QUARTIC EXPONENTIAL MATCHING IN THE UNIT SQUARE

WILL HOWE, GRANT HUFFMAN, JENSEN KOHLMEYER, AND LIAM KRUER

Abstract. Take two independent samples of n uniform points in the unit square. We prove that, with probability at least one half, the points can be paired so that the average of the exponential of the fourth power of the normalized pair distance is at most two. The distance normalization is $L\sqrt{\log n/n}$, with L an absolute constant. One pairing therefore satisfies both coordinate inequalities in the symmetric $(4,4)$ case of Talagrand's matching problem. The argument first controls a vector field transporting smoothed point densities, and then converts that control to a pairing. The main estimate comes from a Gaussian construction whose successive partitions depend only on noise already revealed; a correction using the classical bottleneck matching preserves the required density balance.

hexagon:2611.00270v1

1. Statement and notation

Imagine n red points and n blue points, sampled independently and uniformly in the unit square. A matching pairs every red point with one blue point, using each blue point exactly once. We seek a matching in which long pairs are rare, with a bound that controls the horizontal and vertical errors at the same time.

Put $Q=[0,1]^2$, and write $|\cdot|$ for the Euclidean norm. All logarithms are natural. A permutation π specifies the pairing of the red point X_i with the blue point $Y_{\pi(i)}$. Talagrand's symmetric $(4,4)$ question [4] asks for one permutation satisfying the two inequalities (1.2), one for each coordinate. We prove the Euclidean-distance bound (1.1); each coordinate difference is at most that distance, so it gives the requested simultaneous control.

Theorem 1.1. *There is an absolute constant $L>0$ such that the following holds for every integer $n\ge 2$. Let $X_1,\dots,X_n,Y_1,\dots,Y_n$ be independent, uniformly distributed points of Q. With probability at least $1/2$ there is a permutation π of $\{1,\dots,n\}$ such that*

$$\frac{1}{n}\sum_{i=1}^{n}\exp\left(\frac{|X_i-Y_{\pi(i)}|^4}{L^4a_n^4}\right)\le 2,\qquad a_n=\sqrt{\frac{\log n}{n}}. \tag{1.1}$$

Consequently the same permutation satisfies, for $j=1,2$,

$$\sum_{i=1}^{n}\exp\left(\left(\sqrt{\frac{n}{\log n}}\frac{|X_i^j-Y_{\pi(i)}^j|}{L}\right)^4\right)\le 2n. \tag{1.2}$$

What the estimate says. Write $d_i=|X_i-Y_{\pi(i)}|$ for the lengths of the matched pairs. Inequality (1.1) immediately gives, for every $t\ge 0$,

$$\frac{1}{n}\#\{i: d_i>La_nt\}\le 2e^{-t^4}.$$

It also gives both

$$\frac{1}{n}\sum_i d_i\le La_n(\log 2)^{1/4},\qquad \max_i d_i\le La_n(\log(2n))^{1/4}.$$

Date: November 9, 2026.

The mathematical draft and exposition were developed and written with OpenAI Codex under the direction of the named contributors. Substantial portions of the argument and text are AI-generated.

The first follows from Jensen's inequality, and the second by bounding each summand in (1.1) by $2n$. Thus the same matching has small average distance, a controlled largest distance, and an explicit bound on the fraction of pairs at every intervening scale.

The Leighton–Shor theorem supplies the largest-distance bound at this order, while the AKT theorem supplies the average-distance bound; see [4] and [3, Theorem 4.7.1]. The additional task here is to obtain the fourth-power exponential estimate for one matching. We treat only the symmetric $(4, 4)$ case.

How the proof is organized. The argument has four stages. The purpose of the continuous and Gaussian constructions is to build a pairing of the original finite point sets.

(1) *Replace the points by densities and reduce matching to a flow estimate.* Spread each point over a neighborhood of radius $h \asymp a_n$. With high probability the two resulting densities $\rho_{X,h}$ and $\rho_{Y,h}$ are bounded above and below. A vector field J with $\operatorname{div} J = \rho_{X,h} - \rho_{Y,h}$ and zero normal component at the boundary describes how to move the excess of the first density to fill the deficit of the second. Lemma 2.2 shows that a fourth-power exponential bound on J/a_n gives the desired bound on the pair lengths. Its proof follows the flow, obtains a fractional matching, and chooses one permutation from a convex decomposition of that matching.

(2) *Control a Gaussian model of the density imbalance.* The main construction reveals noise one scale at a time. Before the next independent noise is revealed, it chooses the shapes of the next cells using the field already built. This is the meaning of "predictable" in Section 4. The area-preserving matrix in (4.2) contracts in the direction of the existing vector and expands in the perpendicular direction. Its identity $|A(z)z| \le 1$ controls fluctuations of the squared length even when the vector is large. Section 5 turns this control into an e^{-ct^4} tail up to a cutoff $R = (\log(2/h))^{1/4}$. The exact Gaussian law of the resulting source is checked in Lemma 4.1.

(3) *Pass from Gaussian noise to the actual samples.* The coupling in Section 7 compares the smoothed empirical imbalance with the Gaussian source. Their difference is the divergence of a field whose supremum is $O(a_n)$, which can be absorbed at the desired scale. The result of the Gaussian construction is therefore useful for the sampled points themselves.

(4) *Repair large field values without changing the imbalance.* The stopped estimate does not yet give a full exponential moment at all field values. Section 8 uses the classical bottleneck matching to produce a second field with the same divergence and a uniform bound. The difference of the two fields is divergence-free, so it is represented by a stream function. Truncating that function repairs the exceptional large values while retaining the required divergence. Lemma 8.2 gives the full exponential bound, and Lemma 2.2 then returns a matching of the original points.

The last section assembles these statements and accounts for their exceptional events. We prove the probability threshold $1/2$ in the question; we do not use a stronger probability assertion as an input.

Constants denoted by c, C are positive and absolute, and may change from line to line. When a fixed failure probability is prescribed, they may also depend on that probability. An isonormal process on a real Hilbert space H is a centered Gaussian linear map $W : H \to L^2(\Omega)$ with $\mathbb{E}W(f)W(g) = \langle f, g\rangle$. White noise on Q means the isonormal process on $L^2(Q)$. Its mean-zero part is $f \mapsto W(f - \int_Q f)$. Boundary flux identities are interpreted distributionally after extension by zero; thus no boundary measure is present.

