Fourier computation of empirical diffusion metrics

Scott Schubr
Luminary Oddities

October 30, 2026

Abstract

We give a finite Fourier algorithm for statistical inference on inverse weighted-Laplacian metric entries estimated by empirical heat-kernel Sinkhorn normalization. For every fixed positive overhead exponent, the algorithm uses a deterministic number of exact-real operations bounded by the sample size to the power one plus that exponent. It preserves the population-centered Gaussian limit, consistent joint influence covariance, and the ordinary nonparametric bootstrap of the exact metric estimator. The density is fixed, smooth and strictly positive on a fixed-dimensional flat torus, and the number of prescribed smooth scores is fixed. The proof establishes an empirical spectral gap proportional to bandwidth, a logarithmic number of actual half-scaling iterations, and complete original-law error bounds for the inverse response and all covariance products. Positive normalized Fourier features and reduced Woodbury solves pay the full feature Gram and factorization costs. Higher fixed-order bias extrapolation permits the bandwidth needed for each prescribed operation bound.

Keywords: Sinkhorn normalization, diffusion metric, Fourier features, statistical computation, influence covariance.
Mathematics Subject Classification (2020): 62G20, 65C60, 65T50.

1 Introduction

An empirical diffusion metric involves both kernel normalization and an inverse whose nonconstant spectral gap closes as the bandwidth tends to zero. Fast kernel multiplication alone therefore gives an incomplete computational guarantee. For statistical inference, the numerical metric must approximate its exact counterpart more accurately than the root-sample fluctuation, and the influence calculation must retain the response of the normalization and the inverse. A small error in the metric does not by itself control these quadratic covariance calculations.

We resolve this computational question for the periodic heat-kernel Sinkhorn metric and the population-centered estimator of [7]. The algorithm forms positive finite Fourier features, performs a prescribed logarithmic number of half-scaling iterations from the constant vector, and solves reduced feature systems. It computes metric entries and their complete joint influence covariance under the original empirical law. The main theorem gives the total operation count, including feature Gram formation and factorization, and returns the inferential limits of the exact estimator. The same construction applies to a fixed ordinary bootstrap replicate.

The operation bound is pointwise at each fixed positive smooth density. For every fixed $\eta > 0$, sufficiently high but fixed extrapolation order permits a bandwidth for which the work is $O(n^{1+\eta})$. All constants may depend on the density, dimension, prescribed scores, extrapolation order and chosen exponent. We count real arithmetic and elementary function evaluations at unit cost. This declared model specifies the theorem's computational guarantee.

1.1 Relation to previous work

The symmetry-preserving geometric half iteration is classical; see Knight–Ruiz–Uçar [4, Section 3]. Symmetric Sinkhorn normalization as a diffusion-kernel construction is developed by Kessler–Magnet–Feydy [3]. We use this established normalization and prove a bandwidth-uniform count for the actual sampled trajectory. Fourier acceleration of Sinkhorn iterations is established; Lakshmanan–Pichler–Potts [5] develop an NFFT implementation, while Altschuler–Bach–Rudi–Niles-Weed [1] combine low-rank Nyström approximation with scaling stability and explicit transport-cost guarantees. We use the classical finite-feature and reduced-inverse architecture. The additional issue here is the simultaneous shrinking bandwidth, inverse metric, actual iteration count and original-law influence covariance.

The exact statistical target and its root-n inference are proved in [7]. We state the specific inputs needed from that paper before proving the new numerical return. Differentiability and inference for fixed-regularization Sinkhorn geometric operations are developed by Saidi–Gunsilius–Mena [6]. Dong et al. [2] study sensitivity and conditioning for finite-support inverse cost parameters observed through transition plans. These provide related operator and statistical perspectives; the present theorem concerns fixed-score inverse diffusion energies of an unknown sampled density as the heat bandwidth vanishes.

