

Sharp transport identification with unknown Gaussian covariance

Scott Schubr¹ 

¹Luminary Oddities, Illinois, USA, e-mail: scottschubr@luminaryoddities.com

Abstract: We determine how covariance information controls transport identification under unknown bounded anisotropic Gaussian noise. On a fixed two-sided potential-curvature class, the sharp quadratic Wasserstein modulus is the sum of the latent-covariance uncertainty and the inherited inverse logarithm of order $3/2$. With n noisy observations and m independent pure-noise calibration observations, we prove the sharp minimax rate obtained by replacing that uncertainty with its minimum with $m^{-1/2}$. This identifies the calibration sample size needed to retain the known-noise rate. A Gaussian-divisible fixed reference permits rough log-concave inputs and has an explicit exact compatible-law set. Paired independent noisy repetitions supply a complementary calibration mechanism: they remove the covariance prior and yield sharp full-density estimation throughout the smooth class, together with fixed-reference testing against rough log-concave alternatives. Positive Gaussian completion and covariance alignment transfer an established scalar-noise inverse theorem to these experiments. Gaussian-allocation and moment-corrected oscillatory alternatives give the matching lower bounds.

MSC2020 subject classifications: Primary 62G07, 62G20; secondary 62C20, 60E15.

Keywords and phrases: Gaussian deconvolution, partial identification, covariance calibration, Wasserstein distance, minimax estimation.

1. Models and main conclusions

For a symmetric nonnegative matrix A , let γ_A be the centered Gaussian probability measure with covariance A , including the point mass when $A = 0$. Thus the parameter is a covariance, not a standard deviation. We write W_p for the usual Wasserstein distance with Euclidean cost, and $\|\cdot\|$ and $\|\cdot\|_F$ for operator and Frobenius norms. All dimensions and constants below are fixed.

Fix $d \geq 1$, $S > 0$ and $0 < c < 1 < C < \infty$. Define

$$\begin{aligned} \mathcal{F}_\zeta = \{f = e^{-V} : \int f = 1, \quad \mathbb{E}_f X = 0, \quad V \in C^2(\mathbb{R}^d), \\ cI \leq D^2 V \leq CI, \quad \|\text{Cov}(f) - I\|_F \leq \zeta\}, \quad \mathcal{K}_S = \{0 \leq A \leq SI\}. \end{aligned}$$

Normalization is incorporated in V . The observation law is $f * \gamma_\Sigma$, with $\Sigma \in \mathcal{K}_S$ unknown. For $0 < \delta < 1$, put $\ell(\delta) = [\log(1/\delta)]^{-3/2}$ and set $\ell(0) = 0$.

Theorem 1.1 (Pairwise partial identification). *There are positive k, K, ζ_0, δ_0 , depending only on d, c, C, S , such that for $0 \leq \zeta \leq \zeta_0$ and $0 \leq \delta \leq \delta_0$,*

$$k[\zeta + \ell(\delta)] \leq \sup \{W_2(f, g) : f, g \in \mathcal{F}_\zeta, \Sigma, \Lambda \in \mathcal{K}_S, W_1(f * \gamma_\Sigma, g * \gamma_\Lambda) \leq \delta\} \leq K[\zeta + \ell(\delta)].$$

In particular, exact covariance normalization gives identifiability even when the Gaussian noise covariance is unknown. At identical outputs the sharp ambiguity order under covariance uncertainty ζ is linear in ζ . More precisely, every pair in the supremum satisfies

$$W_2(f, g) \leq K [\min\{\zeta, \|\Sigma - \Lambda\|_F\} + \ell(\delta)]. \quad (1)$$

There are two independent obstructions in this formula: exact Gaussian allocation gives the linear covariance term, while Gaussian smoothing gives the inverse-logarithmic term. The latter exponent and its hard analytic proof come from the known-noise theorem of [10]. Our central question is which additional covariance information removes the former obstruction.

Theorem 1.2 (Sharp calibration-sample tradeoff). *In addition to n independent observations from $f * \gamma_\Sigma$, observe $m \geq 1$ independent samples from γ_Σ , independent of the primary observations. For $r = 1, 2$, all sufficiently large n , every $m \geq 1$, and $0 \leq \zeta \leq \zeta_0$,*

$$\inf_{\hat{f}} \sup_{f \in \mathcal{F}_\zeta, \Sigma \in \mathcal{K}_S} \mathbb{E}_{f, \Sigma} W_2(\hat{f}, f)^r \asymp \left[\min\{\zeta, m^{-1/2}\} + (\log n)^{-3/2} \right]^r. \quad (2)$$

The constants are uniform in m and ζ . Without calibration observations the same statement has ζ in place of the minimum.

Theorem 1.2 is proved in Section 5. It gives the precise criterion $\min\{\zeta, m^{-1/2}\} = O((\log n)^{-3/2})$ for preserving the known-noise minimax order. In the regime $\zeta / (\log n)^{-3/2} \rightarrow \infty$, this is equivalent to $m = \Omega((\log n)^3)$. If the covariance prior is already at that resolution, calibration samples are unnecessary for this rate.

The known latent covariance is information about the model. It cannot be obtained by whitening noisy data without also estimating the noise. If it is removed, the Gaussian pairs

$$(X, Z) \sim N(0, I) \otimes N(0, sI), \quad (X', Z') \sim N(0, (1+t)I) \otimes N(0, (s-t)I)$$

produce the same output whenever $0 < t < s$.

Our third information regime retains two noisy repetitions of each latent sample. Let $0 < a < 1 < b < \infty$ be fixed and define

$$\mathcal{H}_{a,b} = \{f = e^{-V} : \int f = 1, \mathbb{E}_f X = 0, V \in C^2, cI \leq D^2 V \leq CI, aI \leq \text{Cov}(f) \leq bI\}.$$

Write $Q_{f, \Sigma_1, \Sigma_2} = \mathcal{L}(X + Z_1, X + Z_2)$, where $X \sim f$ and the two centered Gaussian errors are mutually independent and independent of X , with unknown $\Sigma_j \in \mathcal{K}_S$. The paired observation space has its Euclidean norm on $\mathbb{R}^{2d}$.

