hexagon:2611.01367v1

THE SHARP INVERSE HÖLDER EXPONENT FOR THE AVERAGED SLICED WASSERSTEIN DISTANCE

WEIMING ZHANG

Abstract. We determine the best power with which the averaged sliced Wasserstein distance of order $p > 1$ controls the ambient Wasserstein distance for arbitrary probabilities on the unit ball in dimension $d \geq 2$. The power is $p/(2p + d - 2)$ below order two, $2/(d + 2)$ from order two through $1 + d/2$, and $1/p$ above $1 + d/2$. It is attained throughout. We prove the comparison by multiscale coupling below order two and by bounding the full transport cost of order $1 + d/2$ by the sliced quadratic distance.

Contents

1. Introduction 1
2. Setting and notation 3
3. Multiscale comparison below order two 4
4. A smooth test for the transport cost 7
5. The critical comparison 10
6. Oscillation in a two-dimensional plane 12
7. An example with small displacements in every projection 15
8. The sharp exponent 18
References 20

1. Introduction

One-dimensional projections determine a probability measure. How far apart can two measures be if transporting one projected measure to the other costs little on average? This question arises when multidimensional transport is estimated from its one-dimensional marginals. A coupling on each line is cheap to compute by monotone rearrangement, but the couplings on different lines need not arise from a single coupling in the ambient space. Rabin, Peyré, Delon and Bernot averaged one-dimensional transport costs over directions to compute sliced barycenters of point clouds [RPDB12].

Fix an integer $d \geq 2$. Let $\mathcal{P}(B_1)$ denote the Borel probability measures on the closed Euclidean unit ball in $\mathbb{R}^d$. We allow atoms and singular measures. For $1 \leq q < \infty$, let W_q be the infimum, over couplings, of the L^q norm of the displacement. Take σ to be probability surface measure on $\mathbb{S}^{d-1}$ and write μ^θ for the image of μ under $x \mapsto x \cdot \theta$. The averaged sliced distance is

$$\mathrm{SW}_q(\mu, \nu) = \left(\int_{\mathbb{S}^{d-1}} W_q(\mu^\theta, \nu^\theta)^q \, d\sigma(\theta) \right)^{1/q}.$$

Projecting an ambient coupling gives $\mathrm{SW}_q \le W_q$. On the space of probabilities on $\mathbb{R}^d$ with finite qth moment, Kitagawa and Takatsu showed that SW_q and W_q induce the same topology and that the ratio W_q/SW_q is unbounded over distinct pairs when $d \ge 2$ [KT26]. For fixed $1 < p < \infty$, call $\alpha > 0$ admissible if the reverse bound $W_p \le C_{d,p,\alpha}\,\mathrm{SW}_p^\alpha$ holds uniformly on $\mathcal{P}(B_1)$. Write $\alpha_*(d,p)$ for the supremum of these exponents and set $s = 1 + d/2$.

Bonnotte established a power comparison on compact supports [Bon13, Chapter 5]. At order one, Carlier, Figalli, Mérigot and Wang proved the endpoint estimate $W_1 \le C_d\,\mathrm{SW}_1^{1/d}$ on B_1 for every $d \ge 2$ by reconstructing a smoothed Lipschitz potential from its projections [CFMW25, Theorem 2.1]. Their compactly supported examples prove sharpness for $d \ge 3$ [CFMW25, Section 2.5]. For the maximum of the projected W_1 distances, Bobkov and Götze obtained the power $2/(d+2)$ on compact supports [BG]. Niles-Weed and Shkrob proved sharpness for that maximum [NS26, Theorem 1.1]. In dimension two, their unit-disk examples satisfy $W_1 \ge \rho$ and $\mathrm{SW}_1 \le \max_\theta W_1(\mu^\theta, \nu^\theta) \le C\rho^2$ for arbitrarily small $\rho > 0$, ruling out every power above $1/2$ for the angular average. They also gave a lattice example for the average [NS26, Lemma 4.1]. For $p > 1$, Letrouit and Mérigot proved a reverse Hölder estimate for SW_p under a uniform qth moment bound with $q > p$ [LM26, Theorem 1.8]. Hopper proved that every admissible exponent satisfies $\alpha \le 1/p$ [Hop26, Section 4.2]. We show that $1/p$ is admissible precisely when $p \ge s$. Park and Slepčev related the sliced quadratic distance to a negative Sobolev norm under density and support hypotheses [PS25, Theorem 5.2]. For general measures at $p > 1$, an estimate on a Lipschitz potential controls the mean displacement, while the cost also depends on how much mass travels a long distance.

Carlier et al. [CFMW25, Introduction] identified the case $p > 1$ as open. Niles-Weed and Shkrob [NS26, footnote 2] subsequently noted that the optimal powers were unknown for both the angular average and the maximum over directions. The following theorem determines the optimal power for the angular L^p average on the unit ball at every order $1 < p < \infty$.

Theorem 1.1. *For every $d \ge 2$ and $1 < p < \infty$, the admissible exponents form the interval $(0, \alpha_*(d,p)]$, where*

$$\alpha_*(d,p) = \begin{cases} p/(2p+d-2), & 1 < p < 2, \\ 2/(d+2), & 2 \le p \le s, \\ 1/p, & p > s. \end{cases}$$

Theorem 1.1 is Theorem 8.2, proved in Section 8. The right endpoint is included even at the two transition orders. In dimension three, for example, the optimal exponents at $p = 3/2$, 2, and 3 are $3/8$, $2/5$, and $1/3$, respectively.

The estimates for the entire range $p \ge 2$ follow from one comparison in which the ambient and projected costs have different orders.

Theorem 1.2. *For every $\mu, \nu \in \mathcal{P}(B_1)$,*

$$W_s(\mu,\nu)^s \le C_d\,\mathrm{SW}_2(\mu,\nu).$$

Theorem 1.2 is Theorem 5.3, proved in Section 5. For $2 \le p \le s$, the critical cost bounds W_p^s by monotonicity. For $p > s$, the diameter of the support lets the same cost bound W_p^p.

Below order two we construct one coupling whose tail at every dyadic scale is bounded by the mass left unmatched after smoothing at that scale. Reconstruction

from the projected distribution functions estimates that mass. Interpolation in the direction variable gives the power $2 + (d-2)/p$ of the inverse smoothing radius. Integrating the tail of the coupling proves that $p/(2p+d-2)$ is admissible. The planar example in Theorem 6.2 rules out every larger exponent.

At the critical order this scale-by-scale bound produces a logarithm. Given a near-optimal pair of Kantorovich potentials u, v, we first raise the upper potential to $v + \varepsilon/2$. On each ball of the cover we choose a constant between u and $v + \varepsilon/2$ on its intersection with B_1. A smooth partition of unity joins these constants into a test f_ε satisfying $u \le f_\varepsilon \le v + \varepsilon/2$ on B_1. The sizes of the balls control the derivatives of the partition, and the Hardy–Littlewood maximal inequality bounds the fractional derivatives of the resulting test. We reconstruct the smoothed test from one-dimensional functions and expand each at the projection of the source point. The first derivatives have a uniformly bounded angular second moment at every source point. This allows us to integrate against an arbitrary source measure, and the quadratic remainder is controlled by the projected transport cost. Letting ε tend to zero removes the error in the cost estimate.

The three sharpness examples distinguish the mechanisms behind the three exponents. The first example oscillates in two coordinates. Its projected distribution functions differ mainly at offsets close to zero. When $d > 2$, their difference also decays away from directions close to the plane of oscillation. A second construction keeps the displacement small in every projection while the ambient transport still detects oscillation. Finally, moving a small atom between two interior points costs a fixed distance per unit mass in the ambient space. A common spherical component allows that mass to move by much shorter distances in projection.

Section 2 fixes notation. Section 3 constructs the multiscale coupling and proves the estimate below order two. Sections 4 and 5 construct the smooth test and prove the critical comparison. Sections 6 and 7 give the two oscillatory examples. Section 8 treats the moving atom and combines the estimates.

2. Setting and notation

We work in $\mathbb{R}^d$, $d \ge 2$, with Euclidean norm $|\cdot|$. Write $B_R = \{x : |x| \le R\}$ and let $\mathcal{P}(B_R)$ denote the Borel probability measures supported on B_R. Let σ be probability surface measure on $\mathbb{S}^{d-1}$. For $\theta \in \mathbb{S}^{d-1}$ set $P_\theta x = x \cdot \theta$ and $\lambda^\theta = (P_\theta)_\# \lambda$. For $1 \le q < \infty$ and probabilities on the same Euclidean space we use

$$W_q(\lambda, \tau) = \inf_{\pi \in \Pi(\lambda,\tau)} \left(\int |x-y|^q \, d\pi(x,y) \right)^{1/q},$$

$$\mathrm{SW}_q(\mu, \nu) = \left(\int_{\mathbb{S}^{d-1}} W_q(\mu^\theta, \nu^\theta)^q \, d\sigma(\theta) \right)^{1/q}.$$

Here $\Pi(\lambda, \tau)$ is the set of couplings and the sliced distance is taken in ambient dimension d. We set $s = 1 + d/2$. Constants may depend on fixed dimensions, orders and fixed smoothing functions.

Definition 2.1. For $1 < p < \infty$, an exponent $\alpha > 0$ is *admissible* if $W_p(\mu, \nu) \le C_{d,p,\alpha} \, \mathrm{SW}_p(\mu, \nu)^\alpha$ uniformly for $\mu, \nu \in \mathcal{P}(B_1)$. We write $\alpha_*(d, p)$ for the supremum of the admissible exponents.

We use $\widehat{f}(\xi) = \int e^{-ix\cdot\xi} f(x) \, dx$ and define $|\nabla|^t$ by the multiplier $|\xi|^t$. For integers $k \ge 0$ and $0 < \beta \le 1$, the norm in $C^{k,\beta}(\mathbb{R}^d)$ includes the supremum norms of the

derivatives through order k and the β-Hölder seminorms of the derivatives of order k. The notation $A \lesssim B$ means $A \le CB$ with the fixed-parameter dependence of the statement in which it occurs.

3. Multiscale comparison below order two

Fournier and Guillin [FG15, Lemma 5] bound the transport cost by a weighted sum of mass discrepancies over dyadic partitions. Here we use smooth convolutions to construct one coupling whose displacement tail is controlled at every dyadic scale. If we smooth $\mu, \nu \in \mathcal{P}(B_1)$ by moving each point at most h, the common part of the two densities can be matched using pairs of original points at distance at most $2h$. We repeat this matching at dyadic scales and use the projections μ^θ, ν^θ to estimate the mass left at each scale. In this section, for $G \in L^1(\mathbb{R}^d)$ and $h > 0$ we write $G_h(x) = h^{-d} G(x/h)$.

