SHARP WASSERSTEIN-1 STABILITY FOR DIFFUSIONS UNDER RADIAL DRIFT BOUNDS

Anonymous authors
Paper under double-blind review

ABSTRACT

We study how drift and initialization errors propagate in diffusions that may expand locally but contract at large distances. Under a radial constraint on the reference drift, we derive a finite-time Wasserstein-1 bound for diffusions with the same additive isotropic noise coefficient bounded away from zero. The bound separates initialization error from expected drift error along the perturbed process. Each term is optimal when the other error is absent. Reflection coupling reduces the analysis to a one-dimensional diffusion absorbed at zero. For fixed radial parameters allowing expansion, the amplification coefficient has a polynomial peak as the noise lower bound decreases. Its integral over all time grows exponentially; accumulating a fixed fraction takes an exponentially long time. A reverse Gaussian mixture shows why a sharp uniform coefficient can still overestimate a specified error: the optimal coefficient grows quadratically with mode separation, while adding a fixed positive constant to the drift gives bounded error per unit amplitude.

1 INTRODUCTION

In score-based sampling, errors in the learned score perturb the drift of the reverse diffusion (Song et al., 2021; Dupuis et al., 2026; Beyler & Bach, 2025). We study how these errors are amplified when the drift may expand locally but contracts at large distances.

Consider a reference diffusion on $\mathbb{R}^d$, $d \geq 1$,

$$\mathrm{d}Y_t = b_t(Y_t)\,\mathrm{d}t + \sigma_t\,\mathrm{d}B_t, \qquad \sigma_t \geq \sigma > 0, \quad 0 \leq t \leq T, \tag{1}$$

where $T > 0$, B is standard Brownian motion, and σ_t is scalar. Let μ_t be its law and $\widehat{\mu}_t$ the law of an approximation with drift $\widehat{b}_t$ and the same noise coefficient. Initial laws have finite first moments. Our objective is to control expectation errors uniformly over 1-Lipschitz observables. Their largest error is $W_1(\widehat{\mu}_T, \mu_T) = \sup_{\mathrm{Lip}(f)\leq 1} |\widehat{\mu}_T f - \mu_T f|$, where $\nu f := \int f \,\mathrm{d}\nu$. The linear transport cost also makes ordered scalar couplings optimal, which we use to establish sharpness.

We assume a radial drift bound that permits local expansion and imposes contraction at large distances:

$$\left\langle b_t(x) - b_t(y), \frac{x-y}{\|x-y\|} \right\rangle \leq -m\|x-y\| + 2A, \qquad t \in [0,T], \quad x \neq y, \tag{2}$$

with $m > 0$ and $A \geq 0$. This holds, for example, when $b_t(x) = -\lambda_t x + u_t(x)$, $\lambda_t \geq m$, and $\|u_t\|_\infty \leq A$. Reverse heat flow for compactly supported data satisfies such a bound on any smoothing interval bounded away from zero (Beyler & Bach, 2025); Section 2 gives the parameters.

In one dimension, the drifts $b_\delta(x) = -mx + A\tanh(x/\delta)$, $A > 0$, all satisfy (2), but $b_\delta'(0) = A/\delta - m$ diverges as $\delta \downarrow 0$. The one-sided Lipschitz estimate for amplification, $e^{(A/\delta - m)t}$, therefore diverges for every $t > 0$. The radial parameters remain fixed. With constant noise, reflection coupling gives contraction in a concave transport cost and a Euclidean W_1 bound after metric comparison (Eberle, 2016). We determine the optimal Euclidean bound at each horizon using only m, A, σ, without a common derivative bound.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

The stability bound. We construct an increasing concave function $\Phi_t(r)$ that bounds the error after time t from an initial Wasserstein distance r. Its slope $\kappa(t) = \Phi_t'(0+)$ governs the propagation of drift error. The main estimate is

$$W_1(\widehat{\mu}_T, \mu_T) \le \Phi_T(W_1(\widehat{\mu}_0, \mu_0)) + \int_0^T \kappa(T-t)\|\widehat{b}_t - b_t\|_{L^1(\widehat{\mu}_t)}\,\mathrm{d}t.$$

Drift error is averaged over states visited by the approximation. Reflection coupling yields a scalar separation process (Section 2.2). Its mean propagates finite initial discrepancies; its boundary slope controls infinitesimal displacements from drift error. With the same drift, sufficiently large initial distances contract even when $\kappa(t) > 1$.

Contributions.

(i) *Optimal stability bounds.* Each error term above is optimal over the prescribed drift and noise class when the other source is absent (Section 2.2).

(ii) *Sharpness with a fixed spatial profile.* Smooth scalar examples approach the drift-error coefficient at vanishing amplitude, with both drifts admissible. One spatial profile, fixed for the time weight and desired accuracy before the amplitude limit, suffices because the same point maximizes amplification at every time (Section 3.2).

(iii) *Time scales and spatial dependence.* For the radial class with $m, A > 0$ fixed, decreasing noise gives polynomial peak amplification but an exponentially long time to accumulate a fixed fraction of the total (Section 4.1). A reverse Gaussian mixture shows a further distinction: the optimal coefficient can grow with mode separation while constant forcing of fixed positive amplitude has bounded error per unit amplitude (Section 4.2).

1.1 Related work

Radial comparison and finite-time bounds. For autonomous diffusions with noise coefficient σ, (2) gives Eberle's curvature bound $2(m - 2A/r)/\sigma^2$ (Eberle, 2016). The resulting concave-cost exponential contraction yields a Euclidean W_1 bound with prefactor $2e^{\chi}$, where $\chi = A^2/(m\sigma^2)$. Stronger curvature information can reduce this prefactor. Lyapunov weights and time-dependent drifts extend this approach (Eberle et al., 2019; Conforti, 2023). Two-point comparison propagates a full modulus of continuity for Neumann heat equations with time-dependent drift on bounded strictly convex domains (Andrews & Clutterbuck, 2011), with curvature-based variants (Andrews & Ni, 2012). Here the absorbed mean gives a sharp finite-time Euclidean W_1 bound on $\mathbb{R}^d$ under the radial and noise constraints (Theorem 2.3).

Perturbation coefficients and their optimality. Gradient–Wasserstein duality and semigroup interpolation integrate expected drift error under the perturbed law against transition Lipschitz coefficients (Kuwada, 2010; Del Moral & Singh, 2022). They also enter perturbation and Stein estimates (Rudolf & Schweizer, 2018; Gorham et al., 2019). Sticky coupling treats spatially uniform discrepancies (Eberle & Zimmer, 2019); stationary perturbation and linear response concern invariant-law error and sensitivity (Huggins & Zou, 2017; Hairer & Majda, 2010; Birrell, 2026).

For scalar diffusions, flow and intertwining formulas identify the Wasserstein Lipschitz coefficient through derivatives of the transition mean (Cloez, 2012; Bonnefont & Joulin, 2014; Monmarché, 2026). The exact scalar Poisson-operator norm is also known (Djellout & Wu, 2011); integrating semigroup norms gives an upper bound on it (Cheng et al., 2020). Under invariant initialization, Lemma 3.2 gives a finite-horizon analogue with time weights. Fixed spatial profiles optimize the time-integrated derivative. In our construction, a common maximizer lets one fixed profile approach the integrated coefficient. Smooth approximation of the extremal drift and preservation of admissibility then establish class-wide optimality (Theorem 2.3).

Score-error guarantees. For reverse-OU models, Dupuis et al. (2026) bound a concave-cost Wasserstein distance using ideal-law L^1 score error under learned-drift dissipativity and a nonincreasing local one-sided bound; they also convert this bound to Euclidean W_1. Their alternative bound uses learned-law error under corresponding bounds on the ideal drift, in the concave-cost distance. Our bound uses reference-drift radial control and perturbed-law error; Corollary 2.4 adds a prescribed local bound.

For L^2 error, W_2 guarantees pair ideal-law error with a learned-score one-sided Lipschitz bound (Kwon et al., 2022), or learned-law error with true-score Lipschitz control (Tang & Zhao, 2026). For reverse heat flow, Beyler & Bach (2025) likewise distinguish ideal-law from numerical-iterate averaging.

2 Setup and Main Results

2.1 Diffusions and drift constraints

For $p \geq 1$, write W_p for Wasserstein distance on $\mathcal{P}_p(\mathbb{R}^d)$, the laws with finite pth moment. Initial laws lie in $\mathcal{P}_1$; assume second moments for L^2 quantities. Norms are Euclidean for vectors and operator norms for derivatives. We identify a density with its law.

Assumption 2.1 (Regularity and growth)**.** We call a coefficient pair $(b, \sigma_\cdot)$ *regular* on $[0, T]$ if:

(i) **Noise.** The function $t \mapsto \sigma_t$ is continuous and bounded between positive constants.
(ii) **Drift.** The drift is twice differentiable in space, and b, ∇b, and $\nabla^2 b$ are jointly continuous in (t, x). For some finite K and every $(t, x) \in [0, T] \times \mathbb{R}^d$,

$$\|b_t(x)\| \leq K(1 + \|x\|), \qquad \|\nabla b_t(x)\| + \|\nabla^2 b_t(x)\| \leq K.$$

These conditions give global strong well-posedness and differentiable stochastic flows. The constant K may depend on the coefficient pair. The sufficient decomposition in the introduction extends to $b_t(x) = -M_t x + c_t + u_t(x)$ with $\mathrm{sym}(M_t) \succeq mI$ and $\|u_t\|_\infty \leq A$, where $\mathrm{sym}(M) = (M + M^\top)/2$.

Definition 2.2 (Diffusions under a radial drift bound)**.** For $m > 0$, $A \geq 0$, and $\sigma > 0$, let $\mathcal{D}_T^d(m, A, \sigma)$ be the coefficient pairs $(b, \sigma_\cdot)$ that satisfy Assumption 2.1, obey (2), and have $\inf_{0 \leq t \leq T} \sigma_t \geq \sigma > 0$.

Let $P_{s,t} f(x) = \mathbb{E}[f(Y_t) \mid Y_s = x]$ and write $\mu P_{s,t}$ for the law evolved from μ. We use $\widehat{P}_{s,t}$ for the perturbed transitions and P_t^b for an autonomous drift with constant noise.

2.2 Sharp stability

Fix $m, \sigma, T > 0$ and $A \geq 0$. Let R solve $\mathrm{d}R_t = (2A - mR_t)\,\mathrm{d}t + 2\sigma\,\mathrm{d}W_t$ until it first reaches zero, and remain zero thereafter, where W is scalar Brownian motion. Define

$$\Phi_t(r) = \mathbb{E}_r R_t, \qquad \kappa(t) = \Phi_t'(0+), \qquad R_0 = r \geq 0. \tag{3}$$

This comparison process uses the radial drift upper bound and twice the noise lower bound; Section 3.1 proves that its mean Φ_t is increasing and concave.

Theorem 2.3 (Sharp stability under a radial drift bound)**.** *Let $(b, \sigma_\cdot) \in \mathcal{D}_T^d(m, A, \sigma)$ and let $(\widehat{b}, \sigma_\cdot)$ be regular. For initial laws $\mu_0, \widehat{\mu}_0 \in \mathcal{P}_1(\mathbb{R}^d)$,*

$$W_1(\widehat{\mu}_T, \mu_T) \leq \Phi_T(W_1(\widehat{\mu}_0, \mu_0)) + \int_0^T \kappa(T - s) \|\widehat{b}_s - b_s\|_{L^1(\widehat{\mu}_s)}\,\mathrm{d}s. \tag{4}$$

The two terms are separately optimal over the class, as specified below.

Optimality. With the same drift, for every $0 \leq s \leq t \leq T$ and $\mu, \nu \in \mathcal{P}_1(\mathbb{R}^d)$,

$$W_1(\mu P_{s,t}, \nu P_{s,t}) \leq \Phi_{t-s}(W_1(\mu, \nu))\,. \tag{5}$$

For each prescribed initial distance this is the optimal bound over the class. The least uniform linear coefficient is $\kappa(t - s)$. For $t > 0$, the ratio $\Phi_t(r)/r$ is nonincreasing with limit e^{-mt} as $r \to \infty$ (Section E.3), so contraction holds at sufficiently large initial distance even when $\kappa(t) > 1$.

For drift error, fix $0 \leq w \in C([0, T])$. For averaging laws $\nu = (\nu_t)_{0 \leq t \leq T}$, define

$$D_w(\nu) := \inf \left\{ D \geq 0 : \ \|\widehat{b}_t - b_t\|_{L^1(\nu_t)} \leq D w(t) \text{ for every } t \in [0, T] \right\}. \tag{6}$$

Here $\inf \varnothing = +\infty$; the choice $\nu = \widehat{\mu}$ corresponds to (4).

With m, A, σ, T, w fixed, let $C_T^\star(w)$ be the least $C \geq 0$ such that $W_1(\widehat{\mu}_T, \mu_T) \leq C D_w(\widehat{\mu})$ for every pair in Theorem 2.3 with the same initial law and finite $D_w(\widehat{\mu})$, in every dimension. Then

$$C_T^\star(w) = \int_0^T \kappa(T-s) w(s)\, \mathrm{d}s. \tag{7}$$

The same coefficient is optimal when both drifts are required to lie in the radial-bound class, with either $D_w(\widehat{\mu})$ or $D_w(\mu)$.

For $A > 0$ and $w \not\equiv 0$, sharpness is approached by fixing the reference drift and spatial profile for the required accuracy, then sending the amplitude to zero; both drifts remain admissible. Ornstein–Uhlenbeck translations attain the coefficient at $A = 0$ (Section 3.2).

Additional control of local expansion. A separate bound on local expansion requires the radial component of each drift increment at separation r to be at most Lr. Define the one-sided Lipschitz coefficient by

$$\mathrm{osLip}(b) := \sup_{t, x \neq y} \frac{\langle b_t(x) - b_t(y), x - y \rangle}{\|x - y\|^2}. \tag{8}$$

For differentiable drifts, this is the supremum of the largest eigenvalue of $\mathrm{sym}(\nabla b_t(x))$. For $L \geq 0$, let $\mathcal{D}_T^d(m, A, \sigma; L)$ be the subclass of $\mathcal{D}_T^d(m, A, \sigma)$ satisfying $\mathrm{osLip}(b) \leq L$.