Theorem 1.3 (Paired full identification and sharp estimation). *Uniformly over $f, g \in \mathcal{H}_{a,b}$ and all permitted noise matrices, for sufficiently small $\delta \geq 0$,*

$$W_1(Q_{f, \Sigma_1, \Sigma_2}, Q_{g, \Lambda_1, \Lambda_2}) \leq \delta \implies W_2(f, g) \leq K\ell(\delta).$$

From n independent paired observations, for $r = 1, 2$ and all sufficiently large n ,

$$\inf_{\hat{f}} \sup_{\substack{f \in \mathcal{H}_{a,b} \\ \Sigma_1, \Sigma_2 \in \mathcal{K}_S}} \mathbb{E}_{f, \Sigma_1, \Sigma_2} W_2(\hat{f}, f)^r \asymp (\log n)^{-3r/2}. \quad (3)$$

All constants depend only on the displayed fixed class parameters. The upper estimator may take values in $\mathcal{H}_{a,b}$.

Section 6 proves this theorem without a latent covariance prior: the cross-covariance of each retained pair identifies that covariance. The noise matrices may be unequal, noncommuting or singular. Both inputs in the pairwise conclusion satisfy two-sided curvature bounds; the rough-input fixed-reference result below is a complementary regime.

A complementary result concerns a fixed reference and a larger input class. Suppose

$$\begin{aligned} \rho &= \nu * \gamma_{\tau I}, \quad \tau > 0, \quad \nu = Z^{-1} e^{-U}, \\ \mathbb{E}_{\nu} X &= 0, \quad \text{Cov}(\rho) = I, \quad U \in C^2(\mathbb{R}^d), \\ \alpha I &\leq D^2 U \leq \beta I. \end{aligned} \tag{4}$$

Here $0 < \alpha \leq \beta < \infty$. Necessarily $0 < \tau < 1$. Let

$$\mathcal{L}_{\zeta} = \{f : f \text{ is a centered log-concave density on } \mathbb{R}^d, \|\text{Cov}(f) - I\|_{\text{F}} \leq \zeta\}.$$

No differentiability or positive curvature is required of f .

Theorem 1.4 (A Gaussian-divisible fixed reference). *Under (4), for all sufficiently small $\zeta, \delta \geq 0$,*

$$\sup_{\substack{f \in \mathcal{L}_{\zeta}, \Sigma, \Lambda \in \mathcal{K}_S \\ W_1(f * \gamma_{\Sigma}, \rho * \gamma_{\Lambda}) \leq \delta}} W_2(f, \rho) \asymp \zeta + \ell(\delta),$$

with constants depending only on the fixed reference, d and S . Moreover, if $q = \rho * \gamma_{\Lambda_0}$, $\Lambda_0 \in \mathcal{K}_S$, and ζ is sufficiently small, its compatible inputs are exactly

$$\{\nu * \gamma_{\tau I - D} : D = D^{\top}, \|D\|_{\text{F}} \leq \zeta, 0 \leq \Lambda_0 + D \leq SI\}. \tag{5}$$

Theorem 1.1 treats pairs throughout the curvature class. Theorem 1.4 uses the exact positive Gaussian component (4) to admit rough inputs. Section 6 instead uses the joint law of two independent-error repetitions of the same latent sample; this provides the missing covariance information directly and permits a general smooth reference.

The statistical consequences are proved after the deterministic interfaces. Single-measurement estimation and divisible-reference testing retain the sharp covariance floor. Pure-noise calibration replaces it by the minimum in (2); paired measurements remove it in full smooth-class estimation and in rough-input fixed-reference testing. The measurable minimum-distance procedures establish minimax rates in fixed dimension.

1.1. Relation to earlier work

Unknown-variance deconvolution and Gaussian-allocation ambiguity are classical. Matias [9] identifies scalar Gaussian variance by excluding a Gaussian component from the latent density, with losses for variance and pointwise density estimation. Our approximately normalized class permits that allocation and quantifies its sharp transport floor. Gassiat, Le Corff and Lehéricy [7] obtain identification from dependent signal blocks and independent noise blocks under a Laplace-growth condition below order two. This is a different information contract from the covariance calibration used here.

Pure-noise auxiliary samples have a substantial deconvolution literature. Johannes [8] treats the primary-plus-auxiliary experiment with nonparametric ordinary or supersmooth error distributions, Sobolev losses and Fourier source conditions. His results already separate primary-sample and noise-sample errors. Here the noise belongs to a finite-dimensional Gaussian covariance family; the shape-constrained W_2 loss and prior covariance information give the sharp minimum in (2). Scricciolo [11] studies auxiliary-sample deconvolution under

ordinary-smooth errors in univariate W_1 loss. These comparisons identify the model-specific calibration contribution.

Repeated measurements likewise provide classical identification. Delaigle, Hall and Meister [6] show that replicated observations can preserve known-error first-order performance in substantial deconvolution classes. Their density-loss results establish the broader statistical principle; here the curvature-constrained transport loss and Gaussian covariance family determine the precise rate. Capitao-Miniconi, Gassiat and Lehericy [4] treat unknown errors through repeated observations and finite-order Laplace-transform conditions. Their repeated-measurement condition allows arbitrary finite positive growth order. Section 6 uses the Gaussian cross-covariance identity to prove sharp full smooth-class W_2 estimation and a testing law against rough log-concave inputs with bounded covariance.

The difficult inverse exponent $3/2$, reference-uniform constants and exact-moment alternatives are inherited from [10]. The present contribution is the sharp calibrated partial-identification law, its two-sample resource tradeoff, paired smooth-class estimation, and the rough-input fixed-reference extensions. Bobkov, Chistyakov and Götze [2] and Courtade, Fathi and Pananjady [5] provide related information-distance and entropy-power stability results with different observation functionals. Scricciolo [12, Theorem 2] obtains supersmooth logarithmic rates in univariate W_1 deconvolution under known errors and dependent signals. That Sobolev-type distribution-estimation result has different loss, regularity and information assumptions from the paired unknown-covariance W_2 experiment here.

2. Analytic tools and the scalar-noise input

We first state precisely the earlier result used below. The isotropic log-concave class consists of all centered log-concave densities with covariance I , including densities with nonsmooth convex potential or bounded support.

Proposition 2.1 (Scalar-noise inverse and alternatives, [10]). *Fix a dimension and scalar Gaussian variance $B > 0$. For isotropic references g whose C^2 potentials have common positive lower and finite upper Hessian bounds, there are common constants $K, \varepsilon_0 > 0$ such that every isotropic log-concave f satisfies*

$$W_2(f, g) \leq K\ell(W_1(f * \gamma_{BI}, g * \gamma_{BI}))$$

when the output discrepancy is positive and at most ε_0 . At zero discrepancy $f = g$. At every fixed permitted reference g there are centered isotropic log-concave f_T , for every sufficiently large real T , with

$$\begin{aligned} W_2(f_T, g) &\geq kT^{-3}, \\ W_1(f_T * \gamma_{BI}, g * \gamma_{BI}) &\leq KT^{-2}e^{-aT^2}, \\ \chi^2(f_T * \gamma_{BI} \mid g * \gamma_{BI}) &\leq KT^{-4}e^{-2aT^2}. \end{aligned} \tag{6}$$

At a standard Gaussian reference the alternatives can be chosen in any fixed curvature interval $0 < c < 1 < C$ by reducing their fixed amplitude.

