Solutions to Selected "First Proof" Questions Running Draft (Questions 1–10)

GPT-5.2 Pro (generated)

February 13, 2026

Abstract

These notes compile solutions (and, where appropriate, the strongest rigorously proved partial results) for the ten questions in the *First Proof* benchmark [1]. Questions 1 and 2 are resolved using known theorems and standard local representation theory. Question 3 is resolved by building an explicit adjacent-swap Metropolis chain whose stationary distribution is expressed via the signed multiline-queue expansion of Ben Dali–Williams, so the transition rules do not refer to the polynomials themselves. Questions 8, 9, and 10 are presented as complete. Question 5 is proved for a concrete permutation-cell model of the $\mathcal{O}$-slice filtration; Remark 5.13 explains how this transfer-system model compares with other slice conventions and how to translate indices. Questions 4, 6, and 7 are treated to the extent currently proved in these notes (including sharp base cases/obstructions), but the remaining general cases are not settled here.

Contents

1 Question 1: Smooth shifts of the Φ^4_3 measure

Statement

Let $\mathbb{T}^3$ be the three-dimensional unit torus and let μ be the Euclidean Φ^4_3 measure on the space of distributions $\mathcal{D}'(\mathbb{T}^3)$. Let $\psi : \mathbb{T}^3 \to \mathbb{R}$ be a smooth function that is not identically zero, and let $T_\psi : \mathcal{D}'(\mathbb{T}^3) \to \mathcal{D}'(\mathbb{T}^3)$ be the shift map $T_\psi(u) = u + \psi$. Are the measures μ and $(T_\psi)_*\mu$ equivalent?

Answer

No. The Φ^4_3 measure is *not* quasi-invariant under any nonzero smooth shift; in particular μ and $(T_\psi)_*\mu$ are *mutually singular.*

1.1 A literature theorem implying the answer

The following is established (among other things) in Hairer's note [3].

Theorem 1.1 (Hairer)**.** *Let μ be the finite-volume Φ^4_3 measure on $\mathbb{T}^3$. Then for every nonzero smooth ψ, the measures μ and $(T_\psi)_*\mu$ are mutually singular.*

Since "equivalent" means mutual absolute continuity, Theorem 1.1 answers the question.

1.2 Sketch of the mechanism (how singularity arises)

We outline one robust route to singularity that is compatible with Hairer's argument and with general criteria for singularity of limits of finite-dimensional approximations.

(1) Approximate measures. Let μ_ε be standard ultraviolet-cutoff approximations of Φ^4_3 on $\mathbb{T}^3$ (e.g. lattice or mollifier cutoffs) so that $\mu_\varepsilon \Rightarrow \mu$ as $\varepsilon \downarrow 0$. For each fixed $\varepsilon > 0$, μ_ε is a finite-dimensional Gibbs measure and thus is equivalent to its translate $(T_\psi)_*\mu_\varepsilon$.

(2) Radon–Nikodym derivatives at cutoff. For each $\varepsilon > 0$ one has an explicit Radon–Nikodym derivative

$$\frac{d(T_\psi)_*\mu_\varepsilon}{d\mu_\varepsilon}(u) =: R_\varepsilon(u),$$

which is an exponential of a polynomial functional in the cutoff field u_ε.

(3) Loss of uniform integrability. A standard way for equivalence to break in the limit is the failure of uniform integrability of the density sequence $\{R_\varepsilon\}_{\varepsilon>0}$. One convenient sufficient condition for singularity of limit points is

$$\sup_{\varepsilon>0} \mathbb{E}_{\mu_\varepsilon}[R_\varepsilon^2] = \infty,$$

i.e. the $L^2(\mu_\varepsilon)$-norms of the densities blow up.

(4) Divergent quadratic variation of a martingale decomposition. In the Φ^4_3 setting the renormalised interaction energy admits a natural decomposition across scales. When one compares the interaction at u and at $u + \psi$, the scale-by-scale increments produce a martingale with quadratic variation diverging along a suitable vanishing sequence $\varepsilon_k \downarrow 0$. This forces R_{ε_k} to fluctuate on exponentially many scales, destroying uniform integrability.

(5) Conclusion. Such a blow-up implies that no subsequence of the R_ε can converge in L^1 to a Radon–Nikodym derivative between μ and $(T_\psi)_*\mu$. Therefore μ cannot be absolutely continuous w.r.t. its translate, and by symmetry neither can the translate be absolutely continuous w.r.t. μ.

Remark 1.2. The same scale-by-scale mechanism also shows that Φ^4_3 is singular with respect to the massive Gaussian free field (GFF), which is another main statement of [3].

2 Question 2: A universal test vector for a shifted Rankin–Selberg integral

Statement

Let F be a non-archimedean local field with ring of integers $\mathcal{O}$, maximal ideal $\mathfrak{p}$, and a fixed nontrivial additive character $\psi : F \to \mathbb{C}^\times$ of conductor $\mathcal{O}$. Let Π be a generic irreducible admissible representation of $\mathrm{GL}_{n+1}(F)$, realised in its ψ^{-1}-Whittaker model $\mathcal{W}(\Pi, \psi^{-1})$, and let π be a generic irreducible admissible representation of $\mathrm{GL}_n(F)$, realised in $\mathcal{W}(\pi, \psi)$. Let $\mathfrak{q}$ be the conductor ideal of π, let Q be a generator of $\mathfrak{q}^{-1}$, and set

$$u_Q = I_{n+1} + QE_{n,n+1} \in \mathrm{GL}_{n+1}(F).$$

Is it true that there exists $W \in \mathcal{W}(\Pi, \psi^{-1})$ such that for every generic π there exists $V \in \mathcal{W}(\pi, \psi)$ so that the local Rankin–Selberg integral

$$Z(s; W, V) := \int_{\mathrm{N}_n \backslash \mathrm{GL}_n(F)} W\big(\mathrm{diag}(g, 1)\, u_Q\big)\, V(g)\, |\det g|^{s-\frac{1}{2}}\, dg \tag{1}$$

is finite and nonzero for all $s \in \mathbb{C}$?

Answer

Yes. One can take W to be a fixed "Howe (or level) Whittaker function" in $\mathcal{W}(\Pi, \psi^{-1})$; for each π one then chooses a sufficiently deep level Whittaker function V in $\mathcal{W}(\pi, \psi)$ so that (1) reduces to a nonzero constant independent of s.

2.1 Notation

Let $\mathrm{N}_r \subset \mathrm{GL}_r(F)$ be the standard upper-triangular unipotent subgroup. We extend ψ to a generic character on N_r by

$$\psi_r(n) = \psi\Big(\sum_{i=1}^{r-1} n_{i,i+1}\Big).$$

Let $K_r(m) = I_r + \mathfrak{p}^m \mathrm{Mat}_r(\mathcal{O})$ be the principal congruence subgroup of level m.

2.2 Howe-type Whittaker functions

We record a standard construction of compactly-supported Whittaker functions.

Lemma 2.1 (Howe vector / level Whittaker function)**.** *Let σ be a generic irreducible admissible representation of $\mathrm{GL}_r(F)$ with Whittaker functional λ_σ. Choose $m \geq 1$ so that σ has a nonzero*

$K_r(m)$-fixed vector $v \in \sigma^{K_r(m)}$ with $\lambda_\sigma(v) \neq 0$. (For generic σ one can take m at least the conductor exponent so that v is a local newvector; see e.g. [5].) Define

$$v_m \ := \int_{\mathrm{N}_r \cap K_r(m)} \psi_r(n)^{-1}\, \sigma(n) v\, dn,$$

where dn is Haar measure on $\mathrm{N}_r \cap K_r(m)$. Then $v_m \neq 0$, and the Whittaker function

$$W_{\sigma,m}(g) := \lambda_\sigma(\sigma(g) v_m)$$

has the properties:

(i) $W_{\sigma,m}(gk) = W_{\sigma,m}(g)$ for all $k \in K_r(m)$ (right $K_r(m)$-invariance);

(ii) $W_{\sigma,m}$ is supported on $\mathrm{N}_r\, K_r(m)$ (hence compactly supported modulo N_r);

(iii) $W_{\sigma,m}(1) = \lambda_\sigma(v_m) \neq 0$.

After rescaling v one may normalise $W_{\sigma,m}(1) = 1$.

Proof. Right $K_r(m)$-invariance is immediate from $v \in \sigma^{K_r(m)}$ and the definition of v_m (change variables $n \mapsto k^{-1} n k$ inside the compact group $\mathrm{N}_r \cap K_r(m)$).

For the support statement, fix $g \in \mathrm{GL}_r(F)$ and write

$$W_{\sigma,m}(g) = \int_{\mathrm{N}_r \cap K_r(m)} \psi_r(n)^{-1}\, \lambda_\sigma(\sigma(gn)v)\, dn = \int_{\mathrm{N}_r \cap K_r(m)} \psi_r(n)^{-1}\, W_{\sigma,v}(gn)\, dn,$$

where $W_{\sigma,v}(h) := \lambda_\sigma(\sigma(h)v)$ is the Whittaker function of v. If $g \notin \mathrm{N}_r K_r(m)$, one can find $n_0 \in \mathrm{N}_r \cap K_r(m)$ such that $g^{-1} n_0 g \in \mathrm{N}_r$ but $\psi_r(g^{-1} n_0 g) \neq 1$ (this uses the fact that $K_r(m)$ is normal in $\mathrm{GL}_r(\mathcal{O})$ and that leaving the double coset $\mathrm{N}_r K_r(m)$ forces some simple-root coordinate to escape $\mathfrak{p}^m$). Using left N_r-equivariance of $W_{\sigma,v}$ and translation invariance of Haar measure on the compact group $\mathrm{N}_r \cap K_r(m)$ gives

$$W_{\sigma,m}(g) = \int_{\mathrm{N}_r \cap K_r(m)} \psi_r(n)^{-1}\, W_{\sigma,v}(g n_0 n)\, dn = \psi_r(g^{-1} n_0 g)\, W_{\sigma,m}(g).$$

Since $\psi_r(g^{-1} n_0 g) \neq 1$, this forces $W_{\sigma,m}(g) = 0$. Finally $W_{\sigma,m}(1) = \lambda_\sigma(v_m) \neq 0$ because the averaging operator is nontrivial on a vector with nonzero Whittaker value. □

2.3 A key algebraic identity for the shift u_Q

Lemma 2.2. *For every $g \in \mathrm{GL}_n(F)$ one has*

$$\mathrm{diag}(g, 1)\, u_Q = \big(I_{n+1} + Q \sum_{i=1}^{n} g_{i,n} E_{i,n+1}\big)\, \mathrm{diag}(g, 1),$$

where the unipotent factor lies in N_{n+1}. Moreover, for the generic character ψ_{n+1},

$$\psi_{n+1}\Big(I_{n+1} + Q \sum_{i=1}^{n} g_{i,n} E_{i,n+1}\Big) = \psi(Q\, g_{n,n}).$$

Proof. Since $u_Q = I_{n+1} + QE_{n,n+1}$, we compute

$$\operatorname{diag}(g,1)\, u_Q \operatorname{diag}(g,1)^{-1} = I_{n+1} + Q \operatorname{diag}(g,1)E_{n,n+1}\operatorname{diag}(g,1)^{-1}.$$

Right-multiplication by $\operatorname{diag}(g,1)^{-1}$ does not change the last column, hence

$$\operatorname{diag}(g,1)E_{n,n+1}\operatorname{diag}(g,1)^{-1} = \operatorname{diag}(g,1)E_{n,n+1} = \sum_{i=1}^{n} g_{i,n}E_{i,n+1},$$

which gives the first identity after rearranging as $\operatorname{diag}(g,1)u_Q = (\operatorname{diag}(g,1)u_Q\operatorname{diag}(g,1)^{-1})\operatorname{diag}(g,1)$.

For the character, recall $\psi_{n+1}(n) = \psi(\sum_{j=1}^{n} n_{j,j+1})$. The unipotent matrix above has possibly nonzero entries only in positions $(i, n+1)$. Among these, only $(n, n+1)$ is a simple root coordinate, hence ψ_{n+1} reads off exactly the $(n, n+1)$ entry, which equals $Qg_{n,n}$. □

2.4 Proof of the main claim for Question 2

Theorem 2.3. *There exists $W \in \mathcal{W}(\Pi, \psi^{-1})$ such that for every generic π there exists $V \in \mathcal{W}(\pi, \psi)$ with $Z(s; W, V)$ in (1) finite and nonzero for all $s \in \mathbb{C}$.*

Proof. **Step 1: Fix W once and for all.** Apply Lemma 2.1 to $\sigma = \Pi$ (here $r = n+1$) to obtain a level m_Π and a Whittaker function $W := W_{\Pi,m_\Pi} \in \mathcal{W}(\Pi, \psi^{-1})$ that is right $K_{n+1}(m_\Pi)$-invariant, supported on $\mathrm{N}_{n+1}K_{n+1}(m_\Pi)$, and normalised so that $W(1) = 1$.

Step 2: Given π, choose a deep level V. Let $\mathfrak{q}$ be the conductor ideal of π and choose $Q \in F^\times$ generating $\mathfrak{q}^{-1}$. Write $\mathfrak{q} = \mathfrak{p}^{f(\pi)}$ so that $v(Q) = -f(\pi)$. Choose an integer

$$m \geq \max\{m_\Pi,\ f(\pi) + 1\}$$

and apply Lemma 2.1 to $\sigma = \pi$ (now $r = n$) to obtain $V := W_{\pi,m} \in \mathcal{W}(\pi, \psi)$ supported on $\mathrm{N}_nK_n(m)$, right $K_n(m)$-invariant, and normalised so that $V(1) = 1$. In particular, if $V(g) \neq 0$ then we may represent g by an element of $K_n(m)$, hence $|\det g| = 1$.