The proof has three computational obligations. First, a feature-covariance comparison gives an empirical gap of order bandwidth. Second, positive semidefiniteness gives local contraction of the actual half iterations, after one smoothing step has entered a small normalization neighborhood. Third, direct original-law product estimates pay the complete inverse response and covariance, including their mixed terms. Sections 3–5 carry out these steps.

2 Target, exact estimator and main theorem

Let $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ have volume one. Fix a strictly positive $f \in C^\infty(\mathbb{T}^d)$ with $\int f = 1$, put $\mu = f\,\mathrm{d}x$, and fix real $u_1, \dots, u_p \in C^\infty(\mathbb{T}^d)$. Dimension d and the number p of scores are fixed. If $u_a^0 = u_a - \mu u_a$, define

$$D_f = -f^{-1}\operatorname{div}(f\nabla), \qquad D_f\phi_a = u_a^0, \quad \mu\phi_a = 0, \qquad G_{ab}^0 = \mu(u_a^0\phi_b). \tag{1}$$

The prescribed metric is equivalently $G_{ab}^0 = \mu(\nabla\phi_a \cdot \nabla\phi_b)$. Its influence and joint covariance are

$$\begin{aligned} \psi_{ab}^0 &= u_a^0\phi_b + u_b^0\phi_a - \nabla\phi_a\cdot\nabla\phi_b - G_{ab}^0, \\ \Sigma_{ab,cd} &= \mu(\psi_{ab}^0\psi_{cd}^0), \qquad a \le b,\ c \le d. \end{aligned} \tag{2}$$

For $t > 0$, let $H_t = e^{t\Delta/2}$ and write its periodic kernel as

$$k_t(x) = \sum_{r\in\mathbb{Z}^d} e^{-2\pi^2 t|r|^2} e^{2\pi i r\cdot x}. \tag{3}$$

At a probability law ν, the positive symmetric scale and Markov operator satisfy

$$a_{t,\nu}H_t(a_{t,\nu}\nu) = 1, \qquad P_{t,\nu}v = a_{t,\nu}H_t(a_{t,\nu}v\nu). \tag{4}$$

For finite positive laws the scale is the unique finite Sinkhorn scale. Let $\Pi_\nu v = (\nu v)1$ and put

$$T_{t,\nu} = \frac{t}{2}\left\{(I - P_{t,\nu} + \Pi_\nu)^{-1} - (I + P_{t,\nu})^{-1} - \frac{\Pi_\nu}{2}\right\}, \qquad G_{ab}(t,\nu) = \nu(u_a T_{t,\nu} u_b). \tag{5}$$

Thus $T1 = 0$ and, on mean-zero functions, $T = tP(I - P^2)^{-1}$.

For an integer $s > d/2$ define

$$M = 2 + \frac{5s}{2}, \qquad 0 < \beta < \min\left\{\frac{1}{2M}, \frac{\eta}{4d}\right\}, \qquad 2K > M, \quad \frac{1}{4K} < \beta, \tag{6}$$

where K is an integer and $\eta > 0$ is prescribed. Such choices exist: choose β first and then increase K. For $n \geq 2$ set $h_n = n^{-\beta}$ and $t_\ell = h_n/2^\ell$, $0 \leq \ell < K$. The fixed extrapolation weights are determined by

$$\sum_{\ell=0}^{K-1} c_\ell = 1, \qquad \sum_{\ell=0}^{K-1} c_\ell 4^{-\ell j} = 0 \quad (1 \leq j < K). \tag{7}$$

Let $X_1, \ldots, X_n$ be iid from μ and $\mu_n = n^{-1}\sum_i \delta_{X_i}$. The exact estimator is

$$\widehat{G}_n = \sum_{\ell=0}^{K-1} c_\ell G(t_\ell, \mu_n). \tag{8}$$

Theorem 2.1 (Computational inference)**.** *Under the fixed model above, the algorithm in Section 3 returns a matrix $\widetilde{G}_n$, a joint covariance matrix $\widetilde{\Sigma}_n$ and a failure flag. Its deterministic work on every sample is*

