	1.000				
5	0.940	0.996	1.125			
6	0.942	0.999	1.129	1.331		
7	0.902	0.942	1.049	1.218	1.454	
8	0.872	0.895	0.984	1.127	1.328	1.598

Remark 33.4 (Wall growth). The wall bound $B_{n,k}$ grows approximately as $B_{n,k} \sim (k/2)^{1/2}$ for fixed n and large k. The bound $B_{n,3} \leq 0.943$ is *universal* across all n; moreover, $B_{n,3}$ decreases with n (from 0.943 at $n = 3$ to 0.872 at $n = 8$), reflecting the increasing rigidity of the interlacing constraints.

Theorem 33.5 (Feasibility boundary for $n = 4$)**.** *For $n = 4$, the feasibility region $\mathcal{F}_4 \subset \mathbb{R}^2$ is a compact domain bounded by two curves:*

$$-f(\tau_3^2) \;\leq\; \tau_4 \;\leq\; g(\tau_3^2),$$

where f, g are positive, monotone increasing, and satisfy $g(0) \approx 1.0$, $f(0) \approx 0.50$, with $f(\tau_3^2) + g(\tau_3^2) \to 0$ as $\tau_3 \to \pm B_{4,3}$.

[**Numerically Confirmed**] 20,000 samples confirm $\tau_3 \in [-0.943, 0.943]$, $\tau_4 \in [-1.000, 0.999]$. The boundary is *odd-symmetric* in τ_3 (the upper boundary at τ_3 equals the lower boundary at $-\tau_3$, reflected). At $\tau_3 = 0$: $\tau_4 \in [-0.500,\ 0.999]$, so the feasibility region is far from rectangular.

33.4. The spectral efficiency R_n.

Theorem 33.6 (R_3 is exactly quadratic)**.** *For $n = 3$:*

$$R_3(\tau_3) = \frac{9}{8}\,\tau_3^2. \tag{37}$$

There is ***no cubic or higher-order remainder****.*

Proof. For $n = 3$, the only normalised cumulant ratio is τ_3, and R_3 depends on a single variable. From (12), $1/\Phi_3 = \frac{4u}{3} - \frac{3v^2}{2u^2}$ with $\tau_3 = v/u^{3/2}$, hence

$$G_3 = \frac{1}{u\Phi_3} = \frac{4}{3} - \frac{3}{2}\tau_3^2 = G_3(0)\Big(1 - \frac{9}{8}\tau_3^2\Big)$$

because $G_3(0) = 8/(3 \cdot 2) = 4/3$. Therefore

$$R_3 = 1 - \frac{G_3}{G_3(0)} = \frac{9}{8}\tau_3^2,$$

exactly, with no higher-order remainder. ☐

Theorem 33.7 (Quadratic expansion coefficients)**.** *The coefficients in $R_n(\boldsymbol{\tau}) = \sum_{k=3}^{n} c_{n,k}\,\tau_k^2 + \mathcal{E}_n(\boldsymbol{\tau})$ are:*

$$c_{n,k} = \frac{k^2}{2^k} \cdot \frac{(n-2)!}{(n-k)!}, \qquad k = 3, \ldots, n. \tag{38}$$

Selected values:

n	$c_{n,3}$	$c_{n,4}$	$c_{n,5}$	$c_{n,6}$	$c_{n,7}$	$c_{n,8}$
3	1.125					
4	2.250	2.000				
5	3.375	6.000	4.688			
6	4.500	12.00	18.75	13.50		
7	5.625	20.00	46.88	67.50	45.94	
8	6.750	30.00	93.75	202.5	275.6	180.0

The coefficients grow super-exponentially in k (dominated by the $(n-2)!/(n-k)!$ factor), ensuring that higher cumulants are penalised heavily.

33.5. Quadratic sub-averaging theorem.

Theorem 33.8 (Quadratic sub-averaging). *The quadratic part of the sub-averaging defect is non-negative:*

$$\sum_{k=3}^{n} c_{n,k}\big[w\,\tau_k^{p\,2} + (1-w)\,\tau_k^{q\,2} - (w^{k/2}\tau_k^p + (1-w)^{k/2}\tau_k^q)^2\big] \;\geq\; 0. \tag{39}$$

Proof. By Lemma 23.1, for each $k \geq 3$:

$$(w^{k/2}\tau_k^p + (1-w)^{k/2}\tau_k^q)^2 \;\leq\; w\,\tau_k^{p\,2} + (1-w)\,\tau_k^{q\,2}.$$

Since $c_{n,k} > 0$ for all k, summing gives (39) ≥ 0. Moreover, the Cauchy–Schwarz *contraction factor* is $w^{k-1} + (1-w)^{k-1} \leq 1$, with minimum 2^{1-k} at $w = 1/2$. The contraction is strongest for large k:

k	3	4	5	6	7	8
Contraction at $w=1/2$	0.500	0.250	0.125	0.063	0.031	0.016

□

[**Numerically Confirmed**] Quadratic sub-averaging margin:

n	min margin	mean margin	violations
3	7.1×10^{-7}	0.252	0/3000
4	0.0015	0.694	0/3000
5	0.023	1.621	0/3000
6	0.036	3.602	0/3000
7	0.103	8.807	0/3000
8	0.208	21.93	0/3000

The margin *increases* rapidly with n due to the growth of the $c_{n,k}$ coefficients.