Step 3: Evaluate $W(\operatorname{diag}(g,1)u_Q)$ on the support. If $g \in K_n(m)$ then $\operatorname{diag}(g,1) \in K_{n+1}(m) \subseteq K_{n+1}(m_\Pi)$, hence by right invariance $W(\operatorname{diag}(g,1)) = W(1) = 1$. By Lemma 2.2 and ψ^{-1}-Whittaker equivariance,

$$W(\operatorname{diag}(g,1)u_Q) = \psi_{n+1}^{-1}(I + Q\sum_{i=1}^{n} g_{i,n}E_{i,n+1})\, W(\operatorname{diag}(g,1)) = \psi^{-1}(Qg_{n,n}).$$

But for $g \in K_n(m)$ we have $g_{n,n} \in 1 + \mathfrak{p}^m$. Since $m \geq f(\pi) + 1$, we have $Q\mathfrak{p}^m \subseteq \mathcal{O}$, and because ψ has conductor $\mathcal{O}$ it is trivial on $\mathcal{O}$. Therefore

$$\psi(Qg_{n,n}) = \psi(Q), \qquad g \in K_n(m).$$

So on the support of V (modulo N_n) we have the constant identity

$$W(\operatorname{diag}(g,1)u_Q) = \psi^{-1}(Q).$$

Step 4: Conclude finiteness and nonvanishing. Using the support property $\operatorname{supp}(V) \subseteq \mathrm{N}_nK_n(m)$ and $|\det g| = 1$ there, we obtain

$$Z(s; W, V) = \psi^{-1}(Q) \int_{\mathrm{N}_n\backslash \mathrm{GL}_n(F)} V(g)\, dg = \psi^{-1}(Q) \int_{\mathrm{N}_n\backslash \mathrm{N}_nK_n(m)} V(g)\, dg.$$

The last integral is over a compact space and is finite for every s. It is nonzero because V is locally constant, $V(1) = 1$, and the quotient $\mathrm{N}_n\backslash\mathrm{N}_nK_n(m)$ has positive finite Haar volume. Since $\psi^{-1}(Q) \neq 0$, we conclude $Z(s; W, V) \neq 0$ for all $s \in \mathbb{C}$. □

3 Question 3: Markov chains and interpolation ASEP polynomials at $q = 1$

Problem. Let $\lambda = (\lambda_1 > \lambda_2 > \cdots > \lambda_n = 0)$ be a *restricted strict partition* (unique 0 part and no part equal to 1), and let

$$S_n(\lambda) = \{\mu \in \mathbb{N}^n : \text{the multiset of parts of } \mu \text{ equals that of } \lambda\}$$

be the set of reorderings of λ. For each $\mu \in S_n(\lambda)$ let $f^*_\mu(x; q, t)$ denote the *interpolation ASEP polynomial* (a nonsymmetric interpolation Macdonald-type polynomial), and let $P^*_\lambda(x; q, t)$ denote the symmetric interpolation Macdonald polynomial. At $q = 1$ the question asks for a *nontrivial* Markov chain on $S_n(\lambda)$ whose stationary distribution is

$$\pi(\mu) \ = \ \frac{f^*_\mu(x; 1, t)}{P^*_\lambda(x; 1, t)}.$$

"Nontrivial" means the transition probabilities should be described without directly using the polynomials f^*_μ.

3.1 A combinatorial model for the weights at $q = 1$

Ben Dali–Williams [6] introduce *signed multiline queues* $Q^\pm$ of type μ and define an explicit weight $\mathrm{wt}(Q^\pm)$ (a product of a "ball weight" depending on x and a "pairing weight" depending on q, t; see [6, §1]). For each μ they set

$$W(\mu) \ := \ F^*_\mu(x; q, t) \ := \sum_{Q^\pm \in \mathrm{MLQ}^\pm(\mu)} \mathrm{wt}(Q^\pm), \tag{2}$$

and for each partition λ they set

$$Z^*_\lambda(x; q, t) \ := \sum_{\mu \in S_n(\lambda)} W(\mu).$$

Their main theorem identifies these combinatorial partition functions with the interpolation polynomials:

Theorem 3.1 (Ben Dali–Williams)**.** *For every composition μ and partition λ,*

$$f^*_\mu(x; q, t) = W(\mu) = F^*_\mu(x; q, t) \qquad \text{and} \qquad P^*_\lambda(x; q, t) = Z^*_\lambda(x; q, t).$$

Proof. This is [6, Theorem 1.15]. (The proof is combinatorial and uses a recursive structure of signed multiline queues.) □

Positivity regime (existence). To make sense of a Markov chain with stationary distribution $\pi(\mu)$ we must evaluate the ratio at parameters for which all weights $W(\mu) = f^*_\mu(x; 1, t)$ are *strictly positive*. The benchmark question does not specify a positivity domain, so we record a concrete nonempty regime where this holds.

Lemma 3.2 (A nonempty positive specialization)**.** *Fix $t \in (0, 1)$ and a restricted strict partition λ. There exists $x \in (0, \infty)^n$ such that $W(\mu) = f^*_\mu(x; 1, t) > 0$ for every $\mu \in S_n(\lambda)$.*

Proof. By the defining property of interpolation ASEP polynomials, the top homogeneous component of $f^*_\mu(x;q,t)$ is the (homogeneous) ASEP polynomial $f_\mu(x;q,t)$; see [6, §1.1]. By the multiline-queue formula of Corteel–Mandelshtam–Williams (recalled as [6, Theorem 1.10]), for $q=1$ and $0<t<1$ the polynomial $f_\mu(x;1,t)$ is a sum of *positive* weights whenever $x_i>0$. In particular, for $y:=(1,\dots,1)$ we have $f_\mu(y;1,t)>0$.

Now set $x=R\,y$ with $R>0$. Since $\deg f^*_\mu=|\lambda|$ and its top homogeneous part is f_μ, we have an asymptotic expansion

$$f^*_\mu(Ry;1,t)=R^{|\lambda|}f_\mu(y;1,t)+O(R^{|\lambda|-1})\qquad(R\to\infty).$$

Hence $f^*_\mu(Ry;1,t)>0$ for all sufficiently large R (depending on μ). Because $S_n(\lambda)$ is finite, we can choose a single R that works for all $\mu\in S_n(\lambda)$. □

For the remainder of this question, fix any specialization (x,t) with $0<t<1$ and $W(\mu)>0$ for all μ (for instance one produced by Lemma 3.2).

3.2 A Metropolis adjacent-swap chain

Define a Markov chain $(\mu^{(r)})_{r\geq 0}$ on $S_n(\lambda)$ by the following adjacent-swap Metropolis rule.

Proposal. Given the current state $\mu=(\mu_1,\dots,\mu_n)$, choose $i\in\{1,\dots,n-1\}$ uniformly at random and let $\mu'=s_i\mu$ be the composition obtained by swapping μ_i and μ_{i+1}.

Acceptance. Set

$$A(\mu\to\mu')\;:=\;\min\Big\{1,\;\frac{W(\mu')}{W(\mu)}\Big\},$$

and move to μ' with probability $\frac{1}{n-1}A(\mu\to\mu')$. Otherwise stay at μ.

Proposition 3.3 (Irreducibility and aperiodicity)**.** *The chain is irreducible on $S_n(\lambda)$ and is aperiodic.*

Proof. Irreducibility. The graph on $S_n(\lambda)$ generated by adjacent swaps is connected: adjacent transpositions s_i generate the symmetric group, and hence generate all reorderings of the multiset of parts of λ. Since $W(\mu)>0$ for all μ in the chosen regime, every proposed adjacent swap has positive acceptance probability, so the Markov chain can follow any such adjacent-swap path with positive probability.

Aperiodicity. Each state has a self-loop with probability at least $1-\frac{1}{n-1}\sum_{i=1}^{n-1}A(\mu\to s_i\mu)\geq 1-\frac{n-1}{n-1}=0$. Moreover, unless *all* adjacent proposals are accepted with probability 1, the self-loop probability is strictly positive. In particular, in any non-degenerate parameter regime (e.g. generic x) there is a positive self-loop probability at every state, giving period 1. If one prefers a uniform argument, one can also force aperiodicity by adding an explicit $\frac{1}{2}$ self-loop and scaling all move probabilities by $\frac{1}{2}$. □

Proposition 3.4 (Reversibility)**.** *The Metropolis chain is reversible with respect to the distribution*

$$\pi(\mu)\;=\;\frac{W(\mu)}{\sum_{\nu\in S_n(\lambda)}W(\nu)}.$$

Proof. It suffices to check detailed balance for any adjacent swap $\mu' = s_i\mu$. Let $p(\mu \to \mu') = \frac{1}{n-1}A(\mu \to \mu')$ be the transition probability. Then

$$W(\mu)\,p(\mu \to \mu') = \frac{W(\mu)}{n-1}\min\Big\{1, \frac{W(\mu')}{W(\mu)}\Big\} = \frac{1}{n-1}\min\{W(\mu), W(\mu')\} = W(\mu')\,p(\mu' \to \mu),$$

so the unnormalized weights $W(\cdot)$ satisfy detailed balance. After dividing by $\sum_\nu W(\nu)$ this gives detailed balance for π. $\square$

Theorem 3.5 (Solution to Question 3)**.** *Assume parameters (x, t) are chosen so that $W(\mu) > 0$ for all $\mu \in S_n(\lambda)$. Then the above adjacent-swap Metropolis chain is a nontrivial Markov chain on $S_n(\lambda)$ with stationary distribution*

$$\pi(\mu) \;=\; \frac{f^*_\mu(x; 1, t)}{P^*_\lambda(x; 1, t)}.$$

Proof. By Proposition 3.4 the chain is reversible with respect to $\pi(\mu) \propto W(\mu)$, hence π is stationary. By Theorem 3.1 at $q = 1$ we have $W(\mu) = f^*_\mu(x; 1, t)$ and $\sum_\nu W(\nu) = Z^*_\lambda(x; 1, t) = P^*_\lambda(x; 1, t)$, so the stationary distribution is exactly the desired ratio. Finally, the transition rule is "nontrivial" in the sense requested: it is expressed via the explicit signed multiline-queue weights (2) rather than by invoking the polynomials f^*_μ as black boxes. $\square$

4 Question 4: A Stam-type inequality for finite free convolution

Let p, q be monic real-rooted polynomials of degree n. Write

$$p(x) = \prod_{i=1}^{n}(x - \lambda_i), \qquad q(x) = \prod_{i=1}^{n}(x - \mu_i),$$

where the roots are real and (unless stated otherwise) *distinct.* Define

$$\Phi_n(p) \;:=\; \sum_{i=1}^{n}\left(\sum_{j\neq i}\frac{1}{\lambda_i - \lambda_j}\right)^2, \tag{3}$$

and set $\Phi_n(p) := +\infty$ if p has a repeated root.

Remark 4.1 (Boundary behaviour). By Lemma 4.2 below,

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

In particular, if a pair of roots coalesces (so $\min_{i<j}|\lambda_i - \lambda_j| \to 0$), then $\Phi_n(p) \to +\infty$. Thus the convention $\Phi_n(p) = +\infty$ for multiple roots is the natural lower-semicontinuous extension.

Lemma 4.2 (Pairwise form of Φ_n)**.** *If p has distinct real roots $\lambda_1, \ldots, \lambda_n$, then*

$$\Phi_n(p) = \sum_{i\neq j}\frac{1}{(\lambda_i - \lambda_j)^2} = 2\sum_{1\le i<j\le n}\frac{1}{(\lambda_i - \lambda_j)^2}. \tag{4}$$

Proof. Expand the square in (3):

$$\Phi_n(p) = \sum_{i=1}^{n}\sum_{j\neq i} \frac{1}{(\lambda_i-\lambda_j)^2} + \sum_{i=1}^{n} \sum_{\substack{j\neq i\\ k\neq i,\, k\neq j}} \frac{1}{(\lambda_i-\lambda_j)(\lambda_i-\lambda_k)}.$$

The first term is exactly $\sum_{i\neq j}(\lambda_i-\lambda_j)^{-2}$. For the second term, fix three *distinct* indices i, j, k and consider the cyclic sum

$$\frac{1}{(\lambda_i-\lambda_j)(\lambda_i-\lambda_k)} + \frac{1}{(\lambda_j-\lambda_k)(\lambda_j-\lambda_i)} + \frac{1}{(\lambda_k-\lambda_i)(\lambda_k-\lambda_j)}.$$

With the common denominator $(\lambda_i-\lambda_j)(\lambda_j-\lambda_k)(\lambda_k-\lambda_i)$, the numerator becomes $(\lambda_j-\lambda_k)+(\lambda_k-\lambda_i)+(\lambda_i-\lambda_j)=0$. Thus the triple sum vanishes after grouping terms by unordered triples $\{i,j,k\}$, and (4) follows. ☐

Finite free additive convolution

The *symmetric finite free additive convolution* $p \boxplus_n q$ (Marcus–Spielman–Srivastava [2]) is the monic degree-n polynomial $r(x)$ whose coefficients are given by

$$r(x) = \sum_{k=0}^{n} x^{n-k} c_k, \qquad c_k := \sum_{i+j=k} \frac{(n-i)!\,(n-j)!}{n!\,(n-k)!}\, a_i\, b_j, \tag{5}$$

where $p(x) = \sum_{i=0}^{n} x^{n-i} a_i$ and $q(x) = \sum_{j=0}^{n} x^{n-j} b_j$ (so $a_0 = b_0 = 1$). This is equivalent to the permutation/"matching" model below.

