QUADRATIC MIXING FOR THE LEE–SIDFORD DIKIN WALK

BY YIPING LU[1,a]

[1] *Industrial Engineering & Management Sciences, Northwestern University,* [a] *yiping.lu@northwestern.edu*

Let $K = \{x \in \mathbb{R}^d : Ax \geq b\}$ be a bounded full-dimensional polytope with m inequalities. For targets on K with density proportional to $\exp(-\vartheta^\mathsf{T} x)$, we prove a $\widetilde{O}(d^2)$ bound, measured in Markov chain iterations, for the chi-square mixing of the $1/2$-lazy Metropolis Dikin walk based on a polylogarithmically rescaled Lee–Sidford metric. The estimate holds for every initial law with finite chi-square divergence. The proof controls the reverse Gaussian proposal through a quantitative average self-concordance estimate at the required $O(1/d)$ scale. We establish this estimate from uniform analytic derivative bounds for the Lewis-weight subspace and an all-order projection-weighted contraction inequality for Gaussian chaoses. For order j, the latter bounds each base-point polynomial at scale $d^{j/2}$, up to j-dependent and logarithmic factors. Truncating at logarithmic order then absorbs the remaining pathwise loss. **This resolves, in the affirmative, the open problem of whether the Gaussian Lee–Sidford Dikin walk mixes in $\widetilde{O}(d^2)$ iterations: the previously stated quadratic bounds [22, Theorem 1.4], [8, Theorem 1.2] rest on reverse-proposal steps identified as unresolved in [15, Section 1], whose $\widetilde{O}(d^{9/4})$ bound was the best rigorously established rate prior to this work. The proof is found by ChatGPT 5.6 Pro. The author verified the proof and promised to fine-tune the draft later.**

1. Introduction. Interior-point geometry supplies affine-invariant local coordinates for both optimization and sampling. At its current point, a Dikin walk proposes a Gaussian increment with covariance proportional to the inverse of a barrier metric and then applies a Metropolis correction. The metric adapts to the boundary, but the same state dependence creates the principal probabilistic difficulty: the reverse Gaussian proposal is evaluated in a random metric at the proposed point.

The logarithmic-barrier Dikin walk has an iteration bound depending polynomially on the number of inequalities [14, 31, 34]. Lewis weights originate in functional analysis [28], became algorithmic row-sampling and barrier tools [4, 25, 26], and subsequently entered local sampling geometries [3, 22]. We ask whether the resulting Lee–Sidford local-metric chain has a mixing bound quadratic in the ambient dimension, up to logarithmic factors.

1.1. *Problem setting and main results.* Let $K = \{x \in \mathbb{R}^d : Ax \geq b\}$ be bounded and full-dimensional, where $A \in \mathbb{R}^{m\times d}$ has full column rank and no zero row. The target distributions are

$$\pi_\vartheta(\mathrm{d}x) = Z_\vartheta^{-1} e^{-\vartheta^\mathsf{T} x}\mathbf{1}_K(x)\,\mathrm{d}x, \qquad \vartheta \in \mathbb{R}^d, \tag{1.1}$$

and include the uniform distribution when $\vartheta = 0$. For probability measures $\mu \ll \pi$, write

$$\chi^2(\mu\|\pi) := \int \left(\frac{\mathrm{d}\mu}{\mathrm{d}\pi} - 1\right)^2 \mathrm{d}\pi. \tag{1.2}$$

MSC2020 subject classifications. Primary 60J10, 65C05; secondary 52A40, 60G15, 68W20, 90C51.

Keywords and phrases. Dikin walk, Lee–Sidford metric, Lewis weights, average self-concordance, Wiener chaos.

Thus finite chi-square divergence is exactly the condition $\mathrm{d}\mu/\mathrm{d}\pi - 1 \in L^2(\pi)$.

Given a local metric g and a proposal radius $\varrho > 0$, write

$$Q_g^{\varrho}(x,\cdot) = N\left(x, \frac{\varrho^2}{d} g(x)^{-1}\right).$$

The Metropolis Dikin kernel, with the usual Metropolis–Hastings correction [10], proposes $y \sim Q_g^{\varrho}(x,\cdot)$, rejects unless $y \in \operatorname{int} K$, and otherwise accepts with probability

$$1 \wedge \frac{\pi_\vartheta(y) q_y(x)}{\pi_\vartheta(x) q_x(y)}, \qquad q_x(y) = \frac{\det(g(x))^{1/2}}{(2\pi\varrho^2/d)^{d/2}} \exp\left(-\frac{d}{2\varrho^2} \|y-x\|_{g(x)}^2\right). \tag{1.3}$$

The $1/2$-lazy version stays at its current state with probability $1/2$ and otherwise applies this kernel.

The precise question is whether a polylogarithmic rescaling of the Lee–Sidford metric, together with a dimension-independent proposal radius, makes this lazy kernel mix in $\widetilde{O}(d^2)$ iterations, uniformly over ϑ, from every initial law whose density belongs to $L^2(\pi_\vartheta)$. The obstruction is not the forward Gaussian proposal itself but the random quadratic form in its reverse density. We must control that form at scale $O(\varrho^2/d)$ with high probability and then obtain overlap of the accepted densities, without assuming a function whose Hessian is comparable to g. The following quantitative average self-concordance condition isolates the required probabilistic input.

DEFINITION 1.1 (Quantitative average self-concordance). For a metric g and $\varepsilon \in (0,1)$, its quantitative average self-concordance (ASC) radius $R_\varepsilon(g)$ is the supremum of radii $r \ge 0$ such that, for every $x \in \operatorname{int} K$, every $0 < s \le r$, and $z \sim N(x, s^2 g(x)^{-1}/d)$,

$$\mathbb{P}\left(\left|\|z-x\|_{g(z)}^2 - \|z-x\|_{g(x)}^2\right| \le \frac{2\varepsilon s^2}{d},\ z \in \operatorname{int} K\right) \ge 1-\varepsilon. \tag{1.4}$$

The ASC condition in [16] is one-sided and does not place an absolute value around the metric discrepancy. Definition 2.1 of [15] makes the two-sided event and the interior convention explicit. We adopt those two conventions and package the dependence on ε by introducing the maximal admissible radius $R_\varepsilon(g)$. Only the one-sided implication is needed in the mixing reduction; the argument below proves the two-sided form.

THEOREM 1.2 (Quantitative average self-concordance). *There are universal constants $c_{\mathrm{asc}}, C_{\mathrm{asc}} > 0$ such that for every $p \ge 4$ and every $\varepsilon \in (0,1/4)$, the baseline ℓ_p Lee–Sidford metric g_0 defined in (2.2) satisfies*

$$R_\varepsilon(g_0) \ge \underline{r}_\varepsilon := c_{\mathrm{asc}} \varepsilon^3 p^{-C_{\mathrm{asc}}} \left(1 + \log \frac{2(m+d+1)}{\varepsilon}\right)^{-C_{\mathrm{asc}}}. \tag{1.5}$$

The same constants work uniformly for all m, d, p, ε, A, and b.

Set

$$p_m := 4 + 2\log(2m), \qquad \lambda_m := (1+p_m^2)(1+p_m), \qquad \varepsilon_0 := \frac{1}{20}, \tag{1.6}$$

and define

$$\widehat{r}_0 := \min\left\{1,\ c_{\mathrm{asc}} \varepsilon_0^3 p_m^{-C_{\mathrm{asc}}} \left(1 + \log \frac{2(m+d+1)}{\varepsilon_0}\right)^{-C_{\mathrm{asc}}}\right\}, \qquad L_{m,d} := \lambda_m \widehat{r}_0^{\,-2}. \tag{1.7}$$

THEOREM 1.3 (Quadratic mixing in Markov chain iterations). *Let g_0 be the baseline metric defined in* (2.2)*, with $p = p_m$, and set $g_\star = L_{m,d} g_0$. There is a proposal radius $\varrho_\star > 0$, depending only on universal constants, such that for every target* (1.1)*, every $\mu_0 \ll \pi_\vartheta$ with $\chi^2(\mu_0 \| \pi_\vartheta) < \infty$, and every $\delta \in (0,1)$, the law μ_T of the 1/2-lazy walk after any integer $T \geq 0$ satisfies $\chi^2(\mu_T \| \pi_\vartheta) \leq \delta$ whenever*

$$T \geq C d^2 L_{m,d} \log\left(\frac{1 + \chi^2(\mu_0 \| \pi_\vartheta)}{\delta}\right). \tag{1.8}$$

Moreover, $L_{m,d} = \mathrm{polylog}(m, d)$.

Theorem 1.3 counts Markov chain iterations only. It excludes the cost of computing exact or approximate Lewis weights [4, 6] and makes no membership- or evaluation-oracle complexity claim. No affine rounding of K is required. The sole initialization condition is finite chi-square divergence. In particular, if $\mathrm{d}\mu_0 / \mathrm{d}\pi_\vartheta \leq M$, then $1 + \chi^2(\mu_0 \| \pi_\vartheta) \leq M$, so the displayed bound is $C d^2 L_{m,d} \log(M/\delta)$.

1.2. *Prior work and scope.*

Barrier metrics for local sampling.. For a polytope with m inequalities in dimension d, the logarithmic-barrier Dikin walk has a $\widetilde{O}(md)$ warm-start iteration bound for uniform sampling [14, 34]; Narayanan extended the barrier-based viewpoint to more general convex sets and randomized interior-point methods [31]. Lewis weights originate in Banach-space geometry [28]. Their algorithmic computation was developed in [4, 6], while approximate John and Lewis-weight barrier constructions for linear programming were developed in [25, 26]. These weighted geometries were subsequently used to construct local-metric sampling algorithms [3, 22].

Chen, Dwivedi, Wainwright and Yu proved that the Vaidya walk mixes from an M-warm start in $\widetilde{O}(m^{1/2} d^{3/2})$ iterations [3, Theorem 1]. Under the additional condition $m \leq \exp(\sqrt{d})$, their Gaussian walk based on approximate John weights has mixing time $\widetilde{O}(d^{5/2})$ [3, Theorem 2]. They conjectured a $\widetilde{O}(d^2)$ bound for a differently scaled improved John walk [3, Conjecture 5]. These results concern the uniform law, total-variation error and an M-warm initialization. Our kernel instead uses the ℓ_p Lee–Sidford metric, treats every exponential tilt, and proves chi-square contraction from every finite-chi-square initial law; it therefore does not resolve the improved-John conjecture. The exact John-ellipsoid walk of Gustafson and Narayanan is a distinct algorithm: it samples uniformly from the maximum-volume ellipsoid of the symmetrized body and has a $\widetilde{O}(d^7)$ warm-start bound [9].

Quadratic Lee–Sidford claims and the reverse proposal.. Laddha, Lee and Vempala stated a $\widetilde{O}(d^2)$ bound for their Lee–Sidford Dikin walk [22, Theorem 1.4]; their proposal is uniform on a local ellipsoid rather than Gaussian. Gu et al. stated a robust Lee–Sidford bound

$$\widetilde{O}\Big((d^2 + dL^2R^2) \log \frac{w}{\delta}\Big)$$

for an L-Lipschitz potential on a body contained in a radius-R ball, from a w-warm start [8, Theorem 1.2]. Kook [15, Section 1] identifies unresolved reverse-proposal steps in the published proofs of these bounds: in the former argument the current point must be controlled in the reverse Dikin ellipsoid, while the proof sketch of [8, Lemma D.5] does not establish the $O(\varrho^2/d)$ quadratic-form scale required by average self-concordance. We therefore record these bounds as previously stated results but do not use them as inputs. This issue concerns only the reverse-density estimate; the SSC, log-determinant convexity and symmetry lemmas from [22] remain independent inputs below.

For the same Gaussian Lee–Sidford Dikin chain considered here, the recent preprint [15, Theorem 1.1] proves average self-concordance after a $\widetilde{O}(d^{1/4})$ metric rescaling and obtains $\widetilde{O}(d^{9/4})$ iterations for exponential targets. It also introduces the selective higher-order expansion, moving orthonormal frames and Wiener-chaos organization used here, and observes that order-k estimates would lead to exponent $2 + 1/(k+1)$ [15, Section 1.3]. Our contribution is to supply all-order analytic frame and contraction estimates, allowing the rescaling loss to be reduced to polylogarithmic size.

Other local-metric samplers.. Jiang and Chen prove $\widetilde{O}(md + \kappa d)$ iterations for the soft-threshold Dikin walk and $\widetilde{O}(d^{5/2} + \kappa d)$ for their regularized Lewis-weight Dikin walk, in the notation and under the conditions of that work, with extensions to weakly log-concave truncations and beyond-worst-case bounds [13]. Gatmiry, Kelner and Vempala analyze Riemannian HMC using a hybrid of the Lewis-weight and logarithmic barriers and obtain $\widetilde{O}(m^{1/3}d^{4/3})$ mixing for polytope sampling [7]. These are different kernels, and their complexity expressions are not directly comparable to the iteration-only bound proved here.

Oracle samplers and global geometry.. For a convex body specified by a well-defined membership oracle $B_1(x_0) \subseteq K \subseteq B_D(x_0)$, the restart version of In-and-Out, initialized from an M-warm law, achieves for every $q \geq 2$ a finite-order Rényi guarantee using

$$\widetilde{O}\left(qMd^2\Lambda \log^6 \frac{1}{\epsilon}\right), \qquad \Lambda := \|\operatorname{cov} \pi\|_{\mathrm{op}},$$

expected membership queries [20, Theorem 3 and Corollary 4]. For an arbitrary log-concave density $\pi \propto e^{-V}$ given by an evaluation oracle, [19, Theorem 1.1] assumes

$$B(0,1) \subseteq \{x : V(x) - \inf V \leq 10d\}, \qquad M_q := \|\,\mathrm{d}\pi_0/\,\mathrm{d}\pi\|_{L^q(\pi)} \leq 10,$$

and gives $\widetilde{O}(qd^2\Lambda \log^3(1/\epsilon))$ expected evaluation queries, where

$$q \geq \max\Big\{2, \widetilde{\Omega}\big(\log\big(d^2\Lambda \log(1/\epsilon)\big)\big)\Big\}.$$

Algorithmic diffusion and its subsequent cold-start improvements are developed in [17, 18]. Lee and Vempala prove an $O(\sqrt{d})$ Poincaré bound for isotropic log-concave measures and derive a $\widetilde{O}(d^{5/2})$ warm-start bound for the ball walk [27]; this is global Euclidean isoperimetric context, distinct from the local barrier geometry used here.

These oracle results concern different algorithms and include global normalization and covariance parameters. Our result instead counts iterations of an affine-invariant local chain given an explicit inequality description, requires no global rounding, and excludes the arithmetic cost of computing the Lee–Sidford metric. It is therefore neither an oracle-complexity comparison nor a claim to be the first quadratic-dimension sampler for general log-concave targets.

1.3. *Proof strategy and contributions.* The proof first closes the abstract mixing bridge for the present linear targets. Lemma 2.2 works directly with the common accepted density, rather than with two separate rejection estimates. It constructs good events for both starting points, controls the covariance and displaced-mean terms in the comparison of q_x and q_y, intersects the events with a common target halfspace, and keeps the interior condition explicit. This is the step that removes the function-counterpart hypothesis used in the general framework of [16].

The ASC proof then follows the higher-order route of [15, Sections 1.2–1.3 and 3]. That work introduced the selective bottleneck recursion, moving orthonormal frames, and the Wiener-chaos organization, and showed how fixed-order estimates improve the dimension exponent. Our additions are uniform analytic frame jets, an all-order projection-weighted Wick-contraction theorem, and a logarithmic-order summation of the selective recursion. Here H_j is the jth bottleneck polynomial defined in (3.5), and $\rho_p \asymp p^{-3}$ is the analytic radius in (4.29). A direct pathwise bound gives $|H_4(t)| \lesssim d^{5/2}$, whereas at the base point we prove

$$\|H_4(0)\|_{L^2} \le p^{O(1)} d^2, \qquad \|H_j(0)\|_{L^2} \le p^2 (Cj)^{Cj} \rho_p^{-(j-1)} (d+j)^{j/2} \quad (j \ge 2). \tag{1.9}$$

The natural $d^{j/2}$ scale is used only over the logarithmic range of j required by the proof.

The raw-to-Wick expansion, chaos orthogonality and product formula for multiple Wiener integrals are classical [12, 32, 33]. Our new probabilistic input is a projection-weighted contraction estimate adapted to a Parseval frame. If $U^\mathsf{T} U = I_d$, $P = UU^\mathsf{T}$ and $W = \mathrm{Diag}(\mathrm{diag}\, P)$, then

$$P \preceq I_m, \qquad W^{-1/2} P^{\circ 2} W^{-1/2} \preceq I_m, \qquad \sum_i w_i = d.$$

The first two identities control the zero and one contractions. After r Gaussian traces, the double contraction has size $d^{r+1} d^{(n-2r-2)/2} = d^{n/2}$, avoiding the extra factor $d^{1/2}$. This is a standalone weighted Gaussian-polynomial statement under projection and relative-row hypotheses, not a theorem for arbitrary weights.

The proof follows one dependency chain. Proposition 2.4 reduces Theorem 1.3 to strong and lower-trace self-concordance, symmetry, and a dimension-free ASC modulus for the scaled metric. Propositions 2.5 and 2.6 verify the deterministic hypotheses, so the main task is Theorem 1.2. Lemma 3.1 decomposes the metric discrepancy into base-point terms, controlled path terms, and one terminal remainder. Theorem 4.8 supplies all-order derivatives of N_x; together with Theorem 5.1, it yields the base-point estimate in Theorem 6.2. Lemma 7.3 controls the remaining path terms, and Lemma 7.4 chooses a logarithmic truncation order that absorbs the terminal $\sqrt{d}$ loss. Lemma 8.1 and Proposition 2.4 then complete the proof of Theorem 1.3.

Notation.. Unless a subscript indicates otherwise, vector norms are Euclidean. We use $\|\cdot\|_{\mathrm{op}}$, $\|\cdot\|_{\mathrm{F}}$, and $\|\cdot\|_{\mathrm{HS}}$ for operator, Frobenius, and Hilbert–Schmidt norms, and $\|u\|_{g(x)}^2 = u^\mathsf{T} g(x) u$ for a local metric norm. The symbols d_{TV} and D_{KL} denote total variation distance and relative entropy. The identity on $\mathbb{R}^k$ is I_k; u^T is transpose; $\mathrm{Diag}(u)$ forms a diagonal matrix and $\mathrm{diag}(M)$ extracts its diagonal; and $M^{\circ 2} = M \circ M$ denotes the Hadamard square. For symmetric matrices, $B \preceq C$ means that $C - B$ is positive semidefinite. The derivative $D^k F(x)$ is viewed as a symmetric k-linear map. We write $X \lesssim Y$ for $X \le CY$ with a universal C, and $X \asymp Y$ when both comparisons hold. Matrix trace is tr; the iterated tensor trace Tr is defined in Section 5. Constants $C, c > 0$ are universal and may change from line to line. Finally, $\mathrm{polylog}(m, d)$ denotes a quantity bounded by $C_0(1 + \log(m + d + 1))^{C_1}$ for universal $C_0, C_1 > 0$, which may differ between occurrences.

2. Lee–Sidford geometry and the mixing reduction. This section fixes the Lee–Sidford metric, proves the abstract accepted-density overlap and conductance reduction, and then verifies the deterministic Lee–Sidford hypotheses. The remaining random input, quantitative ASC, is proved in Sections 3–7.

