Sharp convergence of heat-kernel Sinkhorn metrics

Scott Schubr
Luminary Oddities

October 21, 2026

Abstract

We study the Sinkhorn metric associated with the periodic heat kernel at a smooth strictly positive probability density. If P_h is the symmetrically normalized heat operator, its metric tensor is $hP_h(I - P_h^2)^{-1}$ on density-mean-zero functions. We prove convergence to the inverse weighted Laplacian in full L^2 operator norm at the optimal rate $O(h)$. The sharper rate from H^r to L^2 is $O(h^{1+r/2})$, $0 \le r \le 2$; every exponent is optimal already at constant density. Parameter-differentiated estimates yield second-order convergence of smooth finite-dimensional pullback metrics and their intrinsic Christoffel symbols and curvature under nondegeneracy. A population normalization residual ρ gives projected tensor error $O(h + \rho/h)$, with an obstruction at $\rho \asymp h$. We also give an explicit projected scale approximation whose tensor differs from the exact tensor by $O(h^2)$. The cost depends on the heat scale; fixed Euclidean squared-cost problems with boundaries are outside the scope.

Keywords: Sinkhorn divergence; heat kernel; weighted Laplacian; operator convergence; Sobolev estimates; metric curvature.

1 Introduction

Convergence of the Sinkhorn divergence does not directly control its local metric: the formula for that metric contains an inverse with a shrinking spectral gap. A fixed-mode consistency expansion need not approximate the inverse uniformly. In the periodic heat model we distinguish the sharp $O(h)$ full L^2 rate from the $O(h^2)$ rate on inputs with two Sobolev derivatives. Frequencies of order $h^{-1/2}$ explain the distinction. We retain it under density-parameter differentiation and quantify the amplification of normalization errors.

The tensor identity and formal density limit are due to Lavenant, Luckhardt, Mordant, Schmitzer and Tamanini [1]. Mordant and Xu [2] prove convergence and higher-derivative convergence for horizontal perturbations of finitely supported measures and explain the difficulty with their density analogue. Our heat cost varies with h; this theorem does not settle their fixed Euclidean-cost boundary-domain problem.

Wormell and Reich [3] study the same symmetric periodic Gaussian normalization, with spectral-window and semigroup convergence. Their results strongly precede qualitative inverse convergence: controlled finite spectral blocks, absence of spurious low spectrum and positivity provide a finite-block and tail argument. Here we establish sharp all-frequency and Sobolev rates and parameter-differentiated consequences. Potential expansions already occur in [4]; approximate normalization and generator consistency are treated in [5]. Symmetric splitting and its operator estimates are classical; see [6]. The self-transport Taylor and inverse mechanism used in [7] also motivates the present scale calculation, which we prove independently in every fixed dimension. Its statistical stopping theorems are not premises of this paper.

2 Model and statements

Let $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ have normalized Lebesgue measure and set $H_h = \exp(h\Delta/2)$, with kernel k_h. Its Fourier multipliers are $e^{-2\pi^2 h|k|^2}$. Let $f \in C^\infty(\mathbb{T}^d)$ be a positive probability density, $s = \sqrt{f}$, and define the symmetric scale and Markov operator by

$$a_h H_h(a_h f) = 1, \qquad P_h u = a_h H_h(a_h f u). \tag{1}$$

Existence for small h and uniqueness follow below. Write $L^2_0(f) = \{u \in L^2(f\,\mathrm{d}x) : \int uf = 0\}$ and let $D_f = -f^{-1}\operatorname{div}(f\nabla)$ be the nonnegative selfadjoint weighted Laplacian. On $L^2_0(f)$ put

$$T_{h,f} = hP_h(I - P_h^2)^{-1}, \qquad G_f = D_f^{-1}. \tag{2}$$

$H^r_0(f)$ denotes $H^r(\mathbb{T}^d) \cap L^2_0(f)$ with the ordinary periodic Sobolev norm.

Proposition 2.1 (Exact identification)**.** *Use entropic parameter $\varepsilon = 2h$ and cost*

$$c_h(x, y) = -2h \log \frac{k_h(x - y)}{k_h(0)}.$$

Define $\mathrm{OT}_{\varepsilon,c}(\mu, \nu)$ as the infimum of $\int c\,\mathrm{d}\pi + \varepsilon\,\mathrm{KL}(\pi \mid \mu \otimes \nu)$ over couplings, and let $S_{\varepsilon,c}$ be its debiased divergence. For smooth $u \in L^2_0(f)$,

$$\lim_{t\to 0} t^{-2} S_{2h,c_h}(f\,\mathrm{d}x, (f + tfu)\,\mathrm{d}x) = \langle u, T_{h,f}u\rangle_{L^2(f)}.$$

The tensor extends continuously to $L^2_0(f)$.