Lemma 4.3 (Permutation model)**.** *Let $\lambda = (\lambda_1,\ldots,\lambda_n)$ and $\mu = (\mu_1,\ldots,\mu_n)$ be the (multi)sets of roots of p and q. Then*

$$(p \boxplus_n q)(x) = \frac{1}{n!}\sum_{\pi\in S_n}\prod_{i=1}^{n}(x-\lambda_i-\mu_{\pi(i)}). \tag{6}$$

Proof. Expanding the product in (6) and averaging over π, the coefficient of x^{n-k} is the expected elementary symmetric sum of order k of the matched sums $\{\lambda_i+\mu_{\pi(i)}\}_{i=1}^{n}$. Choosing i of the λ's and j of the μ's with $i+j=k$, there are $\binom{n}{i}\binom{n-i}{j}$ ways to select the positions and $i!\,j!\,(n-k)!$ permutations compatible with that choice. Dividing by $n!$ gives the weight

$$\frac{\binom{n}{i}\binom{n-i}{j}\, i!\, j!\,(n-k)!}{n!} = \frac{(n-i)!\,(n-j)!}{n!\,(n-k)!},$$

and the corresponding contribution is $a_i b_j$ (since a_i and b_j are the signed elementary symmetric sums). Summing over $i+j=k$ yields exactly (5). ☐

The Stam-type question

Problem (First Proof Q4). Is it true that for all monic real-rooted p, q of degree n,

$$\frac{1}{\Phi_n(p\boxplus_n q)} \ \geq\ \frac{1}{\Phi_n(p)} + \frac{1}{\Phi_n(q)}\ ? \tag{7}$$

We can prove (7) for $n = 2$ (with equality).

The case $n = 2$

Proposition 4.4 (Case $n=2$)**.** *For $n = 2$, (7) holds with equality.*

Proof. Write

$$p(x) = x^2 - a_1x + a_2, \qquad q(x) = x^2 - b_1x + b_2,$$

so the root gaps are $g := \sqrt{a_1^2 - 4a_2}$ and $d := \sqrt{b_1^2 - 4b_2}$. For a quadratic with roots $\alpha > \beta$ one has

$$\Phi_2 = \left(\frac{1}{\alpha-\beta}\right)^2 + \left(\frac{1}{\beta-\alpha}\right)^2 = \frac{2}{(\alpha-\beta)^2}, \qquad \text{so} \qquad \frac{1}{\Phi_2} = \frac{(\alpha-\beta)^2}{2}.$$

Hence $1/\Phi_2(p) = g^2/2$ and $1/\Phi_2(q) = d^2/2$.

Using (5) with $n = 2$, we get

$$(p \boxplus_2 q)(x) = x^2 - (a_1 + b_1)x + \Big(a_2 + b_2 + \tfrac{1}{2}a_1b_1\Big).$$

Its discriminant is

$$\Delta = (a_1+b_1)^2 - 4\Big(a_2 + b_2 + \tfrac{1}{2}a_1b_1\Big) = (a_1^2 - 4a_2) + (b_1^2 - 4b_2) = g^2 + d^2.$$

Thus the root gap of $p \boxplus_2 q$ is $\sqrt{\Delta} = \sqrt{g^2+d^2}$, and therefore

$$\frac{1}{\Phi_2(p \boxplus_2 q)} = \frac{g^2+d^2}{2} = \frac{1}{\Phi_2(p)} + \frac{1}{\Phi_2(q)}.$$

$\square$

Comments and evidence for $n \geq 3$

Remark 4.5 (Degenerate "equality")*.* If $q(x) = (x-t)^n$, then $p \boxplus_n q$ is a translate of p (indeed $p \boxplus_n (x-t)^n = p(x-t)$ from (6)). Here $\Phi_n(q) = +\infty$ and $\Phi_n(p \boxplus_n q) = \Phi_n(p)$, so (7) reduces to the trivial equality $1/\Phi_n(p) = 1/\Phi_n(p) + 0$. More generally, as q approaches a multiple-root polynomial, $\Phi_n(q) \to +\infty$ and the inequality (if true) becomes asymptotically sharp.

Remark 4.6 (Relation to free Stam)*.* If $\mu_\lambda := \frac{1}{n}\sum_{i=1}^n \delta_{\lambda_i}$ is the empirical measure of the roots, then $\Phi_n(p) = n^3\,\Phi^*(\mu_\lambda)$ formally matches Voiculescu's free Fisher information $\Phi^*(\mu) = \int (H\mu)^2\,d\mu$ (principal value Hilbert transform). Finite free convolution of root multisets converges, after scaling, to free additive convolution as $n \to \infty$ ([2] and [18, §4]). In that limit, (7) becomes Voiculescu's free Stam inequality $(\Phi^*(\mu \boxplus \nu))^{-1} \geq (\Phi^*(\mu))^{-1} + (\Phi^*(\nu))^{-1}$ [19]. This provides strong heuristic support for (7), but does not by itself yield a finite-n proof.

Remark 4.7 (Numerical checks)*.* We verified (7) numerically for many random choices of distinct real roots for $n \leq 6$ using the coefficient formula (5) and evaluating Φ_n via Lemma 4.2. The inequality held in all tests and was typically strict for $n \geq 3$ away from degenerate (nearly multiple-root) inputs. These computations are *evidence* only; a general proof (or counterexample) remains open in this write-up.

5 Question 5: An $\mathcal{O}$-slice filtration and geometric fixed point connectivity

Statement

Fix a finite group G. Let $\mathcal{O}$ denote an *incomplete transfer system* associated to an N_∞ operad. Define a slice filtration on the G-equivariant stable category adapted to $\mathcal{O}$ and give a characterization of the $\mathcal{O}$-slice connectivity of a connective G-spectrum in terms of the geometric fixed points.

5.1 Transfer systems and an "$\mathcal{O}$-index"

We isolate the minimal structural features of an incomplete transfer system that we will use.

Definition 5.1 (Transfer system)**.** A *transfer system* $\mathcal{O}$ on G is a relation $\to_{\mathcal{O}}$ on the set of subgroups of G such that:

(i) (*Refines inclusion*) $K \to_{\mathcal{O}} H$ implies $K \leq H$.

(ii) (*Reflexive and transitive*) $H \to_{\mathcal{O}} H$ for all H, and $K \to_{\mathcal{O}} H \to_{\mathcal{O}} L$ implies $K \to_{\mathcal{O}} L$.

(iii) (*Conjugation closed*) $K \to_{\mathcal{O}} H$ implies $gKg^{-1} \to_{\mathcal{O}} gHg^{-1}$ for all $g \in G$.

(iv) (*Restriction closed*) If $K \to_{\mathcal{O}} H$ and $L \leq H$, then $K \cap L \to_{\mathcal{O}} L$.

We call an inclusion $K \leq H$ with $K \to_{\mathcal{O}} H$ an $\mathcal{O}$-*admissible transfer*.

Remark 5.2. For an N_∞ operad $\mathcal{O}$, the associated transfer system records which transitive H-sets H/K are *admissible*; in that dictionary, $K \to_{\mathcal{O}} H$ means H/K is admissible. We only use the axioms in Definition 5.1.

The slice-connectivity bounds are controlled by the largest index of an admissible transfer into a subgroup.

Definition 5.3 ($\mathcal{O}$-index)**.** For a subgroup $H \leq G$ define

$$a_{\mathcal{O}}(H) \;:=\; \max\{[H : K] \;:\; K \to_{\mathcal{O}} H\} \in \{1, 2, \ldots, |H|\}.$$

This is well-defined because $H \to_{\mathcal{O}} H$ always holds, so the set is nonempty.

Remark 5.4 (Extremes). If $\mathcal{O}$ is the *trivial* transfer system (only $H \to H$), then $a_{\mathcal{O}}(H) = 1$ for all H. If $\mathcal{O}$ is the *complete* transfer system (all inclusions are admissible), then $a_{\mathcal{O}}(H) = |H|$ for all H. Thus $a_{\mathcal{O}}$ interpolates between the Postnikov (naïve) and the genuine regular slice scalings.

5.2 $\mathcal{O}$-slice cells and the induced filtration

Let S^V denote the one-point compactification of a real H-representation V. For a finite H-set T, write $\mathbb{R}[T]$ for the real permutation representation on T.

Definition 5.5 ($\mathcal{O}$-slice cells)**.** Let $K \to_{\mathcal{O}} H$ be an admissible transfer. Let

$$\rho_{H/K} := \mathbb{R}[H/K]$$

be the permutation representation of H on the left cosets H/K; it has $\dim(\rho_{H/K}) = [H : K]$ and $\dim(\rho_{H/K}^H) = 1$. For an integer $m \geq 0$, define the associated $\mathcal{O}$-*slice cell*

$$C_{\mathcal{O}}(H, K; m) \;:=\; G/H_+ \;\wedge\; S^{m\rho_{H/K}} \;\in \mathrm{Sp}^G.$$

We define its *$\mathcal{O}$-slice dimension* to be

$$\dim_{\mathcal{O}}(C_{\mathcal{O}}(H,K;m)) \ := \ m\,[H:K].$$

Definition 5.6 ($\mathcal{O}$-slice filtration)**.** For an integer n, let $\tau^{\mathcal{O}}_{\geq n} \subseteq \mathrm{Sp}^G$ be the *localizing* subcategory generated by all $\mathcal{O}$-slice cells of $\mathcal{O}$-slice dimension at least n, i.e. by all $C_{\mathcal{O}}(H,K;m)$ with $K \to_{\mathcal{O}} H$ and $m[H:K] \geq n$. We say that X is *$\mathcal{O}$-slice n-connective* if $X \in \tau^{\mathcal{O}}_{\geq n}$.

Remark 5.7. When $\mathcal{O}$ is complete, taking only the case $K = e$ recovers the *regular slice cells* $G/H_+ \wedge S^{m\rho_H}$ with $\rho_H = \mathbb{R}[H]$ and the usual regular slice filtration. When $\mathcal{O}$ is trivial, the only admissible K is $K = H$, so $C_{\mathcal{O}}(H,H;m) = G/H_+ \wedge S^m$ and the filtration collapses to the ordinary Postnikov connectivity condition simultaneously on all geometric fixed points.

5.3 Geometric fixed points and connectivity

For a subgroup $L \leq G$, write $\Phi^L : \mathrm{Sp}^G \to \mathrm{Sp}$ for the (Lewis–May) geometric fixed point functor. For an ordinary spectrum Y, we say Y is *k-connective* if $\pi_i(Y) = 0$ for all $i < k$.

The key combinatorial input is a lower bound on the number of orbits of subgroup actions on admissible coset sets.

Lemma 5.8 (Orbits are not too large)**.** *Assume $K \to_{\mathcal{O}} H$ and let $L \leq H$. For any coset $hK \in H/K$, the stabilizer of hK in L is $L \cap hKh^{-1}$, and it satisfies*

$$L \cap hKh^{-1} \ \to_{\mathcal{O}} \ L.$$

Consequently every L-orbit in the L-set H/K has cardinality at most $a_{\mathcal{O}}(L)$. In particular,

$$|L\backslash H/K| \ \geq \ \frac{|H/K|}{a_{\mathcal{O}}(L)} \ = \ \frac{[H:K]}{a_{\mathcal{O}}(L)}.$$

Proof. The stabilizer formula is standard: $l \cdot (hK) = hK$ iff $h^{-1}lh \in K$, i.e. $l \in hKh^{-1}$; intersecting with L gives $\mathrm{Stab}_L(hK) = L \cap hKh^{-1}$.

Since $K \to_{\mathcal{O}} H$, conjugation closure gives $hKh^{-1} \to_{\mathcal{O}} H$, and then restriction closure (with $L \leq H$) gives $L \cap hKh^{-1} \to_{\mathcal{O}} L$. By definition of $a_{\mathcal{O}}(L)$, every admissible subgroup $J \to_{\mathcal{O}} L$ has index $[L:J] \leq a_{\mathcal{O}}(L)$. Thus each orbit $L/(L \cap hKh^{-1})$ has size $\leq a_{\mathcal{O}}(L)$.

Finally, the number of orbits is at least the total size divided by the maximal orbit size, giving the inequality. $\square$

Lemma 5.9 (Fixed points of permutation representations)**.** *Let T be a finite L-set. Then*

$$\dim(\mathbb{R}[T]^L) = |L\backslash T|.$$

In particular, for $T = H/K$ viewed as an L-set by restriction along $L \leq H$,

$$\dim(\rho^L_{H/K}) = |L\backslash H/K|.$$

Proof. An element of $\mathbb{R}[T]$ is a function $f : T \to \mathbb{R}$. It is L-invariant iff f is constant on L-orbits. Thus $\mathbb{R}[T]^L$ is naturally identified with functions on the orbit set $L\backslash T$, which has dimension $|L\backslash T|$. $\square$

We can now compute the connectivity of geometric fixed points of the $\mathcal{O}$-slice cells.

Proposition 5.10 (Connectivity of Φ^L on $\mathcal{O}$-slice cells)**.** *Let $C_{\mathcal{O}}(H,K;m)$ be an $\mathcal{O}$-slice cell and let $L \leq G$. Then $\Phi^L(C_{\mathcal{O}}(H,K;m))$ is either contractible or a wedge of spheres of dimension*

$$m \cdot |L\backslash H/K|.$$

In particular, if $m[H:K] \geq n$ then

$$\Phi^L(C_{\mathcal{O}}(H,K;m)) \;\in\; \mathrm{Sp}_{\geq \lceil n/a_{\mathcal{O}}(L)\rceil}.$$

Proof. If L is not subconjugate to H, then $(G/H)^L = \varnothing$ and $\Phi^L(G/H_+) \simeq *$, hence $\Phi^L(C_{\mathcal{O}}(H,K;m)) \simeq *$.

If L is subconjugate to H, then $\Phi^L(G/H_+) \simeq (G/H)^L_+$ is a finite wedge of copies of S^0 indexed by $(G/H)^L$. Since Φ^L is symmetric monoidal and exact, we get

$$\Phi^L(C_{\mathcal{O}}(H,K;m)) \simeq (G/H)^L_+ \;\wedge\; \Phi^L(S^{m\rho_{H/K}}) \simeq \bigvee_{(G/H)^L} S^{m\dim(\rho^L_{H/K})}.$$

By Lemma 5.9 we have $\dim(\rho^L_{H/K}) = |L\backslash H/K|$, proving the wedge description.