Define $C_{T,L}^\star(w)$ by restricting the reference drift to this subclass. The comparison drift becomes $\min\{Lr, 2A - mr\}$. Denote its absorbed mean by $\Phi_t^L(r)$ and its boundary slope by $\kappa_L(t) = (\Phi_t^L)'(0+)$, with noise 2σ as before.

Corollary 2.4 (Additional one-sided Lipschitz bound)**.** *Theorem 2.3 and the optimality statements above hold for $\mathcal{D}_T^d(m, A, \sigma; L)$ with Φ^L, κ_L, and $C_{T,L}^\star$ in place of Φ, κ, and $C_T^\star$. The same sharpness construction applies. In particular,*

$$C_{T,L}^\star(w) = \int_0^T \kappa_L(T-s) w(s)\, \mathrm{d}s. \tag{9}$$

The coefficient satisfies $\kappa_L(t) \leq \min\{\kappa(t), e^{Lt}\}$ for $t \geq 0$. Synchronous coupling gives the exponential bound. The absorbed comparison uses the smaller of the two radial drift bounds at each separation.

From score error to sampling error. Population score matching measures score error under the forward marginals (Song et al., 2021). To use this error in (4), we exchange the roles of the diffusions, so the radial bound must hold for the learned drift. Initialize the exact reverse SDE and its approximation at the forward terminal law, so that μ_t are the forward marginals in reverse order. When only the score is changed, $\widehat{b}_t - b_t = \sigma_t^2(\widehat{s}_t - s_t)$, where $s_t, \widehat{s}_t$ are the true and estimated scores. Under the standing regularity assumptions and $(\widehat{b}, \sigma_\cdot) \in \mathcal{D}_T^d(m, A, \sigma)$, Cauchy–Schwarz then gives

$$W_1(\widehat{\mu}_T, \mu_T) \leq \int_0^T \kappa(T-t) \sigma_t^2 \|\widehat{s}_t - s_t\|_{L^2(\mu_t)}\, \mathrm{d}t.$$

For example, let q_v be the density of $X + \sqrt{v}Z$, where $\|X\| \leq a$ almost surely, $a \geq 0$, and $Z \sim \mathcal{N}(0, I_d)$ is independent of X. Reverse the unit-noise heat flow from $q_{v_{\max}}$ to $q_{v_{\min}}$, with $0 < v_{\min} < v_{\max} < \infty$, so $T = v_{\max} - v_{\min}$ (Beyler & Bach, 2025). Replacing the posterior mean in the Gaussian score formula by an estimator bounded in norm by a gives a learned drift satisfying the radial bound with $m = 1/v_{\max}$ and $A = a/v_{\min}$. The score-error bound above applies whenever the resulting drift is regular. Appendix G gives the score formula and the bounds for the exact drift (Beyler & Bach, 2025).

3 Proof Outline

Reflection coupling gives a distance process with noise coefficient $2\sigma_t$, whereas the comparison process uses 2σ. To compare their expected distances, we use the monotonicity of Φ_t for the drift

term and its concavity for the extra quadratic variation. We establish these properties by scalar order: for suitable odd drifts, the transition mean has a positive derivative that is nonincreasing on the positive half-line. Approximation transfers these properties to the absorbed comparison mean.

Reflection coupling then bounds initial error by this mean; semigroup perturbation uses its boundary derivative to propagate drift error. The same scalar argument also gives a common maximizing point for the derivative at every time. Symmetric initial points approach the initial-error bound, while one spatial profile localized at this common maximizer approaches the integrated drift-error coefficient.

3.1 Scalar comparison and upper bounds

For a transition kernel, define its Wasserstein-1 Lipschitz coefficient by

$$\mathrm{Lip}_{W_1}(P_{s,t}) := \sup_{x \neq y} \frac{W_1(\delta_x P_{s,t}, \delta_y P_{s,t})}{\|x - y\|}. \tag{10}$$

It is the smallest constant c such that $W_1(\mu P_{s,t}, \nu P_{s,t}) \le c W_1(\mu, \nu)$ for all $\mu, \nu \in \mathcal{P}_1$ (Lemma A.1).

Concavity of the comparison mean. For a regular autonomous scalar drift b with constant noise $\sigma > 0$, write $\mathrm{id}(x) = x$, $g_t(x) = \partial_x P_t^b \,\mathrm{id}(x)$, and $\kappa_b(t) = \mathrm{Lip}_{W_1}(P_t^b)$. Stochastic order reduces Wasserstein distance to the difference of means:

$$W_1(\delta_x P_t^b, \delta_y P_t^b) = P_t^b \,\mathrm{id}(y) - P_t^b \,\mathrm{id}(x) = \int_x^y g_t(z)\,\mathrm{d}z, \qquad x < y.$$

Thus $\kappa_b(t) = \|g_t\|_\infty$; the flow formula gives $g_t > 0$ (Cloez, 2012). Let Y^x be the diffusion reflected at zero with drift b, noise σ, and initial state $x \ge 0$ (Sections A.1 and B.1).

Lemma 3.1 (Concavity and a common maximizer)**.** *If b is odd and b' is nonincreasing on $[0,\infty)$, its mean with absorption at zero equals $P_t^b \,\mathrm{id}$ on $[0,\infty)$ and is increasing and concave. Moreover,*

$$g_t(x) = \mathbb{E}\exp\left\{\int_0^t b'(Y_s^x)\,\mathrm{d}s\right\} \quad (x \ge 0). \tag{11}$$

The derivative g_t is even, and nonincreasing on $[0,\infty)$, so $\kappa_b(t) = g_t(0)$ for every $t \ge 0$.

Proof sketch of Lemma 3.1. Let X^x be the whole-line diffusion started at $x \ge 0$, and τ_0 its first hitting time of zero. Oddness gives $\mathbb{E}_0 X_u = 0$, so the strong Markov property yields

$$\mathbb{E}X_t^x = \mathbb{E}[X_t^x \mathbf{1}_{\{t<\tau_0\}}],$$

identifying the whole-line and absorbed means. Since b' is even and $|X^x|$ has the reflected law by Tanaka's formula, differentiating the whole-line flow gives (11).

Synchronously coupled reflected paths satisfy $Y_u^x \le Y_u^y$ for $0 \le x \le y$. Monotonicity of b' then gives $0 < g_t(y) \le g_t(x)$, so the mean is increasing and concave. The whole-line mean is odd, so g_t is even and $\kappa_b(t) = \|g_t\|_\infty = g_t(0)$ (Section B.3). □

Define $b_L^0(x) = -mx + \mathrm{clip}((m+L)x, -A, A)$, where $\mathrm{clip}(u, -A, A) = \max\{-A, \min\{u, A\}\}$. For its semigroup $P_t^{0,L}$ with noise σ, oddness and $2b_L^0(r/2) = \min\{Lr, 2A - mr\}$ for $r \ge 0$ give

$$\Phi_t^L(r) = 2P_t^{0,L}\,\mathrm{id}(r/2), \qquad r \ge 0. \tag{12}$$

Even unimodal mollification gives regular approximations preserving oddness and a nonincreasing derivative on $[0,\infty)$. Mean convergence (Appendix B) transfers concavity to Φ^L and, as $L \to \infty$, to Φ (Lemma B.2).

Comparison of expected distances. Fix $0 \le s < t \le T$. Couple two copies X, Y of the reference diffusion, started at $X_s = x$ and $Y_s = y$, by reflection (Lindvall & Rogers, 1986). Writing $D_u = \|X_u - Y_u\|$, we will prove

$$W_1(\delta_x P_{s,t}, \delta_y P_{s,t}) \le \mathbb{E}D_t \le \Phi_{t-s}(\|x - y\|) = \mathbb{E}_{\|x-y\|} R_{t-s}.$$

Until meeting, $\mathrm{d}D_u = \delta_u\,\mathrm{d}u + 2\sigma_u\,\mathrm{d}W_u$, where W is scalar Brownian motion and the radial drift satisfies $\delta_u \le 2A - mD_u$; thereafter D remains zero.

Set $F(u,r) = \Phi_{t-u}(r)$. Its backward equation is $F_u + (2A - mr)F_r + 2\sigma^2 F_{rr} = 0$, with $F(t,r) = r$ and $F(u,0) = 0$. Since $F_r \ge 0$ and $F_{rr} \le 0$, the Itô drift of $F(u, D_u)$ is

$$(\delta_u - 2A + mD_u)F_r + 2(\sigma_u^2 - \sigma^2)F_{rr} \le 0. \tag{13}$$

The stopped process is a supermartingale; removing the localization and using $F(t,r) = r$ gives the claimed comparison (Section C). Jensen's inequality under an optimal initial coupling yields (5). Since $\Phi_t(0) = 0$, concavity gives $\Phi_t(r) \le \kappa(t)r$ and hence

$$\mathrm{Lip}_{W_1}(P_{s,t}) \le \kappa(t-s). \tag{14}$$

With the bound L, use Φ^L throughout.

Propagation of drift error. Apply the backward equation for the reference transition $P_{r,t}f$ along the perturbed process. The shared noise coefficient cancels the second-order terms, leaving $(\widehat{b}_r - b_r)\cdot\nabla P_{r,t}f$, averaged under $\widehat{\mu}_r$. For 1-Lipschitz f, the gradient is bounded by $\mathrm{Lip}_{W_1}(P_{r,t})$; duality and the triangle inequality give (Section A)

$$W_1(\widehat{\mu}_t, \mu_t) \le W_1(\widehat{\mu}_0 P_{0,t}, \mu_0 P_{0,t}) + \int_0^t \mathrm{Lip}_{W_1}(P_{r,t})\|\widehat{b}_r - b_r\|_{L^1(\widehat{\mu}_r)}\,\mathrm{d}r. \tag{15}$$

Substituting the initial-distance and Lipschitz bounds proves (4).

3.2 Sharpness

For $A > 0$, the smooth approximations from Section 3.1 must also admit small drift perturbations. Proposition E.1 slightly shrinks their bounded residual, preserving derivative monotonicity while creating strict radial and one-sided Lipschitz margins, even at $L = 0$. The resulting confining drifts have positive invariant densities with Gaussian tails (Section D.3).

Initial error. Reflection coupling for b_L^0 from $r/2$ and $-r/2$ keeps the trajectories symmetric until meeting, so their distance follows the absorbed comparison. The mean-difference formula and (12) give

$$W_1(\delta_{r/2}P_t^{0,L}, \delta_{-r/2}P_t^{0,L}) = \Phi_t^L(r), \qquad r > 0. \tag{16}$$

Smooth approximations converge at each fixed separation, with Lipschitz coefficients converging to κ_L in $L^1([0,\infty))$. A separate sequence with $L_n \to \infty$ gives Φ, κ (Proposition E.1). Thus every bounded time-weighted coefficient integral converges.

Drift error. To determine what one spatial profile can achieve, fix a regular autonomous scalar drift b with constant noise $\sigma > 0$, $T > 0$, and a nonzero $0 \le w \in C([0,T])$. Its weighted transition mean is

$$H_{b,w}(x) = \int_0^T w(T-t)P_t^b\,\mathrm{id}(x)\,\mathrm{d}t. \tag{17}$$

Suppose b has an invariant law $\pi_b \in \mathcal{P}_1(\mathbb{R})$ with positive density. For $0 \ne h \in C_c^\infty(\mathbb{R})$, $h \ge 0$, start b and $b + \eta w(t)h$ at π_b with noise coefficient σ, and write $\mu_t = \pi_b$ and $\widehat{\mu}_t^\eta$ for their respective laws. Invariant initialization fixes the reference averaging law at every time. For a fixed profile h, the first-order response therefore involves the time-integrated derivative.

Lemma 3.2 (Fixed and time-dependent spatial perturbations)**.** *With b, h, w, T fixed as $\eta \downarrow 0$,*

$$\lim_{\eta\downarrow 0}\frac{W_1(\widehat{\mu}_T^\eta, \pi_b)}{D_w(\widehat{\mu}^\eta)} = \lim_{\eta\downarrow 0}\frac{W_1(\widehat{\mu}_T^\eta, \pi_b)}{D_w(\mu)} = \frac{\pi_b(hH'_{b,w})}{\pi_b h}. \tag{18}$$

Optimizing these limits over fixed profiles gives

$$\begin{aligned}\sup_{0\ne h\in C_c^\infty,\ h\ge 0}\frac{\pi_b(hH'_{b,w})}{\pi_b h} &= \|H'_{b,w}\|_\infty = \sup_x \int_0^T w(T-t)g_t(x)\,\mathrm{d}t \\ &\le \int_0^T w(T-t)\kappa_b(t)\,\mathrm{d}t.\end{aligned} \tag{19}$$

The last integral is the optimal coefficient under $D_w(\widehat{\mu})$ over regular perturbations with the same noise coefficient and initial law and finite $D_w(\widehat{\mu})$. Nonnegative time-dependent profiles, fixed before the amplitude limit, approach it.

Scalar order and Duhamel give the terminal error $\eta\pi_b(hH'_{b,w}) + o(\eta)$; synchronous stability gives the drift-error amplitude $\eta\pi_b h + o(\eta)$ under either averaging law (Section D.2). Any coefficient valid at every amplitude must bound the limiting ratio. Localizing h where $H'_{b,w}$ is nearly maximal gives the fixed-profile supremum; Proposition D.1 allows time-dependent localization.

For the approximating drifts, Lemma 3.1 gives $g_t(0) = \kappa_b(t)$ for every t, so the supremum and time integral in (19) commute. The derivative monotonicity that gave concavity for the upper bound therefore also makes a fixed spatial profile sufficient for sharpness.