2.1. *The metric and Lewis calculus.* Let

$$K = \{x \in \mathbb{R}^d : Ax \ge b\}$$

be bounded and full-dimensional. For $x \in \mathrm{int}\, K$, define

$$S_x = \mathrm{Diag}(Ax - b), \qquad A_x = S_x^{-1} A.$$

Fix $p \geq 4$ and put

$$\beta = \frac{1}{2} - \frac{1}{p}, \qquad \alpha = 2\beta = 1 - \frac{2}{p}, \qquad c_p = 1 - \frac{2}{p}.$$

The ℓ_p Lewis weights $w_x \in \mathbb{R}^m_{++}$ are the unique positive vector satisfying [26, Section 5]

$$w_x = \operatorname{diag}\left(W_x^{\beta} A_x (A_x^{\mathsf{T}} W_x^{2\beta} A_x)^{-1} A_x^{\mathsf{T}} W_x^{\beta}\right), \qquad W_x = \operatorname{Diag}(w_x). \tag{2.1}$$

The baseline Lee–Sidford metric is

$$g_0(x) = A_x^{\mathsf{T}} W_x^{2\beta} A_x. \tag{2.2}$$

Set

$$P_x = W_x^{\beta} A_x g_0(x)^{-1} A_x^{\mathsf{T}} W_x^{\beta}, \qquad \Lambda_x = W_x - P_x^{\circ 2}, \tag{2.3}$$

where $P^{\circ 2} = P \circ P$, and define

$$\bar{\Lambda}_x = W_x^{-1/2} \Lambda_x W_x^{-1/2}, \qquad N_x = 2\bar{\Lambda}_x (I - c_p \bar{\Lambda}_x)^{-1}. \tag{2.4}$$

The standard Lewis identities [15, Section 2.2] give

$$P_x^{\circ 2} \preceq W_x, \qquad 0 \preceq \bar{\Lambda}_x \preceq I, \qquad 0 \preceq N_x \preceq pI, \qquad \sum_{i=1}^{m} w_{x,i} = d. \tag{2.5}$$

LEMMA 2.1 (Analytic Lewis weights in logarithmic coordinates). *The map $x \mapsto \ell_x := \log w_x$ is real analytic on* $\operatorname{int} K$. *More precisely, for $\ell \in \mathbb{R}^m$ set*

$$W(\ell) = \operatorname{Diag}(e^{\ell_1}, \dots, e^{\ell_m}), \qquad B(x,\ell) = W(\ell)^{\beta} A_x$$

and

$$P(x,\ell) = B(x,\ell)\big(B(x,\ell)^{\mathsf{T}} B(x,\ell)\big)^{-1} B(x,\ell)^{\mathsf{T}}.$$

In a neighborhood of (x, ℓ_x), define

$$\mathcal{L}(x,\ell) := \ell - \log \operatorname{diag} P(x,\ell). \tag{2.6}$$

Then $\mathcal{L}(x,\ell_x) = 0$, and, at a solution,

$$\begin{aligned} D_\ell \mathcal{L}(x,\ell_x) &= I - c_p W_x^{-1} \Lambda_x \\ &= W_x^{-1/2}\big(I - c_p \bar{\Lambda}_x\big) W_x^{1/2}. \end{aligned} \tag{2.7}$$

In particular, this Jacobian is invertible, with

$$\big(D_\ell \mathcal{L}(x,\ell_x)\big)^{-1} = W_x^{-1/2}\big(I - c_p \bar{\Lambda}_x\big)^{-1} W_x^{1/2}. \tag{2.8}$$

PROOF. Every row of A_x is nonzero, so every diagonal entry of $P(x,\ell)$ is positive. Thus (2.6) is real analytic wherever $x \in \operatorname{int} K$ and $\ell \in \mathbb{R}^m$. We compute its weight Jacobian. If a full-column-rank matrix B varies according to $\dot{B} = \operatorname{Diag}(z) B$, then its orthogonal projection $P = B(B^{\mathsf{T}} B)^{-1} B^{\mathsf{T}}$ satisfies

$$\dot{P} = (I - P)\operatorname{Diag}(z) P + P \operatorname{Diag}(z)(I - P). \tag{2.9}$$

Taking diagonal entries gives

$$\operatorname{diag} \dot{P} = 2\big(W - P^{\circ 2}\big) z = 2\Lambda z, \qquad W = \operatorname{Diag}(\operatorname{diag} P). \tag{2.10}$$

For a variation $\dot{\ell} = \zeta$ at a Lewis solution, $\dot{B} = \beta\,\mathrm{Diag}(\zeta)B$. Hence

$$D_\ell \log \mathrm{diag}\, P(x,\ell_x)[\zeta] = 2\beta W_x^{-1}\Lambda_x\zeta = c_p W_x^{-1}\Lambda_x\zeta,$$

which proves (2.7). Since $0 \preceq \bar{\Lambda}_x \preceq I$ and $1 - c_p = 2/p > 0$, $I - c_p\bar{\Lambda}_x$ is positive definite. This proves (2.8).

The real analytic implicit-function theorem [21] now produces a unique analytic solution branch near every $x \in \mathrm{int}\, K$. The positive Lewis weights are unique, so any two local branches agree on their overlap. These branches therefore patch to one real analytic map on $\mathrm{int}\, K$. □

For $r \in \mathbb{R}^d$, write $s_{x,r} = A_x r$ and $v_{x,r} = W_x^{1/2} s_{x,r}$. Differentiating (2.6) also recovers the Lewis derivative identity. Indeed, at fixed ℓ one has $D_x B[r] = -\,\mathrm{Diag}(s_{x,r})B$, so (2.10) gives

$$\big(I - c_p W_x^{-1}\Lambda_x\big) D\ell_x[r] = -2W_x^{-1}\Lambda_x s_{x,r}.$$

Using (2.8), the commutativity of $\bar{\Lambda}_x$ and $(I - c_p\bar{\Lambda}_x)^{-1}$, and (2.4), we obtain

$$D \log w_x[r] = -W_x^{-1/2} N_x v_{x,r},$$

$$DW_x[r] = -\,\mathrm{Diag}\Big(W_x^{1/2} N_x v_{x,r}\Big). \tag{2.11}$$

Existence, uniqueness, strict positivity and the first-derivative identity are part of the Lewis-weight calculus in [26, Section 5, in particular Lemma 24]; the notation and the bounds in (2.5) are recorded explicitly in [15, Section 2.2].

2.2. *Abstract local overlap and mixing reduction.* For reference, a C^2 metric g is strongly self-concordant (SSC) if, for every $x \in \mathrm{int}\, K$ and $h \in \mathbb{R}^d$,

$$\Big\| g(x)^{-1/2} Dg(x)[h] g(x)^{-1/2} \Big\|_{\mathrm{F}} \le 2\,\|h\|_{g(x)},$$

and lower-trace self-concordant (LTSC) if, for the same x and h,

$$\mathrm{tr}\big(g(x)^{-1} D^2 g(x)[h,h]\big) \ge -\,\|h\|_{g(x)}^2.$$

The SSC inequality also implies the usual matrix self-concordance inequality because the operator norm is bounded by the Frobenius norm. It is $\bar{\nu}$-symmetric when

$$D_g(x,1) \subseteq K \cap (2x - K) \subseteq D_g(x, \sqrt{\bar{\nu}})$$

for every $x \in \mathrm{int}\, K$, where $D_g(x,r) = \{y : \|y - x\|_{g(x)} \le r\}$. These sampling regularity notions follow [16, 22].

The next lemma is the point at which linearity of the target replaces the function-counterpart assumption in the general Dikin framework. It proves overlap through one common accepted mass and records every interior and reverse-proposal event explicitly.

LEMMA 2.2 (One-step overlap for a linear exponential target). *There are universal constants $\varepsilon_{\mathrm{ov}} \in (0, 10^{-2})$ and $c_{\mathrm{ov}}, c_{\mathrm{near}} > 0$ with the following property. Let g be strongly self-concordant and lower-trace self-concordant, and suppose that $R_{\varepsilon_{\mathrm{ov}}}(g) \ge a$ for some $a > 0$. Put*

$$0 < \varrho \le c_{\mathrm{ov}}(1 \wedge a).$$

For every $\vartheta \in \mathbb{R}^d$ and $x \in \mathrm{int}\, K$, let P_x be the one-step non-lazy Metropolis Dikin law started at x, with proposal $N(x, \varrho^2 g(x)^{-1}/d)$ and target $\pi_\vartheta \propto e^{-\vartheta^\intercal x}\mathbf{1}_K(x)$. Then, for all $x, y \in \mathrm{int}\, K$,

$$\|x - y\|_{g(x)} \le \frac{c_{\mathrm{near}}\varrho}{\sqrt{d}} \quad \Longrightarrow \quad d_{\mathrm{TV}}(\mathsf{P}_x, \mathsf{P}_y) \le 0.8. \tag{2.12}$$

The constants are independent of ϑ, d, and the instance. No function whose Hessian is comparable to g is assumed.

PROOF. We prove the overlap directly through the common accepted density. Fix

$$\varepsilon_{\mathrm{ov}} = 2^{-10}, \qquad \eta_0 = 2^{-8}, \qquad t_0 = \frac{1}{32}.$$

Take $B = 16$. Since $\mathbb{E}\,\|\xi\|^2 = d$, Markov's inequality gives, for every $d \geq 1$ and $\xi \sim N(0, I_d)$,

$$\mathbb{P}\{\|\xi\| > B\sqrt{d}\} \leq \eta_0. \tag{2.13}$$

We choose c_{ov} and c_{near} below; all restrictions on them are universal.

For $s \in \{x, y\}$, let q_s be the density of

$$N\left(s, \frac{\varrho^2}{d} g(s)^{-1}\right), \qquad u_s = z - s,$$

and set $\phi = \frac{1}{2}\log\det g$. For $z \in \operatorname{int} K$, the logarithm of the Metropolis ratio from s is

$$\begin{aligned} \Lambda_s(z) &:= \log \frac{\pi_\vartheta(z) q_z(s)}{\pi_\vartheta(s) q_s(z)} \\ &= -\vartheta^{\mathsf{T}} u_s + \phi(z) - \phi(s) - \frac{d}{2\varrho^2}\left(\|u_s\|^2_{g(z)} - \|u_s\|^2_{g(s)}\right). \end{aligned} \tag{2.14}$$

We first construct a good event for each starting point. The restrictions below impose $c_{\mathrm{ov}} < 1$, so $\varrho < a$. The set of admissible ASC radii is downward closed; hence Definition 1.1 applies at radius ϱ and gives an event

$$\mathcal{A}_s := \left\{ z \in \operatorname{int} K : \left| \|u_s\|^2_{g(z)} - \|u_s\|^2_{g(s)} \right| \leq \frac{2\varepsilon_{\mathrm{ov}}\varrho^2}{d} \right\}$$

such that

$$q_s(\mathcal{A}_s^c) \leq \varepsilon_{\mathrm{ov}}. \tag{2.15}$$

On $\mathcal{A}_s$, the final term in (2.14) is at least $-\varepsilon_{\mathrm{ov}}$.

We next control the determinant term. Matrix differentiation gives

$$D\phi(w)[v] = \frac{1}{2}\operatorname{tr}\left(g(w)^{-1} Dg(w)[v]\right), \tag{2.16}$$

$$\begin{aligned} D^2\phi(w)[v,v] &= \frac{1}{2}\operatorname{tr}\left(g(w)^{-1} D^2 g(w)[v,v]\right) \\ &\quad - \frac{1}{2}\left\| g(w)^{-1/2} Dg(w)[v] g(w)^{-1/2} \right\|_{\mathrm{F}}^2. \end{aligned} \tag{2.17}$$

LTSC and SSC imply

$$D^2\phi(w)[v,v] \geq -\frac{5}{2}\|v\|^2_{g(w)}, \qquad \|D\phi(w)\|_{g(w)^{-1}} \leq \sqrt{d}. \tag{2.18}$$

Indeed, the first term in (2.17) is at least $-\frac{1}{2}\|v\|^2_{g(w)}$, while SSC bounds the second one below by $-2\|v\|^2_{g(w)}$. The second inequality follows from $|\operatorname{tr} M| \leq \sqrt{d}\,\|M\|_{\mathrm{F}}$.

Define

$$\mathcal{N}_s := \{\|u_s\|_{g(s)} \leq B\varrho\}, \qquad \mathcal{D}_s := \{D\phi(s)[u_s] \geq -t_0\}.$$

Writing $u_s = (\varrho/\sqrt{d})g(s)^{-1/2}\xi$ and using (2.13) gives

$$q_s(\mathcal{N}_s^c) \le \eta_0. \tag{2.19}$$

Moreover, $D\phi(s)[u_s]$ is centered Gaussian and, by (2.18), has variance at most ϱ^2. We may therefore choose c_{ov} sufficiently small that

$$\sup_{0<\varrho\le c_{\mathrm{ov}}} q_s(\mathcal{D}_s^c) \le \eta_0. \tag{2.20}$$

On $\mathcal{A}_s \cap \mathcal{N}_s$, both endpoints of $[s,z]$ lie in $\operatorname{int} K$, and convexity places the whole segment in the interior. Put $\delta_s = \|u_s\|_{g(s)}$. The matrix self-concordance consequence needed here follows directly from SSC. The case $u_s = 0$ is immediate. Otherwise, if $r_s(t) = \|u_s\|_{g(s+tu_s)}$, then differentiation and the SSC bound give

$$\left|(r_s^2)'(t)\right| \le 2r_s(t)^3, \qquad \text{hence} \qquad \left|r_s'(t)\right| \le r_s(t)^2.$$

Since $r_s(0) = \delta_s$, comparison with the scalar differential equation gives, for $0 \le t \le 1$,

$$\|u_s\|_{g(s+tu_s)} \le \frac{\delta_s}{1-t\delta_s}. \tag{2.21}$$

Impose

$$Bc_{\mathrm{ov}} \le \frac{1}{2}, \qquad 5B^2c_{\mathrm{ov}}^2 \le \frac{1}{32}.$$

Taylor's formula, (2.18), and (2.21) yield

$$\begin{aligned} \phi(z) - \phi(s) - D\phi(s)[u_s] &= \int_0^1 (1-t)D^2\phi(s+tu_s)[u_s,u_s]\,\mathrm{d}t \\ &\ge -\frac{5}{2}\delta_s^2 \int_0^1 \frac{1-t}{(1-t\delta_s)^2}\,\mathrm{d}t \ge -5\delta_s^2. \end{aligned} \tag{2.22}$$

Consequently, on $\mathcal{G}_s := \mathcal{A}_s \cap \mathcal{N}_s \cap \mathcal{D}_s$,

$$\phi(z) - \phi(s) \ge -\frac{1}{16}, \qquad q_s(\mathcal{G}_s^c) \le \varepsilon_{\mathrm{ov}} + 2\eta_0 < \frac{1}{64}. \tag{2.23}$$

Thus the two proposal-dependent terms in (2.14) are at least $-1/8$ on $\mathcal{G}_s$.

We also record the full comparison of the two forward Gaussian densities. Set

$$\delta = \|x-y\|_{g(x)} \le \frac{c_{\mathrm{near}}\varrho}{\sqrt{d}}.$$

After decreasing c_{near}, we have $\delta \le 1/4$. Integrating the SSC inequality along $[x,y]$ gives

$$(1-\delta)^2 g(x) \preceq g(y) \preceq (1-\delta)^{-2} g(x). \tag{2.24}$$

To see this without invoking a function counterpart, put $h = y - x$ and $\gamma(t) = x + th$. The preceding scalar comparison gives $\|h\|_{g(\gamma(t))} \le \delta/(1-t\delta)$. For every fixed $v \ne 0$, SSC therefore implies

$$\left|\frac{\mathrm{d}}{\mathrm{d}t}\log\bigl(v^{\mathsf{T}} g(\gamma(t))v\bigr)\right| \le \frac{2\delta}{1-t\delta}.$$

Integration from 0 to 1 proves (2.24). Let $M = g(x)^{-1/2}g(y)g(x)^{-1/2}$. The Gaussian relative entropy formula [5] is

$$D_{\mathrm{KL}}(q_x\|q_y) = \frac{1}{2}\left\{\operatorname{tr} M - d - \log\det M + \frac{d}{\varrho^2}\|x-y\|_{g(y)}^2\right\}. \tag{2.25}$$

Every eigenvalue λ of M lies in $[(1-\delta)^2,(1-\delta)^{-2}]$, so $\lambda-1-\log\lambda\le C\delta^2$. Hence

$$D_{\mathrm{KL}}(q_x\|q_y)\le C\left(d\delta^2+\frac{d\delta^2}{\varrho^2}\right)\le Cc_{\mathrm{near}}^2.$$

Pinsker's inequality [35, Lemma 2.5] gives

$$d_{\mathrm{TV}}(q_x,q_y)\le Cc_{\mathrm{near}}. \tag{2.26}$$

Choose c_{near} so that the right side is at most $1/64$. This calculation controls both the covariance and displaced-mean quadratic terms.

It remains to pass from proposal overlap to accepted-density overlap. For $s\in\{x,y\}$, define the accepted density without evaluating $g(z)$ outside its domain by

$$\begin{aligned} k_s(z) &:= \begin{cases} q_s(z)\big(1\wedge e^{\Lambda_s(z)}\big), & z\in\operatorname{int}K,\\ 0, & z\notin\operatorname{int}K,\end{cases}\\ &= \begin{cases} \min\left\{q_s(z), e^{-\vartheta^{\mathsf T}(z-s)}q_z(s)\right\}, & z\in\operatorname{int}K,\\ 0, & z\notin\operatorname{int}K.\end{cases} \end{aligned} \tag{2.27}$$

Thus

$$\mathsf{P}_s=\left(1-\int_{\mathbb{R}^d}k_s(z)\,\mathrm{d}z\right)\delta_{\{s\}}+k_s(z)\,\mathrm{d}z. \tag{2.28}$$

Here $\delta_{\{s\}}$ is the unit point mass at s.

Choose $s_\star\in\{x,y\}$ so that

$$\vartheta^{\mathsf T}s_\star=\min\{\vartheta^{\mathsf T}x,\vartheta^{\mathsf T}y\},$$

and define the common halfspace

$$H:=\{z:\vartheta^{\mathsf T}z\le\vartheta^{\mathsf T}s_\star\}.$$

For every $z\in H$ and $s\in\{x,y\}$, $-\vartheta^{\mathsf T}(z-s)\ge 0$. If $\vartheta\ne 0$, symmetry of the centered Gaussian proposal gives $q_{s_\star}(H)=1/2$; if $\vartheta=0$, then $H=\mathbb{R}^d$. Thus

$$q_{s_\star}(H)\ge\frac{1}{2}. \tag{2.29}$$

On $H\cap\mathcal{G}_s$, all quantities in (2.14) are defined, $\Lambda_s(z)\ge -1/8$, and therefore

$$k_s(z)\ge e^{-1/8}q_s(z). \tag{2.30}$$

Let $m(z)=\min\{q_x(z),q_y(z)\}$. Since $m\le q_s$,

$$\begin{aligned}\int_{\mathbb{R}^d}\min\{k_x(z),k_y(z)\}\,\mathrm{d}z &\ge e^{-1/8}\int_{H\cap\mathcal{G}_x\cap\mathcal{G}_y}m(z)\,\mathrm{d}z\\ &\ge e^{-1/8}\left(\int_H m(z)\,\mathrm{d}z-q_x(\mathcal{G}_x^c)-q_y(\mathcal{G}_y^c)\right).\end{aligned} \tag{2.31}$$

Moreover,

$$\begin{aligned}\int_H m(z)\,\mathrm{d}z &\ge q_{s_\star}(H)-\int_{\mathbb{R}^d}\big(q_{s_\star}(z)-m(z)\big)\,\mathrm{d}z\\ &= q_{s_\star}(H)-d_{\mathrm{TV}}(q_x,q_y)\ge\frac{1}{2}-\frac{1}{64}.\end{aligned} \tag{2.32}$$

Equations (2.23), (2.31), and (2.32) give

$$\int_{\mathbb{R}^d} \min\{k_x(z), k_y(z)\}\,\mathrm{d}z \ge e^{-1/8}\left(\frac{1}{2} - \frac{3}{64}\right) > \frac{1}{5}.$$

If $x = y$, there is nothing to prove. If $x \ne y$, the rejection atoms in (2.28) are supported at distinct points, while the accepted parts are absolutely continuous. Therefore their common mass is exactly the integral above, and

$$d_{\mathrm{TV}}(\mathsf{P}_x, \mathsf{P}_y) = 1 - \int_{\mathbb{R}^d} \min\{k_x(z), k_y(z)\}\,\mathrm{d}z < \frac{4}{5}.$$

This proves (2.12). The construction used only SSC, LTSC, and the ASC event, including its interior condition; no function counterpart is needed. □

Having obtained common accepted mass, it remains only to convert this local overlap into global conductance through the symmetry geometry.