For the connectivity bound, Lemma 5.8 gives $|L\backslash H/K| \geq [H:K]/a_{\mathcal{O}}(L)$, hence each wedge summand has dimension

$$m \cdot |L\backslash H/K| \;\geq\; m\,\frac{[H:K]}{a_{\mathcal{O}}(L)} \;\geq\; \frac{n}{a_{\mathcal{O}}(L)}.$$

Therefore $\Phi^L(C_{\mathcal{O}}(H,K;m))$ is $\lceil n/a_{\mathcal{O}}(L)\rceil$-connective. □

5.4 Main theorem: characterization of $\mathcal{O}$-slice connectivity

Theorem 5.11 (Geometric fixed point criterion)**.** *Let X be a* connective *G-spectrum (i.e. $\pi_i(\Phi^e X) = 0$ for $i < 0$). Fix an integer n. Then the following are equivalent:*

(a) X is $\mathcal{O}$-slice n-connective, i.e. $X \in \tau^{\mathcal{O}}_{\geq n}$.

(b) For every subgroup $H \leq G$, the geometric fixed point spectrum $\Phi^H X$ is $\lceil n/a_{\mathcal{O}}(H)\rceil$-connective.

Proof. We prove (a)⇒(b) and then (b)⇒(a).

(a)⇒(b). The subcategory $\tau^{\mathcal{O}}_{\geq n}$ is localizing by definition. The functor Φ^H preserves colimits and cofiber sequences, hence it suffices to check the claim on the generators of $\tau^{\mathcal{O}}_{\geq n}$, namely the $\mathcal{O}$-slice cells with $m[H':K'] \geq n$. This is exactly Proposition 5.10.

(b)⇒(a): induction on $|G|$. When G is trivial, $\mathrm{Sp}^G \simeq \mathrm{Sp}$, $a_{\mathcal{O}}(e) = 1$, and $\tau^{\mathcal{O}}_{\geq n}$ is generated by S^m with $m \geq n$, so (b) is precisely ordinary n-connectivity. The claim is clear in this base case.

Assume now $|G| > 1$ and that the theorem holds for all proper subgroups of G (with the restricted transfer system). Let $\mathcal{P}$ be the family of proper subgroups of G. There is the standard isotropy separation cofiber sequence

$$E\mathcal{P}_+ \wedge X \;\longrightarrow\; X \;\longrightarrow\; \widetilde{E}\mathcal{P} \wedge X. \tag{8}$$

Write $Y := E\mathcal{P}_+ \wedge X$ and $Z := \widetilde{E}\mathcal{P} \wedge X$.

Step 1: $Y \in \tau^{\mathcal{O}}_{\geq n}$. The spectrum $E\mathcal{P}_+$ is built (as a G-CW spectrum) from cells G/H_+ with $H < G$. Smashing with X and taking filtered colimits shows Y lies in the localizing subcategory generated by the induced spectra $G/H_+ \wedge X \simeq G_+ \wedge_H \mathrm{Res}^G_H X$ for $H < G$.

For each proper $H < G$, the hypothesis (b) for X implies that for every $K \le H$,

$$\Phi^K(\mathrm{Res}^G_H X) \;\simeq\; \Phi^K(X) \;\in\; \mathrm{Sp}_{\ge \lceil n/a_{\mathcal{O}}(K)\rceil}.$$

By the inductive hypothesis applied to the subgroup H and the restricted transfer system $\mathcal{O}|_H$, we conclude $\mathrm{Res}^G_H X \in \tau^{\mathcal{O}|_H}_{\ge n}$ inside Sp^H.

Finally, induction preserves $\mathcal{O}$-slice connectivity: if $W \in \tau^{\mathcal{O}|_H}_{\ge n}$ then $G_+ \wedge_H W \in \tau^{\mathcal{O}}_{\ge n}$, because induction sends the generating cell $H/L_+ \wedge S^{m\rho_{L/M}}$ (with $M \to_{\mathcal{O}} L$ inside H and $m[L : M] \ge n$) to the G-cell $G/L_+ \wedge S^{m\rho_{L/M}}$ of the same $\mathcal{O}$-slice dimension. Therefore each $G/H_+ \wedge X$ lies in $\tau^{\mathcal{O}}_{\ge n}$, hence so does their localizing closure, namely Y.

Step 2: $Z \in \tau^{\mathcal{O}}_{\ge n}$. By construction, Z is *geometric* in the sense that $\Phi^H(Z) \simeq *$ for every proper subgroup $H < G$, and $\Phi^G(Z) \simeq \Phi^G(X)$. Moreover $\Phi^G(X)$ is $\lceil n/a_{\mathcal{O}}(G)\rceil$-connective by assumption.

Let $\mathrm{Sp}^G_{\mathrm{geom}} \subseteq \mathrm{Sp}^G$ be the full subcategory of geometric G-spectra. The functor $\Phi^G : \mathrm{Sp}^G_{\mathrm{geom}} \to \mathrm{Sp}$ is an equivalence with inverse

$$\Psi(Y) := \widetilde{E}\mathcal{P} \wedge \iota(Y),$$

where $\iota : \mathrm{Sp} \to \mathrm{Sp}^G$ is inflation (trivial G-action). Under this equivalence, the $\mathcal{O}$-slice cell $C_{\mathcal{O}}(G, K; m) = G/G_+ \wedge S^{m\rho_{G/K}}$ corresponds to the ordinary sphere S^m, and its $\mathcal{O}$-slice dimension is $m[G : K]$. Thus, inside geometric spectra, the localizing subcategory $\tau^{\mathcal{O}}_{\ge n}$ corresponds exactly to the ordinary connective subcategory $\mathrm{Sp}_{\ge\lceil n/a_{\mathcal{O}}(G)\rceil}$. Since $\Phi^G(Z) \simeq \Phi^G(X) \in \mathrm{Sp}_{\ge\lceil n/a_{\mathcal{O}}(G)\rceil}$, we get $Z \in \tau^{\mathcal{O}}_{\ge n}$.

Step 3: conclude for X. Because $\tau^{\mathcal{O}}_{\ge n}$ is localizing, it is closed under cofibers. From (8) and Steps 1–2 we conclude $X \in \tau^{\mathcal{O}}_{\ge n}$. □

Remark 5.12 (Comparison with the classical criterion)*.* If $\mathcal{O}$ is complete, then $a_{\mathcal{O}}(H) = |H|$ and Theorem 5.11 specializes to the usual criterion for the regular slice filtration: $X \in \tau_{\ge n}$ iff $\Phi^H(X)$ is $\lceil n/|H|\rceil$-connective for all H. At the other extreme, if $\mathcal{O}$ is trivial then $a_{\mathcal{O}}(H) = 1$ and the condition becomes that *all* geometric fixed points are n-connective, i.e. a uniform Postnikov condition.

Remark 5.13 (Relation to the slice literature and normalization)*.* A transfer system $\mathcal{O}$ coming from an N_∞ operad (equivalently, from an indexing system in the sense of Blumberg–Hill [12]) specifies which orbits H/K are *admissible* for norm/transfer operations. The benchmark asks for an $\mathcal{O}$-adapted slice filtration. There are several closely related slice conventions in the literature (the original slice filtration of Hill–Hopkins–Ravenel [13], the regular variant analyzed by Ullman [11], and expository accounts such as Hill's primer [10]). In this note we use a concrete generating set built from permutation representation spheres $\rho_{H/K} = \mathbb{R}[H/K]$ for $\mathcal{O}$-admissible orbits and we declare the $\mathcal{O}$-slice dimension of $G/H_+ \wedge S^{m\rho_{H/K}}$ to be $m[H : K]$. This choice has the correct complete/trivial specializations and is natural from the transfer-system point of view, but other normalizations (e.g. including -1 shifts or using different generating representation spheres) are possible. Theorem 5.11 should therefore be read as a complete geometric-fixed-point connectivity criterion *for this specific permutation-cell model* of an $\mathcal{O}$-slice filtration; translating to another convention amounts to reindexing by fixed-point dimensions.

6 Question 6

6.1 Statement

Let $G = (V, E)$ be a (simple, undirected) graph on $n \coloneqq |V|$ vertices with Laplacian matrix

$$L = \sum_{\{u,v\}\in E} L_{uv}, \qquad L_{uv} \coloneqq (e_u - e_v)(e_u - e_v)^\mathsf{T}.$$

For $S \subseteq V$ let L_S denote the Laplacian of the induced subgraph $G[S]$, viewed as an $n \times n$ matrix by padding zeros on $V \setminus S$.

A set S is called *ε-light* if

$$L_S \preceq \varepsilon L.$$

Question 6 asks whether there exists a universal constant $c > 0$ such that for every graph G and every $\varepsilon \in (0, 1)$ there is an ε-light set S with

$$|S| \geq c\,\varepsilon\, n.$$

6.2 Two simple necessary conditions

The semidefinite constraint implies a few easy-to-check necessary conditions. These do not solve the problem by themselves, but they help clarify what an ε-light set must "look like."

Lemma 6.1 (Vertexwise degree constraint)**.** *If S is ε-light, then for every $v \in V$,*

$$\deg_{G[S]}(v) \leq \varepsilon \deg_G(v),$$

where $\deg_{G[S]}(v)$ is the degree of v in the induced subgraph $G[S]$ (and is 0 when $v \notin S$).

Proof. Fix $v \in V$ and test the semidefinite inequality $L_S \preceq \varepsilon L$ on the standard basis vector e_v:

$$0 \leq e_v^\mathsf{T}(\varepsilon L - L_S)e_v = \varepsilon\, e_v^\mathsf{T} L e_v - e_v^\mathsf{T} L_S e_v.$$

For any Laplacian, $e_v^\mathsf{T} L e_v = \deg_G(v)$ and $e_v^\mathsf{T} L_S e_v = \deg_{G[S]}(v)$, giving the claim. □

The next necessary condition uses effective resistance. Recall that for a connected graph with Laplacian L, the effective resistance of an edge $e = \{u, v\} \in E$ can be written as

$$R^G_{\text{eff}}(u, v) = (e_u - e_v)^\mathsf{T} L^+ (e_u - e_v),$$

where L^+ is the Moore–Penrose pseudoinverse.

Lemma 6.2 (Internal edges must have small effective resistance)**.** *Assume G is connected. If S is ε-light and $e = \{u, v\} \in E(G[S])$ is an edge of the induced subgraph, then*

$$R^G_{\text{eff}}(u, v) \leq \varepsilon.$$

Proof. Since e is an internal edge, its edge-Laplacian L_{uv} is a summand of L_S, hence

$$L_{uv} \preceq L_S \preceq \varepsilon L.$$

Conjugating by $L^{+1/2}$ on the image of L (equivalently, applying the inequality in the L-energy inner product) gives

$$L^{+1/2} L_{uv} L^{+1/2} \preceq \varepsilon\, L^{+1/2} L L^{+1/2} = \varepsilon\, \Pi,$$

where Π is the orthogonal projection onto $\mathrm{im}(L)$. But $L^{+1/2} L_{uv} L^{+1/2}$ is rank-one and equals

$$L^{+1/2}(e_u - e_v)(e_u - e_v)^{\mathsf{T}} L^{+1/2} \;=\; ww^{\mathsf{T}}, \qquad w := L^{+1/2}(e_u - e_v).$$

The largest eigenvalue of ww^{T} is $\|w\|_2^2 = (e_u - e_v)^{\mathsf{T}} L^+ (e_u - e_v) = R_{\mathrm{eff}}^G(u,v)$. The inequality $ww^{\mathsf{T}} \preceq \varepsilon\Pi$ forces $\|w\|_2^2 \le \varepsilon$, yielding the claim. □

Lemma 6.2 shows that an ε-light set can never contain *both* endpoints of an edge whose effective resistance exceeds ε. This recovers, for instance, the fact that when G is a matching (each edge has effective resistance 1), every ε-light set for $\varepsilon < 1$ must be an independent set.

6.3 A universal upper bound on the best constant: $c \le \frac{1}{2}$

Any putative universal constant c cannot exceed 1/2. The obstruction is simply rounding inside small cliques.

Proposition 6.3 (Clique-union obstruction)**.** *For every $\varepsilon \in (0,1)$ and every integer $m \ge 3$, let G be the disjoint union of t copies of the clique K_m (so $n = tm$). Then any ε-light set S satisfies*

$$|S| \;\le\; t\,\lfloor \varepsilon m \rfloor.$$

In particular, taking m so that $1 < \varepsilon m < 2$ forces $|S| \le t = n/m < (\varepsilon/1)n$ and as $\varepsilon m \uparrow 2$ the ratio $|S|/(\varepsilon n)$ approaches 1/2.

Proof. Because G is a disjoint union of components, both L and L_S are block diagonal with one block per clique. The matrix inequality $L_S \preceq \varepsilon L$ holds if and only if it holds on each component.

Fix one clique component $C \cong K_m$ and write $S_C := S \cap C$ with $|S_C| = s$. On this component L is the Laplacian L_{K_m} and L_S is the Laplacian L_{K_s} padded with zeros on $C \setminus S_C$.

For the complete graph one checks directly that

$$L_{K_s} \;\preceq\; \frac{s}{m}\, L_{K_m} \qquad (0 \le s \le m),$$

and that this is tight on the $(s-1)$-dimensional subspace of vectors supported on S_C with sum zero. (Equivalently: L_{K_s} has nonzero eigenvalue s and L_{K_m} has nonzero eigenvalue m.) Thus $L_{K_s} \preceq \varepsilon L_{K_m}$ holds if and only if $s \le \varepsilon m$. Since s is an integer, $s \le \lfloor \varepsilon m \rfloor$. Summing s over the t components yields $|S| \le t\lfloor \varepsilon m \rfloor$.