Admissible perturbations. For $A > 0$, choose $C \geq 0$ below the target integral, an approximating drift b with $H'_{b,w}(0) > C$, and an even bump h supported where $H'_{b,w} > C$, normalized by $\pi_b h = 1$. Then $\pi_b(hH'_{b,w}) > C$. The strict margins in Proposition E.1 keep both drifts admissible for sufficiently small $\eta > 0$, including at $L = 0$. With b, h fixed before $\eta \downarrow 0$, Lemma 3.2 gives a ratio above C under either averaging law. Taking the supremum over these witnesses proves optimality; Section E.4 gives the amplitude restrictions.

At $A = 0$, the clipped drift is $b(x) = -mx$, with $\Phi_t(r) = \Phi_t^L(r) = e^{-mt}r$. Its translations $\widehat{b}_t(x) = -mx + \eta w(t)$ satisfy both constraints and attain the drift-error coefficient exactly (Section E.4).

4 Time Scales and Spatial Profiles

The peak of κ and the time needed to accumulate a fixed fraction of its integral can have different scales. We determine these scales, then compare the optimal coefficient with the response to constant forcing of a fixed diffusion.

4.1 Amplification and Accumulation

Set $\chi = A^2/(m\sigma^2)$, use dimensionless time $q = mt$, and write

$$k_\chi(q) = \kappa(q/m), \qquad \mathcal{I}_\chi(Q) = \int_0^Q k_\chi(q)\,\mathrm{d}q, \qquad \mathcal{I}_\chi^\infty = \int_0^\infty k_\chi(q)\,\mathrm{d}q.$$

For a constant drift-error bound, the coefficient in physical time is

$$C_T^\star(1) = \frac{\mathcal{I}_\chi(mT)}{m}. \tag{20}$$

For $\alpha \in (0,1)$, let $Q_\alpha(\chi)$ be the time at which $\mathcal{I}_\chi$ reaches the fraction α of its total; the corresponding physical horizon is Q_α/m.

Proposition 4.1 (Peak amplification and accumulation time)**.** *As $\chi \to \infty$,*

$$\sup_{q\geq 0} k_\chi(q) = 2\chi + O(\sqrt{\chi}), \qquad \mathcal{I}_\chi^\infty \sim 2\sqrt{\pi\chi}\,e^\chi. \tag{21}$$

For every fixed $\alpha \in (0,1)$,

$$Q_\alpha(\chi) \sim T_\chi \log\frac{1}{1-\alpha}, \qquad T_\chi := \frac{\sqrt{\pi}\,e^\chi}{\sqrt{\chi}}. \tag{22}$$

The equivalent is uniform for α in compact subsets of $(0,1)$.

Thus $\mathcal{I}_\chi^\infty \sim (\sup_q k_\chi(q))T_\chi$: the exponential total comes from the duration of amplification, while the peak grows linearly in χ. On a fixed dimensionless horizon $Q > 0$, the integral is of order χ (Equation (103)):

$$\mathcal{I}_\chi(Q) = 2\chi(Q - 1 + e^{-Q}) + O_Q(\sqrt{\chi}). \tag{23}$$

For comparison, consider contraction in a fixed concave cost. Let $\varphi \in C^2([0,\infty))$ be increasing and concave, with $\varphi(0) = 0$, $\varphi'(0) = 1$, and $a_\varphi := \lim_{r\to\infty} \varphi'(r) > 0$. Suppose φ satisfies

$$2\sigma^2\varphi''(r) + (2A - mr)\varphi'(r) \le -c\varphi(r) \quad (r > 0)$$

for some $c > 0$. Dividing by $\varphi' \ge a_\varphi > 0$ and integrating gives $\log \varphi'(2A/m) \le -\chi$, since $\int_0^{2A/m}(2A - mr)\,\mathrm{d}r = 2A^2/m$. Hence $a_\varphi \le e^{-\chi}$. Using $a_\varphi r \le \varphi(r) \le r$ to convert this contraction estimate to Euclidean W_1 gives the coefficient $a_\varphi^{-1}e^{-ct}$. Its prefactor is at least e^{χ}, whereas the optimal peak grows only linearly in χ.

Proof sketch of Proposition 4.1. Rescale separation by $z = \sqrt{m}\,r/\sigma$. The comparison diffusion Z has drift $2\sqrt{\chi} - z$, noise 2, and absorption time τ_0 at zero. Its mean equation retains the inward linear drift; survival bounds then give the peak estimate and (23) (Equations (52) and (110)).

To relate accumulation to a hitting time, set $\rho_\chi(z) = e^{\sqrt{\chi}z - z^2/4}$ and let P_q be the semigroup of Z killed at zero. For $u_Q = \int_0^Q P_q\,\mathrm{id}\,\mathrm{d}q$, the integrated backward equation gives

$$2\mathcal{I}_\chi(Q) = \int_0^\infty \rho_\chi(z)\big[z - P_Q\,\mathrm{id}(z)\big]\,\mathrm{d}z.$$

Self-adjointness in the weight ρ_χ transfers P_Q from id to 1, and $P_Q 1(z) = \mathbb{P}_z(\tau_0 > Q)$. Since the Poisson formula gives $2\mathcal{I}_\chi^\infty = \int_0^\infty z\rho_\chi(z)\,\mathrm{d}z$, we obtain

$$\frac{\mathcal{I}_\chi(Q)}{\mathcal{I}_\chi^\infty} = \mathbb{P}_{\eta_\chi}(\tau_0 \le Q), \qquad \eta_\chi(\,\mathrm{d}z) = \frac{z\rho_\chi(z)\,\mathrm{d}z}{\int_0^\infty u\rho_\chi(u)\,\mathrm{d}u}. \tag{24}$$

The law η_χ concentrates near the stable drift zero $2\sqrt{\chi}$, far from absorption. Gaussian integration gives the total; τ_0/T_χ converges to a mean-one exponential law, with the Ornstein–Uhlenbeck escape scale (Giorgini et al., 2020). Lemma F.1 and Proposition F.2 prove the identity and limit, respectively, giving (22). □

One-sided Lipschitz bound. For the clipped-linear family b_L^0, set $\lambda = L/m$, $k_{\chi,\lambda}(q) = \kappa_L(q/m)$, and define $\mathcal{I}_{\chi,\lambda}$, $\mathcal{I}_{\chi,\lambda}^\infty$, and $Q_\alpha(\chi,\lambda)$ as above.

For $A, L > 0$, let $U_L' = -b_L^0$ and $p_L \propto e^{-2U_L/\sigma^2}$. Varying L at fixed m, A preserves the minima at $\pm A/m$ but changes the barrier at zero:

$$\log\frac{p_L(A/m)}{p_L(0)} = \frac{2[U_L(0) - U_L(A/m)]}{\sigma^2} = \frac{\chi\lambda}{1+\lambda}. \tag{25}$$

For fixed $\lambda > 0$, the Poisson formula and absorption-time analysis give, as $\chi \to \infty$,

$$\mathcal{I}_{\chi,\lambda}^\infty \sim 2\sqrt{\pi\chi}\,e^{\chi\lambda/(1+\lambda)}, \qquad Q_\alpha(\chi,\lambda) \sim \frac{\pi}{\sqrt{\lambda}}e^{\chi\lambda/(1+\lambda)}\log\frac{1}{1-\alpha}. \tag{26}$$

The quantile equivalent is uniform on compact subsets of $\alpha \in (0,1)$ (Proposition F.2). In contrast, at fixed m, σ, L, T, $C_{T,L}^\star(1) \to \int_0^T e^{Lt}\,\mathrm{d}t$ as $A \to \infty$ (Equation (107)); every fixed horizon captures a vanishing fraction of the total.

At $L = 0$, the potential is flat on $[-A/m, A/m]$. Here $\mathcal{I}_{\chi,0}^\infty = 1 + \sqrt{\pi\chi} + \chi$ and $Q_\alpha(\chi,0) = \Theta_\alpha(\chi)$ as $\chi \to \infty$ (Section F.2).

4.2 Fixed diffusions and constant forcing

To isolate the effect of the perturbation, fix a scalar pair $(b, \sigma.) \in \mathcal{D}_T^1(m, A, \sigma)$, its $\mathcal{P}_1$ initial law, and w. The reference may depend on time. Regular perturbed drifts share its noise and initialization, have finite $D_w(\widehat{\mu})$, and may have time-dependent profiles without preserving the radial bound. The least coefficient $C_T(P, w)$ for this reference diffusion satisfies, for $0 < D_w(\widehat{\mu}) < \infty$ (Proposition D.1),

$$\frac{W_1(\widehat{\mu}_T, \mu_T)}{D_w(\widehat{\mu})} \le C_T(P,w) = \int_0^T w(t)\,\mathrm{Lip}_{W_1}(P_{t,T})\,\mathrm{d}t \le C_T^\star(w). \tag{27}$$

Under $\mathrm{osLip}(b) \le L$, the last bound improves to $C_{T,L}^\star(w)$.

Scalar order makes id optimal for pairs of transition laws; time-dependent localized profiles approach the derivative maxima at vanishing amplitude (Proposition D.1).

Constant forcing in equilibrium. Let b be regular and autonomous in the scalar radial class, with noise σ and invariant density p_b. The centered Poisson integral converges since $\kappa_b \le \kappa$ is integrable:

$$G_b(x) = \int_0^\infty \big(P_t^b\,\mathrm{id}(x) - \bar{x}\big)\,\mathrm{d}t, \qquad \bar{x} = \pi_b\,\mathrm{id}\,.$$

Start b and $b+\eta$ at π_b, with perturbed law $\widehat{\mu}_t^\eta$. Adding the constant η leaves drift increments unchanged, preserving the constraints, and gives $D_1(\mu) = D_1(\widehat{\mu}^\eta) = \eta$. First let $\eta \downarrow 0$ at fixed T, then $T \to \infty$ (Section D.3):

$$\lim_{T\to\infty}\lim_{\eta\downarrow 0}\frac{W_1(\widehat{\mu}_T^\eta, \pi_b)}{\eta} = \frac{2\,\mathrm{Var}_{\pi_b}(X)}{\sigma^2} = \pi_b(G_b') \le \|G_b'\|_\infty \le \int_0^\infty \kappa_b(t)\,\mathrm{d}t.$$

The last two quantities are the infinite-horizon limits in Lemma 3.2.

For odd b with b' nonincreasing on $[0,\infty)$, the common maximizer at zero and the absorbed Poisson formula (Section D.3) give

$$\int_0^\infty \kappa_b(t)\,\mathrm{d}t = \|G_b'\|_\infty = G_b'(0) = \frac{\mathbb{E}_{\pi_b}|X|}{\sigma^2 p_b(0)}. \tag{28}$$

Along the smooth approximations of the clipped-linear wells with $L_n \to \infty$, this maximum tends to $\mathcal{I}_\chi^\infty/m$ (Appendix E).

In the double-well regime, the potential barrier suppresses the midpoint density and determines the exponential scale of the absorption time (Section 4.1). For the smooth approximations, (28) shows how the small density at zero enters the optimal perturbation coefficient. Under an average drift-error constraint, the error can concentrate near zero, where G_b' is maximal. Constant forcing instead averages this derivative under π_b, giving the variance formula above.

Gaussian mixture. A reverse Gaussian mixture makes the same distinction at a fixed horizon and fixed positive forcing amplitude. Take $a > \sqrt{2}$ and $p_s = \frac{1}{2}\mathcal{N}(-a, 1+s) + \frac{1}{2}\mathcal{N}(a, 1+s)$, $0 \le s \le 1$. Its reverse drift is $b_t(x) = -x/(2-t) + [a/(2-t)]\tanh(ax/(2-t))$. Start b_t and $b_t + \varepsilon$ at p_1, with unit noise on $[0,1]$ and fixed $\varepsilon > 0$. Write $\widehat{\mu}_t^\varepsilon$ for the perturbed law and $W_a(\varepsilon) = W_1(\widehat{\mu}_1^\varepsilon, p_0)$. Both averaging laws give $D_1 = \varepsilon$, and both drifts satisfy the radial bound with $m = 1/2, A = a, \sigma = 1$. Here $\mathrm{osLip}(b) = a^2 - 1$ grows with separation (Section G.1).

For the terminal mean $M_t = P_{t,1}\,\mathrm{id}$, Gaussian conditioning gives the derivative; scalar order and Duhamel give the response (Section G.1):

$$M_t'(x) = \frac{1}{2-t} + \frac{a^2(1-t)}{(2-t)^2}\,\mathrm{sech}^2\!\left(\frac{ax}{2-t}\right), \qquad \frac{W_a(\varepsilon)}{\varepsilon} = \int_0^1 \widehat{\mu}_t^\varepsilon(M_t')\,\mathrm{d}t. \tag{29}$$

Equations (27) and (23), together with Section G.2, give, as $a \to \infty$,

$$\begin{aligned} C_1^\star(1) &= 2\mathcal{I}_{2a^2}(\tfrac{1}{2}) = 8(e^{-1/2} - \tfrac{1}{2})a^2 + O(a), \\ C_1(P,1) &= \int_0^1 M_t'(0)\,\mathrm{d}t = \log 2 + a^2(\log 2 - \tfrac{1}{2}), \\ \frac{W_a(\varepsilon)}{\varepsilon} &\longrightarrow \log 2 \qquad (\varepsilon > 0 \text{ fixed}). \end{aligned} \tag{30}$$

The quadratic gap to constant forcing persists even when the reference diffusion is fixed. The growing part of M_t' is concentrated near the midpoint between the modes. Constant forcing averages it over a law that assigns little mass to this region (Section G.2).

5 Conclusion

In the perturbation identity, drift error is paired with gradients of propagated observables. Replacing this pairing by a uniform gradient bound discards where errors occur relative to large gradients and how their directions align. The Gaussian mixture shows that this loss of spatial information can change the order of the bound even for a fixed reference diffusion. Estimating the original pairing from score-learning guarantees could give bounds that retain this spatial dependence.

Reproducibility Statement

This is a theoretical paper. The main text and appendices provide complete proofs of the stability bounds and sharpness constructions (Appendices A–E), the time-scale estimates (Appendix F), and the Gaussian-mixture comparison (Appendix G). The supplementary LaTeX source and bibliography reproduce the manuscript.