$$O(nR_n^2 + R_n^3 + nR_n \log n) = O(n^{1+\eta}), \qquad R_n = O(h_n^{-d/2}(\log n)^{d/2}), \tag{9}$$

and its memory is $O(nR_n + R_n^2)$. Real arithmetic, elementary exponentials, logarithms, trigonometric functions, square roots and prescribed score evaluations have unit cost. At the fixed density f, the probability of a zero failure flag tends to one, and

$$\sqrt{n}\,\|\widetilde{G}_n - \widehat{G}_n\| \longrightarrow 0, \qquad \widetilde{\Sigma}_n \longrightarrow \Sigma \quad \textit{in probability.} \tag{10}$$

Consequently

$$\sqrt{n}\,\operatorname{vec}(\widetilde{G}_n - G^0) \Longrightarrow N(0, \Sigma), \tag{11}$$

where vectorization uses entries $a \leq b$. Every fixed contrast with positive limiting variance admits the corresponding asymptotically standard normal studentization.

For one ordinary nonparametric bootstrap replicate, use the same bandwidths, weights and iteration budget, merging repeated sites. If $\widetilde{G}_n^*$ is its output, then

$$d_{\mathrm{BL}}\left(\mathcal{L}[\sqrt{n}\,\operatorname{vec}(\widetilde{G}_n^* - \widetilde{G}_n) \mid X_1, \ldots, X_n], N(0, \Sigma)\right) \longrightarrow 0 \tag{12}$$

in probability. Each replicate has the work bound (9)*.*

The theorem fixes p, K, η before n varies. The covariance may be singular; studentization is asserted for fixed positive-variance contrasts. The implementation uses sampled positions and score values, with its bandwidth and budgets set in advance.

2.1 Statistical inputs from the exact estimator

We use the following results from [7, Sections 3–7]. They identify the statistical problem to which the computational return will be applied.

Proposition 2.2 (Exact-estimator inputs)**.** *For the fixed model and choices* (6)*:*

1. *The population scale obeys $a_{t,\mu} = f^{-1/2} + O(t)$ uniformly. Its normalized positive selfadjoint Markov operator has a nonconstant $L^2(\mu)$ gap at least ct for small t.*

2. *With $q = 1 + s/2$ and $\epsilon = \|\nu - \mu\|_{H^{-s}} \le ct^{s+q}$,*

$$\|a_{t,\nu}/a_{t,\mu} - 1\|_{H^s} \le C\epsilon t^{-s}. \tag{13}$$

The local scale at a positive atomic law agrees with its finite scale.

3. *The exact estimator* (8) *has asymptotic linear representation with influence* (2) *and covariance* Σ*, and its exact multinomial bootstrap is consistent. For a scalar score* u*, let* $g = u - \nu u$ *and*

$$z_- = (I - P_\nu)^{-1} g, \quad \nu z_- = 0, \qquad z_+ = (I + P_\nu)^{-1} g, \qquad C_\nu = (I - P_\nu)(I + P_\nu)^{-1}. \tag{14}$$

The exact centered influence is the centered version of

$$\Psi_{t,\nu} = \frac{t}{2} \left\{ g(z_- - z_+) + C_\nu (z_- P_\nu z_- + z_+ P_\nu z_+) \right\}. \tag{15}$$

Products here are pointwise. At the base bandwidth h_n*, polarization of these exact empirical influences has a consistent joint covariance under* μ_n*; each fixed influence has* $L^2(\mu_n)$ *norm* $O_{\mathbb{P}}(1)$*.*

The local neighborhood is reached because $\mathbb{E}\|\mu_n - \mu\|^2_{H^{-s}} = O(n^{-1})$ and $M > s + q$. These inputs retain the simultaneous vanishing bandwidth and the moving empirical normalization. The remaining proof concerns the finite algorithm and its error relative to these exact quantities.

