

ON THE ASYMPTOTIC W_1 COST OF TWO-DIMENSIONAL SEMI-DISCRETE MATCHING

BY

GPT-5.6 PRO AND HENG MA

ABSTRACT

Let $X_1, X_2, \dots$ be independent uniform points in the unit square and let λ denote Lebesgue probability measure. We prove that the normalized expected W_1 distance between the empirical measure and Lebesgue measure converges to a positive constant:

$$\lim_{N \rightarrow \infty} \sqrt{\frac{N}{\log N}} \mathbb{E} W_1 \left(\frac{1}{N} \sum_{i=1}^N \delta_{X_i}, \lambda \right) = c_\star.$$

The constant is characterized by the convex viscosity equation

$$\partial_t U = \frac{1}{4\pi} \sqrt{\det D^2 U}, \quad U(0, q) = |q|.$$

Its solution U is unique under certain growth and regularity conditions, and the limiting constant $c_\star = U(1, 0)/\sqrt{2}$.

AI USAGE

This manuscript was substantially generated and subsequently checked by AI systems. At the time of this version, the arguments and citations have not yet been independently verified in full by the human contributor. **A digested and more streamlined version will replace this preliminary version on both Hexagon and arXiv in the coming months.**

CONTENTS

1. Introduction	2
2. Transport formulation and Gaussian reduction	5
3. Dyadic decomposition and scale grouping	9
4. The nonlinear cell operator	11
5. A quantitative two-scale comparison	18
6. Analytic consistency estimates	27
7. The convex viscosity limit	37
8. Completion of the Gaussian and empirical limits	43
References	46

Date: September 30, 2026.

Only Heng Ma takes responsibility for this paper.

1. Introduction

Let $X_1, X_2, \dots$ be independent uniform points in $[0, 1]^2$, let λ be Lebesgue probability measure on this square, and write $\mu_N = N^{-1} \sum_{i=1}^N \delta_{X_i}$. We study the leading constant in the expected transport distance from μ_N to λ . The Ajtai–Komlós–Tusnády matching theory and Talagrand’s general-measure formulation give the critical estimate

$$\mathbb{E}W_1\left(\sum_{i=1}^N \delta_{X_i}, N\lambda\right) \asymp \sqrt{N \log N};$$

see Talagrand [Tal92]. This determines the order of the expectation but leaves open whether its ratio to $\sqrt{N \log N}$ converges. We prove convergence and characterize the limit by a deterministic evolution equation. The theorem is stated in the equivalent empirical-process formulation; the equivalence with $N\mathbb{E}W_1(\mu_N, \lambda)$ is recorded in Lemma 2.2.

THEOREM 1.1 (Main result). *Let $X_1, X_2, \dots$ be independent random variables uniformly distributed on $[0, 1]^2$, and let λ be Lebesgue probability measure on this square. There is a finite positive constant $c_\star$ such that*

$$\lim_{N \rightarrow \infty} \frac{1}{\sqrt{N \log N}} \mathbb{E} \sup_{\text{Lip}(f) \leq 1} \left| \sum_{i=1}^N \left(f(X_i) - \int_{[0,1]^2} f d\lambda \right) \right| = c_\star.$$

Moreover, $c_\star = \kappa/\sqrt{2}$, where $\kappa = U(1, 0)$ and U is the unique viscosity solution in the convex class specified in Subsection 1.2 of

$$\partial_t U = \frac{1}{4\pi} \sqrt{\det D_q^2 U}, \quad U(0, q) = |q|.$$

The equation characterizes $c_\star$ but does not give a numerical value. We outline the reduction from transport to this equation, then specify the solution class used in the theorem.

1.1. PROOF STRATEGY. Write A_N for the expectation in Theorem 1.1. The proof uses the flux formulation of transport. For a balanced source σ on a square Q , it reads

$$\|\sigma\|_{\text{KR}} = \inf_{-\text{div } v = \sigma, v \cdot n = 0} |v|(Q).$$

For an ordinary vector field, $|v|(Q) = \int_Q |v(x)| dx$. The divergence constraint specifies the source transported by v , and the boundary condition prevents flux from leaving Q . Proposition 2.1 gives the precise formulation for vector measures and its dual formulation with Lipschitz potentials.

Ambrosio, Stra, and Trevisan [AST19, Sec. 6] suggested another possible PDE approach to the case $p = 1$, based on the q -Laplacian approximation of Evans and Gangbo [EG99]. For sufficiently regular source and target densities $f^\pm$, this approximation solves $-\text{div}(|\nabla u_q|^{q-2} \nabla u_q) = f^+ - f^-$ and lets $q \rightarrow \infty$, yielding a 1-Lipschitz potential u and a transport density a with $-\text{div}(a \nabla u) = f^+ - f^-$. That proposal was not carried out there for the empirical W_1 asymptotic constant. Here we use the flux formulation to track how successive Gaussian scales change the convex cost of a background flux. The elliptic problem in the quadratic response below is linear; the argument does not take a q -Laplacian limit or construct a transport density.

Fluxes can be added when sources are added, but the cost of their sum depends on their directions. Thus the optimal costs of the individual scales do not determine the cost of

their sum. At each scale we instead retain a convex function of the flux supplied by the coarser scales.

From empirical sampling to a Gaussian field. For a $K \times K$ grid in the unit square, let c_z be the cell centers, let G_z be independent standard Gaussians, and set

$$\Gamma_K = \mathbb{E} \left\| \sum_z (G_z - \bar{G}) \delta_{c_z} \right\|_{\text{KR}}, \quad \bar{G} = K^{-2} \sum_z G_z.$$

Poissonization makes the cell counts independent. Coupling these counts to Gaussians and transporting the balanced coupling error gives Lemma 2.4:

$$\left| A_N - \sqrt{N/K^2} \Gamma_K \right| \leq C \left(\sqrt{N} + \frac{N}{K} + K \sqrt{\log(2K)} \right).$$

For a dyadic choice $K_N \asymp \sqrt{N}/(\log N)^a$, with $0 < a < 1/2$, these errors are negligible compared with $\sqrt{N \log N}$. Thus it suffices to prove

$$\frac{\Gamma_{2^J}}{2^J \sqrt{J \log 2}} \longrightarrow \kappa. \quad (\text{R})$$

The relation $\log K_N / \log N \rightarrow 1/2$ then accounts for the factor $1/\sqrt{2}$ in the empirical constant.

Removing fine scales changes the cost of the coarse flux. After dilation to $Q_{2^J} = [0, 2^J]^2$ and spreading each Gaussian mass over its unit cell, we obtain a balanced density ρ_{2^J} . It is a sum of J independent Haar details, each recording fluctuations between dyadic children and their parent. The expected transport cost of spreading the masses is $O(2^{2J})$, so the Gaussian target becomes

$$\frac{\mathbb{E} \|\rho_{2^J}\|_{\text{KR}}}{2^{2J} \sqrt{J \log 2}} \longrightarrow \kappa.$$

Suppose that the current cost of a flux $p \in \mathbb{R}^2$ is $\Psi(p)$. On a region where the coarse flux is almost constant, a fine source is transported by a correction u , and the cost is applied to $p + u$. Optimizing the fine correction and averaging over the fine source therefore produces a new function of p , rather than a single number. We denote this update by $\mathcal{H}_{h,L} \Psi$, where h is the number of fine levels being removed and L is an auxiliary cell period. Definition 4.1 constructs the update through the dual potential problem with compatible boundary values.

Under the scale and alignment assumptions of Lemma 5.1, the comparison is

$$\mathbb{E} J_\Psi^Q(\eta + \zeta) \approx \mathbb{E} J_{\mathcal{H}_{h,L} \Psi}^Q(\zeta).$$

Here J_Ψ^Q is the transport cost with integrand Ψ , η is the fine source, and ζ is the independent coarse source. The error per unit area in this comparison is bounded by (24). Starting from the Euclidean cost, repeated replacement gives

$$\Phi_0(p) = |p|, \quad \Phi_{k+1} = \mathcal{H}_{h,L} \Phi_k.$$

After the last source is removed, the remaining flux has zero divergence and zero normal trace, hence zero mean. Convexity then makes its minimum cost per unit area exactly $\Phi_m(0)$. This explains why the original random optimization can be reduced to evaluating a deterministic iteration at the origin.

Scale separation and common boundary values justify the replacement. Two issues must be controlled in the preceding reduction. The coarse flux must vary little on a fine cell, and independently optimized cell interiors must fit together. To obtain the required separation, we keep groups of h consecutive levels and omit b levels between groups. Section 3 constructs a centered linear flux for each Haar detail, with zero normal trace and uniformly bounded pointwise second moment. Independence gives an expected cost at most $C2^{2J}\sqrt{D}$ for any D omitted levels. Consequently, omitting $D = o(J)$ levels does not affect the desired limit.

For the boundary issue, prescribing a mean slope alone does not specify the values of a potential along a shared face. The cell operator retains the full boundary function, including the compatibility of opposite faces. It chooses that boundary function after taking expectation, while the optimized interior may depend on the fine source. The original optimizer can choose its boundary values using the fine source as well. Lemma 4.3 controls this additional freedom by a finite approximation of the boundary functions and Gaussian concentration. For the upper bound, the proof smooths the coarse flux, approximates it by a constant on each block, and adds fine fluxes with zero normal trace there. For the lower bound, it applies the same blockwise approximation to a dual potential and pastes the fine potentials using their common boundary values.

The quadratic response determines the equation. To identify the evolution of Φ_k , consider first a smooth convex cost near a background flux, with positive-definite Hessian H . The quadratic model for a unit outer block is

$$\mathbb{E}_{-\operatorname{div} u = \eta_h, u \cdot n = 0} \inf_{u \cdot n = 0} \frac{1}{2} \int_{Q_1} u^T H u = \frac{1}{2} \operatorname{Tr}(\Pi_h(L_H^N)^{-1}), \quad L_H = -\operatorname{div}(H^{-1}\nabla).$$

Here η_h is the normalized h -level Gaussian band, Π_h is its covariance projection, and L_H^N has Neumann boundary conditions. The inverse matrix H^{-1} appears because the dual quadratic cost is $\frac{1}{2} \nabla f^T H^{-1} \nabla f$. In two dimensions the inverse elliptic eigenvalues have a logarithmic sum. The covariance projection cuts this sum off at spatial resolution 2^{-h} , and the elliptic eigenvalue count contributes $\sqrt{\det H}$. Lemma 6.2 computes the coefficient:

$$\frac{1}{2} \operatorname{Tr}(\Pi_h(L_H^N)^{-1}) = \frac{h \log 2}{4\pi} \sqrt{\det H} + o(h).$$

The Dirichlet quadratic problem has the same leading term. These two problems bound the quadratic cell value, so the coefficient is independent of the auxiliary period L .

The calculation suggests a time increment $h \log 2$ per group and a generator $\mathcal{G}(H) = (4\pi)^{-1} \sqrt{\det H}$. To justify this evolution, Section 6 derives one-sided estimates on smooth tests. The upper estimate uses a Neumann flux. The lower estimate requires a Dirichlet potential whose gradient respects the Lipschitz constraint, followed by a global convex barrier that makes a local contact usable in the cell operator. These estimates are formulated for viscosity tests because the iterated costs need not be smooth.

Diffusive rescaling and the accumulated error. The stability estimate $|p| \leq \Phi_k(p) \leq |p| + C\sqrt{k}h$ suggests rescaling both the cost and its flux variable by the square root of time. For a large time scale r , the relevant function is

$$r^{-1/2} \Phi_{\lfloor sr/(h \log 2) \rfloor}(\sqrt{r} q).$$

For the spatial tests ψ covered by Section 6, the rescaled one-sided estimates have leading term $(h \log 2/r) \mathcal{G}(D_q^2 \psi)$, with time step $h \log 2/r$. In the joint regime $h \rightarrow \infty$, $L \rightarrow \infty$, and $h/r \rightarrow 0$, Lemma 7.1 identifies its limit with $U(s, q)$. The consistency estimates give

the equation; comparison in the convex class makes the limit unique. In particular, at the retained time $t_J = mh \log 2 \sim J \log 2$, this yields $\Phi_m(0)/\sqrt{t_J} \rightarrow U(1, 0)$.

The scale choices must make this limit compatible with all earlier replacements. If e_J is the error per unit area in one replacement, the required conditions are

$$D = o(J), \quad me_J = o(\sqrt{J}), \quad h/t_J \rightarrow 0.$$

Section 8 uses $h = \lfloor J^{3/4} \rfloor$ and $b = \lceil (\log J)^2 \rceil$, together with intermediate geometric scales, to verify these conditions in one sequence. Thus

$$2^{-2J} \mathbb{E} \|\rho_{2J}\|_{\text{KR}} = \Phi_m(0) + o(\sqrt{J}),$$

which gives (R) and then the empirical limit.

1.2. VISCOSITY FORMULATION AND NOTATION. The solution class consists of continuous functions $U : [0, \infty) \times \mathbb{R}^2 \rightarrow \mathbb{R}$ such that $U(s, \cdot)$ is convex and 1-Lipschitz for every s , and $U(\cdot, q)$ is nondecreasing for every q . For each finite S ,

$$\sup_{0 \leq s \leq S, q \in \mathbb{R}^2} |U(s, q) - |q|| < \infty, \quad \sup_{q \in \mathbb{R}^2} |U(s, q) - U(t, q)| \leq \omega_S(|s - t|) \quad (0 \leq s, t \leq S),$$

where $\omega_S(\delta) \rightarrow 0$ as $\delta \downarrow 0$. We use the strong convex viscosity formulation: every smooth upper test satisfies the subsolution inequality, and every smooth lower test with positive-semidefinite spatial Hessian satisfies the supersolution inequality, including when that Hessian is singular. Lemma 7.1 records the precise formulation and proves comparison in this class.

Notation. We write $Q_L = [0, L]^2$, so $|Q_L| = L^2$, and all vector norms are Euclidean. Constants denoted by C may change from line to line. A vector measure on a square has zero normal trace when its extension by zero has no boundary divergence measure; this convention is made precise in Proposition 2.1. A source is *balanced* on a square if its total mass there is zero. We write $B_r(a)$ for the Euclidean ball of radius r centered at a , set $B_r := B_r(0)$, and write $\mathbb{S}_+^2$ for the cone of symmetric positive-semidefinite 2×2 matrices.

2. Transport formulation and Gaussian reduction

We first express the empirical process as a transport cost and then compare it with a Gaussian grid. The comparison keeps track of the grid size, so the eventual Gaussian limit will imply convergence for every integer N .

2.1. TRANSPORT COSTS AND FLUXES.

PROPOSITION 2.1 (Transport costs and duality). *For a finite signed measure σ of total mass zero on a bounded convex set $Q \subset \mathbb{R}^2$, put*

$$\|\sigma\|_{\text{KR}} := \sup_{\text{Lip}(f) \leq 1} \int_Q f \, d\sigma.$$

Let $\mathcal{M}(Q; \mathbb{R}^2)$ be the vector Radon measures on Q . For $v \in \mathcal{M}(Q; \mathbb{R}^2)$, denote its extension by zero to $\mathbb{R}^2$ by $\tilde{v}$. We write

$$-\text{div } v = \sigma, \quad v \cdot n = 0,$$

to mean the single distributional identity $-\operatorname{div} \tilde{v} = \tilde{\sigma}$ on $\mathbb{R}^2$. Thus no boundary measure is present; this is the definition of zero normal trace used below. Then

$$\|\sigma\|_{\text{KR}} = \inf \{ |v|(Q) : -\operatorname{div} v = \sigma, \quad v \cdot n = 0 \text{ on } \partial Q \}. \quad (1)$$

More generally, for a convex function $\Psi : \mathbb{R}^2 \rightarrow \mathbb{R}$ with recession function $|\cdot|$, define

$$J_{\Psi}^Q(\sigma) := \inf_{-\operatorname{div} v = \sigma, \, v \cdot n = 0} \left\{ \int_Q \Psi(v^a(x)) \, dx + |v^s|(Q) \right\},$$

where $v = v^a \, dx + v^s$ is the Lebesgue decomposition. This is the relaxation of the integral with linear growth. Its dual form is

$$J_{\Psi}^Q(\sigma) = \sup_{f \in C^1(\bar{Q})} \left\{ \int_Q f \, d\sigma - \int_Q \Psi^*(\nabla f) \right\}. \quad (2)$$

The same value is obtained with $W^{1,\infty}(Q)$ potentials.

Proof. We prove (2) by separating the relaxed primal epigraph and then take $\Psi(z) = |z|$ to obtain (1). Working with vector measures retains singular fluxes and preserves the zero-normal-trace constraint under weak-* limits. The required closedness is verified next.

Set

$$I(v) = \int_Q \Psi(v^a) \, dx + |v^s|(Q).$$

The epigraph

$$\mathcal{E} = \{(-\operatorname{div} \tilde{v}, r) : v \in \mathcal{M}(Q; \mathbb{R}^2), \, I(v) \leq r\}$$

is convex and closed for distributional convergence in the first coordinate and ordinary convergence in the second. Indeed a bounded energy sequence is weak-* compact as measures, the divergence constraint passes to the limit, and the displayed relaxed integral is weak-* lower semicontinuous (this last assertion follows directly by writing it as the supremum of $\int a \cdot dv - \int \Psi^*(a)$ over continuous a).

Separate $(\tilde{\sigma}, r)$, with $r < \inf I$, from $\mathcal{E}$. The separating distribution is a smooth test function f , and the coefficient of the epigraph coordinate is positive; normalize it to one. Minimizing the resulting affine functional over absolutely continuous and singular parts of v gives respectively $-\int \Psi^*(\nabla f)$ and the constraint $|\nabla f| \leq 1$. This proves the reverse inequality in (2); integration by parts gives the forward inequality. Approximation of Lipschitz functions by smooth ones on a convex square gives the $W^{1,\infty}$ version. Taking $\Psi(z) = |z|$, for which Ψ^* is the indicator of the closed unit ball, proves (1). ■

LEMMA 2.2 (The empirical process as a transport cost). *Let*

$$\nu_N = \sum_{i=1}^N \delta_{X_i}, \quad \sigma_N = \nu_N - N\lambda,$$

and let A_N denote the expectation in Theorem 1.1. Then

$$A_N = \mathbb{E} \|\sigma_N\|_{\text{KR}} = \mathbb{E} W_1(\nu_N, N\lambda). \quad (3)$$

Proof. For a fixed realization, the displayed empirical process is $\int f \, d\sigma_N$. The Lipschitz class is invariant under $f \mapsto -f$, so the absolute value does not change the supremum. The last equality is the equal-mass Kantorovich–Rubinstein duality. ■

2.2. COMPARISON WITH THE GAUSSIAN GRID.

DEFINITION 2.3 (Balanced Gaussian grid). Partition $[0, 1]^2$ into a $K \times K$ grid, let c_z be the cell centres, let G_z be independent standard Gaussians, and put

$$\bar{G} = K^{-2} \sum_z G_z, \quad \Gamma_K = \mathbb{E} \left\| \sum_z (G_z - \bar{G}) \delta_{c_z} \right\|_{\text{KR}}. \quad (4)$$

Subtraction of $\bar{G}$ makes the total mass zero.

LEMMA 2.4 (Quantitative Gaussian approximation). *Uniformly for $K^2 \leq N$,*

$$\left| A_N - \sqrt{\frac{N}{K^2}} \Gamma_K \right| \leq C \left(\sqrt{N} + \frac{N}{K} + K \sqrt{\log(2K)} \right). \quad (5)$$

Consequently, it is enough to prove the existence of

$$\lim_{J \rightarrow \infty} \frac{\Gamma_{2^J}}{2^J \sqrt{J \log 2}}. \quad (6)$$

If this dyadic limit is γ , then

$$\lim_{N \rightarrow \infty} \frac{A_N}{\sqrt{N \log N}} = \frac{\gamma}{\sqrt{2}}. \quad (7)$$

Proof. Poissonization and snapping produce independent cell counts at costs $O(\sqrt{N})$ and $O(N/K)$. After coupling these counts to Gaussians, we construct a grid flow for the balanced error; its energy controls the third term in (5). The final choice of K makes all three errors negligible on the scale $\sqrt{N \log N}$.

Step 1. Poissonize and snap to the grid.