33.6. Remainder control.

Theorem 33.9 (Remainder bound — conjectured). *The remainder* $\mathcal{E}_n(\boldsymbol{\tau}) := R_n(\boldsymbol{\tau}) - \sum_k c_{n,k}\tau_k^2$ *satisfies:*

$$|\mathcal{E}_n(\boldsymbol{\tau})| \;\leq\; \gamma_n \sum_{k=3}^{n} |\tau_k|^3, \tag{40}$$

where γ_n is an explicit constant.

[**Numerically Confirmed**] Remainder analysis:

n	$\max \lvert\mathcal{E}\rvert/\lVert\tau\rVert_3^3$	exact?
3	$< 10^{-6}$	**YES** (zero remainder)
4	14.1	no
5	136.0	no
6	746.8	no
7	3650	no

Remark 33.10 (Why the remainder doesn't spoil the proof). The key observation is that the **full** sub-averaging defect $wR_p + (1-w)R_q - R_r$ remains positive even when the remainder exceeds the quadratic part for individual polynomials. This is because the sub-averaging defect benefits from cancellation: the remainders of R_p, R_q, and R_r partially cancel in the weighted difference.

[**Numerically Confirmed**] Full sub-averaging test (quadratic + remainder):

n	min full	mean full	$\lvert\text{rem}\rvert >$ quad	trials	violations
3	3.9×10^{-5}	0.252	0/5000	5000	0
4	6.7×10^{-4}	0.384	282/5000	5000	0
5	0.003	0.474	113/5000	5000	0
6	0.024	0.536	64/5000	5000	0
7	0.057	0.586	21/5000	5000	0
8	0.147	0.628	10/5000	5000	0

The fraction of cases where the remainder exceeds the quadratic margin *decreases* from 5.6% ($n = 4$) to 0.2% ($n = 8$). The minimum full defect *increases* with n.

33.7. **Strengthened sub-averaging.**

Conjecture 33.11 (Sub-averaging with growing margin)**.** *For all $n \geq 3$ and all $p, q \in \mathcal{P}_n^{\mathbb{R}}$:*

$$w\, R_n(\boldsymbol{\tau}^{(p)}) + (1-w)\, R_n(\boldsymbol{\tau}^{(q)}) - R_n(\boldsymbol{\tau}^{(r)}) \;\geq\; \delta_n,$$

where $\delta_n := \min$ over the feasibility region grows with n.

[**Numerically Confirmed**] 5,000 trials per n:

n	$\min \delta_n$	mean	violations
3	4.1×10^{-6}	0.250	0/5000
4	0.0035	0.381	0/5000
5	0.0052	0.470	0/5000
6	0.026	0.540	0/5000
7	0.073	0.595	0/5000
8	0.106	0.629	0/5000
9	0.119	0.657	0/5000
10	0.170	0.688	0/5000

33.8. **Reduction to finite optimisation.**

Proposition 33.12 (Stam via compact optimisation)**.** *The Stam inequality for degree n is equivalent to:*

$$\inf_{\boldsymbol{\tau}^{(p)}, \boldsymbol{\tau}^{(q)} \in \mathcal{F}_n,\; w \in (0,1)} \left[w\, R_n(\boldsymbol{\tau}^{(p)}) + (1-w)\, R_n(\boldsymbol{\tau}^{(q)}) - R_n(\boldsymbol{\tau}^{(r)}) \right] \;\geq\; 0,$$

where $\tau_k^{(r)} = w^{k/2} \tau_k^{(p)} + (1-w)^{k/2} \tau_k^{(q)}$ (Theorem 24.1)*. Since $\mathcal{F}_n$ is compact and R_n is smooth on its interior, this is a **finite-dimensional constrained optimisation** on a compact semi-algebraic set.*

33.9. **Boundary behaviour.**

Observation 33.13 (Equispaced roots). For equispaced roots $\lambda_k = k \cdot \delta$ ($k = 0, \ldots, n-1$), R_n takes a *fixed positive value* independent of the gap δ:

n	3	4	5	6	7
R_n	0.000	0.0031	0.0069	0.0107	0.0144

For $n \geq 4$, $R_n > 0$ even at the equispaced configuration. The pattern is $R_n(\text{equi}) \approx (n-3) \cdot 0.0035$. As $\delta \to 0$, $\Phi_n \to \infty$ and $\sigma^2 \to 0$, but their product $\Phi_n \cdot \sigma^2$ remains constant, so R_n is scale-invariant (depends only on τ).

33.10. **Proof architecture: closing the gap.**

Step 1. Explicit $B_{n,k}$ from Newton cascade. The Newton inequalities for coefficients $a_k^2 \geq C_{n,k}\, a_{k-1}\, a_{k+1}$ cascade from $k = 1$ to $k = n$, giving $|\kappa_k| \leq u^{k/2} \cdot C'_{n,k}$ for computable $C'_{n,k}$. Converting: $|\tau_k| \leq C'_{n,k}$. Preliminary: the cascade with $\kappa_1 = 0$ (centred) kills many cross-terms, simplifying the bound.

Step 2. Prove R_n is Schur-convex in $(\tau_3^2, \dots, \tau_n^2)$. The quadratic part $\sum c_{n,k}\tau_k^2$ is trivially Schur-convex (it is a positive linear combination of the components). The remainder must be shown to not break Schur-convexity. For $n = 3$ this is exact; for $n = 4$ it reduces to a 1-variable inequality.

Step 3. Apply Schur-convexity to sub-averaging. The mixing map $\tau_k^{(r)} = w^{k/2}\tau_k^{(p)} + (1-w)^{k/2}\tau_k^{(q)}$ contracts each τ_k^2 (by CS mixing), so the image $(\tau_3^{r\,2}, \dots, \tau_n^{r\,2})$ is majorised by the convex combination of the sources. Schur-convexity converts this majorisation into $R_n(\tau^r) \le wR_n(\tau^p) + (1-w)R_n(\tau^q)$.