2. Compact smoothing and transport

Let $T = (\mathbb{R}/2\mathbb{Z})^2$ with its ordinary area measure. A scalar field on Q is extended evenly in both coordinates to T; a vector field is extended so that its normal component is odd and its tangential component even at each side. Choose a nonnegative smooth function η of integral one, even in

each coordinate, supported in $[-2,2]^2$, and bounded below by a positive constant on $[-1,1]^2$. For $0<h<1/4$ let η_h be its periodization at scale h. Define

$$k_h(x,z)=\sum_{g\in\{\pm1\}^2}\eta_h(x-gz),\qquad x,z\in Q. \tag{2.1}$$

Each of the integrals of this kernel in one variable is one. Moreover $k_h(x,z)>0$ implies $|x-z|\le 2\sqrt{2}h$. Indeed the circle distance between x_j and $\pm z_j$ is at least $|x_j-z_j|$ for $x_j,z_j\in[0,1]$.

For a sample $U=(U_i)_{i\le n}$ put $\rho_{U,h}(x)=n^{-1}\sum_i k_h(x,U_i)$. Choose a sufficiently large fixed A, and set

$$m=\lfloor (Aa_n)^{-1}\rfloor,\qquad h=m^{-1}. \tag{2.2}$$

We first take n large enough that $m\ge 4$. Then $Aa_n\le h\le 2Aa_n$. The finitely many other values of n will be handled at the end.

Lemma 2.1. *With probability $1-o(1)$ both $\rho_{X,h}$ and $\rho_{Y,h}$ lie between fixed positive constants c_0,C_0 on Q. On the same event, smoothing either sample at any radius $s\ge h$, with $s<1/4$, gives a density at most an absolute constant.*

Proof. Partition Q into the m^2 squares of side h. Each count has binomial law with mean nh^2. The exponential Markov inequality for a binomial variable gives

$$\mathbb{P}\{N_C\notin[nh^2/2,2nh^2]\}\le 2e^{-cnh^2}.$$

A union bound over both samples is at most $4m^2e^{-cnh^2}=o(1)$ when A is sufficiently large. A point's own cell gives the lower density bound using the lower bound on η, and a bounded number of neighboring cells give the upper bound. At radius s, the support meets $O((s/h)^2)$ cells and the kernel is $O(s^{-2})$, proving the last assertion, including reflected cells. □

Lemma 2.2 (Flux-to-matching reduction)**.** *Suppose $c_0\le\rho_{X,h},\rho_{Y,h}\le C_0$ and J is a smooth field on Q with zero normal component and* $\operatorname{div}J=\rho_{X,h}-\rho_{Y,h}$. *If*

$$\int_Q\exp((|J|/b)^4)\le D, \tag{2.3}$$

then a permutation satisfies (1.1) *with La_n replaced by $C\max(h,b)$, where C depends only on c_0,C_0,D.*

Proof. Set $\rho_t=(1-t)\rho_{X,h}+t\rho_{Y,h}$ and $v_t=J/\rho_t$. The smooth vector field v_t is tangent to the boundary and its flow T_t stays in Q. The identity $\partial_t\rho_t+\operatorname{div}(\rho_t v_t)=0$, together with the Jacobian formula for the flow, gives $(T_t)_\#(\rho_{X,h}\,dx)=\rho_t\,dx$. For any convex nondecreasing Ψ with $\Psi(0)=0$, Jensen's inequality along trajectories gives

$$\begin{aligned}\int_Q\Psi(c_0|T_1(z)-z|/b)\rho_{X,h}(z)\,dz&\le\int_0^1\int_Q\Psi(c_0|v_t(x)|/b)\rho_t(x)\,dx\,dt\\&\le C_0\int_Q\Psi(|J(x)|/b)\,dx.\end{aligned}$$

Choose a uniform label i, move X_i to z with density $k_h(z,X_i)$, then to $w=T_1(z)$, and finally choose label k with probability $k_h(w,Y_k)/(n\rho_{Y,h}(w))$. The last label is uniform because the law of w is $\rho_{Y,h}\,dw$. Both smoothing moves have length at most $2\sqrt{2}h$. Applying convexity to these three displacements with $\Psi(t)=e^{t^4}-1$ bounds the mean label-pair cost at scale $C\max(h,b)$ by a constant depending only on the displayed parameters.

The matrix of label-pair probabilities, multiplied by n, is doubly stochastic. For completeness, it is a convex combination of permutation matrices: the support of a doubly stochastic matrix satisfies Hall's condition, since rows in a set of size r have total mass r and each neighboring column has mass at most one. Hall's matching criterion follows by the usual augmenting-path argument: if a

maximum matching leaves a row unmatched, the rows reachable from it by alternating paths have fewer neighbors than rows. Thus a support permutation exists. Subtract the smallest entry on this permutation, normalize the remaining matrix, and repeat; at each nonterminal step an additional positive entry vanishes. This proves the decomposition in finitely many steps.

Some component has no greater total cost than the fractional plan. Finally, if $n^{-1}\sum_i e^{t_i} \le D_1$, then $n^{-1}\sum_i e^{t_i/q} \le D_1^{1/q}$ by concavity, for $q \ge 1$. Taking q large enough makes the bound two and completes the proof. □

3. A local Gaussian flux on affine Whitney cells

Let $Q_0 = (-1/2, 1/2)^2$, and put $d(u) = \operatorname{dist}(u, \partial Q_0)$. Fix a sufficiently small absolute $\varepsilon > 0$. If D is a real matrix with determinant one and $q = \|D^{-1}\|$, take the maximal dyadic squares S in DQ_0 satisfying, on their closures,

$$\ell(S) \le \varepsilon q^{-1} \inf_{w \in S} d(D^{-1}w). \tag{3.1}$$

Here $q \ge 1$. Half-open representatives partition DQ_0 up to a null set, since every interior point off dyadic boundaries has arbitrarily small eligible squares. The function $d(D^{-1}w)$, extended by zero, is q-Lipschitz. If a square's parent is ineligible or meets the complement, this Lipschitz bound and maximality imply

$$c_\varepsilon d(u)/q \le \ell(S) \le \varepsilon d(u)/q, \qquad u \in C = D^{-1}S. \tag{3.2}$$

The children C are parallelograms. One of $D^{-1}e_1, D^{-1}e_2$ has length at least $q/\sqrt{2}$; choose its unit direction e. All children have an edge parallel to e, of length $l_C \asymp_\varepsilon d(u)$ on C. Their diameters, and the norms of their affine linear parts as images of Q_0, are also comparable to $d(u)$. By reducing ε, all these linear parts have norm at most a fixed $\rho < 1$. No lower bound on the short edge is used.