Proof. The Gibbs kernel is $k_h/k_h(0)$ and its symmetric factors are $\sqrt{k_h(0)}a_h$. Thus its normalized self-coupling density relative to $(f\,\mathrm{d}x)^{\otimes 2}$ is $a_h(x)k_h(x - y)a_h(y)$. All heat Fourier coefficients are positive; this gives positive definiteness, universality and $k_h(z) \le k_h(0)$, so $c_h \ge 0$. Apply [1, Theorem 3.4 and Definition 4.1], in the compact-space $m = 0$ case. The coefficient is $(\varepsilon/2)P_h(I - P_h^2)^{-1}$, hence (2). Smooth tangents give positive affine curves for small two-sided t. Their density and boundedness of the tensor give its continuous extension. □

Theorem 2.2 (Sharp rates)**.** *For every fixed $d \ge 1$, positive smooth f, and $0 \le r \le 2$,*

$$\|T_{h,f} - G_f\|_{H^r_0(f)\to L^2(f)} \le C_r h^{1+r/2}, \qquad 0 < h < h_0. \tag{3}$$

At $r = 0$ this gives the full $L^2_0(f)$ operator-norm estimate, using equivalence of weighted and unweighted L^2 norms. Each exponent is optimal already for $f = 1$, where the norm is at least $c_r h^{1+r/2}$ for all small h. Constants are uniform on classes with a common positive lower bound and common bounds on the finitely many spatial derivatives used in the proof.

For parameter families compare operators on the fixed space $L^2(\mathrm{d}x)$. Let f_θ be jointly smooth in x, θ on a neighborhood of a compact $K \subset \mathbb{R}^p$, with mass one, a common positive lower bound and uniform derivative bounds there. Put $s_\theta = \sqrt{f_\theta}$ and $U_\theta u = s_\theta u$. Let $\mathcal{T}_{h,\theta} = U_\theta T_{h,f_\theta} U_\theta^{-1}$ and $\mathcal{G}_\theta = U_\theta G_{f_\theta} U_\theta^{-1}$, both extended by zero on $\operatorname{span}(s_\theta)$.

Theorem 2.3 (Parameters and intrinsic geometry)**.** *For every fixed integer $k \ge 0$ and $0 \le r \le 2$,*

$$\max_{|\alpha|\le k} \sup_{\theta\in K} \|\partial_\theta^\alpha(\mathcal{T}_{h,\theta} - \mathcal{G}_\theta)\|_{H^r\to L^2} \le C_{k,r} h^{1+r/2}. \tag{4}$$

Set $\eta_i = \partial_i f_\theta/s_\theta$, $g^h_{ij} = \langle \eta_i, \mathcal{T}_{h,\theta}\eta_j\rangle$ and $g^0_{ij} = \langle \eta_i, \mathcal{G}_\theta\eta_j\rangle$. Then $g^h - g^0 = O_{C^k(K)}(h^2)$ for each fixed k. If the family is an immersion on a neighborhood of K, g^0 is uniformly positive definite on K. For $k \ge 2$, its coordinate Christoffel symbols converge in C^{k-1} and intrinsic Riemann curvature components in C^{k-2} at rate $O(h^2)$.

This is intrinsic geometry of the finite-dimensional family, not ambient infinite-dimensional curvature or convergence of induced path distances. Smoothness is convenient; every fixed claim uses finitely many derivatives.

3 Scaling and heat comparisons

Let M_v denote multiplication by v and put

$$b_h = a_h s, \quad V = \frac{\Delta s}{2s}, \quad S = -\Delta/2 + V, \quad Q_h = M_{b_h} H_h M_{b_h}, \quad E_h = e^{-hS}. \tag{5}$$

Then $Q_h = U_f P_h U_f^{-1}$ and $S = U_f (D_f/2) U_f^{-1}$. The form of S is $\frac{1}{2}\int f|\nabla(v/s)|^2 \,\mathrm{d}x$. Thus $S \geq 0$, its kernel is $\mathrm{span}(s)$ and weighted Poincare gives a positive gap on $s^\perp$.