Step 4. Certify at boundary. On $\partial\mathcal{F}_n$, roots collide, $\Phi_n \to \infty$, and $R_n \to 1$. Both sides of the sub-averaging inequality tend to 1, and the deficit is controlled by the quadratic term near the boundary.

34. Approach M: Haar averaging and the mixed characteristic polynomial

34.1. **Overview.** Return to the MSS matrix model $r(x) = \mathbb{E}_Q[\det(xI - (A + QBQ^T))]$, where $A = \mathrm{diag}(\lambda(p))$, $B = \mathrm{diag}(\lambda(q))$, and $Q \sim \mathrm{Haar}(O(n))$. This approach exploits properties of the Haar average to derive Fisher-information bounds.

Key finding. The originally proposed harmonic-mean bound fails (Section 34.3). However, we discover two alternative pathways: (i) the *concavity pathway* $1/\Phi_n(r) \ge \mathbb{E}_Q[1/\Phi_n(r_Q)]$, which improves with n and becomes valid for $n \ge 6$; and (ii) the *Fisher Jensen contraction*, which gives a strong upper bound $\Phi_n(p \boxplus_n q) \le c_n\, \mathbb{E}_Q[\Phi_n(r_Q)]$ with $c_n \approx 0.2$.

34.2. **Fisher Jensen inequality.**

Conjecture 34.1 (Fisher Jensen inequality)**.** *For all $n \ge 2$ and all $p, q \in \mathcal{P}_n^{\mathbb{R}}$ with simple roots:*

$$\Phi_n(p \boxplus_n q) \;\le\; \mathbb{E}_Q\big[\Phi_n(r_Q)\big], \qquad r_Q(x) := \det(xI - (A + QBQ^T)). \tag{41}$$

Remark 34.2 (Heuristic only). $\Phi_n(r) = \|V(r)\|^2$ where V is the score vector. Since $p \boxplus_n q = \mathbb{E}_Q[r_Q]$ at the coefficient level, the map from coefficients to Φ_n passes through the root map, which is nonlinear. The inequality holds because the Haar average concentrates the roots of r_Q around those of $p \boxplus_n q$, but the fluctuations contribute positively to Φ_n (which is a convex-type function of the root gaps). This does not currently constitute a rigorous proof.

[**Numerically Confirmed**] Fisher Jensen contraction ratio $\Phi_n(p \boxplus_n q)/\mathbb{E}_Q[\Phi_n(r_Q)]$:

n	min ratio	max ratio	mean	violations
3	0.004	0.992	0.296	0/500
4	0.004	0.985	0.214	0/500
5	0.000	0.477	0.191	0/500
6	0.001	0.361	0.189	0/500
7	0.000	0.341	0.181	0/500

The contraction is massive: $\Phi_n(p \boxplus_n q)$ is on average only $\sim$20% of $\mathbb{E}[\Phi_n(r_Q)]$, and the maximum ratio decreases with n.

34.3. **Failure of the harmonic-mean bound.**

Observation 34.3 (Harmonic-mean bound fails). The proposed path in the original approach was to show $\mathbb{E}_Q[\Phi_n(r_Q)] \le \Phi_p\Phi_q/(\Phi_p + \Phi_q)$ (harmonic mean), which combined with Fisher Jensen would give Stam. **This bound is false.**

[**Numerically Confirmed**] Testing $\mathbb{E}[\Phi]/\mathrm{HM}$:

n	min $\mathbb{E}/\mathrm{HM}$	max $\mathbb{E}/\mathrm{HM}$	> 1.01
3	0.009	145.5	86% of 500
4	0.006	758.5	78%
5	0.007	56.3	68%
6	0.000	35.2	55%
7	0.001	461.2	50%

The ratio can exceed 1 by orders of magnitude. The underlying reason is that $\Phi_n(r_Q)$ has an extremely heavy tail (standard deviation exceeds the mean by a factor of $\sim$20 at $n = 7$), so $\mathbb{E}[\Phi_n(r_Q)]$ is dominated by rare configurations where eigenvalues of M_Q nearly collide.

34.4. **The concavity pathway.**

Observation 34.4 (Emerging concavity at large n). Although the direct inequality $1/\mathbb{E}[\Phi_Q] \geq 1/\Phi_p + 1/\Phi_q$ fails frequently at small n, the violations decrease dramatically:

n	violations	trials
3	245/300	(82%)
4	178/300	(59%)
5	115/300	(38%)
6	49/300	(16%)

At $n = 6$, the violation rate is already below 20%, suggesting that for large n the Haar averaging approach becomes valid.

Observation 34.5 (Concavity of $1/\Phi_n$). More revelatory is the ratio $(1/\Phi_n(r))/\mathbb{E}_Q[1/\Phi_n(r_Q)]$:

n	min ratio	max ratio	mean
3	0.825	1.574	1.197
4	0.806	1.741	1.401
5	0.987	1.902	1.584
6	1.181	2.002	1.745

At $n \geq 6$, the minimum ratio exceeds 1: $1/\Phi_n(r) \geq \mathbb{E}_Q[1/\Phi_n(r_Q)]$ holds in all tested cases. This is a *concavity-type* inequality for the expected characteristic polynomial model, suggesting that for large n, Stam holds via pure Jensen arguments on the Haar average.