3 A finite algorithm

We describe one band t. The same construction is used at all K bands. Merge repeated sample sites and write the resulting law as $\nu = \sum_{i=1}^m w_i \delta_{x_i}$, with $m \le n$ and positive integer multiples w_i of $1/n$. Put $W = \operatorname{diag}(w_i)$, $e = (\sqrt{w_i})_{i=1}^m$ and $\mathcal{P} = I - ee^T$.

3.1 Positive normalized features

Fix $A = 30$, $B = 20$ and $C_0 = A + d + 1$. Let

$$L = \left\lceil \sqrt{C_0 \log n/t} \right\rceil, \qquad a = 2\pi^2 t, \qquad E_t = d(1 + \sqrt{\pi/a})^{d-1} \frac{e^{-aL^2}}{aL}. \tag{16}$$

Let $k_{t,L}$ be (3) restricted to $|r_j| \le L$ for all j, and define

$$q_t = \frac{k_{t,L} + E_t + n^{-A-2}}{1 + E_t + n^{-A-2}}. \tag{17}$$

Its real sine–cosine features give an $m \times R$ matrix F such that $q_t(x_i - x_j) = (FF^T)_{ij}$. The constant feature is one. For one representative of each pair $\{r, -r\}$ of nonzero retained frequencies, the two other features are

$$\sqrt{\frac{2e^{-2\pi^2 t|r|^2}}{1 + E_t + n^{-A-2}}} \cos(2\pi r \cdot x), \qquad \sqrt{\frac{2e^{-2\pi^2 t|r|^2}}{1 + E_t + n^{-A-2}}} \sin(2\pi r \cdot x).$$

3.2 Scaling and projected solves

Use the symmetry-preserving half iteration [4], starting at $b_0 = 1$, and perform exactly

$$J = 1 + \left\lceil \frac{(B+2)\log n}{\log(4/3)} \right\rceil, \qquad b_{j+1} = \left\{ \frac{b_j}{F[F^T(w\, b_j)]} \right\}^{1/2} \quad (0 \le j < J), \tag{18}$$

componentwise. Form

$$V_0 = W^{1/2}\,\mathrm{diag}(b_J)F, \qquad V = \mathcal{P}V_0, \qquad H = V^T V, \qquad S = VV^T. \tag{19}$$

Cholesky-factor $I - H$ and $I + H$. If $I - H$ is not positive definite, return zero metric and covariance matrices with a nonzero failure flag. Otherwise use

$$\widetilde{R}_- = I + V(I-H)^{-1}V^T, \qquad \widetilde{R}_+ = I - V(I+H)^{-1}V^T, \qquad \widetilde{C} = 2\widetilde{R}_+ - I - ee^T. \tag{20}$$

Only their actions are formed. The last projector makes $\widetilde{C}e = 0$.

For each scalar score needed in polarization, compute

$$\begin{aligned} g &= u - \nu u, \qquad g_E = W^{1/2} g, \\ z_{-,E} &= \widetilde{R}_- g_E, \qquad z_{+,E} = \widetilde{R}_+ g_E, \qquad z_\pm = W^{-1/2} z_{\pm,E}, \\ \widetilde{\Phi}_t(u) &= \frac{t}{2} g_E^T (z_{-,E} - z_{+,E}), \\ U &= z_-(W^{-1/2} S z_{-,E}) + z_+(W^{-1/2} S z_{+,E}), \\ \widetilde{\psi}_E(u) &= \mathcal{P}\frac{t}{2}\left\{ W^{1/2}[g(z_- - z_+)] + \widetilde{C} W^{1/2} U \right\}. \end{aligned} \tag{21}$$

All products in U and in the square brackets are products of function values. Thus the conversion by $W^{-1/2}$ occurs before multiplication.

Polarize with $(u_a + u_b, u_a - u_b)$, dividing the differences by four, to obtain $\widetilde{G}_{ab}(t,\nu)$ and $\widetilde{\psi}_{ab,E}(t,\nu)$. Combine the metrics with (7). At the base band h_n form

$$\widetilde{\Sigma}_{ab,cd} = \widetilde{\psi}_{ab,E}^T \widetilde{\psi}_{cd,E}. \tag{22}$$

This is the original-weight empirical covariance, because the Euclidean influence vectors include $W^{1/2}$. A one-site law has centered inputs zero and returns zero quantities by these same formulas.