Lemma 3.1 (The scale)**.** *For every fixed integer $m > d/2$,*

$$b_h = e^{-hV/2} + O_{H^m}(h^3). \tag{6}$$

The solution is positive and unique among positive continuous scales. For the parameter families above the remainder holds in every fixed parameter derivative order. Positive-order derivatives of b_h are $O_{H^m}(h)$.

Proof. In the Sobolev algebra H^m use $F_h(b) = bH_h(bs)/s - 1$. Its derivative at one is

$$DF_h(1) = M_{1/s}(I + H_h)M_s + M_{H_h s/s - 1}.$$

The first summand has inverse $M_{1/s}(I + H_h)^{-1} M_s$, uniformly bounded by the Fourier multiplier bound and Sobolev multiplier estimates. The second summand is $O(h)$. Hence the derivative has a uniform inverse, and F_h has a uniform second-derivative bound on a fixed neighborhood. For $A = \Delta/2$ insert $b_* = 1 + hb_1 + h^2 b_2$ into $b_* H_h(b_* s)$. The coefficients give $b_1 = -As/(2s) = -V/2$, $b_2 = V^2/8$: the $A^2 s$ terms at second order cancel. Heat Taylor remainders, using sufficient spatial derivatives, give residual $O_{H^m}(h^3)$ also for $b_* = e^{-hV/2}$. A quantitative Newton contraction in a ball of radius Ch^3 about b_* proves (6). Positivity follows by Sobolev embedding. No operator-norm convergence $H_h \to I$ is assumed.

For uniqueness put $w_i = a_i f$, so $w_i H_h w_i = f$. Their difference z satisfies $H_h z = -fz/(w_1 w_2)$. Its pairing with z is nonnegative on the left and strictly negative on the right unless $z = 0$. Parameter differentiation of the residual and equation, with the uniform inverse, gives the remainder estimates inductively. Derivatives of $e^{-hV/2}$ of positive order are $O(h)$. □

Lemma 3.2 (Full-norm comparison)**.** *Q_h, E_h are positive selfadjoint operators fixing s, and for all fixed k, $\|Q_h - E_h\|_{C^k(K;L^2 \to L^2)} \leq C_k h^2$.*

Proof. The fixed vectors follow from the scale equation and $Ss = 0$. Reduce by Lemma 3.1 to comparison with $R_h = M_{e^{-hV/2}} H_h M_{e^{-hV/2}}$. Lift V periodically to $\mathbb{R}^d$. With $z = y - x$, a Brownian bridge is $X_t = x + (t/h)z + \beta_t$, with centered covariance $t(h-t)I/h$. The Feynman–Kac factor is $\mathbb{E}\exp(-\int_0^h V(X_t)\,\mathrm{d}t)$. Bounded second derivatives imply

$$\left| \mathbb{E}\int_0^h V(X_t)\,\mathrm{d}t - \frac{h}{2}(V(x) + V(y)) \right| \leq C(h|z|^2 + h^2).$$

Taylor expansion on the segment cancels the linear terms; the bridge covariance contributes Ch^2. Both exponents are bounded by Ch, so replacement by one minus the exponent costs Ch^2. Integrate the absolute kernel difference using $\int |z|^2 k_h^{\mathbb{R}^d}(z)\,\mathrm{d}z = dh$. Both row and column bounds are Ch^2. Summing lifts preserves the bounds, and Schur's test proves the assertion.

For parameter derivatives the linear term satisfies the same estimate with $\partial^\alpha V$. Differentiated exponential remainders contain at least two time integrals, hence remain $O(h^2)$. The scale remainder has the stated derivative control. □

Lemma 3.3 (Strong comparison)**.** *For each fixed k, $\|Q_h - E_h\|_{C^k(K;H^2\to L^2)} \le C_k h^3$.*

Proof. Reduce to symmetric splitting by Lemma 3.1. Set $A = \Delta/2$, $B = -M_V$, $C = [A, B]$ and $R(t) = e^{tB/2}e^{tA}e^{tB/2}$. $[A, C] : H^2 \to L^2$ is bounded and $[C, B]$ is a bounded multiplier: their differential orders are at most two and zero. For smooth inputs, remove the bounded outer $e^{tB/2}$ factors from $R'(t) - (A + B)R(t)$. The remaining defect is

$$-\int_0^{t/2} e^{-rB}Ce^{rB}\,\mathrm{d}r\, e^{tA} + \frac{1}{2}\int_0^t e^{(t-r)A}Ce^{rA}\,\mathrm{d}r. \tag{7}$$