Write $\mu_n = \sum_{i=1}^n (\delta_{X_i} - \lambda)$. Couple $M \sim \text{Pois}(N)$ to the same infinite sample sequence. Since $\|\delta_x - \lambda\|_{\text{KR}} \leq \text{diam } Q_1$,

$$\mathbb{E} \|\mu_N - \mu_M\|_{\text{KR}} \leq C \mathbb{E} |M - N| \leq C \sqrt{N}. \quad (5a)$$

Let $\lambda_K = K^{-2} \sum_z \delta_{c_z}$, and let $c(x)$ be the centre of the grid cell containing x . Replacing each $\delta_{X_i} - \lambda$ by $\delta_{c(X_i)} - \lambda_K$ costs at most C/K . Thus the expected snapping error in the Poissonized process is CN/K .

Step 2. Couple the Poisson counts to Gaussians.

The cell counts P_z are now independent $\text{Pois}(q)$ variables, $q = N/K^2$, while the balanced snapped measure is

$$\sum_z (P_z - \bar{P}) \delta_{c_z}, \quad \bar{P} = K^{-2} \sum_z P_z. \quad (5b)$$

Put $Y_z = P_z - q$. Couple the Y_z 's independently to $\sqrt{q} G_z$ with

$$\mathbb{E} |P_z - q - \sqrt{q} G_z|^2 \leq C. \quad (8)$$

We use the following normal-approximation estimate to obtain (8).

Normal-approximation result. If $V_n = \sum_i \xi_i$, the ξ_i are independent and centred, and $\mathbb{E} V_n^2 = 1$, then for every $p > 1$

$$W_p(\mathcal{L}(V_n), N(0, 1)) \leq C_p \left(\sum_i \mathbb{E} |\xi_i|^{p+2} \right)^{1/p}.$$

Here C_p depends only continuously on p .

This is Lemma 3.4 of Fang [Fan19], which restates Bobkov's Theorem 1.1. To check the hypotheses without a limiting argument, let $r = \lceil q \rceil$, $\theta = q/r \in [1/2, 1]$, and use infinite divisibility to write P_q as a sum of r independent $\text{Pois}(\theta)$ variables. After normalization to variance one, the sum of fourth moments appearing in the $p = 2$ estimate is $(1 + 3\theta)/q \leq 4/q$. Hence $W_2((P_q - q)/\sqrt{q}, N(0, 1)) \leq C/\sqrt{q}$, which is exactly (8), uniformly for $q \geq 1$.

Let $e_z = Y_z - \sqrt{q}G_z$ and $\tilde{e} = e - \bar{e}\mathbf{1}$. The couplings are independent and centred, so $\text{Cov}(e) \leq CI$ and $\text{Cov}(\tilde{e}) \leq C\Pi_0$, where Π_0 is orthogonal projection off the constants. Notice that projection is exactly what turns Y_z into $P_z - \bar{P}$ and G_z into $G_z - \bar{G}$; it introduces no unbalanced measure.

Step 3. Transport the balanced coupling error.

The errors in (8) have uniformly bounded variances. Summing their absolute values over K^2 cells would discard cancellation. We instead use the discrete inverse Laplacian to construct a flow whose divergence is exactly the balanced error.

Give the $K \times K$ nearest-neighbor grid unit edge weights and let $B : \mathbb{R}^E \rightarrow \mathbb{R}^{K^2}$ be its signed incidence matrix, $L = BB^T$. The flow

$$m = B^T L^+ \tilde{e}$$

satisfies $Bm = \tilde{e}$. Sending $|m_e|$ units of mass along the corresponding Euclidean edge of length $1/K$ gives

$$\left\| \sum_z \tilde{e}_z \delta_{c_z} \right\|_{\text{KR}} \leq K^{-1} \sum_{e \in E} |m_e|. \quad (5c)$$

Moreover

$$\mathbb{E} \|m\|_2^2 = \text{Tr}(B^T L^+ \text{Cov}(\tilde{e}) L^+ B) \leq C \text{Tr} L^+.$$

For $0 \leq r, s < K$ with $(r, s) \neq (0, 0)$, the nonzero eigenvalues of the path-grid Neumann Laplacian are

$$\lambda_{r,s} = 4 \sin^2\left(\frac{\pi r}{2K}\right) + 4 \sin^2\left(\frac{\pi s}{2K}\right),$$

including when one index is zero. Therefore

$$\text{Tr} L^+ \leq CK^2 \sum_{0 \leq r, s < K, (r,s) \neq (0,0)} (r^2 + s^2)^{-1} \leq CK^2 \log(2K). \quad (5d)$$

Since $|E| \leq 2K^2$, Cauchy-Schwarz in (5c) yields expected KR error at most $CK \sqrt{\log(2K)}$. Together with (5a) and the snapping bound this proves (5).

Step 4. Choose the dyadic grid scale and return to A_N .

Fix $0 < a < 1/2$. For every N , choose a dyadic integer K_N such that

$$K_N \leq \frac{\sqrt{N}}{(\log N)^a} < 2K_N.$$

All three errors in (5) are $o(\sqrt{N \log N})$, while $\log K_N / \log N \rightarrow 1/2$. Substitution of (6) into (5) therefore proves (7), completing the reduction from the empirical problem to the dyadic Gaussian limit.

■

2.3. FROM GAUSSIAN ATOMS TO A CELLWISE DENSITY. The next comparison puts the Gaussian field on a square with unit cells, where its dyadic decomposition is convenient. The resulting error is of lower order than the target cost $K^2\sqrt{\log K}$.

LEMMA 2.5 (Spreading the Gaussian masses over cells). *Let $\widehat{\Gamma}_K$ be the Gaussian KR expectation after dilating the unit square and the grid by K . Thus*

$$\widehat{\Gamma}_K = K\Gamma_K.$$

Let $C_z = z + [0, 1)^2 \subset Q_K$ and set

$$\rho_K(x) = \sum_z (G_z - \bar{G}) \mathbf{1}_{C_z}(x).$$

Then

$$\left| \widehat{\Gamma}_K - \mathbb{E} \|\rho_K\|_{\text{KR}, Q_K} \right| \leq CK^2. \quad (9)$$

Proof. Dilation multiplies every transport distance by K , proving the first identity. For each cell, move its signed mass between the centre and the uniform measure on that cell. This costs at most $C|G_z - \bar{G}|$. Summing and taking expectations proves (9).

■

3. Dyadic decomposition and scale grouping

We group the Gaussian density into independent bands of dyadic scales. Each band has the same law after its outer blocks are rescaled to unit squares. To separate consecutive retained bands, we will omit some Haar levels; the estimate below shows that the resulting cost grows only as the square root of their number. Its proof constructs local fluxes whose pointwise second moments also control the later cell comparisons.

LEMMA 3.1 (Independent Haar bands and the cost of omitted levels). *Assume $K = 2^J$, and write $G = \sum_z G_z \mathbf{1}_{C_z}$ for the unit-cell Gaussian field. On cellwise-constant functions in Q_K , let P_j be conditional expectation on dyadic squares of side 2^j . Then*

$$\rho_K = \sum_{j=0}^{J-1} \xi_j, \quad \xi_j = (P_j - P_{j+1})G, \quad (10)$$

and the ξ_j are independent Gaussian fields.

If $h \geq 1$, then a group $\eta_{r,h} = (P_r - P_{r+h})G$ is balanced on every dyadic square of side 2^{r+h} , and different such outer squares carry independent copies. After rescaling an outer square to the unit square, its source is

$$\eta_h = n \sum_{C \in \mathcal{P}_n} (Z_C - \bar{Z}) \mathbf{1}_C, \quad n = 2^h, \quad (11)$$

where $\mathcal{P}_n$ is the $n \times n$ partition into cells of side $1/n$, the Z_C are i.i.d. $N(0, 1)$, and $\bar{Z} = n^{-2} \sum_C Z_C$. In particular, the rescaled law depends on h , but not on r .

We call a square of side 2^{r+h} on which $\eta_{r,h}$ is balanced an outer block of the band. After rescaling, an outer block has side one and contains $2^h \times 2^h$ fine cells.

If D dyadic levels are omitted, their sum ζ_D satisfies

$$\mathbb{E} \|\zeta_D\|_{\text{KR}, Q_K} \leq CK^2 \sqrt{D}. \quad (12)$$

Proof. By (1), it suffices to construct a flux for the omitted levels with expected L^1 norm at most $CK^2\sqrt{D}$. We construct a zero-normal flux for each Haar detail and use independence to add its second moment across levels. First we identify the law of each band.

Step 1. Decompose the field into independent dyadic bands.

The spaces $\text{ran}(P_j - P_{j+1})$ are mutually orthogonal. Orthogonal projections of a Gaussian vector are independent, proving (10). The average of $(P_r - P_{r+h})G$ on each square of side 2^{r+h} is zero. Its restrictions to distinct such squares depend on disjoint sets of the independent unit-cell Gaussians, so they are independent. Within one outer square, each square of side 2^r has $P_r G = 2^{-r} Z_C$, with independent standard Gaussians Z_C . Hence $\eta_{r,h} = 2^{-r}(Z_C - \bar{Z})$ on that square. Under $x = 2^{r+h}y$, fluxes are unchanged in dimension two and source densities acquire the factor 2^{r+h} , giving (11).

Step 2. Construct a local right inverse on one Haar parent.

The local flux must solve the divergence equation and have matching normal components at every child interface; otherwise pasting would create an additional source there. Work on one level- j parent $B = x_0 + [0, 2\ell]^2$, $\ell = 2^j$. Denote its four child averages by $A_{\sigma\tau}$, where $\sigma = 1$ means left, $\tau = 1$ means lower, and $\sigma, \tau \in \{-1, 1\}$. They are independent $N(0, \ell^{-2})$. Put

$$\alpha = \frac{1}{4} \sum_{\sigma, \tau} \sigma A_{\sigma\tau}, \quad \beta = \frac{1}{4} \sum_{\sigma, \tau} \tau A_{\sigma\tau}, \quad \gamma = \frac{1}{4} \sum_{\sigma, \tau} \sigma\tau A_{\sigma\tau}.$$

Thus α, β, γ are independent and each has variance $(4\ell^2)^{-1}$. Let $s_\ell = 1$ on $[0, \ell)$ and -1 on $[\ell, 2\ell)$, and let

$$\tau_\ell(t) = \begin{cases} t, & 0 \leq t \leq \ell, \\ 2\ell - t, & \ell < t \leq 2\ell, \end{cases} \quad \tau'_\ell = s_\ell \quad \text{a.e.}$$

The Haar detail on B is

$$\xi_j = \alpha s_\ell(x_1 - x_{0,1}) + \beta s_\ell(x_2 - x_{0,2}) + \gamma s_\ell(x_1 - x_{0,1})s_\ell(x_2 - x_{0,2}),$$

and define

$$\begin{aligned} (R_j \xi_j)_1(x) &= -\tau_\ell(x_1 - x_{0,1})\{\alpha + \gamma s_\ell(x_2 - x_{0,2})\}, \\ (R_j \xi_j)_2(x) &= -\beta \tau_\ell(x_2 - x_{0,2}). \end{aligned} \tag{12a}$$

Direct differentiation gives $-\text{div } R_j \xi_j = \xi_j$. The first component vanishes on the two vertical parent faces and the second on the two horizontal parent faces, so the outer normal trace is zero. At the internal vertical child interface the normal first component is continuous; at the internal horizontal interface the normal second component is continuous, while only the tangential first component jumps. Hence there is no interface measure in the divergence.

Step 3. Paste the local fields and retain a pointwise bound.

Paste (12a) over all level- j parents. Across a common parent face both normal traces are zero, so the pasted field lies in $H(\text{div})$ and again creates no interface source. Choose the representative zero on the union of dyadic interfaces. Since $0 \leq \tau_\ell \leq \ell$, at every point

$$\mathbb{E}|R_j \xi_j(x)|^2 \leq \frac{3}{4}, \quad \mathbb{E} \int_{\text{parent}} |R_j \xi_j|^2 \leq C\ell^2. \tag{12b}$$

The Gaussian Haar projections at different levels are orthogonal and therefore independent; R_j is deterministic and linear. Consequently, for every set I of levels,

$$\sup_x \mathbb{E} \left| \sum_{j \in I} R_j \xi_j(x) \right|^2 = \sup_x \sum_{j \in I} \mathbb{E} |R_j \xi_j(x)|^2 \leq C|I|. \quad (12c)$$

The pointwise estimate (12c), including the chosen representatives at interfaces, is used in (16), (17), and (22). The construction is unchanged by dilation of an outer block: fluxes are scale invariant in dimension two.

Step 4. Sum the omitted levels and conclude the KR estimate.

For the D omitted levels, the same independence gives

$$\mathbb{E} \int_{Q_K} \left| \sum_{j \text{ omitted}} R_j \xi_j \right|^2 = \sum_{j \text{ omitted}} \mathbb{E} \int_{Q_K} |R_j \xi_j|^2 \leq CDK^2.$$

Jensen followed by Cauchy–Schwarz in space bounds the expected L^1 norm by $CK^2\sqrt{D}$. Formula (1) now gives (12), completing the proof.

■

4. The nonlinear cell operator

After the finer scales have been treated, their contribution is encoded by a convex cost profile $\Psi(p)$ for a base flux p . The operator defined below incorporates one more Haar band and returns the updated profile. We define it through the dual potential: first prescribe its values on all four cell faces, optimize the interior for each fine-source realization, and take expectation; only then optimize the boundary trace. This order keeps the trace deterministic with respect to the fine source.

4.1. DEFINITION AND BASIC PROPERTIES.

DEFINITION 4.1 (Cell operator). Let $\mathcal{P}_M$ be the class of convex functions $\Psi : \mathbb{R}^2 \rightarrow \mathbb{R}$ satisfying

$$|p| \leq \Psi(p) \leq |p| + M. \quad (13)$$

Every $\Psi \in \mathcal{P}_M$ is 1-Lipschitz, and

$$-M \leq \Psi^*(a) \leq 0 \quad (|a| \leq 1), \quad \Psi^*(a) = +\infty \quad (|a| > 1). \quad (14)$$

For an integer $L \geq 1$, tile Q_L by unit squares (the normalized outer blocks) and put an independent copy of (11) in each square; call the resulting source $\eta_h^{(L)}$. To retain the boundary data needed for pasting cell interiors, let $\mathcal{T}_L$ be the restrictions to all four faces of ∂Q_L of 1-Lipschitz functions on Q_L , modulo constants and represented by the normalization $c(0) = 0$. Since Q_L is convex, McShane extension shows that this is exactly the compact set of anchored 1-Lipschitz boundary functions.

The slope a will record the mean gradient of the dual potential. For $|a| \leq 1$, let $\mathcal{A}_L(a)$ be the functions $F \in W_{\text{loc}}^{1,\infty}(\mathbb{R}^2)$ satisfying

$$|\nabla F| \leq 1, \quad F(x + Le_i) = F(x) + La_i \quad (i = 1, 2), \quad F(0) = 0. \quad (14a)$$

Let $\mathcal{T}_L^{\text{per}}(a) \subset \mathcal{T}_L$ be the restrictions of these functions to all four faces of ∂Q_L . A trace $c \in \mathcal{T}_L^{\text{per}}(a)$ specifies the values on all four faces, including the opposite-face increments in

(14a) and their agreement at the corners. These data allow a cell extension to be repeated over the plane. For every $c \in \mathcal{T}_L$, let

$$K_{h,L,\Psi}(c) = \mathbb{E} \sup_{\substack{F|_{\partial Q_L=c} \\ |\nabla F| \leq 1}} \left\{ \int_{Q_L} F \eta_h^{(L)} - \int_{Q_L} \Psi^*(\nabla F) \right\}. \quad (15)$$

The source has zero total mass, so adding a constant to c leaves this value unchanged. The best expected cell value at slope a and the updated cost at base flux p are, respectively,

$$\mathcal{R}_{h,L,\Psi}(a) = L^{-2} \max_{c \in \mathcal{T}_L^{\text{per}}(a)} K_{h,L,\Psi}(c), \quad (\mathcal{H}_{h,L}\Psi)(p) = \sup_{|a| \leq 1} \{p \cdot a + \mathcal{R}_{h,L,\Psi}(a)\}. \quad (16)$$

When h, L, Ψ are fixed, we abbreviate $\mathcal{R} = \mathcal{R}_{h,L,\Psi}$.

Extension to a compact slope domain. Let Ψ be proper closed convex and suppose $K_\Psi = \text{dom } \Psi^* \subseteq B_1$ is compact and Ψ^* is finite and continuous on K_Ψ . In (15), retain the same ambient trace space but interpret the value as $-\infty$ unless the trace has an extension with $\nabla F(x) \in K_\Psi$ almost everywhere; equivalently the extended integrand $\Psi^* = +\infty$ off K_Ψ enforces this condition. In (16), set $\mathcal{R}(a) = -\infty$ for $a \notin K_\Psi$. These conventions define $\mathcal{H}_{h,L}\Psi = (-\mathcal{R})^*$. This extension is used only for the lower barrier with $|a_0| < 1$ in Lemma 7.1.

PROPOSITION 4.2 (Properties of the cell operator). *For $\Psi \in \mathcal{P}_M$, the map $\mathcal{H}_{h,L}$ is order preserving, commutes with addition of constants, is a contraction in the uniform norm, and maps convex 1-Lipschitz profiles to convex 1-Lipschitz profiles. Moreover,*

$$\Psi \leq \mathcal{H}_{h,L}\Psi \leq \Psi + C\sqrt{h}. \quad (16a)$$

The bound is uniform in L . If Ψ is invariant under the dihedral group D_4 of the square, so is $\mathcal{H}_{h,L}\Psi$. On the extension to a compact slope domain in Definition 4.1, the map is also order preserving and constant-equivariant.

Proof. We first establish the variational properties on $\mathcal{P}_M$ and bound the update by the expected size of a Haar flux. We then verify that the same definition remains well posed on a compact slope domain.

Step 1. Establish the variational properties on $\mathcal{P}_M$.

The bounds in (14) follow from (13) and convex conjugacy. A convex function whose recession function is $|\cdot|$ has all subgradients in the closed unit ball and is therefore 1-Lipschitz.

The space $\mathcal{T}_L$, and the closed set of pairs (a, c) , with $|a| \leq 1$ and $c \in \mathcal{T}_L^{\text{per}}(a)$, are compact in the uniform topology by Arzela–Ascoli. Repeating a trace and any one of its cell extensions produces a continuous piecewise $W^{1,\infty}$ function on the plane; its weak gradient has norm at most one, so it is globally 1-Lipschitz. Thus the opposite-face conditions are also sufficient, not only necessary. For fixed trace, the admissible interior potentials are also compact after normalization. The linear source term is uniformly continuous and $F \mapsto -\int \Psi^*(\nabla F)$ is upper semicontinuous under weak-* convergence in $W^{1,\infty}$. Anchoring F and using $|\nabla F| \leq 1$ and (14) bound the absolute payoff in (15) by $CL\|\eta_h^{(L)}\|_{L^1(Q_L)} + ML^2$, which is integrable. Thus the maxima in (15)–(16) are attained. Convex combinations of traces and their admissible interiors show that $\mathcal{R}$ is concave. Compactness of the admissible pairs and upper semicontinuity of the payoff give upper semicontinuity of $\mathcal{R}$. Thus (16) is convex and 1-Lipschitz and $(\mathcal{H}_{h,L}\Psi)^* = -\mathcal{R}$ on the

closed unit ball. If $\Psi \leq \Phi$, then $\Psi^* \geq \Phi^*$, which proves order preservation; the constant rule and contraction follow in the same way.

Step 2. Bound the update and preserve the symmetries.