Lemma 3.1. *Let $d \ge 2$ and put $h_j = 2^{-j}$ for integers $j \ge 0$. There exists a radial probability density $K \in C_c^\infty(\mathbb{R}^d)$ supported in B_1 such that for every $\mu, \nu \in \mathcal{P}(B_1)$ there is a coupling $\pi \in \Pi(\mu, \nu)$ satisfying*

$$\pi\{(x,y) : |x-y| > 2h_j\} \le \tfrac{1}{2} \|(\mu - \nu) * K_{h_j}\|_{L^1(\mathbb{R}^d)} \qquad \text{for every } j \ge 0.$$

Proof. We choose K so that one convolution takes us from h_j to $2h_j$. This lets us match the overlap of the densities at successively coarser scales. A limit of these finite matchings gives the coupling.

A convolution identity. Fix a radial smooth probability density φ supported in B_1. Let Z_ℓ, $\ell \ge 1$, be independent with this density. The series $Z = \sum_{\ell \ge 1} 2^{-\ell} Z_\ell$ converges absolutely in B_1, and its law is rotation invariant. It has a smooth density K, since it is the convolution of $\varphi_{1/2}$ with the law of the independent remainder. The sum $2Z - Z_1$ has the law of Z and is independent of Z_1. Dilating the law of $2Z$ therefore gives

$$K_{2h} = K_h * \varphi_h \qquad (h > 0). \tag{3.1}$$

The unmatched mass at each scale. Fix $J \ge 1$. At scale h_J we keep track of the original points x, y as well as the position z after smoothing. The two measures on original and current positions are $M_z(dx)\,dz$ and $N_z(dy)\,dz$, initially with

$$M_z(dx) = K_{h_J}(z - x)\,\mu(dx), \qquad N_z(dy) = K_{h_J}(z - y)\,\nu(dy).$$

At level j, let $a(z), b(z)$ be the masses of M_z, N_z and put $c(z) = \min\{a(z), b(z)\}$. Matching the common mass $c(z)\,dz$ by a product at each z gives the measure of matched pairs

$$\int_{\{c>0\}} c(z)\, \frac{M_z}{a(z)} \otimes \frac{N_z}{b(z)}\, dz.$$

At each z with $c(z) > 0$ we remove the fraction $c(z)/a(z)$ of M_z and the fraction $c(z)/b(z)$ of N_z. This leaves current densities $a - c$ and $b - c$. For $j > 0$, we convolve both remaining measures in the z coordinate with φ_{h_j} and repeat at level $j - 1$.

The subtraction preserves the difference $a - b$. At the finest level this difference is $(\mu - \nu) * K_{h_J}$, and by (3.1) each convolution takes $(\mu - \nu) * K_{h_j}$ to $(\mu - \nu) * K_{h_{j-1}}$. The remaining masses on the two sides are equal and $(a - c) + (b - c) = |a - b|$. Thus the mass left on either side after matching at level j is

$$m_j = \tfrac{1}{2} \|(\mu - \nu) * K_{h_j}\|_{L^1(\mathbb{R}^d)}.$$

Each original point starts within h_J of its current position. The convolution at level j adds at most h_j to this distance, so the bound h_j becomes $2h_j = h_{j-1}$ at the next level. Every pair matched at level j is therefore at distance at most $2h_j$.

One coupling for all scales. After level zero, if $m_0 > 0$, we add the product of the remaining original marginals divided by m_0. The resulting measure π_J has marginals μ, ν, since each match removes exactly its own marginals. For $0 \le j \le J$, the levels J through j have matched a mass $1 - m_j$ at distances at most $2h_j$. By compactness of $B_1 \times B_1$, a subsequence of π_J converges weakly as $J \to \infty$ to a coupling π with the same marginals. For each fixed j, the set $\{|x - y| > 2h_j\}$ is open, so its π-mass is at most the lower limit of its masses along the subsequence, hence at most m_j. □

Lemma 3.2. *Let $d \ge 2$, $1 \le q \le 2$, and let $K \in C_c^\infty(\mathbb{R}^d)$ be radial and supported in B_1. Then for all $0 < h \le 1$ and $\mu, \nu \in \mathcal{P}(B_1)$,*

$$\|(\mu - \nu) * K_h\|_{L^1(\mathbb{R}^d)} \le C_{d,q,K}\, h^{-2-(d-2)/q}\, \mathrm{SW}_q(\mu, \nu).$$

Proof. We reconstruct the smoothed density from the distribution functions of the projections. Integrating the reconstruction kernel in space gives an angular L^1 bound, and Plancherel gives an angular L^2 bound. Splitting the directions between these two estimates gives the power for intermediate q.

Reconstruction from projected distribution functions. Put $\eta = \mu - \nu$. For $\theta \in \mathbb{S}^{d-1}$ and $t \in \mathbb{R}$ set

$$F_\theta(t) = \eta\{x : P_\theta x \le t\}, \qquad a(\theta) = \int_{\mathbb{R}} |F_\theta(t)|\, dt.$$

The half-space indicator is jointly Borel in (θ, t, x), so $(\theta, t) \mapsto F_\theta(t)$ is measurable. Each F_θ vanishes off $[-1, 1]$. For any coupling $\gamma \in \Pi(\mu^\theta, \nu^\theta)$,

$$a(\theta) \le \int_{\mathbb{R}^2} \int_{\mathbb{R}} \left|\mathbf{1}_{\{u \le t\}} - \mathbf{1}_{\{v \le t\}}\right| dt\, d\gamma(u, v) = \int_{\mathbb{R}^2} |u - v|\, d\gamma(u, v).$$

Hölder's inequality and the infimum over γ give $a(\theta) \le W_q(\mu^\theta, \nu^\theta)$. In particular, $\|a\|_{L^q(\sigma)} \le \mathrm{SW}_q(\mu, \nu)$.

Integrating first in t over $[-1, 1]$ and using $\eta(B_1) = 0$ gives

$$ir\widehat{F}_\theta(r) = \int_{B_1} \left(e^{-irP_\theta x} - e^{-ir}\right) d\eta(x) = \widehat{\eta}(r\theta).$$

Write $\kappa(r) = \widehat{K}(re_1)$ and let $c_d = |\mathbb{S}^{d-1}|/(2\pi)^d$, with $|\mathbb{S}^{d-1}|$ the surface area. For a measurable set $E \subset \mathbb{S}^{d-1}$ define

$$U_E(x) = c_d \int_E \int_0^\infty e^{irP_\theta x}\, ir^d \kappa(hr) \widehat{F}_\theta(r)\, dr\, d\sigma(\theta).$$

The integral converges absolutely since $|\widehat{F}_\theta| \le a(\theta)$ and κ is rapidly decreasing. Radiality of $\widehat{K}$ and Fourier inversion in polar coordinates give $U_{\mathbb{S}^{d-1}} = \eta * K_h$. Its support lies in $B_{1+h} \subset B_2$.

Two estimates on a set of directions. We show that

$$\|U_E\|_{L^1(B_2)} \lesssim \min\left\{ h^{-d} \int_E a\, d\sigma,\ h^{-s} \left(\int_E a^2\, d\sigma \right)^{1/2} \right\}. \tag{3.2}$$

For the first bound put $g(r) = ir^d \kappa(r)$ on $[0, \infty)$ and $k(w) = \int_0^\infty e^{irw} g(r)\, dr$. The smoothness and rapid decay of κ and its derivatives give $g, g', g'' \in L^1(0, \infty)$.

Moreover $g(0) = g'(0) = 0$ because $d \geq 2$. Two integrations by parts give $|k(w)| \leq |w|^{-2}\|g''\|_1$ for $w \neq 0$. Together with $|k(w)| \leq \|g\|_1$ this gives $k \in L^1(\mathbb{R})$. Substituting the Fourier integral for $\widehat{F}_\theta$ and rescaling in r, we obtain

$$U_E(x) = c_d h^{-d-1} \int_E \int_{\mathbb{R}} k\big((P_\theta x - t)/h\big) F_\theta(t)\, dt\, d\sigma(\theta).$$

The sections of B_2 perpendicular to θ have uniformly bounded $(d-1)$-dimensional volume. Hence $\int_{B_2} |k((P_\theta x - t)/h)|\, dx \leq C_d h \|k\|_1$ uniformly in θ, t, giving the first bound.

For the second bound, polar coordinates express U_E as the inverse Fourier transform of $\mathbf{1}_E(\theta)\, ir\kappa(hr)\widehat{F}_\theta(r)$ at frequencies $r\theta$ with $r > 0$. Plancherel and $|\widehat{F}_\theta| \leq a(\theta)$ give

$$\|U_E\|^2_{L^2(\mathbb{R}^d)} \lesssim \int_E a^2\, d\sigma \int_0^\infty r^{d+1} |\kappa(hr)|^2\, dr \lesssim h^{-d-2} \int_E a^2\, d\sigma.$$

Cauchy–Schwarz on B_2 gives the second bound in (3.2).