Subtract Ce^{tA} from both integrands; their integrals cancel. Use

$$e^{-rB}Ce^{rB} - C = \int_0^r e^{-vB}[C, B]e^{vB}\,\mathrm{d}v,$$
$$e^{(t-r)A}Ce^{rA} - Ce^{tA} = \int_0^{t-r} e^{(t-r-v)A}[A, C]e^{(v+r)A}\,\mathrm{d}v.$$

The defect is $O(t^2)$ from H^2 to L^2. Variation of constants, boundedness of $e^{t(A+B)}$ on L^2, and density give the $O(h^3)$ local error. Every parameter derivative of the commutator coefficients has the same differential order. Smooth multipliers preserve H^2; derivatives of the potential exponentials and of $e^{t(A+B)}$ are bounded on the required spaces by iterated Duhamel integrals. Differentiating the displayed cancellation retains the $O(t^2)$ defect and proves the assertion for every fixed k. There is no assumed bounded L^2 double commutator. □

4 Singular inverses and parameter estimates

P_h is Markov and contracts $L^2(f)$, so Q_h is a positive compact contraction. Its spectrum is in $[0, 1]$, with positive nonzero eigenvalues. The gap of S and Lemma 3.2 give, on $s^\perp$,

$$\|Q_h\|, \|E_h\| \le 1 - ch, \qquad \|(I - Q_h)^{-1}\| + \|(I - E_h)^{-1}\| \le C/h. \tag{8}$$

The plus inverses have norm at most one. The exact identity

$$hR(I - R^2)^{-1} = \frac{h}{2}[(I - R)^{-1} - (I + R)^{-1}] \tag{9}$$

and the resolvent identity bound the Q_h versus E_h tensor difference by Ch on L^2. The latter tensor is $h/[2\sinh(hS)]$. For $\lambda > 0$,

$$g_h(\lambda) = \frac{h}{2\sinh(h\lambda)} - \frac{1}{2\lambda}, \quad |g_h(\lambda)| \le Ch, \quad |g_h(\lambda)| \le Ch^2\lambda. \tag{10}$$

These are scalar Taylor bounds near $h\lambda = 0$ and elementary bounds away from zero. This proves the L^2 estimate.

The graph norm of S is equivalent to H^2. Its commutation with E_h shows that $(I - E_h)^{-1}$ is $O(1/h)$ on $H^2 \cap s^\perp$ also. Hence

$$\frac{h}{2}(I - Q_h)^{-1}(Q_h - E_h)(I - E_h)^{-1}$$

is $O(h^2)$ from $H^2 \cap s^\perp$ to L^2 by Lemma 3.3. The plus-inverse difference contributes $O(h^4)$ in this norm. The second bound in (10) proves the reference strong estimate. Projection onto $s^\perp$ is bounded on every H^r; extend by this projection and interpolate to obtain $0 \le r \le 2$. Multiplication by s and $1/s$ transfers the bounds to $H^r_0(f)$.

4.1 Moving nullspace

On the fixed $L^2(\,\mathrm{d}x)$ let $\Pi_\theta = |s_\theta\rangle\langle s_\theta|$, $J_\theta = I - \Pi_\theta$, and

$$B_Q = I - Q_h + h\Pi_\theta, \qquad B_E = I - E_h + h\Pi_\theta.$$

Their inverse norms are $O(1/h)$ on L^2. By graph-norm equivalence and the smooth finite-rank projection, B_E^{-1} is $O(1/h)$ on H^2 as well. Positive-order parameter derivatives of B_Q are $O(h)$ on L^2 by Lemma 3.1. Those of B_E are $O(h)$ on L^2 and H^2 by Duhamel, since derivatives of V are bounded smooth multipliers on these spaces. Repeated inverse differentiation therefore gives

$$\partial^\alpha B_Q^{-1} = O_{L^2\to L^2}(h^{-1}), \qquad \partial^\alpha B_E^{-1} = O_{H^j\to H^j}(h^{-1}), \quad j = 0, 2.$$