For clarity, the alternatives in [10] are actual positive densities, not signed formal perturbations. They add a narrow Gaussian times $T^{-2}[\cos(Tx_1) - b_T \cos(x_1) + c_T]$, with b_T, c_T

chosen to cancel mass and covariance exactly. Uniform curvature bounds control the perturbation, and the Gaussian envelope gives the chi-square bound in (6). We use the theorem and its statistical alternatives as established inputs; all covariance transfers needed in the present paper are proved below.

Lemma 2.2 (Moments and covariance interpolation). *Centered log-concave laws with covariance bounded above have uniformly bounded fourth moments in fixed dimension. Consequently, for two such laws p, q with a common fourth-moment bound and $\delta = W_1(p, q) \in (0, 1)$,*

$$\|\text{Cov}(p) - \text{Cov}(q)\|_F \leq K\delta^{2/3}. \quad (7)$$

The same bound applies after adding Gaussian noise with uniformly bounded covariance.

Proof. For a centered one-dimensional marginal with variance at most v^2 , Chebyshev gives upper-tail probabilities at least $3/4$ at $-2v$ and at most $1/4$ at $2v$. The survival function is log-concave, by the marginal/integration property of log-concave laws. Its logarithm therefore has a uniform negative secant slope, which extends to the right by concavity. If that tail is already zero, the conclusion is immediate. Applying the same argument to the negative marginal gives exponential tails on both sides. Integration and the coordinate inequality $|x|^4 \leq d \sum_i |x_i|^4$ prove the fourth-moment bound. Gaussian addition preserves it by $|x + y|^4 \leq 8(|x|^4 + |y|^4)$. For (7), multiply $x_i x_j$ by a smooth cutoff equal to one on the radius- R ball and zero outside the radius- $2R$ ball. Its Lipschitz constant is $O(R)$; the omitted second moment is $O(R^{-2})$. Kantorovich duality bounds each entry difference by $K(R\delta + R^{-2})$. Set $R = \delta^{-1/3}$ and sum the finitely many entries. $\square$

Lemma 2.3 (Curvature and forward Gaussian displacement). *Suppose $f = e^{-V}$ is centered and $cI \leq D^2V \leq CI$. Then*

$$C^{-1}I \leq \text{Cov}(f) \leq c^{-1}I, \quad I(f) := \int |\nabla \log f|^2 f \leq dC.$$

*If $A \geq 0$, $f_A = f * \gamma_A$ has potential Hessian between $c(1 + c\|A\|)^{-1}I$ and CI , and*

$$W_2(f, f_A) \leq \frac{1}{2}\sqrt{dC} \|A\|. \quad (8)$$

Proof. The Gaussian tails and linear gradient growth justify integration by parts: $\mathbb{E}\nabla V = 0$, $\mathbb{E}[X\nabla V^\top] = I$, and $\mathbb{E}[\nabla V \nabla V^\top] = \mathbb{E}D^2V \leq CI$. Directional Cauchy–Schwarz gives the covariance lower bound. The upper is Brascamp–Lieb [3]; the Fisher bound follows by taking the trace. For $A > 0$, the posterior curvature is at least $cI + A^{-1}$. Brascamp–Lieb and

$$D^2(-\log f_A) = A^{-1} - A^{-1} \text{Cov}(X | X + Z) A^{-1}$$

give the stronger lower bound $(c^{-1}I + A)^{-1}$. The alternative identity $D^2(-\log f_A) = \mathbb{E}[D^2V | X + Z] - \text{Cov}(\nabla V | X + Z)$ gives the upper. Singular A follows by replacing it with $A + \epsilon I$ and taking the C^2 limit; the first two density derivatives are integrable by the stated envelopes. Finally $f_t = f * \gamma_{tA}$ solves the continuity equation with velocity $-A\nabla \log f_t/2$. Conditional Jensen gives $I(f_t) \leq I(f)$. Its transport action on $0 \leq t \leq 1$ is at most $dC\|A\|^2/4$, so the dynamic Wasserstein bound [1] proves (8). Positive regularization treats singular A . $\square$

Lemma 2.4 (Covariance matching and Gaussian coupling). *On any fixed positive spectral interval, matrix square roots and inverse square roots are Frobenius-Lipschitz. In particular, if C_f, C_g lie in that interval and $L = C_g^{1/2} C_f^{-1/2}$, then*

$$LC_f L^\top = C_g, \quad \|L - I\|_F \leq K \|C_f - C_g\|_F, \quad W_2(f, L_\# f) \leq K \|C_f - C_g\|_F$$

for a centered law f with covariance C_f . If $A, D \geq aI$, $a > 0$, then

$$W_1(\gamma_A, \gamma_D) \leq \|A^{1/2} - D^{1/2}\|_F \leq \frac{\|A - D\|_F}{2\sqrt{a}}. \quad (9)$$

Proof. For $X = A^{1/2} - D^{1/2}$ the Sylvester identity $A^{1/2}X + XD^{1/2} = A - D$ holds without commutation. In left and right orthonormal eigenbases every entry is divided by a sum of square roots at least $2\sqrt{a}$. This proves the square-root estimate. The identity for inverses gives the analogous inverse-square-root bound on a compact positive interval. Also $L - I = (C_g^{1/2} - C_f^{1/2})C_f^{-1/2}$. The same-input coupling proves the transport bound for L , and coupling the same standard Gaussian proves (9). $\square$

Lemma 2.5 (Known scalar noise without isotropy). *For centered densities with common two-sided positive potential-curvature bounds and fixed scalar noise BI , the inverse upper in Proposition 2.1 holds uniformly without requiring isotropy.*

Proof. Put $\delta = W_1(f * \gamma_{BI}, g * \gamma_{BI})$. The moments are uniform by Lemmas 2.2–2.3, so $\|C_f - C_g\|_F = O(\delta^{2/3})$. Replace f by $L_\# f$, $L = C_g^{1/2} C_f^{-1/2}$, paying $O(\delta^{2/3})$ in W_2 and therefore in output W_1 . Apply the common map $C_g^{-1/2}$. Both inputs become isotropic and their potential curvatures remain in one fixed positive interval. The common noise covariance becomes $BC_g^{-1} \leq BCI$. Convolve both outputs with the same Gaussian of covariance $(BC + 1)I - BC_g^{-1}$. The discrepancy only decreases, and the noise is now the exact fixed scalar noise in Proposition 2.1. Undo the uniformly bounded affine map and pay the matching cost. Since $\ell(K\delta^{2/3}) \leq K'\ell(\delta)$ for sufficiently small δ , the result follows. At zero discrepancy the covariances agree and Gaussian injectivity gives $f = g$. $\square$