4 Gap and scaling bounds

Lemma 4.1 (Feature accuracy)**.** *For a sufficiently large fixed C_0, the kernel* (17) *is real, symmetric, strictly positive, positive semidefinite and has integral one. Uniformly over the fixed bands,*

$$R = O(t^{-d/2}(\log n)^{d/2}), \qquad \|q_t - k_t\|_\infty \le C n^{-A}. \tag{23}$$

Proof. For $a > 0$, monotonicity and a Gaussian integral bound give

$$\sum_{j\in\mathbb{Z}} e^{-aj^2} \le 1 + \sqrt{\pi/a}, \qquad 2\sum_{j>L} e^{-aj^2} \le \frac{e^{-aL^2}}{aL}.$$

A union bound over the coordinates therefore makes E_t a bound for the discarded Fourier absolute sum. The added constant makes the retained series strictly positive since $k_t > 0$. All its Fourier

coefficients are nonnegative, giving positive semidefiniteness. The denominator fixes the constant coefficient to one. Also $\|k_{t,L}\|_\infty \le Ct^{-d/2}$; hence the normalization error is bounded by a constant times $(E_t + n^{-A-2})(1 + t^{-d/2})$. The choice $C_0 = A + d + 1$ gives $E_t \le Ct^{-d/2}n^{-2\pi^2 C_0}$; using $\beta d < 1$ proves (23). Counting the retained cube gives the rank bound. □

Lemma 4.2 (Empirical gap)**.** *With probability tending to one, the exact empirical normalized heat operator at every fixed band has a nonconstant $L^2(\mu_n)$ gap at least $ct/2$, for a constant $c > 0$ depending on the fixed model.*

Proof. Write $a_t = a_{t,\mu}$ and $a_{t,n} = a_{t,\mu_n}$. Sobolev embedding and (13) give

$$\|a_{t,n}/a_t - 1\|_\infty = O_{\mathbb{P}}(n^{-1/2}t^{-s}). \tag{24}$$

Use the heat feature vector $\varphi_t(x) = (e^{-\pi^2 t|r|^2}e^{2\pi i r\cdot x})_r \in \ell^2$ and put $g_t(x) = a_t(x)\varphi_t(x)$. The covariance $C_t = \mathbb{E}[g_t(X) \otimes g_t(X)]$ has the population Markov operator's nonzero spectrum. Its largest eigenvalue is one and its second is at most $1 - ct$, by Proposition 2.2. Since $\|g_t(x)\|^2 \le Ck_t(0) \le Ct^{-d/2}$, Hilbert-space variance gives

$$\mathbb{E}\|C_{t,n} - C_t\|_{\mathrm{HS}}^2 \le Cn^{-1}t^{-d}, \quad C_{t,n} = n^{-1}\sum_i g_t(X_i) \otimes g_t(X_i).$$

The nonzero spectrum of $C_{t,n}$ is that of the sampled Gram matrix with population scales. Replacing them by $a_{t,n}$ is a diagonal congruence by $I + O_{\mathbb{P}}(n^{-1/2}t^{-s})$. Its norm error is of the same order, and $s > d/2$ absorbs the covariance error. Weyl comparison thus gives

$$\lambda_2(P_{t,\mu_n}) \le 1 - ct + O_{\mathbb{P}}(n^{-1/2}t^{-s}).$$

The exact normalized matrix is a positive selfadjoint Markov matrix. Its constant eigenvalue is exactly one. Finally $\beta(s + 1) < 1/2$, so the last error is $o_{\mathbb{P}}(t)$. A fixed union over the bands proves the assertion. □