Finally, if $1 < \varepsilon m < 2$ then $\lfloor \varepsilon m \rfloor = 1$, so $|S| \le t = n/m$. Taking m with εm arbitrarily close to 2 from below makes n/m arbitrarily close to $(\varepsilon/2)n$, proving that no universal constant can exceed 1/2. □

6.4 A paving lemma and completion of the argument

The necessary conditions above (especially Lemma 6.2) explain why one cannot hope to take c arbitrarily large. What remains is a lower bound: showing that *some* universal $c > 0$ always works.

The cleanest way to finish the problem is via the following vertex-partitioning statement.

Lemma 6.4 (Vertex paving for graph Laplacians)**.** *For every graph $G = (V, E)$ with Laplacian L and every integer $r \ge 2$, there exists a partition*

$$V = V_1 \sqcup \cdots \sqcup V_r$$

such that for each $i = 1, \ldots, r$,

$$L_{V_i} \preceq \frac{2}{r}\, L. \tag{9}$$

Remark 6.5. Lemma 6.4 is a graph-valued instance of the "paving" phenomenon that appears in the Kadison–Singer problem. One can deduce statements of this type from the Marcus–Spielman–Srivastava theorem on partitions of sums of rank-one positive semidefinite matrices (often stated as Weaver's KS_r conjecture). Since that theory is well beyond the scope of these notes, we record Lemma 6.4 as an external input.

Theorem 6.6. *There exists a universal constant $c > 0$ such that for every graph $G = (V, E)$ on n vertices and every $0 < \varepsilon < 1$, there exists an ε-light set $S \subseteq V$ with*

$$|S| \geq c\,\varepsilon n.$$

In particular, one may take $c = 1/3$.

Proof. Fix $0 < \varepsilon < 1$ and set $r := \lceil 2/\varepsilon \rceil$. Apply Lemma 6.4 to obtain a partition $V = V_1 \sqcup \cdots \sqcup V_r$ satisfying (9). Let S be a largest part, so $|S| \geq n/r$. Since $r \leq 2/\varepsilon + 1 = (2 + \varepsilon)/\varepsilon$, we have

$$|S| \geq \frac{n}{2/\varepsilon + 1} = \frac{\varepsilon n}{2 + \varepsilon} \geq \frac{\varepsilon n}{3}.$$

Moreover, (9) gives

$$L_S \preceq \frac{2}{r} L \preceq \varepsilon L,$$

so S is ε-light. $\square$

Remark 6.7. Proposition 6.3 shows that no universal constant can exceed $1/2$. The argument above gives an explicit positive constant ($c = 1/3$) once one accepts Lemma 6.4.

7 Question 7: Uniform lattices with 2-torsion and $\mathbb{Q}$-acyclic universal covers

Statement

Let Γ be a *uniform lattice* in a real semisimple Lie group, and assume that Γ contains some 2-torsion. Is it possible for Γ to be the fundamental group of a compact manifold M without boundary such that the universal cover $\widetilde{M}$ is $\mathbb{Q}$-*acyclic*, i.e.

$$\widetilde{H}_k(\widetilde{M}; \mathbb{Q}) = 0 \qquad \text{for all } k \geq 0\ ?$$

7.1 Two finiteness properties

It is convenient to isolate two standard group-theoretic properties.

Definition 7.1 (FH($\mathbb{Q}$) and Mfld($\mathbb{Q}$))**.** A group Γ is said to be **FH($\mathbb{Q}$)** if it admits a free, properly discontinuous, cocompact action on a finite-dimensional CW complex X such that $\widetilde{H}_*(X; \mathbb{Q}) = 0$.

We say $\Gamma \in$ **Mfld($\mathbb{Q}$)** if there exists a closed (topological) manifold M with $\pi_1(M) \cong \Gamma$ and $\widetilde{H}_*(\widetilde{M}; \mathbb{Q}) = 0$.

Evidently $\Gamma \in \mathrm{Mfld}(\mathbb{Q})$ implies $\Gamma \in \mathrm{FH}(\mathbb{Q})$ by taking $X = \widetilde{M}$ with its lifted CW structure.

7.2 Chain-level consequences of $\mathbb{Q}$-acyclicity

Lemma 7.2 (Chain-level description)**.** *Let M be a finite CW complex with fundamental group Γ and universal cover $\widetilde{M}$. Then the cellular chain complex $C_*(\widetilde{M};\mathbb{Q})$ is a bounded complex of finitely generated free $\mathbb{Q}\Gamma$-modules. Moreover,*

$$\widetilde{H}_*(\widetilde{M};\mathbb{Q}) = 0 \text{ in degrees } > 0 \quad\Longleftrightarrow\quad C_*(\widetilde{M};\mathbb{Q}) \to \mathbb{Q} \to 0 \text{ is a free } \mathbb{Q}\Gamma\text{-resolution of } \mathbb{Q}.$$

Proof. The CW structure on M lifts to a Γ-equivariant CW structure on $\widetilde{M}$ with finitely many cell orbits in each degree. Hence each chain group is a finite rank free module $C_k(\widetilde{M};\mathbb{Q}) \cong (\mathbb{Q}\Gamma)^{c_k}$, where c_k is the number of k-cells of M. The augmentation map $C_0(\widetilde{M};\mathbb{Q}) \to \mathbb{Q}$ exhibits $\mathbb{Q}$ as the 0th homology of $C_*(\widetilde{M};\mathbb{Q})$, and vanishing of homology in positive degrees is exactness of the augmented complex. □

Proposition 7.3 (Rational Poincaré duality for Γ)**.** *If M is a closed n-manifold with $\pi_1(M) \cong \Gamma$ and $\widetilde{H}_*(\widetilde{M};\mathbb{Q}) = 0$, then Γ is a $\mathbb{Q}$-Poincaré duality group of dimension n: for every $\mathbb{Q}\Gamma$-module V there are natural isomorphisms*

$$H^k(\Gamma;V) \;\cong\; H_{n-k}(\Gamma;V \otimes O),$$

where O is the $\mathbb{Q}\Gamma$-module given by the orientation character of M. Equivalently, the classifying map $M \to B\Gamma$ is a $\mathbb{Q}$-homology equivalence.

Proof. By Lemma 7.2, $C_*(\widetilde{M};\mathbb{Q})$ is a finite free resolution of $\mathbb{Q}$. Poincaré duality for the manifold yields a chain homotopy equivalence

$$C_*(\widetilde{M};\mathbb{Q}) \;\simeq\; \mathrm{Hom}_{\mathbb{Q}\Gamma}(C_{n-*}(\widetilde{M};\mathbb{Q}), \mathbb{Q}\Gamma) \otimes O,$$

the standard algebraic formulation of $\mathbb{Q}$-Poincaré duality for Γ. Applying $\mathrm{Hom}_{\mathbb{Q}\Gamma}(-,V)$ and taking homology gives the stated duality. The final claim follows because group (co)homology with $\mathbb{Q}$-coefficients is computed from the resolution $C_*(\widetilde{M};\mathbb{Q})$, and coinvariants identify $C_*(\widetilde{M};\mathbb{Q}) \otimes_{\mathbb{Q}\Gamma} \mathbb{Q} \cong C_*(M;\mathbb{Q})$. □

Remark 7.4 (Euler characteristics). If $\Gamma \in \mathrm{FH}(\mathbb{Q})$ one can define a rational Euler characteristic $\chi(\Gamma)$ using any finite free $\mathbb{Q}\Gamma$-resolution of $\mathbb{Q}$. When $\Gamma = \pi_1(M)$ for a finite CW complex with $\mathbb{Q}$-acyclic universal cover, one gets $\chi(\Gamma) = \chi(M) \in \mathbb{Z}$.

7.3 Known obstruction: odd torsion

The following theorem is due to Fowler [8, Thm. 1.3.2 and Thm. 5.1.1] (see also [9]).

Theorem 7.5 (Fowler: odd torsion obstruction)**.** *Let Γ be a uniform lattice in a semisimple Lie group. If Γ contains an element of odd prime order (equivalently, p-torsion for some $p \neq 2$), then there does* not *exist an ANR $\mathbb{Q}$-homology manifold X with $\pi_1(X) \cong \Gamma$ and $\widetilde{H}_*(X;\mathbb{Q}) = 0$. In particular, such a Γ cannot lie in* $\mathrm{Mfld}(\mathbb{Q})$*.*

Remark 7.6 (Heuristic of the obstruction). Very roughly, the proof compares the symmetric signature class of a hypothetical $\mathbb{Q}$-homology manifold model for $B\Gamma$ to an equivariant signature class of the locally symmetric orbifold $\Gamma\backslash G/K$. The mismatch is detected by a ρ-invariant associated to lens spaces normal to generic points of orbifold singular strata; these ρ-invariants are nonzero for odd primes.

7.4 The pure 2-torsion case (status)

If Γ has any odd-order torsion, Theorem 7.5 gives a complete negative answer. The remaining case is when all torsion in Γ is 2-primary.

Theorem 7.7 (Status for pure 2-torsion)**.** *Let Γ be a uniform lattice in a real semisimple Lie group and assume that every torsion element of Γ has order a power of 2. It is currently an* open problem *whether such a Γ can lie in* $\mathrm{Mfld}(\mathbb{Q})$ *(equivalently, be realized as π_1 of a closed manifold with $\mathbb{Q}$-acyclic universal cover); see, for example, [20].*

7.4.1 A necessary integrality constraint

Even without resolving the 2-primary case, there are simple constraints any positive example must satisfy.

Proposition 7.8 (Integrality of the $\mathbb{Q}$-Euler characteristic)**.** *Assume $\Gamma \in \mathrm{Mfld}(\mathbb{Q})$. Then the rational Euler characteristic $\chi(\Gamma)$ (defined whenever Γ is of type $\mathrm{FP}(\mathbb{Q})$) is an integer. In particular, if Γ is a uniform lattice whose rational Euler characteristic is not an integer, then $\Gamma \notin \mathrm{Mfld}(\mathbb{Q})$.*

Proof. Choose a closed manifold M with $\pi_1(M) = \Gamma$ and $\widetilde{M}$ $\mathbb{Q}$-acyclic. By Lemma 7.2, $C_*(\widetilde{M};\mathbb{Q})$ is a finite free $\mathbb{Q}\Gamma$-resolution of $\mathbb{Q}$. The Euler characteristic $\chi(\Gamma)$ is the alternating sum of the free ranks in this resolution, which are exactly the numbers of cells of M. Hence $\chi(\Gamma) = \chi(M) \in \mathbb{Z}$. □

Remark 7.9. For uniform lattices, $\chi(\Gamma)$ is computable in principle via classical formulas (Harder, Gauss–Bonnet, etc.) and coincides with the orbifold Euler characteristic of $\Gamma\backslash G/K$. This can sometimes rule out $\mathrm{Mfld}(\mathbb{Q})$ if denominators occur.

Conclusion (at the level of this note)

- If a uniform lattice Γ contains odd-order torsion, then the answer to Question 7 is *no* by Theorem 7.5.
- The remaining "pure 2-torsion" case is recorded (in the attached note) as open.
- Any positive example must satisfy integrality constraints such as Proposition 7.8.

Remark 7.10 (Why the pure 2-torsion case is different)*.* Fowler's obstruction for odd primes is detected by ρ–invariants coming from lens spaces normal to singular strata in the locally symmetric orbifold $\Gamma\backslash G/K$; these invariants are nontrivial for odd prime torsion and lead to Theorem 7.5. For groups whose torsion is purely 2–primary, the same strategy runs into genuinely 2–local phenomena in surgery theory (e.g. 2–primary L–groups and related *UNil* terms), and a comparable general obstruction is not currently available. In particular, it remains open whether a uniform lattice with only 2–torsion can lie in $\mathrm{Mfld}(\mathbb{Q})$; see [20].

8 Question 8: Smoothing quadrivalent polyhedral Lagrangian surfaces

8.1 Problem statement

A *polyhedral Lagrangian surface* $K \subset \mathbb{R}^4$ is a finite polyhedral complex all of whose faces are Lagrangians, and which is a topological submanifold of $\mathbb{R}^4$. A *Lagrangian smoothing* of K is a

Hamiltonian isotopy K_t of smooth Lagrangian submanifolds, parameterised by $(0, 1]$, extending to a topological isotopy, parametrised by $[0, 1]$, with endpoint $K_0 = K$.

Let K be a polyhedral Lagrangian surface with the property that exactly 4 faces meet at every vertex. Does K necessarily have a Lagrangian smoothing?

8.2 Answer

Yes. One can smooth K by a compactly supported Hamiltonian isotopy, by smoothing each edge and vertex in Darboux charts where the polyhedral Lagrangian is the graph of the differential of a piecewise-quadratic function.

8.2.1 Symplectic preliminaries

We fix the standard identification $\mathbb{R}^4 \cong T^*\mathbb{R}^2$ with coordinates (q_1, q_2, p_1, p_2) and symplectic form

$$\omega_0 = dq_1 \wedge dp_1 + dq_2 \wedge dp_2,$$

and Liouville one–form $\lambda_0 = p_1 dq_1 + p_2 dq_2$.

The following fact is standard: graphs of exact one–forms move by Hamiltonian isotopy.