3. Pairwise identification under covariance uncertainty

Proof of Theorem 1.1: upper bound. Write $p = f * \gamma_\Sigma$, $q = g * \gamma_\Lambda$, and $D = \Sigma - \Lambda$. By Lemma 2.2,

$$\|D\|_F \leq 2\zeta + K\delta^{2/3}, \quad \delta = W_1(p, q). \quad (10)$$

Let D_+, D_- be the spectral positive and negative parts, so $D = D_+ - D_-$. Put $B = S + 1$. The matrix

$$K_0 = BI - \Sigma + D_+ = BI - \Lambda + D_- \geq I \quad (11)$$

is the SAME positive covariance in both formulas, even for noncommuting Σ, Λ . Thus

$$p * \gamma_{K_0} = (f * \gamma_{D_+}) * \gamma_{BI}, \quad q * \gamma_{K_0} = (g * \gamma_{D_-}) * \gamma_{BI}.$$

The new output discrepancy is at most δ . Since $\|D_\pm\| \leq S$, Lemma 2.3 puts both smoothed inputs in common curvature bounds. Lemma 2.5 bounds their W_2 distance by $K\ell(\delta)$. Returning

to f, g by (8) costs at most $K\|D\|_F$. This directly gives $K[\|D\|_F + \ell(\delta)]$; alternatively use (10) and absorb $\delta^{2/3}$ into $\ell(\delta)$. Taking the better bound proves both the stated upper and (1). If $\delta = 0$, the two smoothed laws agree by Gaussian injectivity and the same displacement argument gives (1) directly. $\square$

Matching lower bounds. Choose $0 < s < S$ and $t = \zeta/\sqrt{d}$. For small ζ , the centered Gaussian priors of covariances I and $(1+t)I$ belong to $\mathcal{F}_\zeta$. Noise covariances sI and $(s-t)I$ give identical outputs. Their W_2 distance is $\sqrt{d}(\sqrt{1+t} - 1) \geq k\zeta$. For the logarithmic term, use (6) at the standard Gaussian reference with a common fixed noise sI . Both priors have covariance exactly I and lie in $\mathcal{F}_\zeta$ after reducing the amplitude. For every small δ , choosing real T of order $\sqrt{\log(1/\delta)}$ gives the logarithmic lower. The maximum of these lower bounds pays half their sum. $\square$

4. Gaussian-divisible anchors and exact compatible laws

Theorem 1.4: upper bound. Write $p = f * \gamma_\Sigma$, $q = \rho * \gamma_\Lambda$ and $D = \Sigma - \Lambda$. For $\zeta \leq 1/2$, both priors have bounded covariance, so Lemma 2.2 gives

$$\|D\|_F \leq \zeta + K\delta^{2/3}, \quad \delta = W_1(p, q).$$

Choose the thresholds so this is at most $\min(\tau/2, 1/4)$, and define the actual probability density

$$\rho_D = \nu * \gamma_{\tau I - D}.$$

Its Gaussian covariance lies between $\tau I/2$ and $3\tau I/2$. Lemma 2.3 gives common lower and upper curvature bounds for ρ_D . Moreover

$$\rho_D * \gamma_\Sigma = \rho * \gamma_\Lambda, \quad \text{Cov}(f) - \text{Cov}(\rho_D) = \text{Cov}(p) - \text{Cov}(q). \quad (12)$$

Convolve both observed laws with $\gamma_{BI-\Sigma}$, $B = S + 1$. They now have the common scalar noise BI and latent laws f, ρ_D , with output discrepancy at most δ . The covariance difference in (12) is $O(\delta^{2/3})$, with NO ζ loss.

Let $C_D = I - D$, $C_f = \text{Cov}(f)$, and replace f by $L_\# f$, $L = C_D^{1/2} C_f^{-1/2}$. This costs $O(\delta^{2/3})$ by Lemma 2.4. Apply $C_D^{-1/2}$ to both laws. Both inputs are isotropic; $L_\# f$ remains log-concave and the reference keeps uniform two-sided curvature. The common noise is $BC_D^{-1} \leq 2BI$. Adding the same Gaussian covariance $(2B+1)I - BC_D^{-1}$ now gives the exact fixed scalar-noise model of Proposition 2.1. Its full log-concave input scope is used here, rather than Lemma 2.5. Undoing the maps yields $W_2(f, \rho_D) \leq K\ell(\delta)$. Finally couple the common ν variable and a standard Gaussian to get

$$W_2(\rho_D, \rho) \leq \|(\tau I - D)^{1/2} - \sqrt{\tau}I\|_F \leq K_\tau \|D\|_F.$$

This proves the upper, without a heat-displacement estimate for the rough input f . $\square$

Exact set and lower bounds. If $f * \gamma_\Sigma = \rho * \gamma_{\Lambda_0}$, covariance equality gives $D = \Sigma - \Lambda_0 = I - \text{Cov}(f)$, so $\|D\|_F \leq \zeta$ and $\tau I - D > 0$ for small ζ . Dividing the nonvanishing Gaussian characteristic multiplier identifies f with $\nu * \gamma_{\tau I - D}$. Conversely every matrix in (5) gives an actual centered log-concave density with covariance $I - D$ and precisely the required

output. This proves (5); no arbitrary backward Gaussian transform is declared positive. For the ambiguity lower take $f = \rho * \gamma_{tI}$, $t = \zeta/\sqrt{d}$, and noise covariances $(s-t)I$, sI , $0 < s < S$. Outputs agree exactly. The reverse triangle inequality in L^2 for any coupling gives

$$W_2(f, \rho) \geq \left| \sqrt{\mathbb{E}_f |X|^2} - \sqrt{\mathbb{E}_\rho |X|^2} \right| = \sqrt{d}(\sqrt{1+t} - 1) \geq k\zeta.$$

The alternatives (6) at the fixed reference ρ remain isotropic log-concave, hence lie in every $\mathcal{L}_\zeta$. They supply the logarithmic lower at every small output discrepancy. The maximum pays half the sum. At zero discrepancy the exact set and the Gaussian coupling above give the upper $K\zeta$ directly. $\square$

5. Statistical consequences

Throughout this section $Y_1, \dots, Y_n$ are independent observations from the stated noisy law, and $Q_n = n^{-1} \sum_i \delta_{Y_i}$.