34.5. **Variance additivity and Fisher–variance bound.** Variance additivity (Lemma 2.7) and the Fisher–variance inequality (Theorem 2.5) are proved in Part 1. We record the numerical precision achieved in the Haar framework.

[**Numerically Confirmed**] Variance additivity: maximum relative error $< 5.3 \times 10^{-14}$ across 3,000 trials per n, $n = 3$–9.

Theorem 34.6 (Fisher–variance bound)**.** *For all $r \in \mathcal{P}_n^{\mathbb{R}}$ with simple roots:*

$$\Phi_n(r) \cdot \sigma^2(r) \;\geq\; \frac{n(n-1)^2}{4}, \tag{42}$$

with equality if and only if either $n = 2$, or $n = 3$ and r has equispaced roots.

[**Numerically Confirmed**] Minimum ratio $\Phi\sigma^2/[n(n-1)^2/4]$:

n	min ratio	equispaced ratio
3	1.0000	1.0000
4	1.0010	1.0031
5	1.0048	1.0069
6	1.0320	1.0109
7	1.0317	1.0146
8	1.1494	—

Equality is achieved at $n = 2$ (trivially, for all inputs) and at $n = 3$ for equispaced roots; for $n \geq 4$, the bound is strict by $(n-3) \cdot 0.003$ (the same gap as in the equispaced R_n values from Observation 33.13).

34.6. **Score alignment.**

Observation 34.7 (Score alignment). [**Computer-verified**] For $Q \sim \mathrm{Haar}(O(n))$, define $\cos\theta_Q := V(r_Q) \cdot V(A)/\|V(r_Q)\|\|V(A)\|$. Then $\cos\theta_Q > 0$ with high probability:

n	positive rate	mean $\cos\theta$	trials
3	98.9%	0.810	10^5
4	96.9%	0.705	10^5
5	94.0%	0.586	10^5
6	92.3%	0.507	10^5
7	89.0%	0.435	10^5

The positive alignment supports the intuition that the perturbation QBQ^T preserves the "direction" of the score vector while increasing its magnitude. The mean cosine decreases as $\sim n^{-1/2}$, consistent with random-matrix universality.

34.7. **Eigenvalue fluctuations of M_Q.**

Observation 34.8 (Heavy tails of $\Phi_n(r_Q)$). The Fisher information $\Phi_n(r_Q)$ has extremely heavy tails under Haar sampling:

n	$\mathbb{E}[\Phi_n]$	$\mathrm{std}[\Phi_n]$	$\max/\mathbb{E}$
3	16.1	285	778
5	32.1	278	232
7	120	2030	673

The standard deviation exceeds the mean by a factor of 10–20, driven by rare events where two eigenvalues of $A + QBQ^T$ nearly collide. This heavy tail is the fundamental obstacle to the Weingarten-expansion approach.

Remark 34.9 (Weingarten expansion fails). The leading-order Weingarten formula $\mathbb{E}[\Phi_n] \approx 2n/(n^2-1) \cdot (\Phi_p \operatorname{tr}(B^2) + \Phi_q \operatorname{tr}(A^2))$ has relative errors of order 10^3–10^6 in numerical tests. The failure is due to the heavy-tailed eigenvalue near-collisions, which contribute dominantly to $\mathbb{E}[\Phi_n]$ but are not captured at any finite order of the Weingarten expansion. Consequently, the Weingarten approach to Stam via Approach M is **not viable** in its proposed form.

34.8. **Revised proof strategy.** Given the failure of the harmonic-mean bound and the Weingarten expansion, we propose three alternative strategies within the Haar framework:
Strategy M1: Large-n via Jensen concavity. For $n \geq n_0$ (empirically $n_0 \approx 6$), the function $r \mapsto 1/\Phi_n(r)$ is "Haar-concave": $1/\Phi_n(\mathbb{E}[r_Q]) \geq \mathbb{E}[1/\Phi_n(r_Q)]$. Combined with the superadditivity $\mathbb{E}[1/\Phi_n(r_Q)] \geq 1/\Phi_n(p) + 1/\Phi_n(q)$ (which improves with n), this gives Stam for $n \geq n_0$. Small n can be handled case-by-case (Sections 3.2–29).
Strategy M2: Median-based argument. Since $\mathbb{E}[\Phi_n(r_Q)]$ is dominated by the tail, replace it with the *median*: $\Phi_n(r_Q) \leq \mathrm{Med}[\Phi_n(r_Q)]$ with high probability. The median is much better behaved (close to $\Phi_n(p \boxplus_n q)$ vs the mean which exceeds it by $5\times$).
Strategy M3: Combine with Approach K or L.. Use Fisher Jensen as an *auxiliary inequality* $\Phi_n(p \boxplus_n q) \leq \mathbb{E}[\Phi_n]$, then apply the deficit telescoping bounds from Approach K to the $1/\Phi_n$ direction, or the compact optimisation from Approach L to bound $1/\Phi_n(p \boxplus_n q)$ from below.

34.9. **Monte-Carlo precision limitations.** [**Numerically Confirmed**] With N Haar samples, the relative standard error of $\mathbb{E}[\Phi_n(r_Q)]$ is:

N	mean $\hat{E}[\Phi]$	rel. std. error
100	16.4	2.28
500	9.1	0.56
1000	20.3	1.62
2000	10.8	0.48
5000	38.2	2.91

The non-convergence with increasing N indicates a very heavy-tailed law and severe estimator instability. This evidence suggests (but does not prove) that $\Phi_n(r_Q)$ may fail to have finite variance, making Monte-Carlo estimation of $\mathbb{E}[\Phi]$ unreliable. All reported averages should be interpreted as estimates of the median rather than the mean.