AI Use Statement

We used generative AI tools to improve the clarity and presentation of the manuscript and to assist with consistency checks. The authors take full responsibility for the final manuscript and supplementary material.

References

Larbi Alili, Pierre Patie, and Jesper Lund Pedersen. Representations of the first hitting time density of an Ornstein–Uhlenbeck process. *Stochastic Models*, 21(4):967–980, 2005. doi: 10.1080/15326340500294702.

Brian D. O. Anderson. Reverse-time diffusion equation models. *Stochastic Processes and their Applications*, 12(3):313–326, 1982. doi: 10.1016/0304-4149(82)90051-5.

Ben Andrews and Julie Clutterbuck. Proof of the fundamental gap conjecture. *Journal of the American Mathematical Society*, 24(3):899–916, 2011. doi: 10.1090/S0894-0347-2011-00699-1. URL https://arxiv.org/abs/1006.1686v2.

Ben Andrews and Lei Ni. Eigenvalue Comparison on Bakry–Emery Manifolds. *Communications in Partial Differential Equations*, 37(11):2081–2092, 2012. doi: 10.1080/03605302.2012.668602. URL https://arxiv.org/abs/1111.4967v1.

Eliot Beyler and Francis Bach. Convergence of deterministic and stochastic diffusion-model samplers: A simple analysis in Wasserstein distance, 2025. URL https://arxiv.org/abs/2508.03210v2.

Jeremiah Birrell. Sensitivity of SDE solutions to perturbations of the diffusion and drift. *arXiv preprint arXiv:2606.06231*, 2026. URL https://arxiv.org/abs/2606.06231v1.

S. G. Bobkov and F. Götze. Exponential Integrability and Transportation Cost Related to Logarithmic Sobolev Inequalities. *Journal of Functional Analysis*, 163(1):1–28, 1999. doi: 10.1006/jfan.1998.3326.

Michel Bonnefont and Aldéric Joulin. Intertwining Relations for One-Dimensional Diffusions and Application to Functional Inequalities. *Potential Analysis*, 41(4):1005–1031, 2014. doi: 10.1007/s11118-014-9408-7. URL https://arxiv.org/abs/1304.3595v1.

Lingyan Cheng, Ruinan Li, and Liming Wu. Exponential Convergence in the Wasserstein Metric W_1 for One Dimensional Diffusions. *Discrete and Continuous Dynamical Systems*, 40(9):5131–5148, 2020. doi: 10.3934/dcds.2020222. URL https://arxiv.org/abs/1703.00507v1.

Bertrand Cloez. Wasserstein Decay of One Dimensional Jump-Diffusions, 2012. arXiv:1202.1259v2.

Giovanni Conforti. Coupling by reflection for controlled diffusion processes: Turnpike property and large time behavior of Hamilton–Jacobi–Bellman equations. *The Annals of Applied Probability*, 33(6A):4608–4644, 2023. doi: 10.1214/22-AAP1927.

Pierre Del Moral and Sumeetpal Sidhu Singh. Backward Itô–Ventzell and stochastic interpolation formulae. *Stochastic Processes and their Applications*, 154:197–250, 2022. doi: 10.1016/j.spa.2022.09.007.

Hacene Djellout and Liming Wu. Lipschitzian Norm Estimate of One-Dimensional Poisson Equations and Applications. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 47(2):450–465, 2011. doi: 10.1214/10-AIHP360.

Benjamin Dupuis, Tyler Farghly, Maxime Haddouche, Alain Durmus, and Umut Simsekli. Tightening the Score Matching Gap for Diffusion Models, 2026. URL https://arxiv.org/abs/2607.04442v1.

Andreas Eberle. Reflection couplings and contraction rates for diffusions. *Probability Theory and Related Fields*, 166(3–4):851–886, 2016. URL https://arxiv.org/abs/1305.1233v4.

Andreas Eberle and Raphael Zimmer. Sticky couplings of multidimensional diffusions with different drifts. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 55(4):2370–2394, 2019. doi: 10.1214/18-AIHP951. URL https://arxiv.org/abs/1612.06125v1.

Andreas Eberle, Arnaud Guillin, and Raphael Zimmer. Quantitative Harris-type theorems for diffusions and McKean–Vlasov processes. *Transactions of the American Mathematical Society*, 371(10):7135–7173, 2019. doi: 10.1090/tran/7576. URL https://arxiv.org/abs/1606.06012v2.

Ludovico T. Giorgini, Woosok Moon, and John S. Wettlaufer. Analytical survival analysis of the Ornstein–Uhlenbeck process. *Journal of Statistical Physics*, 181:2404–2414, 2020. doi: 10.1007/s10955-020-02669-y. URL https://arxiv.org/abs/2006.13197v2.

I. V. Girsanov. On transforming a certain class of stochastic processes by absolutely continuous substitution of measures. *Theory of Probability & Its Applications*, 5(3):285–301, 1960. doi: 10.1137/1105027.

Jackson Gorham, Andrew B. Duncan, Sebastian J. Vollmer, and Lester Mackey. Measuring sample quality with diffusions. *The Annals of Applied Probability*, 29(5):2884–2928, 2019. doi: 10.1214/19-AAP1467.

Martin Hairer and Andrew J. Majda. A simple framework to justify linear response theory. *Nonlinearity*, 23(4):909–922, 2010. doi: 10.1088/0951-7715/23/4/008. URL https://arxiv.org/abs/0909.4313v3.

Ulrich G. Haussmann and Etienne Pardoux. Time reversal of diffusions. *The Annals of Probability*, 14(4):1188–1205, 1986.

Jonathan Huggins and James Zou. Quantifying the accuracy of approximate diffusions and Markov chains. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, volume 54 of *Proceedings of Machine Learning Research*, pp. 382–391. PMLR, 2017. URL https://proceedings.mlr.press/v54/huggins17a.html.

Kazumasa Kuwada. Duality on gradient estimates and Wasserstein controls. *Journal of Functional Analysis*, 258(11):3758–3774, 2010. doi: 10.1016/j.jfa.2010.01.010. URL https://arxiv.org/abs/0910.1741v1.

Dohyun Kwon, Ying Fan, and Kangwook Lee. Score-based generative modeling secretly minimizes the Wasserstein distance. In *Advances in Neural Information Processing Systems*, volume 35, 2022.

Torgny Lindvall and L. C. G. Rogers. Coupling of multidimensional diffusions by reflection. *The Annals of Probability*, 14(3):860–872, 1986.

Nikiforos Mimikos-Stamatopoulos, Benjamin J. Zhang, and Markos A. Katsoulakis. Score-based generative models are provably robust: An uncertainty quantification perspective. In *Advances in Neural Information Processing Systems*, volume 37, 2024.

Pierre Monmarché. Long-time L^p Wasserstein contraction for diffusion processes without global dissipativity, 2026. arXiv:2603.00773v2.

Daniel Rudolf and Nikolaus Schweizer. Perturbation theory for Markov chains via Wasserstein distance. *Bernoulli*, 24(4A):2610–2639, 2018. doi: 10.3150/17-BEJ938. URL https://arxiv.org/abs/1503.04123v3.

Saeed Saremi, Ji Won Park, and Francis Bach. Chain of log-concave Markov chains. In *International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=yiMB2DOjsR.

Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *International Conference on Learning Representations*, 2021.

Wenpin Tang and Hanyang Zhao. Contractive diffusion probabilistic models. *SIAM Journal on Imaging Sciences*, 19(2):1364–1389, 2026. doi: 10.1137/25M1777992. URL https://arxiv.org/abs/2401.13115v3.

Appendix guide. Section A gives the perturbation estimates, and Sections B and C prove the scalar and radial comparison bounds. Fixed-diffusion coefficients and Poisson formulas are treated in Section D; Section E constructs the smooth approximations and completes the boundary cases of the sharpness proof in Section 3.2. Section F proves the time-scale estimates; Section G treats Gaussian smoothing, noise schedules, and the mixture example.

A Semigroup Perturbation Bounds

Under Assumption 2.1, synchronous coupling and the linear-growth bound give continuous dependence of transition laws in W_1 on the starting point and times. The supremum defining $\mathrm{Lip}_{W_1}(P_{s,t})$ can therefore be taken over a countable dense set of distinct initial pairs, so it is measurable in (s,t). For each coefficient pair, the derivative bound also gives a finite exponential bound on every finite time interval.

A.1 Synchronous stability and flow derivatives

Consider two well-posed additive-noise equations in $\mathbb{R}^d$ with the same initial state and noise. For $K, \delta \geq 0$, if b_t is K-Lipschitz and $\sup_{t,x} \|\widetilde{b}_t(x) - b_t(x)\| \leq \delta$, subtraction and Gronwall give, pathwise,

$$\sup_{0\leq r\leq t} \|\widetilde{X}_r - X_r\| \leq \delta \int_0^t e^{K(t-u)}\,\mathrm{d}u \leq \delta t e^{Kt}. \tag{31}$$

This estimate applies in particular to the globally Lipschitz clipped drifts used below.

For a drift satisfying Assumption 2.1, put $K_j = \sup_{r,x} \|\nabla^j b_r(x)\|$, $j = 1, 2$, on $[0, T]$, using operator norms. For the solution $X_r^{s,x}$, its first two derivatives with respect to the initial state, $J_r = D_x X_r^{s,x}$ and $N_r = D_x^2 X_r^{s,x}$, satisfy

$$\dot{J}_r = \nabla b_r(X_r^{s,x}) J_r, \qquad \dot{N}_r = \nabla b_r(X_r^{s,x}) N_r + \nabla^2 b_r(X_r^{s,x})[J_r, J_r], \quad (J_s, N_s) = (I, 0).$$

Hence $\|J_r\| \leq e^{K_1(r-s)}$ and $\|N_r\| \leq K_2 \int_s^r e^{K_1(r-u)} e^{2K_1(u-s)}\,\mathrm{d}u \leq K_2(r-s)e^{2K_1(r-s)}$. These deterministic bounds justify differentiation under expectation for smooth tests with bounded first two derivatives. In dimension one, J_r is positive, and for the coordinate $f(x) = x$ they yield

$$\partial_x P_{s,t} f(x) = \mathbb{E}_x \exp\left\{\int_s^t \partial_x b_r(X_r^{s,x})\,\mathrm{d}r\right\}, \qquad 0 < \partial_x P_{s,t} f(x) \leq e^{K_1(t-s)}, \tag{32}$$

$$\|\partial_{xx} P_{s,t} f\|_\infty \leq K_2(t-s)e^{2K_1(t-s)}. \tag{33}$$

The estimates include $s = t$ and are uniform in the starting point and times on this fixed horizon.

A.2 Duhamel identity

Proof of (15). Fix $t \leq T$ and first let f be bounded, smooth, and 1-Lipschitz, with bounded derivatives. Under Assumption 2.1, $u(r,x) = P_{r,t} f(x)$ solves

$$\partial_r u + b_r \cdot \nabla u + \frac{\sigma_r^2}{2} \Delta u = 0, \qquad u(t, \cdot) = f.$$

Kantorovich–Rubinstein duality gives

$$\|\nabla u(r, \cdot)\|_\infty \leq \mathrm{Lip}_{W_1}(P_{r,t}). \tag{34}$$

Indeed, $|u(r,x) - u(r,y)| \leq W_1(\delta_x P_{r,t}, \delta_y P_{r,t})$.

Apply Itô's formula along the perturbed process $\widehat{X}_r$, started from $\widehat{\mu}_0$, and first stop on bounded spatial regions. The gradient bound is uniform on $[0,t]$, so the stochastic integral is square-integrable. Linear growth gives $\mathbb{E}\sup_{r\leq t} \|\widehat{X}_r\| < \infty$ for $\widehat{\mu}_0 \in \mathcal{P}_1$; thus the drift residual is integrable. Boundedness of u and dominated convergence remove the stops, giving

$$\widehat{\mu}_t f - \widehat{\mu}_0 P_{0,t} f = \int_0^t \int \langle \widehat{b}_r(x) - b_r(x), \nabla u(r,x) \rangle \widehat{\mu}_r(\,\mathrm{d}x)\,\mathrm{d}r. \tag{35}$$

Therefore

$$|\widehat{\mu}_t f - \widehat{\mu}_0 P_{0,t} f| \le \int_0^t \mathrm{Lip}_{W_1}(P_{r,t}) \|\widehat{b}_r - b_r\|_{L^1(\widehat{\mu}_r)} \, \mathrm{d}r. \tag{36}$$

For a general 1-Lipschitz f with $f(0) = 0$, clip its values to $[-R, R]$ and mollify. Both operations preserve the Lipschitz bound; finite first moments justify sending the mollification scale to zero and then R to infinity. Taking the supremum in (36) therefore bounds $W_1(\widehat{\mu}_t, \widehat{\mu}_0 P_{0,t})$. The triangle inequality and $\mu_t = \mu_0 P_{0,t}$ give (15). □

For $d = 1$, (35) also holds for $f(x) = x$ under the same $\mathcal{P}_1$ assumption. The flow bounds (32)–(33) give a $C^{1,2}$ backward solution with linear growth and bounded first two spatial derivatives. Its stopped values are dominated by $C_t(1 + \sup_{r \le t} |\widehat{X}_r|) \in L^1$, and the drift residual is integrable by linear growth. The stochastic integrand $\sigma_r \partial_x u(r, \widehat{X}_r)$ is bounded, so its martingale is square-integrable without a second-moment assumption on the initial law. These bounds remove the stops. Exchanging the reference and perturbed diffusions gives the identity with transition operators $\widehat{P}$ and discrepancy averaged under μ_r.

For coefficients smooth in time with bounded spatial derivatives through order three, Del Moral & Singh (2022) give this semigroup interpolation formula, also allowing different diffusion matrices. Equal diffusion matrices remove the Hessian residual. The localized backward-equation proof above requires only the time continuity in Assumption 2.1.