The affine trace $F(x) = a \cdot x$ gives $\mathcal{R}(a) \geq -\Psi^*(a)$, so biconjugacy gives the lower bound in (16a). For the upper bound, we convert the source pairing into a perturbation of the base flux. Paste on the unit outer blocks the linear zero-normal flux Q from (12a), so that $-\operatorname{div} Q = \eta_h^{(L)}$, $\mathbb{E}Q(x) = 0$, and $\sup_x \mathbb{E}|Q(x)|^2 \leq Ch$. For every $F \in \mathcal{A}_L(a)$, integration by parts and $\oint \nabla F = a$ give

$$p \cdot a + L^{-2} \left(\int F \eta_h^{(L)} - \int \Psi^*(\nabla F) \right) = \oint [(p + Q) \cdot \nabla F - \Psi^*(\nabla F)] \leq \oint \Psi(p + Q). \quad (16b)$$

Taking the fixed-trace maximum and expectation in (16b), then using that Ψ is 1-Lipschitz, yields

$$(\mathcal{H}_{h,L}\Psi)(p) \leq \mathbb{E} \oint \Psi(p + Q) \leq \Psi(p) + C\sqrt{h}. \quad (16c)$$

This proves (16a) directly from the dual cell problem. The same argument, without the p -term, also gives

$$0 \leq L^{-2}K_{h,L,\Psi}(c) \leq C(M + \sqrt{h}), \quad 0 \leq \mathcal{R}_{h,L,\Psi}(a) \leq C(M + \sqrt{h}). \quad (16d)$$

For the lower bound, use any deterministic extension of c : its expected source pairing is zero and $-\Psi^* \geq 0$. The upper bound follows from integration by parts against the pasted local flux Q and Fenchel's inequality. These estimates also imply the integrable domination needed when traces converge.

Applying a coordinate reflection or permutation simultaneously to the source, flux, and trace preserves their laws and the variational value; this proves the D_4 -invariance assertion.

Step 3. Treat a compact slope domain.

The additional issue is to retain the constraint $\nabla F \in K_\Psi$ when traces and interior potentials converge. Write $g = \Psi^*$, extended by $+\infty$ off K_Ψ . The compatible traces, and the set of compatible affine-cocycle pairs (a, c) , are closed. Indeed, if F_n are normalized extensions with $\nabla F_n \in K_\Psi$, uniform convergence and weak- * compactness in $W^{1,\infty}$ give an extension F . For every affine functional ℓ nonpositive on K_Ψ , weak- * passage in $\int_E \ell(\nabla F_n)$ gives $\int_E \ell(\nabla F) \leq 0$ for every measurable E ; finite-dimensional separation therefore gives $\nabla F \in K_\Psi$ almost everywhere. The trace and cocycle identities pass to the limit. Thus both compatible sets are compact.

The functional $\int g(\nabla F)$ is weak- * lower semicontinuous, so the payoff in (15) is upper semicontinuous and the fixed-trace maximum is attained. For a fixed source its value is upper semicontinuous in the compatible trace: take maximizing extensions along a convergent sequence of traces and use the preceding compactness. The maximum is measurable in the finite Gaussian coefficient vector by the measurable maximum theorem. Moreover, after anchoring one value,

$$\left| \max_F \left\{ \int F \eta_h^{(L)} - \int g(\nabla F) \right\} \right| \leq C_{K_\Psi, L, g} (1 + \|\eta_h^{(L)}\|_{L^1(Q_L)}), \quad (16s)$$

uniformly over compatible traces. The right side is integrable. Reverse Fatou therefore shows that the expected fixed-trace value is upper semicontinuous. Compactness of

the affine-cocycle pairs now proves attainment and upper semicontinuity of the outer maximum defining $\mathcal{R}$.

Convex combinations of compatible potentials remain compatible, and the payoff is concave in the potential. Mixing affine-cocycle potentials proves that $\mathcal{R}$ is concave. Every compatible potential of slope a satisfies

$$a = \oint_{Q_L} \nabla F \in K_\Psi,$$

whereas $F(x) = a \cdot x$ is compatible for every $a \in K_\Psi$. Consequently $\mathcal{R}$ is finite precisely on K_Ψ , is upper semicontinuous there, and equals $-\infty$ off it. Hence $-\mathcal{R}$ is a proper closed convex extended-real function and $\mathcal{H}_{h,L}\Psi = (-\mathcal{R})^*$, as asserted.

Incompatible traces have value $-\infty$. Thus, if $\Psi \leq \Phi$, extended-real conjugacy gives $\Psi^* \geq \Phi^*$ pointwise even when their compact domains differ; the payoff inequality may be maximized over the common ambient trace space and proves successively the order of the fixed-trace values, $\mathcal{R}$, and $\mathcal{H}$. Finally $(\Psi + C)^* = \Psi^* - C$, which proves constant equivariance.

■

4.2. A FINITE APPROXIMATION OF COMMON BOUNDARY TRACES. A global optimizing potential can use the fine source to choose its trace on every cell boundary. To compare this value with the cell operator, we approximate all of those boundary values simultaneously by a finite family of compatible traces. For each fixed trace the cell interiors optimize separately, and Gaussian concentration bounds the gain from choosing the trace after seeing the source. The lemma is stated for a constant base flux; the proof also records the conditional estimates needed when that constant is chosen from an independent coarse field.

LEMMA 4.3 (Common-trace comparison). *Fix an h -level source and a convex profile $\Psi \in \mathcal{P}_M$. Let $\Omega = Q_R$, with R/L an integer, and let Σ_L be the union of all boundaries of its L -squares. Let $\mathfrak{T}_{\Omega,L}$ be the compact space of restrictions of global 1-Lipschitz potentials to all of Σ_L , modulo one additive constant. Let η consist of independent copies of $\eta_h^{(L)}$ in these cells. Fix a deterministic constant base vector $p \in \mathbb{R}^2$ and define*

$$\mathcal{V}_\Omega(p, \eta) = \sup_{\substack{F \in W^{1,\infty}(\Omega), \\ |\nabla F| \leq 1 \\ F(0)=0}} \left\{ \int_\Omega F \eta + \int_\Omega p \cdot \nabla F - \int_\Omega \Psi^*(\nabla F) \right\}. \quad (17a)$$

For $c \in \mathfrak{T}_{\Omega,L}$, let $\mathcal{V}_c(p, \eta)$ be (17a) with $F|_{\Sigma_L} = c$. If D is an L -cell, write c_D for the restricted trace and

$$m_D(c) = \int_D \nabla F = \int_{\partial D} c_D n d\mathcal{H}^1, \quad (17b)$$

which depends only on the trace. After translation to the reference cell we denote the same quantity by $m_L(c_D)$. For $0 < \varepsilon < 1/10$ there is a finite net $\mathcal{N} \subset \mathfrak{T}_{\Omega,L}$ such that

$$\begin{aligned} \mathbb{E} \mathcal{V}_\Omega(p, \eta) &\leq \max_{c \in \mathcal{N}} \sum_D \{p \cdot m_D(c) + K_{h,L,\Psi}(c_D)\} \\ &\quad + |\Omega| \left[C(M + |p| + \sqrt{h}) \left(\varepsilon + \frac{1}{\varepsilon L} \right) + C \sqrt{\frac{\log(C/\varepsilon)}{L}} \right]. \end{aligned} \quad (17)$$

Conversely, for every deterministic $c \in \mathfrak{T}_{\Omega,L}$,

$$\mathbb{E}\mathcal{V}_{\Omega}(p, \eta) \geq \sum_D \{p \cdot m_D(c) + K_{h,L,\Psi}(c_D)\}. \quad (17c)$$

In Lemma 5.1, the coarse field is first frozen to a constant p_D on each block before this estimate is applied.

For the reference L -cell, write $K(c) = K_{h,L,\Psi}(c)$ for the expected optimized fine dual value at boundary trace c . This function is concave in c . If

$$\mathcal{R}(a) = L^{-2} \max\{K(c) : c \text{ has affine-periodic slope } a\}, \quad |a| \leq 1, \quad (18)$$

then $\mathcal{R}$ is concave and upper semicontinuous and

$$H(p) = \sup_{|a| \leq 1} \{p \cdot a + \mathcal{R}(a)\}, \quad H^* = -\mathcal{R} \text{ on } \overline{B}_1. \quad (19)$$

With the notation of Definition 4.1, $H = \mathcal{H}_{h,L}\Psi$ exactly. The cell size remains finite throughout this comparison.

Proof. We prove (17) by approximating the random trace and then comparing the maximum over that finite family with its expected fixed-trace values. To make its later conditional use explicit, we carry out the construction with an independent coarse field.

On a coarse probability space let P be a strongly $L^1(\Omega; \mathbb{R}^2)$ -measurable field with $\mathbb{E} \int_{\Omega} |P| < \infty$. Independently, on a product fine probability space, put independent copies η_D of the active fine source in the L -cells and set $\eta = \sum_D \eta_D$. With

$$\mathcal{A}_{\Omega} = \{F \in C(\overline{\Omega}) : F(0) = 0, \text{Lip}(F) \leq 1\},$$

set

$$\mathcal{V}_{\Omega}(P, \eta) = \max_{F \in \mathcal{A}_{\Omega}} \left\{ \langle \eta, F \rangle + \int_{\Omega} P \cdot \nabla F - \int_{\Omega} \Psi^*(\nabla F) \right\}, \quad \mathbf{D}_{R,L} = \mathbb{E}_P \mathbb{E}_{\eta} \mathcal{V}_{\Omega}(P, \eta). \quad (17a')$$

The maximum is the free-outer-boundary dual problem; no periodic outer trace is imposed. Let Q be the zero-normal flux obtained by pasting (12a) over the unit outer blocks, so $-\text{div } Q = \eta$. Integration by parts and (14) bound the objective in (17a') by $\int_{\Omega} (|P| + |Q| + M)$. By (12c), $\mathbb{E}_{\eta} \int_{\Omega} |Q| \leq C|\Omega|\sqrt{h}$. The assumed integrability of P therefore makes $\mathbf{D}_{R,L}$ finite.

Let $\mathcal{V}_c(P, \eta)$ denote (17a') with $F|_{\Sigma_L} = c$. For $P \equiv p$, these definitions agree with those in the statement.

Step 1. Construct a finite family of compatible traces.

The zero normal trace of Q allows integration by parts in (17a'). Its objective becomes

$$\sup_F \int [(P + Q) \cdot \nabla F - \Psi^*(\nabla F)], \quad |\nabla F| \leq 1. \quad (20)$$

Rounding and interpolating each face separately would not ensure agreement at junctions or compatibility around cycles. We therefore round values on the whole skeleton and use one McShane extension. The preliminary contraction by $\theta = 1 - 4\varepsilon$ supplies the slack for rounding.

Normalize F at one skeleton vertex. Scale it by $\theta = 1 - 4\varepsilon$, sample Σ_L at unit spacing (including every junction), and round the sampled values to a lattice of mesh $\varepsilon/8$. Distinct

sample vertices have Euclidean distance at least one, and therefore their rounded values $q(v)$ satisfy

$$|q(v) - q(w)| \leq \theta|v - w| + \varepsilon/4 \leq (1 - 2\varepsilon)|v - w|.$$

Define the whole skeleton trace, rather than interpolating face by face, by

$$c_q(x) = \inf_v \{q(v) + (1 - 2\varepsilon)|x - v|\}, \quad x \in \Sigma_L. \quad (20a)$$

This McShane formula agrees with the node data and is globally $(1 - 2\varepsilon)$ -Lipschitz in the Euclidean metric. Since c_q is one function on Σ_L , its restrictions agree on every shared edge and vertex. Comparison with the nearest node on the same unit skeleton edge gives $\|c_q - \theta F\|_{L^\infty(\Sigma_L)} \leq 1$. These traces form the net $\mathcal{N}$. The skeleton graph has at most $C|\Omega|/L$ unit edges. Once one anchor is fixed, each new vertex has at most C/ε possible increments (cycle constraints only reduce this number). Hence

$$\log |\mathcal{N}| \leq C \frac{|\Omega|}{L} \log(C/\varepsilon). \quad (21)$$

Step 2. Match the optimizer to the common trace.

Preserve the scaled optimizer $F_{\text{core}} = \theta F$ on the core at distance $r = C/\varepsilon$ from the skeleton, and use $F_{\text{skel}} = c_q$ on the skeleton. The bound $\|c_q - \theta F\|_{L^\infty(\Sigma_L)} \leq 1$ gives, for a core point x and a skeleton point y ,

$$|F_{\text{core}}(x) - F_{\text{skel}}(y)| \leq C + (1 - 4\varepsilon)|x - y| \leq (1 - \varepsilon)|x - y|.$$

Thus the union of all core data and the common skeleton trace has a McShane 1-Lipschitz extension. In the intervening corridors, we use the zero-normal flux Q on each unit outer block. Let $\mathcal{C}_{\Omega, L, \varepsilon}$ be the resulting union of whole unit outer blocks. It has deterministic relative area at most $C/(\varepsilon L)$, and, because every unit-outer-block source is balanced,

$$\mathbb{E}_\eta \int_{\mathcal{C}_{\Omega, L, \varepsilon}} |Q| \leq C|\Omega| \frac{\sqrt{h}}{\varepsilon L}. \quad (21a)$$

On the core, put $v = \nabla F$. Concavity gives $-\Psi^*(\theta v) \geq \theta[-\Psi^*(v)] + (1 - \theta)[- \Psi^*(0)]$. By (14) and $1 - \theta = 4\varepsilon$, replacing v by θv loses at most $C\varepsilon(M + |P| + |Q|)$ per point. On the corridor, both old and new integrands have magnitude at most $|P| + |Q| + M$. Consequently this modification gives, for every realized coarse field P ,

$$\begin{aligned} 0 &\leq \mathbb{E}_\eta [\mathcal{V}_\Omega(P, \eta) - \max_{c \in \mathcal{N}} \mathcal{V}_c(P, \eta) \mid P] \\ &\leq C|\Omega|(M + \sqrt{h}) \left(\varepsilon + \frac{1}{\varepsilon L} \right) + C\varepsilon \int_\Omega |P| + C \int_{\mathcal{C}_{\Omega, L, \varepsilon}} |P|. \end{aligned} \quad (21a')$$

The last two terms retain the spatial dependence of the realized field P . If $\sup_x \mathbb{E}|P(x)| \leq C\sqrt{T}$, only now averaging in P and using the deterministic corridor fraction turns the last two terms into $C|\Omega|\sqrt{T}(\varepsilon + 1/(\varepsilon L))$. For $P \equiv p$, the same calculation gives the geometric error in (17).

Step 3. Decompose the fixed-trace problem cell by cell.

For a realized coarse field P , put

$$\begin{aligned} Y_D(P, c_D; \eta_D) &= \max_{\substack{F_D | \partial D = c_D \\ |\nabla F_D| \leq 1}} \left\{ \langle \eta_D, F_D \rangle + \int_D P \cdot \nabla F_D - \int_D \Psi^*(\nabla F_D) \right\}, \\ K_{D, P}(c_D) &= \mathbb{E}_{\eta_D} Y_D(P, c_D; \eta_D). \end{aligned}$$

Every family of cell interiors whose boundary values agree with one $c \in \mathfrak{T}_{\Omega,L}$ pastes to a global 1-Lipschitz function: restrict the straight segment between two points to its successive cell pieces and sum the cellwise Lipschitz bounds. Conversely every global potential restricts to such a family. Thus, realization by realization,

$$\mathcal{V}_\Omega(P, \eta) = \max_{c \in \mathfrak{T}_{\Omega,L}} \sum_D Y_D(P, c_D; \eta_D), \quad \mathbb{E}_\eta[\mathcal{V}_c(P, \eta) \mid P] = \sum_D K_{D,P}(c_D). \quad (21c)$$

The second equality follows by linearity of conditional expectation over the finite sum. In general $K_{D,P}$ depends on the whole restriction $P|_D$, so it cannot be replaced by a function of $\int_D P$. If $P = p_D$ is constant on D , however, integration by parts gives the exact specialization

$$K_{D,P}(c_D) = K_{h,L,\Psi}(c_D) + p_D \cdot m_D(c). \quad (21d)$$

In particular (21d) proves the fixed-trace identity used in (17), and it also applies conditionally when the constants p_D are random but measurable with respect to the independent coarse sigma-field.

Step 4. Control optimization over the finite trace family.

For a fixed skeleton trace, the optimized fine value in one L -cell is CL -Lipschitz in its standard Gaussian coefficients. Indeed, in each unit outer block subtract the average of F ; balance makes this subtraction invisible, and the unit-block Poincare inequality gives

$$\sum_{C \text{ fine}} \left(n \int_C (F - \int_{Q_1} F) \right)^2 \leq C \int_{Q_1} |\nabla F|^2 \leq C. \quad (21b)$$

Summing (21b) over the L^2 outer blocks gives Gaussian Lipschitz radius CL . Values in different L -cells depend on disjoint Gaussian blocks. For fixed c , $\mathcal{V}_c(P, \eta)$ therefore has Gaussian Lipschitz radius at most $C\sqrt{|\Omega|}$, uniformly in the realized coarse field P . The elementary Gaussian estimate

$$\mathbb{E} \max_{\alpha \leq m} (Y_\alpha - \mathbb{E} Y_\alpha) \leq \sigma \sqrt{2 \log m}$$

for σ -subgaussian variables, combined with (21), gives, pointwise in P ,

$$\mathbb{E}_\eta \max_{c \in \mathcal{N}} \mathcal{V}_c(P, \eta) \leq \max_{c \in \mathcal{N}} \sum_D K_{D,P}(c_D) + C|\Omega| \sqrt{\frac{\log(C/\varepsilon)}{L}}. \quad (21e)$$

Specialize to $P \equiv p$, use (21d), and combine (21a') with (21e) to prove (17). Formula (17c) follows by restricting the original supremum to any fixed common trace. For random cellwise constants, (21d)–(21e) remain valid conditionally on the coarse sigma-field. We will take expectation in that field after the deterministic comparison in the next section.

Step 5. Identify the deterministic one-cell operator.

If c_1, c_2 are two traces, convex combinations of their admissible interiors are admissible, and the integrand in (20) is concave in ∇F . Hence K is concave. On an $m \times m$ torus of L -cells, write a deterministic trace pattern on the universal cover as c , with $c(x + mL e_i) - c(x) = mL a_i$. For the cell indexed by z , put $c_z(x) = c(x + Lz) - c(Lz)$ on the reference cell boundary. Shifting z by e_i in the average $m^{-2} \sum_z c_z$ cancels the interior terms; the wraparound terms give the increment La_i between its opposite faces. Thus the averaged trace is one-cell affine-periodic, and cellwise concavity gives

$$m^{-2} \sum_z K(c_z) \leq K \left(m^{-2} \sum_z c_z \right).$$

Conversely such a trace repeats globally. This proves (18)–(19) on a torus; the explicit finite-square affine-periodicity correction used later is proved in Lemma 5.1. Equations (15)–(16) say exactly that $H = \mathcal{H}_{h,L}\Psi$. This identifies the deterministic cell functional in (18)–(19).

■

5. A quantitative two-scale comparison

We now remove one fine band η from the transport problem by replacing Ψ with $\mathcal{H}_{h,L}\Psi$. The remaining source is the independent coarse field ζ . The previous section treats a constant base flux, so we first need room to smooth the coarse problem, freeze it on blocks, and join the resulting cell competitors. With the outer blocks of the fine band normalized to side one, the scales satisfy

$$1 \ll L \ll B \ll A \ll S \leq R :$$

Here L is the cell side used by $\mathcal{H}_{h,L}$, and B is the side of a block on which the smoothed flux or slope is approximated by a constant. Smoothing on the larger scale A makes this approximation accurate, while $A \ll S$ keeps its effect on the coarse source small. The ambient square has side R .