Lemma 8.1 (Hamiltonian isotopy of exact graphs)**.** *Let $U \subset \mathbb{R}^2$ be open and $f_t : U \to \mathbb{R}$ a smooth one–parameter family, $t \in [0, 1]$, with $\partial_t f_t$ compactly supported in U. Set $L_t := \{(q, df_t(q)) \mid q \in U\} \subset T^*U$. Then (L_t) is the image of L_0 under a compactly supported Hamiltonian isotopy.*

Proof. Define a time–dependent Hamiltonian on T^*U by

$$H_t(q, p) := -\partial_t f_t(q).$$

Its Hamiltonian vector field is $X_{H_t} = \sum_i (\partial_{q_i} H_t)\, \partial_{p_i}$ (purely vertical), hence its flow exists for all time and is compactly supported. A direct computation shows that the flow sends $\mathrm{graph}(df_0)$ to $\mathrm{graph}(df_t)$. ☐

Remark 8.2. In practice we will work in Darboux charts on small balls in $\mathbb{R}^4$; since these balls are contractible, any symplectomorphism between them can be assumed exact (after adjusting by a constant), and Lemma 8.1 applies to produce an *ambient* Hamiltonian isotopy supported in the chart.

8.2.2 Local structure of a polyhedral Lagrangian edge

Let e be an open edge of K (so e contains no vertices). Exactly two faces meet along e. Near a point of e, K is the union of two Lagrangian half–planes glued along the common line ℓ tangent to e.

Lemma 8.3 (Edge chart as a kinked exact graph)**.** *Let $x \in e$ be a point on an edge away from vertices. There exists a small ball $B_x \subset \mathbb{R}^4$ centered at x and a symplectic affine chart $\Phi_x : B_x \to T^*U$ (for $U \subset \mathbb{R}^2$ open) such that $\Phi_x(K \cap B_x)$ is the graph of df for a continuous function $f : U \to \mathbb{R}$ which is smooth on $U \setminus \{q_2 = 0\}$ and is affine–quadratic on each half–space $\{q_2 > 0\}$ and $\{q_2 < 0\}$.*

Proof. Translate so $x = 0$. Let $P_\pm$ be the two Lagrangian planes containing the two faces meeting along e. They intersect along the line ℓ spanned by the edge direction. Choose a Lagrangian plane Q transverse to both $P_\pm$ (this is possible because the set of Lagrangian planes non–transverse to a fixed Lagrangian plane has codimension one in the Lagrangian Grassmannian, so a generic Q avoids both). Choose the complementary Lagrangian plane Q^* so $\mathbb{R}^4 \cong Q \oplus Q^*$.

With respect to this splitting, each $P_\pm$ is the graph of a symmetric linear map $A_\pm : Q \to Q^*$. Moreover, because the two faces coincide along the edge line, the resulting "slope" data agree along the projected edge, so the union of the two half–planes is the graph of a continuous piecewise–affine map $p(q)$ over a neighborhood $U \subset Q$ with a crease along a line. After a linear change of basis in Q we may arrange that this line is $\{q_2 = 0\}$. Since each $A_\pm$ is symmetric, $p(q)$ is the gradient of a (piecewise–quadratic) potential on each side; continuity of p across $\{q_2 = 0\}$ implies these primitives can be chosen to glue to a continuous function f with $df = p$. Setting Φ_x to be the affine symplectic identification $Q \oplus Q^* \cong T^*Q$ completes the claim. □

8.2.3 Local structure at a quadrivalent vertex

Now let v be a vertex of K. By assumption exactly four faces meet at v, hence exactly four edges emanate from v in cyclic order.

Lemma 8.4 (Vertex chart as a piecewise–quadratic exact graph)**.** *Let v be a vertex where exactly four faces meet. There exists a small ball $B_v \subset \mathbb{R}^4$ about v and a symplectic affine chart $\Phi_v : B_v \to T^*U$ such that $\Phi_v(K \cap B_v)$ is the graph of df for a continuous function $f : U \to \mathbb{R}$ with the following properties:*

(i) *$U \subset \mathbb{R}^2$ is a neighborhood of the origin and is subdivided by four rays from 0 into four closed sectors $S_1, \dots, S_4$ meeting along those rays.*

(ii) *f is smooth on the interior of each sector and (on each S_i) is given by a quadratic polynomial.*

(iii) *df is continuous on U (equivalently, the piecewise–affine map $q \mapsto df(q)$ extends continuously across the rays).*

Proof. Translate so $v = 0$. Let $P_1, \dots, P_4$ be the four Lagrangian planes containing the faces meeting at v. Choose a Lagrangian plane Q transverse to all four P_i. This is possible because $\mathrm{Lag}(2) \cong U(2)/O(2)$ is a smooth real 3–manifold, while for a fixed Lagrangian plane P the non-transversality locus $\{Q : Q \cap P \neq 0\}$ is a real codimension–one subset; avoiding finitely many such "walls" leaves a dense open set. Let Q^* be the complementary Lagrangian plane, so $\mathbb{R}^4 \cong Q \oplus Q^*$.

For sufficiently small B_v, the intersection $K \cap B_v$ is a cone on a PL circle (the *link*)

$$L := K \cap \partial B_v,$$

subdivided into four linear arcs, one per face. Because each P_i is transverse to Q, the projection $\pi_Q : Q \oplus Q^* \to Q$ restricts to a homeomorphism from each face sector to its image in Q.

It remains to ensure that these four projected sectors do not overlap. Since L has only four edges, the only possible failure of injectivity of $\pi_Q|_L$ is an intersection between the images of two *nonadjacent* edges. For any fixed pair of nonadjacent edges $E, E' \subset L$, the condition that $\pi_Q(E)$ meets $\pi_Q(E')$ in their interiors is a closed semialgebraic condition on Q of positive codimension among Lagrangian planes transverse to all P_i (it amounts to the existence of nonzero vectors $u \in P_i$ and $u' \in P_j$ from the corresponding face directions with $\pi_Q(u) = \pi_Q(u')$). Since there are only finitely many such pairs, we may choose Q so that π_Q is injective on L, hence (after shrinking B_v if needed) on the whole cone $K \cap B_v$.

Thus π_Q identifies $K \cap B_v$ with the graph of a continuous map $p : U := \pi_Q(K \cap B_v) \to Q^*$ which is affine on each sector.

Because each face lies in a Lagrangian plane, the affine part of p on each sector has symmetric linear part. Equivalently, the one–form $p \cdot dq$ is closed on the interior of each sector and has compatible traces across the boundary rays (because p is continuous there). Since U is contractible, we can define a primitive

$$f(q) := \int_{\gamma_q} p \cdot dq,$$

where γ_q is any piecewise smooth path in U from 0 to q. The closedness implies this integral is path independent, hence $df = p$. On each sector p is affine, so f is quadratic on that sector. Finally, the continuity of p implies df extends continuously across the rays. $\square$

8.2.4 Smoothing a piecewise–quadratic graph

We next show that a graph as in Lemma 8.3 or Lemma 8.4 can be smoothed through Hamiltonian isotopies, with support in an arbitrarily small neighborhood of the singular locus.

Lemma 8.5 (Local smoothing in a cotangent chart)**.** *Let $f : U \to \mathbb{R}$ be continuous and piecewise C^∞ on a finite sector decomposition, with df continuous on U and such that f agrees with a smooth function near ∂U. Then for any sufficiently small neighborhood W of the sector boundaries (the rays/lines where $D^2 f$ jumps) there exists $\widetilde{f} : U \to \mathbb{R}$ with:*

- *(a) $\widetilde{f}$ is smooth on all of U;*
- *(b) $\widetilde{f} = f$ on $U \setminus W$;*
- *(c) the graphs $\mathrm{graph}(df)$ and $\mathrm{graph}(d\widetilde{f})$ are related by a Hamiltonian isotopy supported in $\pi^{-1}(W) \subset T^*U$.*

Proof. Choose a smooth cutoff function $\rho : U \to [0, 1]$ supported in W and equal to 1 on a smaller neighborhood $W' \subset W$ of the sector boundaries. Extend f to a slightly larger open set $U' \supset \overline{U}$ by any C^1 extension (possible since df is continuous up to the boundary and ∂U is piecewise smooth). Let η_ε be a standard mollifier on $\mathbb{R}^2$ of scale ε. For $\varepsilon > 0$ small, the convolution $f_\varepsilon := f * \eta_\varepsilon$ is smooth on U and converges to f in C^1 on compact subsets of U. Define

$$\widetilde{f} := f + \rho\,(f_\varepsilon - f).$$

Then $\widetilde{f}$ is smooth (because f_ε is smooth and ρ is smooth), and $\widetilde{f} = f$ where $\rho = 0$, i.e. outside W. This proves (a)–(b).

For (c), consider the homotopy $f_t := f + t\,\rho\,(f_\varepsilon - f)$. Then $\partial_t f_t$ is compactly supported in W and Lemma 8.1 gives a Hamiltonian isotopy between $\mathrm{graph}(df)$ and $\mathrm{graph}(d\widetilde{f})$ supported in $\pi^{-1}(W)$. $\square$

8.2.5 Global smoothing of K

We now smooth K in finitely many steps.

Theorem 8.6 (Quadrivalent polyhedral Lagrangians admit smoothing)**.** *Let $K \subset \mathbb{R}^4$ be a polyhedral Lagrangian surface which is a topological submanifold and such that exactly four faces meet at each vertex. Then K admits a Lagrangian smoothing: there exists a Hamiltonian isotopy K_t of smooth Lagrangian submanifolds for $t \in (0, 1]$ which extends to a topological isotopy on $[0, 1]$ with $K_0 = K$.*

Proof. Because K is a finite polyhedral complex, it has finitely many vertices and edges. Choose pairwise disjoint small balls B_v around each vertex v. For each open edge segment e with the vertex balls removed, cover it by finitely many balls B_x (as in Lemma 8.3) disjoint from the vertex balls. After shrinking, we may assume all these balls are pairwise disjoint.

Step 1: smooth along edges away from vertices. On each B_x the set $K \cap B_x$ is, in a Darboux chart, the graph of *df* for a function with a single crease along a line (Lemma 8.3). Choose W a thin neighborhood of that line, supported away from the boundary of the chart. Lemma 8.5 gives a Hamiltonian isotopy supported in B_x replacing $K \cap B_x$ by a smooth Lagrangian patch; since the supports are disjoint over different B_x, we can do this simultaneously for all such balls. After this step, the resulting polyhedral Lagrangian (still denoted K) is smooth along every edge except possibly inside the vertex balls.

Step 2: smooth at vertices. For each vertex ball B_v, Lemma 8.4 identifies $K \cap B_v$ with the graph of *df* for a piecewise–quadratic potential on a four–sector decomposition. Apply Lemma 8.5 with W a small neighborhood of the rays to obtain a smooth potential $\widetilde{f}$ equal to f near ∂U, hence producing a smooth Lagrangian patch which agrees with K near ∂B_v. Again, since the vertex balls are disjoint, the corresponding Hamiltonian isotopies can be performed simultaneously. After this step the resulting Lagrangian is smooth everywhere.

Step 3: assemble the isotopy and extend to $t = 0$. Each local modification was achieved by an ambient Hamiltonian isotopy supported in a disjoint union of balls, so their composition is a global Hamiltonian isotopy on $\mathbb{R}^4$ carrying the original polyhedral K to a smooth Lagrangian K_1. To obtain the family K_t for $t \in (0, 1]$, run the same construction with smoothing scale proportional to t, so that as $t \to 0$ the modified potentials converge to the original piecewise–quadratic ones in C^1 away from the singular locus and in C^0 globally. This yields a continuous family of embedded Lagrangians which extends, in the topology of compact subsets (or equivalently Hausdorff topology on each fixed compact set), to the original polyhedral surface K at $t = 0$. Because the isotopies are supported in arbitrarily small neighborhoods of the 1–skeleton, this extension is a topological isotopy with endpoint $K_0 = K$. □

Remark 8.7. The argument above is local and constructive. The hypothesis "four faces meet at every vertex" was used to ensure that the link of each vertex is a four–gon so that one can choose a Lagrangian projection for which the projected sectors do not overlap; the same method works for higher valence provided one can arrange a non–overlapping Lagrangian projection in each vertex chart.

9 Question 9: Polynomial tests for separability in scaled quadrifocal tensors

Restated statement

Fix $n \geq 5$. Let $A(1), \ldots, A(n) \in \mathbb{R}^{3\times 4}$ be Zariski–generic. For $\alpha, \beta, \gamma, \delta \in [n]$ define the $3 \times 3 \times 3 \times 3$ tensor $Q(\alpha\beta\gamma\delta)$ by

$$Q(\alpha\beta\gamma\delta)_{ijk\ell} := \det \begin{bmatrix} A(\alpha)(i,:) \\ A(\beta)(j,:) \\ A(\gamma)(k,:) \\ A(\delta)(\ell,:) \end{bmatrix}, \qquad 1 \leq i, j, k, \ell \leq 3.$$

Let $\lambda \in \mathbb{R}^{n\times n\times n\times n}$ satisfy $\lambda_{\alpha\beta\gamma\delta} \neq 0$ iff $\alpha, \beta, \gamma, \delta$ are *not all identical*.

Does there exist a polynomial map

$$F : \mathbb{R}^{81n^4} \to \mathbb{R}^N \qquad (N \text{ allowed to depend on } n)$$

with the properties:

(a) F does *not* depend on $A(1), \dots, A(n)$;

(b) the degrees of the coordinate functions of F are bounded by a constant independent of n; and

(c) $F(\{\lambda_{\alpha\beta\gamma\delta} Q(\alpha\beta\gamma\delta)\}) = 0$ holds iff λ is separable, i.e. $\lambda_{\alpha\beta\gamma\delta} = u_\alpha v_\beta w_\gamma x_\delta$ for some $u, v, w, x \in \mathbb{R}^n$.

Answer: yes (existence, bounded degree)

We give an existence proof, based on (i) elimination and (ii) *equivariant Noetherianity* (Noetherianity up to symmetry). The construction is non-constructive: it produces F as a finite list of generators of an elimination ideal; the degree bound comes from a symmetry theorem.