A.3 Wasserstein Lipschitz coefficients and initial laws

Lemma A.1 (Extension from points to probability laws)**.** *Let P be a Borel Markov kernel on $\mathbb{R}^d$ with $P(x, \cdot) \in \mathcal{P}_1(\mathbb{R}^d)$ for every x. Let $\omega : [0, \infty) \to [0, \infty)$ be continuous, nondecreasing, and concave, with $\omega(0) = 0$. If*

$$W_1(\delta_x P, \delta_y P) \le \omega(\|x - y\|) \qquad (x, y \in \mathbb{R}^d),$$

then P maps $\mathcal{P}_1$ into $\mathcal{P}_1$, and

$$W_1(\mu P, \nu P) \le \omega(W_1(\mu, \nu)) \qquad (\mu, \nu \in \mathcal{P}_1).$$

Proof. Concavity and monotonicity give $\omega(r) \le \omega(1)(1 + r)$. Applying the pointwise bound to the 1-Lipschitz norm function yields $P\|\cdot\|(x) \le P\|\cdot\|(0) + \omega(\|x\|)$, so $\mu P, \nu P$ have finite first moments. For any 1-Lipschitz f with $f(0) = 0$ and any coupling π of μ, ν, duality and Jensen's inequality give

$$|(\mu P - \nu P) f| \le \int \omega(\|x - y\|) \pi(\mathrm{d}x, \mathrm{d}y) \le \omega\left(\int \|x - y\| \pi(\mathrm{d}x, \mathrm{d}y) \right). \tag{37}$$

The first-moment bounds justify these integrals. Taking the supremum over f and an optimal initial coupling proves the claim. In particular, $\omega(r) = cr$ gives the linear extension for $\mathrm{Lip}_{W_1}(P) \le c$. □

The duality step in (34) is the $p = 1$ case of Kuwada (2010): a constant multiple of the Euclidean metric gives the corresponding Lipschitz operator bound, without the additional metric-measure assumptions needed for the $p > 1$ converse. The analogous discrete-time telescoping argument uses reference transitions and averages error under the perturbed law (Rudolf & Schweizer, 2018), assuming geometric Wasserstein ergodicity and a Lyapunov bound for the perturbed kernel.

The same choice of reference transitions and perturbed-law averaging appears in Mimikos-Stamatopoulos et al. (2024). Their equation $\partial_t m - \Delta m - \mathrm{div}(mb) = 0$ has SDE drift $-b$ and noise $\sqrt{2}$ in our convention. The propagating coefficient is b^1, while the squared residual is averaged under m^2. Their W_1 assertion is on the torus and uses an integrated L^2 bound.

A.4 First-order perturbation estimates

Let $T > 0$ and $0 \le w \in C([0, T])$. Fix a regular coefficient pair $(b, \sigma.)$, $\mu_0 \in \mathcal{P}_1(\mathbb{R}^d)$, and a bounded vector field h_t, continuous in time and smooth in space, with jointly continuous bounded spatial derivatives. Set $\widehat{b}_t = b_t + \eta w(t) h_t$, $\eta \ge 0$, with the same noise coefficient and initial law. With K_1, K_2

from Section A.1, put $H_0 = \sup_t \|h_t\|_\infty$ and $H_1 = \sup_t \|\nabla h_t\|_\infty$. With $A_T = T\|w\|_\infty H_0 e^{K_1 T}$, the norm inequality in (31) gives

$$\sup_{t\le T} \|\widehat{X}_t^\eta - X_t\| \le \eta A_T. \tag{38}$$

For a smooth bounded test f with bounded derivatives, put $F_t = h_t \cdot \nabla P_{t,T} f$. The flow bounds in Section A.1 and the chain and product rules give

$$K_F := \sup_{t\le T} \mathrm{Lip}(F_t) \le H_1 e^{K_1 T}\|\nabla f\|_\infty + H_0 e^{2K_1 T}\big(\|\nabla^2 f\|_\infty + K_2 T\|\nabla f\|_\infty\big). \tag{39}$$

These bounds and the Duhamel identity also apply to $f(x) = x$ in dimension one with $\mu_0 \in \mathcal{P}_1$ (Section A.2). Since $\mathrm{Lip}(\|h_t\|) \le H_1$, coupling gives

$$\sup_{t\le T} |\widehat{\mu}_t^\eta(\|h_t\|) - \mu_t(\|h_t\|)| \le \eta A_T H_1, \qquad \sup_{t\le T} |\widehat{\mu}_t^\eta F_t - \mu_t F_t| \le \eta A_T K_F. \tag{40}$$

Substitute $\widehat{b}_t - b_t = \eta w(t) h_t$ into Duhamel's identity and apply (40) to obtain

$$\left| (\widehat{\mu}_T^\eta - \mu_T) f - \eta \int_0^T w(t)\mu_t F_t \,\mathrm{d}t \right| \le \eta^2 A_T K_F \int_0^T w(t)\,\mathrm{d}t. \tag{41}$$

For $w \not\equiv 0$, let $d_h = \sup_{t:w(t)>0} \mu_t(\|h_t\|)$. Since $\|\widehat{b}_t - b_t\|_{L^1(\nu_t)} = \eta w(t)\nu_t(\|h_t\|)$ for every probability law ν_t,

$$D_w(\mu) = \eta d_h, \qquad \left| \frac{D_w(\widehat{\mu}^\eta)}{\eta} - d_h \right| \le \eta A_T H_1 \quad (\eta > 0). \tag{42}$$

In particular, $d_h > 0$ ensures a positive denominator for small η. The constants depend on the fixed b, h, w, T, f. Boundedness of h_t and its derivatives suffices throughout; the later localization constructions use compactly supported profiles. If $w \equiv 0$, $\widehat{\mu}_T = \mu_T$ and both drift-error amplitudes vanish.

Nonnegative scalar perturbations. In dimension one, suppose $h_t \ge 0$ and $\mu_0 \in \mathcal{P}_1$. Scalar comparison gives $\widehat{X}_t^\eta \ge X_t$ under synchronous coupling. Ordered transport (46) and the coordinate Duhamel identity therefore give, for every $\eta \ge 0$,

$$W_1(\widehat{\mu}_T^\eta, \mu_T) = \eta \int_0^T w(t)\widehat{\mu}_t^\eta(h_t\, \partial_x P_{t,T}\,\mathrm{id})\,\mathrm{d}t. \tag{43}$$

The bounds above apply with $f = \mathrm{id}$, so (41) is also an expansion of this Wasserstein distance. In particular, for $h_t \equiv 1$ and nonzero w, $D_w(\mu) = D_w(\widehat{\mu}^\eta) = \eta$.

B Scalar Comparison Diffusions

Rescaled comparisons. Set $q = mt$, $z = \sqrt{m}\, r/\sigma$, $\chi = A^2/(m\sigma^2)$, and $\lambda = L/m$. The physical comparison processes defining Φ and Φ^L become

$$\gamma_\chi(z) = 2\sqrt{\chi} - z, \qquad \gamma_{\chi,\lambda}(z) = \min\{\lambda z, 2\sqrt{\chi} - z\}, \qquad \mathrm{d}Z_q = \gamma(Z_q)\,\mathrm{d}q + 2\,\mathrm{d}W_q, \tag{44}$$

with absorption at zero. For either drift γ, start at $Z_0 = z \ge 0$ and write τ_0 for the first hit of zero. Define $v_\chi(q, z) = \mathbb{E}_z Z_q$ and $k_\chi(q) = \sup_{z>0} v_\chi(q, z)/z$, using subscripts χ, λ for the constrained process. The concavity argument below identifies this supremum with $\partial_z v_\chi(q, 0+)$ and gives the analogous constrained identity. Brownian scaling then yields

$$\Phi_t(r) = \frac{\sigma}{\sqrt{m}} v_\chi\left(mt, \frac{\sqrt{m}}{\sigma} r\right), \qquad \kappa(t) = k_\chi(mt), \tag{45}$$

with the same identities relating Φ^L, κ_L to $v_{\chi,\lambda}, k_{\chi,\lambda}$. Throughout this appendix, $a_\chi := 2\sqrt{\chi}$ is the scalar drift parameter.

B.1 Stochastic order and Wasserstein distance

Stochastic order and synchronous solutions. For stochastically ordered $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$ with $\mu \succeq_{\mathrm{st}} \nu$, the ordered quantile coupling and the coordinate test give matching transport bounds:

$$W_1(\mu,\nu) = \int x\,\mu(\mathrm{d}x) - \int x\,\nu(\mathrm{d}x). \tag{46}$$

Every coupling (U, V) with $U \geq V$ has this cost. Synchronous solutions of the same scalar SDE preserve their initial order: a crossing requires a meeting, after which pathwise uniqueness forces coalescence.

We also use comparison for different drifts. If $\widetilde{b}_t \geq b_t$, b_t is K-Lipschitz, and $X_0 \leq \widetilde{X}_0$, drive both equations with the same additive-noise path. Their difference is absolutely continuous, and $R_t = (X_t - \widetilde{X}_t)^+$ satisfies $R_t' \leq K R_t$ almost everywhere. Gronwall gives $R_t = 0$. In particular, a nonnegative drift perturbation orders the terminal laws, so (46) identifies its Wasserstein error with a difference of means. This argument requires no monotonicity of b_t.

B.2 Brownian survival comparison

For either absorbed scalar drift γ, $|\gamma(y)| \leq a_\chi + y$ on $y \geq 0$. Stopping also at an upper level, pathwise Gronwall and Doob's bound $\mathbb{E}\sup_{s\leq Q}|W_{s\wedge\tau}|^2 \leq 4Q$ give

$$\mathbb{E}_z \sup_{s\leq Q}(Z_s^\gamma)^2 \leq 3e^{2Q}\big(z^2 + a_\chi^2 Q^2 + 16Q\big) < \infty. \tag{47}$$

Indeed, the stopped supremum is at most $e^Q(z + a_\chi Q + 2\sup_{s\leq Q}|W_{s\wedge\tau}|)$, with τ the first boundary hit. The estimate is independent of the upper level, so nonexplosion and Fatou's lemma remove that stop. It also holds for the absorbed Brownian comparison below. Set $M_\chi := 1 + 2\chi + 2\sqrt{2\chi/\pi}$.

Lemma B.1 (Brownian bounds for survival and the absorbed mean)**.** *For $q \geq 0$,*

$$k_\chi(q) \leq 1 + 2\sqrt{2\chi q/\pi} + 2\chi q, \tag{48}$$

$$k_\chi(q) \leq e^{-q} + 2\chi(1 - e^{-q}) + \sqrt{2\chi/\pi}\int_0^q e^{-(q-s)}s^{-1/2}\,\mathrm{d}s \leq M_\chi. \tag{49}$$

The same upper bounds hold for $k_{\chi,\lambda}$ for every $\lambda \geq 0$.

Proof. Let C^z be Brownian motion with drift a_χ and diffusion coefficient 2, started at $z > 0$ and absorbed at zero. Until absorption, its drift dominates $a_\chi - Z_q$; synchronous one-dimensional comparison, continued with absorption, gives $Z_q^z \leq C_q^z$ and $\mathbb{P}_z(\tau_Z > s) \leq \mathbb{P}_z(\tau_C > s)$. Write Φ_{N} and ϕ_{N} for the standard Gaussian distribution and density. The killed transition density of C is

$$p_{a_\chi}(s,z,y) = \frac{e^{a_\chi(y-z)/4 - a_\chi^2 s/8}}{\sqrt{8\pi s}}\left[e^{-(y-z)^2/(8s)} - e^{-(y+z)^2/(8s)}\right], \quad y > 0.$$

Exponential tilting of the Dirichlet heat kernel on $(0,\infty)$ gives this density. Integrating in y gives

$$S_{a_\chi}(s,z) := \mathbb{P}_z(\tau_C > s) = \Phi_{\mathrm{N}}\left(\frac{z + a_\chi s}{2\sqrt{s}}\right) - e^{-a_\chi z/2}\Phi_{\mathrm{N}}\left(\frac{a_\chi s - z}{2\sqrt{s}}\right). \tag{50}$$

For $s > 0$, the identity $e^{-a_\chi z/2}\phi_{\mathrm{N}}((a_\chi s - z)/(2\sqrt{s})) = \phi_{\mathrm{N}}((a_\chi s + z)/(2\sqrt{s}))$ yields

$$\partial_z S_{a_\chi}(s,z) = \frac{1}{\sqrt{s}}\phi_{\mathrm{N}}\left(\frac{z + a_\chi s}{2\sqrt{s}}\right) + \frac{a_\chi}{2}e^{-a_\chi z/2}\Phi_{\mathrm{N}}\left(\frac{a_\chi s - z}{2\sqrt{s}}\right) \leq \frac{1}{\sqrt{2\pi s}} + \frac{a_\chi}{2}.$$

Since $S_{a_\chi}(s,0) = 0$, integration over the initial point gives

$$\frac{\mathbb{P}_z(\tau_Z > s)}{z} \leq \frac{S_{a_\chi}(s,z)}{z} \leq \frac{1}{\sqrt{2\pi s}} + \frac{a_\chi}{2}. \tag{51}$$

Retain the inward linear term in the absorbed equation:

$$\mathrm{d}Z_q = (-Z_q + a_\chi \mathbf{1}_{\{q<\tau_Z\}})\,\mathrm{d}q + 2\mathbf{1}_{\{q<\tau_Z\}}\,\mathrm{d}W_q.$$

The finite-horizon bound (47) justifies the integrating factor and expectation. With $\tau_0 = \tau_Z$, this gives

$$v_\chi(q,z) = e^{-q}z + 2\sqrt{\chi}\int_0^q e^{-(q-s)}\mathbb{P}_z(\tau_0 > s)\,\mathrm{d}s. \tag{52}$$

Using (51) proves the first inequality in (49). If $J(q) = \int_0^q e^{-(q-s)}s^{-1/2}\,\mathrm{d}s$, then $J(q) \le 2\sqrt{q}$. The substitution $s = y^2$ also gives, for $q > 0$,

$$J(q) = 2\int_0^{\sqrt{q}} e^{-(\sqrt{q}-y)(\sqrt{q}+y)}\,\mathrm{d}y \le \frac{2(1-e^{-q})}{\sqrt{q}} \le \frac{2}{\sqrt{q}}.$$