Let P be orthogonal projection in $L^2(Q_0)$ onto child-constant functions, and P_0 projection onto constants. Choose a nonnegative smooth bump b of integral one supported in a fixed central ball of Q_0. The Bogovskiĭ kernel, in dimension two, is

$$K(u,v) = (u-v)\int_1^\infty b(v + t(u-v))t\,dt. \tag{3.3}$$

This is the classical integral construction; see [1]. The estimates needed here are proved below.

Lemma 3.1. *There is an absolute constant C, independent of D, such that*

$$\|(P - P_0)K(u,\cdot)\|_2 \le C, \tag{3.4}$$

$$\|(P - P_0)(K(u,\cdot) - K(u',\cdot))\|_2 \le C \min\{1, (|u-u'|/d_C)^{1/3}\} \tag{3.5}$$

for u, u' in the closure of the same child C, where $d_C = d(c_C)$ and c_C is its center. Projections of the integrable kernel mean its cell averages. For isonormal W define the two components of $H(u)$ by applying W to the two rows of $(P - P_0)K(u,\cdot)$. Then

$$\operatorname{div} H = W\big((P - P_0)\,\cdot\big)$$

with zero normal boundary condition. On each child a continuous modification satisfies, with $\gamma = 1/6$,

$$\mathbb{P}\left\{ sup_C|H| + d_C^\gamma [H]_{C^\gamma(C)} > t \right\} \le Ce^{-ct^2}, \qquad t \ge 1. \tag{3.6}$$

All statements remain valid conditionally if D is determined before the independent process W is sampled.

Proof. Directly from (3.3),

$$|K(u,v)| \le C|u-v|^{-1}, \quad |\partial_u K(u,v)| \le C|u-v|^{-2}, \quad K(u,v) = 0 \quad \text{if } |u-v| > Cd(u). \tag{3.7}$$

For the first two bounds the integration is restricted to $t \le C/|u-v|$; differentiating b gives one additional factor t. For the third, if $z = v + t(u-v)$ is in the central ball, then $u = (1-t^{-1})v + t^{-1}z$.

Signed distances to the four sides are affine on this segment, so $d(u) \ge c/t$ while $|u-v| = |z-v|/t \le C/t$.

We claim, for $0 < r \le cd(u)$,

$$\left\| P(|K(u,\cdot)|\mathbf{1}_{|u-\cdot|\le r}) \right\|_2^2 \le Cr/d(u). \tag{3.8}$$

For a nonnegative function k, use cross-sections of a child along e. Every such cross-section has length l_C. If I_C is the projection onto $e^\perp$, Cauchy–Schwarz gives

$$\frac{(\int_C k)^2}{|C|} \le \frac{1}{l_C} \int_{I_C} \left(\int_{C\cap\{v_\perp = s\}} k \, d\ell \right)^2 ds.$$

For a child meeting the ball in (3.8), $l_C \ge c_\varepsilon d(u)$. At a fixed transverse coordinate its contributing sections are disjoint intervals of at least this length, each meeting an interval of length at most $2r$. There are consequently only $O_\varepsilon(1)$ of them. The first bound in (3.7) bounds each line integral by $C\log(2 + r/|s-u_\perp|)$, and it vanishes when $|s-u_\perp| > r$. The integral of the square of this logarithm is $O(r)$. Sum over the children to prove (3.8). This also controls the signed vector components.

Outside the ball of radius $cd(u)$, (3.7) gives a bounded L^2 norm: the possible radii range between two constant multiples of $d(u)$. Projection contracts L^2. The parent average is the constant projection of the child average. This proves (3.4), and the same argument bounds the projection of the absolute kernel.

For (3.5) let $\delta = |u-u'| \ll d_C$ and take $r = d_C(\delta/d_C)^{2/3}$. Near the two points, (3.8) gives an L^2 bound $C(r/d_C)^{1/2}$. Outside these balls the derivative bound in (3.7), integrated in L^2 along the segment, gives $C\delta/r$. These bounds agree at $C(\delta/d_C)^{1/3}$. The remaining values of δ are covered by (3.4).

If $g \in L^2(Q_0)$ has mean zero, $Bg(u) = \int K(u,v)g(v)\,dv$ obeys $\operatorname{div} Bg = g$, including after extension by zero. Indeed pairing with $\nabla\varphi$ and using $z = v + t(u-v)$ gives

$$\int Bg \cdot \nabla\varphi = \int g(v) \int b(z) \int_0^1 (z-v)\cdot\nabla\varphi(v + s(z-v))\, ds\, dz\, dv = -\int g\varphi.$$

The term $\varphi(z)$ vanishes by the mean-zero condition. Contributing segments stay in Q_0, and the size bound makes the pairing absolutely integrable. Apply this identity to $(P-P_0)g$. The row bounds and the corresponding absolute-kernel bound justify interchanging W and the u-integral, or equivalently first using finite sums of cell averages and then passing in L^2. This proves the asserted divergence identity.