1. Universal parameterization

Let $\Bbbk$ be an algebraically closed field of characteristic 0 (e.g. $\mathbb{C}$); we work Zariski–geometrically over $\Bbbk$. Define the affine space

$$\mathcal{A}_n := (\Bbbk^{3\times 4})^n \cong \Bbbk^{12n}$$

parametrizing camera matrices $A(1), \dots, A(n)$. Let

$$G_n := (\Bbbk^*)^{4n}$$

be the algebraic torus with coordinates (u, v, w, x), where $u = (u_1, \dots, u_n)$ etc.

Let

$$V_n := (\Bbbk^{3\times 3\times 3\times 3})^{n^4} \cong \Bbbk^{81n^4}$$

be the ambient space of collections $T = \{T(\alpha\beta\gamma\delta)\}_{\alpha,\beta,\gamma,\delta\in[n]}$. Define a polynomial map

$$\Phi_n : \mathcal{A}_n \times G_n \longrightarrow V_n \tag{10}$$

by

$$(\Phi_n(A, u, v, w, x))(\alpha\beta\gamma\delta)_{ijk\ell} := u_\alpha v_\beta w_\gamma x_\delta \ \det \begin{bmatrix} A(\alpha)(i,:) \\ A(\beta)(j,:) \\ A(\gamma)(k,:) \\ A(\delta)(\ell,:) \end{bmatrix}.$$

Thus Φ_n simultaneously (i) builds all quadrifocal tensors from A and (ii) scales each tensor by the rank–one array $u \otimes v \otimes w \otimes x$.

Let

$$X_n := \overline{\operatorname{im}(\Phi_n)} \subseteq V_n$$

be the Zariski closure of the image. Then X_n is an affine variety whose defining ideal

$$I_n := I(X_n) \subseteq \Bbbk[V_n]$$

is independent of any particular choice of A. In particular, any set of generators of I_n defines a polynomial map F satisfying property (a).

2. Uniform bounded degree via Noetherianity up to symmetry

We now explain why one can choose generators of I_n in degrees bounded uniformly in n.

Symmetric group action. There is a natural action of $\mathfrak{S}_n$ on V_n by relabeling camera indices: for $\sigma \in \mathfrak{S}_n$ and $T \in V_n$ define

$$(\sigma \cdot T)(\alpha\beta\gamma\delta) := T(\sigma^{-1}\alpha\ \sigma^{-1}\beta\ \sigma^{-1}\gamma\ \sigma^{-1}\delta).$$

This action is linear and extends to an action on the coordinate ring $\Bbbk[V_n]$. The map Φ_n is $\mathfrak{S}_n$–equivariant (permute the $A(i)$ and simultaneously permute the entries of u, v, w, x), so X_n is $\mathfrak{S}_n$–stable and I_n is an $\mathfrak{S}_n$–stable ideal.

Passage to $n = \infty$. Let V_∞ be the direct limit of the V_n under the natural inclusions (extend a collection by adding zero tensors supported on indices $> n$). Equivalently, its coordinate ring is the polynomial ring

$$R := \Bbbk\big[x_{\alpha\beta\gamma\delta,ijk\ell} : \alpha, \beta, \gamma, \delta \in \mathbb{N},\ 1 \le i, j, k, \ell \le 3\big].$$

The infinite symmetric group $\mathfrak{S}_\infty$ acts on R by permuting the indices $\alpha, \beta, \gamma, \delta$. For each n, $\Bbbk[V_n]$ identifies with the subring of R generated by variables with $\alpha, \beta, \gamma, \delta \le n$.

The varieties X_n are compatible with these inclusions (by functoriality of Φ_n), so the ideals I_n form a directed system; let

$$I_\infty := \bigcup_{n \ge 1} I_n \subseteq R.$$

Then I_∞ is an $\mathfrak{S}_\infty$–stable ideal.

Equivariant Noetherianity theorem. A deep theorem due to Draisma (and refined by Sam–Snowden in the language of twisted commutative algebras) asserts that R is Noetherian *up to symmetry*: any $\mathfrak{S}_\infty$–stable ideal is generated by finitely many $\mathfrak{S}_\infty$–orbits of polynomials. Concretely:

Theorem 9.1 (Noetherianity up to symmetry, informal form)**.** *Let R be a polynomial ring in countably many variables equipped with the natural action of $\mathfrak{S}_\infty$ permuting the variables in finitely many "index slots". Then every $\mathfrak{S}_\infty$–stable ideal $J \subseteq R$ is generated by finitely many $\mathfrak{S}_\infty$–orbits. In particular, there exists a finite set $\{f_1, \dots, f_m\} \subseteq J$ such that*

$$J = \langle \mathfrak{S}_\infty \cdot f_1, \dots, \mathfrak{S}_\infty \cdot f_m \rangle.$$

Applying Theorem 9.1 to $J = I_\infty$, we obtain polynomials $f_1, \dots, f_m \in I_\infty$ involving only finitely many camera indices. Let

$$D := \max_{1 \le r \le m} \deg(f_r).$$

Then D is a constant independent of n. For each n, restricting the $\mathfrak{S}_\infty$–orbits to indices $\le n$ shows that I_n is generated by polynomials of degree $\le D$. This yields property (b).

3. The polynomial map F

Fix the finite generating set $\{f_1, \dots, f_m\}$ above. For each n, let $\mathcal{O}_n(f_r)$ denote the (finite) $\mathfrak{S}_n$–orbit of the restriction of f_r to $\Bbbk[V_n]$. Define

$$F_n : V_n \to \Bbbk^{N(n)}$$

to be the map whose coordinate functions are all polynomials in the union $\bigcup_{r=1}^m \mathcal{O}_n(f_r)$. Then:

- F_n depends only on n (and the chosen finite list of generators), not on any particular A;
- every coordinate of F_n has degree $\le D$, independent of n;
- $F_n(T) = 0$ iff $T \in X_n$.

Thus properties (a) and (b) are immediate.

4. Specialization to a generic A gives property (c)

Fix a Zariski–generic $A \in \mathcal{A}_n$, and write $Q_A(\alpha\beta\gamma\delta)$ for the corresponding quadrifocal tensors. Define the linear embedding

$$\iota_A : \Bbbk^{n^4} \hookrightarrow V_n, \qquad (\lambda_{\alpha\beta\gamma\delta}) \mapsto \{\lambda_{\alpha\beta\gamma\delta} Q_A(\alpha\beta\gamma\delta)\}.$$

(We use all 3^4 entries of each tensor, so this is an embedding on the open set where at least one entry of each $Q_A(\alpha\beta\gamma\delta)$ is nonzero.)

"If" direction. If λ is separable, $\lambda = u \otimes v \otimes w \otimes x$, then

$$\iota_A(\lambda) = \Phi_n(A, u, v, w, x) \in \mathrm{im}(\Phi_n) \subseteq X_n$$

so $F_n(\iota_A(\lambda)) = 0$.

"Only if" direction (generic identifiability argument). Assume λ has the stated nonvanishing pattern and that

$$F_n(\iota_A(\lambda)) = 0.$$

Then $\iota_A(\lambda) \in X_n$. Because $\lambda_{\alpha\beta\gamma\delta} \neq 0$ whenever not all indices are identical, the point $\iota_A(\lambda)$ lies in the Zariski–open subset $U \subset V_n$ where all coordinates corresponding to *non-forced* determinants are nonzero. (For generic A the only forced zeros come from repeating a row vector in the determinant.) On this open set, taking Zariski closures does not add new points, so

$$X_n \cap U = \mathrm{im}(\Phi_n) \cap U.$$

Hence there exist $(A', u, v, w, x) \in \mathcal{A}_n \times G_n$ such that

$$\iota_A(\lambda) = \Phi_n(A', u, v, w, x). \tag{11}$$

Unwinding definitions, (11) says that for every $\alpha, \beta, \gamma, \delta$,

$$\lambda_{\alpha\beta\gamma\delta}\, Q_A(\alpha\beta\gamma\delta) \;=\; (u_\alpha v_\beta w_\gamma x_\delta)\, Q_{A'}(\alpha\beta\gamma\delta). \tag{12}$$

Define the blockwise scalar ratio

$$\mu_{\alpha\beta\gamma\delta} := \frac{\lambda_{\alpha\beta\gamma\delta}}{u_\alpha v_\beta w_\gamma x_\delta}.$$

Then (12) is equivalent to

$$Q_{A'}(\alpha\beta\gamma\delta) = \mu_{\alpha\beta\gamma\delta}\, Q_A(\alpha\beta\gamma\delta) \qquad \text{for all } \alpha, \beta, \gamma, \delta. \tag{13}$$

At this point we use the Zariski–genericity of A (and $n \geq 5$). The relevant identifiability of generic multiview data from determinantal (Plücker) coordinates is standard; see for instance Aholt–Sturmfels–Thomas [14] for multiview varieties and their generic fibers. The family of tensors $\{Q_A(\alpha\beta\gamma\delta)\}$ encodes the full set of 4×4 minors of the $4 \times (3n)$ matrix whose columns are the transposed row vectors of the $A(i)$. For a Zariski–generic configuration of $3n$ vectors in $\Bbbk^4$, those Plücker coordinates determine the vectors up to the standard right action of GL_4 and scaling of each vector. A relation of the form (13) therefore forces μ to come from a coordinatewise rescaling of the underlying vectors. Because $\mu_{\alpha\beta\gamma\delta}$ is constant across the 3^4 choices of row indices inside each

$Q(\alpha\beta\gamma\delta)$, this rescaling must in turn be constant on each camera index in each slot; equivalently, μ is separable:

$$\mu_{\alpha\beta\gamma\delta} = \tilde{u}_\alpha \tilde{v}_\beta \tilde{w}_\gamma \tilde{x}_\delta.$$

(One can make this explicit by choosing five indices to fix a basis of $\Bbbk^4$ and applying Cramer's rule to compare determinants across overlapping quadruples; the resulting multiplicative cocycle conditions integrate to the displayed outer product.)

Finally, since $\lambda = (u \otimes v \otimes w \otimes x) \cdot \mu$ and both factors are separable, λ is separable. This proves property (c).

Conclusion

Taking F_n to be any bounded-degree generating set of $I(X_n)$ constructed above gives a polynomial map independent of A whose coordinate degrees are uniformly bounded in n and which vanishes exactly on scaled data coming from separable λ (for Zariski–generic A and the given nonvanishing pattern).

Remark 9.2. The proof is intentionally "existential." In principle one can compute explicit equations (and hence an explicit F_n) by eliminating the camera variables and the torus variables from the defining equations of Φ_n. Uniform bounded degree is a genuinely additional input: it comes from the equivariant Noetherianity theorem.

Remark 9.3 (On genericity and identifiability). The "only if" direction should be read in the Zariski–generic sense stated in the question: outside a proper algebraic subset of camera matrices A, the multiview parameterization is generically identifiable (up to the standard projective gauge). This is a classical theme in multiview geometry; one algebraic treatment is via the multiview varieties studied by Aholt–Sturmfels–Thomas [14].

10 Question 10: PCG for the RKHS-mode ALS subproblem with missing data

Statement

We are given a d-way tensor with missing entries, and we consider an alternating optimization (ALS-type) method for a CP decomposition of rank r where mode k is infinite-dimensional and constrained to lie in an RKHS. In the mode-k subproblem, all other factor matrices are fixed and we solve for A_k.

Let $n := n_k$ and $M := \prod_{i \neq k} n_i$ so that the mode-k unfolding is a matrix $T \in \mathbb{R}^{n \times M}$ (with unobserved entries set to 0). Let $N = nM$ and let $q \ll N$ be the number of observed entries. Let $S \in \mathbb{R}^{N \times q}$ be the selection matrix such that $S^T \operatorname{vec}(T)$ is the vector of observed entries.

Let

$$Z \;=\; A_d \odot \cdots \odot A_{k+1} \odot A_{k-1} \odot \cdots \odot A_1 \in \mathbb{R}^{M \times r}$$

be the Khatri–Rao product of the other factor matrices, and let $B = TZ$ be the (sparse) MTTKRP. Assume $A_k = KW$ where $K \in \mathbb{R}^{n \times n}$ is the PSD kernel matrix for mode k and $W \in \mathbb{R}^{n \times r}$ is the unknown. We must solve the linear system (size $nr \times nr$)

$$\Big[(Z \otimes K)^T S S^T (Z \otimes K) \;+\; \lambda (I_r \otimes K)\Big] \operatorname{vec}(W) \;=\; (I_r \otimes K) \operatorname{vec}(B), \tag{14}$$

with $\lambda > 0$, without forming any $N \times N$ objects and avoiding any $\Theta(N)$ work.

Solution

10.1 Basic structure: SPD normal equations

Let $P_\Omega := SS^T \in \mathbb{R}^{N\times N}$. Since S is a selection matrix, P_Ω is a diagonal mask: it has a 1 on observed entries and 0 elsewhere. In particular P_Ω is symmetric positive semidefinite.

Define the coefficient matrix

$$A \;:=\; (Z\otimes K)^T P_\Omega (Z\otimes K) \;+\; \lambda(I_r\otimes K).$$

If K is positive definite (or if we add a tiny "nugget" $K \leftarrow K+\delta I$), then A is symmetric positive definite (SPD), so we can solve (14) using Conjugate Gradients (CG). Even when K is only PSD, A is SPD on the natural range space induced by the ridge term and in practice the same PCG approach is used after a small regularization.

10.2 Key identities for fast matrix–vector products

The core idea is that CG/PCG never needs to form A; it only needs a routine that computes $y \leftarrow Ax$.