Lemma 5.1 (A sufficient empirical estimate). *For laws Q with a common second-moment bound and $b = 1/[4(d+1)]$,*

$$\mathbb{E}W_1(Q_n, Q) \leq Kn^{-b}, \quad \Pr\{W_1(Q_n, Q) > n^{-b/2}\} \leq Kn^{-b/2}.$$

Proof. Move the part outside $[-R, R]^d$ to zero and quantize the rest at mesh η , including zero as a site. The expected cost is at most $K\eta + K/R$. The grid has $O((1+R/\eta)^d)$ sites and diameter $O(R)$. Cauchy–Schwarz and the empirical Bernoulli variances bound its expected total variation error by half the square root of its size divided by n . Thus

$$\mathbb{E}W_1(Q_n, Q) \leq K[\eta + R^{-1} + R\sqrt{(1+R/\eta)^d/n}].$$

Take $R = n^b$, $\eta = n^{-b}$. The last term is $O(n^{-1/4})$, and Markov proves the probability bound. This deliberately elementary estimate suffices for a logarithmic inverse rate. $\square$

Theorem 5.2 (Estimation with approximately known latent covariance). *For $r = 1, 2$, every sequence $0 \leq \zeta_n \leq \zeta_0$, and all sufficiently large n ,*

$$\inf_{\hat{f}_n} \sup_{f \in \mathcal{F}_{\zeta_n}, \Sigma \in \mathcal{K}_S} \mathbb{E}_{f, \Sigma} W_2(\hat{f}_n, f)^r \asymp [\zeta_n + (\log n)^{-3/2}]^r.$$

The constants are uniform in the sequence. The infimum is over measurable probability-valued estimators, and an upper estimator may take values in $\mathcal{F}_{\zeta_n}$.

Proof. For each n choose a countable dense subset of $\mathcal{F}_{\zeta_n} \times \mathcal{K}_S$ in W_2 times Frobenius distance. Such a subset exists because the ambient spaces are separable. Gaussian convolution is continuous into W_1 , including singular covariances, by same-Gaussian square-root coupling. Select the first joint candidate whose output W_1 distance from Q_n is within $\eta_n = n^{-b/2}$ of the countable infimum. The distances are continuous in the observations, so selection is measurable. Density gives

$$W_1(p_{\hat{f}_n, \hat{\Sigma}_n}, p_{f, \Sigma}) \leq 2W_1(Q_n, p_{f, \Sigma}) + \eta_n.$$

On the good event in Lemma 5.1, Theorem 1.1 bounds the input error by $K[\zeta_n + (\log n)^{-3/2}]$. The complement has polynomially small probability. The common centered covariance bound

gives a uniform deterministic W_2 diameter by independent coupling, proving both upper risks. For a lower of order ζ_n^r , use the identical-output Gaussian pair in the proof of Theorem 1.1. For the logarithmic lower use (6) with common fixed noise and $T^2 = A \log n$, A sufficiently large. Then $n\chi^2 \rightarrow 0$ and product total variation tends to zero. The two-point estimation inequality gives a lower $k(\log n)^{-3r/2}$. Their maximum is comparable to the stated sum to power r . $\square$

Proof of Theorem 1.2. For $m \geq 1$ let $\widehat{\Sigma}$ be the Frobenius projection of $m^{-1} \sum_{j=1}^m Z_j Z_j^\top$ onto $\mathcal{K}_S$. The Gaussian fourth-moment identity and nonexpansiveness of the projection give, uniformly for all permitted covariances,

$$\mathbb{E} \|\widehat{\Sigma} - \Sigma\|_F^2 \leq K/m. \quad (13)$$

Indeed $\sum_{i,j} \text{Var}(Z_i Z_j) = (\text{tr } \Sigma)^2 + \text{tr}(\Sigma^2) \leq K$ for one observation. This formula covers singular covariances.

Set $b = 1/[4(d+1)]$, $\eta_n = n^{-b/2}$ and $A_n = \{W_1(Q_n, f * \gamma_\Sigma) \leq \eta_n\}$. By Lemma 5.1, $\Pr(A_n^c) \leq K n^{-b/2}$. Choose a deterministic countable dense list $(f_k, \Sigma_k) \in \mathcal{F}_\zeta \times \mathcal{K}_S$. Among indices satisfying

$$W_1(f_k * \gamma_{\Sigma_k}, Q_n) < 2\eta_n, \quad (14)$$

let d_n be the infimum of $\|\Sigma_k - \widehat{\Sigma}\|_F$ and select the first eligible index for which this distance is less than $d_n + \eta_n$. If the eligible set is empty, return the standard Gaussian. All choices are measurable because the list is countable and each displayed distance is measurable. On A_n , density provides eligible approximants to the true parameter, so $d_n \leq \|\Sigma - \widehat{\Sigma}\|_F$. The selected pair then has

$$\begin{aligned} W_1(f_k * \gamma_{\Sigma_k}, f * \gamma_\Sigma) &< 3\eta_n, \\ \|\Sigma_k - \Sigma\|_F &\leq 2\|\widehat{\Sigma} - \Sigma\|_F + \eta_n. \end{aligned} \quad (15)$$

Apply (1) and $\min\{\zeta, 2x + \eta_n\} \leq 2 \min\{\zeta, x\} + \eta_n$. On A_n the resulting error is at most

$$K \left[\min\{\zeta, \|\widehat{\Sigma} - \Sigma\|_F\} + (\log n)^{-3/2} \right].$$

Equation (13) bounds its first and second moments at the desired minimum. On A_n^c all returned laws remain in $\mathcal{F}_\zeta$ and have uniformly bounded centered covariance, hence uniformly bounded W_2 diameter. The polynomial exceptional probability is smaller than both logarithmic risk rates. This proves the upper in (2) uniformly in m, ζ .

For the calibration lower, fix $0 < s < S$ and put $t = a \min\{\zeta, m^{-1/2}\}$, for a sufficiently small fixed $a > 0$. Use latent Gaussians with covariances $I, (1+t)I$ and noise covariances $sI, (s-t)I$. The priors obey all curvature and covariance conditions and $0 \leq t \leq s/2$ after reducing a and ζ_0 . Their primary output laws are identical. For one pure-noise observation the Kullback–Leibler divergence is

$$\frac{d}{2} \left[\frac{s}{s-t} - 1 - \log \frac{s}{s-t} \right] \leq K t^2.$$

Thus the divergence for m calibration observations is at most $K m t^2 \leq K a^2$. Choose a so Pinsker's inequality bounds their total variation by $1/2$. Their latent W_2 separation is at least kt , so the two-point estimation inequality gives the lower $k \min\{\zeta, m^{-1/2}\}^r$.