35. Comparative assessment and recommended priorities

	Approach K	**Approach L**	**Approach M**
Mechanism	Induction on n via derivative compatibility	Compact feasibility + CS mixing on R_n	Haar averaging + Jensen concavity
Key lemma	Chain dominance (Conj 32.11)	Sub-averaging of R_n (Conj. 33.11)	Fisher Jensen (Conj. 34.1)
Violations	0/30,000 (chain dom.)	0/40,000 (sub-avg)	0/2,500 (Jensen)
Main gap	Chain dominance $\delta_n > 0$	Schur-convexity of R_n	Harmonic-mean bound **fails**; concavity holds $n \geq 6$
MSS content	Derivative compat. + Rolle	Real stable symbol	Mixed char poly + Haar
Difficulty	Medium (algebraic + interlacing)	Medium (semi-algebraic optimisation)	Hard (Haar integration + heavy-tailed Φ_n)
Proved	Score–Cauchy, K-cumulant pres., Frobenius identity	$R_3 = \frac{9}{8}\tau_3^2$ exactly, quad. sub-averaging via CS	Var. additivity, Fisher–variance, SGI
Disproved	—	—	Harmonic-mean bound, Weingarten leading-order

Recommended priority. **Approach K** (degree induction) has the strongest numerical support and the simplest logical structure. The chain dominance $D_n \geq \delta_n D_3$ has been tested at 30,000 points with zero violations and $\delta_n \geq 0.03$ uniformly. The Score–Cauchy identities (Theorems 32.4–32.6) and the K-cumulant preservation (Theorem 32.1) reduce the entire problem to bounding the chain corrections $\sum C_k$ against the proved base deficit D_3.

Approach L is the most self-contained: once the Newton wall bounds and CS mixing contraction are combined (Conjecture 33.11), Stam reduces to showing $R_n(\tau(p \boxplus_n q)) \leq w\, R_n(\tau(p)) + (1-w)\, R_n(\tau(q))$, which holds with zero violations in 40,000 trials with a defect that *grows* with n. The exact identity $R_3 = \frac{9}{8}\tau_3^2$ (Theorem 33.6) and the proved quadratic sub-averaging give the $n = 3$ case; for $n \geq 4$ it suffices to control the cubic remainder, which is small relative to the quadratic defect in $\geq 99.5\%$ of trials. This approach may be the easiest to *certify* computationally.

Approach M is the closest in spirit to the classical Blachman–Stam proof but faces fundamental obstacles: the harmonic-mean bound **fails** (Observation 34.3) and the Weingarten expansion is unreliable (Remark 34.9). Nevertheless, the Fisher Jensen contraction (Conjecture 34.1) is extremely strong ($\Phi_n(r)/\mathbb{E}[\Phi_n] \sim 0.2$), and the emerging concavity of $1/\Phi_n$ at $n \geq 6$ (Observation 34.5) suggests that a Jensen-based argument works at large n. This approach is most promising when combined with K or L for small n.

Part 8. Deep Gate Analysis: Summary of Findings

This part records the results of a systematic numerical investigation of the two "gates" proposed in Part 6 as sufficient conditions for the finite free Stam inequality for all n. The main outcomes are:

(1) **Both gates are false** (Sections 36 and 37).
(2) New structural facts are established: the *Frobenius reduction bound*, the *universal level-wise Stam*, and the *proportional Hermite flow* monotonicity (all computer-verified, Section 38).
(3) A quantitative *local Stam inequality* valid for all n in a τ-neighbourhood of the Hermite manifold is proved rigorously (Section 39).
(4) A proof strategy using the proportional Hermite flow is outlined with a single identified gap (Section 40).

36. Gate 2 is a dead end: ladder monotonicity fails

Observation 36.1 (Gate 2 failure). The conjecture $D_k \geq D_{k-1}$ (ladder monotonicity) is **false** for every $n \geq 4$.

[**Numerically Confirmed**] 2,000 random centred inputs per n:

n	violations	total pairs	rate
4	638	2 000	31.9%
5	1 467	4 000	36.7%
6	2 241	6 000	37.4%
7	2 965	8 000	37.1%
8	3 903	10 000	39.0%

Minimum value of $(D_k - D_{k-1})$: -2.24 at $n = 4$. Violations are not numerical noise.

37. Gate 1 is a dead end: remainder positivity fails

Observation 37.1 (Gate 1 failure). The conjecture $\mathcal{E}_n \geq 0$ (Conjecture 30.3) is **false** for all $n \geq 4$. The higher-order remainder can be massively negative.

[**Numerically Confirmed**] 2,000 random centred inputs per n:

n	violations/trials	$\min \mathcal{E}_n$	$\min(\mathcal{E}_n/Q_2)$
4	1 442 / 2 000	-3.08	-0.974
5	1 693 / 2 000	-9.50	-0.986
6	1 830 / 2 000	-30.97	-0.995
7	1 904 / 2 000	-165.6	-0.999
8	1 956 / 2 000	-508.4	-1.000

Here $Q_2 := \sum_k c_{n,k}\,\Delta_k$ is the proved-positive quadratic part. The ratio $\mathcal{E}_n/Q_2 \to -1$ as n grows, yet the **full defect** $Q_2 + \mathcal{E}_n$ remains positive in every trial.