Lemma 4.3 (Actual scaling trajectory)**.** *For the fixed iteration count* (18)*, uniformly over the fixed bands,*

$$\|\log b_J - \log a_{t,n}\|_\infty = O_{\mathbb{P}}(n^{-B}). \tag{25}$$

Proof. First use the exact heat matrix K_t with the empirical weights. In logarithmic coordinates its half map is $\mathcal{F}_k(u) = \{u - \log(K_t(we^u))\}/2$. It is sup-norm nonexpansive: its derivative is half of identity minus a stochastic matrix and has row absolute sums at most one. If $u_* = \log a_{t,n}$, heat smoothing $H^{-s} \to H^s \to C^0$ and (24) show that its first iterate from zero satisfies

$$\|u_1 - u_*\|_\infty = O_{\mathbb{P}}(t + n^{-1/2}t^{-s}) = o_{\mathbb{P}}(1).$$

Both u_1 and u_* approach $-\log f/2$. The whole later trajectory stays in the same sup-norm box about u_*.

Let P_* be the exact normalized operator in $L^2(w)$. At $u = u_* + v$, the stochastic matrix in the derivative is

$$P_u = \mathrm{diag}((P_*e^v)^{-1})P_*\,\mathrm{diag}(e^v).$$

Positive semidefiniteness and the Markov property give $\|I - P_*\|_{\mathrm{op}} \le 1$. For $\|v\|_\infty \le r$,

$$\|P_u - P_*\|_{\mathrm{op}} \le e^{2r} - 1, \qquad \|D\mathcal{F}_k(u)\|_{\mathrm{op}} \le e^{2r}/2.$$

Choose a fixed small r for which the last bound is at most $3/4$. The straight segment to u_* lies in the box, so the actual iterates contract in $L^2(w)$ there. Since $w_i \geq 1/n$,

$$\|u_J - u_*\|_\infty \leq \sqrt{n}\, r(3/4)^{J-1} = O(n^{-B}).$$

The fixed iteration count is independent of the density-dependent asymptotic onset of this event.

Now compare the actual finite-kernel map $\mathcal{F}_q$ with $\mathcal{F}_k$ along the exact-kernel trajectory. All scaling vectors on that trajectory are bounded above and below on the good event, and $H_t\mu_n$ is uniformly bounded below. Equation (23) therefore gives $\|\mathcal{F}_q(u_j) - \mathcal{F}_k(u_j)\|_\infty \leq Cn^{-A}$. The finite-kernel map is also sup-norm nonexpansive. Induction over all actual steps yields path difference at most CJn^{-A}. Since $A > B + 3$, this proves (25). □

5 Complete inverse and covariance return

Lemma 5.1 (Reduced inverse accuracy)**.** *On an event of probability tending to one, the Cholesky test succeeds at every band. In original function coordinates, write $P_0 = P_* - \Pi_w$ and use S for its trial counterpart under the isometry $W^{1/2}$. For $\delta_0 = Cn^{-B}$,*

$$\begin{aligned} \|S - P_0\|_{\mathrm{op}} &\leq \delta_0, \\ \|\widetilde{R}_- - R_-\|_{\mathrm{op}} \leq C\delta_0 t^{-2}, &\qquad \|\widetilde{R}_+ - R_+\|_{\mathrm{op}} \leq C\delta_0, \\ \|\widetilde{C} - C_*\|_{\mathrm{op}} &\leq C\delta_0. \end{aligned} \tag{26}$$

The resolvents in the second line are compared on centered inputs; the response in the last line vanishes on constants.

Proof. Before projection, the finite scaled Gram equals DP_*D plus a kernel error, with $D = \mathrm{diag}(b_J/a_{t,n}) = I + O_{\mathbb{P}}(n^{-B})$. The bounded scales and the weighted Schur test bound that kernel error by Cn^{-A}. Orthogonal projection has norm one, giving the first line. Also $S \succeq 0$, and Lemma 4.2 gives $\|S\| \leq 1 - ct/4$ on the good event. Nonzero eigenvalues of VV^T and V^TV agree, so $I - H$ is positive definite. The Woodbury identities (20) are therefore valid.