Consequently $J(q) \le 2$, proving the uniform bound. The elementary estimate (48) follows from the same damped bound using $e^{-q} \le 1$, $1 - e^{-q} \le q$, and $J(q) \le 2\sqrt{q}$. All integrations are controlled by the explicit integrable function in (51). Finally $\gamma_{\chi,\lambda}(z) \le a_\chi - z$ on $z > 0$, so comparison with the same noise and absorption at zero gives $k_{\chi,\lambda} \le k_\chi$. □

B.3 Odd flows, concavity, and a common maximizer

Proof of Lemma 3.1. Let X^x solve

$$\mathrm{d}X_t^x = b(X_t^x)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t, \qquad X_0^x = x \ge 0,$$

and let Z^x be its absorption at zero. Write $v(t,x) = \mathbb{E}Z_t^x$. The reflected process starts from x and satisfies

$$\mathrm{d}Y_t^x = b(Y_t^x)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t + \mathrm{d}K_t^x, \quad Y_t^x \ge 0, \quad \int_0^t \mathbf{1}_{\{Y_s^x > 0\}}\,\mathrm{d}K_s^x = 0, \tag{53}$$

where K^x is continuous, nondecreasing, and starts at zero. We prove

$$v(t,x) = \mathbb{E}X_t^x, \qquad \partial_x v(t,x) = \mathbb{E}\exp\left\{\int_0^t b'(Y_s^x)\,\mathrm{d}s\right\}. \tag{54}$$

Write τ_0 for the first hit of zero. Oddness and uniqueness imply that X^0 has a symmetric law and mean zero at every time. The strong Markov property therefore gives

$$\mathbb{E}[X_t^x \mathbf{1}_{\{\tau_0 \le t\}}] = \mathbb{E}[\mathbf{1}_{\{\tau_0 \le t\}}\mathbb{E}_0 X_{t-\tau_0}] = 0.$$

All variables are integrable by linear growth, proving the first identity. Applying (32) differentiates this whole-line mean. Since b' is even, its integral is unchanged when X^x is replaced by $|X^x|$. Tanaka's formula and the nondegenerate noise show that $|X^x|$ has the law of the reflected solution (53). This proves the second identity.

Reflected solutions driven by the same Brownian path preserve their initial order: a crossing requires a meeting, after which pathwise uniqueness makes them coalesce. Thus $Y_s^x \le Y_s^y$ for $0 \le x \le y$. Monotonicity of b' makes the integrand in (54) no smaller at x than at y. Consequently $\partial_x v(t,x)$ is positive and nonincreasing. Together with $v(t,0) = 0$, this proves concavity. The whole-line mean is odd, so its derivative is even. By ordered transport, transition distances are differences of these means; taking secants identifies the Lipschitz coefficient with $\sup_x g_t(x) = g_t(0)$. Thus, for $T > 0$ and $0 \le w \in C([0,T])$,

$$\mathrm{Lip}_{W_1}(P_t^b) = g_t(0), \qquad \sup_x \int_0^T w(T-t)g_t(x)\,\mathrm{d}t = \int_0^T w(T-t)\,\mathrm{Lip}_{W_1}(P_t^b)\,\mathrm{d}t. \tag{55}$$

□

Gradient formulas and diffusion normalization. For constant noise and no jumps, Cloez (2012) gives the fixed-process gradient formula used in (54): the auxiliary process is the original diffusion and the potential is minus its drift derivative. For noise coefficient σ, Bonnefont & Joulin (2014) specializes to the same formula by taking their generator coefficient $\sigma_{\mathrm{BJ}} = \sigma/\sqrt{2}$ and gradient weight $a_{\mathrm{BJ}} = 1$. The resulting potential $-b'$ is bounded for the smooth confining approximations; their Gaussian tails put the coordinate in the invariant Dirichlet-form domain.

Under a C^2 drift with bounded second derivative and a one-sided Lipschitz bound, Monmarché (2026) likewise identifies the one-dimensional W_1 coefficient with the synchronous linearized response, using the coordinate for equality. For odd b with b' nonincreasing on $[0,\infty)$, reflected order additionally places this maximum at zero for every elapsed time.

Smooth approximation preserving concavity. The following smooth approximations also give the sharpness construction in Section E. Fix $m > 0$, $A \geq 0$, and $L \geq 0$, and choose an even, nonnegative C_c^∞ probability density ρ, supported in $[-1, 1]$ and nonincreasing on $[0, \infty)$. Put $\rho_\varepsilon(x) = \varepsilon^{-1}\rho(x/\varepsilon)$ and

$$h_L(x) = \mathrm{clip}((m + L)x, -A, A), \qquad b_L^0(x) = -mx + h_L(x). \tag{56}$$

For $\varepsilon > 0$ and $0 \leq \theta < 1$, define

$$b_{\varepsilon,\theta}(x) = -mx + (1 - \theta)(h_L * \rho_\varepsilon)(x). \tag{57}$$

The residual is smooth, odd, nondecreasing, and bounded in absolute value by $(1 - \theta)A$. With $a = A/(m + L)$, differentiation gives

$$b'_{\varepsilon,\theta}(x) = -m + (1 - \theta)(m + L)(\mathbf{1}_{[-a,a]} * \rho_\varepsilon)(x).$$

The convolution factor is even, takes values in $[0, 1]$, and has derivative $\rho_\varepsilon(x + a) - \rho_\varepsilon(x - a) \leq 0$ for $x \geq 0$. Thus $b'_{\varepsilon,\theta}$ is even and nonincreasing on the positive half-line, with values in $[-m, L - \theta(m + L)]$. The first two derivatives are bounded for each fixed construction. In particular, the drift is regular and obeys the radial bound $2(1 - \theta)A - mr$ at separation r. Since h_L is $(m + L)$-Lipschitz and ρ_ε has support in $[-\varepsilon, \varepsilon]$,

$$\|b_{\varepsilon,\theta} - b_L^0\|_\infty \leq (m + L)\varepsilon + A\theta. \tag{58}$$

These are strict radial and one-sided margins when $A > 0$ and $\theta > 0$, including at $L = 0$. At $A = 0$ the residual vanishes and all drifts equal $-mx$.

Mean convergence and transfer of concavity. For the dimensionless comparison, apply this family with parameters $(m, L, A) = (1, \lambda, a_\chi)$ and $\theta = 0$. Its unsmoothed drift is the odd extension $\Gamma_\lambda(x) = -x + \mathrm{clip}((1 + \lambda)x, -a_\chi, a_\chi)$ of $\gamma_{\chi,\lambda}$. Since the even mollifier has zero first moment, the smoothed drift is exactly $\Gamma_\lambda * \rho_\varepsilon$. The first identity in (54) uses only oddness, linear growth, and uniqueness, so it also holds for the nonsmooth globally Lipschitz limit Γ_λ. The uniform drift error (58) and (31), with the common Lipschitz constant $\max\{1, \lambda\}$, give convergence of whole-line means uniformly on finite time intervals. These are also the absorbed means. Hence Lemma 3.1, applied at noise 2, transfers monotonicity and concavity to $v_{\chi,\lambda}(q, \cdot)$. For $\chi = 0$, this mean is $e^{-q}z$.

B.4 Limit as the one-sided Lipschitz bound increases

Fix $\chi \geq 0$ and write $z_\lambda = 2\sqrt{\chi}/(1 + \lambda)$ for the switching point. Set $\overline{k}_\chi(q) = \sup_{0 \leq s \leq q} k_\chi(s)$, which is finite by Lemma B.1. This bound controls the error after the comparison path enters $[0, z_\lambda]$.

Lemma B.2 (Convergence as the one-sided Lipschitz bound increases)**.** *For every fixed $q, z > 0$,*

$$v_{\chi,\lambda}(q, z) \uparrow v_\chi(q, z) \qquad (\lambda \to \infty). \tag{59}$$

For $q > 0$ and $z > z_\lambda$,

$$0 \leq v_\chi(q, z) - v_{\chi,\lambda}(q, z) \leq \overline{k}_\chi(q) z_\lambda. \tag{60}$$

Proof. On $(0, \infty)$ the scalar drifts $\gamma_{\chi,\lambda}$ increase pointwise to $2\sqrt{\chi} - z$; at zero every process remains absorbed. One-dimensional comparison gives monotonicity of the absorbed means. Couple the comparison processes before absorption using the same Brownian motion. If $z > z_\lambda$, their coefficients agree until the common path first reaches z_λ. After that time the process with drift $\gamma_{\chi,\lambda}$ remains below the one with drift γ_χ. The strong Markov property bounds the expected difference at time q by $\sup_{0 \leq s \leq q} v_\chi(s, z_\lambda) \leq \overline{k}_\chi(q) z_\lambda$. This proves (59)–(60).

□

Consequently,

$$k_{\chi,\lambda}(q) \uparrow k_\chi(q), \qquad \mathcal{I}_{\chi,\lambda}(Q) \uparrow \mathcal{I}_\chi(Q), \qquad \mathcal{I}_{\chi,\lambda}^\infty \uparrow \mathcal{I}_\chi^\infty. \tag{61}$$

Since $k_{\chi,\lambda} \leq k_\chi$, while for every $z > 0$

$$\liminf_{\lambda \to \infty} k_{\chi,\lambda}(q) \geq \lim_{\lambda \to \infty} \frac{v_{\chi,\lambda}(q, z)}{z} = \frac{v_\chi(q, z)}{z},$$

taking the supremum over z gives pointwise convergence of the Lipschitz coefficients. The integral statements follow from monotone convergence.

Concavity and boundary slopes. The shape-preserving approximation proves that $v_{\chi,\lambda}(q,\cdot)$ is increasing and concave. Its pointwise limit $v_\chi(q,\cdot)$ has the same properties by Lemma B.2. Both means vanish at zero. For either model, abbreviate its mean and coefficient by v, k; concavity makes $v(q,z)/z$ nonincreasing in z, and Lemma B.1 bounds it on each finite time interval. Hence

$$v(q,z) \le k(q)z, \qquad k(q) = \lim_{z\downarrow 0} \frac{v(q,z)}{z} < \infty. \tag{62}$$

The same vanishing separation therefore approaches the maximal quotient at every elapsed time. Monotone convergence gives, for bounded nonnegative measurable ω and $0 < Q \le \infty$,

$$\int_0^Q \omega(q)k(q)\,\mathrm{d}q = \lim_{z\downarrow 0} \int_0^Q \omega(q)\frac{v(q,z)}{z}\,\mathrm{d}q. \tag{63}$$

The integrals may initially be infinite; the Poisson argument below proves finiteness on the infinite horizon. The initial separation is fixed throughout each integral before the limit is taken.

B.5 Absorbed Poisson representation

The two scalar drifts agree for all sufficiently large positive arguments. A bounded-derivative solution of the Poisson equation identifies their occupation potentials using only finite-horizon moment control.

For either absorbed diffusion in (44), with $\chi, \lambda \ge 0$ and noise 2, write $Z^\gamma, v_\gamma, k_\gamma$ for its process, mean, and coefficient. Define the Poisson candidate by

$$\Psi_\gamma(z) = \frac{1}{2}\int_0^z \gamma(u)\,\mathrm{d}u,$$
$$H_\gamma(0) = 0, \qquad H_\gamma'(z) = \frac{1}{2}e^{-\Psi_\gamma(z)}\int_z^\infty y e^{\Psi_\gamma(y)}\,\mathrm{d}y. \tag{64}$$

Lemma B.3 (Poisson potential and integrated amplification)**.** *The derivative H_γ' is positive, continuous, bounded, and tends to one at infinity. The function H_γ is the unique C^2 solution of $2H_\gamma'' + \gamma H_\gamma' = -z$, $H_\gamma(0) = 0$, with bounded derivative, and*

$$H_\gamma(z) = \mathbb{E}_z\int_0^{\tau_0} Z_q^\gamma\,\mathrm{d}q = \int_0^\infty v_\gamma(q,z)\,\mathrm{d}q, \tag{65}$$
$$\int_0^\infty k_\gamma(q)\,\mathrm{d}q = H_\gamma'(0) = \frac{1}{2}\int_0^\infty z e^{\Psi_\gamma(z)}\,\mathrm{d}z < \infty. \tag{66}$$

Proof. For either drift, $\Psi_\gamma(z) = \sqrt{\chi}z - z^2/4 +$ constant for all sufficiently large z. The integral in (64) is finite; its integrating factor gives

$$(e^{\Psi_\gamma}H_\gamma')' = -\tfrac{1}{2}ze^{\Psi_\gamma}, \qquad 2H_\gamma'' + \gamma H_\gamma' = -z.$$

Writing H_χ for the affine potential, both derivatives agree on the common tail. Integration by parts gives

$$H_\chi'(z) = \frac{1}{2}e^{z^2/4-\sqrt{\chi}z}\int_z^\infty y e^{\sqrt{\chi}y - y^2/4}\,\mathrm{d}y = 1 + \sqrt{\chi}\,e^{z^2/4-\sqrt{\chi}z}\int_z^\infty e^{\sqrt{\chi}y-y^2/4}\,\mathrm{d}y. \tag{67}$$

It holds on all of $[0,\infty)$ in the affine case. For $z > 2\sqrt{\chi}$, the last tail ratio is at most $(z/2 - \sqrt{\chi})^{-1}$: put $y = z + s$ and drop the nonpositive term $-s^2/4$ in the exponent. Thus $H_\gamma'(z) \to 1$. Positivity and continuity on compact intervals prove boundedness, and $0 \le H_\gamma(z) \le \|H_\gamma'\|_\infty z$. The displayed ODE gives $H_\gamma \in C^2$, including the switching point of $\gamma_{\chi,\lambda}$. For two solutions, the difference has derivative $ce^{-\Psi_\gamma}$; boundedness at infinity forces $c = 0$, and the zero boundary value fixes the additive constant.