The angular L^q estimate. Taking $E = \mathbb{S}^{d-1}$ gives the assertion for $q = 1, 2$. For $1 < q < 2$, write $A = \|a\|_{L^q(\sigma)}$. If $A = 0$, the reconstruction integral vanishes. Otherwise, put $\lambda = Ah^{-(d-2)/q}$ and $E = \{\theta : a(\theta) > \lambda\}$. On E we have $\int_E a\, d\sigma \leq \lambda^{1-q} A^q$. On its complement $\int_{E^c} a^2\, d\sigma \leq \lambda^{2-q} A^q$. Applying the first bound in (3.2) to U_E and the second to U_{E^c} therefore yields

$$\|\eta * K_h\|_1 \leq \|U_E\|_{L^1(B_2)} + \|U_{E^c}\|_{L^1(B_2)} \lesssim h^{-d}\lambda^{1-q}A^q + h^{-s}\lambda^{1-q/2}A^{q/2}.$$

Both terms equal $h^{-2-(d-2)/q}A$, and $A \leq \mathrm{SW}_q(\mu, \nu)$. □

Theorem 3.3. *Let $1 < p < 2$. For all $\mu, \nu \in \mathcal{P}(B_1)$,*

$$W_p(\mu, \nu) \leq C_{d,p}\, \mathrm{SW}_p(\mu, \nu)^{p/(2p+d-2)}.$$

Proof. We integrate the tail of the displacement $|x - y|$ under a single coupling. Put $D = \mathrm{SW}_p(\mu, \nu)$ and $b = 2 + (d-2)/p$. For the kernel K and the coupling π of Lemma 3.1, apply Lemma 3.2 with $q = p$ and $h = h_j = 2^{-j}$. Since K is fixed by the dimension, we obtain

$$\pi\{|x - y| > 2h_j\} \leq C_{d,p} D h_j^{-b} \qquad (j \geq 0).$$

For $0 < t < 2$ choose $j \geq 1$ with $2h_j \leq t < 4h_j$. The tail is decreasing in t and vanishes for $t \geq 2$, so for some C depending only on d, p,

$$\pi\{|x - y| > t\} \leq \min\{1, CDt^{-b}\} \qquad (t > 0).$$

If $D = 0$, the coupling is concentrated on the diagonal. For $D > 0$ put $\rho = (CD)^{1/b}$. Since $b \geq 2 > p$, integration of the tail gives

$$\begin{aligned} W_p(\mu, \nu)^p &\leq \int |x - y|^p\, d\pi = \int_0^\infty pt^{p-1} \pi\{|x - y| > t\}\, dt \\ &\leq \int_0^\rho pt^{p-1}\, dt + \rho^b \int_\rho^\infty pt^{p-b-1}\, dt = \frac{b}{b-p}\, \rho^p. \end{aligned}$$

Taking the pth root gives the exponent $1/b = p/(2p + d - 2)$. □

4. A smooth test for the transport cost

For $\mu, \nu \in \mathcal{P}(B_1)$, the dual problem for $W_p(\mu,\nu)^p$ gives two potentials. We construct a single smooth function between them after a small upward shift of the upper potential. We use the compact-space form of Kantorovich duality [Vil03, Theorem 1.3].

Theorem 4.1 (Kantorovich duality)**.** *For a compact metric space X, a continuous cost $c : X \times X \to \mathbb{R}$ and probabilities λ, τ on X, the infimum of $\int c\, d\pi$ over $\Pi(\lambda,\tau)$ equals the supremum of $\int u\, d\lambda - \int v\, d\tau$ over continuous u, v on X satisfying $u(x) - v(y) \le c(x,y)$ for all $x, y \in X$.*

For the power cost, raising v by $\delta > 0$ gives $u(a) \le v(b) + \delta$ whenever $|a-b| \le \delta^{1/p}$. Thus a constant fits between u and $v + \delta$ on the portion of any ball of diameter at most $\delta^{1/p}$ in B_1. The potentials may allow a larger radius at a given center. Our derivative estimate at these varying radii uses the maximal inequality [Ste70, Chapter I, Theorem 1(c)].

Theorem 4.2 (Hardy–Littlewood maximal inequality)**.** *For $1 < q < \infty$, the centered maximal operator of ball averages of the absolute value maps $L^q(\mathbb{R}^d)$ boundedly into $L^q(\mathbb{R}^d)$.*

Write $B(z,r) = \{x : |x-z| < r\}$ for an open ball and M for the centered maximal operator. Since M takes a supremum over radii, the same norm bound applies when the radius of each average depends on its center.

Lemma 4.3. *Let $0 < \gamma < 2$, let $\rho : \mathbb{R}^d \to (0,1]$ be 1-Lipschitz, and let $G \in C_c^\infty(B(0,3))$. Suppose that for some $A > 0$ and every $z \in \mathbb{R}^d$,*

$$|G(z)| \le A\rho(z)^\gamma, \qquad |D^2 G(z)| \le A\rho(z)^{\gamma-2}.$$

Then $\||\nabla|^\gamma G\|_{L^q(\mathbb{R}^d)} \le C_{d,\gamma,q} A$ for every $1 < q < \infty$.

Proof. The Hessian controls second differences up to the local radius. At larger distances we use the size of G and the maximal inequality.

Second differences at small distances. The Fourier multiplier has the integral representation

$$|\nabla|^\gamma G(z) = c_{d,\gamma} \int_{\mathbb{R}^d} \frac{2G(z) - G(z+h) - G(z-h)}{|h|^{d+\gamma}}\, dh.$$

Indeed, the L^1 norm in z of the numerator is at most $\min\{|h|^2 \|D^2 G\|_1, 4\|G\|_1\}$ by Taylor's formula and translation invariance. This permits Fourier transformation under the integral. The resulting multiplier before normalization is the integral of $(2 - 2\cos(h\cdot\xi))|h|^{-d-\gamma}$ in h. This integral is finite for $0 < \gamma < 2$ and positive for $|\xi| = 1$. Rotation and scaling give its value at ξ as $c_{d,\gamma}^{-1}|\xi|^\gamma$.

For $|h| \le \rho(z)/2$, the Lipschitz bound gives $\rho \ge \rho(z)/2$ along both segments from z to $z \pm h$. The Hessian bound thus controls the absolute integral over these h by

$$C_d A \rho(z)^{\gamma-2} \int_0^{\rho(z)/2} t^{1-\gamma}\, dt \le C_{d,\gamma} A.$$

This part vanishes for $|z| > 4$ since G is supported in $B(0,3)$ and $\rho \le 1$.

Ball averages at large distances. On $|h| > \rho(z)/2$, the contribution from $2G(z)$ is bounded by $C_{d,\gamma}|G(z)|\rho(z)^{-\gamma} \le C_{d,\gamma} A \mathbf{1}_{B_3}(z)$. For $y = z \pm h$ we have $\rho(y) \le$

$\rho(z) + |h| \le 3|z - y|$. The two translated terms therefore contribute in absolute value at most $C_{d,\gamma} A H(z)$, with

$$H(z) = \int_{B_3} \frac{\rho(y)^\gamma}{(\rho(y) + |z - y|)^{d+\gamma}}\, dy.$$

Fix $1 < q < \infty$, put $q' = q/(q-1)$, and take a nonnegative $b \in C_c^\infty(\mathbb{R}^d)$. For fixed y put $r = \rho(y)$. Integrating the kernel in H against $b(z)\, dz$ over $B(y, r)$ gives at most $C_d M b(y)$. On $2^j r < |z - y| \le 2^{j+1} r$, the kernel is at most $2^{-j(d+\gamma)} r^{-d}$ and the integral of b is at most $|B(y, 2^{j+1} r)| M b(y)$. Summing over $j \ge 0$ and integrating in y gives

$$\int_{\mathbb{R}^d} H(z) b(z)\, dz \le C_{d,\gamma} \int_{B_3} M b(y)\, dy \le C_{d,\gamma} |B_3|^{1/q} \|M b\|_{q'} \le C_{d,\gamma,q} \|b\|_{q'}$$

by Theorem 4.2. Duality bounds $\|H\|_q$ by the same constant. Together the two distance ranges give $||\nabla|^\gamma G| \le C_{d,\gamma} A(\mathbf{1}_{B_4} + H)$, as required. □

Theorem 4.4. *Let $2 \le p < \infty$, $k = \lceil p \rceil - 1$ and $\beta = p - k$. For $\mu, \nu \in \mathcal{P}(B_1)$ and $0 < \varepsilon < 1$, there exists a real $f \in C_c^\infty(\{x : |x| < 3\})$ satisfying*

$$W_p(\mu, \nu)^p - \varepsilon \le \int f\, d(\mu - \nu), \qquad \|f\|_{C^{k,\beta}} \le C_{d,p}.$$

The same f satisfies $\||\nabla|^p f\|_{L^q(\mathbb{R}^d)} \le C_{d,p,q}$ for every $1 < q < \infty$.

At $p = 2$, Bołbotowski and Bouchitté [BB26, Theorem 3.1] proved the related bound $W_2^2 \le 4 Z_2$ for measures with the same barycenter and finite second moments. The Zolotarev distance Z_2 is the supremum of $\int f\, d(\mu - \nu)$ over functions whose gradients are 1-Lipschitz. At the critical order $p = s$, Section 5 uses both the Hölder and Fourier derivative bounds of Theorem 4.4 to estimate this pairing for the chosen test function. Letting ε tend to zero then bounds $W_s(\mu, \nu)^s$.

Proof. We find the radii on which a constant can separate the dual potentials. On overlapping balls we choose constants whose difference is controlled by the p-th power of the radius. This controls the derivatives when we join the constants with a partition of unity.

Radii for separating the potentials. By Theorem 4.1 for the cost $|a - b|^p$ on B_1, choose a feasible pair u_0, v_0 with dual value at least $W_p(\mu, \nu)^p - \varepsilon/2$. A common translation gives $\min v_0 = 0$. Taking b at this minimum in the dual constraint gives $u_0(a) \le 2^p$. Put $u = \max\{u_0, 0\}$ and $v = \min\{v_0, 2^p\}$. Both take values in $[0, 2^p]$. At any pair (a, b) where a truncation acts, $u(a) \le v(b)$. Thus the pair is feasible and the truncations can only increase its dual value. With $w = v + \varepsilon/2$, we have

$$u(a) - w(b) \le |a - b|^p - \varepsilon/2, \qquad \int u\, d\mu - \int w\, d\nu \ge W_p(\mu, \nu)^p - \varepsilon. \tag{4.1}$$

For $z \in \mathbb{R}^d$ and $r > 0$, let $I(z, r)$ be the set of constants in $[0, 2^p]$ lying between u and w throughout $B_1 \cap B(z, r)$. It is nonempty exactly when $u(a) \le w(b)$ for all a, b in this intersection. Define

$$\rho(z) = \sup\{0 < r \le 1 : I(z, r) \ne \varnothing\}.$$

For $r_0 = (\varepsilon/2)^{1/p}/2$, any $a, b \in B_1 \cap B(z, r_0)$ satisfy $|a - b|^p \le \varepsilon/2$. By (4.1), $I(z, r_0)$ is nonempty and $r_0 \le \rho(z) \le 1$. If $I(z, r)$ is nonempty and $r > |z - z'|$, then

$B(z', r - |z - z'|) \subset B(z, r)$. Taking suprema over $r \le 1$ and interchanging centers shows that ρ is 1-Lipschitz.