The resolvent identity pays two minus inverses, each of norm $O(t^{-1})$, and two bounded plus inverses. On the mean-zero space this proves the second line of (26). The exact zero-on-constants response is $2(I + P_0)^{-1} - I - \Pi_w$; hence its error is bounded by the plus-resolvent error. This also verifies the last line. □

Lemma 5.2 (Metric and full influence errors)**.** *At each band, for every fixed scalar score and its centered exact and trial influences,*

$$\begin{aligned} |\widetilde{\Phi}_t - \Phi_t| &\leq C\delta_0/t, \\ \|\widetilde{\psi}_t - \psi_t\|_{L^2(w)} &\leq C\sqrt{n}\,\delta_0 t^{-2} \end{aligned} \tag{27}$$

on the good event. The resulting original-law joint covariance error is $o_{\mathbb{P}}(1)$.

Proof. The metric bound follows from the prefactor $t/2$, a bounded centered score and the minus-resolvent error. For the influence write all quantities in $L^2(w)$. Exact and trial minus outputs have norm $O(t^{-1})$, their difference is $O(\delta_0 t^{-2})$, and plus outputs are bounded with difference $O(\delta_0)$. For every vector $\|z\|_\infty \leq \sqrt{n}\|z\|_{L^2(w)}$.

For pointwise products the exact identity

$$zPz - xSx = (z - x)Pz + xP(z - x) + x(P - S)x$$

pays both output changes and the operator change. Applied to $U = z_- P_* z_- + z_+ P_* z_+$ and its trial version it gives

$$\|U - \widetilde{U}\|_2 \le C\sqrt{n}\,\delta_0 t^{-3}, \qquad \|U\|_2 + \|\widetilde{U}\|_2 \le C\sqrt{n}\,t^{-2}.$$

Together with (26) and the prefactor $t/2$, this proves the second line of (27). Centering at the same law is an $L^2(w)$ contraction.

The exact influence norms are $O_{\mathbb{P}}(1)$ by Proposition 2.2. For two scalar or polarized entries, the covariance difference is bounded by

$$\|\widetilde{\psi}_a - \psi_a\|_2\|\psi_b\|_2 + \|\widetilde{\psi}_b - \psi_b\|_2\|\psi_a\|_2 + \|\widetilde{\psi}_a - \psi_a\|_2\|\widetilde{\psi}_b - \psi_b\|_2.$$

This includes the two linear mixed terms and their product. Since $\sqrt{n}\,n^{-B}h_n^{-2} \to 0$, it proves covariance consistency after the fixed polarization. Every pairing uses the original law w. □

Proof of Theorem 2.1. The fixed multiband metric error is $O_{\mathbb{P}}(n^{-B}h_n^{-1})$ by Lemma 5.2, so it is $o_{\mathbb{P}}(n^{-1/2})$. The same lemma and the exact covariance input prove (10). Failure has probability tending to zero by Lemma 5.1, so the specified fallback leaves these convergences unchanged. The exact estimator's asymptotic linear representation and Slutsky's lemma give (11) and the stated studentization.

For completeness the operation budget is

Operation at one band	Work
Real feature formation and storage	$O(nR)$
J actual scaling multiplications	$O(nR\log n)$
Projection and feature Gram V^TV	$O(nR^2)$
Two reduced Cholesky factorizations	$O(R^3)$
Fixed metric and influence right-hand sides	$O(nR + R^2)$
Fixed joint covariance	$O(n)$

The projection is performed as $V_0 - e(e^TV_0)$, and S is applied as $V(V^Tx)$; neither needs a dense sample-by-sample matrix. There are only K fixed bands. The stated memory and deterministic work bounds follow, including on fallback samples. Now $nR_n^2 = O(n^{1+\beta d}(\log n)^d)$ and $R_n^3 = O(n^{3\beta d/2}(\log n)^{3d/2})$. Equation (6) absorbs both powers and logarithms into $O(n^{1+\eta})$.