Finally cover a child by nets of mesh $2^{-j}d_C$ with $O(2^{2j})$ points, join points to a nearest point in the preceding net, and include nearby same-level pairs. By (3.5), the probability that any level-j link exceeds $t2^{-j\gamma}$ is bounded by

$$C2^{4j}\exp\{-ct^2 2^{2j(1/3-\gamma)}\}.$$

The sum over j is Ce^{-ct^2} for $t \ge 1$. Add a base value using (3.4); summing links gives the modification and the Hölder seminorm bound. Nets can be chosen from a fixed countable dense set in square coordinates, so they are measurable in D. Countably many children can be treated simultaneously. Values from different children need not be identified on their boundaries. □

4. Predictable partitions and exact white noise

Fix $0 < h < 1/10$, and put

$$K_* = \log(2/h), \qquad R = K_*^{1/4}. \tag{4.1}$$

For $z \in \mathbb{R}^2$ write $s_z = 1 + |z|^2$ and define

$$A(z) = \sqrt{s_z}\, I - \frac{z z^{\mathsf{T}}}{\sqrt{s_z}}. \tag{4.2}$$

Its eigenvalues in the radial and tangential directions are respectively $s_z^{-1/2}$ and $s_z^{1/2}$. In particular

$$\det A(z) = 1, \qquad |A(z)z| \le 1, \qquad \|A(z)\| = \sqrt{1+|z|^2}. \tag{4.3}$$

Let p_R denote Euclidean projection onto the radius-R ball. Every metric state below is capped by using $A(p_R z)$.

Start with Q as one cell, old matrix I, and field $Z_0 = 0$. A current cell has the form $P = c + sA_{\mathrm{old}}Q_0$. If $\operatorname{diam} P \le h$, leave it unchanged thereafter. Otherwise evaluate the previous field at c, set $A_{\mathrm{new}} = A(p_R Z_{i-1}(c))$, and use (3.1) with $D = A_{\mathrm{new}}^{-1}A_{\mathrm{old}}$ in its old coordinates. Each child is $C = c' + s\ell(S)A_{\mathrm{new}}Q_0$. If $B_P = sA_{\mathrm{old}}$ and $B_C = s\ell(S)A_{\mathrm{new}}$ are the affine linear parts, then

$$\|B_P^{-1}B_C\| \le \rho < 1. \tag{4.4}$$

Thus every active branch has length at most $N = \lceil CK_* \rceil$ for a fixed C; choose it so that $\sqrt{2}\rho^N \le h$. In particular every final cell has diameter at most h. This remains true with arbitrarily small Whitney children.

Let P_i be projection onto constants on the partition after step i, and P_0 onto global constants. Use independent isonormal processes W_i on $L^2(Q)$, $1 \le i \le N+1$. The step-i partition is chosen before W_i is used. Expose the source

$$\xi_i(g) = W_i((P_i - P_{i-1})g).$$

On each active parent use the field of Lemma 3.1. The map $g \mapsto sg(c + sA_{\mathrm{old}}\,\cdot)$ is the relevant L^2 isometry, so the physical source is s^{-1} times the normalized source. Push the normalized field H forward as $A_{\mathrm{old}}H$; its physical divergence is $s^{-1}\operatorname{div} H$. Add this physical flux, divided by $\sqrt{K_*}$, to Z_{i-1} to obtain Z_i.

The fields of previous steps are continuous on the interior of each current parent. Hence the center evaluations are defined. At each step all children have countable dyadic addresses. Kernel rows are measurable L^2 functions of their parameters, and their Gaussian pairings may be constructed from fixed orthonormal-basis expansions. The countable nets in Lemma 3.1 give measurable point evaluations. These facts give a joint measurable construction. Since $\|A_{\mathrm{old}}\| \le CR$, (3.4) gives finite $L^2(\Omega \times Q)$ norm at each step; summing the zero-normal identities over the parents produces no boundary term.

Lemma 4.1. *Define*

$$\xi_{\mathrm{res}}(g) = W_{N+1}((I - P_N)g), \qquad \xi = \sum_{i=1}^{N} \xi_i + \xi_{\mathrm{res}}.$$

Then ξ has exactly the law of mean-zero white noise on Q, and

$$\operatorname{div}(\sqrt{K_*}Z_N) = \xi - \xi_{\mathrm{res}}. \tag{4.5}$$

Conditionally on the preceding construction the residual has the stated independent isonormal representation.

Proof. For deterministic g, the increments are conditionally centered normal variables with predictable variances $v_i = \|(P_i - P_{i-1})g\|_2^2$ and final variance $v_{N+1} = \|(I - P_N)g\|_2^2$. Nested orthogonal projections give, pathwise,

$$\sum_{i=1}^{N+1} v_i = \|(I - P_0)g\|_2^2.$$

The conditional Gaussian formula makes $\exp(i\sum_{i\le j}\xi_i(g)+\frac{1}{2}\sum_{i\le j}v_i)$ a complex martingale of initial mean one, including the residual as the last step. Its modulus is bounded by $\exp(\|g\|_2^2/2)$, so the final characteristic function is $\exp(-\|(I-P_0)g\|_2^2/2)$. Apply this to all real linear combinations of any finite collection of test functions to obtain the joint Gaussian law. The divergence formula follows by summing the local identities. The last assertion follows directly from independence of W_{N+1}. □

5. The quartic stopped estimate

In this section a uniform point $x\in Q$, independent of all the Gaussian processes, is included in the initial information. With probability one it misses all cell boundaries. All probabilities and expectations in this section include this point. Write $Z_i=Z_i(x)$. For its child at active step i, let G_i be the random Hölder bound in (3.6); set $G_i=0$ when inactive. The child and its matrix are known before the fresh noise is sampled. Conditional iteration of the exponential-square bound gives, after choosing a sufficiently large fixed constant,

$$\mathbb{P}\left\{\sum_{i=1}^{N}G_i^2>CK_*\right\}\le Ce^{-cK_*}. \tag{5.1}$$

The exponent in this estimate can be made any prescribed fixed multiple of K_* by enlarging C.

Stop the pointwise calculation at the first step with $|Z_i|>R$, retaining that step. Subsequent terms in the calculation are zero. At an active step let $\widehat{z}_i$ be the capped state used for its old metric. For $i\ge 2$ it is the cap of Z_{i-2} at the center of the preceding parent; for $i=1$ it is zero. Before the crossing put $\delta_i=|Z_{i-1}-\widehat{z}_i|$, and put it zero at inactive or subsequently stopped steps. Both vectors in this difference lie in the radius-R ball, so $\delta_i\le 2R$.