Compatible constants on overlapping balls. Choose a maximal disjoint family of balls $B(z_i, \rho_i/40)$, where $\rho_i = \rho(z_i)$. Since $r_0 \le \rho_i \le 1$, only finitely many of their centers lie in any bounded set. For each z, maximality gives an i such that $B(z, \rho(z)/40)$ meets $B(z_i, \rho_i/40)$. Then

$$40|z - z_i| < \rho(z) + \rho_i \le |z - z_i| + 2\rho_i.$$

Thus the balls $B(z_i, \rho_i/16)$ cover $\mathbb{R}^d$. The larger balls $V_i = B(z_i, \rho_i/8)$ have bounded overlap. Indeed, at $z \in V_i$ we have $7\rho_i/8 \le \rho(z) \le 9\rho_i/8$. The disjoint balls of radii $\rho_i/40$ for these indices have volumes bounded below by $c_d\rho(z)^d$ and all lie in $B(z, 2\rho(z))$.

For each i, the interval $I_i = I(z_i, \rho_i/4)$ is closed and nonempty by the definition of ρ_i and the nesting of the intervals. When $\rho_i = 1$, choose any $c_i \in I_i$. When $\rho_i < 1$, a radius strictly between ρ_i and $\min\{2\rho_i, 1\}$ has empty interval. We can therefore choose $a_i, b_i \in B_1 \cap B(z_i, 2\rho_i)$ with $A_i = u(a_i) > w(b_i)$. Let c_i be the closest point of I_i to A_i. For $y \in B_1 \cap B(z_i, \rho_i/4)$, (4.1) gives

$$u(y) \le w(b_i) + (9\rho_i/4)^p \le A_i + (9\rho_i/4)^p, \qquad w(y) \ge A_i - (9\rho_i/4)^p.$$

Since $A_i \in [0, 2^p]$, these bounds give $|c_i - A_i| \le (9\rho_i/4)^p$.

If V_i meets V_j, their radii are comparable and $|z_i - z_j| \le (\rho_i + \rho_j)/8$. When both $\rho_i, \rho_j < 1$, we have $|a_i - b_j| \le C\rho_i$ and hence

$$A_i - A_j \le u(a_i) - w(b_j) \le |a_i - b_j|^p \le C_p\rho_i^p.$$

Interchanging i, j bounds the opposite difference. The estimates for $|c_i - A_i|$ and $|c_j - A_j|$ then give $|c_i - c_j| \le C_p\rho_i^p$. If either radius equals one, both are bounded below by $7/9$ and c_i, c_j lie in $[0, 2^p]$. Thus in all cases

$$|c_i - c_j| \le C_p\rho_i^p \qquad \text{whenever } V_i \cap V_j \ne \varnothing. \tag{4.2}$$

A smooth function between the potentials. Fix $0 \le \eta \in C_c^\infty(B(0, 1/8))$ equal to one on $B_{1/16}$. Set $\eta_i(z) = \eta((z - z_i)/\rho_i)$ and $\psi_i = \eta_i / \sum_j \eta_j$. The sum in the denominator is locally finite and at least one. At $z \in V_i$ only terms with $z \in V_j$ can contribute to its derivatives. There are at most C_d such terms, with ρ_j comparable to ρ_i. Differentiating the quotient gives $|\partial^\alpha \psi_i| \le C_{d,\alpha}\rho_i^{-|\alpha|}$.

The function $F = \sum_i c_i\psi_i$ takes values in $[0, 2^p]$. At any $z \in B_1$ with $\psi_i(z) > 0$, we have $z \in B(z_i, \rho_i/4)$, so $u(z) \le c_i \le w(z)$. The same inequalities hold for $F(z)$ since the weights sum to one. For $|\alpha| \ge 1$ and $z \in V_j$, we have

$$\partial^\alpha F(z) = \sum_{i:\, z \in V_i} (c_i - c_j)\partial^\alpha \psi_i(z).$$

By (4.2) and bounded overlap, this is bounded in absolute value by $C_{d,p,\alpha}\rho(z)^{p-|\alpha|}$. Choose a smooth cutoff $0 \le \chi \le 1$ equal to one on B_2 and compactly supported in $B(0, 3)$, and put $f = \chi F$. Where $|z| \ge 2$ we have $\rho(z) = 1$ because $B(z, 1)$ misses B_1. The product rule gives

$$\|f\|_\infty \le 2^p, \qquad |\partial^\alpha f(z)| \le C_{d,p,\alpha}\rho(z)^{p-|\alpha|} \quad (|\alpha| \ge 1). \tag{4.3}$$

On B_1 we have $u \le f \le w$, so (4.1) gives $\int f\, d(\mu - \nu) \ge W_p(\mu, \nu)^p - \varepsilon$.

Hölder and Fourier bounds. Estimate (4.3) bounds the derivatives through order k in supremum norm. Fix $|\alpha| = k$ and $t = |z - z'| > 0$. If $t < \rho(z)/2$, the radius is at

least $\rho(z)/2$ on the segment, so the order-$(k+1)$ estimate gives $|\partial^\alpha f(z)-\partial^\alpha f(z')|\le C_{d,p}t\rho(z)^{\beta-1}\le C_{d,p}t^\beta$. If $t\ge\rho(z)/2$, then $\rho(z)\le 2t$ and $\rho(z')\le 3t$. The order-k estimate at the two endpoints gives the same bound. This proves the $C^{k,\beta}$ estimate, including $\beta=1$.

For even p, the operator $|\nabla|^p=(-\Delta)^{p/2}$ is local. Its value on f is supported in B_3 and bounded by (4.3) with $|\alpha|=p$, which gives the L^q estimate. For the remaining orders, write $p=m+\gamma$ with m even and $0<\gamma<2$. Here $m\ge 2$. For $G=(-\Delta)^{m/2}f$, the orders m and $m+2$ in (4.3) give

$$|G(z)|\le C_{d,p}\rho(z)^\gamma,\qquad |D^2G(z)|\le C_{d,p}\rho(z)^{\gamma-2}.$$

Since $|\nabla|^p f=|\nabla|^\gamma G$, Lemma 4.3 applied to this fixed G gives the claimed bound for every $1<q<\infty$. □

5. The critical comparison

We use the Fourier convention of Section 2 and write $|\mathbb{S}^{d-1}|$ for the surface area of $\mathbb{S}^{d-1}$. To bound W_s^s by SW_2, we write a smooth test as an average of functions of $P_\theta x$ and expand each function at the projected source point. Differentiating the polar Fourier formula in the projected variable changes the radial weight from r^{d-1} to r^d.

Lemma 5.1. *Let $d\ge 2$, let u be a Schwartz function on $\mathbb{R}^d$, and put $g=|\nabla|^s u$. Then for every $x\in\mathbb{R}^d$,*

$$\left(\int_{\mathbb{S}^{d-1}}\left|\int_0^\infty e^{irP_\theta x}\widehat{u}(r\theta)r^d\,dr\right|^2 d\sigma(\theta)\right)^{1/2}\le C_d\big(\|g\|_{3/2}+\|g\|_3\big).$$

Proof. Replacing u by $u(\cdot+x)$ accounts for the phase and preserves both norms of g, so we may take $x=0$. Since $s+d/2-1=d$, the radial integrand is $\widehat{g}(r\theta)r^{d/2-1}$. In the squared modulus of the integral, write the two radii as tr and r. Polar coordinates give

$$\begin{aligned}&\int_{\mathbb{S}^{d-1}}\left|\int_0^\infty \widehat{g}(r\theta)r^{d/2-1}\,dr\right|^2 d\sigma(\theta)\\&\qquad=\frac{1}{|\mathbb{S}^{d-1}|}\int_0^\infty\left(\int_{\mathbb{R}^d} t^{d/2}\widehat{g}(t\xi)\overline{\widehat{g}(\xi)}\,d\xi\right)\frac{dt}{t}.\end{aligned}$$

The change of variables is justified by absolute convergence, since $\widehat{g}(\xi)=|\xi|^s\widehat{u}(\xi)$ is bounded near zero and decays rapidly at infinity. Its weak derivatives of orders $j\le d$ are integrable, being $O_u(|\xi|^{s-j})$ near zero. Fourier inversion then gives $g(z)=O_u((1+|z|)^{-d})$, so both Lebesgue norms are finite. Set $g_t(z)=t^{-d/2}g(z/t)$ for $t>0$. By Plancherel the inner integral is $(2\pi)^d\int_{\mathbb{R}^d}g_t(z)\overline{g(z)}\,dz$. We have $\|g_t\|_{3/2}=t^{d/6}\|g\|_{3/2}$ and $\|g_t\|_3=t^{-d/6}\|g\|_3$. Applying Hölder's inequality in both orders gives

$$\left|\int_{\mathbb{R}^d} g_t(z)\overline{g(z)}\,dz\right|\le\min\{t^{d/6},t^{-d/6}\}\,\|g\|_{3/2}\|g\|_3.$$

The minimum is integrable against dt/t. Taking square roots and using the arithmetic–geometric mean inequality proves the bound. □

The point x is fixed across directions in this estimate. We can therefore integrate its square against the source measure, even when that measure has atoms.

Proposition 5.2. *Put $k = \lceil s \rceil - 1$ and $\beta = s - k$. For $R > 0$, $\mu, \nu \in \mathcal{P}(B_1)$ and real $f \in C_c^\infty(\{x : |x| < R\})$,*

$$\left| \int f \, d(\mu - \nu) \right| \le C_{d,R} \, \mathrm{SW}_2(\mu, \nu) \left(\|f\|_{C^{k,\beta}} + \||\nabla|^s f\|_{3/2} + \||\nabla|^s f\|_3 \right).$$

Proof. We smooth the test at scale h and bound the first two derivatives of its one-dimensional profiles. In the Taylor expansion along a projected coupling, the first derivative is evaluated at the source and the second derivative is paired with the squared displacement.