LEMMA 5.1 (Two-scale comparison). *Let η be one retained h -level dyadic band, normalized so that its outer blocks have side one. Let ζ be an independent sum of at most T coarser Haar levels whose smallest parent side, in this normalization, is the dyadic integer S . Work on Q_R , where R is a common multiple of L, B, A, S , and all dyadic parents are aligned with its boundary. For brevity write $Q := Q_R$ throughout this lemma. Let P be the sum of the scaled linear right inverses (12a), so that $-\operatorname{div} P = \zeta$ and $P \cdot n = 0$. It satisfies*

$$\sup_{x \in Q_R} \mathbb{E}|P(x)| \leq C\sqrt{T}, \quad \mathbb{E} \int_{\mathbb{T}_{2R}^2} |P^{\operatorname{ref}}(x+z) - P^{\operatorname{ref}}(x)| dx \leq C|z|/S, \quad (22)$$

for $|z| \leq A \ll S$, where P^{ref} is the covariant reflection defined after (22b1) below. The same estimate, without reflection, holds on every aligned interior translate. For $1 \ll L \ll B \ll A \ll S$, $0 < \varepsilon < 1/10$, and a D_4 -invariant $\Psi \in \mathcal{P}_M$, put

$$\delta = \varepsilon + \frac{1}{\varepsilon L} + \frac{L}{\varepsilon B} + \frac{B}{\varepsilon A} + \frac{A}{S}. \quad (23)$$

Assume also

$$\varepsilon L \geq C, \quad L + B^2/A \leq c\varepsilon B. \quad (23a)$$

Both inequalities hold for the schedule used below. On this aligned square,

$$\begin{aligned} \frac{1}{|Q|} |\mathbb{E} J_{\Psi}^Q(\eta + \zeta) - \mathbb{E} J_{\mathcal{H}_{h,L}\Psi}^Q(\zeta)| &\leq C\delta \left(1 + \frac{\mathbb{E} J_{\mathcal{H}_{h,L}\Psi}^Q(\zeta)}{|Q|} + M + \sqrt{T} + \sqrt{h} \right) \\ &\quad + C\sqrt{\frac{\log(C/\varepsilon)}{L}}. \end{aligned} \quad (24)$$

The estimate is uniform for every iterated profile occurring below.

Proof. Write $H = \mathcal{H}_{h,L}\Psi$. For the upper bound, we start with a flux for $J_H^Q(\zeta)$ and add fine-source corrections with zero normal trace on B -blocks. The local estimate (26) will compare their cost with H at the frozen flux. For the lower bound, we start with a dual

potential for $J_H^Q(\zeta)$ and insert fine cell optimizers while preserving common values on the B -block boundaries.

Both constructions smooth on Q while preserving the boundary constraints. We first estimate translations of the coarse flux and then use reflection to smooth fluxes and potentials, including measure-valued fluxes. The separate square-root term in (24) comes from Gaussian concentration in Lemma 4.3.

Step 1. Control translations of the coarse flux.

We use the following explicit right inverse in (22). On a Haar parent $[0, r]^2$, for a mean-zero density f , set

$$\begin{aligned} m(y) &= r^{-1} \int_0^r f(s, y) ds, \\ P_1(x, y) &= - \int_0^x (f(s, y) - m(y)) ds, \\ P_2(x, y) &= - \int_0^y m(t) dt. \end{aligned} \tag{22a0}$$

Then $-\operatorname{div} P = f$ and $P \cdot n = 0$ on all four faces. For a normalized Haar detail on a parent of side r , every pointwise Gaussian coefficient in (22a0) is $O(1)$, while the expected total variation of DP per unit area is $O(r^{-1})$. The BV translation inequality bounds the mean difference under a shift by $|z|$ times the mean total variation; the pointwise bound controls larger shifts. Consequently

$$\mathbb{E} \int |P_j(\cdot + z) - P_j| \leq C \min\{1, |z|/r_j\}. \tag{22a}$$

The fields P_j are independent and centred, so their pointwise second moments add, giving the first estimate in (22). Since the scales $r_j \geq S$ grow geometrically, summing (22a) gives the second. The explicit normal component vanishes on every aligned parent face; covariant reflection therefore creates no jump there, and the same BV translation estimate holds on the reflected torus.

Step 2. Smooth without changing the boundary constraint.

An effective minimizing flux may have a singular part. We therefore smooth in the space of vector measures and preserve the divergence and normal-trace constraints before estimating the energy. A vector Radon measure $\mu \in \mathcal{M}(\overline{Q}_R; \mathbb{R}^2)$ has source ρ and zero outer normal trace precisely when

$$\int_{\overline{Q}_R} \nabla \phi \cdot d\mu = \langle \rho, \phi \rangle \quad \text{for every } \phi \in C^1(\overline{Q}_R). \tag{22b0}$$

In particular, if $w = P\mathcal{L}^2 + v$ and $P\mathcal{L}^2$ have the same source and zero normal trace, then v is divergence-free and the left side of (22b0) is zero.

Identify $\mathbb{T}_{2R}^2$ with $[-R, R]^2$. For $\epsilon = (\epsilon_1, \epsilon_2) \in \{\pm 1\}^2$, set $S_\epsilon x = (\epsilon_1 x_1, \epsilon_2 x_2)$ and $O_\epsilon = \operatorname{diag}(\epsilon_1, \epsilon_2)$. Define the covariant reflection, including any boundary-supported tangential part, by

$$\int_{\mathbb{T}_{2R}^2} \Phi(y) \cdot dv^{\operatorname{ref}}(y) = \sum_{\epsilon \in \{\pm 1\}^2} \int_{\overline{Q}_R} \Phi(S_\epsilon x) \cdot O_\epsilon dv(x). \tag{22b1}$$

For every smooth periodic scalar χ , the chain rule and (22b0) give

$$\int_{\mathbb{T}_{2R}^2} \nabla \chi \cdot dv^{\text{ref}} = \sum_{\epsilon} \int_{\overline{Q}_R} \nabla(\chi \circ S_{\epsilon}) \cdot dv = 0. \quad (22b2)$$

Thus the reflected measure remains divergence-free across the faces, periodic seams, and corners. For a density P , define $P^{\text{ref}}(S_{\epsilon}x) = O_{\epsilon}P(x)$. The same calculation reflects its scalar source evenly. Thus $(P\mathcal{L}^2 + v)^{\text{ref}} = P^{\text{ref}}\mathcal{L}^2 + v^{\text{ref}}$.

Let φ_A be a nonnegative smooth radial probability kernel on the torus, supported at scale $A < R/4$, and put

$$v_A = (\varphi_A * v^{\text{ref}})|_{Q_R}, \quad P_A = (\varphi_A * P^{\text{ref}})|_{Q_R}. \quad (22b3)$$

Convolution preserves (22b2), and covariance under the reflections forces the classical normal component of v_A to vanish on every side. Hence v_A is an exact smooth divergence-free zero-normal field on Q_R , and

$$\int_{Q_R} |Dv_A| \leq CA^{-1}|v|(\overline{Q}_R). \quad (22b4)$$

To control the energy of this smoothing, let H be any convex profile with $|z| \leq H(z) \leq |z| + M_H$. Its recession function is $|z|$, and for $\mu = \mu^a\mathcal{L}^2 + \mu^s$ define

$$\mathcal{F}_H(\mu; U) = \int_U H(\mu^a) dx + |\mu^s|(\overline{U}).$$

The perspective

$$\widehat{H}(t, z) = \begin{cases} tH(z/t), & t > 0, \\ |z|, & t = 0, \end{cases}$$

is convex, lower semicontinuous, and positively one-homogeneous. If λ dominates $\mathcal{L}^2$ and μ , and $(a, b) = d(\mathcal{L}^2, \mu)/d\lambda$, then

$$\mathcal{F}_H(\mu) = \int \widehat{H}(a, b) d\lambda. \quad (22b5)$$

Applying convexity pointwise to the convolution of the pair $(\mathcal{L}^2, \mu)$, followed by Fubini, proves the measure-level Jensen inequality

$$\mathcal{F}_H(\varphi_A * \mu; \mathbb{T}_{2R}^2) \leq \mathcal{F}_H(\mu; \mathbb{T}_{2R}^2). \quad (22b6)$$

This argument includes singular measures and does not presuppose an L^1 minimizer.

Assume now that H is invariant under the signed coordinate reflections, as is every profile in the iteration. The reflected measure has energy at most four times the original energy (overlapping face masses use the triangle inequality for the recession norm), while the convolved density has four equal quadrant energies. From (22b6),

$$\int_{Q_R} H(P_A + v_A) dx \leq \mathcal{F}_H(P\mathcal{L}^2 + v; Q_R). \quad (22b7)$$

Moreover (22) and Fubini give

$$\mathbb{E} \int_{Q_R} |P_A - P| \leq C(A/S)|Q_R|. \quad (22b8)$$

Since H is 1-Lipschitz, (22b7)–(22b8) yield

$$\mathbb{E} \int_{Q_R} H(P + v_A) \leq \mathbb{E} \mathcal{F}_H(P\mathcal{L}^2 + v; Q_R) + C(A/S)|Q_R|. \quad (22b9)$$

Consequently, we may smooth the divergence-free correction on the whole square at expected cost CA/S per unit area.

Step 3. Upper bound: smooth and freeze the flux.

For the upper bound in (24), use the relaxed linear-growth functional on vector Radon measures and take a near-minimizer $w = P\mathcal{L}^2 + v$ for $J_{\mathcal{H}_{h,L}\Psi}^Q(\zeta)$, where $\operatorname{div} v = 0$ and v has zero outer normal trace. (The singular part is charged by the recession function $|\cdot|$.) Convolve only v on scale A . Equations (22b7)–(22b9) change the effective energy by at most CA/S per unit area. Since $H \geq |\cdot|$, $|v|(Q) \leq \mathcal{F}_H(P\mathcal{L}^2 + v; Q) + \int_Q |P|$. Let $p_D = f_D(P + v_A)$ on each B -block. By Poincaré's inequality and (22b4),

$$\sum_D \int_D \left| v_A - \oint_D v_A \right| \leq CB \int_Q |Dv_A| \leq C(B/A)|v|(Q).$$

Averaging the translation bound (22) over pairs in each B -block gives the corresponding $C(B/S)|Q|$ bound for P . Thus

$$R^{-2}\mathbb{E} \sum_D \int_D |P + v_A - p_D| \leq C \left(\frac{B}{A} + \frac{B}{S} \right) \left(1 + R^{-2}\mathbb{E}|v|(Q_R) + \sqrt{T} \right). \quad (25)$$

The remaining local task is to insert the fine source at constant base p_D . We next prove a deterministic comparison uniform in this base, so that it can subsequently be applied conditionally to the coarse field.

Step 4. Compare the B -block cost with H .

For a constant base p , we compare the expected B -block cost with $H(p)$. Repeating an optimal affine-periodic L -cell trace shows that the block cost is at least $H(p)$. For the reverse bound, Lemma 4.3 bounds the gain from choosing a trace after observing the fine source by optimization over deterministic common traces and a concentration error. We then average a maximizing trace over L -cell translations and correct its opposite-face increments to obtain a trace admissible for H .

The fine source is exactly balanced on every B -block. To formulate the local problem without suppressing singular parts, for $p \in \mathbb{R}^2$ set

$$\begin{aligned} I_B(p, \eta_D) &= \min \{ \mathcal{F}_\Psi(p\mathcal{L}^2 + u; D) : -\operatorname{div} u = \eta_D, [u \cdot n]_{\partial D} = 0 \}, \\ N_B(p) &= B^{-2}\mathbb{E}_\eta I_B(p, \eta_D). \end{aligned} \quad (25a0)$$

Here the constraint is on the correction u , not on the constant base $p\mathcal{L}^2$. Apply the same separation argument as in (2) to this affine constraint. Fenchel's inequality reads $\Psi(p+u) \geq \nabla F \cdot (p+u) - \Psi^*(\nabla F)$, while the zero-normal constraint gives $\int_D \nabla F \cdot du = \langle \eta_D, F \rangle$. Separation gives the reverse inequality and hence, realization by realization,

$$I_B(p, \eta_D) = \max_{F \in \mathcal{A}_B} \left\{ \langle \eta_D, F \rangle + \int_D p \cdot \nabla F - \int_D \Psi^*(\nabla F) \right\}. \quad (25a1)$$

We first prove the lower side of the finite comparison. Fix $|a| \leq 1$ and a trace $c \in \mathcal{T}_L^{\text{per}}(a)$ attaining $\mathcal{R}(a)$. Repeat its universal-cover affine-periodic lift through the B -box and, for each fine realization, fill every L -cell by an optimizer with that complete trace. The interiors paste, and every cell has mean gradient a . Substitution in (25a1), expectation, and maximization over a give

$$(\mathcal{H}_{h,L}\Psi)(p) \leq N_B(p). \quad (25a2)$$

For the reverse side apply (21a')–(21e) to (25a1) with the constant realized field $P \equiv p$. Since (21d) applies exactly,

$$N_B(p) \leq \mathfrak{A}_{B,L}(p) + C(M + \sqrt{h} + |p|) \left(\varepsilon + \frac{1}{\varepsilon L} \right) + C \sqrt{\frac{\log(C/\varepsilon)}{L}}, \quad (25a3)$$

$$\mathfrak{A}_{B,L}(p) = B^{-2} \max_{c \in \mathfrak{I}_{B,L}} \sum_z \{K(c_z) + p \cdot m_L(c_z)\}. \quad (25a4)$$

Every constant is uniform in the possibly unbounded vector p .

Choose a maximizing trace in (25a4), one global extension F , and write $n = B/L$,

$$g(s) = n^{-2} \sum_{z \in \{0, \dots, n-1\}^2} F(Lz + s), \quad s \in Q_L, \quad q = B^{-2} \int_{Q_B} \nabla F. \quad (25a5)$$

Additive constants do not affect K , so concavity and the divergence theorem give

$$n^{-2} \sum_z K(c_z) \leq K(g|_{\partial Q_L}), \quad B^{-2} \sum_z p \cdot m_L(c_z) = p \cdot q. \quad (25a6)$$

For s on the face $s_j = 0$, let $D_j(s) = g(s + Le_j) - g(s)$. Telescoping in direction e_j leaves only the two outer rows of the B -box. Each row has n terms, weighted by n^{-2} , so D_j is $2/n$ -Lipschitz. Its face average is Lq_j by the divergence theorem; hence

$$\|D_j - Lq_j\|_\infty \leq 2L/n. \quad (25a7)$$

The averaged function g is already Lipschitz, but (25a7) gives only approximate affine-periodicity. We repeat g with the desired affine increments and then remove the resulting small seam jumps by a Lipschitz envelope. Put $\theta = 1 - 4\varepsilon$, $d = 2L/n$, and, on half-open cells of the universal cover, define

$$G_0(s + Lk) = \theta g(s) + \theta Lq \cdot k.$$

Its seam jumps are at most d . A segment of Euclidean length r crosses at most $2 + \sqrt{2}r/L$ grid faces: the sum of its two coordinate displacements is at most $\sqrt{2}r$. Since $\sqrt{2}d/L = 2\sqrt{2}/n \leq \varepsilon$ when $n\varepsilon \geq 8$, we obtain

$$|G_0(x) - G_0(y)| \leq (1 - 3\varepsilon)|x - y| + 4d. \quad (25a8)$$

Since $|\theta q| \leq \theta < 1 - \varepsilon/2$, the following infimum is finite. Define

$$\widehat{G}(x) = \inf_y \{G_0(y) + (1 - \varepsilon/2)|x - y|\}.$$

The change of variable $y \mapsto y + Lk$ proves the exact cocycle $\widehat{G}(x + Lk) = \widehat{G}(x) + \theta Lq \cdot k$, while (25a8) gives $G_0 - 4d \leq \widehat{G} \leq G_0$. Thus $\widehat{c} = \widehat{G}|_{\partial Q_L}$ is an admissible affine-periodic trace of slope θq and differs from θg by at most $4d$.

For a fine realization take an optimizer with trace g , keep its θ -multiple on every unit outer block farther than $C(1 + d/\varepsilon)$ from ∂Q_L , prescribe $\widehat{c}$ on the boundary, and use McShane extension in the collar. Since $d = 2L^2/B$, the collar occupies at most a fraction $C(1/(\varepsilon L) + L/(\varepsilon B))$ of the cell. Its balanced unit-block sources are controlled by the flux estimate used in (21a'). On the retained blocks, contraction and $-\Psi^*(\theta a) \geq -\theta\Psi^*(a) - (1 - \theta)\Psi^*(0)$ give

$$L^{-2}K(g) \leq \mathcal{R}(\theta q) + C(M + \sqrt{h}) \left(\varepsilon + \frac{1}{\varepsilon L} + \frac{L}{\varepsilon B} \right). \quad (25b)$$

Since $p \cdot q \leq p \cdot (\theta q) + 4\varepsilon|p|$, equations (25a3)–(25b) and the definition of $\mathcal{H}$ give, for every p ,

$$\boxed{(\mathcal{H}_{h,L}\Psi)(p) \leq N_B(p) \leq (\mathcal{H}_{h,L}\Psi)(p) + C(M + \sqrt{h} + |p|) \left(\varepsilon + \frac{1}{\varepsilon L} + \frac{L}{\varepsilon B} \right) + C\sqrt{\frac{\log(C/\varepsilon)}{L}}.} \quad (26)$$

Estimate (26) is uniform in the possibly unbounded base vector p , so it applies conditionally to the random block averages p_D .

Step 5. Upper bound: paste the local flux corrections.

Now condition on the coarse sigma-field. Each $p_D = f_D(P + v_A)$ is then deterministic and independent of the fine field, and the local sources are independent. Using the same correction u at two bases changes its normalized energy by at most $|p - p'|$, so N_B is 1-Lipschitz and hence Borel. Equations (25a0) and (26) give the conditional estimate

$$\begin{aligned} \mathbb{E}_\eta \left[\sum_D I_B(p_D, \eta_D) \mid \mathcal{G} \right] &= \sum_D |D| N_B(p_D) \\ &\leq \sum_D |D| H(p_D) + C\alpha \left\{ (M + \sqrt{h})|Q| + \sum_D |D| |p_D| \right\} \\ &\quad + C|Q| \sqrt{\frac{\log(C/\varepsilon)}{L}}, \quad \alpha = \varepsilon + \frac{1}{\varepsilon L} + \frac{L}{\varepsilon B}, \end{aligned} \quad (26i)$$

Here $H = \mathcal{H}_{h,L}\Psi$.

Fix $\tau_{\text{sel}} > 0$ and choose measurable $\tau_{\text{sel}}|D|$ -minimizers $u_D^{\tau_{\text{sel}}}$ in (25a0) as justified at the end of the proof. The zero extensions have no interface divergence. Add them to the unfrozen smoothed flux to obtain the global competitor

$$\mu^{\tau_{\text{sel}}} = (P + v_A)\mathcal{L}^2 + \sum_D u_D^{\tau_{\text{sel}}},$$

which has source $\eta + \zeta$ and zero outer normal trace. The constants p_D enter only the energy comparison below. The 1-Lipschitz property of Ψ and subadditivity of its recession norm give

$$\mathcal{F}_\Psi(\mu^{\tau_{\text{sel}}}; Q) \leq \sum_D \mathcal{F}_\Psi(p_D \mathcal{L}^2 + u_D^{\tau_{\text{sel}}}; D) + \sum_D \int_D |P + v_A - p_D|. \quad (26j)$$

This remains true if local corrections have tangential mass on a common face. Take conditional expectation, use (26i), Jensen $\sum_D |D| H(p_D) \leq \int_Q H(P + v_A)$, and $\sum_D |D| |p_D| \leq \int_Q H(P + v_A)$ since $H \geq |\cdot|$. Equations (22b9) and (25), followed by $\tau_{\text{sel}} \downarrow 0$, prove the upper inequality in (24).

Step 6. Lower bound: smooth the potential and freeze its slopes.

For the lower bound in (24), fix $\tau_{\text{sel}} > 0$, use (2) for $J_{\mathcal{H}_{h,L}\Psi}^Q(\zeta)$, and take a $\tau_{\text{sel}}|Q|$ -maximizing global potential F , normalized at one point. Extend it by successive even reflections across the square sides and then $2R$ -periodically, and convolve at scale A . Even reflection and convolution preserve the unit Lipschitz bound: explicitly, the coordinate-folding map from the torus to Q_R is 1-Lipschitz. If $F_A = \varphi_A * F^{\text{ref}}$, then $\|D^2 F_A\|_\infty \leq C/A$.

Reflecting P covariantly as in (22b1), Fubini gives the exact identity

$$\int_{\mathbb{T}_{2R}^2} P^{\text{ref}} \cdot \nabla F_A = \int_{\mathbb{T}_{2R}^2} P_A \cdot \nabla F^{\text{ref}}.$$

For $\mathcal{R} = -H^*$, concavity and reflection invariance give $\mathcal{R}(\nabla F_A) \geq \varphi_A * \mathcal{R}(\nabla F^{\text{ref}})$. Divide the resulting torus inequality by four. By (22b8), replacing P_A by P in the pairing changes its expected value by at most $CA|Q|/S$. The reflected convolution uses the entire square, so no outer collar is discarded.