Put $C = \|H_\gamma'\|_\infty$. Extend H_γ as a C^2 function across zero and apply Itô's formula up to $t \wedge \tau_0 \wedge \tau_M$, where τ_M is the first hit of $M > z$. The stopped stochastic integral has quadratic variation at most $4C^2t$ and thus zero expectation. The boundary value $H_\gamma(0) = 0$ and absorption make the terminal

value vanish after τ_0. The bound (47), together with $0 \le H_\gamma(y) \le Cy$, gives uniform integrability of the terminal values as $M \to \infty$; the occupation terms converge by monotone convergence. Therefore, for every finite t,

$$H_\gamma(z) = \mathbb{E}_z H_\gamma(Z_t^\gamma) + \int_0^t v_\gamma(s,z)\,\mathrm{d}s. \tag{68}$$

Nonnegativity implies $\int_0^\infty v_\gamma(s,z)\,\mathrm{d}s \le H_\gamma(z) < \infty$. For each fixed z, choose deterministic $t_j \to \infty$ with $v_\gamma(t_j,z) \to 0$. Then

$$0 \le \mathbb{E}_z H_\gamma(Z_{t_j}^\gamma) \le C v_\gamma(t_j,z) \longrightarrow 0.$$

Sending $j \to \infty$ in (68) identifies H_γ with the integral of v_γ. Tonelli's theorem gives the occupation identity (65). Finally, (63) with $\omega \equiv 1$ and $Q = \infty$ gives $\int_0^\infty k_\gamma(q)\,\mathrm{d}q = \lim_{z\downarrow 0} H_\gamma(z)/z = H_\gamma'(0)$. This establishes both the value and finiteness in (66). □

Write $\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-u^2}\,\mathrm{d}u$ for the complementary error function. In the affine case, evaluating (67) at zero and completing the square gives

$$\mathcal{I}_\chi^\infty = H_\chi'(0) = 1 + \sqrt{\pi\chi}\, e^\chi \operatorname{erfc}(-\sqrt{\chi}). \tag{69}$$

Affine hitting-time interpretation. For $\chi > 0$, Tonelli applied to (52) gives the finite identity

$$H_\chi(z) = z + 2\sqrt{\chi}\,\mathbb{E}_z \tau_0. \tag{70}$$

Thus $\mathcal{I}_\chi^\infty$ is one plus $2\sqrt{\chi}$ times the boundary slope of the mean absorption time. Before absorption, $(Z_q - 2\sqrt{\chi})/2$ is a unit-noise Ornstein–Uhlenbeck process with hitting level $-\sqrt{\chi}$; see Alili et al. (2005) for hitting-time representations.

The case $\chi = 0$ and monotonicity in χ. At time zero, $v_\chi(0,z) = z$, so $k_\chi(0) = 1$. At $\chi = 0$, (52) gives $v_0(q,z) = e^{-q}z$ and $k_0(q) = e^{-q}$. Synchronous absorbed comparison also gives $v_{\chi_1} \le v_{\chi_2}$ for $\chi_1 \le \chi_2$, and therefore $k_{\chi_1} \le k_{\chi_2}$. Finally, $\operatorname{erfc}(-\sqrt{\chi}) \to 2$ in (69) yields $\mathcal{I}_\chi^\infty \sim 2\sqrt{\pi\chi}e^\chi$ as $\chi \to \infty$.

C Reflection Coupling and the Initial-Distance Bound

C.1 Radial reduction

Fix $0 \le s < t \le T$ and initial points x, y. If $x = y$, take identical paths. Otherwise, before meeting use reflection coupling:

$$\begin{aligned} \mathrm{d}X_u &= b_u(X_u)\,\mathrm{d}u + \sigma_u\,\mathrm{d}B_u, \\ \mathrm{d}Y_u &= b_u(Y_u)\,\mathrm{d}u + \sigma_u(I - 2e_u e_u^\top)\,\mathrm{d}B_u, \qquad e_u := \frac{X_u - Y_u}{\|X_u - Y_u\|}. \end{aligned}$$

The coefficients of this coupled equation are locally Lipschitz off the diagonal. Construct it up to lower separation cutoffs and upper spatial cutoffs. The reflection matrix is orthogonal, so its stochastic integral is Brownian and both coordinates have the prescribed marginal equations. Linear growth and bounded noise prevent explosion as the upper cutoffs increase. At a finite limiting meeting time τ, the stochastic integrals have continuous limits and $X_\tau = Y_\tau$; continue with synchronous increments. Pathwise uniqueness then keeps the coordinates together. For $D_u := \|X_u - Y_u\|$ and $u < \tau$, Itô's formula gives

$$\mathrm{d}D_u = \langle b_u(X_u) - b_u(Y_u), e_u\rangle\,\mathrm{d}u + 2\sigma_u\,\mathrm{d}W_u, \tag{71}$$

where W is one-dimensional Brownian motion. Under (2), the radial drift is at most $2A - mD_u$. With $\operatorname{osLip}(b) \le L$, it is at most $\min\{LD_u, 2A - mD_u\}$.

C.2 Backward comparison and radial cutoffs

For either mean $v = v_\chi$ or $v_{\chi,\lambda}$, with its drift γ from (44), localizing the linear terminal datum and using the finite-horizon moment bound (47) gives the classical backward equation on compact strips away from zero. For $q, z > 0$,

$$\partial_q v = 2\partial_{zz} v + \gamma(z)\partial_z v, \qquad v(q,0) = 0, \qquad v(0,z) = z. \tag{72}$$

The concavity and boundary-slope argument gives $v(q,z) \le k(q)z$ and the derivative signs used below.

Comparison bounds for Theorem 2.3 and Corollary 2.4. Use $\beta(r) = 2A - mr$ under the radial bound alone, or $\beta(r) = \min\{Lr, 2A - mr\}$ with the additional one-sided Lipschitz bound. For $u \in [s,t]$ and $r \ge 0$, define

$$F(u,r) := \frac{\sigma}{\sqrt{m}}\, v\left(m(t-u), \frac{\sqrt{m}}{\sigma} r\right). \tag{73}$$

By scaling (72),

$$\partial_u F + \beta(r)\partial_r F + 2\sigma^2 \partial_{rr} F = 0, \qquad F(t,r) = r, \qquad F(u,0) = 0. \tag{74}$$

By concavity and monotonicity, $F_r \ge 0$ and $F_{rr} \le 0$. For $\sigma_u \ge \sigma$, concavity controls the additional generator term $2(\sigma_u^2 - \sigma^2)F_{rr} \le 0$.

Apply Itô's formula to $F(u, D_u)$ up to the meeting time and standard lower/upper radial localizations. Using (71), the drift bound, and $\sigma_u \ge \sigma$,

$$\begin{aligned} &\partial_u F + \langle b_u(X_u) - b_u(Y_u), e_u\rangle F_r + 2\sigma_u^2 F_{rr} \\ &\le \partial_u F + \beta(D_u)F_r + 2\sigma^2 F_{rr} = 0. \end{aligned}$$

Stop at $\theta = t \wedge \tau_\varepsilon \wedge \tau^M$, where

$$\tau_\varepsilon = \inf\{u \ge s : D_u \le \varepsilon\}, \qquad \tau^M = \inf\{u \ge s : D_u \ge M\}. \tag{75}$$

On each compact strip away from zero, the stopped martingale has zero mean, so

$$\mathbb{E}F(\theta, D_\theta) \le F(s, D_s). \tag{76}$$

One may first stop before t and then use the continuous terminal datum. For fixed M, continuity and $F(u,0) = 0$ allow $\varepsilon \downarrow 0$; after meeting the copies are synchronized. Linear growth and bounded noise, applied separately to X and Y on this finite horizon, give

$$\mathbb{E} \sup_{s \le u \le t} D_u^2 < \infty. \tag{77}$$

Since $0 \le F(u,r) \le M_\chi r$ by Lemma B.1, this supplies uniform integrability as $M \to \infty$. With $F(t,r) = r$, we obtain

$$\mathbb{E}D_t \le F(s, \|x-y\|) = \frac{\sigma}{\sqrt{m}} v\left(m(t-s), \frac{\sqrt{m}}{\sigma}\|x-y\|\right). \tag{78}$$

The tangent bound (62) gives $\mathbb{E}D_t \le k(m(t-s))\|x-y\|$, proving (14) and, in the constrained case,

$$\mathrm{Lip}_{W_1}(P_{s,t}) \le \kappa_L(t-s) = k_{\chi,\lambda}(m(t-s)). \tag{79}$$

Lemma A.1 with $\omega = \Phi_{t-s}$ or Φ^L_{t-s} gives the corresponding bound for initial laws. □

Two-point modulus comparison. Andrews & Clutterbuck (2011) propagate $2\phi(|x-y|/2)$ for smooth Neumann solutions of $f_\tau = \Delta f + X_\tau \cdot \nabla f$ on bounded smooth strictly convex domains. Their modulus satisfies $\phi_\tau \ge \phi_{ss} + \omega\phi_s$, $\phi_s > 0$, and $\phi(0,\tau) \ge 0$, where $2\omega(s)$ bounds the radial drift at separation $2s$. For constant noise, the operator normalization is

$$\tau = \sigma^2 t/2, \qquad X_\tau = 2b_t/\sigma^2, \qquad \omega(s) = 2(A - ms)/\sigma^2.$$

For a backward evolution one first reverses time. The rescaled modulus $2\phi(\sigma^2 t/2, r/2)$ has operator $2\sigma^2\partial_{rr} + (2A - mr)\partial_r$. Their drift is already time-dependent and need not be a gradient. Here localization and concavity give the comparison on the whole space for unbounded tests and a positive noise lower bound. Absorption occurs when the coupled paths meet. The curvature specialization of Andrews & Ni (2012) uses $\partial_{ss} - as\partial_s$.

Writing κ_{E} for the curvature in Eberle (2016), our autonomous constant-noise bound gives $\kappa_{\mathrm{E}}(r) \ge 2(m - 2A/r)/\sigma^2$. It satisfies positive limiting curvature and $\int_0^1 r\kappa_{\mathrm{E}}(r)^-\,\mathrm{d}r < \infty$. Eberle et al. (2019) extend the contraction argument with Lyapunov weights.

Sticky coupling and spatially uniform drift error. Eberle & Zimmer (2019) dominate the distance of time-dependent unit-noise diffusions by a sticky scalar process with drift $M + r\ell(r)$, using ℓ for their radial coefficient. Here M bounds the pointwise drift discrepancy, and ℓ is Lipschitz and a negative constant outside a bounded interval. Dividing the radial bound by r gives $2A/r - m$, which diverges as $r \downarrow 0$ when $A > 0$. A fixed regular drift can use its spatial derivative bound near zero and the radial bound at large distances; that derivative bound is not uniform across the class.

D Optimal Coefficients for a Fixed Diffusion

Throughout, $T > 0$ and $0 \le w \in C([0,T])$.

D.1 Time-dependent perturbations

Fix a regular coefficient pair $(b, \sigma_\cdot)$ and $\mu_0 \in \mathcal{P}_1(\mathbb{R}^d)$. Let $\mathcal{F}$ consist of smooth bounded 1-Lipschitz functions with bounded derivatives. Clipping and mollification make $\mathcal{F}$ a determining class for W_1 on $\mathcal{P}_1$; spatial gradient norms are Euclidean. The next proposition extends the scalar coefficient $C_T(P, w)$ in Section 4.2 to any dimension.

Proposition D.1 (Optimal coefficient for a fixed diffusion)**.** *Over regular perturbed drifts with the same noise coefficient, common initialization, and finite $D_w(\widehat{\mu})$, the least nonnegative constant in $W_1(\widehat{\mu}_T, \mu_T) \le C D_w(\widehat{\mu})$ is*

$$C_T(P, w) = \sup_{f \in \mathcal{F}} \int_0^T w(t) \|\nabla P_{t,T} f\|_\infty \, \mathrm{d}t \le \int_0^T w(t) \operatorname{Lip}_{W_1}(P_{t,T}) \, \mathrm{d}t. \tag{80}$$

For $w \not\equiv 0$, the coefficient is approached by $\widehat{b}_t = b_t + \eta w(t) h_t$, with h_t of common compact spatial support, satisfying the regularity conditions of Section A.4, and fixed before $\eta \downarrow 0$. In dimension one,

$$C_T(P, w) = \int_0^T w(t) \|\partial_x P_{t,T} \, \mathrm{id} \, \|_\infty \, \mathrm{d}t = \int_0^T w(t) \operatorname{Lip}_{W_1}(P_{t,T}) \, \mathrm{d}t, \tag{81}$$

and nonnegative h_t suffice.

Proof. Retain the terminal test function in (35). For common initialization,

$$|(\widehat{\mu}_T - \mu_T) f| \le D_w(\widehat{\mu}) \int_0^T w(t) \|\nabla P_{t,T} f\|_\infty \, \mathrm{d}t.$$

Taking the supremum over $\mathcal{F}$ proves the first upper bound. Gradient–Wasserstein duality gives the second. Both integrands are bounded on this fixed horizon by the exponential bound from the spatial Lipschitz constant of b.

The case $w \equiv 0$ follows from the upper bound. For the lower bound with $w \not\equiv 0$, keep one $f \in \mathcal{F}$ fixed throughout time and write $\mu_t = \mu_0 P_{0,t}$. For every $t > 0$, μ_t has full support. Indeed, the additive Gaussian path $\int_0^t \sigma_s \, \mathrm{d}B_s$ has full support among continuous paths starting at zero. The solution map is continuous in that path and the initial point by Gronwall. A deterministic continuous control can follow any prescribed smooth path to any endpoint, which gives positive probability to every open endpoint set, also after averaging over μ_0.