Approximation to order h^s. Fix a real even $\kappa \in C_c^\infty(\mathbb{R})$ equal to one near zero. Let ρ be the inverse Fourier transform on $\mathbb{R}^d$ of $\xi \mapsto \kappa(|\xi|)$. For $0 < h \le 1$ set $\rho_h(z) = h^{-d}\rho(z/h)$ and $f_h = \rho_h * f$. The Schwartz function ρ has integral one, and all its moments of positive order vanish because its Fourier transform is constant near zero. Taylor expansion of $f(x - hz)$ to order k thus gives

$$\|f - f_h\|_\infty \le C_d h^s \|f\|_{C^{k,\beta}} \int_{\mathbb{R}^d} |z|^s |\rho(z)| \, dz. \tag{5.1}$$

First derivatives at a fixed source point. For $\theta \in \mathbb{S}^{d-1}$ and $t \in \mathbb{R}$, define

$$B_\theta(t) = \frac{|\mathbb{S}^{d-1}|}{(2\pi)^d} \int_0^\infty e^{irt} \kappa(hr) \widehat{f}(r\theta) r^{d-1} \, dr.$$

Fourier inversion gives $f_h(x) = \int_{\mathbb{S}}^{d-1} B_\theta(P_\theta x) \, d\sigma(\theta)$. Differentiation in t inserts the factor ir. We apply Lemma 5.1 to f_h and use $|\nabla|^s f_h = \rho_h * (|\nabla|^s f)$ and $\|\rho_h\|_1 = \|\rho\|_1$. With $g = |\nabla|^s f$ this yields

$$\sup_{x \in \mathbb{R}^d} \left(\int_{\mathbb{S}}^{d-1} |B'_\theta(P_\theta x)|^2 \, d\sigma(\theta) \right)^{1/2} \le C_d \big(\|g\|_{3/2} + \|g\|_3 \big). \tag{5.2}$$

Uniform control of the second derivatives. Let $U_\theta(t) = \int_{\theta^\perp} f(t\theta + z) \, dz$, with Lebesgue measure on $\theta^\perp$. Fubini gives $\widehat{U}_\theta(r) = \widehat{f}(r\theta)$. Differentiating k times under the integral yields $\|U_\theta^{(k)}\|_\infty \le C_{d,R} \|f\|_{C^{k,\beta}}$, since the integrand is supported where $|z| \le R$. Put $m = d + 1 - k = \lfloor s \rfloor$, so $2 \le m \le s$. The multiplier taking $U_\theta^{(k)}$ to B''_θ is a dimensional constant times $\mathbf{1}_{(0,\infty)}(r) r^m \kappa(hr)$, as $\widehat{U_\theta^{(k)}}(r) = (ir)^k \widehat{f}(r\theta)$. At $h = 1$ this compactly supported multiplier has an integrable second weak derivative, since its value and first derivative vanish at zero from the right. Two integrations by parts give an inverse transform bounded by $C_d(1 + |t|)^{-2}$. Dilation then bounds the L^1 norm of the kernel at scale h by $C_d h^{-m}$. Convolution with $U_\theta^{(k)}$ gives

$$\sup_{\theta, t} |B''_\theta(t)| \le C_{d,R} h^{-s} \|f\|_{C^{k,\beta}}. \tag{5.3}$$

The estimate along projected couplings. Write $S = \mathrm{SW}_2(\mu, \nu)$ and set $A(\theta) = (\int |B'_\theta(t)|^2 \, d\mu^\theta(t))^{1/2}$. For each coupling of μ^θ and ν^θ, expand $B_\theta(v) - B_\theta(t)$ at the source coordinate t. The remainder is bounded in modulus by $\frac{1}{2}\|B''_\theta\|_\infty |v - t|^2$. Cauchy–Schwarz bounds the integrated linear term by $A(\theta)$ times the square root of the coupling's quadratic cost. Taking the infimum over couplings at this fixed θ gives

$$\left| \int B_\theta \, d(\mu^\theta - \nu^\theta) \right| \le A(\theta) W_2(\mu^\theta, \nu^\theta) + \tfrac{1}{2} \|B''_\theta\|_\infty W_2(\mu^\theta, \nu^\theta)^2.$$

By the source marginal and Fubini,

$$\int_{\mathbb{S}}^{d-1} A(\theta)^2 \, d\sigma(\theta) = \int_{B_1} \int_{\mathbb{S}}^{d-1} |B'_\theta(P_\theta x)|^2 \, d\sigma(\theta) \, d\mu(x).$$

Thus angular Cauchy–Schwarz and (5.2) control the linear term by $C_d S(\|g\|_{3/2} + \|g\|_3)$. By (5.3) the integrated remainder is bounded by $C_{d,R} h^{-s} \|f\|_{C^{k,\beta}} S^2$. Integrating the profiles to recover f_h and applying (5.1) to both marginals gives

$$\left| \int f \, d(\mu - \nu) \right| \le C_{d,R} \left[(h^s + h^{-s} S^2) \|f\|_{C^{k,\beta}} + S\big(\|g\|_{3/2} + \|g\|_3\big) \right].$$

For $0 < S \le 1$, choose $h = S^{1/s}$ so both errors have order S. If $S = 0$, let h decrease to zero. For $S > 1$ the left-hand side is at most $2\|f\|_\infty \le 2S\|f\|_{C^{k,\beta}}$. □

Theorem 5.3. *For every $\mu, \nu \in \mathcal{P}(B_1)$,*

$$W_s(\mu, \nu)^s \le C_d \, \mathrm{SW}_2(\mu, \nu).$$

In dimension two this bounds the full quadratic cost by the root mean square of the projected Wasserstein distances.

Proof. Since $s \ge 2$, Theorem 4.4 applies at $p = s$. For each $0 < \varepsilon < 1$ its test f has support in $\{x : |x| < 3\}$. The Hölder bound and the two bounds with $q = 3/2$ and $q = 3$ hold for this same f, with constants depending only on d. Taking $R = 3$ in Proposition 5.2 gives

$$W_s(\mu, \nu)^s - \varepsilon \le \int f \, d(\mu - \nu) \le C_d \, \mathrm{SW}_2(\mu, \nu).$$

Letting ε decrease to zero proves the comparison. □

6. Oscillation in a two-dimensional plane

We write $x = (y, z) \in \mathbb{R}^2 \times \mathbb{R}^{d-2}$ and decompose each projection direction as $\theta = (\theta_\parallel, \theta_\perp)$ in the same coordinates. For $d = 2$ we give $\mathbb{R}^0 = \{0\}$ mass one. We construct two measures whose densities differ by an angular oscillation in y. The difference of their distribution functions after projection by P_θ is the signed mass of a half-space with normal θ.

Lemma 6.1. *Let $d \ge 2$, let $a \in C_c^\infty((1/16, 1/8))$ be nonnegative and nonzero, and let b be a smooth radial probability density supported in $\{z \in \mathbb{R}^{d-2} : |z| < 1/8\}$. Extend a by zero. For each positive integer n, set $g_n(y) = a(r)\cos(n\phi)$ for $y = r(\cos\phi, \sin\phi) \ne 0$, set $g_n(0) = 0$, and put*

$$G_n(x) = g_n(y) b(z), \qquad D_{n,\theta}(t) = \int_{\{P_\theta x \le t\}} G_n(x) \, dx \quad (\theta \in \mathbb{S}^{d-1}, \ t \in \mathbb{R}).$$

Then for every integer $K \ge 1$ there is a constant C_K depending only on d, K, a, b such that, for every integer $n \ge K + 2$ and $h = 1/n$,

$$|D_{n,\theta}(t)| \le C_K h \left(1 + \frac{|t| + |\theta_\perp|}{h}\right)^{-K} \qquad (\theta \in \mathbb{S}^{d-1}, \ t \in \mathbb{R}).$$

For fixed $c > 0$, a half-space carrying signed mass of absolute value at least ch must have its boundary within $O_c(h)$ of the origin and its normal within $O_c(h)$ of the plane.

Proof. We first estimate the half-plane integral by integration by parts. Its vanishing moments then give the decay in the transverse direction.

Cancellation in the plane. For $e \in \mathbb{S}^1$ let $H_e(t) = \int_{\{y\cdot e \le t\}} g_n(y)\,dy$. For every integer $L \ge 0$ we claim that

$$|H_e(t)| \le C_L h(1+|t|/h)^{-L}, \qquad \int_{\mathbb{R}} t^j H_e(t)\,dt = 0 \quad (j = 0,\dots,n-2). \tag{6.1}$$

Let $R(y_1,y_2) = (-y_2, y_1)$. Since $(Ry)\cdot\nabla = \partial_\phi$, the vector field $ha(r)(Ry)\sin(n\phi)$ has divergence g_n. Its flux across the line $y = te + wRe$ gives

$$H_e(t) = -h\int_{\mathbb{R}} w\, a\big(\sqrt{t^2+w^2}\big)\sin\big(n\phi(te+wRe)\big)\,dw.$$

The amplitude has bounded integral in absolute value, uniformly in t, so $|H_e(t)| \le Ch$. For $t \ne 0$, the phase on the line satisfies $(t^2+w^2)\partial_w \exp(in\phi) = int\,\exp(in\phi)$. Integrating by parts L times gains $(n|t|)^{-L}$ in absolute value. Each integration replaces the amplitude A by $\partial_w((t^2+w^2)A)$ up to a constant factor. These amplitudes are smooth on a fixed compact annulus in (t,w), so their $L^1(dw)$ norms are bounded uniformly in t. This proves the first bound in (6.1).

The angular integral of g_n vanishes, so H_e has compact support. Its derivative is $\int_{\mathbb{R}} g_n(te+wRe)\,dw$. Integration by parts gives

$$\int_{\mathbb{R}} t^j H_e(t)\,dt = -\frac{1}{j+1}\int_{\mathbb{R}^2} (e\cdot y)^{j+1} g_n(y)\,dy.$$

At fixed r, the power $(e\cdot y)^{j+1}$ is a trigonometric polynomial in ϕ of degree at most $j+1$. For $j+1 < n$ its integral against $\cos(n\phi)$ is zero, proving the moment identities.

Decay away from the plane. Put $u = |\theta_\parallel|$ and $v = |\theta_\perp|$. If $u = 0$, integration in y gives $D_{n,\theta} = 0$. If $v = 0$, then $D_{n,\theta} = H_{\theta_\parallel}$ and (6.1) gives the claim. For $uv > 0$ put $e = \theta_\parallel/u$ and $J(t) = H_e(t/u)$. The density κ of the projection of $b(z)\,dz$ along $\theta_\perp/v$ is independent of this unit direction by radiality. Integration of b over the perpendicular hyperplanes gives a smooth density supported in $[-1/8, 1/8]$. In transverse dimension one it is b itself. Integrating first in y now gives

$$D_{n,\theta} = J * \kappa_v, \qquad \kappa_v(t) = v^{-1}\kappa(t/v).$$

In particular $|D_{n,\theta}| \le Ch$, since κ_v has mass one.