Write $H = \mathcal{H}_{h,L}\Psi$, $\mathcal{R} = -H^*$, and $M_h = M + C\sqrt{h}$. Concavity of $\mathcal{R}$, even reflection, and (22) give for the smoothed potential

$$\mathbb{E} \int_Q \{P \cdot \nabla F_A + \mathcal{R}(\nabla F_A)\} \geq \mathbb{E} J_H^Q(\zeta) - C(A/S + \tau_{\text{sel}})|Q|, \quad (26a)$$

where $\tau_{\text{sel}}|Q|$ is the initial near-maximization error.

For every B -block D , put

$$a_D = \oint_D \nabla F_A.$$

Then $|a_D| \leq 1$, and Jensen's inequality for the concave function $\mathcal{R}$ gives

$$|D|\mathcal{R}(a_D) \geq \int_D \mathcal{R}(\nabla F_A). \quad (26b)$$

Choose measurably a trace $c_D \in \mathcal{T}_L^{\text{per}}(a_D)$ attaining $\mathcal{R}(a_D)$, and repeat one of its cell extensions as $a_D \cdot x + u_D(x)$. Since u_D is L -periodic and $\text{Lip}(u_D) \leq 2$, we may fix its additive constant so that $\|u_D\|_\infty \leq CL$. Let ℓ_D be affine with gradient a_D and the value of F_A at the centre of D , and put $f_D = \ell_D + u_D$, with phase aligned to the global L -grid. Since $\|D^2 F_A\|_\infty \leq C/A$,

$$\|F_A - f_D\|_{L^\infty(D)} \leq C(L + B^2/A). \quad (26c)$$

Step 7. Lower bound: join the fine potentials.

The potentials f_D may disagree on adjacent block boundaries. We prescribe the common value θF_A there and interpolate to θf_D inside each block. The contraction provides the slack needed for the interpolation gradient; away from these seams, the affine-periodic traces remain available for fine optimization.

Set $\theta = 1 - 4\varepsilon$ and

$$w = C(L + B^2/A)/\varepsilon < B/4$$

by (23a). Choose a cutoff χ_D which is zero on ∂D , one outside the width- w seam, and satisfies $|\nabla \chi_D| \leq C/w$. Define

$$G_D = \theta\{(1 - \chi_D)F_A + \chi_D f_D\}. \quad (26d)$$

The first part of its gradient is θ times a convex combination of two vectors in $\overline{B}_1$, while (26c) makes the cutoff term at most ε . Hence $|\nabla G_D| \leq 1 - 3\varepsilon$. Because $G_D = \theta F_A$ on every ∂D , the G_D 's paste into one global Lipschitz function. The L -cells meeting a width- w seam occupy at most a proportion $C(w + L)/B$ of each B -block. By the choice of w , this is at most

$$C\left(\frac{L}{\varepsilon B} + \frac{B}{\varepsilon A}\right). \quad (26e)$$

If an L -cell does not meet the width- w seam, its trace is, up to an additive constant, a translate of θc_D . Concavity of K and $K(0) \geq 0$ from (16d) give

$$K(\theta c_D) \geq \theta K(c_D) = \theta L^2 \mathcal{R}(a_D). \quad (26f)$$

Let $t_C = G_D|_{\partial C}$ be the trace on each L -cell and, conditionally on the coarse sigma-field, fill it by a measurable realization-wise optimizer for the independent fine Gaussian vector in that cell. Complete traces make the interiors paste exactly to a global potential $\tilde{F}$. The fixed-trace identity (21c) now gives, without interchanging a random maximum and expectation,

$$\mathbb{E}_\eta \left[\langle \eta, \tilde{F} \rangle - \int_Q \Psi^*(\nabla \tilde{F}) \Big| \mathcal{G} \right] = \sum_C K(t_C). \quad (26g)$$

If an L -cell meets the seam, its deterministic trace extension is admissible in $K(t_C)$. Its expected pairing with the centred fine source is zero, and $-\Psi^* \geq 0$, so $K(t_C) \geq 0$. Therefore (26b), (26e), and (26f), with $0 \leq \mathcal{R} \leq M_h$, imply

$$\sum_C K(t_C) \geq \theta \int_Q \mathcal{R}(\nabla F_A) - CM_h |Q| \left(\varepsilon + \frac{L}{\varepsilon B} + \frac{B}{\varepsilon A} \right). \quad (26h)$$

It remains to control the coarse-source pairing after replacing the interiors. Put $P_D = f_D P$. The potentials $\tilde{F}$ and θF_A have identical traces on ∂D , and hence

$$\int_D P_D \cdot \nabla(\tilde{F} - \theta F_A) = 0, \quad \int_D P \cdot \nabla \tilde{F} \geq \theta \int_D P \cdot \nabla F_A - 2 \int_D |P - P_D|. \quad (26k)$$

Averaging (22) over pairs in each B -block gives

$$\mathbb{E} \sum_D \int_D |P - P_D| \leq C(B/S) |Q|. \quad (26l)$$

Finally, multiplying the expected integral in (26a) by $\theta = 1 - 4\varepsilon$ changes its value by at most $C\varepsilon(M_h + \sqrt{T})|Q|$, because $0 \leq \mathcal{R} \leq M_h$ and $\sup_x \mathbb{E}|P(x)| \leq C\sqrt{T}$. Combine (26a), (26g)–(26l), and let $\tau_{\text{sel}} \downarrow 0$. This proves the lower inequality in (24) with the terms in (23).

The constructions above use choices depending on the coarse field or on individual fine-source realizations. We finish by verifying the measurability needed for their conditional expectations.

Step 8. Measurable selections.

We use the following selection principle. If S is standard Borel, X is a countable union of compact metrizable sets, $\Gamma \subset S \times X$ is Borel with nonempty closed sections in the relevant compact sets, and f is jointly Borel, lower semicontinuous in x , and has compact parameterwise sublevels, then the infimum is universally measurable and there is a universally measurable τ_{sel} -minimizer for every $\tau_{\text{sel}} > 0$. Apply this to $-f$ for upper semicontinuous maxima. Indeed, intersect rational sublevel sets with the compact exhaustion and select successively the first member of a countable finite-ball cover whose section is nonempty; the nested balls give the selector. On our completed Gaussian spaces universal measurability is sufficient.

For a closed block D , take $X_D = \mathcal{M}(\overline{D}; \mathbb{R}^2)$ with its weak- $*$ Borel structure. Since $C(\overline{D}; \mathbb{R}^2)$ is separable, every total-variation ball is compact metrizable, and these balls exhaust X_D . Choose a countable C^1 -dense family $\{\phi_k\}$. If the finite Gaussian vector g

parametrizes $\eta_D(g)$, the graph

$$\Gamma_D = \left\{ (g, u) : \int_{\overline{D}} \nabla \phi_k \cdot du = \langle \eta_D(g), \phi_k \rangle \text{ for every } k \right\} \quad (26m)$$

is closed and is exactly the divergence plus zero-normal constraint by density. The function

$$(p, g, u) \mapsto \mathcal{F}_\Psi(p\mathcal{L}^2 + u; D) \quad (26n)$$

is jointly Borel and weak-* lower semicontinuous by the perspective formula (22b5). Its sublevels are compact because

$$|u|(\overline{D}) \leq \mathcal{F}_\Psi(p\mathcal{L}^2 + u; D) + |p||D|. \quad (26o)$$

The explicit zero-normal Haar right inverse makes every constraint section nonempty. The principle therefore supplies a jointly universally measurable $u_D^{\tau_{\text{sel}}}(p, g)$ satisfying the constraints and

$$\mathcal{F}_\Psi(p\mathcal{L}^2 + u_D^{\tau_{\text{sel}}}; D) \leq I_B(p, \eta_D(g)) + \tau_{\text{sel}}|D|. \quad (26p)$$

This proves the measurability asserted in (26j), including composition with an unbounded coarse-measurable p_D .

For the effective primal problem use the same measure space on Q , the closed subspace

$$\mathcal{V}_Q = \left\{ v : \int_{\overline{Q}} \nabla \phi_k \cdot dv = 0 \text{ for every } k \right\},$$

and parameter space $L^1(Q; \mathbb{R}^2)$ for P . The map $(P, v) \mapsto \mathcal{F}_H(P\mathcal{L}^2 + v; Q)$ is jointly Borel, weak-* lower semicontinuous in v , and coercive exactly as in (26n)–(26o). We obtain a universally measurable $\tau_{\text{sel}}|Q|$ -minimizer $v^{\tau_{\text{sel}}}(P)$. Reflection, convolution, block averaging, and finite sums are Borel operations, so the entire competitor for the upper bound is measurable.

For dual selections, the spaces

$$\mathcal{A}_L = \{F \in C(\overline{Q}_L) : F(0) = 0, \text{Lip}(F) \leq 1\}, \quad \mathcal{T}_L = \{F|_{\partial Q_L} : F \in \mathcal{A}_L\}$$

are compact Polish in the uniform norm, and the fixed-trace graph $\{(c, F) : F|_{\partial Q_L} = c\}$ is closed. Uniform convergence in $\mathcal{A}_L$ implies weak-* convergence of gradients. Hence, for the cell Gaussian vector g ,

$$(g, F) \mapsto \langle \eta(g), F \rangle - \int_{Q_L} \Psi^*(\nabla F) \quad (26q)$$

is jointly Borel and upper semicontinuous in F . The maximum is attained on each compact section, and the selection principle above gives a jointly measurable exact optimizer $F^{\text{opt}}(c, g)$. For an unanchored trace, subtract its value at a fixed corner, apply the selector, and add the value back. This proves the exact conditional identity (26g), not merely an inequality between outer expectations.

The full affine-periodic trace graph

$$\mathfrak{G}_L = \{(a, c) \in \overline{B}_1 \times \mathcal{T}_L : c(L, t) - c(0, t) = La_1, c(t, L) - c(t, 0) = La_2\} \quad (26r)$$

is compact. The value function $K(c)$ is upper semicontinuous: use the upper semicontinuity of the realization-wise maximum and the pathwise, trace-independent domination

$$K^{\text{raw}}(c, \eta) \leq \int_{Q_L} (|Q| + M),$$

where K^{raw} denotes the maximum before expectation, in reverse Fatou. Its maximizing sections are nonempty compact, so there is a measurable $a \mapsto c^*(a)$ with $K(c^*(a)) = L^2\mathcal{R}(a)$. A measurable periodic interior extension is supplied explicitly by

$$E(c)(x) = \inf_{y \in \partial Q_L} \{c(y) + |x - y|\}$$

and affine-cocycle repetition; $c \mapsto E(c)$ is Lipschitz in the uniform norm on compact sets. This justifies c_D, f_D , and every cutoff in the lower bound. For the extension to a compact slope domain in Definition 4.1, one does not use the McShane formula. Instead restrict the potential correspondence to $\nabla F \in K_\Psi$; weak-* closure of the closed convex gradient constraint gives a compact closed graph, and the same selection principle supplies a compatible interior.

Finally, on the compact Polish space $\mathcal{A}_Q$, the effective dual objective

$$(P, F) \mapsto \int_Q P \cdot \nabla F + \int_Q \mathcal{R}(\nabla F) \quad (26s)$$

is jointly Borel and upper semicontinuous in F : the first term is continuous for $P \in L^1$, and $-\mathcal{R}$ is convex lower semicontinuous. It therefore has a coarse-measurable maximizing (or τ_{sel} -maximizing) potential. Thus all conditional expectations in (21c), (26i), and (26g) are legitimate. Letting the deterministic $\tau_{\text{sel}} \downarrow 0$ in the two competitor constructions proves (24).

■

6. Analytic consistency estimates

One group of h levels changes a profile viewed on the spatial scale $\sqrt{T}$ through a quadratic transport problem: a local Hessian $H/\sqrt{T}$ gives energy of order $h/\sqrt{T}$. Lemma 6.2 computes its coefficient and supplies moment bounds for the nonlinear errors. The upper estimate uses a Neumann flux and a local expansion, with Gaussian tails controlling the flux outside the expansion region. The lower estimate uses a zero-boundary truncation to keep a Dirichlet potential where the dual cost is quadratic. Lemma 6.5 then turns a local lower contact into the global order needed for monotonicity of the cell operator.

6.1. LIPSCHITZ TRUNCATION WITH ZERO BOUNDARY DATA. The truncation below is used in Lemma 6.4. Besides controlling the gradient, it preserves the zero boundary values that allow Dirichlet potentials on different blocks to be pasted.

LEMMA 6.1. *Let D be a square or flat torus and let $g \in W^{1,4}(D)$; in the square case assume $g \in W_0^{1,4}(D)$. For every $\lambda > 0$ there are an open set $E \subset D$ and a function g_λ such that, in the square case,*

$$g_\lambda \in W^{1,\infty}(D) \cap W_0^{1,4}(D) \quad \text{and its zero extension is globally Lipschitz,}$$

and in either case

$$\|\nabla g_\lambda\|_\infty \leq C\lambda, \quad g_\lambda = g \quad \text{on } D \setminus E, \quad (27)$$

and

$$|E| \leq C\lambda^{-4} \int_D |\nabla g|^4, \quad \int_D |\nabla g_\lambda - \nabla g|^2 \leq C\lambda^{-2} \int_D |\nabla g|^4. \quad (28)$$

Proof. We keep g where its gradient maximal function is at most λ and extend these values with a controlled Lipschitz constant. For a square, the extension must also take the prescribed value zero at every exterior point. We verify this compatibility before constructing the extension; the maximal estimate will then control both errors in (28).

Step 1. Lipschitz control on the good set.

In the boundary case let $\tilde{g}$ be the zero extension of g to $\mathbb{R}^2$. Since $g \in W_0^{1,4}(D)$, $\tilde{g} \in W^{1,4}(\mathbb{R}^2)$ and $\nabla \tilde{g} = 1_D \nabla g$. Set

$$\mathcal{O} = \{x \in \mathbb{R}^2 : \mathcal{M}(|\nabla \tilde{g}|)(x) > \lambda\}, \quad E = D \cap \mathcal{O}.$$

At Lebesgue points of any Sobolev function v , telescoping its averages over the balls $B(x, 2^{-k}|x - y|)$, and then comparing the averages at x and y on one ball of radius $3|x - y|$, gives

$$|v(x) - v(y)| \leq C|x - y|\{\mathcal{M}(|\nabla v|)(x) + \mathcal{M}(|\nabla v|)(y)\}.$$

Consequently, if $x, y \in D \setminus E$,

$$|g(x) - g(y)| \leq C\lambda|x - y|. \quad (28a)$$

To include zero exterior data in the extension, we also need a Lipschitz bound between a good interior point and a point outside D .

If $x \in D$ and $d = \text{dist}(x, D^c)$, then $B(x, 2d) \setminus D$ occupies a fixed fraction of $B(x, 2d)$. Telescoping from x to the average on that ball, followed by L^1 -Poincare and $\tilde{g} = 0$ on the exterior fraction, gives

$$|\tilde{g}(x)| \leq Cd\mathcal{M}(|\nabla \tilde{g}|)(x).$$

Thus, if also $x \notin \mathcal{O}$ and $y \in D^c$,

$$|g(x)| \leq C\lambda d \leq C\lambda|x - y|. \quad (28b)$$

Step 2. Extension with exact zero trace.

It follows that the prescribed data

$$g \text{ on } D \setminus \mathcal{O}, \quad 0 \text{ on } \mathbb{R}^2 \setminus D$$

are $C\lambda$ -Lipschitz on their union. Let $A = (D \setminus \mathcal{O}) \cup D^c$, let f denote these data, and take the McShane extension

$$\tilde{g}_\lambda(x) = \inf_{a \in A} \{f(a) + C\lambda|x - a|\}.$$

Every exterior point was prescribed zero, so $\tilde{g}_\lambda = 0$ on $\mathbb{R}^2 \setminus D$. Consequently $g_\lambda = \tilde{g}_\lambda|_D$ belongs to $W^{1,\infty}(D)$, agrees with g off E , and satisfies $\|\nabla g_\lambda\|_\infty \leq C\lambda$. Moreover $|g_\lambda(x)| \leq C\lambda \text{dist}(x, D^c)$. If χ_k is zero at distance $< 1/k$, one at distance $> 2/k$, and $|\nabla \chi_k| \leq Ck$, then $\chi_k g_\lambda \rightarrow g_\lambda$ in $W^{1,4}(D)$: both gradient-error terms are bounded and supported on a boundary strip whose measure tends to zero, while $|g_\lambda \nabla \chi_k| \leq C\lambda$ there. Hence after mollifying each compactly supported Lipschitz cutoff, $g_\lambda \in W_0^{1,4}(D)$. The required zero-trace property is therefore in $W^{1,4}$; no closure assertion in the $W^{1,\infty}$ -norm is needed.

Step 3. Exceptional-set and gradient estimates.

The strong L^4 maximal estimate followed by Chebyshev gives the desired measure bound:

$$|E| \leq C\lambda^{-4} \int_D |\nabla g|^4. \quad (28c)$$

Since the gradients agree off E ,

$$\begin{aligned} \int_D |\nabla g_\lambda - \nabla g|^2 &\leq C\lambda^2 |E| + C \int_E |\nabla g|^2 \\ &\leq C\lambda^{-2} \int_D |\nabla g|^4, \end{aligned}$$

where the last step is Cauchy–Schwarz followed by (28c). On a flat torus use the periodic maximal function and the periodic McShane extension of the good-set data; the same calculation proves (27)–(28).

■

6.2. QUADRATIC ENERGY AND GAUSSIAN MOMENTS. The quadratic cost with matrix H has dual cost with matrix H^{-1} , so its optimizer solves the elliptic equation with operator L_H below. In dimension two, the trace of the inverse operator over the cell-constant modes grows linearly in h as the cell resolution 2^{-h} decreases. Its coefficient will be the generator of the limiting equation. We must show that this coefficient is independent of the boundary trace and the period L , and bound moments of the optimizers to estimate the nonlinear remainders.

LEMMA 6.2. *Let H be a positive definite 2×2 matrix and $n = 2^h$. On the unit square let Π_h be orthogonal projection onto functions constant on the n^2 small squares, followed by removal of the constants, and put*

$$L_H = -\operatorname{div}(H^{-1}\nabla).$$

If $E_{h,H}^N$ is the expected minimum of $\frac{1}{2} \int u^T H u$ under $-\operatorname{div} u = \eta_h$ and zero normal boundary condition, and $E_{h,H}^D$ is the corresponding Dirichlet dual value, then

$$E_{h,H}^N = \frac{1}{2} \operatorname{Tr}(\Pi_h(L_H^N)^{-1}), \quad E_{h,H}^D = \frac{1}{2} \operatorname{Tr}(\Pi_h(L_H^D)^{-1}), \quad (29)$$

and

$$E_{h,H}^D, E_{h,H}^N = \frac{h \log 2}{4\pi} \sqrt{\det H} + o(h). \quad (30)$$

The error is uniform for H in compact subsets of the positive definite cone. More precisely, define the deterministic full-trace quadratic value

$$\mathcal{E}_{h,L}^{\operatorname{tr}}(H) = L^{-2} \max_c \mathbb{E} \sup_{f|_{\partial Q_L}=c} \left\{ \int f \eta_h^{(L)} - \frac{1}{2} \int \nabla f^T H^{-1} \nabla f \right\}, \quad (30a)$$

where c ranges over all affine-periodic complete traces (the affine slope is optimized as part of c). Then

$$E_{h,H}^D \leq \mathcal{E}_{h,L}^{\operatorname{tr}}(H) \leq E_{h,H}^N, \quad (30b)$$

so it has the leading term (30), uniformly in the auxiliary period L .

The Gaussian primal and dual optimizers used in (29) also satisfy

$$\mathbb{E} \oint |u|^r \leq C_{H,r} h^{r/2}, \quad r = 2, 3, 4. \quad (31)$$

Proof. We first express each expected quadratic energy as a trace and compute its logarithmic leading term. The Dirichlet and Neumann values then bound the complete-trace value, uniformly in L . The remaining task is (31): a pointwise variance estimate for the Gaussian optimizers will imply all three moment bounds.

Step 1. Variational trace identities.