LEMMA 2.3 (Symmetry, cross-ratio separation, and conductance). *Let π be a log-concave probability distribution on a bounded full-dimensional convex body K, and let g be $\bar{\nu}$-symmetric. Suppose a reversible Markov kernel K with stationary law π satisfies, for some $\Delta \in (0,1)$ and all $x, y \in \operatorname{int} K$,*

$$\|x-y\|_{g(x)} \le \Delta \quad \Longrightarrow \quad d_{\mathrm{TV}}(\mathsf{K}_x, \mathsf{K}_y) \le 0.9. \tag{2.33}$$

Then its global stationary conductance satisfies

$$\Phi(\mathsf{K}) \ge \frac{c\Delta}{\sqrt{\bar{\nu}}} \tag{2.34}$$

for a universal constant $c > 0$.

PROOF. Symmetry yields the required metric isoperimetry. For distinct $x, y \in \operatorname{int} K$, let the line through them meet ∂K at p, q, ordered as p, x, y, q, and let $d_K(x,y)$ denote the cross-ratio distance

$$d_K(x,y) = \frac{\|x-y\|_2\,\|p-q\|_2}{\|p-x\|_2\,\|y-q\|_2}.$$

Let u be the unit vector along this line and let t_x be the smaller of the two Euclidean distances from x to the endpoints. Both $x + t_x u$ and $x - t_x u$ lie in K, so $t_x u \in K \cap (2x - K) - x$. The outer symmetry inclusion gives $t_x \|u\|_{g(x)} \le \sqrt{\bar{\nu}}$. Directly from the displayed cross-ratio formula,

$$d_K(x,y) \ge \frac{\|x-y\|_2}{t_x} \ge \frac{\|x-y\|_{g(x)}}{\sqrt{\bar{\nu}}}. \tag{2.35}$$

The first inequality holds whichever of the two endpoint distances defines t_x.

The cross-ratio isoperimetric inequality for log-concave measures [30, Theorem 2.5], in the form used in [16, proof of Lemma B.6], says that if $K = S_1 \sqcup S_2 \sqcup S_3$, then

$$\pi(S_3) \ge c\, d_K(S_1, S_2)\pi(S_1)\pi(S_2). \tag{2.36}$$

Equations (2.35) and (2.36) give the required global isoperimetry in the local norm of g.

For the conductance partition, in the standard local-overlap form of [16, 29], fix S with $0 < \pi(S) \le 1/2$ and choose a fixed $\alpha < 0.05$. Put

$$\begin{aligned} Q(S,S^c) &= \int_S \mathsf{K}(x,S^c)\pi(\mathrm{d}x), \\ S_1 &= \{x \in S : \mathsf{K}(x,S^c) < \alpha\}, \\ S_2 &= \{y \in S^c : \mathsf{K}(y,S) < \alpha\}. \end{aligned}$$

For $x \in S_1$ and $y \in S_2$, $d_{\mathrm{TV}}(\mathsf{K}_x, \mathsf{K}_y) > 1 - 2\alpha > 0.9$. Hence (2.33) implies $\|x-y\|_{g(x)} > \Delta$, and (2.35) separates S_1, S_2 by at least $\Delta/\sqrt{\bar\nu}$ in cross ratio. If either S_1 or S_2 has lost a fixed fraction of its side, the definitions and reversibility already give $Q(S,S^c) \ge c\pi(S)$. Otherwise both retain a fixed fraction, and (2.36) gives

$$\pi\big(K \setminus (S_1 \cup S_2)\big) \ge \frac{c\Delta}{\sqrt{\bar\nu}}\pi(S).$$

The complement is contained in $(S \setminus S_1) \cup (S^c \setminus S_2)$, whose stationary mass is at most $2Q(S,S^c)/\alpha$. Thus $Q(S,S^c) \ge c\Delta\pi(S)/\sqrt{\bar\nu}$. Taking the infimum over S proves (2.34). This is the local-overlap partition of [16, Lemma B.2], written here with its cross-ratio separation explicit. $\square$

PROPOSITION 2.4 (Chi-square corollary of the Dikin mixing framework). *Let $K \subset \mathbb{R}^d$ be bounded, convex and full-dimensional, and let $g : \operatorname{int} K \to \mathbb{S}^d_{++}$ be a C^2 metric that is strongly self-concordant, lower-trace self-concordant and $\bar\nu$-symmetric. Suppose that $R_\varepsilon(g) \ge \mathfrak{a}(\varepsilon) > 0$ for every $\varepsilon \in (0,1)$, where $\mathfrak{a}$ is independent of the dimension and of the instance. There are universal constants $\varepsilon_{\mathrm{mix}} \in (0, 10^{-2})$ and $c_0 \in (0,1)$ and $c_1, c_2 > 0$ with the following property. If*

$$a_0 := 1 \wedge \mathfrak{a}(\varepsilon_{\mathrm{mix}}), \qquad \varrho_{\mathrm{KV}} := c_0 a_0, \tag{2.37}$$

then, for every $\vartheta \in \mathbb{R}^d$, the $1/2$-lazy Metropolis Dikin kernel K with proposal radius ϱ_{KV} is reversible with stationary law π_ϑ, has global conductance

$$\Phi(\mathsf{K}) := \inf_{0 < \pi_\vartheta(S) \le 1/2} \frac{\int_S \mathsf{K}(x,S^c)\pi_\vartheta(\mathrm{d}x)}{\pi_\vartheta(S)} \ge \frac{c_1 a_0}{\sqrt{d\bar\nu}}, \tag{2.38}$$

and has $L^2(\pi_\vartheta)$ spectral gap at least $c_2 a_0^2/(d\bar\nu)$. Consequently, for every $\mu_0 \ll \pi_\vartheta$ with finite chi-square divergence and every integer $T \ge 0$,

$$\chi^2(\mu_0 \mathsf{K}^T \| \pi_\vartheta) \le \exp\left(-\frac{c_2 a_0^2 T}{d\bar\nu}\right) \chi^2(\mu_0 \| \pi_\vartheta). \tag{2.39}$$

Consequently, for every $\delta \in (0,1)$, the left side is at most δ whenever

$$T \ge C a_0^{-2} d\bar\nu \log\left(\frac{1 + \chi^2(\mu_0\|\pi_\vartheta)}{\delta}\right). \tag{2.40}$$

PROOF. We derive the chi-square statement from the total-variation framework of [16]. Take $\varepsilon_{\mathrm{mix}} = \varepsilon_{\mathrm{ov}}$ from Lemma 2.2. Since

$$\varrho_{\mathrm{KV}} = c_0 a_0, \qquad a_0 \le 1 \wedge R_{\varepsilon_{\mathrm{mix}}}(g),$$

choosing $c_0 \le c_{\mathrm{ov}}$ lets that lemma apply directly. Thus the non-lazy kernels satisfy (2.12), uniformly in ϑ. For the lazy kernels $\mathsf{K}_x = (\delta_x + \mathsf{P}_x)/2$, the triangle inequality gives

$$d_{\mathrm{TV}}(\mathsf{K}_x, \mathsf{K}_y) \le \frac{1}{2} + \frac{1}{2}(0.8) = 0.9.$$

Lemma 2.3, with $\Delta = c_{\mathrm{near}}\varrho_{\mathrm{KV}}/\sqrt{d}$, now gives

$$\Phi(\mathsf{K}) \geq \frac{c\varrho_{\mathrm{KV}}}{\sqrt{d\bar{\nu}}} = \frac{ca_0}{\sqrt{d\bar{\nu}}}.$$

This is the global stationary conductance in (2.38), not a warm or restricted conductance, and proves that display after adjusting c_1.

From conductance to chi-square contraction. Metropolis detailed balance makes K self-adjoint on $L^2(\pi_\vartheta)$. Since $\mathsf{K} = (I+\mathsf{P})/2$, its spectrum lies in $[0,1]$. Cheeger's inequality for reversible Markov kernels [23, Theorem 2.1] therefore yields

$$\mathrm{gap}(\mathsf{K}) \geq \frac{\Phi(\mathsf{K})^2}{2} \geq \frac{ca_0^2}{d\bar{\nu}}.$$

If $f_0 = \mathrm{d}\mu_0/\mathrm{d}\pi_\vartheta - 1$, detailed balance gives

$$\frac{\mathrm{d}(\mu_0\mathsf{K}^T)}{\mathrm{d}\pi_\vartheta} - 1 = \mathsf{K}^T f_0.$$

The function f_0 has π_ϑ-mean zero, and hence

$$\begin{aligned}\chi^2(\mu_0\mathsf{K}^T\|\pi_\vartheta) &= \left\|\mathsf{K}^T f_0\right\|^2_{L^2(\pi_\vartheta)}\\ &\leq \left(1 - \frac{ca_0^2}{d\bar{\nu}}\right)^{2T} \|f_0\|^2_{L^2(\pi_\vartheta)}\\ &\leq \exp\left(-\frac{c_2 a_0^2 T}{d\bar{\nu}}\right)\chi^2(\mu_0\|\pi_\vartheta).\end{aligned}$$

No warmness bound is used: finite chi-square divergence is needed only for $f_0 \in L^2(\pi_\vartheta)$. Solving the last display for T gives (2.40); replacing the numerator $\chi^2(\mu_0\|\pi_\vartheta)$ by $1 + \chi^2(\mu_0\|\pi_\vartheta)$ also covers the stationary initialization without a separate case. The same chi-square formulation is recorded in [15, Lemma 2.2]. $\square$

The mixing argument is now reduced to deterministic Lee–Sidford regularity and the quantitative ASC estimate of Theorem 1.2.

2.3. *Deterministic Lee–Sidford geometry and scaling.*

PROPOSITION 2.5 (Established Lee–Sidford regularity). *Let $p \geq \max\{4, 2(1+\log m)\}$ and set*

$$\lambda_p := (1+p^2)(1+p), \qquad H_p := \lambda_p g_0. \tag{2.41}$$

Then H_p is strongly self-concordant and lower-trace self-concordant, and it is $\bar{\nu}_p$-symmetric with

$$\bar{\nu}_p \leq e\lambda_p d. \tag{2.42}$$

PROOF. Under the dictionary $q = p$ and $A_x^{\mathsf{T}} W_x^{1-2/q} A_x = g_0(x)$, (2.41) is exactly the LS matrix of [22, Definition 6]. Strong self-concordance and convexity of $\log\det H_p$ are [22, Lemmas 4.2 and 4.3]. The latter gives

$$\mathrm{tr}(H_p^{-1}D^2H_p[h,h]) \geq \mathrm{tr}(H_p^{-1}DH_p[h]H_p^{-1}DH_p[h]) \geq 0,$$

hence lower-trace self-concordance. The proof of [22, Lemma 4.3] also gives

$$\|y-x\|^2_{H_p(x)} \leq \lambda_p \sum_i w_{x,i}^{1-2/p} \leq \lambda_p d^{1-2/p} m^{2/p} \leq e\lambda_p d$$

for $y \in K \cap (2x - K)$; here $p \geq 2(1 + \log m)$ implies $m^{2/p} \leq e$. The reverse inclusion follows directly from the leverage identity

$$a_{x,i}^{\mathsf{T}} g_0(x)^{-1} a_{x,i} = w_{x,i}^{2/p}.$$

Here $a_{x,i}^{\mathsf{T}}$ denotes the ith row of A_x. Since $w_{x,i} = (P_x)_{ii} \leq 1$, if $\|y - x\|_{H_p(x)} \leq 1$, Cauchy–Schwarz gives

$$\left| a_{x,i}^{\mathsf{T}}(y - x) \right| \leq \lambda_p^{-1/2} w_{x,i}^{1/p} \leq 1$$

for every i, and hence $y \in K \cap (2x - K)$. This proves (2.42). See also [15, Section 2.2] for this normalization. □

PROPOSITION 2.6 (Metric scaling). *For every $L > 0$,*

$$R_\varepsilon(Lg) = L^{1/2} R_\varepsilon(g). \tag{2.43}$$

If $L \geq 1$ and g is $\bar{\nu}$-symmetric, then Lg is $L\bar{\nu}$-symmetric. Scaling by $L \geq 1$ also preserves strong self-concordance and lower-trace self-concordance.

PROOF. A radius-r proposal for Lg is a radius-$r/L^{1/2}$ proposal for g. Its discrepancy is multiplied by L, while

$$L \frac{2\varepsilon (r/L^{1/2})^2}{d} = \frac{2\varepsilon r^2}{d},$$

which proves (2.43). The symmetry statement follows from $\|u\|_{Lg(x)}^2 = L \|u\|_{g(x)}^2$. In the defining differential inequalities for strong self-concordance, the left side is unchanged while the local norm on the right is multiplied by $L^{1/2}$. For lower-trace self-concordance, the trace term is unchanged while the negative lower bound is multiplied by L; hence the condition is preserved when $L \geq 1$. □

2.4. *Affine covariance.*

LEMMA 2.7 (Affine covariance). *Let $\Psi(\widetilde{x}) = x_c + T\widetilde{x}$, where $T \in \mathbb{R}^{d \times d}$ is invertible, and pull the polytope and metric back through Ψ. Then the Lewis weights and the matrices P and N are unchanged at corresponding points, while*

$$\widetilde{g}_0(\widetilde{x}) = T^{\mathsf{T}} g_0(\Psi(\widetilde{x})) T. \tag{2.44}$$

Moreover,

$$D^k \widetilde{N}_{\widetilde{x}}[r_1, \ldots, r_k] = D^k N_{\Psi(\widetilde{x})}[Tr_1, \ldots, Tr_k], \tag{2.45}$$

$$\|r\|_{\widetilde{g}_0(\widetilde{x})} = \|Tr\|_{g_0(\Psi(\widetilde{x}))}. \tag{2.46}$$

The Gaussian proposal and the discrepancy in (1.4) are carried into their counterparts by Ψ.

PROOF. The pulled-back inequality matrix is $\widetilde{A} = AT$, and the slack vectors at $\widetilde{x}$ and $\Psi(\widetilde{x})$ coincide. Hence

$$\widetilde{A}_{\widetilde{x}} = A_{\Psi(\widetilde{x})} T.$$

Substitution in (2.1) cancels T and T^{T} inside the inverse, so the Lewis fixed point, and therefore P and N, is unchanged. The same substitution in (2.2) proves (2.44). The chain rule gives (2.45), and (2.46) follows by congruence. Finally,

$$T \widetilde{g}_0(\widetilde{x})^{-1} T^{\mathsf{T}} = g_0(\Psi(\widetilde{x}))^{-1},$$

so corresponding Gaussian proposals have the same law after applying Ψ. The two squared local norms in (1.4) are preserved by (2.44). □

The covariance lemma allows every ASC estimate to be proved at an arbitrary base point after imposing $x = 0$ and $g_0(x) = I_d$.

3. Selective expansion of the metric discrepancy. The normalization, path variables, and first bottleneck decomposition in this section follow [15, Sections 1.2 and 3.1]. We retain that notation to make the comparison transparent. Kook iterates the recursion to fixed order. Lemma 3.1 records its exact arbitrary-order form, which will later be used with a logarithmic truncation order.

By Lemma 2.7, the metric transforms covariantly, while the ASC discrepancy and radius are affine invariant. Fix a base point x and apply an affine change of coordinates so that

$$x = 0, \qquad g_0(x) = I_d. \tag{3.1}$$

Let $h \sim N(0, I_d)$, set $\eta = r/\sqrt{d}$, and define

$$F(t) = h^\top g_0(x + th)h, \qquad 0 \le t \le \eta. \tag{3.2}$$

For $x_t = x + th$, abbreviate $A_t = A_{x_t}$, $W_t = W_{x_t}$ and $N_t = N_{x_t}$, and set

$$s_t = A_t h, \quad u_t = W_t^\beta s_t, \quad v_t = W_t^{1/2} s_t, \quad \zeta_t = W_t^{-1/p} s_t, \quad q_t = W_t^{-1/2} u_t^{\circ 2}. \tag{3.3}$$

Then $F(t) = \|u_t\|^2$. Since $s_t' = -s_t^{\circ 2}$ and (2.11) holds, if $\ell_t = \log w_t$, then

$$F'(t) = -2E_1(t) - 2\beta H_1(t), \qquad E_1(t) = \langle s_t, W_t^\alpha s_t^{\circ 2}\rangle, \qquad H_1(t) = q_t^\top N_t v_t. \tag{3.4}$$

For $j \ge 1$, define

$$H_j(t) = q_t^\top N_t^{(j-1)} v_t, \tag{3.5}$$

where $N_t^{(k)} = D^k N_{x_t}[h, \dots, h]$. Differentiation gives

$$H_j'(t) = G_{j+1}(t) + H_{j+1}(t), \quad G_{j+1}(t) = (q_t')^\top N_t^{(j-1)} v_t + q_t^\top N_t^{(j-1)} v_t'. \tag{3.6}$$

LEMMA 3.1 (Selective expansion). *For every integer $J \ge 1$,*

$$\begin{aligned}\int_0^\eta H_1(t)\,\mathrm{d}t &= \sum_{j=1}^J \frac{\eta^j}{j!} H_j(0) + \sum_{j=2}^J \int_0^\eta \frac{(\eta - t)^{j-1}}{(j-1)!} G_j(t)\,\mathrm{d}t \\ &\quad + \int_0^\eta \frac{(\eta - t)^J}{J!}\big(G_{J+1}(t) + H_{J+1}(t)\big)\,\mathrm{d}t.\end{aligned} \tag{3.7}$$

PROOF. Iterate

$$H_j(t) = H_j(0) + \int_0^t \big(G_{j+1}(s) + H_{j+1}(s)\big)\,\mathrm{d}s$$

and apply Fubini's theorem. After j integrations the kernel is $(\eta - t)^{j-1}/(j-1)!$. □

The expansion separates the proof into three tasks. Analytic frame estimates control the derivatives of N_x; the weighted Wick theorem controls the base-point polynomials $H_j(0)$; and pathwise localization controls the $G_j(t)$ terms and the terminal remainder. The next four sections treat these tasks in that order.

4. Analytic horizontal-frame jets. Following [15, Section 1.2 and Appendix B.3], we express Lewis-weight derivatives through moving orthonormal frames. A horizontal gauge and a complex relative-row neighborhood yield dimension-free Cauchy bounds for all derivatives of N_x, uniformly in m, d and the smallest Lewis weight.

4.1. *Horizontal-frame coordinates.*

LEMMA 4.1 (Algebraic identities for a physical frame). *Fix $y \in \operatorname{int} K$, put $\widehat{B}_y = W_y^{\beta} A_y$, and let U be any real orthonormal frame for $\operatorname{col}(\widehat{B}_y)$. Then*

$$UU^{\mathsf{T}} = \widehat{B}_y(\widehat{B}_y^{\mathsf{T}}\widehat{B}_y)^{-1}\widehat{B}_y^{\mathsf{T}} = P_y, \tag{4.1}$$

$$\operatorname{diag}(UU^{\mathsf{T}}) = w_y. \tag{4.2}$$

In particular, every row of U is nonzero.