Lemma 5.1. *On the good event $\sum_{i=1}^{N}G_i^2\le CK_*$ from* (5.1),

$$\sum_i\delta_i^2\le CR^2. \tag{5.2}$$

Proof. Projection onto the cap ball cannot increase the distance from $Z_{i-1}\in\overline{B}(0,R)$. The last increment at x has size at most $CRG_{i-1}/\sqrt{K_*}$. For an earlier increment $j\le i-2$, both x and the center used in step $i-1$ belong to the jth child. In the jth parent's coordinates their distance is at most

$$C\|B_{P_{j-1}}^{-1}B_{P_j}\|\rho^{i-2-j}.$$

The first factor is comparable to the d_C of that child by (3.2). Apply (3.6) to the oscillation, and use $\|A_{\mathrm{old}}\|\le CR$. It follows that

$$\delta_i\le\frac{CR}{\sqrt{K_*}}\left(G_{i-1}+\sum_{j\le i-2}\rho^{\gamma(i-2-j)}G_j\right).$$

The first two steps obey the same formula with empty sums. The geometric sequence has bounded ℓ^1 norm. Its convolution therefore has ℓ^2 norm at most a constant times that of (G_j), which, together with (5.1), proves (5.2). □

Proposition 5.2. *For $1\le r\le R$ the stopped process satisfies*

$$\mathbb{P}\{\max_{i\le \textit{crossing or } N}|Z_i|>r\}\le Ce^{-cr^4}+Ce^{-cK_*}. \tag{5.3}$$

There is an absolute $\theta>0$ such that, if

$$\Phi_R(t)=\begin{cases}e^{\theta t^4}, & 0\le t\le R,\\ e^{\theta R^4}\big(1+4\theta R^3(t-R)\big), & t>R,\end{cases} \tag{5.4}$$

then the continued, unstopped field satisfies

$$\mathbb{E}\int_Q \Phi_R(|Z_N(x)|)\,dx \le C. \tag{5.5}$$

Proof. At an active step the increment has the form

$$\Delta_i = K_*^{-1/2} A(\widehat{z}_i) U_i,$$

where U_i is conditionally centered Gaussian with covariance at most CI. At inactive steps it is zero. All terms below are zero after the crossing. Let $S_i = 1 + |Z_i|^2$, frozen after crossing, and put $M_i = 2Z_{i-1}\cdot\Delta_i$. From (4.3),

$$\mathbb{E}(M_i^2 \mid \mathcal{F}_{i-1}) \le \frac{C}{K_*}(1 + R^2\delta_i^2), \qquad |\Delta_i|^2 \le \frac{C}{K_*}(S_{i-1} + \delta_i^2)|U_i|^2.$$

By Lemma 5.1, the sum of the conditional variance bounds is at most $C(1 + R^4/K_*) = C$, except on the event in (5.1). Conditional Gaussian exponential martingales imply

$$\mathbb{P}\left\{\max_\ell \left|\sum_{i\le\ell} M_i\right| > u\right\} \le Ce^{-cu^2} + Ce^{-cK_*}. \tag{5.6}$$

Here the exponential martingale is stopped when its predictable variance exceeds the fixed bound. Thus this argument restricts an event; it does not condition on the future event (5.1).

The conditional Gaussian-square formula also gives

$$\mathbb{P}\{\sum_i |U_i|^2 > CK_*\} \le Ce^{-cK_*}.$$

Furthermore, for $u \ge 1$,

$$\mathbb{P}\left\{\sum_i \delta_i^2|U_i|^2/K_* > u,\ \sum_i \delta_i^2 \le CR^2\right\} \le C\exp(-cK_*u/R^2). \tag{5.7}$$

To verify this, δ_i is predictable and at most $2R$. For a small fixed c, the conditional logarithm of the moment generating function of $(c/R^2)\delta_i^2|U_i|^2$ is at most $C\delta_i^2/R^2$. Iteration gives a compensated exponential bound; restrict its compensator to $\sum_i \delta_i^2 \le CR^2$ and apply Markov's inequality. This proves (5.7) without conditioning on that future restriction.

The exact identity $S_i - S_{i-1} = M_i + |\Delta_i|^2$ and discrete Gronwall give, on $\sum_i |U_i|^2 \le CK_*$,

$$\max_i S_i \le C\left(1 + max_\ell \left|\sum_{i\le\ell} M_i\right| + \sum_i \delta_i^2|U_i|^2/K_*\right).$$

Indeed the remaining nonnegative coefficient of S_{i-1} sums to a constant and the resulting Gronwall factor is its exponential. Use (5.6) at $u \asymp r^2$ and (5.7) at the same threshold. Since $K_*/R^2 = R^2 \ge r^2$, this proves (5.3), with enlarged constants for bounded r. In particular the crossing probability is at most Ce^{-cK_*}.

For the continued field, conditional orthogonality of increments and (3.4) give $\mathbb{E}|Z_N|^2 \le CR^2$. The function (5.4) is positive, nondecreasing and convex. On no crossing, integrate (5.3) up to R, with θ smaller than the tail constant, to bound its expectation. On crossing use the global bound

$$\Phi_R(t) \le e^{\theta R^4}(1 + 4\theta R^3 t).$$

Cauchy–Schwarz bounds its contribution by $Ce^{(\theta - c/2)K_*}(1 + R^4)$. Reduce θ again. This is bounded, and averaging over the independent uniform x proves (5.5). □

6. Smoothing and completion of the Gaussian source

Let G be the mean-zero Green function of Δ on T, so that $\Delta G = \delta_0 - 1/4$, and put $L_h = \nabla G * \eta_h$. For $x, z \in Q$ set

$$F_h(x,z) = \sum_{g\in\{\pm1\}^2} L_h(x - gz). \tag{6.1}$$

It has zero mean in z, zero x-normal component, and $\operatorname{div}_x F_h(x,z) = k_h(x,z) - 1$.