For the Neumann problem, introduce a multiplier f , minimize first in u , and obtain

$$\sup_{\int f=0} \left\{ \langle \eta_h, f \rangle - \frac{1}{2} \int \nabla f^T H^{-1} \nabla f \right\}.$$

The maximizing potential solves $L_H f = \eta_h$, and the value is $\frac{1}{2} \langle \eta_h, L_H^{-1} \eta_h \rangle$. Since the covariance of η_h is Π_h , taking expectation proves the first identity in (29). The same computation with zero boundary values proves the Dirichlet identity.

Step 2. Heat cutoff and the logarithmic coefficient.

It suffices to replace the cell projection in (29) by a heat cutoff at time n^{-2} with bounded error. This exposes the logarithmic contribution of the scales between n^{-2} and 1. Write P_n for projection onto the n^2 cell-constant functions and P_0 for the constants, so $\Pi_h = P_n - P_0$. Cell Poincare and ellipticity give, for either boundary condition,

$$\|(I - P_n)f\|_2^2 \leq Cn^{-2} \|\nabla f\|_2^2 \leq C_H n^{-2} \langle L_H f, f \rangle. \quad (32a)$$

Consequently

$$\text{Tr}((I - P_n)e^{-tL_H}) = \|(I - P_n)e^{-tL_H/2}\|_{\text{HS}}^2 \leq C_H n^{-2} \text{Tr}(L_H e^{-tL_H}). \quad (32b)$$

In the Neumann case all traces omit the constant mode. The contribution of $0 < t < n^{-2}$ to $\int_0^\infty \text{Tr}(P_n e^{-tL_H}) dt$ is at most one. Integrating (32b) from n^{-2} to infinity gives

$$C_H n^{-2} \int_{n^{-2}}^\infty \text{Tr}(L_H e^{-tL_H}) dt = C_H n^{-2} \text{Tr}(e^{-n^{-2}L_H}) = O_H(1). \quad (32c)$$

For Dirichlet conditions, replacing P_n by $P_n - P_0$ subtracts the bounded rank-one term

$$|Q_1|^{-1} \langle 1, (L_H^D)^{-1} 1 \rangle.$$

The spectral gap absorbs the tail $t \geq 1$. Therefore

$$\text{Tr}(\Pi_h L_H^{-1}) = \int_{n^{-2}}^1 \text{Tr}'(e^{-tL_H}) dt + O_H(1). \quad (32d)$$

The heat-trace coefficient $\sqrt{\det H}/(4\pi t)$, together with the factor $1/2$ in the energy and the cutoff interval $[n^{-2}, 1]$, gives

$$\frac{1}{2} \frac{\sqrt{\det H}}{4\pi} \int_{n^{-2}}^1 \frac{dt}{t} = \frac{\sqrt{\det H}}{4\pi} \log n.$$

The factor $\sqrt{\det H}$ comes from the area of the frequency ellipse for L_H . We verify the heat-trace coefficient by the following bracketing argument. The linear change of variables turning L_H into the Laplacian maps the square to a parallelogram. Tile it by squares of side r and bracket by Dirichlet and Neumann direct sums over squares contained in or meeting the parallelogram. Interior lattice counts are Riemann sums for

$$(2\pi)^{-2} \int_{\{\xi^T H^{-1} \xi \leq 1\}} d\xi = \frac{\sqrt{\det H}}{4\pi}.$$

The squares meeting the boundary contribute at most $C_H(rR + r^{-1}\sqrt{R} + r^{-2})$ eigenvalues below energy R . Divide by R , first send $R \rightarrow \infty$, and then $r \downarrow 0$. Thus, for both boundary conditions,

$$N_H(R) = \frac{\sqrt{\det H}}{4\pi} R + o_H(R), \quad (32e)$$

uniformly on compact ellipticity classes. Equivalently, $t \operatorname{Tr}'(e^{-tL_H}) \rightarrow \sqrt{\det H}/(4\pi)$. Logarithmic Cesaro averaging in (32d) yields

$$\frac{1}{2} \operatorname{Tr}(\Pi_h L_H^{-1}) = \frac{\sqrt{\det H}}{4\pi} \log n + o_H(\log n). \quad (32)$$

Since $\log n = h \log 2$, (30) follows.

Step 3. Dirichlet–Neumann squeeze.

For the lower bound in (30b), paste the zero-boundary Dirichlet potentials on the unit outer blocks. Their common outer trace is zero, so they are admissible in (30a), and their expected values add. For the upper bound, paste the zero-normal Neumann flows instead. Against any admissible full trace, integration by parts and quadratic completion bound the dual value by the sum of their energies. Both bounds are proportional to the number L^2 of unit blocks, so the normalization in (30a) makes the squeeze uniform in L . No random spatial mean is subtracted: the fields retain their zero-normal boundary condition.

Step 4. Pointwise Gaussian moments.

The optimizers are linear images of the Gaussian source, so their pointwise moments are determined by the sum of squared coefficients. We will show that this sum is $O_H(1 + \log n)$. The spatial estimate needed for both boundary conditions is the Green-gradient bound $C_H|x - y|^{-1}$ in (32f): summing its squared cell contributions over concentric annuli will produce the logarithm in (32h).

We first record the Green estimates and check that they apply after the anisotropic change of variables.

Dirichlet Green estimate. Diening, Rolfes and Salgado [DRS24, Proposition 4.1] state: if $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$, is a convex polytope and $G : \Omega \times \Omega \rightarrow \mathbb{R}$ is its Green function (in the paper's $W_0^{1,2}$ framework, hence the Dirichlet Green function of the Laplacian), then for every i ,

$$|\partial_{x_i} G(x, \xi)| \leq \frac{C_\Omega}{|x - \xi|^{d-1}}. \quad (32fD)$$

In addition, there is $\alpha \in (0, 1]$, depending only on the inner angles of Ω , such that for every i, j and all $x, y, \xi \in \Omega$ off the displayed diagonals,

$$\begin{aligned} \frac{|\partial_{x_i} G(x, \xi) - \partial_{y_i} G(y, \xi)|}{|x - y|^\alpha} &\leq C_\Omega (|x - \xi|^{-d-\alpha+1} + |y - \xi|^{-d+\alpha+1}), \\ \frac{|\partial_{x_i} \partial_{\xi_j} G(x, \xi) - \partial_{y_i} \partial_{\xi_j} G(y, \xi)|}{|x - y|^\alpha} &\leq C_\Omega (|x - \xi|^{-d-\alpha} + |y - \xi|^{-d+\alpha}). \end{aligned}$$

Only (32fD), the first assertion, is used below. The proof of the cited proposition points to the convex-polytope Green estimates identified there; no smooth-boundary hypothesis is present.

Neumann heat-gradient estimate. Wang and Yan [WY10, Theorem 1.2] use $P_t = e^{t\Delta_N}$ for the Neumann heat semigroup on Ω , $p_t(x, y)$ for its kernel, and $\|\cdot\|_{1 \rightarrow \infty}$ for the L^1 -to- L^∞ operator norm. Its complete assertion is: if Ω is a compact convex domain in $\mathbb{R}^d$, then there are $C, c > 0$ such that, for every $t > 0$,

$$\|\nabla P_t\|_{1 \rightarrow \infty} \leq Ct^{-(d+1)/2}, \quad |\nabla_x p_t(x, y)| \leq Ct^{-(d+1)/2} \exp\left(-\frac{|x - y|^2}{ct}\right). \quad (32fN)$$

The theorem explicitly covers nonsmooth convex domains. Its proof constructs reflecting Brownian motion on a convex domain, obtains the gradient contraction and the short-time heat-kernel bound, and uses the Neumann spectral gap for large time. Section 4.3 of the same paper defines the mean-zero Neumann Green operator and records the Green-gradient consequence. Indeed, in dimension two, with the normalization

$$G_{\Omega}^N(z, w) = \int_0^{\infty} \{p_t(z, w) - |\Omega|^{-1}\} dt,$$

the derivative of the constant vanishes and (32fN) is integrable for $z \neq w$; differentiation first on a truncated time interval and then dominated convergence give

$$|\nabla_z G_{\Omega}^N(z, w)| \leq C \int_0^{\infty} t^{-3/2} e^{-|z-w|^2/(ct)} dt = \frac{C'}{|z-w|}. \quad (32fN')$$

We now check the hypotheses and the change of variables. Let $S = H^{1/2}$ and $\mathcal{P}_H = SQ_1$. The variational change $z = Sx$ transforms $L_H = -\operatorname{div}_x(H^{-1}\nabla_x)$ into the Euclidean Laplacian on the parallelogram $\mathcal{P}_H$, preserving respectively Dirichlet and natural Neumann boundary conditions. The parallelogram is a compact convex polygon, so (32fD) and (32fN') apply, including at its corners. Green-kernel normalization gives, for either boundary condition,

$$G_H^{\#}(x, y) = (\det S) G_{\mathcal{P}_H}^{\#}(Sx, Sy).$$

Therefore

$$|H^{-1}\nabla_x G_H^{\#}(x, y)| = \det(S) |H^{-1/2}\nabla_z G_{\mathcal{P}_H}^{\#}(Sx, Sy)| \leq C_H |x - y|^{-1}. \quad (32f)$$

Thus (32f) holds for the Dirichlet kernel and the mean-zero Neumann kernel. Its constant is allowed to depend on the fixed positive-definite matrix H , as in (31). We now apply this decay bound to the cell coefficients.

If $e_C = n\mathbf{1}_C$, the coefficient of the cell Gaussian Z_C in the primal optimizer is

$$k_C(x) = n \left(\int_C H^{-1}\nabla_x G_H^{\#}(x, y) dy - n^{-2} \int_{Q_1} H^{-1}\nabla_x G_H^{\#}(x, y) dy \right). \quad (32g)$$

For the mean-subtraction term, $n^{-1} \int_{Q_1} |x - y|^{-1} dy \leq C/n$ uniformly in x . Its squared contribution summed over all n^2 coefficients is $O(1)$. For every cell meeting $B(x, 2/n)$, polar integration gives $\int_C |x - y|^{-1} dy \leq C/n$; hence its first term in (32g) is $O_H(1)$, and the boundedly many such cells contribute $O_H(1)$ to $\sum_C |k_C(x)|^2$. In the annulus $j/n \leq \operatorname{dist}(x, C) < (j+1)/n$, there are $O(j)$ cells and (32f) gives $n \int_C |x - y|^{-1} dy \leq C/j$. Together with the $O(1/n)$ mean term, this gives $|k_C(x)| \leq C_H/j$. Summing $\sum_{j \leq n} j \cdot j^{-2}$ gives

$$\sup_x \sum_C |k_C(x)|^2 \leq C_H(1 + \log n). \quad (32h)$$

Gaussian moment identities, followed by spatial integration, prove (31) for both primal and dual optimizers.

■

6.3. UPPER CONSISTENCY FROM A NEUMANN FLUX. The upper estimate uses the profile only near the point being tested. A quadratic Neumann optimizer has size of order $\sqrt{h}$, whereas the expansion in (33a) holds up to radius $r_0\sqrt{T}$. When $h/T \rightarrow 0$, its Gaussian tail outside this radius is negligible. The argument also applies when the contact slope has norm one.

LEMMA 6.3. *Put*

$$\mathcal{G}(H) = \frac{1}{4\pi} \sqrt{\det H}, \quad H \in \mathbb{S}_+^2. \quad (33)$$

Let F_T be a convex 1-Lipschitz profile, $p_T = \sqrt{T}q_0$, and suppose that, for $|z| \leq r_0\sqrt{T}$,

$$F_T(p_T + z) - F_T(p_T) \leq a \cdot z + \frac{1}{2\sqrt{T}} z^T H z + C \frac{|z|^3}{T}. \quad (33a)$$

The case $|a| = 1$ is included. If $H > 0$, $h \rightarrow \infty$, and $h/T \rightarrow 0$, then, uniformly in L ,

$$\begin{aligned} (\mathcal{H}_{h,L}F_T)(p_T) - F_T(p_T) &\leq \frac{h \log 2}{\sqrt{T}} \mathcal{G}(H) + \frac{o_H(h)}{\sqrt{T}} \\ &\quad + C_H \frac{h^{3/2}}{T} + C_H \sqrt{h} e^{-c_H T/h}. \end{aligned} \quad (33b)$$

For singular H , the same upper consistency follows by using $H + \gamma I$, taking the scaling limit first, and then $\gamma \downarrow 0$.

Proof. For every realization and every zero-normal flux u with $-\operatorname{div} u = \eta_h^{(L)}$, integration by parts and Fenchel's inequality give the general primal domination

$$(\mathcal{H}_{h,L}F_T)(p) \leq \mathbb{E} \int_{Q_L} F_T(p + u(x)) dx. \quad (33c)$$

Choose on every unit outer block the Neumann quadratic minimizer for $\frac{1}{2} \int u^T H u$, and paste these zero-normal fields. This is a linear centred Gaussian function of the source; hence $\mathbb{E} \int u = 0$. Let $B = \{(\omega, x) : |u(\omega, x)| > r_0\sqrt{T}\}$. On B^c , insert (33a); on B , use the 1-Lipschitz bound. The full linear term vanishes in expectation, so its contribution on B^c can be estimated on B through the identity

$$\mathbb{E} \int 1_{B^c} a \cdot u = -\mathbb{E} \int 1_B a \cdot u.$$

This cancellation uses the whole probability and spatial domain; the linear term need not vanish on B^c alone.

Apply (33a) with $z = tv$ and $z = -tv$ and let $t \downarrow 0$. The 1-Lipschitz bound gives $|a \cdot v| \leq |v|$ for every v , so $|a| \leq 1$. Hence the transferred linear term and the Lipschitz contribution on B are both bounded by $\mathbb{E} \int 1_B |u|$. The positive quadratic term on B^c is at most its full-space value, evaluated by (30), while (31) bounds the cubic remainder by $C_H h^{3/2}/T$. At every fixed x , (32h) gives $\operatorname{Var} u(x) \leq C_H h$. Gaussian tails applied to the two terms supported on B give $C_H \sqrt{h} e^{-c_H T/h}$. This proves (33b) for $H > 0$. If H is singular, the Taylor quadratic is bounded above by the one with $H + \gamma I$; apply the preceding estimate with fixed $\gamma > 0$, use (30), and only then send $\gamma \downarrow 0$.

■

6.4. CONSISTENCY FROM A TRUNCATED DIRICHLET POTENTIAL. For a lower bound we need an admissible dual potential, not just a bound on its expected gradient. The conjugate cost D in (34) equals the quadratic cost Q near the contact slope and satisfies $D \geq Q$ on its effective domain. The latter inequality bounds the cell value by the quadratic value. For the reverse bound, we truncate a Dirichlet optimizer so that every gradient lies where $D = Q$. These profiles are used for the lower barriers in the next subsection.

LEMMA 6.4. Use $\mathcal{G}$ from (33). Let ψ be either a global convex 1-Lipschitz profile satisfying $|q| \leq \psi(q) \leq |q| + M$ for some M , or a proper closed convex profile whose conjugate domain is compactly contained in B_1 , let $q_0 \in \mathbb{R}^2$, and suppose $a_0 \in \partial\psi(q_0)$ lies in the open unit ball. Assume that, on $\bar{B}_1 - a_0$, its conjugate has the exact local form

$$\psi^*(a_0 + b) = \psi^*(a_0) + q_0 \cdot b + D(b), \quad (34)$$

where $D(0) = 0$, D is convex, and, for some $\rho > 0$,

$$D(b) = \frac{1}{2}b^T H^{-1}b \quad \text{if } |b|_{H^{-1}} \leq \rho. \quad (34a)$$

Assume on its whole effective domain that

$$D(b) \geq Q(b) := \frac{1}{2}b^T H^{-1}b;$$

thus D is the quadratic near zero plus a convex exterior wall. Put $\Psi_T(p) = \sqrt{T} \psi(p/\sqrt{T})$ and $p_T = \sqrt{T}q_0$. If $H > 0$, $h \rightarrow \infty$, and $T/h \rightarrow \infty$, then, uniformly in L ,

$$\begin{aligned} & (\mathcal{H}_{h,L}\Psi_T)(p_T) - \Psi_T(p_T) \\ &= \frac{h \log 2}{\sqrt{T}} \mathcal{G}(H) + \frac{o_H(h)}{\sqrt{T}} + O_H\left(\frac{h^2}{T^{3/2}}\right). \end{aligned} \quad (35)$$

This lemma is used only with $H > 0$ and an interior dual slope.

Proof. After removing the affine part in (34), we will bound the increment between the Dirichlet and Neumann quadratic values, divided by $\sqrt{T}$, with an additional lower-bound error of order $h^2/T^{3/2}$. Since both quadratic values have the leading term (30), this proves (35). The only change needed in the quadratic Dirichlet competitor is a truncation that keeps its gradient in the set where $D = Q$.

Step 1. Dual reduction and quadratic value.

Since $\Psi_T^* = \sqrt{T}\psi^*$, the definition (16) and (34) turn the increment on the left of (35) exactly into

$$\mathcal{V}_D = L^{-2} \max_c \mathbb{E} \sup_{f|_{\partial Q_L=c}} \left\{ \int f \eta_h^{(L)} - \sqrt{T} \int D(\nabla f) \right\}. \quad (36)$$

The term $p_T \cdot a$ cancels the spatial average of the linear term in (34), and the remaining affine source term has expectation zero. For $Q(b) = \frac{1}{2}b^T H^{-1}b$, the change of variable $g = \sqrt{T}f$ and (30a) give

$$\mathcal{V}_Q = T^{-1/2} \mathcal{E}_{h,L}^{\text{tr}}(H). \quad (37)$$

Lemma 6.2 evaluates this quadratic value. To obtain a lower bound for the nonlinear value, we now keep the Dirichlet competitor inside the region where the costs agree.

Step 2. Truncated Dirichlet competitor.

Let f_D be the zero-boundary Dirichlet dual optimizer for the unscaled quadratic problem, pasted over the unit blocks, and put $b_D = T^{-1/2} \nabla f_D$. Equation (31) and the pointwise bound (32h) give

$$\mathbb{E} \int |b_D|^4 \leq C_H (h/T)^2, \quad \sup_x \mathbb{E} |b_D(x)|^2 \leq C_H h/T. \quad (37a)$$

These moment bounds do not ensure that the Gaussian optimizer stays pointwise in the set where $D = Q$. Choose $\lambda > 0$, independently of h and T , so that the bound $C\lambda$ is smaller than both $\rho/\|H^{-1/2}\|$ and $\text{dist}(a_0, \partial B_1)$. Apply Lemma 6.1 to $f_D/\sqrt{T}$ and denote the resulting potential on each block by g_λ . Each g_λ belongs to $W_0^{1,4}$ on its block and has a

globally Lipschitz zero extension, so the truncated block potentials paste with exactly zero trace. On each unit block, quadratic completion shows that the dual objective at $f_D/\sqrt{T}$ exceeds its value at g_λ by

$$\frac{\sqrt{T}}{2} \int (\nabla g_\lambda - b_D)^T H^{-1} (\nabla g_\lambda - b_D).$$

The integral is over that block. By (28) and (37a), the expected squared gradient error is at most $C_H(h/T)^2$ per unit area. The lost value is therefore at most $C_H h^2/T^{3/2}$.

Step 3. Comparison of the two values.

Since $D \geq Q$, (37) and the upper bound in (30b) bound $\mathcal{V}_D$ by $T^{-1/2} E_{h,H}^N$. The truncated competitor has zero trace and stays everywhere in the region where $D = Q$, so its value is at least $T^{-1/2} E_{h,H}^D$ minus the preceding error. Applying (30) to these two bounds proves (35). The construction and the error estimate are blockwise, hence uniform in L .

■

6.5. GLOBAL BARRIERS FOR LOCAL LOWER CONTACTS. Monotonicity of the cell operator compares profiles on the whole space, whereas a viscosity test initially gives only a local order relation. We therefore replace a local lower test by a global convex profile below the spatial slice. It must retain almost the same Hessian at the contact, meet the hypotheses of Lemma 6.4, and grow strictly more slowly than the slice at infinity. The last property keeps the later minimizing points in a compact set. The uniform estimate (38c) is needed because these points will approach, rather than equal, the original contact point.