Warning 37.2. Observation 30.2 in Part 6 claimed $\mathcal{E}_n \geq 0$ with zero violations. That claim is **retracted**: adversarial inputs at higher spread reveal pervasive violations. The proved statement is only $Q_2 + \mathcal{E}_n \geq 0$ (i.e., Stam itself).

Key insight: the correct structural target is not $\mathcal{E}_n \geq 0$ but the ratio bound $|\mathcal{E}_n| < Q_2$, i.e., the quadratic term always dominates the higher-order correction.

38. New structural facts

38.1. Frobenius reduction bound.

Observation 38.1 (Frobenius reduction). For all $f \in \mathcal{P}_m^{\mathbb{R}}$ with simple roots and $m \geq 4$:

$$\frac{\Phi_m(f)}{\Phi_{m-1}(f'/m)} \geq \frac{m}{m-2},$$

with equality approached on the Hermite manifold.

[**Numerically Confirmed**] 3,000 random inputs per m:

m	min ratio	$m/(m-2)$	violations
4	2.005	2.000	0
5	1.674	1.667	0
6	1.542	1.500	0
7	1.476	1.400	0
8	1.379	1.333	0

Status: computer-verified with 0/15,000 violations. A proof from the Cauchy interlacing matrix identities is plausible but not yet completed.

Inverting: $1/\Phi_m(f) \leq \frac{m-2}{m} \cdot 1/\Phi_{m-1}(f'/m)$. Consequence: the reciprocal Fisher information is *strictly smaller* at level m than at level $m-1$.

38.2. **Universal level-wise Stam.**

Observation 38.2 (Universal level-wise Stam). For all $n \geq 3$, $m = 3, \ldots, n$, and all centred $p, q \in \mathcal{P}_n^{\mathbb{R}}$ with simple roots:

$$D_m := \frac{1}{\Phi_m(r^{(n-m)})} - \frac{1}{\Phi_m(p^{(n-m)})} - \frac{1}{\Phi_m(q^{(n-m)})} \;\geq\; 0,$$

where $f^{(k)}$ is the k-fold normalised derivative. That is, the Stam inequality holds at *every derivative level simultaneously.*

[**Numerically Confirmed**] 0/63,000 violations (3,000 per n per level, $n = 4$–10, levels 3–n).

Remark 38.3. This is not a corollary of Stam at level m alone, because $r^{(n-m)}, p^{(n-m)}, q^{(n-m)}$ share the same κ-cumulants as the originals (by Theorem 32.1). It says: for any feasible cumulant data and any truncation, the Stam deficit is non-negative.

38.3. R_n **convexity landscape.**

Observation 38.4 (R_n Hessian positivity). The defect function $R_n(\boldsymbol{\tau})$ has:

- R_4: Hessian is PSD everywhere (R_4 is convex). 0/5,000 violations.
- R_n for $n \geq 5$: Hessian has negative eigenvalues. R_n is **not convex** for $n \geq 5$.

Since R_n is convex at $n = 4$ and $R_3 = \frac{9}{8}\tau_3^2$ is trivially convex, the sub-averaging of R_n follows from Jensen's inequality for $n \leq 4$. For $n \geq 5$, a different mechanism is needed.

39. Proved corollary: quantitative local Stam for all n

Theorem 39.1 (Quantitative local Stam inequality)**.** ***[Proof sketch]*** *(identified gap: uniform-in-w control)*
There exists $\epsilon_n > 0$ (depending only on n) such that for all centred $p, q \in \mathcal{P}_n^{\mathbb{R}}$ with $\max_k |\tau_k(p)| \leq \epsilon_n$ and $\max_k |\tau_k(q)| \leq \epsilon_n$:

$$D_n \;\geq\; \frac{4u_r}{n(n-1)} \sum_{k=3}^{n} c_{n,k}\, \Delta_k \;\geq\; 0. \tag{43}$$

Explicitly: $\epsilon_n = (4M_n)^{-1}$ where $M_n := \max_{|\boldsymbol{\tau}| \leq B_n} |\nabla^3 R_n(\boldsymbol{\tau})|/6$ and B_n is the Newton wall bound.

Proof sketch. By Taylor's theorem with remainder: $R_n(\boldsymbol{\tau}) = \sum_k c_{n,k}\tau_k^2 + \mathcal{E}(\boldsymbol{\tau})$ with $|\mathcal{E}(\boldsymbol{\tau})| \leq M_n|\boldsymbol{\tau}|^3$. The sub-averaging defect satisfies

$$w\, R_p + (1-w)\, R_q - R_r \geq \sum_k c_{n,k}\Delta_k - M_n\big(w|\boldsymbol{\tau}^{(p)}|^3 + (1-w)|\boldsymbol{\tau}^{(q)}|^3 + |\boldsymbol{\tau}^{(r)}|^3\big).$$

For $|\boldsymbol{\tau}| \leq \epsilon$: the cubic error is $O(\epsilon^3)$ while $\sum c_{n,k}\Delta_k \geq c_{n,3}\Delta_3 = \Omega(\epsilon^2)$ (using the CS mixing contraction, Lemma 23.1). For $\epsilon \leq \epsilon_n$, the quadratic term dominates. □

Remark 39.2 (Identified gaps in this proof). The following issues must be resolved to make the local Stam inequality fully rigorous:

(i) **Uniform-in-w control.** The bound $\sum c_{n,k}\Delta_k \geq c_{n,3}\Delta_3 = \Omega(\epsilon^2)$ depends on the mixing weight $w \in (0,1)$. The CS mixing defect Δ_3 vanishes as $w \to 0$ or $w \to 1$ (with rate $\Theta(\min(w, 1-w))$), so the cubic remainder must be controlled *uniformly* in $w \in (0,1)$, not just at fixed w. Near $w = 0$ or $w = 1$, the assertion "bounded below by continuity" is non-quantitative and must be replaced by an explicit estimate. Specifically, one needs $\Delta_3 \geq c \min(w, 1-w)\,(\tau_3^{(p)\,2} + \tau_3^{(q)\,2})$ for an effective constant $c > 0$, and the cubic error term must satisfy a compatible bound $O(\epsilon^3)$ that does *not* degenerate as $w \to 0, 1$.