Expand $\kappa_v(t-w)$ in w at zero through degree $K-1$. By dilation, the moment identities in (6.1) cancel every polynomial term against $J(w)\,dw$. The Taylor remainder and the bound in (6.1) with $L = K+2$ give

$$|D_{n,\theta}(t)| \le C_K v^{-K-1}\int_{\mathbb{R}} |w|^K |J(w)|\,dw \le C_K h\left(\frac{uh}{v}\right)^{K+1}.$$

Here $\|\kappa_v^{(K)}\|_\infty = v^{-K-1}\|\kappa^{(K)}\|_\infty$, and the last integral is estimated by scaling w by uh. Together with $|D_{n,\theta}| \le Ch$ and $u \le 1$, this gives $|D_{n,\theta}| \le C_K h(1+v/h)^{-K}$. This suffices when $|t| < v/4$. When $|t| \ge v/4$, every w in the support of κ_v satisfies $|t-w| \ge |t|/2$. The convolution and (6.1) then give $|D_{n,\theta}(t)| \le C_K h(1+|t|/h)^{-K}$. In this range $|t| + v \le 5|t|$, which proves the asserted joint bound. □

Theorem 6.2. *Let $1 < p < \infty$. For all sufficiently large integers n there exist absolutely continuous $\mu_n, \nu_n \in \mathcal{P}(B_1)$ such that*

$$W_1(\mu_n,\nu_n) \ge c_{d,p}\, n^{-1}, \qquad \mathrm{SW}_p(\mu_n,\nu_n) \le C_{d,p}\, n^{-2-(d-2)/p}.$$

The projected W_p is $O(n^{-2})$ in every direction. Its decay away from the plane gives the additional factor $n^{-(d-2)/p}$ upon averaging.

Proof. We add the angular oscillation to a common radial density and test it with a Lipschitz function. On each projected line we use the common density to bound the cost of a quantile coupling, then average this bound over the sphere.

Construction and ambient lower bound. Fix a, b as in Lemma 6.1, depending only on d, and use its functions $g_n, G_n, D_{n,\theta}$. Put $h = 1/n \le 1/8$. The functions G_n are uniformly bounded, are supported in $B_{1/4}$, and have integral zero by angular integration. Let $U = \mathbf{1}_{B_1}/|B_1|$ be the uniform probability density on the ball. With Z_h the normalizing integral, define the probability density

$$Q_h(x) = \frac{h\,\mathbf{1}_{B_1}(x)}{Z_h(h+|x|)^{d+1}}.$$

Scaling x by h gives $Z_h \le \int_{\mathbb{R}^d}(1+|x|)^{-d-1}\,dx$. Fix $\delta > 0$ small enough that $\delta|G_n| \le U/4$ for all n. Take μ_n, ν_n with respective densities $(U+Q_h)/2 + \delta G_n$ and $(U+Q_h)/2 - \delta G_n$. Each integrates to one and dominates $U/4 + Q_h/2$ on the ball.

The function $f_n(y,z) = hg_n(y)$ has Lipschitz constant at most $h\|a'\|_\infty + \sup_{r>0}|a(r)|/r$, by differentiating in the radial and unit angular directions. Since $\int_0^{2\pi}\cos^2(n\phi)\,d\phi = \pi$,

$$\int f_n\,d(\mu_n - \nu_n) = 2\delta\pi h\int_0^\infty a(r)^2 r\,dr.$$

For every coupling the left-hand side is bounded by the Lipschitz constant times the mean displacement. Taking the infimum gives $W_1(\mu_n,\nu_n) \ge c_d h$.

Transport on a projected line. Fix $\theta \in \mathbb{S}^{d-1}$, write $v = |\theta_\perp|$, and let q_h^θ be the density of the projection of $Q_h(x)\,dx$. For $|t| \le 1/4$, integrate on $t\theta + \theta^\perp$ over the disk $|w| \le h + |t|$ in $\theta^\perp$. There $t\theta + w \in B_1$ and $h + |t\theta + w| \le 2(h+|t|)$. The disk has $(d-1)$-dimensional volume proportional to $(h+|t|)^{d-1}$, so the upper bound on Z_h gives

$$q_h^\theta(t) \ge c_d\,\frac{h}{(h+|t|)^2} \qquad (|t| \le 1/4). \tag{6.2}$$

Let F_0, F_1 be the distribution functions of $\mu_n^\theta, \nu_n^\theta$, respectively. Their densities are positive on $(-1,1)$ by the projected $U/4$ lower bound. They are continuous there by dominated convergence in the hyperplane integral, since the boundary of each ball cross-section has measure zero in $\theta^\perp$. For $0 \le \ell \le 1$ put $F_\ell = (1-\ell)F_0 + \ell F_1$ and $\rho_\ell = F_\ell'$. Each ρ_ℓ dominates $q_h^\theta/2$. For $0 < \gamma < 1$, the quantile $X_\ell(\gamma) = F_\ell^{-1}(\gamma)$ satisfies

$$\partial_\ell X_\ell(\gamma) = \frac{(F_0 - F_1)(X_\ell(\gamma))}{\rho_\ell(X_\ell(\gamma))}$$

by implicit differentiation. Taking γ uniform on $(0,1)$ couples $X_0(\gamma)$ to $X_1(\gamma)$. Jensen's inequality in ℓ and the change of variables $\gamma = F_\ell(t)$ give

$$W_p(\mu_n^\theta,\nu_n^\theta)^p \le \int_0^1\int_{-1}^1 |F_0(t) - F_1(t)|^p \rho_\ell(t)^{1-p}\,dt\,d\ell.$$

The difference $F_0 - F_1 = 2\delta D_{n,\theta}$ vanishes for $|t| > 1/4$. Put $m = d-2$, take $K = \lceil 2 + m/p\rceil$, and suppose $n \ge K+2$. Lemma 6.1 controls the numerator, and

(6.2) bounds the density from below. We obtain

$$W_p(\mu_n^\theta, \nu_n^\theta)^p \le C_{d,p} h^{2p-1} \int_{-1/4}^{1/4} \frac{(1+|t|/h)^{2p-2}}{(1+(|t|+v)/h)^{Kp}}\, dt$$
$$\le C_{d,p} h^{2p} (1+v/h)^{-m-1}. \tag{6.3}$$

For the last inequality the integrand is bounded by $(1+(v+|t|)/h)^{-m-2}$, whose integral over $\mathbb{R}$ is $2h(1+v/h)^{-m-1}/(m+1)$.

The angular average. For $m > 0$, the transverse component of θ under σ is uniform on the unit ball of $\mathbb{R}^m$. To see this, parametrize the sphere up to a set of surface measure zero by

$$\theta = \big(\sqrt{1-|z|^2}(\cos\psi, \sin\psi), z\big), \qquad |z| < 1, \quad 0 \le \psi < 2\pi.$$

The ψ derivative has squared length $1-|z|^2$ and is orthogonal to the z derivatives. The Gram matrix of the z derivatives is $I + zz^\top/(1-|z|^2)$ with determinant $(1-|z|^2)^{-1}$, so the surface element is $d\psi\, dz$. It follows by scaling z that

$$\int_{\mathbb{S}^{d-1}} (1+|\theta_\perp|/h)^{-m-1}\, d\sigma(\theta) \le C_d h^m \int_{\mathbb{R}^m} (1+|z|)^{-m-1}\, dz \le C_d' h^m.$$

For $m = 0$ we have $v = 0$ in every direction, and the angular integral is $1 = h^m$. Thus the line estimate (6.3) and the angular integral give $\mathrm{SW}_p(\mu_n, \nu_n)^p \le C_{d,p} h^{2p+m}$. Taking the p-th root and using the ambient lower bound $c_d h$ with $h = 1/n$ proves the theorem. □

7. An example with small displacements in every projection

The mass of the half-space $\{P_\theta x \le t\}$ is the value of the projected distribution function at t. We construct a perturbation whose half-space integrals concentrate on the scale $|t| \lesssim h$, uniformly in θ. A common radial density supplies enough mass at these offsets to keep the projected quantiles within distance $O(h^s)$.

Lemma 7.1. *Let $d \ge 2$. There are constants $c_d, C_d, h_d > 0$ such that, for every $0 < h < h_d$, there is a real $g_h \in C_c^\infty(\mathbb{R}^d)$ supported in $B_{1/8}$ with the following properties.*

(i) The integral of g_h is zero and

$$\int_{\mathbb{R}^d} g_h^2\, dx \ge c_d, \qquad \|g_h\|_\infty + h\|\nabla g_h\|_\infty \le C_d.$$

(ii) For every $\theta \in \mathbb{S}^{d-1}$ and $t \in \mathbb{R}$,

$$\left| \int_{\{P_\theta x \le t\}} g_h(x)\, dx \right| \le C_d h^{d/2} (1+|t|/h)^{-2}.$$

Thus hg_h has Lipschitz constant at most C_d, and its integral against $g_h\, dx$ is at least $c_d h$.

Proof. We prescribe a Fourier phase constant along rays and fix the L^2 mass of the resulting function. Angular oscillation bounds its amplitude in space, and restriction of the Fourier transform to each ray controls its half-space integrals. Spatial decay then allows us to cut off the function.

Fixed L^2 mass. Fix nonzero real bumps $a_0 \in C_c^\infty((1,2))$ and $a_j \in C_c^\infty((-1,1))$ for $1 \le j < d$. For $\eta_d > 0$, put $r = \eta_d$ and $v_j = \eta_j / r$. For a parameter $\kappa > 0$ set

$$a(\eta) = a_0(r) \prod_{j=1}^{d-1} a_j(v_j), \qquad \psi(\eta) = \kappa \left(v_1 + \sum_{j=2}^{d-1} v_j^2 \right).$$

The support of a is compact in $\{\eta_d > 0\}$, so we may fix $\kappa > 0$ small enough that $|\nabla \psi| \le 1/32$ on a neighborhood of this support. We extend products with a by zero to $\{\eta_d \le 0\}$. For $0 < h < 1/16$ put

$$U_h(x) = (2\pi)^{-d} h^{-d/2} \int_{\mathbb{R}^d} a(\eta) e^{i(x\cdot\eta + \psi(\eta))/h} \, d\eta, \qquad f_h = 2 \operatorname{Re} U_h.$$

Fourier inversion gives $\widehat{U}_h(\xi) = h^{d/2} a(h\xi) e^{i\psi(h\xi)/h}$. This smooth compactly supported transform and its reflected conjugate have disjoint supports away from zero. Thus f_h is a real Schwartz function of integral zero. By Plancherel and the change of variables $\eta = h\xi$,

$$\int_{\mathbb{R}^d} f_h^2 \, dx = 2(2\pi)^{-d} \int_{\mathbb{R}^d} |a(\eta)|^2 \, d\eta =: L > 0.$$

Uniform size and gradient bounds. In the coordinates (r, v) the Jacobian is r^{d-1} and the phase $x \cdot \eta + \psi(\eta)$ is

$$r x_d + (r x_1 + \kappa) v_1 + \sum_{j=2}^{d-1} (\kappa v_j^2 + r x_j v_j).$$

For the v_1 integral, the L^1 bound for a_1 and two integrations by parts give the majorant $C(1 + |r x_1 + \kappa|/h)^{-2}$. Its integral over $1 \le r \le 2$ is $O(h)$ uniformly in x_1. Indeed, $|r x_1 + \kappa| \ge \kappa/2$ when $|x_1| \le \kappa/4$. When $|x_1| > \kappa/4$, the change of variable to $(r x_1 + \kappa)/h$ bounds the integral by $Ch/|x_1|$.