Every differentiated term has one more inverse than differentiated B factors, so its power is h^{-1}. The identity $B_Q^{-1} - B_E^{-1} = B_Q^{-1}(Q_h - E_h)B_E^{-1}$ is thus $O_{C^k}(1)$ on L^2 and $O_{C^k}(h)$ from H^2 to L^2. The exact zero-extended Q tensor is

$$\frac{h}{2}[B_Q^{-1} - h^{-1}\Pi_\theta - J_\theta(I + Q_h)^{-1}J_\theta]. \tag{11}$$

Subtract the E expression. Projections cancel exactly. The plus-inverse differences are $O_{C^k}(h^2)$ on L^2 and $O_{C^k}(h^3)$ from H^2 to L^2. We obtain tensor differences $O_{C^k}(h)$ and $O_{C^k}(h^2)$ respectively.

4.2 The differentiated scalar filter

Set $A_\theta = S_\theta + \Pi_\theta$, uniformly strictly positive and selfadjoint, with common domain H^2. Its positive-order parameter derivatives are bounded L^2 operators. g_h is analytic on each sector $|\arg z| < \gamma < \pi/2$, with a removable value at zero and $|g_h(z)| \le C_\gamma h$. Take a contour on sector-boundary rays joined by a fixed small arc below the common spectral lower bound. Its resolvent satisfies $\|(z - A_\theta)^{-1}\| \le C/(1+|z|)$. The spectral theorem controls the zeroth derivative. Every positive-order parameter derivative has at least two resolvents and bounded derivatives of A_θ, so its contour integral is bounded by $Ch\int_0^\infty (1+r)^{-2}\,\mathrm{d}r$. These uniformly convergent derivative integrals justify the functional-calculus differentiation, first on truncated contours.

For the strong norm define

$$j(w) = \frac{1}{w}\left(\frac{1}{2\sinh w} - \frac{1}{2w}\right).$$

It is bounded analytic on the sector, $j(0) = -1/12$, and decays at infinity. The spectral theorem bounds $j(hA_\theta)$; the same differentiated contour argument, with bounded $j(hz)$, gives $\partial^\alpha j(hA_\theta) = O(1)$ on L^2. Now

$$g_h(A_\theta) = h^2 j(hA_\theta)A_\theta$$

is $O_{C^k}(h^2)$ from H^2 to L^2 by the product rule. $J_\theta g_h(A_\theta)J_\theta$ is the reference tensor remainder. Combining with (11) and interpolation proves (4).

Contract the strong estimate with the smooth η_i. They satisfy $\langle \eta_i, s_\theta\rangle = \int \partial_i f_\theta = 0$. Their Gram matrix in the positive inverse-Laplacian form is positive definite for an immersion, and compactness gives its uniform lower bound. Standard coordinate formulas for inverse metrics, Christoffel symbols and curvature finish Theorem 2.3.

4.3 Sharpness

At $f = 1$, a mode with $\mu_k = 4\pi^2|k|^2$ has tensor multiplier $h/[2\sinh(h\mu_k/2)]$. Choose $k_h = (\text{round}\,\sqrt{t/(2\pi^2 h)}, 0, \ldots, 0)$. The absolute error divided by h tends to $1/(2t) - 1/(2\sinh t) > 0$. Its H^r norm grows like $h^{-r/2}$, giving the lower bound $ch^{1+r/2}$. For $r = 2$ a fixed nonzero mode also works: its error is $-h^2\mu_k/24 + O(h^4)$. This completes Theorem 2.2.

5 Normalization and an explicit approximation

Proposition 5.1 (Population residual)**.** *Fix $m > d/2$ and f. For positive $\widetilde{b}$ in a sufficiently small fixed H^m neighborhood of one, set $\rho = \|\widetilde{b}H_h(\widetilde{b}s)/s - 1\|_{H^m}$. Then $\|\widetilde{b} - b_h\|_{H^m} \le C\rho$. Let $J = I - |s\rangle\langle s|$ and $\overline{Q} = JM_{\widetilde{b}}H_hM_{\widetilde{b}}J|_{s^\perp}$. If $\rho \le c_0 h$ with c_0 sufficiently small, $\overline{Q}$ is a positive strict contraction and*

$$\|h\overline{Q}(I - \overline{Q}^2)^{-1} - (2S)^{-1}\|_{s^\perp \to s^\perp} \le C(h + \rho/h).$$