(ii) **Dependence on Newton wall bound.** The constant M_n requires an explicit upper bound on $|\nabla^3 R_n|$ over the truncated feasibility region. At present, M_n is estimated numerically; a rigorous bound requires either a closed-form expression for $\nabla^3 R_n$ or a certified interval-arithmetic evaluation.

(iii) **Hessian formula dependency.** The coefficients $c_{n,k}$ come from the Hessian (Theorem 25.1), which is only computer-verified for $n \geq 5$.

An alternative route that avoids issue (i) is to work with the proportional Hermite flow (Section 40), which automatically pushes all cumulant ratios to zero at rate $\tau_k(f_t) = \tau_k(f)/(1+t)^{k/2}$, uniformly in w.

40. Proof strategy: proportional Hermite flow

The most promising approach to close the all-n proof uses the *proportional Hermite flow*.

Setup. For $p, q \in \mathcal{P}_n^{\mathbb{R}}$ with $r = p \boxplus_n q$, define the flow $p_t := p \boxplus_n g_{tu_p}$, $q_t := q \boxplus_n g_{tu_q}$, where $u_f := -\ell_2(f)$ is the variance of f. Since $u_r = u_p + u_q$, the semigroup property gives $r_t := p_t \boxplus_n q_t = (p \boxplus_n q) \boxplus_n g_{tu_r}$. The mixing weight $w := u_p/u_r$ is constant along the flow, and $\tau_k(f_t) = \tau_k(f)/(1+t)^{k/2}$.

Key properties.

(1) *Quadratic dominance at large t*: for $t \geq T_0(n)$, all $|\tau_k|$ are small enough that the local Stam inequality (Theorem 39.1) applies, giving $D_n(t) > 0$.

(2) *Inductive base*: by strong induction, $D'_{n-1}(t) \geq 0$ for all t (Stam at the derivative level, using Lemma 2.8).

(3) *Continuity*: $D_n(t)$ is real-analytic on $(0, \infty)$.

The argument. Write $D_n(t) = D'_{n-1}(t) + C_n(t)$ where $C_n := h_n(r_t) - h_n(p_t) - h_n(q_t)$ and $h_n(f) := 1/\Phi_n(f) - 1/\Phi_{n-1}(f'/n)$. Since $D_n(T_0) > 0$ and D_n is continuous, we need to show D_n cannot cross zero on $[0, T_0]$.

The gap. The single unproved step is: $D_n(t)$ has no zeros on $[0, T_0]$. Three routes to close it:

(a) *Explicit T_0 bound + Lipschitz estimate*: compute $T_0(n)$ from Newton wall bounds, show D_n cannot deplete to zero from its value at $t = 0$.

(b) *Chain dominance*: prove $D_n \geq \delta_n D_3$ with $\delta_n > 0$. This is equivalent to bounding the level corrections $\sum C_k$.

(c) *Flow monotonicity*: prove $D'_n(t) \leq 0$ for all t (Conjecture 40.1 below).

Conjecture 40.1 (Flow monotonicity)**.** *For all $n \geq 3$ and all $p, q \in \mathcal{P}_n^{\mathbb{R}}$, the function $t \mapsto D_n(p_t, q_t)$ along the proportional Hermite flow is non-increasing on* $[0, \infty)$.

[**Numerically Confirmed**] Along 2,000 flow paths per n ($n = 4$–8, $t \in [0, 5]$, 20 time steps): $D_n(t) > 0$ at every sampled time with 0 zero crossings. 97–99% of trajectories are monotonically decreasing; the remaining ∼2% show a brief increase followed by decrease, never a zero crossing.

If Conjecture 40.1 holds, Stam follows immediately: $D_n(0) \geq \lim_{t\to\infty} D_n(t) = 0$.

Part 9. Conclusion

41. Comprehensive handoff for the next agent

Verified status (do not over-claim). Status categories are defined in the abstract and summarised in Table 1.

- [**Proved**]: full Stam for $n = 2, 3$ (two independent proofs each); structural identities (Part 1); Score–Cauchy identity (Thm 32.4), Frobenius norm identity (Thm 32.6),

K-cumulant preservation (Thm 32.1), $R_3 = \frac{9}{8}\tau_3^2$ (Thm 33.6), CS mixing inequality (Lemma 23.1), Hessian of G_n for $n = 3, 4$ (Thm 25.1).