For a bootstrap law μ_n^*, conditional Hilbert-space variance and the original empirical bound give $\mathbb{E}\|\mu_n^* - \mu\|_{H^{-s}}^2 \le C/n$. The heat feature-covariance discrepancy likewise has joint squared mean $O(n^{-1}t^{-d})$: condition on the original sample for the bootstrap variance and add the original sampling variance. Every preceding good-event argument therefore holds under the joint original/bootstrap law. Merged weights remain at least $1/n$. Consequently $\sqrt{n}(\widetilde{G}_n^* - \widehat{G}_n^*) \to 0$ jointly in probability, as does the original numerical error. For each $\epsilon > 0$, Markov's inequality applied to the conditional error probability turns this into conditional negligibility in original-sample probability. The exact bootstrap input then gives (12). Each replicate repeats the same finite budget. □

6 An explicit circle target and reproducible algebra

The uniform circle gives a direct normalization example. Let $d = 1$, $f = 1$ and $u(x) = \cos(2\pi x)$. Then $D_f = -\partial_x^2$, $\phi = u/(4\pi^2)$ and

$$G^0 = \frac{1}{8\pi^2}, \qquad \psi^0(x) = \frac{3\cos(4\pi x)}{8\pi^2}, \qquad \Sigma = \frac{9}{128\pi^4}. \tag{28}$$

Indeed $\mu u^2 = 1/2$ and $2u\phi - |\phi'|^2 - G^0 = 3(2\cos^2(2\pi x) - 1)/(8\pi^2)$. At finite population bandwidth, $P_t u = e^{-2\pi^2 t} u$ and hence

$$G(t, \mu) = \frac{t}{4\sinh(2\pi^2 t)}. \tag{29}$$

This gives the target, covariance and even-bias normalization in closed form.

The accompanying Python reference implementation follows (17)–(22), including merged multiplicities, the constant projector and conversion to function coordinates. Its optional dense comparison checks the same finite-feature metric and influence actions by a second linear-algebra route. The analytic proof above establishes the asymptotic and operation claims; the floating-point comparisons illustrate the finite formulas. The program prints its output and writes a file only when the caller supplies an explicit output path.

AI assistance. AI tools assisted mathematical exploration, verification code, literature comparison and manuscript preparation. The author is responsible for the stated results and references.

References

[1] J. Altschuler, F. Bach, A. Rudi and J. Niles-Weed, Massively scalable Sinkhorn distances via the Nyström method, *Advances in Neural Information Processing Systems* **32** (2019). `https://arxiv.org/abs/1812.05189v3`.

[2] H. Dong, J. Li, Y. Gong, R. Li and Y. Liu, Sinkhorn linearization and the spectral proxy: unifying the statistical and algorithmic theory of feature-parameterized inverse optimal transport via a single spectral sandwich, preprint (2026). `https://arxiv.org/abs/2608.13201v1`.

[3] N. Kessler, R. Magnet and J. Feydy, Sinkhorn normalization of diffusion kernels, *Proceedings of Machine Learning Research* **306** (2026), 56764–56795. `https://proceedings.mlr.press/v306/kessler26a.html`.

[4] P. A. Knight, D. Ruiz and B. Uçar, A symmetry preserving algorithm for matrix scaling, *SIAM J. Matrix Anal. Appl.* **35** (2014), 931–955. doi:10.1137/110825753.

[5] R. Lakshmanan, A. Pichler and D. Potts, Nonequispaced fast Fourier transform boost for the Sinkhorn algorithm, preprint, revised March 4, 2023. `https://arxiv.org/abs/2201.07524v5`.

[6] T. L. Saidi, F. Gunsilius and G. Mena, Second-order Sinkhorn geometry and $\sqrt{n}$-inference, preprint (2026). `https://arxiv.org/abs/2610.36488v1`.

[7] S. Schubr, Statistical inference for empirical heat-kernel Sinkhorn metrics, Luminary Research Archive, edition 3 (2026). `https://research.luminaryoddities.com/statistical-inference-empirical-heat-kernel-sinkhorn-metrics/v/3`.