Proof. The derivative of F_h at b_h has a uniform inverse. Taylor expansion and absorption of the quadratic remainder in the fixed small neighborhood give the scale bound. Sobolev multiplication yields $\|\overline{Q} - Q_h|_{s^\perp}\| \le C\rho$. Compression preserves positivity and the subspace, and $\rho \le c_0 h$ preserves the gap. The minus-resolvent difference is $O(\rho/h^2)$, so (9) gives tensor error $O(\rho/h)$. Add Theorem 2.2. □

Consequently $\rho = o(h)$ suffices for convergence and $\rho = O(h^2)$ retains the full-norm $O(h)$ rate. This is a known-density population Sobolev statement with explicit projection, not an empirical-residual or stopping-time guarantee. The scale is critical in a worst-case sense. At $f = 1$ set $\widetilde{b} = \sqrt{1 + ch}$, $0 < c < \mu_1/2$. Its residual is ch, and compressed eigenvalues are $(1 + ch)e^{-h\mu/2} < 1$ for small h, yet their tensor multipliers tend to $(\mu - 2c)^{-1}$ instead of μ^{-1}. Thus a vanishing order-h residual alone is insufficient.

Corollary 5.2 (Explicit projected scale)**.** *Let $b_h^0 = e^{-hV/2}$, and let $\mathcal{T}_{h,\theta}^0$ be the tensor of $J_\theta M_{b_h^0} H_h M_{b_h^0} J_\theta$ restricted to $s_\theta^\perp$, extended by zero. For every fixed k,*

$$\|\mathcal{T}_{h,\theta}^0 - \mathcal{T}_{h,\theta}\|_{C^k(K;L^2\to L^2)} \le C_k h^2.$$

In particular it approximates the limiting tensor at $O(h^2)$ from H^2 to L^2, also with parameter derivatives.

Proof. Lemma 3.1 gives $Q_h^0 - Q_h = O_{C^k}(h^3)$ on L^2. After compression the gap persists. On the full fixed space augment the compression by Π_θ. The augmented operator differs from Q_h by $J_\theta(Q_h^0 - Q_h)J_\theta = O_{C^k}(h^3)$. It fixes s_θ; its positive-order derivatives are $O(h)$. The regularized inverse proof of (11) then gives $h \cdot h^{-1} \cdot h^3 \cdot h^{-1} = h^2$ with all fixed parameter derivatives. Apply Theorem 2.3 for the last assertion. □

The explicit scale uses the true density, its derivatives and its ground-state projection. It avoids exact scale solution in this population construction; no empirical algorithm or runtime bound follows.

6 Concluding questions

The sharp norm scale explains why fixed-mode asymptotics alone cannot quantify uniform local geometry. Smooth finite-dimensional models inherit the second-order rate even though full L^2 approximation is only first order. Normalization accuracy must also be judged relative to the shrinking gap. Fixed squared costs with boundaries, vanishing densities, empirical inverse tensors and full ambient curvature remain separate problems. The present constants and domains do not transfer automatically.

Computational assistance. AI tools assisted literature research, derivation, proof checking and manuscript preparation. The proofs are analytic and do not depend on numerical experiments or internal review reports.

References

[1] H. Lavenant, J. Luckhardt, G. Mordant, B. Schmitzer and L. Tamanini, *The Riemannian geometry of Sinkhorn divergences*, arXiv:2405.04987v2, 2025.

[2] G. Mordant and L. Xu, *Convergence of the Sinkhorn Riemannian metric for finitely supported measures*, arXiv:2609.14651v1, 2026.

[3] C. L. Wormell and S. Reich, *Spectral convergence of diffusion maps: improved error bounds and an alternative normalisation*, arXiv:2006.02037v3, 2021.

[4] R. Yao, Y.-H. Kim and G. Schiebinger, *Quantitative Stability of Many-Marginal Schrödinger Bridge*, arXiv:2604.15191v1, 2026.

[5] X. Cheng and B. Landa, *Bi-stochastically normalized graph Laplacian: convergence to manifold Laplacian and robustness to outlier noise*, arXiv:2206.11386v3, 2024.

[6] T. Ichinose and H. Tamura, *Results on Convergence in Norm of Exponential Product Formulas and Pointwise of the Corresponding Integral Kernels*, arXiv:0905.3458, 2009.

[7] S. Schubr, *First-Hit Stopping and Bias-Corrected Score Inference in Dimension Three*, manuscript 2.1, public edition 2, 2026. Luminary Research Archive.