For $2 \le j < d$, the derivative of the quadratic phase in v_j is $2\kappa(v_j - v_*)$, where $v_* = -r x_j / (2\kappa)$. The portion $|v_j - v_*| \le \sqrt{h}$ contributes at most $2\sqrt{h} \|a_j\|_\infty$. On the complement, integration by parts bounds the boundary terms and the remaining integral together by

$$C_\kappa h \left(\frac{\|a_j\|_\infty + \|a_j'\|_1}{\sqrt{h}} + \|a_j\|_\infty \int_{|w| > \sqrt{h}} \frac{dw}{w^2} \right) \lesssim \sqrt{h}.$$

The weight $r^{d-1} a_0(r)$ is bounded. The factor h from integration in r and the $d-2$ factors $\sqrt{h}$ cancel the prefactor $h^{-d/2}$ in U_h, giving $\|U_h\|_\infty \lesssim 1$. For $d = 2$ only the integrations in v_1 and r occur. Differentiation by $h\partial_{x_\ell}$ inserts the factor $i\eta_\ell$. This multiplies the r weight by ir and, for $\ell < d$, the bump $a_\ell(v_\ell)$ by v_ℓ. The same bounds give

$$\|f_h\|_\infty + h \|\nabla f_h\|_\infty \lesssim 1.$$

Half-space integrals. For an integrable function b write $F_b^\theta(t) = \int_{\{P_\theta x \le t\}} b(x) \, dx$. By Fubini, the Fourier transform of $t \mapsto \int_{\theta^\perp} U_h(t\theta + y) \, dy$ is $\tau \mapsto \widehat{U}_h(\tau\theta)$. This transform vanishes near $\tau = 0$. The inverse transform of $\widehat{U}_h(\tau\theta)/(i\tau)$ has the projected density as derivative and tends to zero at $-\infty$, so it equals $F_{U_h}^\theta$. Setting $z = h\tau$ gives

$$F_{U_h}^\theta(t) = \frac{h^{d/2}}{2\pi} \int_{\mathbb{R}} e^{izt/h} e^{i\psi(z\theta)/h} \frac{a(z\theta)}{iz} \, dz.$$

On each half-line meeting the support of $a(z\theta)$, the ratios η_j/η_d at $\eta = z\theta$ are constant in z. Thus the phase factor $e^{i\psi(z\theta)/h}$ is constant there. The amplitude $a(z\theta)/z$ is supported where $1 \le |z| \le 2\sqrt{d}$, and its derivatives through order two are uniformly bounded since $|\theta| = 1$. For $|t| \le h$ we bound the integral absolutely. For $|t| > h$ we integrate twice by parts in z, gaining $(h/|t|)^2$. Taking twice the real part gives

$$|F^\theta_{f_h}(t)| \lesssim h^{d/2}(1 + |t|/h)^{-2}.$$

Localization in $B_{1/8}$. For $|x| \ge 1/16$ the phase $\Phi_x(\eta) = x \cdot \eta + \psi(\eta)$ satisfies $|\nabla_\eta \Phi_x| \ge |x| - 1/32 \gtrsim 1 + |x|$ near the support of a. Put $V_x = \nabla_\eta \Phi_x / |\nabla_\eta \Phi_x|^2$. The operator $(h/i)V_x \cdot \nabla_\eta$ fixes $e^{i\Phi_x/h}$. The derivatives of ψ are bounded on this support, so V_x and each fixed order of its η-derivatives are $O((1 + |x|)^{-1})$ there. Integrating by parts $d + 2$ times gives

$$|f_h(x)| \lesssim h^{d/2+2}(1 + |x|)^{-d-2} \qquad (|x| \ge 1/16).$$

Choose a smooth cutoff $0 \le \chi \le 1$ supported in $B_{1/8}$ and equal to one on $B_{1/16}$. Choose a smooth probability density b_0 supported in $B_{1/8}$. Put

$$m_h = \int_{\mathbb{R}^d} \chi f_h \, dx, \qquad g_h = \chi f_h - m_h b_0.$$

Then $\int g_h \, dx = 0$, and $\int f_h \, dx = 0$ gives $|m_h| \le \|(1 - \chi) f_h\|_1$. The tail estimate yields

$$\|g_h - f_h\|_1 \le 2\|(1 - \chi) f_h\|_1 \lesssim h^{d/2+2}.$$

The product rule gives $\|g_h\|_\infty + h\|\nabla g_h\|_\infty \lesssim 1$. Consequently

$$\left| \int_{\mathbb{R}^d} g_h^2 \, dx - L \right| \le (\|g_h\|_\infty + \|f_h\|_\infty)\|g_h - f_h\|_1 \lesssim h^{d/2+2},$$

which proves (i) for small h. For $|t| \le 1/8$ the half-space integrals change by at most $\|g_h - f_h\|_1$. Since $h + |t| \le 1$ on this interval, $h^{d/2+2} \le h^{d/2}(1 + |t|/h)^{-2}$, so the bound for $F^\theta_{f_h}$ also gives (ii) there. For $|t| \ge 1/8$ the support and zero integral of g_h give $F^\theta_{g_h}(t) = 0$. □

Theorem 7.2. *For all sufficiently small $h > 0$ there are $\mu_h, \nu_h \in \mathcal{P}(B_1)$ with smooth densities such that $W_1(\mu_h, \nu_h) \ge c_d h$. In each direction θ the marginals $\mu^\theta_h, \nu^\theta_h$ admit a coupling supported on $\{(t, v) : |t - v| \le C_d h^s\}$.*

For the same pair the coupling bound gives $\mathrm{SW}_p(\mu_h, \nu_h) \le C_d h^s$ simultaneously for all $1 \le p < \infty$.

Proof. We add opposite copies of the perturbation from Lemma 7.1 to a common radial background. We test the ambient cost against the perturbation and bound the projected displacements by differentiating the quantiles.

Mass available in each projection. Take h in the range of the lemma with $h < 1/16$. Fix a smooth radial probability density b supported in B_1 and positive in its interior. Let $0 \le \zeta \le 1$ be a smooth radial cutoff equal to one on $B_{1/2}$ and supported in $B_{3/4}$. Choose Z_h so that

$$k_h(x) = \frac{\zeta(x)}{Z_h} \frac{h}{(h^2 + |x|^2)^{(d+1)/2}}$$

has integral one. After $x = hy$, Z_h is the integral of $\zeta(hy)(1 + |y|^2)^{-(d+1)/2}$. It lies between the integrals of $(1 + |y|^2)^{-(d+1)/2}$ over B_1 and over $\mathbb{R}^d$, both finite

and positive. For $|t| \le 1/8$, the disk $\{t\theta + y : y \in \theta^\perp,\ |y| \le h + |t|\}$ lies in $B_{1/2}$, and $h^2 + |t\theta + y|^2 \le 2(h + |t|)^2$ on the disk. Its $(d-1)$-dimensional volume is $c_d(h + |t|)^{d-1}$, so the projected density satisfies

$$K_h^\theta(t) := \int_{\theta^\perp} k_h(t\theta + y)\, dy \ge c_d \frac{h}{(h + |t|)^2} = c_d h^{-1}(1 + |t|/h)^{-2}. \tag{7.1}$$

A Lipschitz test for the ambient cost. Let g_h be given by Lemma 7.1. Its uniform supremum bound allows us to fix $\varepsilon > 0$ with $\varepsilon \|g_h\|_\infty \le \frac{1}{4} \min_{B_{1/8}} b$ for all these h. Define

$$d\mu_h = \left(\frac{b + k_h}{2} + \varepsilon g_h\right) dx, \qquad d\nu_h = \left(\frac{b + k_h}{2} - \varepsilon g_h\right) dx.$$

Both densities dominate $b/4 + k_h/2$ since g_h is supported in $B_{1/8}$. They are smooth, are supported in B_1, and have integral one because $\int g_h\, dx = 0$. For every $\pi \in \Pi(\mu_h, \nu_h)$,

$$2\varepsilon \int_{\mathbb{R}^d} g_h^2\, dx = \int (g_h(x) - g_h(y))\, d\pi(x, y) \le \|\nabla g_h\|_\infty \int |x - y|\, d\pi(x, y).$$

Part (i) of the lemma and the infimum over π give $W_1(\mu_h, \nu_h) \ge c_d h$.

Displacement of the projected quantiles. Fix $\theta \in \mathbb{S}^{d-1}$ and let H_0, H_1 be the distribution functions of $\mu_h^\theta, \nu_h^\theta$, respectively. For $0 \le \tau \le 1$ put $H_\tau = (1-\tau)H_0 + \tau H_1$ and $\rho_\tau = \partial_t H_\tau$. Integrating the smooth compactly supported densities on the hyperplanes $\{P_\theta x = t\}$ gives smooth projected densities. Each is positive for $-1 < t < 1$ because $b(t\theta) > 0$ and the ambient densities dominate $b/4$. Thus $\rho_\tau > 0$ on $(-1, 1)$, and H_τ goes from zero to one on $[-1, 1]$. For every $0 < u < 1$ there is a unique $Q_\tau(u) \in (-1, 1)$ with $H_\tau(Q_\tau(u)) = u$. Implicit differentiation gives

$$\partial_\tau Q_\tau(u) = \frac{(H_0 - H_1)(Q_\tau(u))}{\rho_\tau(Q_\tau(u))}.$$

The difference $H_0(t) - H_1(t)$ is 2ε times the half-space integral of g_h and vanishes for $|t| \ge 1/8$. For $|t| \le 1/8$, part (ii) of the lemma bounds its modulus by $C_d h^{d/2}(1 + |t|/h)^{-2}$. The density satisfies $\rho_\tau(t) \ge K_h^\theta(t)/2$. By (7.1) the factors $(1 + |t|/h)^{-2}$ cancel in the quotient, giving $|\partial_\tau Q_\tau(u)| \le C_d h^{1+d/2} = C_d h^s$. Integrating over τ gives this bound for $|Q_1(u) - Q_0(u)|$. For U uniform on $(0, 1)$ the law of $(Q_0(U), Q_1(U))$ is therefore a coupling of $\mu_h^\theta, \nu_h^\theta$ supported on the asserted strip. □

8. The sharp exponent

We work on B_1 with probability surface measure on the directions. For $p > s$, the sharp restriction on $\alpha_*(d, p)$ comes from moving a small atom a fixed distance.