PROOF. Two full-column-rank matrices with the same column space have the same orthogonal projection. This gives (4.1); the Lewis fixed-point equation (2.1) then gives (4.2). Strict positivity of every Lewis weight and (4.2) imply that no row of U vanishes. $\square$

Fix a real base point satisfying (3.1) and a direction $e \in \mathbb{R}^d$ with $\|e\| = 1$. Along $y_\tau = x + \tau e$, let

$$\widehat{B}_\tau = W_{y_\tau}^{\beta} A_{y_\tau}.$$

Lemma 2.1 makes this a real analytic, full-rank matrix path. Orthonormalize it and apply the horizontal-gauge construction of Lemma B.1. We thereby obtain the physical horizontal frame U_τ, choosing its initial gauge so that

$$U_0 = \widehat{B}_0 = W_x^{\beta} A_x. \tag{4.3}$$

This is possible because $\widehat{B}_0^{\mathsf{T}}\widehat{B}_0 = g_0(x) = I_d$. The frame satisfies

$$U_\tau^{\mathsf{T}} U_\tau = I_d, \qquad U_\tau^{\mathsf{T}} U_\tau' = 0. \tag{4.4}$$

The horizontal gauge and its evolution equation are proved in Appendix B.

Write u_i^{T} for the ith row of U and define, initially on the real path,

$$w_i = u_i^{\mathsf{T}} u_i, \qquad \omega_i = w_i^{1/2}, \qquad \phi_i = \frac{u_i \otimes u_i}{\omega_i}. \tag{4.5}$$

Let Φ be the matrix whose ith row is ϕ_i^{T}, after identifying $\mathbb{R}^d \otimes \mathbb{R}^d$ with $\mathbb{R}^{d^2}$. Then

$$G := \Phi\Phi^{\mathsf{T}} = W^{-1/2} P^{\circ 2} W^{-1/2}, \qquad P = UU^{\mathsf{T}}, \tag{4.6}$$

and, with

$$a = \frac{2}{p}, \qquad c_p = 1 - \frac{2}{p},$$

we have

$$N(U) = 2(I - G)(aI + c_p G)^{-1}. \tag{4.7}$$

Indeed, $aI + c_p G = I - c_p(I - G) = I - c_p\bar{\Lambda}$. For the physical frame just constructed, Lemma 4.1 and the row definitions give

$$G(U_\tau) = W_{y_\tau}^{-1/2} P_{y_\tau}^{\circ 2} W_{y_\tau}^{-1/2}, \qquad N(U_\tau) = N_{y_\tau}. \tag{4.8}$$

Let

$$s_i(\tau) = (A_{y_\tau} e)_i.$$

Because each slack is affine,

$$s_i(\tau) = \frac{s_i(0)}{1 + \tau s_i(0)}. \tag{4.9}$$

At the base point, $A_x = W_0^{-\beta} U_0$, and hence

$$|s_i(0)| = w_i^{-\beta} \left| u_i^\top e \right| \le w_i^{1/2-\beta} = w_i^{1/p} \le 1, \qquad \left\| W_0^{1/2} s(0) \right\| \le 1. \tag{4.10}$$

The second inequality follows from $W_0^{1/2} s(0) = W_0^{1/p} U_0 e$ and $0 < W_0 \preceq I$.

Define

$$v(U,\tau) = W(U)^{1/2} s(\tau), \qquad z(U,\tau) = -s(\tau) - \beta W(U)^{-1/2} N(U) v(U,\tau), \tag{4.11}$$

where $W(U) = \operatorname{Diag}(u_i^\top u_i)$. From (2.11), $\widehat{B}'_\tau = \operatorname{Diag}(z_\tau) \widehat{B}_\tau$. The horizontal condition therefore gives the closed equation

$$U' = \mathcal{F}(U,\tau) := (I - UU^\top) \operatorname{Diag}(z(U,\tau)) U, \qquad U(0) = U_0. \tag{4.12}$$

The converse identification of a solution of this closed equation with the physical frame is proved after the analytic neighborhood has been established.

4.2. *The weighted complex neighborhood.* We complexify all matrices and use algebraic transpose $A^\top$, rather than the Hermitian adjoint $A^* = \overline{A}^\top$, in the Stiefel equations and the holomorphic maps $UU^\top$ and $\Phi\Phi^\top$. Norms are the Hermitian Euclidean, operator and Frobenius norms. For these norms,

$$\left\| A^\top \right\|_{\mathrm{op}} = \| A \|_{\mathrm{op}}, \qquad \left\| A^\top \right\|_{\mathrm{F}} = \| A \|_{\mathrm{F}}; \tag{4.13}$$

the first identity follows because the nonzero eigenvalues of $\overline{A} A^\top$ and $A^\top \overline{A}$ agree, while $A^\top \overline{A} = (A^* A)^\top$. Loewner-order statements are used only at the real base point; away from it, invertibility follows by operator-norm perturbation and a Neumann series. Fix $W_0 = \operatorname{Diag}(w_{0,i})$ and define

$$\| V \|_{\mathcal{X}} := \| V \|_{\mathrm{F}} + \max_{1 \le i \le m} \frac{\left\| e_i^\top V \right\|_2}{\sqrt{w_{0,i}}}. \tag{4.14}$$

For an $m \times q$ matrix X, set

$$\| X \|_{\mathrm{row},0} := \max_i \frac{\left\| e_i^\top X \right\|_2}{\sqrt{w_{0,i}}}, \tag{4.15}$$

and, for $y \in \mathbb{C}^m$, define

$$\| y \|_{\mathrm{adm},0} = \| y \|_2 + \left\| W_0^{-1/2} y \right\|_\infty. \tag{4.16}$$

LEMMA 4.2 (Analytic row map). *There are universal constants $c_0, C > 0$ such that, if $\| U - U_0 \|_{\mathcal{X}} \le \delta \le c_0$, then each $u_i^\top u_i$ lies in $\{ z : |z - w_{0,i}| \le w_{0,i}/2 \}$. The square-root branch agreeing with $\sqrt{w_{0,i}}$ is analytic, and*

$$\| \omega(U) - \omega(U_0) \|_2 \le C \| U - U_0 \|_{\mathrm{F}}, \tag{4.17}$$

$$\| \Phi(U) - \Phi(U_0) \|_{\mathrm{F}} + \| \Phi(U) - \Phi(U_0) \|_{\mathrm{row},0} \le C \| U - U_0 \|_{\mathcal{X}}, \tag{4.18}$$

$$\| G(U) - G(U_0) \|_{\mathrm{op}} + \| G(U) - G(U_0) \|_{\mathrm{row},0} \le C \| U - U_0 \|_{\mathcal{X}}. \tag{4.19}$$

The maps $U \mapsto \omega(U)$, $U \mapsto \Phi(U)$ and $U \mapsto G(U)$ are holomorphic on this neighborhood.

PROOF. Write $u_i = u_{0,i} + \Delta_i$. By (4.14), $\|\Delta_i\| \le \delta\sqrt{w_{0,i}}$, so

$$\left|u_i^\mathsf{T} u_i - w_{0,i}\right| \le 2\,\|u_{0,i}\|\,\|\Delta_i\| + \|\Delta_i\|^2 \le (2\delta+\delta^2)w_{0,i}.$$

For sufficiently small c_0, this is at most $w_{0,i}/2$.

Consider

$$\mathcal{R}(u) = \frac{u\otimes u}{(u^\mathsf{T} u)^{1/2}}.$$

After writing $u = \sqrt{w_{0,i}}\,\widehat{u}$, all normalized rows lie in one fixed complex neighborhood of the real unit sphere. On that neighborhood the selected branch of $\widehat{u} \mapsto (\widehat{u}^\mathsf{T}\widehat{u})^{1/2}$ and the normalized row map $\widehat{u} \mapsto \widehat{u}\otimes\widehat{u}/(\widehat{u}^\mathsf{T}\widehat{u})^{1/2}$ are holomorphic with universally bounded derivatives. Both unnormalized maps are homogeneous of degree one. Thus integration along the segment from $u_{0,i}$ to u_i gives, with no dependence on $w_{0,i}$,

$$|\omega_i(U) - \omega_i(U_0)| + \|\mathcal{R}(u_i) - \mathcal{R}(u_{0,i})\| \le C\,\|\Delta_i\|\,.$$

Squaring and summing proves (4.17) and the Frobenius part of (4.18); division by $\sqrt{w_{0,i}}$ proves the row part.

At the real base point, $G_0 = \Phi_0\Phi_0^\mathsf{T} \preceq I$, hence $\|\Phi_0\|_{\mathrm{op}} \le 1$. Equation (4.18) gives $\|\Phi\|_{\mathrm{op}} \le 1 + C\delta$. From

$$G - G_0 = (\Phi-\Phi_0)\Phi^\mathsf{T} + \Phi_0(\Phi-\Phi_0)^\mathsf{T}$$

we obtain

$$\|G-G_0\|_{\mathrm{op}} \le \|\Phi-\Phi_0\|_{\mathrm{op}}\left(\|\Phi\|_{\mathrm{op}} + \|\Phi_0\|_{\mathrm{op}}\right) \le C\,\|U-U_0\|_{\mathcal{X}}\,.$$

For the ith row, use

$$e_i^\mathsf{T}(G-G_0) = (\phi_i - \phi_{0,i})^\mathsf{T}\Phi^\mathsf{T} + \phi_{0,i}^\mathsf{T}(\Phi-\Phi_0)^\mathsf{T}$$

and $\|\phi_{0,i}\| = \sqrt{w_{0,i}}$. The already proved Frobenius and relative-row bounds yield

$$\left\|e_i^\mathsf{T}(G-G_0)\right\|_2 \le C\sqrt{w_{0,i}}\,\|U-U_0\|_{\mathcal{X}}\,,$$

which proves the row part of (4.19). Holomorphy follows from the selected square-root branches and matrix multiplication. □

LEMMA 4.3 (Resolvent stability). *Let*

$$R(U) = (aI + c_p G(U))^{-1}.$$

There are universal $c_1, C > 0$ such that, whenever $\|U-U_0\|_{\mathcal{X}} \le \delta \le c_1 a$,

$$\|R(U)\|_{\mathrm{op}} \le \frac{2}{a}. \tag{4.20}$$

For every $m\times q$ matrix X,

$$\|R(U)X\|_{\mathrm{row},0} \le Ca^{-1}\,\|X\|_{\mathrm{row},0} + Ca^{-2}\,\|X\|_{\mathrm{op}}\,. \tag{4.21}$$

Furthermore,

$$\|N(U)-N(U_0)\|_{\mathrm{op}} + \|N(U)-N(U_0)\|_{\mathrm{row},0} \le Ca^{-3}\,\|U-U_0\|_{\mathcal{X}}\,. \tag{4.22}$$

Finally, for every $y\in\mathbb{C}^m$,

$$\|N(U)y\|_2 \le Ca^{-1}\,\|y\|_2\,, \tag{4.23}$$

$$\left\|W_0^{-1/2}N(U)y\right\|_\infty \le Ca^{-1}\left\|W_0^{-1/2}y\right\|_\infty + Ca^{-2}\,\|y\|_2\,. \tag{4.24}$$

PROOF. At U_0, G_0 is real symmetric and $0 \preceq G_0 \preceq I$. Thus the smallest singular value of $aI + c_p G_0$ is at least a. Lemma 4.2 and a Neumann series prove (4.20) when $C\delta \le a/2$. Here and below $1/2 \le c_p < 1$, because $p \ge 4$.

The identity

$$R = a^{-1} I - \frac{c_p}{a} G R \tag{4.25}$$

will be used for row estimates. Since $G = \Phi\Phi^{\mathsf{T}}$, Lemma 4.2 gives

$$\left\| e_i^{\mathsf{T}} G \right\|_2 \le \|\phi_i\|_2 \left\| \Phi^{\mathsf{T}} \right\|_{\mathrm{op}} \le C\sqrt{w_{0,i}}. \tag{4.26}$$

Applying (4.25) to X, and using $\|RX\|_{\mathrm{op}} \le 2a^{-1} \|X\|_{\mathrm{op}}$, yields row by row

$$\frac{\left\|e_i^{\mathsf{T}} R X\right\|_2}{\sqrt{w_{0,i}}} \le a^{-1} \frac{\left\|e_i^{\mathsf{T}} X\right\|_2}{\sqrt{w_{0,i}}} + \frac{c_p}{a} \frac{\left\|e_i^{\mathsf{T}} G\right\|_2}{\sqrt{w_{0,i}}} \|RX\|_{\mathrm{op}} \le a^{-1} \|X\|_{\mathrm{row},0} + C a^{-2} \|X\|_{\mathrm{op}}.$$

Because $a + c_p = 1$ and R commutes with G,

$$N(U) = 2(I - G)R = \frac{2}{c_p}(R - I). \tag{4.27}$$

The resolvent identity gives

$$N(U) - N(U_0) = -2R(U)(G(U) - G_0)R(U_0). \tag{4.28}$$

The operator part of (4.22) follows from (4.19) and (4.20). For the row part, the right resolvent first gives

$$\|(G - G_0)R_0\|_{\mathrm{op}} \le C a^{-1} \|U - U_0\|_{\mathcal{X}}$$

and, directly on each row,

$$\|(G - G_0)R_0\|_{\mathrm{row},0} \le \|G - G_0\|_{\mathrm{row},0} \|R_0\|_{\mathrm{op}} \le C a^{-1} \|U - U_0\|_{\mathcal{X}}.$$

Apply (4.21) to

$$X = (G - G_0)R_0.$$

The two terms are respectively bounded by $Ca^{-2} \|U - U_0\|_{\mathcal{X}}$ and $Ca^{-3} \|U - U_0\|_{\mathcal{X}}$. This proves the row part of (4.22); the weaker common bound Ca^{-3} also covers the operator part since $a \le 1/2$. Finally, for a one-column matrix y,

$$\|y\|_{\mathrm{row},0} = \left\| W_0^{-1/2} y \right\|_{\infty}, \qquad \|y\|_{\mathrm{op}} = \|y\|_2.$$

Thus (4.27), (4.20), (4.21), and $2/c_p \le 4$ give (4.23)–(4.24). □

PROPOSITION 4.4 (Identification of the horizontal physical frame). *Let $y_\tau = x + \tau e$ range over a real interval $I \ni 0$ contained in* $\operatorname{int} K$. *The physical horizontal frame constructed above is the unique real analytic orthonormal frame for* $\operatorname{col}(\widehat{B}_\tau)$ *with initial value* (4.3) *and gauge* (4.4). *Every row of this frame is nonzero. On every subinterval on which it lies in the row-map neighborhood of Lemma 4.2, it solves* (4.12). *Conversely, a real solution of* (4.12) *with initial value* (4.3) *agrees with the physical horizontal frame on every connected interval on which it remains in that neighborhood. If a finite endpoint lies in I and the solution remains in a closed subball of the neighborhood up to that endpoint, the identification continues across it.*

PROOF. Lemma 2.1 shows that $\widehat{B}_\tau = W_{y_\tau}^{\beta} A_{y_\tau}$ is real analytic and full column rank on I. The analytic orthonormal frame

$$\widetilde{U}_\tau = \widehat{B}_\tau \big(\widehat{B}_\tau^{\mathsf{T}} \widehat{B}_\tau\big)^{-1/2}$$

spans its column space and satisfies $\widetilde{U}_0 = \widehat{B}_0 = U_0$. Applying the gauge construction in Lemma B.1, with initial gauge I_d, gives the stated frame. If U and $\overline{U}$ are two such frames, then $\overline{U} = UO$ for an orthogonal matrix path O with $O(0) = I$. Their horizontal conditions give $O' = 0$, so $O = I$; this proves uniqueness without invoking any choice of orthonormalization. Lemma 4.1 gives the nonzero-row claim and (4.8) identifies its row maps with the physical Lewis quantities.

The slack identity (4.9) gives $A'_{y_\tau} = -\operatorname{Diag}(s(\tau))A_{y_\tau}$. Moreover, (2.11), followed by (4.8), gives

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\log w_{y_\tau} = -W(U_\tau)^{-1/2}N(U_\tau)v(U_\tau,\tau).$$

The product rule therefore gives

$$\widehat{B}'_\tau = \operatorname{Diag}\Big(-s(\tau) - \beta W(U_\tau)^{-1/2}N(U_\tau)v(U_\tau,\tau)\Big)\,\widehat{B}_\tau.$$

The horizontal-gauge formula in Lemma B.1 now yields precisely (4.12). Lemmas 4.2 and 4.3 make its right side real analytic, hence locally Lipschitz in the $\mathcal{X}$ norm, throughout the stated neighborhood. Thus a real solution of the closed equation and the physical frame coincide near zero by ODE uniqueness. Reapplying uniqueness at every point of their common interval proves identification on its connected component. If a finite endpoint in I is approached inside a closed subball, the rational slack functions are regular there and all U-dependent denominators stay uniformly away from zero. The vector field is bounded on a slightly larger compact neighborhood, so the closed solution has a limit. Local existence and uniqueness at that limiting point extend both solutions and the identification across the endpoint. □

REMARK 4.5. A uniform bound on $\|N(U)\|_{\mathrm{row},0}$ is generally false. On the null space of G, N has an identity component of size comparable to p. Lemma 4.3 controls instead differences of N, derivatives of N, and the action of N on vectors whose Euclidean and relative-coordinate norms are both controlled.

4.3. *Complex existence and Cauchy estimates.* The elementary analytic ODE lemma used next is proved in Appendix A.

LEMMA 4.6 (Uniform analytic flow). *There are universal constants $c, C > 0$ such that the solution of* (4.12) *extends holomorphically to*

$$|\tau| < 2\rho_p, \qquad \rho_p = cp^{-3} = c'a^3, \tag{4.29}$$

and satisfies

$$\|U(\tau) - U_0\|_{\mathcal{X}} \le c_1 a/2. \tag{4.30}$$

On the same disk,

$$U(\tau)^{\mathsf{T}}U(\tau) = I_d. \tag{4.31}$$

PROOF. Decrease c_1 if necessary so that $c_1 a \le c_0$, where c_0 is the row-map radius in Lemma 4.2, and set

$$\delta = c_1 a, \qquad \mathcal{D} = \{(U,\tau) : \|U - U_0\|_{\mathcal{X}} < \delta,\ |\tau| < 1/4\}.$$

For $(U,\tau) \in \mathcal{D}$, Lemma 4.2 gives

$$2^{-1/2}\sqrt{w_{0,i}} \le |\omega_i(U)| \le (3/2)^{1/2}\sqrt{w_{0,i}}, \qquad 1 \le i \le m. \tag{4.32}$$

Indeed, $\omega_i(U)^2 = u_i^\top u_i$ lies in the disk of radius $w_{0,i}/2$ around $w_{0,i}$.