For the logarithmic lower use (6) at the standard Gaussian, with the same fixed noise sI in both hypotheses. All calibration laws then coincide. Taking $T^2 = A \log n$, A sufficiently large, makes the product primary total variation vanish and gives the lower $k(\log n)^{-3r/2}$. The maximum of the two lower bounds is comparable to the sum in (2) to power r . When $\zeta = 0$ the calibration lower is zero, and the logarithmic lower still applies. Without calibration observations Theorem 5.2 gives the asserted result. $\square$

Theorem 5.3 (Blind-noise fixed-reference testing). *Under (4), define*

$$\mathcal{T}_n(r; \zeta) = \inf_{\psi} \left[\sup_{\Lambda \in \mathcal{K}_S} \Pr(\psi = 1) + \sup_{\substack{f \in \mathcal{L}_\zeta, \Sigma \in \mathcal{K}_S \\ W_2(f, \rho) \geq r}} \Pr(\psi = 0) \right].$$

There are fixed $0 < k < K$ such that, for every sufficiently small known sequence $\zeta_n \geq 0$ and $r_n = \zeta_n + (\log n)^{-3/2}$,

$$\mathcal{T}_n(Kr_n; \zeta_n) \longrightarrow 0, \quad \mathcal{T}_n(kr_n; \zeta_n) \longrightarrow 1.$$

Proof. Reject when $\inf_{\Lambda \in \mathcal{K}_S} W_1(Q_n, \rho * \gamma_\Lambda) > 2\eta_n$, $\eta_n = n^{-b/2}$. Compactness of $\mathcal{K}_S$ and continuity of convolution make the statistic measurable. Under the null the empirical estimate controls rejection. On the good event for the true output, acceptance provides a permitted reference output within $3\eta_n$ of it. Theorem 1.4 then forces $W_2(f, \rho) \leq K_0 r_n$. A strictly larger separation constant makes both errors tend uniformly to zero. For the lower, when $\zeta_n \geq (\log n)^{-3/2}$ the identical-output alternative from Theorem 1.4 has separation at least kr_n for a sufficiently small fixed k , and the two errors sum to one. Otherwise use the common-noise oscillatory alternative with $T^2 = A \log n$. Its product total variation tends to zero and its separation is at least kr_n . A common small constant treats alternating cases. Constant tests give total error at most one, proving the limit. $\square$

Corollary 5.4 (Operational covariance completion under exact isotropy). *If $\zeta = 0$, let $\widehat{\Sigma}$ be the Frobenius projection of $n^{-1} \sum_i Y_i Y_i^\top - I$ onto $\mathcal{K}_S$, and $B = S + 1$. Then*

$$\mathbb{E} \|\widehat{\Sigma} - \Sigma\|_F^2 \leq K/n.$$

*The completed empirical probability measure $\widetilde{Q}_n = Q_n * \gamma_{BI - \widehat{\Sigma}}$ obeys, pathwise,*

$$W_1(\widetilde{Q}_n, f * \gamma_{BI}) \leq W_1(Q_n, f * \gamma_\Sigma) + \frac{1}{2} \|\widehat{\Sigma} - \Sigma\|_F.$$

Proof. Projection onto the closed convex covariance set is nonexpansive; summing entry variances and using the fourth-moment bound gives the first estimate. For the second, first convolve both empirical and true outputs with the SAME estimated complementary Gaussian. This contracts W_1 even though the kernel depends on the data. Compare its true-output convolution with the ideal completion using (9); both complementary covariances are at least I . $\square$

6. Paired measurements as covariance calibration

6.1. Full smooth-class estimation

Proof of Theorem 1.3. Put $C_f = \text{Cov}(f)$ and $C_g = \text{Cov}(g)$. Uniform fourth moments and Lemma 2.2, applied to the joint laws in dimension $2d$, bound their covariance difference by $K\delta^{2/3}$. Their off-diagonal blocks are respectively C_f and C_g , hence $\|C_f - C_g\|_F \leq K\delta^{2/3}$. Use the same analysis-only linear map $L = C_g^{-1/2}$ on both inputs and first-channel outputs. Then $\|L\| \leq a^{-1/2}$, $\|L^{-1}\| \leq \sqrt{b}$, $\text{Cov}(L_{\#}g) = I$, and

$$\|\text{Cov}(L_{\#}f) - I\|_F \leq Ka^{-1}\delta^{2/3}.$$

The transformed potentials have Hessians between caI and CbI ; their noise covariances lie between zero and $(S/a)I$. First-marginal output W_1 is at most $a^{-1/2}\delta$. Thus Theorem 1.1, with these fixed enlarged class constants, gives

$$W_2(L_{\#}f, L_{\#}g) \leq K[\delta^{2/3} + \ell(a^{-1/2}\delta)] \leq K'\ell(\delta).$$

Mapping back proves the pairwise assertion. At $\delta = 0$, equality of joint covariances gives $C_f = C_g$ and $\Sigma_1 = \Lambda_1$. Equality of first-marginal convolutions and the nonvanishing Gaussian characteristic function give $f = g$ directly.

For estimation choose a deterministic countable dense list in $\mathcal{H}_{a,b} \times \mathcal{K}_S^2$ under W_2 and Frobenius metrics. This space is a subspace of a separable metric space and contains the standard Gaussian with zero noise. The map to joint outputs is W_1 -continuous: couple latent variables in W_2 and use common independent standard Gaussians for both errors. Continuity of matrix square roots includes singular covariances. Let Q_n be the empirical paired law and $h = 1/[4(2d+1)]$, $\eta_n = n^{-h/2}$. Select the first candidate whose W_1 distance to Q_n is within η_n of the countable infimum. The index is measurable, and density gives

$$W_1(Q_{\hat{\theta}}, Q_{\theta}) \leq 2W_1(Q_n, Q_{\theta}) + \eta_n.$$

Lemma 5.1 in dimension $2d$ bounds the probability of $W_1(Q_n, Q_{\theta}) > \eta_n$ by $Kn^{-h/2}$. On its complement, the pairwise bound gives error at most $K(\log n)^{-3/2}$. On the exceptional event, independent coupling gives $W_2(\hat{f}, f)^2 \leq 2db$. Its polynomially small probability is negligible for either asserted risk, proving the upper bound.