LEMMA 6.5. *Let U be convex and 1-Lipschitz, with $|p| - C \leq U(p) \leq |p| + C$. Suppose a convex C^2 function ϕ touches U strictly from below at q_0 , $|a_0| := |\nabla \phi(q_0)| < 1$, and $H = D^2 \phi(q_0) > 0$. For every $0 < \delta < \lambda_{\min}(H)$ there is a global proper convex barrier $B_\delta \leq U$, touching at q_0 , such that*

$$B_\delta(q) < U(q) \quad (q \neq q_0), \quad D^2 B_\delta(q_0) = H - \delta I, \quad (38a)$$

and

$$\text{dom } B_\delta^* \Subset B_1, \quad U(q) - B_\delta(q) \longrightarrow +\infty \quad \text{as } |q| \rightarrow \infty. \quad (38b)$$

Near a_0 , its conjugate is exactly quadratic with Hessian $(H - \delta I)^{-1}$. Its conjugate Bregman gap—that is, $B_\delta^*(a+b) - B_\delta^*(a) - \nabla B_\delta^*(a) \cdot b$ —based at every slope sufficiently near a_0 , is bounded below globally by this same quadratic form and equals it for small increments. Consequently, for $(B_\delta)_T(p) = \sqrt{T} B_\delta(p/\sqrt{T})$,

$$(\mathcal{H}_{h,L}(B_\delta)_T)(\sqrt{T}q_0) - (B_\delta)_T(\sqrt{T}q_0) = \frac{h \log 2}{\sqrt{T}} \mathcal{G}(H - \delta I) + \frac{o_{H,\delta}(h)}{\sqrt{T}} + O_{H,\delta} \left(\frac{h^2}{T^{3/2}} \right). \quad (38)$$

Moreover, there is a neighborhood V of q_0 on which

$$\sup_{q \in V} \left| \frac{\sqrt{T}}{h \log 2} \left\{ (\mathcal{H}_{h,L}(B_\delta)_T)(\sqrt{T}q) - (B_\delta)_T(\sqrt{T}q) \right\} - \mathcal{G}(H - \delta I) \right| \longrightarrow 0 \quad (38c)$$

whenever $h \rightarrow \infty$ and $h/T \rightarrow 0$, uniformly in L . No global upper barrier is asserted or needed.

Proof. Write $g = U^*$. Since conjugation reverses order, it is enough to construct a convex function above g whose conjugate has the desired contact. The local test gives a quadratic upper bound for g near a_0 . We extend that bound over a compact ball inside B_1 , with

strict inequality away from a_0 , and set it equal to $+\infty$ outside the ball. This choice will give both the global order and the slower growth required in (38b).

Step 1. The dual contact and a strict quadratic bound.

The conjugate g is finite and continuous in B_1 and is $+\infty$ off $\overline{B_1}$. The local lower contact and convexity give $a_0 \in \partial U(q_0)$, so q_0 maximizes

$$g(a_0) = \sup_q \{a_0 \cdot q - U(q)\}. \quad (39)$$

It is the unique maximizer. Indeed, a second maximizer would make U affine with slope a_0 on the segment joining it to q_0 . The positive-definite quadratic term of the lower test would then put ϕ strictly above that affine function immediately away from q_0 , contrary to the local lower contact.

Choose $r < 1$ so that a neighborhood of a_0 is contained in B_r . From the lower cone bound,

$$a \cdot q - U(q) \leq -(1-r)|q| + C \quad (a \in B_r). \quad (40)$$

Thus all maximizers defining $g(a)$ for a near a_0 lie in one compact set. Uniqueness in (39) further implies that these maximizers converge to q_0 as $a \rightarrow a_0$. They consequently lie, after the slope neighborhood is reduced, in a neighborhood V of q_0 where $U \geq \phi$. If $q_a \in V$ maximizes the expression defining $g(a)$, then

$$g(a) = a \cdot q_a - U(q_a) \leq \sup_{q \in V} \{a \cdot q - \phi(q)\}.$$

The local Legendre transform of ϕ has Hessian H^{-1} at a_0 . Comparing first with $H - \frac{\delta}{2}I$, and only then with $H - \delta I$, gives, for all sufficiently small v ,

$$\begin{aligned} g(a_0 + v) - g(a_0) - q_0 \cdot v &\leq \frac{1}{2}v^T(H - \frac{\delta}{2}I)^{-1}v \\ &\leq \frac{1}{2}v^T(H - \delta I)^{-1}v - c_\delta|v|^2 \end{aligned} \quad (41)$$

for some $c_\delta > 0$. The strict second inequality is the ingredient that prevents any further contact.

Step 2. Extension beyond the contact slope.

We now extend the quadratic bound while leaving it unchanged near a_0 . Choose the radius of the quadratic region and the compact dual domain first; then choose the coefficient of a convex term to dominate the remaining annulus.

Put $A_\delta = (H - \delta I)^{-1}$ and $Q_\delta(v) = \frac{1}{2}v^T A_\delta v$. Choose

$$0 < 2\rho < d < \text{dist}(a_0, \partial B_1)$$

so that (41) holds for $|v| \leq 2\rho$, and, on $|v| \leq d$, set

$$D(v) = Q_\delta(v) + C_\rho(|v| - \rho)_+^2. \quad (41a)$$

This is convex and equals Q_δ on $|v| \leq \rho$. Choose $C_\rho > 0$ strictly larger than

$$\max \left\{ 0, \sup_{2\rho \leq |v| \leq d} \frac{g(a_0 + v) - g(a_0) - q_0 \cdot v - Q_\delta(v)}{(|v| - \rho)^2} \right\}. \quad (41b)$$

The supremum is finite by continuity on the compact annulus. For $0 < |v| \leq 2\rho$, (41) gives the strict inequality below; for $2\rho \leq |v| \leq d$, it follows from the strict choice of C_ρ

in (41b). Thus

$$D(v) > g(a_0 + v) - g(a_0) - q_0 \cdot v \quad (0 < |v| \leq d). \quad (41c)$$

Step 3. The global lower barrier.

Define the proper closed convex function

$$\tilde{g}(a_0 + v) = g(a_0) + q_0 \cdot v + D(v) \quad (|v| \leq d), \quad \tilde{g} = +\infty \quad (|v| > d), \quad (41d)$$

and set $B_\delta = \tilde{g}^*$. We have $\tilde{g} \geq g$, with equality among its effective slopes only at a_0 ; hence $B_\delta \leq U$. At q_0 , the expression defining $B_\delta(q_0)$ is $U(q_0) - D(v)$, so its unique maximizing slope is a_0 and $B_\delta(q_0) = U(q_0)$.

If equality $B_\delta(q) = U(q)$ held elsewhere, a maximizing slope for the compact-domain conjugate $B_\delta(q)$ would have to be a slope where $\tilde{g} = g$, hence a_0 . But $\tilde{g}$ is differentiable at a_0 , with gradient q_0 , so a_0 can maximize its conjugate only at $q = q_0$. This proves the strict assertion in (38a). Since (41d) is exactly affine plus Q_δ near a_0 , local conjugacy gives $D^2 B_\delta(q_0) = H - \delta I$. Also $\text{dom } B_\delta^* = \overline{B}_a(a_0) \Subset B_1$, so $B_\delta(q) \leq r_\delta |q| + C_\delta$ for some $r_\delta < 1$. The lower cone bound for U proves (38b).

Step 4. Uniform Bregman gap and consistency.

The discrete contact will occur at slopes near a_0 , so consistency at a_0 alone is insufficient. We check that the conjugate Bregman gap is at least $Q_\delta(b)$, with equality for small b , uniformly at these nearby slopes. Write $W(v) = C_\rho(|v| - \rho)_+^2$ for the second term in (41a). If $|a - a_0| < \rho/4$, then, for every increment b for which the left side is finite,

$$\begin{aligned} B_\delta^*(a + b) - B_\delta^*(a) - \nabla B_\delta^*(a) \cdot b \\ = Q_\delta(b) + \{W(a - a_0 + b) - W(a - a_0) - \nabla W(a - a_0) \cdot b\} \geq Q_\delta(b). \end{aligned} \quad (41e)$$

For increments outside the compact conjugate domain the left side is $+\infty$. If $|a - a_0| < \rho/4$ and $|b| < \rho/2$, both $W(a - a_0)$ and $W(a - a_0 + b)$ vanish, so equality holds in (41e). Thus the hypotheses and all constants in Lemma 6.4 hold uniformly for the corresponding spatial points q near q_0 . The extension of $\mathcal{H}_{h,L}$ to profiles with compact conjugate domain in Definition 4.1 has an attained cell value and is monotone, as proved in Proposition 4.2. Lemma 6.4 gives (38), and, after division by $h \log 2/\sqrt{T}$, gives (38c), because its two normalized errors are $o(1)$ and $O(h/T)$.

■

7. The convex viscosity limit

We now show that many small cell-operator increments converge to the evolution with generator $\mathcal{G}$. There are two tasks: identify the equation satisfied in the limit, and prove uniqueness in the class of convex profiles with bounded corrections to $|q|$. The first uses the local Gaussian upper estimate and the global lower barriers from the preceding section. The second requires a comparison argument on $\mathbb{R}^2$, where a maximizing pair of spatial points can escape to infinity together.

LEMMA 7.1. *Let $h_m \rightarrow \infty$, $r_m \rightarrow \infty$, and $h_m/r_m \rightarrow 0$, and let L_m be arbitrary integers tending to infinity. For each m , define the m th row of the triangular monotone scheme by*

$$F_0^{(m)}(p) = |p|, \quad F_{k+1}^{(m)} = \mathcal{H}_{h_m, L_m} F_k^{(m)}, \quad t_k^{(m)} = k h_m \log 2. \quad (42)$$

Thus h_m and L_m are fixed during the iteration in row m , while the row parameters vary with m . Use piecewise-constant interpolation in the physical time t and then rescale time by r_m ; equivalently, U_m below is constant on each interval $[t_k^{(m)}/r_m, t_{k+1}^{(m)}/r_m)$. Thus

$$U_m(s, q) = r_m^{-1/2} F_{\lfloor sr_m/(h_m \log 2) \rfloor}^{(m)}(\sqrt{r_m} q) \quad (43)$$

converges locally uniformly, at every fixed positive time, to the unique viscosity solution in the class specified in Subsection 1.2 of

$$\partial_s U = \mathcal{G}(D_q^2 U) = \frac{1}{4\pi} \sqrt{\det D_q^2 U}, \quad U(0, q) = |q|. \quad (44)$$

Here the viscosity formulation is the strong convex one: $U(s, \cdot)$ is convex; every upper test, and every lower test whose spatial Hessian is positive semidefinite, satisfies the corresponding evolution inequality, including when that Hessian is singular. In particular, with $\kappa = U(1, 0)$, every integer sequence k_m satisfying the conditions below obeys

$$\frac{F_{k_m}^{(m)}(0)}{\sqrt{t_{k_m}^{(m)}}} \longrightarrow \kappa \quad \text{whenever } t_{k_m}^{(m)} \rightarrow \infty, \quad h_m/t_{k_m}^{(m)} \rightarrow 0. \quad (45)$$

Proof. We first obtain the time modulus and identify the two viscosity inequalities. For an upper contact the Gaussian flux uses only the local test. For a lower contact we minimize the difference from a global spatial barrier; the resulting scalar function of time retains the contact derivative and satisfies the discrete inequality (47i). We then prove comparison. Translating a maximizing pair by its common center allows us to remove the spatial localization while retaining the matrix order supplied by the relative quadratic penalty. Uniqueness and scaling will give (45).

Step 1. Stability, compactness, and the initial trace.

The one-step bound (16a), summed directly, is too large for the time modulus we need. We instead iterate its primal bound using independent centered fluxes, whose second moments add. Let Q be the pasted linear fine flux in (16b), let X be uniform on Q_{L_m} , independent of the source, and put $Z = Q(X)$. Then $\mathbb{E}Z = 0$ and $\mathbb{E}|Z|^2 \leq Ch_m$. Equation (16b), before using the Lipschitz estimate, gives the stronger domination

$$(\mathcal{H}_{h_m, L_m} \Psi)(p) \leq \mathbb{E} \Psi(p + Z) \quad (45a)$$

for every admissible profile Ψ . Use independent copies of the fine source, the uniform point, and hence Z , at successive iterations. Repeated application of (45a) within row m gives

$$F_{k+\ell}^{(m)}(p) \leq \mathbb{E} F_k^{(m)}(p + Z_1 + \cdots + Z_\ell).$$

Since every $F_k^{(m)}$ is 1-Lipschitz,

$$0 \leq F_{k+\ell}^{(m)}(p) - F_k^{(m)}(p) \leq C\sqrt{\ell h_m}, \quad |p| \leq F_k^{(m)}(p) \leq |p| + C\sqrt{k h_m}. \quad (46)$$

Consequently, at grid times,

$$|U_m(s, q) - U_m(t, q)| \leq C\sqrt{|s - t|} + C\sqrt{h_m/r_m}, \quad (46a)$$

while the spatial Lipschitz constant is one. Since the interpolation is constant between grid times and $h_m/r_m \rightarrow 0$, its jumps vanish uniformly. Together with the spatial Lipschitz bound, (46a) yields locally uniform subsequential limits. Moreover, $0 \leq U_m(s, q) - |q| \leq C\sqrt{s} + o(1)$, which gives the initial trace uniformly on $\mathbb{R}^2$.

Step 2. Upper half-relaxed limit and the subsolution inequality.

For consistency put $\Delta_m = h_m \log 2/r_m$ and $U_m^k(q) := U_m(k\Delta_m, q)$, and define

$$(S_m\varphi)(q) = r_m^{-1/2} \mathcal{H}_{h_m, L_m}[p \mapsto \sqrt{r_m} \varphi(p/\sqrt{r_m})](\sqrt{r_m}q). \quad (46b)$$

The map is monotone and constant-equivariant. We use the standard half-relaxed limits

$$\overline{U}(s, q) = \limsup_{\substack{m \rightarrow \infty \\ (s', q') \rightarrow (s, q)}} U_m(s', q'), \quad \underline{U}(s, q) = \liminf_{\substack{m \rightarrow \infty \\ (s', q') \rightarrow (s, q)}} U_m(s', q').$$

Let a smooth convex ϕ touch the upper half-relaxed limit strictly from above at (s_0, q_0) , $s_0 > 0$. We must bound its time derivative by $\mathcal{G}(D_q^2\phi)$ using only this local contact. Select grid maxima $(s_m, q_m) = (k_m\Delta_m, q_m) \rightarrow (s_0, q_0)$ in a fixed small parabolic cylinder. If V_m is the pasted Neumann quadratic flux for $D_q^2\phi(s_0, q_0) + \gamma I$, the primal domination (33c), expressed in the rescaled variables, gives

$$U_m^{k_m}(q_m) \leq \mathbb{E} \int U_m^{k_m-1}(q_m + V_m(x)/\sqrt{r_m}) dx. \quad (46c)$$

On $|V_m| \leq r_0\sqrt{r_m}$, maximality in the cylinder bounds the right side by the same expectation with U_m replaced by ϕ , plus the contact constant. On the complementary event, the cone bounds (46), the spatial Lipschitz bound, and (32h) give $C\sqrt{h_m/r_m} \exp(-cr_m/h_m)$, which is $o(\Delta_m)$. Taylor expansion of

$$\sqrt{r_m}\{\phi(s_m - \Delta_m, q_m + z/\sqrt{r_m}) - \phi(s_m - \Delta_m, q_m)\}$$

has exactly the form (33a). The centred linear term vanishes. Lemma 6.3 therefore yields

$$\frac{\phi(s_m, q_m) - \phi(s_m - \Delta_m, q_m)}{\Delta_m} \leq \mathcal{G}(D_q^2\phi(s_0, q_0) + \gamma I) + o(1). \quad (46d)$$

First let $m \rightarrow \infty$ and then $\gamma \downarrow 0$. This proves

$$\phi_s(s_0, q_0) \leq \mathcal{G}(D^2\phi(s_0, q_0)), \quad (47)$$

including unit slopes and singular upper jets. No global upper barrier is used.

Step 3. Lower contact and a time-dependent minimum.

For a lower contact, the test itself need not lie below the preceding time slice on the whole space, so it cannot be inserted directly into the monotone scheme. We use Lemma 6.5 to fix a global spatial barrier and allow only its additive constant to vary with time. For every $S < \infty$, the stability estimates pass to the limit on $0 \leq s, t \leq S$:

$$|\underline{U}(s, q) - \underline{U}(t, q)| \leq C\sqrt{|s - t|}, \quad |q| \leq \underline{U}(s, q) \leq |q| + C_S. \quad (47a)$$

Let a smooth convex ϕ touch $\underline{U}$ strictly from below at (s_0, q_0) , with $s_0 > 0$, and put

$$\tau = \phi_s(s_0, q_0), \quad a_0 = D_q\phi(s_0, q_0), \quad H = D_q^2\phi(s_0, q_0) > 0. \quad (47b)$$

First suppose $|a_0| < 1$. Fix $0 < \delta < \lambda_{\min}(H)$ and apply Lemma 6.5 to the spatial slice $q \mapsto \underline{U}(s_0, q)$, using $q \mapsto \phi(s_0, q)$ as its lower test. Denote the resulting barrier by B_δ .

The largest constant that can be added to this barrier at time s is

$$c(s) = \min_{q \in \mathbb{R}^2} \{\underline{U}(s, q) - B_\delta(q)\}. \quad (47c)$$

The minimum is attained, uniformly for s near s_0 , because (47a) and (38b) make the expression uniformly coercive. Strict global contact gives

$$c(s_0) = 0, \quad \operatorname{argmin}_q \{\underline{U}(s_0, q) - B_\delta(q)\} = \{q_0\}. \quad (47d)$$

The common time modulus, strictness on compact annuli, and coercivity then imply that every minimizer in (47c) converges to q_0 as $s \rightarrow s_0$.

To preserve the original time derivative τ , we need a smooth lower test for c with that derivative at s_0 . More precisely, there are $K, r > 0$ such that

$$c(s) \geq \ell(s) := \tau(s - s_0) - K(s - s_0)^2 \quad \text{for } |s - s_0| \leq r, \quad c(s_0) = \ell(s_0). \quad (47e)$$

Indeed, all relevant minimizers lie in the parabolic neighborhood where $\underline{U} \geq \phi$. Put $z = q - q_0$ and $\sigma = s - s_0$. The exact local quadratic formula for B_δ and the spacetime Taylor expansion of ϕ give

$$\phi(s, q) - B_\delta(q) = \tau\sigma + \frac{\delta}{2}|z|^2 + \sigma D_{sq}^2 \phi(s_0, q_0) \cdot z + O(\sigma^2) + o(|z|^2). \quad (47f)$$

After shrinking r , absorb the last spatial remainder into $\delta|z|^2/4$. Completing the square absorbs the mixed term at a cost of order σ^2 , which proves (47e). The small Hessian loss δI in the barrier is what permits this absorption.

Step 4. Discrete contacts and the one-step inequality.

At a discrete minimum, global spatial order and monotonicity bound its time increment below by the increment of B_δ ; the lower time test bounds it above. Uniform consistency (38c) then gives the inequality for τ .

At grid times put

$$c_m^k = \min_q \{U_m^k(q) - B_\delta(q)\}, \quad (47g)$$

and choose a minimizer q_m^k . These minima are attained, and their minimizers are uniformly bounded near s_0 , since

$$U_m^k(q) - B_\delta(q) \geq (1 - r_\delta)|q| - C_\delta.$$

The lower half-relaxed limit of the scalar grid functions c_m is exactly c . For the lower bound, take minimizing points, use their compactness, and invoke the definition of $\underline{U}$. For the reverse bound, take a minimizer of (47c), use a recovery sequence for $\underline{U}$ there, and move its time coordinate to the nearest grid time using (46a).

Fix $\nu > 0$ and replace ℓ on $[s_0 - r, s_0 + r]$ by $\ell_\nu(s) = \ell(s) - \nu(s - s_0)^4$. Then $c - \ell_\nu$ has its unique minimum at s_0 , and $\ell'_\nu(s_0) = \tau$. Choose a recovery subsequence of grid points for $c(s_0)$ and relabel its row indices by m . Minimizing $c_m^k - \ell_\nu(k\Delta_m)$ over the grid points in this interval and using the preceding half-relaxed-limit equality produces interior indices k_m such that

$$t_m = k_m \Delta_m \longrightarrow s_0, \quad c_m^{k_m} - \ell_\nu(t_m) \leq c_m^j - \ell_\nu(j\Delta_m) \quad (47h)$$

for every grid point $j\Delta_m$ in the interval. If $q_m = q_m^{k_m}$, every subsequential limit of q_m , together with the recovery upper bound implicit in this selection, attains the minimum in (47d). Hence $q_m \rightarrow q_0$.