By (4.10), $|s_i(0)| \le 1$. Hence $|1 + \tau s_i(0)| \ge 3/4$ on $|\tau| \le 1/4$, and (4.9) gives

$$\|s(\tau)\|_\infty \le 2, \qquad \left\|W_0^{1/2} s(\tau)\right\|_2 \le 2. \tag{4.33}$$

The second bound follows term by term from

$$\sum_i w_{0,i}\,|s_i(\tau)|^2 \le (4/3)^2 \sum_i w_{0,i}\,|s_i(0)|^2 \le (4/3)^2.$$

Since $v_i = \omega_i(U) s_i(\tau)$, (4.32) and (4.33) imply

$$\|v\|_2 \le C\left\|W_0^{1/2} s(\tau)\right\|_2 \le C, \qquad \left\|W_0^{-1/2} v\right\|_\infty \le C\,\|s(\tau)\|_\infty \le C. \tag{4.34}$$

Put $y = N(U)v$. Equations (4.23), (4.24) and (4.34) give

$$\|y\|_2 \le C a^{-1}, \qquad \left\|W_0^{-1/2} y\right\|_\infty \le C a^{-2}. \tag{4.35}$$

The diagonal operator $W_0^{1/2} W(U)^{-1/2}$ has operator norm at most C by (4.32). Consequently

$$\left\|W(U)^{-1/2} y\right\|_\infty \le C\left\|W_0^{-1/2} y\right\|_\infty \le C a^{-2}, \qquad \left\|W_0^{1/2} W(U)^{-1/2} y\right\|_2 \le C\,\|y\|_2 \le C a^{-1}.$$

Since $0 < \beta < 1/2$, (4.11) and (4.33) now give separately in the two relevant norms

$$\|z\|_\infty \le C a^{-2}, \qquad \left\|W_0^{1/2} z\right\|_2 \le C a^{-1}. \tag{4.36}$$

Let $Z = \mathrm{Diag}(z)$. Since U remains in a fixed operator-norm neighborhood of U_0,

$$\|u_i\|_2 \le \|u_{0,i}\|_2 + \|u_i - u_{0,i}\|_2 \le (1+\delta)\sqrt{w_{0,i}} \le C\sqrt{w_{0,i}},$$

and therefore

$$\|ZU\|_{\mathrm{F}}^2 = \sum_i |z_i|^2\,\|u_i\|_2^2 \le C\left\|W_0^{1/2} z\right\|_2^2 \le C a^{-2}.$$

For the relative row norm,

$$e_i^\top \mathcal{F}(U,\tau) = z_i u_i^\top - u_i^\top (U^\top Z U). \tag{4.37}$$

Here

$$\|U\|_{\mathrm{op}} \le \|U_0\|_{\mathrm{op}} + \|U - U_0\|_{\mathrm{F}} \le 1 + c_1 a \le C$$

and, by (4.13),

$$\left\|U^\top Z U\right\|_{\mathrm{op}} \le \left\|U^\top\right\|_{\mathrm{op}} \|ZU\|_{\mathrm{op}} \le C\,\|ZU\|_{\mathrm{F}} \le C a^{-1}.$$

Dividing (4.37) by $\sqrt{w_{0,i}}$ and using the last two displays gives, for every i,

$$\frac{\left\|e_i^\top \mathcal{F}(U,\tau)\right\|_2}{\sqrt{w_{0,i}}} \le C\,|z_i| + C\left\|U^\top Z U\right\|_{\mathrm{op}} \le C a^{-2}.$$

The Frobenius estimate follows directly from

$$\|\mathcal{F}(U,\tau)\|_{\mathrm{F}} \le \left(1+\|U\|_{\mathrm{op}}^2\right)\|ZU\|_{\mathrm{F}} \le Ca^{-1}.$$

Thus

$$\sup_{\|U-U_0\|_{\mathcal{X}}<\delta,\ |\tau|<1/4} \|\mathcal{F}(U,\tau)\|_{\mathcal{X}} \le Ca^{-2}. \tag{4.38}$$

On $\mathcal{D}$, every denominator $1+\tau s_i(0)$ is nonzero, every chosen square-root branch remains inside its defining disk, and $aI + c_pG(U)$ is invertible by Lemma 4.3. Therefore s, ω, $W(U)^{-1/2}$, N, v, z and $\mathcal{F}$ are jointly holomorphic on $\mathcal{D}$. Applying Lemma A.1 with

$$\tau_0 = \frac{1}{4}, \qquad \delta = c_1 a, \qquad M = Ca^{-2}$$

gives

$$\min\{\tau_0, \delta/M\} \ge c_2 a^3.$$

Here $a \le 1/2$, and every constant used above is universal: $\|U_0\|_{\mathrm{op}} = 1$, while the relative-row normalization absorbs all dependence on m, d, and $\min_i w_{0,i}$. Thus the ODE lemma gives a holomorphic solution on a disk of radius $c_3 a^3$ that stays in the half ball. Define $\rho_p = c_3 a^3/2$, after decreasing c_3 so that $2\rho_p < 1/4$. This proves (4.30) on $|\tau| < 2\rho_p$, with a dimension-free constant.

To verify the complex Stiefel constraint, set $Q = U^{\mathsf{T}}U$ and $\mathcal{M} = U^{\mathsf{T}}ZU$. Directly from (4.12),

$$Q' = \mathcal{M}(I-Q) + (I-Q)\mathcal{M}.$$

Thus $E = I - Q$ solves the homogeneous linear matrix equation $E' = -\mathcal{M}E - E\mathcal{M}$ with $E(0) = 0$. Uniqueness gives $E \equiv 0$, proving (4.31) throughout the disk. □

LEMMA 4.7 (Kernel identity on the complex flow). *On the disk of Lemma 4.6,*

$$G(U(\tau))\omega(U(\tau)) = \omega(U(\tau)), \qquad N(U(\tau))\omega(U(\tau)) = 0. \tag{4.39}$$

PROOF. By (4.31), $\sum_j u_j u_j^{\mathsf{T}} = U^{\mathsf{T}}U = I$. Therefore

$$(G\omega)_i = \frac{1}{\omega_i}\sum_j (u_i^{\mathsf{T}}u_j)^2 = \frac{1}{\omega_i} u_i^{\mathsf{T}}\left(\sum_j u_j u_j^{\mathsf{T}}\right) u_i = \omega_i.$$

Thus $G\omega = \omega$. Since $a + c_p = 1$, we also have $R\omega = (aI + c_pG)^{-1}\omega = \omega$, and (4.7) gives $N\omega = 2(I-G)R\omega = 0$. □

THEOREM 4.8 (Analytic horizontal-frame jets). *Write $\omega_x = w_x^{1/2}$ coordinatewise. There are universal constants $c, C > 0$ such that, for every real base point $x \in \operatorname{int} K$, every $k \ge 1$, and every $r_1, \dots, r_k \in \mathbb{R}^d$,*

$$\begin{aligned} &\left\|D^k N_x[r_1,\dots,r_k]\right\|_{\mathrm{op}} + \left\|W_x^{-1/2} D^k N_x[r_1,\dots,r_k]\right\|_{2\to\infty} \\ &\qquad \le Cp^2 k^k \rho_p^{-k} \prod_{j=1}^k \|r_j\|_{g_0(x)}, \end{aligned} \tag{4.40}$$

$$\left\|D^k \omega_x[r_1,\dots,r_k]\right\|_2 \le Ck^k \rho_p^{-k} \prod_{j=1}^k \|r_j\|_{g_0(x)}. \tag{4.41}$$

For repeated directions,

$$\left\|D^k N_x[r^k]\right\|_{\mathrm{op}} + \left\|W_x^{-1/2} D^k N_x[r^k]\right\|_{2\to\infty} \le Cp^2 k! \rho_p^{-k} \|r\|_{g_0(x)}^k, \tag{4.42}$$

$$\left\|D^k \omega_x[r^k]\right\|_2 \le Ck! \rho_p^{-k} \|r\|_{g_0(x)}^k. \tag{4.43}$$

PROOF. We first prove the repeated-direction estimates. Given x and $r \neq 0$, use the affine map

$$\Psi(\widetilde{x}) = x + g_0(x)^{-1/2}\widetilde{x}.$$

Then $\widetilde{g}_0(0) = I_d$ and the direction corresponding to r is $\widetilde{r} = g_0(x)^{1/2} r$, whose Euclidean norm is $\|r\|_{g_0(x)}$. Equations (2.45) and (2.46), together with invariance of the Lewis weights at corresponding points, show that it suffices by homogeneity to work under (3.1) with $\|r\| = 1$.

Let $U(\tau)$ be the analytic solution in Lemma 4.6, with initial frame (4.3). For real τ near zero, Proposition 4.4 and real ODE uniqueness give

$$N(U(\tau)) = N_{x+\tau r}, \qquad \omega(U(\tau)) = w_{x+\tau r}^{1/2}. \tag{4.44}$$

Indeed, (4.10) implies $1 + \tau s_i(0) > 0$ for real $|\tau| < 2\rho_p < 1/4$, so this real segment remains in $\operatorname{int} K$ and the physical frame is defined there. The local identification near zero gives

$$\frac{\mathrm{d}^k}{\mathrm{d}\tau^k} N(U(\tau))\Big|_{\tau=0} = D^k N_x[r,\dots,r], \qquad \frac{\mathrm{d}^k}{\mathrm{d}\tau^k} \omega(U(\tau))\Big|_{\tau=0} = D^k \omega_x[r,\dots,r].$$

Since (4.30) places the circle $|\tau| = \rho_p$ in the resolvent neighborhood, Lemma 4.3 gives

$$\sup_{|\tau|\le\rho_p} \left(\|N(U(\tau)) - N(U_0)\|_{\mathrm{op}} + \|N(U(\tau)) - N(U_0)\|_{\mathrm{row},0}\right) \le Ca^{-3}(c_1 a/2) \le Cp^2.$$

For $k \ge 1$, subtracting $N(U_0)$ does not change the derivative. Cauchy's formula on the circle $|\tau| = \rho_p$, strictly inside $|\tau| < 2\rho_p$, in the matrix Banach norm $\|M\|_{\mathrm{op}} + \left\|W_x^{-1/2} M\right\|_{2\to\infty}$ gives the following estimate; here its second summand is exactly $\|M\|_{\mathrm{row},0}$:

$$\left\|\frac{\mathrm{d}^k}{\mathrm{d}\tau^k} N(U(\tau))\big|_{\tau=0}\right\|_{\mathrm{op}} + \left\|W_x^{-1/2} \frac{\mathrm{d}^k}{\mathrm{d}\tau^k} N(U(\tau))\big|_{\tau=0}\right\|_{2\to\infty} \le Cp^2 k! \rho_p^{-k},$$

which proves (4.42). Equation (4.17) and (4.30) give $\sup_{|\tau|\le\rho_p} \|\omega(U(\tau)) - \omega(U_0)\|_2 \le Ca \le C$. The same Cauchy argument proves (4.43) without a prefactor depending on p.

It remains to pass from diagonal to mixed directions. The maps $D^k N_x$ and $D^k \omega_x$ are symmetric k-linear maps into their respective Banach spaces. For unit vectors $r_1, \dots, r_k$, the real polarization identity and the triangle inequality give

$$\|T[r_1, \dots, r_k]\| \le \frac{1}{2^k k!} \sum_{\epsilon \in \{-1,1\}^k} \left\|T[(\epsilon_1 r_1 + \dots + \epsilon_k r_k)^k]\right\| \le \frac{k^k}{k!} \sup_{\|r\|\le 1} \left\|T[r^k]\right\|.$$

Scaling each input separately gives the same bound multiplied by $\prod_j \|r_j\|$. Combining this with the diagonal $k!$ bounds and undoing the affine normalization proves (4.40) and (4.41). □

5. Projection-weighted Wick contractions. Kook [15, Section 3.3] estimated the first Wiener-chaos contractions case by case. We isolate their Parseval-frame structure in an all-order theorem whose double-contraction bound removes the residual $d^{1/2}$ loss. Throughout this section all tensors are real, and Tr denotes the unnormalized trace

$$(\mathrm{Tr}^r T)_{a_1 \dots a_{n-2r}} = \sum_{b_1,\dots,b_r=1}^{d} T_{b_1 b_1 \dots b_r b_r a_1 \dots a_{n-2r}}.$$

For tensors S, T, the contraction $S \otimes_\ell T$ sums the last ℓ indices of each tensor, and $S \widetilde{\otimes}_\ell T$ is its average over all permutations of the remaining indices. Thus symmetrization is an orthogonal projection in Hilbert–Schmidt norm and introduces no additional combinatorial factor. Let $I_n(T)$ denote the order-n Wick polynomial, normalized by

$$\mathbb{E}[I_n(S) I_m(T)] = \mathbf{1}_{\{m=n\}} n! \langle S, T \rangle_{\mathrm{HS}}. \tag{5.1}$$

The raw-to-Wick and product formulas [32, Chapter 1] are

$$T[h^{\otimes n}] = \sum_{r=0}^{\lfloor n/2 \rfloor} a_{n,r} I_{n-2r}(\mathrm{Tr}^r T), \qquad a_{n,r} = \frac{n!}{2^r r! (n-2r)!}, \tag{5.2}$$

$$I_a(S) I_b(T) = \sum_{\ell=0}^{a \wedge b} \ell! \binom{a}{\ell} \binom{b}{\ell} I_{a+b-2\ell}(S \widetilde{\otimes}_\ell T), \tag{5.3}$$

where $\widetilde{\otimes}_\ell$ is the symmetrized ℓ-fold contraction. Appendix C records these identities. We use the conventions that $I_0(c) = c$, a contraction whose order exceeds one of the tensor orders is absent, and the Hilbert–Schmidt norm of a scalar is its absolute value. These conventions include without exception the cases in which a trace leaves a tensor of order zero or one.

Let $U \in \mathbb{R}^{m \times d}$ satisfy $U^\mathsf{T} U = I_d$ and assume that every row is nonzero. Write the column representation of its ith row as

$$u_i = \sqrt{w_i}\, k_i, \qquad w_i = \|u_i\|^2 > 0, \qquad \|k_i\| = 1, \qquad \sum_i w_i = d, \tag{5.4}$$

and set $K_i = k_i \otimes k_i$. Define

$$\Gamma_{ij} = \sqrt{w_i w_j} \langle K_i, K_j \rangle_{\mathrm{HS}} = \frac{P_{ij}^2}{\sqrt{w_i w_j}}, \qquad P = U U^\mathsf{T}. \tag{5.5}$$

The Schur-product theorem [11, Section 5.2] gives

$$0 \preceq \Gamma \preceq I_m. \tag{5.6}$$

Indeed, $P \preceq I$ implies $P \circ (I - P) \succeq 0$, which is $W - P^{\circ 2} \succeq 0$; conjugation by $W^{-1/2}$ proves (5.6).

THEOREM 5.1 (Weighted Wick contraction). *Let $n \geq 1$ and let $T_i \in \mathrm{Sym}^n(\mathbb{R}^d)$, $1 \leq i \leq m$. Suppose that, for some $L \geq 0$,*

$$\left(\sum_i \left| T_i[x^{\otimes n}] \right|^2 \right)^{1/2} \leq L \|x\|^n, \tag{5.7}$$

$$|T_i[x_1, \dots, x_n]| \leq L \sqrt{w_i} \prod_{j=1}^{n} \|x_j\|, \tag{5.8}$$

$$\left| \sum_i \sqrt{w_i} T_i[x^{\otimes n}] \right| \leq L \|x\|^n \tag{5.9}$$

for all indicated vectors. If $h \sim N(0, I_d)$, then

$$\left\| \sum_i \sqrt{w_i}(k_i^\mathsf{T} h)^2 T_i[h^{\otimes n}] \right\|_{L^2} \le (Cn)^{Cn} L(d+n)^{n/2}. \tag{5.10}$$

PROOF. Set

$$\mathcal{P}(h) = \sum_i \sqrt{w_i}(k_i^\mathsf{T} h)^2 T_i[h^{\otimes n}].$$

Since $\|k_i\| = 1$,

$$(k_i^\mathsf{T} h)^2 = 1 + I_2(K_i). \tag{5.11}$$

The noncentered part

$$\mathcal{P}_0(h) = \sum_i \sqrt{w_i} T_i[h^{\otimes n}]$$

satisfies, by (5.9) and the Gaussian radial moment bound,

$$\|\mathcal{P}_0\|_{L^2} \le L(\mathbb{E}\,\|h\|^{2n})^{1/2} \le (Cn)^{n/2} L(d+n)^{n/2}. \tag{5.12}$$

For the centered part, put $S_{i,r} = \operatorname{Tr}^r T_i$ and $m_r = n - 2r$. Then

$$\mathcal{P}_c = \sum_{r=0}^{\lfloor n/2 \rfloor} a_{n,r} \sum_i \sqrt{w_i} I_2(K_i) I_{m_r}(S_{i,r}). \tag{5.13}$$

Only the contractions $\ell = 0, 1, 2$ occur. The $\ell = 1$ term is absent when $m_r = 0$, and the $\ell = 2$ term is absent when $m_r < 2$. Thus, when $n = 1$, only $\ell = 0, 1$ occur, while for $n = 2$ the $\ell = 2$ contraction is a scalar. The estimates below use the same formulas in these edge cases.

For $\ell = 0$, an unsymmetrized kernel is

$$A_{r,0} = \sum_i \sqrt{w_i} K_i \otimes S_{i,r}.$$

Using (5.6) coordinate by coordinate,

$$\|A_{r,0}\|_{\mathrm{HS}}^2 = \sum_{i,j} \Gamma_{ij} \langle S_{i,r}, S_{j,r} \rangle_{\mathrm{HS}} \le \sum_i \|S_{i,r}\|_{\mathrm{HS}}^2. \tag{5.14}$$

Symmetrization cannot increase the Hilbert–Schmidt norm.

For $\ell = 1$, an unsymmetrized representative is

$$A_{r,1} = \sum_i u_i \otimes S_{i,r}(k_i, \cdot).$$

For a multi-index $\gamma = (a_1, \dots, a_{m_r - 1})$, put

$$c_{i,\gamma} = \sum_{b=1}^d (k_i)_b (S_{i,r})_{b, a_1, \dots, a_{m_r - 1}}.$$

Then $(A_{r,1})_{a_0,\gamma} = \sum_i (u_i)_{a_0} c_{i,\gamma}$. For each fixed γ, the Parseval identity for the rows of U gives

$$\sum_{a_0=1}^d \left| \sum_i (u_i)_{a_0} c_{i,\gamma} \right|^2 = c_\gamma^\mathsf{T} P c_\gamma \le \sum_i |c_{i,\gamma}|^2.$$

Summing over γ and using $\|k_i\| = 1$ gives

$$\|A_{r,1}\|_{\mathrm{HS}}^2 \le \sum_i \|S_{i,r}(k_i,\cdot)\|_{\mathrm{HS}}^2 \le \sum_i \|S_{i,r}\|_{\mathrm{HS}}^2 . \tag{5.15}$$

The same bound holds after symmetrization.

For $\ell = 2$, which is present only when $m_r \ge 2$, the kernel is

$$A_{r,2} = \sum_i \sqrt{w_i} S_{i,r}(k_i, k_i, \cdot).$$

For a fixed output multi-index $(a_1, \dots, a_{m_r-2})$, expanding the r traces and applying (5.8) gives

$$\begin{aligned} &\left|S_{i,r}(k_i, k_i, e_{a_1}, \dots, e_{a_{m_r-2}})\right| \\ &\quad \le \sum_{b_1,\dots,b_r=1}^{d} |T_i[e_{b_1}, e_{b_1}, \dots, e_{b_r}, e_{b_r}, k_i, k_i, e_{a_1}, \dots]| \le L d^r \sqrt{w_i}. \end{aligned}$$

Thus each output coordinate of $A_{r,2}$ is at most

$$L d^r \sum_i w_i = L d^{r+1}. \tag{5.16}$$

There are d^{m_r-2} output coordinates, and hence

$$\|A_{r,2}\|_{\mathrm{HS}} \le L d^{r+1} d^{(m_r-2)/2} = L d^{n/2}, \tag{5.17}$$

because

$$r + 1 + \frac{n-2r-2}{2} = \frac{n}{2}. \tag{5.18}$$

It remains to track the coefficients. Regard $\mathbf{T}(h) = (T_i[h^{\otimes n}])_{i=1}^m$ as an $\mathbb{R}^m$-valued polynomial. The coordinatewise raw-to-Wick identity is

$$\mathbf{T}(h) = \sum_{r=0}^{\lfloor n/2 \rfloor} a_{n,r} \big(I_{m_r}(S_{i,r})\big)_{i=1}^m.$$

The chaos orders m_r are pairwise distinct. Orthogonality and (5.1) therefore give the exact identity

$$\mathbb{E}\,\|\mathbf{T}(h)\|_2^2 = \sum_{r=0}^{\lfloor n/2 \rfloor} a_{n,r}^2 m_r! \sum_i \|S_{i,r}\|_{\mathrm{HS}}^2 . \tag{5.19}$$

On the other hand, (5.7) and the Gaussian radial-moment bound imply

$$\sum_{r=0}^{\lfloor n/2 \rfloor} a_{n,r}^2 m_r! \sum_i \|S_{i,r}\|_{\mathrm{HS}}^2 \le L^2 (C(d+n))^n. \tag{5.20}$$

This is the only point in the proof at which orthogonality between different values of r is used. After multiplication by $I_2(K_i)$, the output order is

$$q(r,\ell) = m_r + 2 - 2\ell = n - 2r + 2 - 2\ell, \tag{5.21}$$

and different pairs (r, ℓ) can have the same value of $q(r, \ell)$. We therefore make no orthogonality assertion for the terms in (5.13); from this point onward they are combined by the triangle inequality.