For the lower fix $0 < s < S$ and restrict both noise matrices to sI . The invertible transformation

$$(Y_1, Y_2) \mapsto ((Y_1 + Y_2)/2, Y_1 - Y_2)$$

gives independent laws $f * \gamma_{(s/2)I}$ and γ_{2sI} . The latter is common to all parameters, so the chi-square divergence equals that of the first coordinate. Use the exact-isotropic alternatives of Proposition 2.1 at the standard Gaussian reference and variance $s/2$. They belong to the fixed curvature interval and have covariance $I \in [aI, bI]$. Their separation is at least kT^{-3} and their one-observation chi-square divergence is at most $KT^{-4}e^{-2AT^2}$ for a fixed $A > 0$. Taking $T^2 = B \log n$ with sufficiently large B makes product total variation tend to zero. The two-point estimation bound proves the lower bound for each $r = 1, 2$. $\square$

6.2. Rough-input fixed-reference testing

Let ρ be any centered isotropic density whose C^2 potential has fixed two-sided positive Hessian bounds. Define

$$\mathcal{G} = \{f : f \text{ is a centered log-concave density, } \frac{1}{2}I \leq \text{Cov}(f) \leq \frac{3}{2}I\}.$$

The endpoints can be replaced by any fixed positive bounds containing I . For $\Sigma_1, \Sigma_2 \in \mathcal{K}_S$, write

$$Q_{f, \Sigma_1, \Sigma_2} = \mathcal{L}(X + Z_1, X + Z_2), \quad (16)$$

where $X \sim f$ and the two Gaussian errors are mutually independent and independent of X . Their coordinates may be correlated and their covariances may be unequal or singular. The joint pairing is retained.

Theorem 6.1 (Paired fixed-reference inverse and sharp testing). *For all sufficiently small $\delta \geq 0$,*

$$W_1(Q_{f, \Sigma_1, \Sigma_2}, Q_{\rho, \Lambda_1, \Lambda_2}) \leq \delta \implies W_2(f, \rho) \leq K\ell(\delta) \quad (17)$$

uniformly over $f \in \mathcal{G}$ and permitted noise covariances. Here W_1 on paired laws uses the Euclidean norm on $\mathbb{R}^{2d}$. For n independent paired observations define the sum-of-errors testing risk

$$\mathcal{P}_n(r) = \inf_{\chi} \left[\sup_{\Lambda_1, \Lambda_2 \in \mathcal{K}_S} \Pr_{\rho, \Lambda_1, \Lambda_2}(\chi = 1) + \sup_{\substack{f \in \mathcal{G}, \Sigma_1, \Sigma_2 \in \mathcal{K}_S \\ W_2(f, \rho) \geq r}} \Pr_{f, \Sigma_1, \Sigma_2}(\chi = 0) \right].$$

There are constants $0 < k < K$ such that

$$\mathcal{P}_n(K(\log n)^{-3/2}) \longrightarrow 0, \quad \mathcal{P}_n(k(\log n)^{-3/2}) \longrightarrow 1. \quad (18)$$

Proof. All joint laws have uniform fourth moments, by Lemma 2.2. Truncated quadratic tests on $\mathbb{R}^{2d}$ therefore bound their full covariance difference by $K\delta^{2/3}$. The covariance matrix in (16) is

$$\begin{pmatrix} C_f + \Sigma_1 & C_f \\ C_f & C_f + \Sigma_2 \end{pmatrix}, \quad C_f = \text{Cov}(f).$$

Comparing its cross block and then its first diagonal block with the reference law gives

$$\|C_f - I\|_F \leq K\delta^{2/3}, \quad \|D\|_F \leq K\delta^{2/3}, \quad D = \Sigma_1 - \Lambda_1. \quad (19)$$

The first marginals have W_1 discrepancy at most δ . Put $B = S + 1$ and convolve both with the same positive Gaussian of covariance $BI - \Sigma_1 \geq I$. They become $f * \gamma_{BI}$ and $\rho * \gamma_{BI-D}$; the discrepancy contracts. Both Gaussian covariances are at least I , so Lemma 2.4 and (19) imply

$$W_1(f * \gamma_{BI}, \rho * \gamma_{BI}) \leq K\delta^{2/3}.$$

Let $L = C_f^{-1/2}$. Its coupling cost satisfies $W_2(f, L_{\#}f) \leq K\delta^{2/3}$. The law $L_{\#}f$ is exactly isotropic and log-concave. Replacing f by this law in the common scalar-noise comparison costs the same amount in output W_1 . Proposition 2.1 therefore gives $W_2(L_{\#}f, \rho) \leq$

$K\ell(K\delta^{2/3}) \leq K'\ell(\delta)$. The triangle inequality proves (17). At $\delta = 0$, (19) gives $C_f = I$ and $\Sigma_1 = \Lambda_1$; first-marginal Gaussian injectivity gives $f = \rho$ directly.

For the testing upper, apply Lemma 5.1 in dimension $2d$ to the empirical paired law. With the corresponding polynomial tolerance η_n , reject when its W_1 distance from $\{Q_{\rho, \Lambda_1, \Lambda_2} : \Lambda_1, \Lambda_2 \in \mathcal{K}_S\}$ exceeds $2\eta_n$. This compactly parametrized family is continuous under same-Gaussian coupling, including singular covariances, so the statistic is measurable. On the good empirical event, acceptance places the true paired law within $3\eta_n$ of some null law. Equation (17) then gives $W_2(f, \rho) \leq K_0(\log n)^{-3/2}$. A larger constant proves the upper, and the same event controls the null error.

For the lower, fix $0 < s < S$ and use $\Sigma_1 = \Sigma_2 = sI$ under both hypotheses. The invertible map

$$(Y_1, Y_2) \mapsto ((Y_1 + Y_2)/2, Y_1 - Y_2)$$

gives independent coordinates with laws $f * \gamma_{(s/2)I}$ and γ_{2sI} . The second law is the same under both hypotheses. Thus the paired chi-square divergence equals that of the first coordinate. Use the every-reference alternatives in (6), now with variance $s/2$, and take $T^2 = A \log n$ for large fixed A . Their product total variation tends to zero and their latent separation is at least $k(\log n)^{-3/2}$. The testing lower bound and a constant test prove the second limit in (18). $\square$

Only observing the separate marginal laws is a different experiment. If all channel covariances have a common positive floor tI , then replacing f by $f * \gamma_{uI}$ and every covariance Σ_j by $\Sigma_j - uI$, $0 < u < t$, preserves every marginal output. The paired cross-covariance in (19) removes precisely this allocation ambiguity.