Let $x = \mathrm{vec}(X)$ with $X \in \mathbb{R}^{n\times r}$. We use the Kronecker identity

$$(A\otimes B)\,\mathrm{vec}(X) = \mathrm{vec}(BXA^T).$$

With $A = Z$ and $B = K$, this gives

$$(Z\otimes K)\,\mathrm{vec}(X) \;=\; \mathrm{vec}(KXZ^T). \tag{15}$$

The matrix KXZ^T has size $n\times M$, i.e. length $N = nM$ when vectorized, and cannot be formed explicitly. However, we do not need all of its entries: the mask P_Ω (equivalently S^T) extracts only the q observed positions.

Observed-entry indexing. Let $\Omega \subset [n]\times[M]$ denote the observed set in the mode-k unfolding. Write $\Omega = \{(i_t, j_t)\}_{t=1}^q$ so that the tth observed entry corresponds to row $i_t \in [n]$ and column $j_t \in [M]$ in the unfolding. Then for any $Y \in \mathbb{R}^{n\times M}$ we have

$$S^T\,\mathrm{vec}(Y) \;=\; (\,Y_{i_1 j_1},\, Y_{i_2 j_2},\, \ldots,\, Y_{i_q j_q}\,)^T.$$

Therefore, if $Y = KXZ^T$ then each observed entry is

$$Y_{i_t j_t} \;=\; (KX)_{i_t,:}\cdot Z_{j_t,:} \qquad (t = 1,\ldots,q), \tag{16}$$

a dot product of two length-r vectors.

10.3 Exact matvec $y = Ax$ using only Ω (no $\Theta(N)$ work)

We now show how to compute $y = Ax$ exactly with cost proportional to q and to kernel multiplies.

Step A: apply $Z\otimes K$ and then select observed entries. Given $x = \mathrm{vec}(X)$, compute

$$U := KX \in \mathbb{R}^{n\times r}.$$

Then compute the length-q vector

$$y_\Omega := S^T(Z\otimes K)x \in \mathbb{R}^q$$

via (16):

$$(y_\Omega)_t \;=\; U_{i_t,:}\cdot Z_{j_t,:}, \qquad t = 1,\ldots,q.$$

This costs $\Theta(qr)$ once we can read the rows $Z_{j_t,:}$.

Avoid forming Z explicitly. When M is huge we do not form $Z \in \mathbb{R}^{M\times r}$. Because Z is a Khatri–Rao product, each row $Z_{j,:}$ can be computed from the corresponding multi-index $(\alpha_1, \ldots, \alpha_{k-1}, \alpha_{k+1}, \ldots, \alpha_d)$ as

$$Z_{j,:} \;=\; A_1(\alpha_1,:) \odot \cdots \odot A_{k-1}(\alpha_{k-1},:) \odot A_{k+1}(\alpha_{k+1},:) \odot \cdots \odot A_d(\alpha_d,:),$$

an elementwise product of $d-1$ length-r vectors.

Recovering the multi-index from j. Assume the unfolding columns are ordered in the standard lexicographic (row-major) way: $j = 1 + \sum_{m\neq k}(\alpha_m - 1)\, s_m$ with strides $s_m := \prod_{\ell<m,\ \ell\neq k} n_\ell$ (so the last mode index varies fastest). Then one can recover α_m by repeated division:

$$\alpha_m \;=\; 1 + \Big\lfloor \frac{j-1}{s_m} \Big\rfloor \bmod n_m, \qquad (m \neq k),$$

which costs $\Theta(d)$ integer operations (or $\Theta(1)$ if one precomputes the tuples for the distinct j_t that occur).

Thus each needed $Z_{j_t,:}$ can be generated on the fly in $\Theta((d-1)r)$ time and (optionally) cached if the same j_t repeats.

Step B: scatter back and apply $(Z \otimes K)^T$. We must next apply $(Z \otimes K)^T S$ to y_Ω. Let $Y_\Omega \in \mathbb{R}^{n\times M}$ be the sparse matrix whose only nonzeros are

$$(Y_\Omega)_{i_t j_t} = (y_\Omega)_t, \qquad t = 1, \ldots, q.$$

Then $P_\Omega \operatorname{vec}(KXZ^T) = \operatorname{vec}(Y_\Omega)$ and

$$(Z \otimes K)^T P_\Omega (Z \otimes K)x = (Z^T \otimes K)\operatorname{vec}(Y_\Omega) = \operatorname{vec}(K\, Y_\Omega\, Z),$$

again by the Kronecker identity. We can compute the $n \times r$ matrix $H := Y_\Omega Z$ without forming Y_Ω: initialize $H = 0 \in \mathbb{R}^{n\times r}$ and accumulate

$$H_{i_t,:} \;\mathrel{+}=\; (y_\Omega)_t\, Z_{j_t,:} \qquad \text{for } t = 1, \ldots, q. \tag{17}$$

This is another $\Theta(qr)$ pass over the observed entries. Then compute $G := KH$ and obtain

$$(Z \otimes K)^T P_\Omega (Z \otimes K)x \;=\; \operatorname{vec}(G).$$

Step C: add the ridge term. Finally,

$$\lambda(I_r \otimes K)\operatorname{vec}(X) = \lambda \operatorname{vec}(KX) = \lambda \operatorname{vec}(U),$$

so the full matvec is

$$Ax = \operatorname{vec}(G + \lambda U).$$

Right-hand side without $\Theta(N)$ work. The RHS is $(I_r \otimes K)\operatorname{vec}(B) = \operatorname{vec}(KB)$. We compute $B = TZ$ as a sparse MTTKRP: if t_t denotes the observed tensor value at (i_t, j_t) in the unfolding, then

$$B_{i_t,:} \;\mathrel{+}=\; t_t\, Z_{j_t,:} \qquad (t = 1, \ldots, q),$$

which costs $\Theta(qr)$, and then multiply by K once to obtain KB.

10.4 PCG and a practical preconditioner

Why CG/PCG is appropriate. CG solves SPD linear systems using only: (i) one matvec $x \mapsto Ax$ per iteration, and (ii) vector inner products and saxpy operations, all in $\Theta(nr)$ time. PCG replaces the Euclidean inner product with the one induced by a preconditioner $M \simeq A$, which can greatly reduce the number of iterations.

A Kronecker preconditioner that avoids forming Z. A common and effective preconditioner approximates the masked normal matrix by a separable Kronecker form. Let

$$G_Z := Z^T Z \in \mathbb{R}^{r \times r}.$$

Even if M is enormous, G_Z can be computed *without forming* Z using the Khatri–Rao Gram identity: if $Z = \odot_{m \neq k} A_m$, then

$$G_Z = Z^T Z \; = \; (A_1^T A_1) \circ \cdots \circ (A_{k-1}^T A_{k-1}) \circ (A_{k+1}^T A_{k+1}) \circ \cdots \circ (A_d^T A_d), \tag{18}$$

where $\circ$ denotes the Hadamard (entrywise) product. Computing each $A_m^T A_m$ costs $\Theta(n_m r^2)$ and the Hadamard product costs $\Theta(dr^2)$.

We then take the preconditioner

$$M \; := \; (G_Z + \lambda I_r) \otimes (K + \delta I_n), \tag{19}$$

with a small $\delta \geq 0$ (often $\delta = 0$ if K is strictly PD). This captures: (i) the coupling across rank components via G_Z, and (ii) the dominant RKHS geometry via K, while remaining easy to apply.

Applying M^{-1} cheaply. Given a vector $b = \mathrm{vec}(B_0)$ with $B_0 \in \mathbb{R}^{n \times r}$, solving $M\,x = b$ is equivalent to solving the matrix equation

$$(K + \delta I)\, X\, (G_Z + \lambda I_r) = B_0, \qquad x = \mathrm{vec}(X).$$

This can be done by two small solves:

1. Solve $(K + \delta I)\, Y = B_0$ for $Y \in \mathbb{R}^{n \times r}$ (e.g. using a precomputed Cholesky of $K + \delta I$).
2. Solve $X\,(G_Z + \lambda I_r) = Y$ for X (i.e. right-multiply by $(G_Z + \lambda I_r)^{-1}$, via an $r \times r$ Cholesky).

The cost is $\Theta(\mathrm{solve}(K) \cdot r + nr^2)$, which is typically far smaller than forming or factoring the $nr \times nr$ matrix A.

PCG outline. With the matvec routine from §3 and the preconditioner (19), PCG proceeds as follows: start with w_0 (often 0), set $r_0 = b - Aw_0$, solve $z_0 = M^{-1} r_0$, set $p_0 = z_0$, and iterate

$$\alpha_k = \frac{\langle r_k, z_k \rangle}{\langle p_k, Ap_k \rangle}, \quad w_{k+1} = w_k + \alpha_k p_k, \quad r_{k+1} = r_k - \alpha_k A p_k, \quad z_{k+1} = M^{-1} r_{k+1}, \quad \beta_k = \frac{\langle r_{k+1}, z_{k+1} \rangle}{\langle r_k, z_k \rangle}, \quad p_{k+1} = z_{k+1}$$

until $\|r_k\| / \|b\|$ is below tolerance.

10.5 Complexity (avoiding $\Theta(N)$ work)

Let $\mathrm{mv}(K)$ denote the cost of multiplying K by one vector; for a dense kernel matrix this is $\Theta(n^2)$, and for structured/approximate kernels it can be much smaller. One PCG iteration requires:

- two kernel multiplies by K on r right-hand sides (to form $U = KX$ and $G = KH$), costing $\Theta(r\,\mathrm{mv}(K))$ each;
- two passes over the q observed entries: one to compute $(y_\Omega)_t = U_{i_t,:} \cdot Z_{j_t,:}$ and one to scatter (17), each $\Theta(qr)$;
- vector operations in $\Theta(nr)$.

Thus the matvec cost is

$$\Theta\big(qr + r\,\mathrm{mv}(K)\big),$$

up to constant factors, and it never scales with $N = nM$. Applying the Kronecker preconditioner costs

$$\Theta\big(r\,\mathrm{solve}(K) + nr^2\big),$$

after a one-time factorization of $K + \delta I$ and $G_Z + \lambda I_r$. If PCG converges in I iterations, the total solve cost is

$$\Theta\Big(I\big(qr + r\,\mathrm{mv}(K)\big)\Big) \quad \text{(plus preconditioner applications).}$$

Since $n, r < q \ll N$, this replaces the prohibitive $\Theta((nr)^3)$ direct solve and avoids any $\Theta(N)$ work or storage.

Remark 10.1 (Practical notes).

- If K is very large, one often uses a fast kernel approximation (Nyström, random features) so that both $\mathrm{mv}(K)$ and $\mathrm{solve}(K)$ are sub-quadratic in n.
- The preconditioner (19) is intentionally simple and robust; stronger preconditioners can incorporate sampling weights from the mask (row/column counts of Ω), apply Jacobi (diagonal) scaling, or use better approximations to the masked Gram $(Z \otimes K)^T P_\Omega (Z \otimes K)$.

References

[1] First Proof benchmark paper. *First Proof: A Benchmark for Mathematical Proving*, arXiv:2602.05192v1 (2026).

[2] A. Marcus, D. A. Spielman, and N. Srivastava. Finite free convolutions of polynomials. *Probability Theory and Related Fields* **168** (2017), 659–727. (Preprint: arXiv:1504.00350.)

[3] M. Hairer. Φ^4_3 is orthogonal to the Gaussian free field (and in particular not quasi-invariant under nonzero smooth shifts). Unpublished note (Sept. 16, 2022), available at `hairer.org/Phi4.pdf`.

[4] W. Casselman and J. Shalika. The unramified principal series of p-adic groups. II. The Whittaker function. *Compositio Math.* **41** (1980), no. 2, 207–231.

[5] H. Jacquet, I. Piatetski-Shapiro, and J. Shalika. Conducteur des représentations du groupe linéaire. *Math. Ann.* **256** (1981), 199–214.

[6] H. Ben Dali and L. K. Williams. A combinatorial formula for interpolation Macdonald polynomials. arXiv:2510.02587 (2025).

[7] A. Ayyer, J. Martin, and L. Williams. t-PushTASEP and Macdonald polynomials. arXiv:2403.10485 (2024).

[8] J. Fowler. *Poincaré duality groups and homology manifolds*. PhD thesis, University of Chicago, 2009. Available at `kisonecat.com/research/thesis.pdf`.

[9] J. Fowler. Finiteness properties for some rational Poincaré duality groups. arXiv:1204.4667 (2012).

[10] M. A. Hill. The equivariant slice filtration: a primer. arXiv:1107.3582 (2011).

[11] J. Ullman. On the slice spectral sequence. arXiv:1206.0058 (2012).

[12] A. Blumberg and M. Hill. Operadic multiplications in equivariant spectra, norms, and transfers. *Adv. Math.* **285** (2015), 658–708.

[13] M. Hill, M. Hopkins, and D. Ravenel. On the nonexistence of elements of Kervaire invariant one. *Ann. of Math.* **184** (2016), no. 1, 1–262.

[14] C. Aholt, B. Sturmfels, and R. Thomas. A Hilbert scheme in computer vision. *Canad. J. Math.* **65** (2013), no. 5, 961–988.

[15] J. Draisma. Noetherianity up to symmetry. *Compositio Math.* **153** (2017), no. 6, 1235–1274.

[16] S. Sam and A. Snowden. GL-equivariant modules over polynomial rings in infinitely many variables. *Trans. Amer. Math. Soc.* **368** (2016), no. 2, 1097–1158.

[17] J. A. Mingo and R. Speicher. *Free Probability and Random Matrices.* Fields Institute Monographs, vol. 34. Springer, 2017.

[18] A. W. Marcus. Polynomial convolutions and (finite) free probability. *arXiv preprint* arXiv:2108.07054, 2021.

[19] D. Voiculescu. The analogues of entropy and of Fisher's information measure in free probability theory, I. *Comm. Math. Phys.* **155** (1993), 71–92.

[20] S. Weinberger. *Variations on a Theme of Borel.* Draft manuscript, available at `math.uchicago.edu/~shmuel/VTBdraft.pdf`.