Proposition 8.1. *Let $1 < p < \infty$. Regard σ as a probability on $\mathbb{R}^d$ supported on $\mathbb{S}^{d-1}$, and put $y = e_1/4$. For $0 < t \le 1/2$ set $\mu_t = (1-t)\sigma + t\delta_0$ and $\nu_t = (1-t)\sigma + t\delta_y$. Then*

$$W_p(\mu_t, \nu_t) = t^{1/p}/4, \qquad c_{d,p}\, t \le \mathrm{SW}_p(\mu_t, \nu_t) \le C_d t.$$

In each projection the common surface measure fills the interval between the two atom locations. This lets us couple the projected laws with every displacement bounded by $C_d t$.

Proof. We compute the ambient cost from the target support. For the slices we construct a coupling for the upper bound and use the projected means for the lower bound.

Ambient cost. Every target point is at distance at least $1/4$ from the origin, so every coupling incurs cost at least $t4^{-p}$ for the mass at 0. Fixing the common sphere mass and sending 0 to y attains this cost.

A displacement bound in each direction. Reflection preserves the cost and interchanges the directions θ and $-\theta$, so we may take $b = P_\theta y \geq 0$. By rotation invariance the density of σ^θ on $(-1, 1)$ is

$$f_d(z) = k_d(1 - z^2)^{(d-3)/2},$$

with k_d chosen to give integral one. Indeed, in the parametrization $(z, \sqrt{1 - z^2}\,\omega)$, $\omega \in \mathbb{S}^{d-2}$, the area factor in the angular variables is $(1 - z^2)^{(d-2)/2}$ and the length factor in z is $(1 - z^2)^{-1/2}$. For $d = 2$ the two values of ω give the two semicircles. In particular $m_d = \min_{[0,1/4]} f_d > 0$.

Put $a = (1 - t)m_d$. Since $b \leq 1/4$, the common projected measure $(1 - t)\sigma^\theta$ dominates $\gamma = a\mathbf{1}_{[0,b]}(z)\,dz$. Couple $(1-t)\sigma^\theta - \gamma$ to itself by the identity. To couple the rest, put $\ell = t/a$ and let $T(u) = \min\{b, \max\{0, u\}\}$. Push forward $a\,du$ on $[-\ell, b]$ by $u \mapsto (T(u), T(u + \ell))$. The arguments of T range over $[-\ell, b]$ and $[0, b + \ell]$, respectively. In both marginals T leaves density a on $[0, b]$. The excess interval of length ℓ gives mass $a\ell = t$ at 0 in the first marginal and at b in the second. Since T is 1-Lipschitz, every displacement in the resulting coupling is at most $\ell \leq 2t/m_d$.

Two-sided sliced bound. For any coupling of the projected laws, the L^p norm of the displacement is at least the absolute difference $t|P_\theta y|$ of their means. Taking the angular L^p norm of this lower bound and the uniform upper bound gives

$$t\left(\int_{\mathbb{S}^{d-1}} |P_\theta y|^p \, d\sigma(\theta)\right)^{1/p} \leq \mathrm{SW}_p(\mu_t, \nu_t) \leq 2t/m_d.$$

The integral is positive since $|P_\theta y|$ is positive in a neighborhood of e_1. □

Theorem 8.2. *For every $d \geq 2$ and $1 < p < \infty$, the admissible exponents form the interval $(0, \alpha_*(d, p)]$, where*

$$\alpha_*(d, p) = \begin{cases} p/(2p + d - 2), & 1 < p < 2, \\ 2/(d + 2), & 2 \leq p \leq s, \\ 1/p, & p > s. \end{cases}$$

In dimension two the critical order is 2, so the formula reads $\alpha_*(2, p) = 1/\max\{2, p\}$.

Proof of Theorem 8.2. We match an endpoint estimate with an example in each range of p, then use bounded support to pass to smaller powers. Fix $\mu, \nu \in \mathcal{P}(B_1)$ for the estimates and let α be any admissible exponent as in Definition 2.1. For $1 \leq q \leq r < \infty$, Hölder's inequality on each probability coupling gives $W_q(\mu, \nu) \leq W_r(\mu, \nu)$ after taking infima. The same comparison on each line followed by Hölder's inequality on the probability space $(\mathbb{S}^{d-1}, \sigma)$ gives $\mathrm{SW}_q(\mu, \nu) \leq \mathrm{SW}_r(\mu, \nu)$.

The endpoint below order two. For $1 < p < 2$, Theorem 3.3 gives admissibility of $p/(2p+d-2)$. For the measures of Theorem 6.2, order monotonicity and admissibility of α give

$$n^{-1} \lesssim W_1(\mu_n, \nu_n) \leq W_p(\mu_n, \nu_n) \lesssim \mathrm{SW}_p(\mu_n, \nu_n)^\alpha \lesssim n^{-\alpha(2+(d-2)/p)}.$$

Letting $n \to \infty$ forces $\alpha \leq p/(2p + d - 2)$.

The critical power for $2 \le p \le s$. Order monotonicity and Theorem 5.3 give

$$W_p(\mu,\nu)^s \le W_s(\mu,\nu)^s \le C_d\ \mathrm{SW}_2(\mu,\nu) \le C_d\ \mathrm{SW}_p(\mu,\nu).$$

Thus $1/s = 2/(d+2)$ is admissible. In each direction the coupling provided by Theorem 7.2 has displacements bounded by $C_d h^s$, so $\mathrm{SW}_p(\mu_h,\nu_h) \le C_d h^s$. The ambient lower bound from the same theorem gives $W_p(\mu_h,\nu_h) \ge W_1(\mu_h,\nu_h) \ge c_d h$. Admissibility forces $h \lesssim h^{s\alpha}$ for all sufficiently small $h > 0$, hence $\alpha \le 1/s$.

The power $1/p$ *above the critical order.* For $p > s$, the diameter bound on B_1 gives $|x-y|^p \le 2^{p-s}|x-y|^s$ for every pair of support points. Taking infima over couplings and using Theorem 5.3 gives

$$W_p(\mu,\nu)^p \le 2^{p-s} W_s(\mu,\nu)^s \le 2^{p-s} C_d\ \mathrm{SW}_p(\mu,\nu).$$

This proves admissibility of $1/p$. Conversely, Proposition 8.1 gives $t^{1/p}/4 \le C_{d,p,\alpha}(C_d t)^\alpha$ for $0 < t \le 1/2$. Letting $t \downarrow 0$ forces $\alpha \le 1/p$.

The planar oscillations below 2, the directional couplings on $[2, s]$, and the moving atom above s therefore give matching upper bounds in the three ranges. Let a be the endpoint obtained for the given (d, p). All projections are supported on $[-1, 1]$, so $\mathrm{SW}_p(\mu,\nu) \le 2$. For every $0 < \eta < a$ we have

$$\mathrm{SW}_p(\mu,\nu)^a \le 2^{a-\eta}\,\mathrm{SW}_p(\mu,\nu)^\eta.$$

The estimate at a thus gives the full interval $(0, a]$. □

References

[BG] S. G. Bobkov and F. Götze, *Quantified Cramér–Wold continuity theorem for the Kantorovitch transport distance*, Ann. Fac. Sci. Toulouse Math. (6), to appear.

[BB26] K. Bołbotowski and G. Bouchitté, *Sharp inequalities between Zolotarev and Wasserstein distances in* $\mathcal{P}_2(\mathbb{R}^d)$, Probab. Theory Related Fields **196** (2026), 3139–3155.

[Bon13] N. Bonnotte, *Unidimensional and evolution methods for optimal transportation*, doctoral thesis, Université Paris-Sud and Scuola Normale Superiore, Pisa, 2013.

[CFMW25] G. Carlier, A. Figalli, Q. Mérigot and Y. Wang, *Sharp comparisons between sliced and standard* 1*-Wasserstein distances*, preprint, arXiv:2510.16465v1, 2025.

[FG15] N. Fournier and A. Guillin, *On the rate of convergence in Wasserstein distance of the empirical measure*, Probab. Theory Related Fields **162** (2015), 707–738.

[Hop26] J. S. Hopper, *Sliced Wasserstein geodesics and equivalence of metrics*, preprint, arXiv:2407.07219v2, 2026.

[KT26] J. Kitagawa and A. Takatsu, *Sliced optimal transport: Is it a suitable replacement?*, Indiana Univ. Math. J. **75** (2026), no. 2, 487–524.

[LM26] C. Letrouit and Q. Mérigot, *Embedding large subsets of Wasserstein spaces into Banach spaces*, preprint, arXiv:2610.12029v1, 2026.

[NS26] J. Niles-Weed and J. Shkrob, *Nearly sharp comparison results for sliced and max-sliced Wasserstein distances*, preprint, arXiv:2608.13374v1, 2026.

[PS25] S. Park and D. Slepčev, *Geometry and analytic properties of the sliced Wasserstein space*, J. Funct. Anal. **289** (2025), no. 7, article 110975.

[RPDB12] J. Rabin, G. Peyré, J. Delon and M. Bernot, *Wasserstein barycenter and its application to texture mixing*, in *Scale Space and Variational Methods in Computer Vision*, Lecture Notes in Comput. Sci., vol. 6667, Springer, Berlin, 2012, 435–446.

[Ste70] E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Mathematical Series, vol. 30, Princeton University Press, Princeton, NJ, 1970.

[Vil03] C. Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, Providence, RI, 2003.