Lemma 6.1. *For all $x \in Q$,*

$$\|F_h(x,\cdot)\|_\infty \le C/h, \qquad \|F_h(x,\cdot)\|_2^2 \le C\log(2/h), \tag{6.2}$$

$$\|\nabla_z F_h(x,\cdot)\|_2^2 \le C/h^2, \qquad \|D_x^j F_h\|_\infty \le C_j h^{-1-j} \quad (0 \le j \le 3). \tag{6.3}$$

If P is projection onto constants of any convex partition with cell diameters at most h, then

$$\|(I-P)F_h(x,\cdot)\|_2^2 \le C. \tag{6.4}$$

Proof. Cut off $(2\pi)^{-1}\log|w|$ near zero in T. Its Laplacian is δ_0 plus a smooth function; subtract the required mean, and solve the smooth remainder by Fourier series. This constructs G and gives $|\nabla G(w)| \le C(1 + \operatorname{dist}_T(w,0)^{-1})$ and, away from zero, $|D^j\nabla G(w)| \le C_j(1 + \operatorname{dist}_T(w,0)^{-1-j})$. For the smoothed kernel split the convolution into distances $O(h)$ and its complement. On the former move derivatives to the bump; integration of $|w|^{-1}$ on a radius-Ch disk is $O(h)$. On the latter use the unsmoothed derivatives. Thus

$$|L_h(w)| \le \frac{C}{h + \operatorname{dist}_T(w,0)}, \qquad |\nabla L_h(w)| \le \frac{C}{(h + \operatorname{dist}_T(w,0))^2},$$

and the indicated derivative supremum bounds follow in the same way. Integration in dimension two proves (6.2)–(6.3), with the fixed reflection sum included.

For z, z' in one cell, $|z - z'| \le h$. Along their segment the denominator $h + \operatorname{dist}_T(x - gz, 0)$ changes by at most a constant factor. Consequently

$$|L_h(x - gz) - L_h(x - gz')| \le \frac{Ch}{(h + \operatorname{dist}_T(x - gz, 0))^2}.$$

Average over z' in that cell, square, and integrate in z. Summing over cells and reflections gives a bound by a constant times $\int_T h^2(h + \operatorname{dist}_T(w,0))^{-4}\,dw$, which is bounded. This is (6.4). □

Proposition 6.2. *Given a fixed $\varepsilon_0 > 0$, for all sufficiently small h there is a joint law of mean-zero white noise ξ and a smooth zero-normal field V_h such that*

$$\operatorname{div} V_h = \xi(k_h(x,\cdot) - 1), \qquad \int_Q \Phi_R\left(\frac{|V_h|}{C\sqrt{K_*}}\right) \le C \tag{6.5}$$

with probability at least $1 - \varepsilon_0$. Here K_*, R are as in (4.1), and C may depend on ε_0.

Proof. Reflect Z_N as a vector and convolve by η_h. The distributional zero-normal identity makes reflection and convolution commute with divergence. Jensen and (5.5) retain the bounded expected Φ_R integral. Its source, before normalization by $\sqrt{K_*}$, is the smoothing of $\xi - \xi_{\mathrm{res}}$.

Add the residual field

$$E_h(x) = W_{N+1}((I - P_N)F_h(x,\cdot)).$$

All final cells have diameter at most h, so (6.4) bounds its conditional pointwise variance by C. By the derivative estimates, its conditional expected squared $H^3(Q)$ norm is $O(h^{-8})$, uniformly in the random partition. The Sobolev embedding $H^3(Q) \hookrightarrow C^1(Q)$ and Markov's inequality imply, for example, a C^1 norm at most h^{-6} except with probability $O(h^4)$. On an h^8-mesh, the Gaussian union bound then gives

$$\|E_h\|_\infty \le C\sqrt{K_*}$$

with probability $1 - o(1)$; interpolation contributes at most h^2. Higher Sobolev orders give smooth versions. The divergence is exactly the missing residual source and its normal component is zero.

Adding a bounded normalized error preserves the estimate after a fixed rescaling: if $\|E\|_\infty \le B$, convexity gives

$$\Phi_R(|J + E|/2) \le \tfrac{1}{2}\Phi_R(|J|) + \tfrac{1}{2}\Phi_R(B).$$

For fixed B, the last term is uniformly bounded. Apply this observation and Markov's inequality to the expected integral, with failure budget ε_0, to prove the proposition. □

7. The empirical–Gaussian coupling

We record all details of the transfer because the source must remain exact, and the two empirical samples must retain their independent product law. The only non-elementary probabilistic input is the following local coupling: for every $\lambda \ge 1$ there is a coupling of $N_\lambda \sim \mathrm{Pois}(\lambda)$ and $Z \sim N(0,1)$ such that

$$\mathbb{P}\{|N_\lambda - \lambda - \sqrt{\lambda}Z| > t\} \le Ce^{-ct} \qquad (t \ge 1). \tag{7.1}$$

This is [2, Lemma 4.5]. The error has mean zero, since both marginals have their stated means.

Lemma 7.1. *For any fixed $\varepsilon_0 > 0$, h chosen by (2.2), and sufficiently large n, one can jointly realize independent uniform samples X, Y, mean-zero white noise ξ, and a smooth zero-normal error field E_n so that*

$$\rho_{X,h} - \rho_{Y,h} = \sqrt{2/n}\,\xi(k_h(x,\cdot) - 1) + \operatorname{div} E_n, \qquad \|E_n\|_\infty \le Ca_n \tag{7.2}$$

with probability at least $1 - \varepsilon_0$.

Proof. First treat one uniform sample U. Let P_h be projection onto the constants of the regular h-grid, and define $b_C(x) = h^{-2}\int_C F_h(x,z)\,dz$. Lemma 6.1 gives

$$\sum_C b_C(x) = 0, \quad \max_C |b_C(x)| \le C/h, \quad \sum_C |b_C(x)|^2 \le Ch^{-2}\log(2/h),$$
$$\mathrm{Lip}_x\, b_C \le C/h^2, \qquad \|(I - P_h)F_h(x,\cdot)\|_2^2 \le C.$$

We compare successive fluxes; applying div to each always gives the corresponding exact difference of smoothed mean-zero sources.