7. Covariance information and exact Gaussian components

The linear covariance term is an exact identification floor. The calibration theorem quantifies two ways to control it: a prior covariance bound and independent noise observations. Paired data provide a third information regime, recover full smooth-class estimation, and permit rough-input fixed-reference testing using any reference in the stated curvature class. The hard inverse logarithm is common to these conclusions and inherited from the known-noise theory.

The Gaussian component in (4) has a specific mathematical role: it makes every small signed covariance adjustment an actual probability density. Approximate transport proximity alone does not provide that property. The following explicit example records the distinction.

Proposition 7.1 (Smooth nearby laws with zero Gaussian component). *There are centered isotropic C^∞ densities q_j on $\mathbb{R}$, with $\frac{1}{2} \leq (-\log q_j)'' \leq \frac{3}{2}$, such that $W_2(q_j, \gamma_1) \rightarrow 0$ and no q_j has a positive scalar Gaussian convolution component.*

Proof. Let ϕ be the standard Gaussian density, $t_k = 2^k$, $b_k = e^{-t_k}$, and $G(x) = \sum_{k \geq 1} b_k \cos(t_k x)$. Every derivative series converges uniformly. Write $B_r = \sum_k b_k t_k^r < \infty$ and, for small $\varepsilon > 0$, put

$$Z_\varepsilon = 1 + \varepsilon \sum_k b_k e^{-t_k^2/2}, \quad p_\varepsilon(x) = \phi(x) \frac{1 + \varepsilon G(x)}{Z_\varepsilon}.$$

These are positive even probability densities. Direct differentiation gives

$$|(-\log p_\varepsilon)'' - 1| \leq \frac{\varepsilon B_2}{1 - \varepsilon B_0} + \frac{\varepsilon^2 B_1^2}{(1 - \varepsilon B_0)^2}.$$

Their variance is $v_\varepsilon = 1 - \varepsilon Z_\varepsilon^{-1} \sum_k b_k t_k^2 e^{-t_k^2/2}$. For uniformly small ε it belongs to $[9/10, 1]$, and $q_\varepsilon(x) = \sqrt{v_\varepsilon} p_\varepsilon(\sqrt{v_\varepsilon} x)$ is isotropic with the displayed common curvature interval.

Its unscaled characteristic function has only positive summands:

$$\widehat{p}_\varepsilon(t) = Z_\varepsilon^{-1} \left[e^{-t^2/2} + \frac{\varepsilon}{2} \sum_k b_k \{e^{-(t-t_k)^2/2} + e^{-(t+t_k)^2/2}\} \right].$$

In particular $\widehat{p}_\varepsilon(t_k) \geq \varepsilon e^{-t_k^2}/(2Z_\varepsilon)$. If $p_\varepsilon = \mu * \gamma_a$ for any probability measure μ and $a > 0$, its characteristic modulus would be at most $e^{-at^2/2}$, a contradiction along t_k . Scaling preserves the absence of every positive Gaussian factor.

Finally put $\beta = 2\varepsilon B_0 < 1$. Since $p_\varepsilon/\phi \geq (1 - \varepsilon B_0)/(1 + \varepsilon B_0) \geq 1 - \beta$, write $p_\varepsilon = (1 - \beta)\phi + \beta h_\varepsilon$. The residual density has second moment at most one. Couple the common component identically and the residual component independently with ϕ to obtain $W_2(p_\varepsilon, \phi)^2 \leq 2\beta$. The scaling coupling costs $W_2(q_\varepsilon, p_\varepsilon) \leq |1 - \sqrt{v_\varepsilon}| = O(\varepsilon)$. Taking any $\varepsilon_j \downarrow 0$ proves the claim. $\square$

Acknowledgements and disclosure

OpenAI tools assisted with literature comparison, proof development, exposition and verification. The author assumes responsibility for all content.

Funding

The author received no external research funding for this work.

Competing interests

The author declares no competing interests.

References

- [1] J.-D. Benamou and Y. Brenier, A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem, *Numer. Math.* **84** (2000), 375–393.
- [2] S. G. Bobkov, G. P. Chistyakov and F. Götze, Stability of Cramér’s characterization of normal laws in information distances, in *High Dimensional Probability VII*, Progress in Probability 71, Birkhäuser, 2016, pp. 3–35. [doi:10.1007/978-3-319-40519-3_1](https://doi.org/10.1007/978-3-319-40519-3_1).
- [3] H. J. Brascamp and E. H. Lieb, On extensions of the Brunn–Minkowski and Prékopa–Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation, *J. Funct. Anal.* **22** (1976), 366–389.
- [4] J. Capitao-Miniconi, E. Gassiat and L. Lehericy, Deconvolution of repeated measurements corrupted by unknown noise, *Electron. J. Statist.* **19** (2025). <https://doi.org/10.1214/25-EJS2415>.

- [5] T. A. Courtade, M. Fathi and A. Pananjady, Wasserstein stability of the entropy power inequality for log-concave random vectors, *2017 IEEE International Symposium on Information Theory*, 2017, pp. 659–663. https://people.eecs.berkeley.edu/~courtade/pdfs/W2Stability_EPI.pdf.
- [6] A. Delaigle, P. Hall and A. Meister, On deconvolution with repeated measurements, *Ann. Statist.* **36** (2008), 665–685. <https://doi.org/10.1214/009053607000000884>.
- [7] E. Gassiat, S. Le Corff and L. Lehericy, Deconvolution with unknown noise distribution is possible for multivariate signals, *Ann. Statist.* **50** (2022), 303–323. <https://arxiv.org/abs/2006.14226v2>.
- [8] J. Johannes, Deconvolution with unknown error distribution, *Ann. Statist.* **37** (2009), 2301–2323. <https://doi.org/10.1214/08-AOS652>.
- [9] C. Matias, Semiparametric deconvolution with unknown noise variance, *ESAIM Probab. Stat.* **6** (2002), 271–292. [doi:10.1051/ps:2002015](https://doi.org/10.1051/ps:2002015).
- [10] S. Schubr, *Sharp Wasserstein Stability for Gaussian Deconvolution of Log-Concave Laws*, edition 2, Luminary Oddities research archive, 2026. <https://research.luminaryoddities.com/a-sharp-logarithmic-inverse-bound-for-gaussian-estimation/v/2>.
- [11] C. Scricciolo, Adaptive minimax-optimal Wasserstein deconvolution with unknown error distributions, *Statist. Probab. Lett.* **230** (2026), article110589. <https://doi.org/10.1016/j.spl.2025.110589>.
- [12] C. Scricciolo, Minimax rates for Wasserstein–Kantorovich distribution deconvolution under known ordinary smooth errors and dependent signal processes, *Statist. Methods Appl.* (2026). <https://doi.org/10.1007/s10260-026-00858-9>.