For $\ell=0$, the L^2 norm of the contribution for a fixed r is at most

$$a_{n,r}\sqrt{(m_r+2)!}\left(\sum_i \|S_{i,r}\|_{\mathrm{HS}}^2\right)^{1/2} \le (n+2)\left(a_{n,r}^2 m_r! \sum_i \|S_{i,r}\|_{\mathrm{HS}}^2\right)^{1/2}.$$

For $\ell=1$, the product coefficient is $2m_r$, the output order is m_r, and the corresponding bound is at most

$$2m_r\left(a_{n,r}^2 m_r! \sum_i \|S_{i,r}\|_{\mathrm{HS}}^2\right)^{1/2}.$$

Taking the triangle inequality first, and then Cauchy–Schwarz in r, the combined $\ell=0,1$ contribution is at most

$$3(n+2)\sqrt{1+\lfloor n/2\rfloor}\left(\sum_{r=0}^{\lfloor n/2\rfloor} a_{n,r}^2 m_r! \sum_i \|S_{i,r}\|_{\mathrm{HS}}^2\right)^{1/2}.$$

Here the absent $\ell=1$ term for $m_r=0$ is simply omitted. By (5.20), the last display is bounded by $(Cn)^{Cn}L(d+n)^{n/2}$.

For $\ell=2$, the product coefficient is $m_r(m_r-1)$ and the output order is m_r-2. Equations (5.17) and the elementary estimates

$$a_{n,r}\le (Cn)^n, \qquad m_r(m_r-1)\le n^2, \qquad \sqrt{(m_r-2)!}\le (Cn)^{n/2}$$

show that each such contribution is at most $(Cn)^{Cn}Ld^{n/2}$. There are at most $1+\lfloor n/2\rfloor$ values of r. The triangle inequality now proves the same bound for $\mathcal{P}_c$. Combining it with (5.12) proves (5.10). □

REMARK 5.2. Condition (5.7) controls the global tensor energy, (5.8) controls each row relative to its leverage weight, and (5.9) supplies the weighted mean cancellation. Thus the theorem is tailored to Gaussian polynomials indexed by a Parseval frame, rather than to arbitrary weighted families.

The double contraction is the only point at which the number of rows could enter. The factor $\sqrt{w_i}$ in (5.8) turns the row sum into $\sum_i w_i = d$, and (5.18) exactly compensates for the two contracted Gaussian indices.

6. Base-point bottleneck polynomials. This section combines Theorems 4.8 and 5.1 to control the base-point terms $H_j(0)$ in the selective expansion (3.7).

Return to the normalization (3.1). Write

$$U=W^{\beta}A_x, \qquad U^{\mathsf{T}}U=I_d, \qquad B=W^{1/2}A_x=W^{1/p}U, \qquad \|B\|_{\mathrm{op}}\le 1. \tag{6.1}$$

Let $u_i=\sqrt{w_i}k_i$ be the rows of U. For $h\sim N(0,I_d)$,

$$q_i=\sqrt{w_i}(k_i^{\mathsf{T}}h)^2, \qquad v=Bh. \tag{6.2}$$

LEMMA 6.1 (Cubic base term). *There is a universal $C>0$ such that*

$$\|E_1(0)+\beta H_1(0)\|_{L^2}\le Cp^2\sqrt{d}. \tag{6.3}$$

PROOF. Let $u=Uh$ and define

$$M=-W^{-1/2}(I+\beta N)W^{1/p}U. \tag{6.4}$$

Since $W^{-\beta} = W^{-1/2} W^{1/p}$, substitution into (3.4) gives

$$E_1(0) + \beta H_1(0) = -\langle Mh, u^{\circ 2}\rangle =: -T(h). \tag{6.5}$$

The polynomial T is odd and centered. The Gaussian Poincaré inequality [2, 24] gives

$$\mathbb{E}T^2 \le \mathbb{E}\,\|\nabla T(h)\|^2 \le 2\mathbb{E}\left\|M^{\mathsf{T}} u^{\circ 2}\right\|^2 + 8\mathbb{E}\left\|U^{\mathsf{T}}(u \circ Mh)\right\|^2. \tag{6.6}$$

If m_i^{T} is the ith row of M, then

$$\mathbb{E}\left\|U^{\mathsf{T}}(u \circ Mh)\right\|^2 \le \sum_i \mathbb{E}[u_i^2 (m_i^{\mathsf{T}} h)^2] \le 3\sum_i w_i \|m_i\|^2 = 3\,\mathrm{tr}(M^{\mathsf{T}} W M). \tag{6.7}$$

Let $C_M = MM^{\mathsf{T}}$. Since u is Gaussian with covariance $P = UU^{\mathsf{T}}$,

$$\begin{aligned}\mathbb{E}\left\|M^{\mathsf{T}} u^{\circ 2}\right\|^2 &= w^{\mathsf{T}} C_M w + 2\,\mathrm{tr}(C_M P^{\circ 2}) \\ &\le \left\|M^{\mathsf{T}} w\right\|^2 + 2\,\mathrm{tr}(M^{\mathsf{T}} W M),\end{aligned} \tag{6.8}$$

where $P^{\circ 2} \preceq W$ was used.

Now

$$W^{1/2} M = -(I + \beta N) W^{1/p} U.$$

Using $0 \preceq N \preceq pI$, $W^{1/p} \preceq I$ and $\|U\|_{\mathrm{F}}^2 = d$,

$$\mathrm{tr}(M^{\mathsf{T}} W M) \le Cp^2 d. \tag{6.9}$$

Also, using $P \preceq I$,

$$\begin{aligned}\left\|M^{\mathsf{T}} w\right\|^2 &= w^{\mathsf{T}} W^{-1/2}(I + \beta N) W^{1/p} P W^{1/p} (I + \beta N) W^{-1/2} w \\ &\le Cp^2 w^{\mathsf{T}} W^{-1} w = Cp^2 \sum_i w_i = Cp^2 d.\end{aligned} \tag{6.10}$$

Combining (6.6)–(6.10) proves (6.3). □

For $j \ge 2$, define the symmetric j-tensor

$$\mathcal{T}_i^{(j)}[x_1, \ldots, x_j] = \frac{1}{j} \sum_{r=1}^{j} e_i^{\mathsf{T}} D^{j-1} N_x[x_1, \ldots, \widehat{x_r}, \ldots, x_j] B x_r. \tag{6.11}$$

On the diagonal,

$$\mathcal{T}_i^{(j)}[h^{\otimes j}] = e_i^{\mathsf{T}} D^{j-1} N_x[h^{j-1}] Bh. \tag{6.12}$$

THEOREM 6.2 (Natural scale of every bottleneck polynomial). *For every $j \ge 2$,*

$$\|H_j(0)\|_{L^2} \le p^2 (Cj)^{Cj} \rho_p^{-(j-1)} (d + j)^{j/2}. \tag{6.13}$$

In particular,

$$\|H_4(0)\|_{L^2} \le p^{O(1)} d^2. \tag{6.14}$$

PROOF. We verify the hypotheses of Theorem 5.1 with $n = j$. First, Theorem 4.8, (6.12) and $\|B\|_{\mathrm{op}} \le 1$ give

$$\left(\sum_i \left|\mathcal{T}_i^{(j)}[h^{\otimes j}]\right|^2\right)^{1/2} \le Cp^2 (j-1)! \rho_p^{-(j-1)} \|h\|^j . \tag{6.15}$$

Second, the mixed row estimate (4.40) gives

$$\left|\mathcal{T}_i^{(j)}[x_1, \dots, x_j]\right| \le Cp^2 j^j \rho_p^{-(j-1)} \sqrt{w_i} \prod_{\ell=1}^{j} \|x_\ell\| . \tag{6.16}$$

For the weighted mean, let $\omega = w^{1/2}$. The identity $N\omega = 0$ holds along the complex line by Lemma 4.7. Differentiating it $j-1$ times in direction h gives

$$D^{j-1}N[h^{j-1}]\omega = -\sum_{\ell=0}^{j-2} \binom{j-1}{\ell} D^\ell N[h^\ell] D^{j-1-\ell}\omega[h^{j-1-\ell}]. \tag{6.17}$$

To make the cancellation explicit, put $k = j-1$. Entrywise differentiation of the identity $N_y^{\mathsf{T}} = N_y$ shows that every matrix $D^\ell N[h^\ell]$ is symmetric at a real base point. In particular,

$$\omega^{\mathsf{T}} D^k N[h^k] Bh = \left(D^k N[h^k]\omega\right)^{\mathsf{T}} Bh. \tag{6.18}$$

Substituting (6.17) in (6.18), and using $\|B\|_{\mathrm{op}} \le 1$, gives

$$\begin{aligned} \left|\sum_i \sqrt{w_i}\mathcal{T}_i^{(j)}[h^{\otimes j}]\right| &= \left|\omega^{\mathsf{T}} D^{j-1} N[h^{j-1}] Bh\right| \\ &\le \|h\| \sum_{\ell=0}^{j-2} \binom{j-1}{\ell} \left\|D^\ell N[h^\ell]\right\|_{\mathrm{op}} \left\|D^{j-1-\ell}\omega[h^{j-1-\ell}]\right\|_2 . \end{aligned} \tag{6.19}$$

The factorial cancellation in this Leibniz sum is exact. Namely, for $1 \le \ell \le k-1$, (4.42) and (4.43) give

$$\begin{aligned} &\binom{k}{\ell} \left\|D^\ell N[h^\ell]\right\|_{\mathrm{op}} \left\|D^{k-\ell}\omega[h^{k-\ell}]\right\|_2 \\ &\quad \le Cp^2 \binom{k}{\ell} \ell!(k-\ell)! \rho_p^{-k} \|h\|^k = Cp^2 k! \rho_p^{-k} \|h\|^k . \end{aligned} \tag{6.20}$$

For $\ell = 0$, the same conclusion, with p in place of p^2, follows from $\|N\|_{\mathrm{op}} \le p$ and (4.43). The case $k = 1$, equivalently $j = 2$, consists only of this $\ell = 0$ term. Multiplication by the remaining factor $\|h\|$ in (6.19), followed by summing at most k terms, therefore yields

$$\left|\sum_i \sqrt{w_i}\mathcal{T}_i^{(j)}[h^{\otimes j}]\right| \le Cp^2 j! \rho_p^{-(j-1)} \|h\|^j \le p^2 (Cj)^{Cj} \rho_p^{-(j-1)} \|h\|^j . \tag{6.21}$$

Thus (5.7)–(5.9) hold with $L = p^2 (Cj)^{Cj} \rho_p^{-(j-1)}$. Finally,

$$H_j(0) = \sum_i \sqrt{w_i} (k_i^{\mathsf{T}} h)^2 \mathcal{T}_i^{(j)}[h^{\otimes j}].$$

Theorem 5.1 proves (6.13). Equation (6.14) follows because $\rho_p^{-3} = p^{O(1)}$. □

7. Pathwise localization and logarithmic summation. Building on the fixed-order localization in [15, Sections 3.1–3.2], we combine the analytic jet bounds with a single Gaussian event that is uniform over the path and the logarithmic truncation order. We retain the normalization (3.1).

7.1. *A uniform path event.* At the base point, let k_i be the unit row directions in (5.4). Then

$$\zeta_{0,i} = k_i^{\mathsf{T}} h. \tag{7.1}$$

For $\varepsilon \in (0, 1/4)$, define

$$\Lambda_\varepsilon = 1 + \log \frac{2(m+d+1)}{\varepsilon}. \tag{7.2}$$

LEMMA 7.1 (Initial Gaussian event). *There is a universal $C > 0$ such that the event*

$$\mathcal{E}_0 = \left\{ \|h\| \le C\sqrt{d\Lambda_\varepsilon}, \quad \|\zeta_0\|_\infty \le C\sqrt{\Lambda_\varepsilon} \right\} \tag{7.3}$$

has probability at least $1 - \varepsilon/8$.

PROOF. Gaussian concentration [1, 24] gives

$$\mathbb{P}\Big(\|h\| > \sqrt{d} + \sqrt{2\log(16/\varepsilon)} \Big) \le \frac{\varepsilon}{16}.$$

Each $k_i^{\mathsf{T}} h$ is standard normal, so the normal tail bound and a union bound give

$$\mathbb{P}\left(\max_{1 \le i \le m} \left| k_i^{\mathsf{T}} h \right| > \sqrt{2\log(32m/\varepsilon)} \right) \le \frac{\varepsilon}{16}.$$

Both displayed thresholds are bounded by those in (7.3) for a sufficiently large universal C, since $\Lambda_\varepsilon \ge 1$. The union bound proves the claim. □

LEMMA 7.2 (Identification on a real proposal segment). *Work under (3.1), let $h \neq 0$, $\lambda = \|h\|$, $e = h/\lambda$, and $\tau = t\lambda$. On $\mathcal{E}_0$, if*

$$0 \le t \le r/\sqrt{d}, \qquad r \le c\rho_p \Lambda_\varepsilon^{-1/2},$$

then, after decreasing the universal constant $c > 0$,

$$0 \le \tau < \min\{\rho_p, 1/2\}.$$

Every slack along $x + \tau e$ is positive, the physical horizontal Lewis frame exists on this segment, and it is the restriction of the holomorphic solution in Lemma 4.6. In particular,

$$C^{-1} w_{x,i} \le w_{x+th,i} \le C w_{x,i}, \qquad 1 \le i \le m. \tag{7.4}$$

PROOF. On $\mathcal{E}_0$, $0 \le \tau \le Cr\sqrt{\Lambda_\varepsilon}$, so the stated radius condition places the segment in $[0, \rho_p) \cap [0, 1/2)$. By (4.10), $|(A_x e)_i| \le 1$, and hence

$$(A(x + \tau e) - b)_i = (Ax - b)_i \big(1 + \tau (A_x e)_i\big) > 0.$$

Thus the segment lies in $\operatorname{int} K$. The physical horizontal frame exists there and, by Proposition 4.4, satisfies the same real ODE and initial condition as the restriction of the holomorphic frame. Local uniqueness identifies them near zero. If their maximal interval of equality ended earlier, (4.30) would keep the common solution in the row-map neighborhood, and local uniqueness at the endpoint would extend it, a contradiction. Lemma 4.2 and (4.2) now give (7.4). □

LEMMA 7.3 (Uniform path bounds). *There are universal constants $c, C > 0$ with the following property. Let*

$$\Gamma_\varepsilon = C_\Gamma p^2 \Lambda_\varepsilon^2. \tag{7.5}$$

Here $C_\Gamma \geq 1$ is a fixed universal constant chosen large enough for all estimates in this lemma. If

$$r \leq c\rho_p \Lambda_\varepsilon^{-1/2}, \tag{7.6}$$

then, on $\mathcal{E}_0$, for every $0 \leq t \leq \eta = r/\sqrt{d}$, the point $x_t = x + th$ belongs to $\operatorname{int} K$ *and*

$$\|\zeta_t\|_\infty + \|s_t\|_\infty \leq \Gamma_\varepsilon, \tag{7.7}$$

$$\|u_t\| + \|v_t\| + \|q_t\| + \left\|q_t'\right\| + \left\|v_t'\right\| \leq \Gamma_\varepsilon \sqrt{d}, \tag{7.8}$$

$$\left|E_1'(t)\right| \leq \Gamma_\varepsilon d, \tag{7.9}$$

$$\|h\|_{g_0(x_t)} \leq \Gamma_\varepsilon \sqrt{d}. \tag{7.10}$$

For every integer $k \geq 1$,

$$\left\|N_t^{(k)}\right\|_{\mathrm{op}} \leq Cp^2 k! \rho_p^{-k} \Gamma_\varepsilon^k d^{k/2}. \tag{7.11}$$

Consequently, for $j \geq 2$,

$$|G_j(t)| \leq Cp^2 \Gamma_\varepsilon^2 (j-2)! \left(\frac{\Gamma_\varepsilon}{\rho_p}\right)^{j-2} d^{j/2}, \tag{7.12}$$

$$|H_j(t)| \leq Cp^2 \Gamma_\varepsilon^2 (j-1)! \left(\frac{\Gamma_\varepsilon}{\rho_p}\right)^{j-1} d^{(j+1)/2}. \tag{7.13}$$

PROOF. If $h = 0$, every assertion is immediate. Otherwise put $\lambda = \|h\|$, $e = h/\lambda$, and $\tau = t\lambda$. Lemma 7.2 gives $x_t \in \operatorname{int} K$ and (7.4) throughout the required interval. We use this identification below.

For the original direction h, the exact formula is

$$s_{t,i} = \frac{s_{0,i}}{1 + ts_{0,i}}, \qquad s_{0,i} = w_{0,i}^{1/p} \zeta_{0,i}. \tag{7.14}$$

The event $\mathcal{E}_0$ gives $\|s_0\|_\infty \leq C\sqrt{\Lambda_\varepsilon}$. Decreasing c if necessary, $|ts_{0,i}| \leq 1/2$. Hence $\|s_t\|_\infty \leq C\sqrt{\Lambda_\varepsilon}$. Also,

$$\zeta_{t,i} = \zeta_{0,i} \left(\frac{w_{t,i}}{w_{0,i}}\right)^{-1/p} (1 + ts_{0,i})^{-1}, \tag{7.15}$$

so (7.4) proves (7.7).

The identities

$$u_{t,i} = \sqrt{w_{t,i}}\zeta_{t,i}, \qquad v_t = W_t^{1/p} u_t, \qquad q_{t,i} = \sqrt{w_{t,i}}\zeta_{t,i}^2 \tag{7.16}$$

and $\sum_i w_{t,i} = d$ imply

$$\|u_t\| + \|v_t\| + \|q_t\| \leq C\Gamma_\varepsilon \sqrt{d}. \tag{7.17}$$

The derivative formulas are

$$v_t' = -s_t \circ v_t - \frac{1}{2}(N_t v_t) \circ s_t, \tag{7.18}$$

$$q_t' = \left(\left(2\beta - \frac{1}{2}\right)\ell_t' - 2s_t\right) \circ q_t, \qquad \ell_t' = -W_t^{-1/2} N_t v_t. \tag{7.19}$$

At the real point x_t, $0 \preceq N_t \preceq pI_m$, so

$$\|N_t v_t\| \le p\,\|v_t\|. \tag{7.20}$$

For q_t', use the exact cancellation

$$\ell_t' \circ q_t = -\left(W_t^{-1/2} N_t v_t\right) \circ \left(W_t^{-1/2} u_t^{\circ 2}\right) = -(N_t v_t) \circ \zeta_t^{\circ 2}. \tag{7.21}$$

Indeed, with $R = C\sqrt{\Lambda_\varepsilon}$ denoting a common bound for $\|s_t\|_\infty$ and $\|\zeta_t\|_\infty$,

$$\left\|v_t'\right\| \le R\left(\|v_t\| + \tfrac{1}{2}\|N_t v_t\|\right) \le CpR^2\sqrt{d},$$

and the sharper identity (7.21) gives

$$\left\|q_t'\right\| \le C\left(\|\zeta_t\|_\infty^2 \|N_t v_t\| + \|s_t\|_\infty \|q_t\|\right) \le CpR^3\sqrt{d}.$$

Together with (7.17), these inequalities prove (7.8) after fixing the stated universal C_Γ sufficiently large.