Replace $F_h(x, U_i)$ by $b_{C(U_i)}(x)$. The differences are independent, centered, bounded by C/h, and have variance at most C. Scalar Bernstein inequalities in the two coordinates give at threshold Da_n for the normalized sum the exponent

$$\frac{cD^2 \log n}{1 + D/A}.$$

A union bound over a mesh of size n^{-8} succeeds for sufficiently large fixed D. The normalized fields have deterministic Lipschitz bound C/h^2, so interpolation is negligible.

Now take $P \sim \mathrm{Pois}(n)$ independent of an infinite iid sequence extending U, and replace the first n points by the first P points, still dividing by n. On $|P - n| \le B\sqrt{n}$ the difference is a sum of at most $B\sqrt{n}$ independent centered b_C variables, conditionally on P. The probability of the complementary size event is at most B^{-2}. Bernstein's exponent at Da_n is at least

$$c\min\{D^2\sqrt{n}/B,\ Dna_nh\} \ge c\min\{D^2\sqrt{n}/B,\ DA\log n\}.$$

Choose B for the desired failure budget and then D for the same mesh argument. Subtracting the constant mode for every point means that a change of sample size creates no unaccounted constant source.

The Poisson cell counts N_C are independent with common mean $\lambda = nh^2$. Couple them independently to standard Gaussians Z_C using (7.1), conditionally on the already sampled counts.

The errors $E_C = N_C - \lambda - \sqrt{\lambda} Z_C$ are independent centered variables with uniform exponential tails. The error flux is $n^{-1} \sum_C E_C b_C$. Weighted Bernstein gives exponent at least

$$c \min \left\{ \frac{D^2 n^2 a_n^2 h^2}{\log(2/h)},\ Dna_n h \right\} \geq c' \min\{D^2 A^2, DA\} \log n.$$

Also $\max_C |E_C| \leq C \log n$ with probability $1 - o(1)$. On this event the flux Lipschitz bound is $O(h^{-2})$, because $h^{-2} \log n / n = O(1)$. A mesh argument again bounds this error uniformly by Ca_n.

Realize Z_C as $W_Q(\mathbf{1}_C/h)$, and fill the orthogonal complement by independent white noise. Then

$$\frac{\sqrt{\lambda}}{n} \sum_C Z_C b_C(x) = \frac{1}{\sqrt{n}} W_Q(P_h F_h(x, \cdot)).$$

The omitted Gaussian field $n^{-1/2} W_Q((I - P_h) F_h(x, \cdot))$ has variance at most C/n. Its expected squared H^3 norm is $O(n^{-1} h^{-8})$. The same Sobolev, net and interpolation argument as above bounds its supremum by Ca_n with probability $1 - o(1)$. For example its Lipschitz norm is at most n^5 with probability $1 - o(1)$, and the n^{-8} mesh suffices.

Combining these four differences proves the one-sample assertion. Use independent copies of the entire construction for X and Y. They retain their required product law, and $\xi = (W_X - W_Y)/\sqrt{2}$ is standard white noise; constants are irrelevant when paired with $k_h(x, \cdot) - 1$. Summing the smooth zero-normal errors proves (7.2). □

8. A bounded comparison field and the quartic splice

The classical Leighton–Shor theorem supplies, with probability $1 - o(1)$, a matching from either sample to the same evenly spread n-point set whose largest edge has length at most $C(\log n)^{3/4}/\sqrt{n}$. Compose the two matchings. Thus, writing $K_* = \log(2/h)$ and $R = K_*^{1/4}$ as before, one obtains a matching of X to Y of maximum length at most $Ca_n R$. The stronger classical failure estimate is $C \exp(-c(\log n)^{3/2})$ [3, Theorem 4.7.1].

Lemma 8.1. *On the count event of Lemma 2.1 and this bottleneck event, there is a smooth vector-reflection-invariant field B such that*

$$\operatorname{div} B = \rho_{X,h} - \rho_{Y,h}, \qquad \|B\|_\infty \leq Ca_n R. \tag{8.1}$$

Proof. Let $s_* = Ca_n R$ bound the lengths in the classical matching. First convolve its straight paths at radius s_* on the reflected torus. The field is

$$B_{s_*}(x) = \frac{1}{n} \sum_i \sum_g \int_0^1 g(Y_{\pi(i)} - X_i) \eta_{s_*}(x - g((1-t)X_i + tY_{\pi(i)}))\, dt.$$

Differentiation in t gives divergence $\rho_{X,s_*} - \rho_{Y,s_*}$. A contributing path starts within Cs_* of x. The count bound implies that at most Cns_*^2 starting points contribute. Each velocity is at most s_* and each kernel at most Cs_*^{-2}, so $\|B_{s_*}\|_\infty \leq Cs_*$.

For each marginal repair the smoothing radius from s_* to h by the dilation paths $gX_i + s(t)w$, where w has density η and $s(t)$ runs linearly between the two radii. Their time densities are the corresponding reflected smoothings, uniformly bounded by Lemma 2.1, and their speeds are $O(s_*)$. Their integrated fluxes are bounded by Cs_* and have exactly the divergence of the endpoint density difference. Choose the orientations and subtract the Y repair to obtain (8.1). All fields have the required vector reflection symmetry and hence zero normal component on Q. □

Lemma 8.2 (Divergence-preserving splice)**.** *Let $R \geq 1$. Suppose smooth vector-reflection-invariant fields J, B on T have the same divergence and*

$$\int_Q \Phi_R(|J|) \leq D, \qquad \|B\|_\infty \leq C_B R.$$

Then a smooth vector-reflection-invariant field J_* with that divergence satisfies

$$\int_Q e^{c_*|J_*|^4} \le C_*, \tag{8.2}$$

where $c_* > 0, C_* < \infty$ depend only on D, C_B, θ.

Proof. Let M be the centered Hardy–Littlewood maximal operator on T, and set $E = \{M|J| \le \lambda\}$, where $\lambda = C_2 R \ge R$ and C_2 is fixed. Since $J - B$ is divergence-free and has zero mean by vector reflection, its nonzero Fourier coefficients give a periodic smooth stream function u with $J - B = \nabla^\perp u$ and mean zero.