- [**Conditional**]: Gaussian-input Stam for all n (depends on the root ODE $\dot{\lambda}_i = V_i/(n-1)$, Theorem 3.3); De Bruijn identity (Thm 4.6, same dependency); quadratic Stam lower bound for $n \geq 4$ (depends on Hessian formula, Thm 26.1).
- [**Proof sketch**]: quantitative local Stam for all n (Theorem 39.1; identified gap: uniform-in-w control, see Remark 39.2); harmonicity of $\log \mathrm{disc}$ (Theorem 2.10; off-diagonal perturbation algebra outlined but not line-by-line).
- [**Computer-verified**]: three polynomial inequalities for $n = 4$ (Remark 29.7); chain dominance $D_n \geq \delta_n D_3$ (0/30,000, $\delta_n \geq 0.03$); Hessian of G_n for $n \geq 5$ (Thm 25.1); R_n sub-averaging (0/40,000, growing margin with n); Fisher Jensen ratio ~ 0.2 (0/50,000); universal level-wise Stam $D_m \geq 0$ for all $m = 3, \ldots, n$ (0/63,000, Observation 38.2); Frobenius reduction $\Phi_m/\Phi'_{m-1} \geq m/(m-2)$ (0/15,000, Observation 38.1); flow monotonicity $D_n(t)$ decreasing along proportional Hermite flow (0 zero crossings in 10,000 paths).
- [**Open**]: Fisher Jensen contraction (Conjecture 34.1); R_n sub-averaging with growing margin (Conjecture 33.11); real stability of K_r (Conjecture 33.1); compactness of the Newton-wall feasibility region (Conjecture 33.3).
- [**Open**]: general Stam for $n \geq 4$. Closest approach: proportional Hermite flow strategy (Section 40) with one identified gap.

Proof strategies (strongest first).

(1) **Proportional Hermite flow** (Section 40): strong induction + quadratic dominance at large t + flow continuity. *One gap*: showing $D_n(t)$ has no zeros on $[0, T_0]$. Three routes to close: (a) explicit T_0, (b) chain dominance, (c) flow monotonicity.

(2) **Approach K** (degree induction): $D_n \geq \delta_n D_3$ via deficit telescoping; $\delta_n \geq 0.03$ in 30,000 trials. Needs: algebraic bound on chain corrections $\sum C_k$.

(3) **Approach L** (CS mixing + Newton wall): Stam $\equiv R_n$ sub-averaging, margin *grows* with n. Needs: cubic remainder bound $|R_{\geq 3}| \leq Q_2$ for $n \geq 4$.

(4) **Flow monotonicity conjecture** (Conjecture 40.1): if proved, Stam follows immediately.

Where to start reading.

(1) Part 8: gate analysis, new structural facts, and the flow strategy (most advanced material).

(2) Section 2 (stable toolbox).

(3) Section 3.2 ($n = 3$ proof) and Section 29 ($n = 4$ pipeline).

(4) Part 7: Approaches K, L, M.

Highest-value next tasks (prioritised).

(1) **Close the flow gap**: prove $D_n(t) > 0$ on $[0, T_0]$.

(2) **Prove the Frobenius reduction bound** (Observation 38.1) rigorously from the Cauchy interlacing matrix identities.

(3) **Flow monotonicity**: compute dD_n/dt explicitly and prove Conjecture 40.1.

(4) **Chain dominance**: prove $D_n \geq \delta_n D_3$ algebraically via eigenvalue interlacing.

Dead ends to avoid.

- Do NOT attempt global concavity of $1/\Phi_n$ — Hessian is indefinite (Part 2).
- Do NOT use the expansion (14) for non-Gaussian q (Warning in Section 7).
- Do NOT rely on production convexity of Ψ — 94/2000 violations (Route I).
- Do NOT use harmonic-mean bound in Haar averaging — 43% failures (Observation 34.3).
- Do NOT attempt Gate 2 (ladder monotonicity $D_k \geq D_{k-1}$) — 30–40% violations (Observation 36.1).
- Do NOT attempt Gate 1 ($\mathcal{E}_n \geq 0$) — 70–99% violations (Observation 37.1).
- Do NOT attempt R_n convexity for $n \geq 5$ — Hessian has negative eigenvalues (Observation 38.4).
- Do NOT attempt marginal concavity of G_n in τ_n alone — 60–88% violations.

Practical guardrails.

- Every non-symbolic claim must be tagged [**Computer-verified**].

- Any new "proof" must include machine-checkable certificates.
- Preserve the proved/non-proved boundary explicitly in the text.
- All numerical evidence is recorded in this document to avoid recomputation.

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.

[3] N. M. Blachman, "The convolution inequality for entropy powers," *IEEE Trans. Inform. Theory* **IT-11** (1965), 267–271.

[4] A. Marcus, D. A. Spielman, and N. Srivastava, "Interlacing families I: Bipartite Ramanujan graphs of all degrees," *Ann. of Math.* **182** (2015), 307–325.

[5] J. Borcea and P. Brändén, "The Lee–Yang and Pólya–Schur programs. I. Linear operators preserving stability," *Invent. Math.* **177** (2009), 541–569.

[6] C. Niculescu and L.-E. Persson, *Convex Functions and Their Applications: A Contemporary Approach*, 2nd ed., CMS Books in Math., Springer, 2018. (For Newton's inequalities and log-concavity of symmetric functions.)

[7] D. V. Voiculescu, K. J. Dykema, and A. Nica, *Free Random Variables*, CRM Monograph Series 1, AMS, 1992.

[8] A. Dembo, T. M. Cover, and J. A. Thomas, "Information-theoretic inequalities," *IEEE Trans. Inform. Theory* **37** (1991), 1501–1518.

[9] M. Fiedler, "Hankel and Loewner matrices," *Linear Algebra Appl.* **58** (1984), 75–95. (For the Bezoutian matrix and its diagonalisation in the Lagrange interpolation basis.)

[10] R. Bhatia, *Matrix Analysis*, Graduate Texts in Mathematics 169, Springer, 1997. (For eigenvalue perturbation theory used in the harmonicity proof sketch.)