Differentiate $E_1(t) = \sum_i w_{t,i}^\alpha s_{t,i}^3$:

$$E_1'(t) = \alpha \sum_i w_{t,i}^\alpha \ell_{t,i}' s_{t,i}^3 - 3\sum_i w_{t,i}^\alpha s_{t,i}^4. \tag{7.22}$$

The second sum is at most $3\|s_t\|_\infty^2 \|u_t\|^2$. For the first, Cauchy–Schwarz and $W_t^{1/2}\ell_t' = -N_t v_t$ give

$$\begin{aligned}\sum_i w_i^\alpha \left|\ell_i'\right| |s_i|^3 &\le \left(\sum_i w_i^\alpha s_i^2\right)^{1/2} \left(\sum_i w_i^\alpha (\ell_i')^2 s_i^4\right)^{1/2} \\ &\le \|u_t\|\,\|\zeta_t\|_\infty \|s_t\|_\infty \|N_t v_t\|.\end{aligned}$$

This proves (7.9). Since $F(t) = \|u_t\|^2 = \|h\|_{g_0(x_t)}^2$, (7.10) follows as well.

Apply (4.42) at x_t in the direction h and use (7.10) to obtain (7.11). Inserting (7.8) and (7.11) into the definitions of G_j and H_j proves (7.12) and (7.13). More explicitly, for $j \ge 3$ the two factors outside $N_t^{(j-2)}$ in G_j contribute at most $\Gamma_\varepsilon^2 d$, and hence

$$|G_j(t)| \le Cp^2\Gamma_\varepsilon^2 d\,(j-2)!\rho_p^{-(j-2)}\Gamma_\varepsilon^{j-2} d^{(j-2)/2},$$

which is (7.12). When $j = 2$, the same conclusion follows directly from $\|N_t\|_{\mathrm{op}} \le p$, so no zeroth-order use of (7.11) is required. Similarly, the two exterior factors in H_j contribute $\Gamma_\varepsilon^2 d$, while $N_t^{(j-1)}$ contributes

$$Cp^2(j-1)!\rho_p^{-(j-1)}\Gamma_\varepsilon^{j-1} d^{(j-1)/2}.$$

Their product is exactly the right side of (7.13); in particular its dimension power is $d\,d^{(j-1)/2} = d^{(j+1)/2}$. □

7.2. *Summing the expansion.* The parameter choice below keeps localization valid, bounds each low-order contribution by $O(\varepsilon)$, absorbs the terminal $\sqrt{d}$ factor, and retains an inverse-polylogarithmic radius.

LEMMA 7.4 (Compatible choice of truncation parameters). *Fix once and for all the constants in all preceding estimates. Let $c_{\mathrm{path}} \in (0,1]$ be a valid constant in (7.6), and write $\rho_p = c_\rho p^{-3}$ with $c_\rho \in (0,1]$. There are universal constants*

$$C_{\mathrm{cl}}, C_2, C_Q, C_J, C_S, A_0, B_0 \ge 1, \qquad c_S, c_{\mathrm{low}} > 0,$$

which can be fixed in the displayed order

$$C_{\text{cl}}\ ;\ C_2\ ;\ C_Q\ ;\ C_J\ ;\ C_S\ ;\ c_S\ ;\ (A_0, B_0, c_{\text{low}}), \tag{7.23}$$

such that the following holds. Define

$$Q_\varepsilon = C_Q p^4 \Gamma_\varepsilon^4, \tag{7.24}$$

$$J = \left\lceil C_J \log\left(\frac{2(d+1)Q_\varepsilon}{\varepsilon}\right)\right\rceil, \tag{7.25}$$

$$r_{\text{cand}} = c_S \varepsilon^3 \rho_p Q_\varepsilon^{-1} J^{-C_S}. \tag{7.26}$$

For every $0 < r \le r_{\text{cand}}$*, if*

$$\theta = \frac{r\Gamma_\varepsilon}{\rho_p}, \qquad \xi = (C_{\text{cl}} J)^{C_2} \frac{r}{\rho_p},$$

then

$$r \le c_{\text{path}} \rho_p \Lambda_\varepsilon^{-1/2}, \qquad \theta \le J^{-4}, \qquad \xi \le \frac{1}{4}, \tag{7.27}$$

$$C_{\text{cl}} p^4 r^2 \varepsilon^{-2} \le \frac{\varepsilon}{8}, \qquad \frac{1}{2}\Gamma_\varepsilon r^2 \le \frac{\varepsilon}{20}, \qquad 2p^2 \rho_p \xi^2 \le \frac{\varepsilon^2}{160}, \tag{7.28}$$

$$C_{\text{cl}} p^2 \Gamma_\varepsilon^2 r^2 \le \frac{\varepsilon}{40}, \qquad C_{\text{cl}} p^2 \Gamma_\varepsilon^2 r \sqrt{d}\, \theta^J \le \frac{\varepsilon}{40}. \tag{7.29}$$

Moreover,

$$r_{\text{cand}} \ge c_{\text{low}} \varepsilon^3 p^{-A_0} \Lambda_\varepsilon^{-B_0}. \tag{7.30}$$

PROOF. Choose C_{cl} to dominate every universal coefficient already fixed in the cubic, jet, Wick and path estimates used below, including the fixed base constant inside $(Cj)^{Cj}$ in (6.13). Next choose C_2 so that, for all integers $J \ge j \ge 2$,

$$\frac{(Cj)^{Cj}}{j!}(1+J)^{j/2} \le (C_{\text{cl}} J)^{C_2 j}, \tag{7.31}$$

where the C on the left is the now fixed constant from (6.13). This is possible by Stirling's inequality. Choose C_Q so large that

$$Q_\varepsilon \ge C_{\text{cl}} \max\{1, p^4, p^2\Gamma_\varepsilon^2, \Gamma_\varepsilon\} \tag{7.32}$$

for every $p \ge 4$ and $\Lambda_\varepsilon \ge 1$. Next choose $C_J \ge 1$ for the terminal estimate below, choose $C_S \ge C_2 + 4$, and only then choose $c_S > 0$ sufficiently small in terms of these fixed universal constants. This is the order in (7.23).

Since $\Gamma_\varepsilon = C_\Gamma p^2 \Lambda_\varepsilon^2$,

$$Q_\varepsilon = C_Q C_\Gamma^4 p^{12} \Lambda_\varepsilon^8.$$

In particular, with

$$X := \frac{2(d+1)Q_\varepsilon}{\varepsilon},$$

one has $X \ge 16$. Thus the ceiling in (7.25) defines an integer $J \ge 3$, and

$$C_J \log X \le J \le C_J \log X + 1 \le 2C_J \log X. \tag{7.33}$$

For $0 < r \le r_{\mathrm{cand}}$,

$$\frac{r\Lambda_\varepsilon^{1/2}}{\rho_p} \le c_S\varepsilon^3 Q_\varepsilon^{-1} J^{-C_S}\Lambda_\varepsilon^{1/2} \le Cc_S\Lambda_\varepsilon^{-15/2},$$

$$\theta \le c_S\varepsilon^3\frac{\Gamma_\varepsilon}{Q_\varepsilon}J^{-C_S} \le c_S J^{-C_S},$$

$$\xi \le c_S C_{\mathrm{cl}}^{C_2}\varepsilon^3 Q_\varepsilon^{-1} J^{C_2-C_S}.$$

Thus a sufficiently small c_S, chosen after C_S, gives all three inequalities in (7.27); in particular $r \le \rho_p \le 1$. The remaining low-order estimates follow from

$$C_{\mathrm{cl}}p^4r^2\varepsilon^{-2} \le C_{\mathrm{cl}}c_S^2\varepsilon^4\frac{p^4\rho_p^2}{Q_\varepsilon^2},$$

$$\frac{1}{2}\Gamma_\varepsilon r^2 \le \frac{1}{2}c_S^2\varepsilon^6\frac{\rho_p^2\Gamma_\varepsilon}{Q_\varepsilon^2},$$

$$2p^2\rho_p\xi^2 \le 2c_S^2C_{\mathrm{cl}}^{2C_2}\varepsilon^6\frac{p^2\rho_p}{Q_\varepsilon^2}J^{2C_2-2C_S},$$

$$C_{\mathrm{cl}}p^2\Gamma_\varepsilon^2r^2 \le C_{\mathrm{cl}}c_S^2\varepsilon^6\frac{p^2\rho_p^2\Gamma_\varepsilon^2}{Q_\varepsilon^2}.$$

Here $p^4\rho_p^2 \le 1$, $p^2\rho_p \le 1$, and every remaining displayed ratio is at most one by (7.32). Since $\varepsilon^4 \le \varepsilon$, $\varepsilon^6 \le \varepsilon^2$, and $\varepsilon^2 \le \varepsilon$ for $\varepsilon \in (0, 1/4)$, decreasing c_S once proves (7.28) and the first inequality in (7.29).

For the last terminal inequality, retain the preceding value of X. Then $Q_\varepsilon\sqrt{d} \le \varepsilon X/2$, and (7.32), $r \le 1$, and $\theta \le J^{-4}$ give

$$C_{\mathrm{cl}}p^2\Gamma_\varepsilon^2 r\sqrt{d}\,\theta^J \le Q_\varepsilon\sqrt{d}\,J^{-4J}.$$

Choose C_J sufficiently large that the integer in (7.25) satisfies $J \ge e$; the inequality $J \ge C_J \log X$ is automatic from (7.33). Then

$$J^{-4J} \le e^{-4J} \le X^{-4C_J}, \qquad Q_\varepsilon\sqrt{d}\,J^{-4J} \le \frac{\varepsilon}{2}X^{1-4C_J} \le \frac{\varepsilon}{40}.$$

This choice of C_J precedes the choices of C_S and c_S, as required.

Finally,

$$\log X \le C\big(\Lambda_\varepsilon + \log p + \log\Lambda_\varepsilon\big) \le Cp\Lambda_\varepsilon,$$

and hence (7.33) gives $J \le Cp\Lambda_\varepsilon$. Since $Q_\varepsilon \le Cp^{12}\Lambda_\varepsilon^8$ and $\rho_p = c_\rho p^{-3}$, substitution in (7.26) gives

$$r_{\mathrm{cand}} \ge c_{\mathrm{low}}\varepsilon^3 p^{-(15+C_S)}\Lambda_\varepsilon^{-(8+C_S)}.$$

Thus one may set $A_0 = 15 + C_S$ and $B_0 = 8 + C_S$, with $c_{\mathrm{low}} > 0$ depending only on the constants already fixed in (7.23). □

PROOF OF THEOREM 1.2. Fix $\varepsilon \in (0, 1/4)$ and the quantities in Lemma 7.4. After decreasing c_{asc} and increasing C_{asc} in (1.5), (7.30) gives $\underline{r}_\varepsilon \le r_{\mathrm{cand}}$. Fix an arbitrary $0 < r \le \underline{r}_\varepsilon$, set $\eta = r/\sqrt{d}$, and put

$$\theta = \frac{r\Gamma_\varepsilon}{\rho_p}. \tag{7.34}$$

Lemma 7.4 supplies all smallness conditions below, including (7.6). By (3.4), Lemma 3.1, and one integration of E_1,

$$
\begin{aligned}
\frac{1}{2}\left|F(\eta)-F(0)\right| \le \eta\left|E_1(0)+\beta H_1(0)\right| &+ \int_0^{\eta}(\eta-t)\left|E_1'(t)\right|\,\mathrm{d}t \\
&+ \beta\sum_{j=2}^{J}\frac{\eta^j}{j!}\left|H_j(0)\right| \\
&+ \beta\sum_{j=2}^{J}\int_0^{\eta}\frac{(\eta-t)^{j-1}}{(j-1)!}\left|G_j(t)\right|\,\mathrm{d}t \\
&+ \beta\int_0^{\eta}\frac{(\eta-t)^J}{J!}\big(\left|G_{J+1}(t)\right|+\left|H_{J+1}(t)\right|\big)\,\mathrm{d}t. \qquad (7.35)
\end{aligned}
$$

By Lemma 6.1 and Chebyshev's inequality,

$$
\mathbb{P}\Big(\eta\left|E_1(0)+\beta H_1(0)\right| > \frac{\varepsilon}{20}\Big) \le C_{\mathrm{cl}}p^4\frac{r^2}{\varepsilon^2} \le \frac{\varepsilon}{8}. \qquad (7.36)
$$

The final inequality is part of (7.28).

On $\mathcal{E}_0$, Lemma 7.3 gives

$$
\int_0^{\eta}(\eta-t)\left|E_1'(t)\right|\,\mathrm{d}t \le \frac{1}{2}\eta^2\Gamma_\varepsilon d = \frac{1}{2}\Gamma_\varepsilon r^2 \le \frac{\varepsilon}{20}. \qquad (7.37)
$$

Let

$$
S_J = \sum_{j=2}^{J}\frac{\eta^j}{j!}\left|H_j(0)\right|.
$$

Theorem 6.2 yields

$$
\mathbb{E}S_J \le p^2\rho_p\sum_{j=2}^{J}\frac{(Cj)^{Cj}}{j!}\left(\frac{r}{\rho_p}\right)^j\left(1+\frac{j}{d}\right)^{j/2}. \qquad (7.38)
$$

For $j \le J$ and $d \ge 1$, the last factor is at most $(1+J)^{j/2}$. The named Stirling estimate (7.31) therefore bounds the factor multiplying $(r/\rho_p)^j$ by $(C_{\mathrm{cl}}J)^{C_2 j}$. Put $\xi = (C_{\mathrm{cl}}J)^{C_2}r/\rho_p$. By (7.27), $\xi \le 1/4$. Consequently,

$$
\mathbb{E}S_J \le p^2\rho_p\sum_{j=2}^{\infty}\xi^j \le 2p^2\rho_p\xi^2 \le \frac{\varepsilon^2}{160}
$$

by (7.28). Thus

$$
\mathbb{E}S_J \le \frac{\varepsilon^2}{160}. \qquad (7.39)
$$

Markov's inequality implies

$$
\mathbb{P}\Big(S_J > \frac{\varepsilon}{20}\Big) \le \frac{\varepsilon}{8}. \qquad (7.40)
$$

On $\mathcal{E}_0$, (7.12) and $\int_0^{\eta}(\eta-t)^{j-1}\,\mathrm{d}t = \eta^j/j$ give

$$
\sum_{j=2}^{J}\int_0^{\eta}\frac{(\eta-t)^{j-1}}{(j-1)!}\left|G_j(t)\right|\,\mathrm{d}t
$$

$$\leq C_{\mathrm{cl}}p^2\Gamma_\varepsilon^2 r^2 \sum_{j=2}^{J} \frac{\theta^{j-2}}{j(j-1)} \leq C_{\mathrm{cl}}p^2\Gamma_\varepsilon^2 r^2 \leq \frac{\varepsilon}{20}. \tag{7.41}$$

The terminal G_{J+1} term is at most

$$C_{\mathrm{cl}}p^2\Gamma_\varepsilon^2 r^2 \frac{\theta^{J-1}}{J(J+1)}. \tag{7.42}$$

Indeed, integrating the order-$J+1$ instance of (7.12) gives

$$\frac{\eta^{J+1}}{(J+1)!} Cp^2\Gamma_\varepsilon^2 (J-1)! \left(\frac{\Gamma_\varepsilon}{\rho_p}\right)^{J-1} d^{(J+1)/2},$$

and $\eta = r/\sqrt{d}$ turns this expression into (7.42); all powers of d cancel. The first inequality in (7.29), together with $\theta \leq J^{-4} \leq 1$ and $J \geq 3$, therefore gives

$$C_{\mathrm{cl}}p^2\Gamma_\varepsilon^2 r^2 \frac{\theta^{J-1}}{J(J+1)} \leq C_{\mathrm{cl}}p^2\Gamma_\varepsilon^2 r^2 \leq \frac{\varepsilon}{40}. \tag{7.43}$$

For the terminal H_{J+1} term, (7.13) gives

$$\int_0^\eta \frac{(\eta-t)^J}{J!} |H_{J+1}(t)| \,\mathrm{d}t \leq C_{\mathrm{cl}}p^2\Gamma_\varepsilon^2 \frac{r\sqrt{d}}{J+1}\theta^J. \tag{7.44}$$

Here the corresponding unsimplified bound is

$$\frac{\eta^{J+1}}{(J+1)!} Cp^2\Gamma_\varepsilon^2 J! \left(\frac{\Gamma_\varepsilon}{\rho_p}\right)^{J} d^{(J+2)/2}.$$

Thus exactly one factor $\sqrt{d}$ remains: $d^{(J+2)/2}d^{-(J+1)/2} = \sqrt{d}$. The terminal estimate therefore retains exactly this factor. The logarithmic choice of J absorbs it through θ^J, as quantified in (7.29). The second inequality in (7.29) gives

$$C_{\mathrm{cl}}p^2\Gamma_\varepsilon^2 r\sqrt{d}\,\theta^J \leq \frac{\varepsilon}{40}. \tag{7.45}$$

Since $(J+1)^{-1} \leq 1$, (7.44) and (7.45) show separately that the terminal H_{J+1} term is at most $\varepsilon/40$. Together with (7.43), this proves that the entire terminal remainder is at most $\varepsilon/20$.

The failures of $\mathcal{E}_0$, (7.36), and (7.40) have total probability at most

$$\frac{\varepsilon}{8} + \frac{\varepsilon}{8} + \frac{\varepsilon}{8} = \frac{3\varepsilon}{8} \leq \varepsilon.$$

On their complement, each of the five contributions on the right side of (7.35) is at most $\varepsilon/20$, and $0 < \beta < 1$. Hence $\frac{1}{2}|F(\eta) - F(0)| \leq \varepsilon/4$, and therefore

$$|F(\eta) - F(0)| \leq \frac{\varepsilon}{2} \leq 2\varepsilon. \tag{7.46}$$

If $z = x + \eta h$, then

$$\|z-x\|_{g_0(z)}^2 - \|z-x\|_{g_0(x)}^2 = \eta^2(F(\eta) - F(0)).$$

Since $\eta^2 = r^2/d$, (7.46) gives the stronger two-sided estimate

$$\left| \|z-x\|_{g_0(z)}^2 - \|z-x\|_{g_0(x)}^2 \right| \leq \frac{\varepsilon r^2}{2d} \leq \frac{2\varepsilon r^2}{d}.$$

Together with $z \in \operatorname{int} K$, which is guaranteed by the path event, this implies the event in (1.4); in particular, it implies the one-sided ASC condition used by the mixing framework. The base point was arbitrary, and Lemma 2.7 transfers both the discrepancy event and its probability back from the normalization (3.1). The same argument holds with any $0 < s \le r$ in place of r. Hence the estimate is uniform in x and s, and the success probability is at least $1 - 3\varepsilon/8 \ge 1 - \varepsilon$, proving Theorem 1.2. □

REMARK 7.5 (The fourth-order loss). At order four, a pathwise Cauchy–Schwarz estimate gives

$$|H_4(t)| \le \|q_t\| \left\| N_t^{(3)} \right\|_{\mathrm{op}} \|v_t\| \lesssim d^{5/2}.$$

Theorem 6.2 instead gives $\|H_4(0)\|_{L^2} \le p^{O(1)} d^2$. The expansion uses the base-point estimate at each retained order and postpones the pathwise bound to order J, where θ^J absorbs the residual $d^{1/2}$ factor.