We recall the pointwise Sobolev estimate

$$|u(x) - u(y)| \le C \operatorname{dist}_T(x, y)(M|\nabla u|(x) + M|\nabla u|(y)).$$

To prove it, telescope the averages over balls of radii $2^{-j} \operatorname{dist}_T(x, y)$ about each point, using the L^1 Poincaré inequality, and compare the two largest balls in a ball of three times the radius. Poincaré on a ball follows by integration on segments, and the same argument in finitely many charts treats large torus distances. On E, $M|\nabla u| \le \lambda + C_B R$, so $u|_E$ is CR-Lipschitz. The McShane formula $v(x) = \inf_{y \in E}(u(y) + CR \operatorname{dist}_T(x, y))$ extends it with the same Lipschitz constant. If E is empty take $v = 0$.

The mean-zero stream function has the determinant reflection symmetry $u(gx) = \det(g)u(x)$. Averaging v with these signs retains its Lipschitz constant and agreement with u on the invariant set E. Consequently $K = B + \nabla^\perp v$ has the desired divergence, $\|K\|_\infty \le CR$, and $K = J$ almost everywhere on E. For the last statement use locality of weak derivatives on the set where two Lipschitz functions agree; it follows, for instance, by applying the one-dimensional absolute-continuity property on almost every parallel line and then to both coordinate derivatives.

Convexity implies $\Phi_R(M|J|) \le M(\Phi_R(|J|))$. The weak $(1, 1)$ maximal inequality gives

$$|E^c| \le CD/\Phi_R(\lambda) \le CDe^{-\theta R^4}. \tag{8.3}$$

Here weak type follows from the disjoint-ball covering argument: choose balls whose averages exceed the threshold, select disjoint ones by decreasing radius, and cover the original set by their fixed dilates. For $x \in E$ with $|J| \le R$, a sufficiently small c_* bounds $e^{c_*|J|^4}$ by $\Phi_R(|J|)$. The part with $R < |J| \le \lambda$ has measure at most $De^{-\theta R^4}$, and its exponential is at most $e^{c_* C_2^4 R^4}$. On E^c use (8.3) and $|K| \le CR$. Choosing c_* below both relevant constant multiples of θ bounds $\int_Q e^{c_*|K|^4}$.

Finally mollify v with an arbitrarily narrow even torus kernel. The resulting field equals $K * \eta_\delta + B - B * \eta_\delta$. Its divergence is still $\operatorname{div} B$, and the last term can have supremum at most one. Jensen and a further fixed reduction of c_* preserve the bound. This proves (8.2) with a smooth field. □

9. Proof of the theorem and the indexing convention

Proof of Theorem 1.1. Use the mesh (2.2). Choose a small fixed failure budget ε_0, for example $1/20$. The white-noise marginal in Proposition 6.2 agrees exactly with the white-noise marginal in Lemma 7.1. Join their two joint laws by conditioning on this common marginal. This is legitimate on countable isonormal coordinates; the fields are also determined by countably many coordinates. Equivalently, regular conditional probabilities on these standard Borel spaces give the usual gluing of probability laws.

Since $\sqrt{K_*/n} \asymp a_n$, the field $J = \sqrt{2/n}\, V_h + E_n$ has exact divergence $\rho_{X,h} - \rho_{Y,h}$ and obeys

$$\int_Q \Phi_R(|J|/(Ca_n)) \le C$$

except on an event of probability at most $2\varepsilon_0 + o(1)$. The bounded normalized error is absorbed by convexity as in the proof of Proposition 6.2. The count and classical bottleneck events have

probability $1 - o(1)$. On their intersection, Lemma 8.1 supplies a same-divergence comparison field. Apply Lemma 8.2 in units of Ca_n to obtain a smooth zero-normal field with

$$\int_Q \exp\left((|J_*|/(C'a_n))^4\right) \le C''.$$

The density bounds and Lemma 2.2 now give a single permutation satisfying (1.1). The intersection probability is larger than $1/2$ for all sufficiently large n. For the finitely many remaining $n \ge 2$, increase L so that the deterministic bound $|X_i - Y_k| \le \sqrt{2}$ makes every summand at most two. Finally $|X_i^j - Y_{\pi(i)}^j| \le |X_i - Y_{\pi(i)}|$ proves (1.2) for both coordinates of the same permutation. □

Corollary 9.1. *There is an absolute $L' > 0$ such that, for each $N \ge 2$ and independent uniform samples $(X_i)_{i=0}^N, (Y_i)_{i=0}^N$, with probability at least $1/2$ a permutation of $\{0, \dots, N\}$ satisfies*

$$\sum_{i=0}^{N} \exp\left(\left(\sqrt{\frac{N}{\log N}}\frac{|X_i^j - Y_{\pi(i)}^j|}{L'}\right)^4\right) \le 2N, \qquad j = 1, 2.$$

Proof. Apply Theorem 1.1 with $n = N + 1$. Multiply the displacement scale by $4^{1/4}$. Concavity of $t \mapsto t^{1/4}$ bounds the new exponential average by $2^{1/4} < 4/3$. Its sum is therefore at most $(4/3)(N+1) \le 2N$. The scales $\sqrt{\log(N+1)/(N+1)}$ and $\sqrt{\log N/N}$ are comparable by absolute constants for $N \ge 2$. Increasing L' absorbs this comparison. This also covers the literal indexing of [4]. □

References

[1] M. Costabel and A. McIntosh, *On Bogovskiĭ and regularized Poincaré integral operators for de Rham complexes on Lipschitz domains*, Math. Z. **265** (2010), no. 2, 297–320. doi:10.1007/s00209-009-0517-8.

[2] R. Lachièze-Rey and S. Muirhead, *Asymptotics for the critical level and a strong invariance principle for high intensity shot noise fields*, arXiv:2111.09180v2 (2022), Lemma 4.5. `https://arxiv.org/abs/2111.09180`.

[3] M. Talagrand, *Upper and Lower Bounds for Stochastic Processes: Decomposition Theorems*, second ed., Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 60, Springer, 2021. doi:10.1007/978-3-030-82595-9.

[4] M. Talagrand, *Matchings*, problem note, available online.

Email address: `howew@purdue.edu`

Email address: `huffma97@purdue.edu`

Email address: `jenwin@purdue.edu`

Email address: `kruerl@purdue.edu`