By the definition of $c_m^{k_m-1}$,

$$U_m^{k_m-1} \geq B_\delta + c_m^{k_m-1}.$$

Using $U_m^{k_m} = S_m U_m^{k_m-1}$, monotonicity and constant equivariance, and then evaluating at the current minimizer q_m , gives the one-step inequality

$$c_m^{k_m} - c_m^{k_m-1} \geq S_m B_\delta(q_m) - B_\delta(q_m). \quad (47i)$$

On the other hand, (47h), applied also at $k_m - 1$, gives

$$c_m^{k_m} - c_m^{k_m-1} \leq \ell_\nu(t_m) - \ell_\nu(t_m - \Delta_m). \quad (47j)$$

Divide by Δ_m . Since $q_m \rightarrow q_0$, the uniform consistency (38c), written in the variables (46b), makes the quotient of the right side of (47i) converge to $\mathcal{G}(H - \delta I)$; the quotient in (47j) converges to $\ell'_\nu(s_0) = \tau$. Therefore

$$\phi_s(s_0, q_0) \geq \mathcal{G}(H - \delta I).$$

Letting $\delta \downarrow 0$ proves the supersolution inequality for positive-definite interior-slope lower jets.

Step 5. Singular Hessians and unit slopes.

If a lower jet has positive-semidefinite singular spatial Hessian, the pointwise scheme monotonicity $U_m^{k+1} \geq U_m^k$ passes to the limit. Testing this temporal monotonicity from the past gives $\phi_s \geq 0 = \mathcal{G}(H)$. Finally, if $|a_0| = 1$, convexity and the 1-Lipschitz bound imply

$$\underline{U}(s_0, q_0 + ta_0) = \underline{U}(s_0, q_0) + t \quad (t \geq 0).$$

The lower-contact inequality along this ray gives $a_0^T H a_0 \leq 0$. For $H \geq 0$, this forces $H a_0 = 0$, so H is singular and the preceding argument applies. Every subsequential limit is therefore a strong convex viscosity solution of (44). To identify all subsequential limits, it remains to prove comparison in this precise class.

Step 6. Whole-space localization for comparison.

For completeness, we prove comparison in the precise class produced by (46)–(46a): the functions are continuous, convex and 1-Lipschitz in space, nondecreasing in time, have bounded cone corrections on finite time intervals, and share a spatially uniform time modulus. Let u be a subsolution and v a strong convex supersolution in this class, with $u(0, \cdot) \leq v(0, \cdot)$, and suppose that $u - v > 0$ somewhere before a time S .

We need an interior maximum to which the doubled-variable lemma applies. The terminal penalty below creates a positive gap between the time derivatives, the relative quadratic controls $x - y$, and the last term makes the common spatial position bounded. Choose the parameters in the order γ , then ϵ , then η . Take $\gamma > 0$ small enough that the violation survives the terminal penalty, and then $\epsilon > 0$ small enough for (48a) to exclude the initial time. Finally choose $\eta > 0$ for spatial coercivity; it will tend to zero with ϵ fixed. Maximize

$$\begin{aligned} \mathcal{M}_{\epsilon, \eta}(t, x, y) = & u(t, x) - v(t, y) - \frac{|x - y|^2}{2\epsilon} - \frac{\gamma}{S - t} \\ & - \eta\{\chi(x) + \chi(y)\}, \quad \chi(x) = \sqrt{1 + |x|^2}, \end{aligned} \quad (48)$$

over $0 \leq t < S$. Since

$$u(t, x) - v(t, y) \leq |x - y| + 2C_S,$$

the relative quadratic controls $x - y$, and the last term is coercive in the common-translation direction. Thus the positive maximum is attained. At $t = 0$, initial ordering and the spatial Lipschitz bound of v give

$$u(0, x) - v(0, y) - \frac{|x - y|^2}{2\epsilon} \leq |x - y| - \frac{|x - y|^2}{2\epsilon} \leq \frac{\epsilon}{2}. \quad (48a)$$

Taking $\epsilon < 2\gamma/S$, the common time modulus keeps positive maximizers a fixed distance from $t = 0$. The term $-\gamma/(S - t)$ keeps them a fixed distance from S .

Before comparing Hessians, we remove the coercive term. Although the common center may then escape, translation by this center leaves the relative quadratic unchanged. Fix ϵ ,

let $\eta_n \downarrow 0$, and let (t_n, x_n, y_n) be maximizers. The preceding cone estimate and positivity of the maximum show that $d_n = x_n - y_n$ is bounded. Pass to a subsequence with

$$t_n \rightarrow t_* \in (0, S), \quad d_n \rightarrow d.$$

The common centres $c_n = (x_n + y_n)/2$ may escape. Translate both functions by this same centre, and normalize only their vertical constants:

$$\widehat{u}_n(t, z) = u(t, z + c_n) - u(t_n, x_n), \quad \widehat{v}_n(t, z) = v(t, z + c_n) - v(t_n, y_n). \quad (48b)$$

The spatial Lipschitz bound, common time modulus, and boundedness of d_n make these families locally uniformly bounded and equicontinuous. After a subsequence they converge locally uniformly to $\widehat{u}, \widehat{v}$.

Spatial translations and addition of constants preserve the equation. The usual nearby-contact argument shows that $\widehat{u}$ is a subsolution and $\widehat{v}$ is a strong convex supersolution. Here are the two boundary cases of that stability argument. At a singular upper test of $\widehat{u}$, add $\sigma|z - z_0|^2$, pass to nearby contacts, and then let $\sigma \downarrow 0$. At a positive-definite lower test of $\widehat{v}$, nearby Hessians remain positive definite. At a singular positive-semidefinite lower test, temporal monotonicity, which also passes to the limit, directly gives $b \geq 0 = \mathcal{G}(Y)$. Thus the strong formulation is closed under this local limit.

Put $\xi_n = d_n/2$ and $\zeta_n = -d_n/2$. Maximality in (48), after the normalizations (48b), says, for every fixed t, z, w ,

$$\begin{aligned} \widehat{u}_n(t, z) - \widehat{v}_n(t, w) - \frac{|z - w|^2}{2\epsilon} - \frac{\gamma}{S - t} \\ \leq -\frac{|\xi_n - \zeta_n|^2}{2\epsilon} - \frac{\gamma}{S - t_n} + \eta_n \{ \chi(c_n + z) + \chi(c_n + w) - \chi(x_n) - \chi(y_n) \}. \end{aligned} \quad (48c)$$

Because χ is 1-Lipschitz, the absolute value of the last line, on compact z, w -sets, is at most

$$\eta_n \{ |z - \xi_n| + |w - \zeta_n| \} \rightarrow 0. \quad (48d)$$

Passing to the limit, while retaining the relative quadratic, shows that $(t_*, d/2, -d/2)$ is an attained global maximum of

$$\widehat{u}(t, z) - \widehat{v}(t, w) - \frac{|z - w|^2}{2\epsilon} - \frac{\gamma}{S - t}. \quad (48e)$$

Step 7. Matrix order and comparison.

At the maximum in (48e), the relative quadratic is now the only spatial penalty. Its Hessian vanishes when both points move by the same vector. We use these directions in the doubled-variable matrix inequality to obtain $X \leq Y$ with no localization error.

We use the doubled-variable lemma [CIL92] in the following form. Apply quadratic sup-convolution to $\widehat{u}$ and quadratic inf-convolution to $\widehat{v}$. The ordinary second-derivative test at the regularized maximum, followed by passage to closed semijets, gives

$$(a, p, X) \in \overline{\mathcal{P}}^{2,+} \widehat{u}, \quad (b, p, Y) \in \overline{\mathcal{P}}^{2,-} \widehat{v}, \quad a - b = \frac{\gamma}{(S - t_*)^2} > 0, \quad (48f)$$

and, for every $\lambda > 0$,

$$\begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \leq A_\epsilon + \lambda A_\epsilon^2, \quad A_\epsilon = \frac{1}{\epsilon} \begin{pmatrix} I & -I \\ -I & I \end{pmatrix}. \quad (48g)$$

The lower block bound in the same regularized second-derivative test keeps the approximate Hessians bounded, so (48g) survives in the closed semijets. Both matrices on its

right annihilate every common direction (z, z) . Testing (48g) on such a direction gives the exact, error-free inequality

$$X \leq Y. \quad (48h)$$

Convexity of $\widehat{u}(t_*, \cdot)$ gives $X \geq 0$: apply each approximating upper quadratic contact in opposite spatial directions and use nonnegativity of the convex second difference. Hence $Y \geq 0$. To apply the supersolution inequality to the closed lower jet, first suppose $Y > 0$. The approximating Hessians are then eventually positive definite, so the inequality passes to the limit by continuity. If Y is singular, temporal monotonicity gives $b \geq 0 = \mathcal{G}(Y)$: at an ordinary lower contact, fixing the spatial point and approaching from earlier times gives a nonnegative time derivative, and this property passes to closed jets. Since determinant is nondecreasing in the Loewner order on $\mathbb{S}_+^2$,

$$a \leq \mathcal{G}(X) \leq \mathcal{G}(Y) \leq b,$$

contradicting the strictly positive time-derivative gap in (48f). Comparison follows. Applied to the upper and lower half-relaxed limits, which have the same initial trace, it shows that they agree; the compactness estimate then gives local uniform convergence of the whole sequence.

Step 8. Scaling and the value at the origin.

To identify the normalization in (45), observe that the equation and its initial datum are invariant under

$$U(s, q) \mapsto a^{-1}U(a^2s, aq).$$

Uniqueness therefore gives $U(s, q) = \sqrt{s}U(1, q/\sqrt{s})$. To obtain (45) for the prescribed indices k_m , apply the convergence just proved with $r_m = t_{k_m}^{(m)}$. The two assumptions in (45) are precisely the conditions required for this choice of r_m . Formula (43) at $(s, q) = (1, 0)$ then gives the claimed limit $U(1, 0) = \kappa$.

■

8. Completion of the Gaussian and empirical limits

We choose one sequence of scales for both the two-scale comparison and the viscosity limit. The omitted levels, accumulated comparison error, and rescaled time step must all be negligible.

LEMMA 8.1 (The dyadic Gaussian limit). *The dyadic Gaussian limit exists, and in the notation of Lemma 7.1 it is*

$$\lim_{J \rightarrow \infty} \frac{\Gamma_{2^J}}{2^J \sqrt{J} \log 2} = \kappa. \quad (49)$$

Proof. It suffices to approximate the Gaussian cost per unit area by the cell iteration with error $o(\sqrt{J})$. Suppose that we retain m groups of h levels, omit D levels, and incur an error at most e_J per unit area at each replacement. Put $t_J = mh \log 2$. The three requirements are

$$\frac{D}{J} \longrightarrow 0, \quad e_J = o\left(\frac{h}{\sqrt{J}}\right) \implies me_J = o(\sqrt{J}), \quad \frac{h}{t_J} \longrightarrow 0.$$

The first allows us to restore omitted levels using (12), the second controls the sum of the comparison errors, and the third permits the viscosity limit. We choose all parameters in the same sequence and then carry out these reductions.

Step 1. Choose and align the retained bands.

Fix a large J , set

$$h = \lfloor J^{3/4} \rfloor, \quad b = \lceil (\log J)^2 \rceil, \quad S = 2^{b+1},$$

and retain consecutive groups of h Haar levels separated by gaps of b omitted levels. Omit also fewer than $h + b$ levels at the fine end and $h + b$ levels at the coarse end. The total number D of omitted levels satisfies

$$D \leq C \left(h + \frac{Jb}{h} \right) = o(J). \quad (50)$$

The coarse-end omission ensures that, after normalizing any retained band $\eta_{r,h}$ to unit outer blocks, the ambient square has side

$$R_r = 2^{J-r-h} \geq 2^{h+b}. \quad (50a)$$

When it exists, the next retained detail starts at level $r + h + b$, whose parent side is $2^{r+h+b+1}$. Hence its smallest parent side in the coordinates obtained by dividing lengths by 2^{r+h} is exactly

$$\frac{2^{r+h+b+1}}{2^{r+h}} = 2^{b+1} = S.$$

Since $h \geq 1$, (50a) also shows that S divides R_r . Thus the ambient square is divisible by every auxiliary dyadic length used below. Orthogonality of the Haar projections makes the active band independent of all bands still present.

Step 2. Make the comparison errors summable.

In Lemma 5.1 choose

$$L = S^{1/8}, \quad B = S^{3/8}, \quad A = S^{5/8}, \quad \varepsilon = S^{-1/32}, \quad (51)$$

rounding to compatible dyadic integers. The parameters are chosen in the order

$$J; \quad h, b, S; \quad L, B, A, \varepsilon.$$

The omitted-level estimate is already (50). To bound e_J uniformly throughout the iteration, we check both the geometry and the sizes of the profiles in (24). The second condition in (23a) follows from

$$L + \frac{B^2}{A} \asymp S^{1/8} \ll S^{11/32} = \varepsilon B.$$

The geometric loss satisfies $\delta \leq CS^{-1/32}$. The concentration error is $O(\sqrt{\log S} S^{-1/16})$. Since $S = 2^{b+1}$ grows faster than any power of J , it remains to bound the other factors polynomially in J , uniformly in the step. If $\Phi_0 = |\cdot|$ and $\Phi_{k+1} = \mathcal{H}_{h,L} \Phi_k$, the construction behind (46) gives

$$\Phi_k(p) \leq \mathbb{E}|p + Z_1 + \cdots + Z_k| \leq |p| + C\sqrt{kh}; \quad (52)$$

hence the height parameter M in (13) is $O(\sqrt{J})$. For the normalized retained band at level r , use the zero-normal flux P from (22). At most J coarser levels remain, so (22) gives $\sup_x \mathbb{E}|P(x)| \leq C\sqrt{J}$. This competitor gives, uniformly in the step,

$$|Q_{R_r}|^{-1} \mathbb{E} J_{\mathcal{H}_{h,L} \Phi_k}^{Q_{R_r}}(\zeta) \leq \mathbb{E} \int_{Q_{R_r}} (|P| + C\sqrt{J}) \leq C\sqrt{J}. \quad (52a)$$

Inserting these bounds into (24) gives the required uniform estimate

$$e_J \leq C\sqrt{J} S^{-1/32} + C\sqrt{\log S} S^{-1/16} = o(h/\sqrt{J}). \quad (52b)$$

Since $m \leq J/h$, the total comparison error satisfies

$$me_J \leq \frac{J}{h} o\left(\frac{h}{\sqrt{J}}\right) = o(\sqrt{J}).$$

Together with $\varepsilon L = S^{3/32} \rightarrow \infty$, the displayed inequality verifies both conditions in (23a). All iterated profiles are D_4 -invariant by Proposition 4.2, so the reflected bounded-square proof applies.

Step 3. Remove the retained bands and evaluate the terminal cost.

Apply (24), successively, from the finest retained band to the coarsest. There are $m \leq J/h$ applications, so their total error per unit area is at most $me_J = o(\sqrt{J})$. In dimension two the rescaling from a band of outer side 2^{r+h} changes both total energy and area by the same factor $2^{2(r+h)}$, so (24) is an error per original unit area with no additional scale factor. Formula (50a) provides exact boundary alignment at every application. For the last band the same argument is used with $\zeta = 0$. After the last replacement no random source remains. For a divergence-free zero-normal field v on a square,

$$\int v_i = \int \operatorname{div}(x_i v) = 0;$$

therefore Jensen's inequality, with equality for $v = 0$, says that the terminal energy is exactly $|Q_{2^J}| \Phi_m(0) = 2^{2J} \Phi_m(0)$. Consequently, if $\rho_{2^J}^{\text{ret}}$ is the retained source in Q_{2^J} ,

$$2^{-2J} \mathbb{E} \|\rho_{2^J}^{\text{ret}}\|_{\text{KR}} = \Phi_m(0) + o(\sqrt{J}). \quad (53)$$

Step 4. Apply the viscosity limit and restore the omitted bands.

The retained logarithmic time is

$$t_J = mh \log 2 = (J - D) \log 2 = J \log 2 + o(J). \quad (54)$$

In particular,

$$t_J \sim J \log 2, \quad \frac{h}{t_J} \longrightarrow 0, \quad L = S^{1/8} \longrightarrow \infty.$$

These relations give the joint limit regime required by Lemma 7.1.

Apply Lemma 7.1 with triangular row index J , $h_J = h$, $L_J = L$, $r_J = t_J$, $F_k^{(J)} = \Phi_k$, and iteration index $k_J = m$. Then $t_{k_J}^{(J)} = t_J$, and (54) verifies the hypotheses of (45). Hence $\Phi_m(0)/\sqrt{t_J} \rightarrow \kappa$. On the other hand, (12) and (50) give

$$\mathbb{E} \|\rho_{2^J} - \rho_{2^J}^{\text{ret}}\|_{\text{KR}} \leq C 2^{2J} \sqrt{D} = o(2^{2J} \sqrt{J}). \quad (55)$$

Combining (53)–(55),

$$\frac{\mathbb{E} \|\rho_{2^J}\|_{\text{KR}, Q_{2^J}}}{2^{2J} \sqrt{J \log 2}} \longrightarrow \kappa. \quad (56)$$

Finally (9) is $o(2^{2J} \sqrt{J})$, and $\widehat{\Gamma}_{2^J} = 2^J \Gamma_{2^J}$. Equation (56) is therefore equivalent to (49), which proves the lemma.

■

Proof of Theorem 1.1. The dyadic limit in (6) equals κ by Lemma 8.1. The approximation in Lemma 2.4 therefore gives, for the full sequence of integers N ,

$$\lim_{N \rightarrow \infty} \frac{1}{\sqrt{N \log N}} \mathbb{E} \sup_{\text{Lip}(f) \leq 1} \left| \sum_{i=1}^N \left(f(X_i) - \int f d\lambda \right) \right| = \frac{\kappa}{\sqrt{2}}.$$

By Lemma 7.1, $\kappa = U(1, 0)$. The AKT bounds recalled in the introduction show that this limit is finite and positive, which proves every assertion of the theorem.

■

References

- [AST19] Luigi Ambrosio, Federico Stra, and Dario Trevisan. A PDE approach to a 2-dimensional matching problem. *Probability Theory and Related Fields*, 173:433–477, 2019. <https://doi.org/10.1007/s00440-018-0837-x>. 2
- [CIL92] Michael G. Crandall, Hitoshi Ishii, and Pierre-Louis Lions. User’s guide to viscosity solutions of second order partial differential equations. *Bulletin of the American Mathematical Society*, 27:1–67, 1992. <https://doi.org/10.1090/S0273-0979-1992-00266-5>. 42
- [DRS24] Lars Diening, Johannes Rolfes, and Abner J. Salgado. Pointwise gradient estimate of the Ritz projection. *SIAM Journal on Numerical Analysis*, 62:1212–1225, 2024. <https://doi.org/10.1137/23M1571800>. 31
- [EG99] Lawrence C. Evans and Wilfrid Gangbo. Differential equations methods for the Monge–Kantorovich mass transfer problem. *Memoirs of the American Mathematical Society*, 137(653), 1999. <https://doi.org/10.1090/memo/0653>. 2
- [Fan19] Xiao Fang. Wasserstein-2 bounds in normal approximation under local dependence. *Electronic Journal of Probability*, 24(35):1–14, 2019. <https://doi.org/10.1214/19-EJP301>. 8
- [Tal92] Michel Talagrand. The Ajtai–Komlós–Tusnády matching theorem for general measures. In *Probability in Banach Spaces*, 8, volume 30 of *Progress in Probability*, pages 39–54. Birkhäuser, 1992. https://doi.org/10.1007/978-1-4612-0367-4_2. 2
- [WY10] Feng-Yu Wang and Lixin Yan. Gradient estimate on the Neumann semigroup and applications. arXiv:1009.1965v2, 2010. <https://arxiv.org/abs/1009.1965>. 31