8. From average self-concordance to mixing. We now rescale the baseline ASC radius into a dimension-free modulus and insert it, together with the deterministic geometry of Section 2, into the abstract mixing reduction.

LEMMA 8.1 (Dimension-free ASC modulus after scaling). *With $p = p_m$ and $L_{m,d}$ from (1.6)–(1.7), the family of metrics $g_\star = L_{m,d} g_0$ has an ASC modulus $\mathfrak{a}_\star : (0,1) \to (0,\infty)$ independent of m and d. One may take*

$$\mathfrak{a}_\star(\varepsilon) = c_\star \begin{cases} \varepsilon^3 (1 + \log(1/\varepsilon))^{-C_{\mathrm{asc}}}, & 0 < \varepsilon \le \varepsilon_0, \\ 1, & \varepsilon_0 < \varepsilon < 1. \end{cases}$$

PROOF. For $0 < \varepsilon \le \varepsilon_0$, put $B = 1 + \log(2(m+d+1)/\varepsilon_0) \ge 1$. Then

$$\frac{\underline{r}_\varepsilon}{\underline{r}_{\varepsilon_0}} = \left(\frac{\varepsilon}{\varepsilon_0}\right)^3 \left(\frac{B}{B + \log(\varepsilon_0/\varepsilon)}\right)^{C_{\mathrm{asc}}} \ge c\varepsilon^3 (1 + \log(1/\varepsilon))^{-C_{\mathrm{asc}}}.$$

Proposition 2.6, Theorem 1.2, and $\widehat{r}_0 \le \underline{r}_{\varepsilon_0}$ give

$$R_\varepsilon(g_\star) \ge L_{m,d}^{1/2} \underline{r}_\varepsilon \ge \frac{\underline{r}_\varepsilon}{\widehat{r}_0} \ge \frac{\underline{r}_\varepsilon}{\underline{r}_{\varepsilon_0}} \ge c\varepsilon^3 (1 + \log(1/\varepsilon))^{-C_{\mathrm{asc}}},$$

which proves the first case. For $\varepsilon > \varepsilon_0$, the event and the required success probability are weaker than at ε_0, and explicitly

$$R_\varepsilon(g_\star) \ge R_{\varepsilon_0}(g_\star) \ge L_{m,d}^{1/2} \underline{r}_{\varepsilon_0} \ge 1.$$

□

PROOF OF THEOREM 1.3. The choice $p = p_m$ satisfies the hypothesis of Proposition 2.5. Thus $H_{p_m} = \lambda_m g_0$ is strongly self-concordant and lower-trace self-concordant, with symmetry parameter at most $e\lambda_m d$. Since

$$g_\star = L_{m,d} g_0 = \widehat{r}_0^{\,-2} H_{p_m} \quad \text{and} \quad \widehat{r}_0^{\,-2} \ge 1,$$

Proposition 2.6 shows that $g_\star$ has the same two differential regularity properties and symmetry parameter

$$\bar{\nu}(g_\star) \le e\lambda_m d \widehat{r}_0^{\,-2} = edL_{m,d}. \tag{8.1}$$

Lemma 8.1 supplies a dimension-free ASC modulus. Hence the constant a_0 in Proposition 2.4 is universal, while (8.1) gives

$$d\bar{\nu}(g_\star) \le ed^2 L_{m,d}.$$

Proposition 2.4 therefore yields an L^2 spectral gap of order $(d^2 L_{m,d})^{-1}$ and the iteration bound (1.8). Its proposal radius depends only on universal constants.

Finally, $p_m = O(1+\log m)$, $\lambda_m = p_m^{O(1)}$, and Theorem 1.2 with fixed ε_0 gives $\widehat{r}_0^{\,-1} = \mathrm{polylog}(m,d)$. Hence $L_{m,d} = \mathrm{polylog}(m,d)$. $\square$

COROLLARY 8.2 (Uniform sampling). *Let π be the uniform distribution on a bounded full-dimensional polytope, and suppose that $\mu_0 \ll \pi$ satisfies*

$$\left\| \frac{\mathrm{d}\mu_0}{\mathrm{d}\pi} \right\|_\infty \le M$$

for some $M \ge 1$. For every $\delta \in (0,1)$, run the Dikin walk with the scaled metric $g_\star$ for

$$T \ge C d^2 L_{m,d} \log\left(\frac{M}{\delta}\right). \tag{8.2}$$

Then its law satisfies $\chi^2(\mu_T \| \pi) \le \delta$. Thus its warm-start iteration complexity is

$$\widetilde{O}\left(d^2 \log \frac{M}{\delta}\right).$$

PROOF. The warmness assumption gives

$$1 + \chi^2(\mu_0 \| \pi) = \int \left(\frac{\mathrm{d}\mu_0}{\mathrm{d}\pi}\right)^2 \mathrm{d}\pi \le M.$$

Apply Theorem 1.3 with $\vartheta = 0$. $\square$

REMARK 8.3 (Logarithmic factors). The powers of p, $\log m$, $\log d$ and ε^{-1} have not been optimized. The analytic argument gives the explicit local scale $\rho_p \asymp p^{-3}$; later estimates deliberately enlarge fixed polynomial powers because they do not affect the dimension exponent.

APPENDIX A: A COMPLEX ANALYTIC ODE LEMMA

LEMMA A.1. *Let $(X, \|\cdot\|)$ be a finite-dimensional complex Banach space. Suppose $F : B(x_0,\delta) \times D(0,\tau_0) \to X$ is holomorphic and*

$$\sup_{B(x_0,\delta)\times D(0,\tau_0)} \|F(x,t)\| \le M.$$

Then the initial-value problem

$$x'(t) = F(x(t),t), \qquad x(0) = x_0,$$

has a unique holomorphic solution on

$$|t| < 2r_0, \qquad r_0 = c\min\{\tau_0, \delta/M\}, \tag{A.1}$$

for a sufficiently small universal $c > 0$, and $\|x(t) - x_0\| \le \delta/2$ there. If H is holomorphic on $B(x_0,\delta)$ and $\sup_{B(x_0,\delta/2)} \|H - H(x_0)\| \le B$, then

$$\left\| \frac{\mathrm{d}^k}{\mathrm{d}t^k} H(x(t))\Big|_{t=0} \right\| \le B k! r_0^{-k}, \qquad k \ge 1. \tag{A.2}$$

PROOF. For $x \in B(x_0, 3\delta/4)$ and $\|h\| = 1$, the one-variable map

$$z \longmapsto F(x + zh, t), \qquad |z| < \delta/4,$$

takes values in the radius-M ball of X. The Banach-valued Cauchy formula therefore gives

$$\|D_x F(x,t)[h]\| \le 4M/\delta,$$

and hence $\|D_x F(x,t)\| \le 4M/\delta$. This argument uses no equivalence between the chosen norm and a Euclidean norm, so its constant is independent of $\dim X$.

Let $\mathcal{Y}$ be the complete metric space of holomorphic functions on $D(0, 2r_0)$, continuous on its closure, satisfying $\sup_{|t| \le 2r_0} \|x(t) - x_0\| \le \delta/2$. Define, with the integral taken along the line segment from 0 to t,

$$(\mathcal{R}x)(t) = x_0 + \int_0^t F(x(s), s)\, \mathrm{d}s.$$

Choose the universal constant in (A.1) so that $2r_0 < \tau_0$, $2r_0 M \le \delta/2$, and $8r_0 M/\delta \le 1/2$. The preceding derivative bound, integrated along the segment between $x(s)$ and $\widetilde{x}(s)$, shows that

$$\sup_{|t| \le 2r_0} \|\mathcal{R}x(t) - \mathcal{R}\widetilde{x}(t)\| \le \frac{1}{2} \sup_{|t| \le 2r_0} \|x(t) - \widetilde{x}(t)\|.$$

Thus $\mathcal{R}$ is invariant and contractive. Its fixed point is the unique holomorphic solution staying in the ball; ordinary local uniqueness then gives uniqueness for the initial-value problem on the whole disk. On $|t| \le r_0$, the sharper estimate $\|x(t) - x_0\| \le r_0 M \le \delta/4$ places the Cauchy circle inside $B(x_0, \delta/2)$. Applying the Banach-valued Cauchy formula to $H(x(t)) - H(x_0)$ proves (A.2). □

APPENDIX B: HORIZONTAL FRAMES

LEMMA B.1 (Horizontal gauge). *Let $E_t \subset \mathbb{R}^m$ be a smooth path of d-dimensional subspaces, and let $\widetilde{U}_t$ be a smooth orthonormal frame for E_t. There is a smooth orthogonal matrix $O_t \in \mathbb{R}^{d\times d}$ such that $U_t = \widetilde{U}_t O_t$ satisfies*

$$U_t^\mathsf{T} U_t = I_d, \qquad U_t^\mathsf{T} U_t' = 0.$$

If $P_t = U_t U_t^\mathsf{T}$ and a full-rank matrix $\widehat{B}_t$ with column space E_t satisfies $\widehat{B}_t' = Z_t \widehat{B}_t$, then

$$U_t' = (I - P_t) Z_t U_t.$$

PROOF. Set $S_t = \widetilde{U}_t^\mathsf{T} \widetilde{U}_t'$. Differentiating $\widetilde{U}_t^\mathsf{T} \widetilde{U}_t = I$ shows that S_t is skew-symmetric. Solve $O_t' = -S_t O_t$, $O_0 = I$. Then

$$(O_t^\mathsf{T} O_t)' = 0, \qquad U_t^\mathsf{T} U_t' = O_t^\mathsf{T} S_t O_t + O_t^\mathsf{T} O_t' = 0,$$

so O_t is orthogonal and U_t is horizontal.

Write $\widehat{B}_t = U_t M_t$ with M_t invertible. From

$$U_t' M_t + U_t M_t' = Z_t U_t M_t$$

left multiplication by $I - P_t$ and right multiplication by M_t^{-1} give

$$(I - P_t) U_t' = (I - P_t) Z_t U_t.$$

The horizontal condition gives $P_t U_t' = U_t (U_t^\mathsf{T} U_t') = 0$, proving the formula. □

APPENDIX C: GAUSSIAN TENSOR IDENTITIES

LEMMA C.1 (Raw monomials and Wick tensors). *For every symmetric n-tensor T and $h \sim N(0, I_d)$,*

$$T[h^{\otimes n}] = \sum_{r=0}^{\lfloor n/2 \rfloor} \frac{n!}{2^r r!(n-2r)!} I_{n-2r}(\mathrm{Tr}^r\, T).$$

Moreover, for symmetric tensors S, T of orders a, b,

$$I_a(S) I_b(T) = \sum_{\ell=0}^{a \wedge b} \ell! \binom{a}{\ell} \binom{b}{\ell} I_{a+b-2\ell}(S \widetilde{\otimes}_\ell T),$$

and

$$\mathbb{E}[I_a(S) I_b(T)] = \mathbf{1}_{\{a=b\}} a! \langle S, T \rangle_{\mathrm{HS}}.$$

Here $I_0(c) = c$, and a summand in the product formula is present only for $0 \le \ell \le \min\{a, b\}$. In particular, an order-zero output is a scalar and its Hilbert–Schmidt norm is its absolute value.

PROOF. The product and isometry formulas are the standard identities for multiple Wiener integrals [32, Chapter 1]. The symmetrized contraction in the product formula is the orthogonal projection of the unsymmetrized contraction onto the symmetric tensor subspace, so

$$\left\| S \widetilde{\otimes}_\ell T \right\|_{\mathrm{HS}} \le \| S \otimes_\ell T \|_{\mathrm{HS}}.$$

This observation also applies when the output order is zero.

The raw-to-Wick formula follows by induction on n. Multiplication by $h_j = I_1(e_j)$ either raises the chaos order by one or contracts e_j with one existing index. A term with exactly r contractions is indexed by a choice of r disjoint unordered pairs among the n tensor slots. Their number is

$$\binom{n}{2r} \frac{(2r)!}{2^r r!} = \frac{n!}{2^r r!(n-2r)!},$$

which gives the asserted coefficient. Distinct values of r produce distinct chaos orders $n - 2r$, and hence are orthogonal by the isometry formula. This last orthogonality statement concerns the raw-to-Wick expansion itself; multiplication by an additional Wick polynomial can make different contraction patterns land in the same output chaos. □

LEMMA C.2 (Gaussian radial moment). *For $h \sim N(0, I_d)$ and every integer $n \ge 1$,*

$$\mathbb{E} \|h\|^{2n} \le (C(d+n))^n. \tag{C.1}$$

PROOF. The variable $\|h\|^2$ has the χ^2_d distribution, and

$$\mathbb{E} \|h\|^{2n} = 2^n \frac{\Gamma(n + d/2)}{\Gamma(d/2)} = \prod_{j=0}^{n-1} (d + 2j) \le (d + 2n)^n.$$

□

REFERENCES

[1] BORELL, C. (1975). The Brunn–Minkowski inequality in Gauss space. *Invent. Math.* **30**, 207–216.

[2] BRASCAMP, H. J. and LIEB, E. H. (1976). On extensions of the Brunn–Minkowski and Prékopa–Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation. *J. Funct. Anal.* **22**(4), 366–389.

[3] CHEN, Y., DWIVEDI, R., WAINWRIGHT, M. J. and YU, B. (2018). Fast MCMC sampling algorithms on polytopes. *J. Mach. Learn. Res.* **19**(55), 1–86.

[4] COHEN, M. B. and PENG, R. (2015). ℓ_p row sampling by Lewis weights. In *Proceedings of the 47th Annual ACM Symposium on Theory of Computing*, 183–192. ACM.

[5] COVER, T. M. and THOMAS, J. A. (2006). *Elements of Information Theory*, second ed. Wiley, Hoboken, NJ.

[6] FAZEL, M., LEE, Y. T., PADMANABHAN, S. and SIDFORD, A. (2022). Computing Lewis weights to high precision. In *Proceedings of the 2022 Annual ACM–SIAM Symposium on Discrete Algorithms*, 2723–2742. SIAM.

[7] GATMIRY, K., KELNER, J. and VEMPALA, S. S. (2024). Sampling polytopes with Riemannian HMC: Faster mixing via the Lewis weights barrier. In *Proceedings of the Thirty Seventh Conference on Learning Theory*, *Proc. Mach. Learn. Res.* **247**, 1796–1881.

[8] GU, Y., KUANG, N. L., MA, Y.-A., SONG, Z. and ZHANG, L. (2024). Log-concave sampling from a convex body with a barrier: A robust and unified Dikin walk. In *Advances in Neural Information Processing Systems* **37**, 69230–69298.

[9] GUSTAFSON, A. and NARAYANAN, H. (2023). John's walk. *Adv. in Appl. Probab.* **55**(2), 473–491.

[10] HASTINGS, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57**(1), 97–109.

[11] HORN, R. A. and JOHNSON, C. R. (2013). *Matrix Analysis*, second ed. Cambridge Univ. Press, Cambridge.

[12] JANSON, S. (1997). *Gaussian Hilbert Spaces*. Cambridge Tracts in Mathematics **129**. Cambridge Univ. Press, Cambridge.

[13] JIANG, M. and CHEN, Y. (2025). Regularized Dikin walks for sampling truncated logconcave measures, mixed isoperimetry and beyond worst-case analysis. In *Proceedings of the Thirty Eighth Conference on Learning Theory*, *Proc. Mach. Learn. Res.* **291**, 3017–3078.

[14] KANNAN, R. and NARAYANAN, H. (2012). Random walks on polytopes and an affine interior point method for linear programming. *Math. Oper. Res.* **37**(1), 1–20.

[15] KOOK, Y. (2026). Beyond the $d^{2.5}$-mixing bound for Dikin walks on polytopes. arXiv preprint arXiv:2607.13943v1.

[16] KOOK, Y. and VEMPALA, S. S. (2024). Gaussian cooling and Dikin walks: The interior-point method for logconcave sampling. In *Proceedings of the Thirty Seventh Conference on Learning Theory*, *Proc. Mach. Learn. Res.* **247**, 3137–3240.

[17] KOOK, Y. and VEMPALA, S. S. (2025). Sampling and integration of logconcave functions by algorithmic diffusion. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, 924–932. ACM.

[18] KOOK, Y. and VEMPALA, S. S. (2025). Faster logconcave sampling from a cold start in high dimension. In *Proceedings of the 66th Annual IEEE Symposium on Foundations of Computer Science*, 997–1006. IEEE.

[19] KOOK, Y. and VEMPALA, S. S. (2026). A unified complexity bound for logconcave sampling. arXiv preprint arXiv:2606.12694v1.

[20] KOOK, Y., VEMPALA, S. S. and ZHANG, M. S. (2026). In-and-Out: Algorithmic diffusion for sampling convex bodies. *Random Structures Algorithms* **68**(3), e70061.

[21] KRANTZ, S. G. and PARKS, H. R. (2002). *The Implicit Function Theorem: History, Theory, and Applications*. Birkhäuser, Boston.

[22] LADDHA, A., LEE, Y. T. and VEMPALA, S. S. (2020). Strong self-concordance and sampling. In *Proceedings of the 52nd Annual ACM Symposium on Theory of Computing*, 1212–1222. ACM.

[23] LAWLER, G. F. and SOKAL, A. D. (1988). Bounds on the L^2 spectrum for Markov chains and Markov processes: A generalization of Cheeger's inequality. *Trans. Amer. Math. Soc.* **309**(2), 557–580.

[24] LEDOUX, M. (2001). *The Concentration of Measure Phenomenon*. Mathematical Surveys and Monographs **89**. Amer. Math. Soc., Providence, RI.

[25] LEE, Y. T. and SIDFORD, A. (2014). Path finding methods for linear programming: Solving linear programs in $\widetilde{O}(\sqrt{\mathrm{rank}})$ iterations and faster algorithms for maximum flow. In *Proceedings of the 55th Annual IEEE Symposium on Foundations of Computer Science*, 424–433. IEEE.

[26] LEE, Y. T. and SIDFORD, A. (2019). Solving linear programs with $\sqrt{\mathrm{rank}}$ linear system solves. arXiv preprint arXiv:1910.08033.

[27] LEE, Y. T. and VEMPALA, S. S. (2024). Eldan's stochastic localization and the KLS conjecture: Isoperimetry, concentration and mixing. *Ann. of Math. (2)* **199**(3), 1043–1092.
[28] LEWIS, D. R. (1978). Finite dimensional subspaces of L_p. *Studia Math.* **63**(2), 207–212.
[29] LOVÁSZ, L. and SIMONOVITS, M. (1993). Random walks in a convex body and an improved volume algorithm. *Random Structures Algorithms* **4**(4), 359–412.
[30] LOVÁSZ, L. and VEMPALA, S. S. (2007). The geometry of logconcave functions and sampling algorithms. *Random Structures Algorithms* **30**(3), 307–358.
[31] NARAYANAN, H. (2016). Randomized interior point methods for sampling and optimization. *Ann. Appl. Probab.* **26**(1), 597–641.
[32] NUALART, D. (2006). *The Malliavin Calculus and Related Topics*, second ed. Springer, Berlin.
[33] PECCATI, G. and TAQQU, M. S. (2011). *Wiener Chaos: Moments, Cumulants and Diagrams: A Survey with Computer Implementation*. Bocconi & Springer Series **1**. Springer, Milan.
[34] SACHDEVA, S. and VISHNOI, N. K. (2016). The mixing time of the Dikin walk in a polytope—a simple proof. *Oper. Res. Lett.* **44**(5), 630–634.
[35] TSYBAKOV, A. B. (2009). *Introduction to Nonparametric Estimation*. Springer Series in Statistics. Springer, New York.