Choose a countable family of nonnegative smooth bumps ρ_j supported in balls with rational centers and arbitrarily small radii, and a countable dense set of unit directions e_j. Include all bump–direction pairs. On any compact time interval inside $(0, T)$, set

$$v_{j,t}(x) = \frac{e_j \rho_j(x)}{\mu_t \rho_j}, \qquad a_j(t) = \mu_t\big(v_{j,t} \cdot \nabla P_{t,T} f\big).$$

The denominators are positive and continuous. Each a_j is continuous, and spatial localization gives $\sup_j a_j(t) = \|\nabla P_{t,T} f\|_\infty$. Therefore finite maxima approximate its weighted integral by dominated convergence. The continuous convex weights $\alpha_j(t) = e^{a_j(t)/\delta} / \sum_{i=1}^N e^{a_i(t)/\delta}$ satisfy

$$\sum_{j=1}^N \alpha_j(t) a_j(t) \ge \max_{j \le N} a_j(t) - \delta \log N.$$

Take $h_t = \sum_j \alpha_j(t) v_{j,t}$, multiplied by a continuous cutoff supported in $(0,T)$ and equal to one on the selected time interval. The discarded endpoint integrals tend to zero as that interval increases to $(0,T)$. Thus, for any $\varepsilon > 0$, a finite choice gives

$$d_h = \sup_{t:w(t)>0} \mu_t(\|h_t\|) \le 1, \qquad \int_0^T w(t)\mu_t\big(h_t \cdot \nabla P_{t,T} f\big)\,\mathrm{d}t \ge \int_0^T w(t)\|\nabla P_{t,T} f\|_\infty \,\mathrm{d}t - \varepsilon. \tag{82}$$

The cutoff can be chosen to vanish on neighborhoods of both endpoints. Hence all denominators used on its support are bounded away from zero; the resulting h_t has common compact spatial support and the regularity required in Section A.4.

Fix this field before choosing η. Equations (41) and (42) give the first-order expansions of the terminal difference and $D_w(\widehat{\mu}^\eta)$ for this f, h_t. A positive integral in (82) implies $d_h > 0$. Since $d_h \le 1$, dividing by $D_w(\widehat{\mu}^\eta)$ gives a limiting test-function ratio at least the selected integral. Choose f and the profile to approximate (80) before sending η to zero. If the coefficient is zero, the upper bound already proves the assertion.

In dimension one, ordered transport and (32) give $g_{t,T}(x) := \partial_x P_{t,T}\,\mathrm{id}(x) > 0$ and $\mathrm{Lip}_{W_1}(P_{t,T}) = \sup_x g_{t,T}(x)$. Apply the same localization directly to $g_{t,T}$ with direction $+1$. The resulting profiles are nonnegative, and the coordinate identity (43) and first-order estimates hold for every $\mathcal{P}_1$ initial law. Their limiting Wasserstein ratios approach $\int_0^T w(t)\|g_{t,T}\|_\infty \,\mathrm{d}t$. Together with the already proved upper bound in (80), this proves (81). □

In (80), the supremum over terminal test functions is outside the time integral; spatial profiles may change with time. Equality in its upper bound is equivalent to the existence of $f_n \in \mathcal{F}$ for which $\|\nabla P_{t,T} f_n\|_\infty \to \mathrm{Lip}_{W_1}(P_{t,T})$ for $w(t)\,\mathrm{d}t$-almost every t, after taking a subsequence if necessary. To see necessity, choose the integrated nonnegative deficits summable; sufficiency follows by domination.

Preserving the drift constraints. The proposition optimizes over regular perturbations. Its localized directions also preserve the drift constraints when the reference drift has strict margins. Sufficient radial margins are a bound $2(A-a) - mr$ with $a > 0$, or a radial bound $2A - (m+c)r$ with $c > 0$. The perturbation changes the radial difference by at most $2\eta\|w\|_\infty \sup_t \|h_t\|_\infty$ in the first case, and by at most $\eta\|w\|_\infty \sup_t \mathrm{Lip}(h_t) r$ in the second. An imposed bound L additionally requires a strict margin in $\mathrm{osLip}(b) \le L$. Small enough η preserves each inequality.

At a boundary point, the allowable directions depend on the constraint. For $b(x) = -mx$ in the class $A = 0$, a compactly supported scalar direction preserving the same radial bound must have $h_t' \le 0$ at every active time $w(t) > 0$, hence must vanish there. Constant translations remain admissible and attain the Ornstein–Uhlenbeck coefficient.

D.2 Fixed spatial perturbations

Use g_t, κ_b from Section 3.1 and $H_{b,w}$ from Section 3.2.

Proof of Lemma 3.2. Apply (43) and the estimates in Section A.4 with $h_t = h$ and $\mu_t = \pi_b$. Then $d_h = \pi_b h > 0$, while Fubini gives $\int_0^T w(t)\pi_b F_t \,\mathrm{d}t = \pi_b(hH'_{b,w})$. The response and D_w estimates yield

$$\begin{aligned} W_1(\widehat{\mu}_T^\eta, \pi_b) &= \eta\pi_b(hH'_{b,w}) + O(\eta^2), \\ D_w(\mu) = \eta\pi_b h, \qquad D_w(\widehat{\mu}^\eta) &= \eta\pi_b h + O(\eta^2). \end{aligned} \tag{83}$$

The remainder constants depend on the fixed b, h, w, T. Since $\pi_b h > 0$, both drift-error amplitudes are positive for all sufficiently small η, and division gives the two limits. The fixed-profile supremum follows by localizing a normalized bump near a nearly maximizing point, using continuity of $H'_{b,w}$ and positivity of the density. The bound $0 < g_t \le e^{T\|b'\|_\infty}$ permits differentiation under the time integral defining $H_{b,w}$; taking a spatial supremum then proves (19).

The time-dependent assertion is the scalar case of Proposition D.1, applied with initial law π_b. Its nonnegative profiles have common compact support and are fixed before the amplitude limit. Equation

(42) gives the same first-order expansion of D_w under μ and $\widehat{\mu}^\eta$, so these profiles also approach the last integral in (19) using $D_w(\mu)$. □

The strict margins above preserve the drift constraints for the time-dependent profiles as well.

Equality in (19) is equivalent to a sequence x_n with $g_t(x_n) \to \kappa_b(t)$ for $w(T-t)\,\mathrm{d}t$-almost every t, after passage to a subsequence if needed. Apply the argument following Proposition D.1 to $\kappa_b(t) - g_t(x_n)$; the bound $0 < g_t \le e^{T\|b'\|_\infty}$ allows dominated convergence.

Strict inequality for a contractive drift. The two profile optimizations can differ even with strict margins in the drift constraints. With unit noise, take

$$b(x) = -x + 1 + \frac{1}{4}\int_0^x e^{-u^2}\,\mathrm{d}u. \tag{84}$$

Here $b'(x) = -1 + e^{-x^2}/4 \le -3/4$, so b belongs to the radial class $m = 1/2$, $A = 0$ and adding a smooth bounded perturbation with derivative at most $1/4$ preserves the radial inequality. Its invariant density has Gaussian tails. For $w \equiv 1$ and all sufficiently small $T > 0$,

$$\left\| \int_0^T g_t\,\mathrm{d}t \right\|_\infty < \int_0^T \|g_t\|_\infty\,\mathrm{d}t. \tag{85}$$

Set $V = b'$ and $\mathcal{A} = \frac{1}{2}\partial_{xx} + b\partial_x + V$. The gradient formula gives $g_t = e^{t\mathcal{A}}1$, so iterated Duhamel expansion yields

$$g_t = 1 + tV + \frac{t^2}{2}K + O_{C_b^1}(t^3), \qquad K = \tfrac{1}{2}b''' + bb'' + (b')^2.$$

For $j \le 3$, $\mathcal{A}^j 1 \in C_b^1$: the terms are constants and products of bounded Gaussian primitives and polynomial–Gaussian factors. Feynman–Kac and the first flow derivative give

$$\|e^{s\mathcal{A}}u\|_{C_b^1} \le e^{(\|V\|_\infty + \|b'\|_\infty)s}(1 + s\|V'\|_\infty)\|u\|_{C_b^1}.$$

Applying this bound to the integral remainder proves the expansion.

Write $J_T(x) = \int_0^T g_t(x)\,\mathrm{d}t$. Then $J_T = T + \frac{1}{2}T^2V + O(T^3)$ uniformly in x. Since V has its unique maximum at zero and $\sup_{|x|\ge 1} V(x) < V(0)$, for sufficiently small T,

$$J_T(0) > \sup_{|x|\ge 1} J_T(x).$$

Thus J_T attains its maximum at a finite point x_T. Equality in (85) would make x_T maximize g_t for almost every $t \in (0,T)$. Along such times tending to zero, the first-order expansion gives $V(x_T) \ge V(0)$, hence $x_T = 0$. But $V'(0) = 0$ and $K'(0) = -1/2$, so

$$g_t'(0) = -t^2/4 + O(t^3) < 0$$

for all sufficiently small positive t. Thus zero cannot maximize those g_t, a contradiction.

D.3 Poisson equation and invariant averages

Let b be a regular autonomous scalar drift in the radial class, with constant noise coefficient σ. The radial condition with one point zero implies $b(x) \le b(0) + 2A - mx$ for $x > 0$ and $b(x) \ge b(0) - 2A - mx$ for $x < 0$. Hence the invariant density

$$p_b(x) = Z_b^{-1}\exp\left\{\frac{2}{\sigma^2}\int_0^x b(y)\,\mathrm{d}y\right\} \tag{86}$$

has Gaussian upper tails and all moments. The zero-flux equation gives invariance, and the contraction bound gives uniqueness. With $\bar{x} = \pi_b\,\mathrm{id}$, the centered Poisson solution for the mean position is

$$\begin{aligned} G_b(x) &= \int_0^\infty (P_t^b\,\mathrm{id}(x) - \bar{x})\,\mathrm{d}t, \\ G_b'(x) &= \int_0^\infty g_t(x)\,\mathrm{d}t = \frac{2}{\sigma^2 p_b(x)}\int_x^\infty (y-\bar{x})p_b(y)\,\mathrm{d}y. \end{aligned} \tag{87}$$

By Section C, $\kappa_b(t) \le k_\chi(mt)$, which is integrable by Lemma B.3. For every Lipschitz f,

$$|P_t^b f(x) - \pi_b f| \le \mathrm{Lip}(f)\kappa_b(t) \int |x-y| \pi_b(\,\mathrm{d}y).$$

This proves absolute and local uniform convergence of the centered Poisson operator $\mathcal{R}_b f = \int_0^\infty (P_t^b f - \pi_b f)\,\mathrm{d}t$, and $\pi_b \mathcal{R}_b f = 0$. The bound $0 < g_t \le \kappa_b(t)$ justifies differentiation for the coordinate. In particular, the weighted transition means from (17) satisfy

$$G_b(x) = \lim_{T\to\infty} \big(H_{b,1}(x) - T\bar{x}\big) = \int_0^\infty \big(P_t^b\,\mathrm{id}(x) - \bar{x}\big)\,\mathrm{d}t,$$
$$\|G_b' - H_{b,1}'\|_\infty \le \int_T^\infty \kappa_b(t)\,\mathrm{d}t \longrightarrow 0.$$

An integrable locally bounded submultiplicative coefficient decays exponentially: choose t_0 with $\kappa_b(t_0) < 1$ and write $t = nt_0 + s$, $0 \le s < t_0$. The finite-horizon Poisson equations therefore converge in distributions to

$$-\frac{\sigma^2}{2} G_b'' - bG_b' = x - \bar{x}, \qquad (p_b G_b')' = -\frac{2}{\sigma^2}(x-\bar{x})p_b.$$

Since $G_b \in C^1$, this equation makes G_b'' continuous. Bounded G_b' and the Gaussian tails give $p_b G_b' \to 0$ at infinity; integrating the second equation proves (87). A homogeneous solution has derivative c/p_b, so bounded derivative and centering determine G_b uniquely.

For $x < y$, ordered transport bounds $|P_t^b f(y) - P_t^b f(x)|$ by $\mathrm{Lip}(f)\int_x^y g_t(z)\,\mathrm{d}z$. Integrating in time and taking the coordinate as a matching test yields

$$\sup_{\mathrm{Lip}(f)\le 1} \mathrm{Lip}(\mathcal{R}_b f) = \|G_b'\|_\infty \le \int_0^\infty \kappa_b(t)\,\mathrm{d}t. \tag{88}$$

This is the unweighted infinite-horizon version of (19), with absolute convergence justified above. It is also the constant-noise, Euclidean-coordinate specialization of Djellout & Wu (2011); the last inequality bounds the norm of the integrated operator by the integral of the semigroup norms (Cheng et al., 2020). For the example (84), the gap $\int_0^\infty \kappa_b(t)\,\mathrm{d}t - \|G_b'\|_\infty$ is at least the positive finite-horizon gap in (85), since $\kappa_b(t) - g_t(x) \ge 0$. Thus (88) is strict.

For simultaneous bounded drift and diffusion perturbations, Birrell (2026) gives invariant-law W_1 bounds under log-Sobolev and concentration assumptions.

Constant forcing. For the regular autonomous radial drift b and noise σ fixed above, start both diffusions at π_b and set $\widehat{b} = b + \eta$, $\eta > 0$. This addition leaves pairwise drift differences and spatial derivatives unchanged, so both drifts satisfy the same radial and any imposed one-sided Lipschitz bounds. The discrepancy is identically η, giving $D_1(\mu) = D_1(\widehat{\mu}^\eta) = \eta$.

With $K_j = \|b^{(j)}\|_\infty$, the scalar response estimate (41) for $h = w = 1$ and (87) give

$$\left|\frac{W_1(\widehat{\mu}_T^\eta, \pi_b)}{\eta} - \pi_b(H_{b,1}')\right| \le \eta K_2 T^3 e^{3K_1 T},$$
$$|\pi_b(H_{b,1}' - G_b')| \le \int_T^\infty \kappa_b(t)\,\mathrm{d}t.$$

Taking $\eta \downarrow 0$ first and then $T \to \infty$ gives

$$\lim_{T\to\infty}\lim_{\eta\downarrow 0} \frac{W_1(\widehat{\mu}_T^\eta, \pi_b)}{\eta} = \pi_b(G_b').$$

Finally, boundedness of G_b' and the Gaussian upper tails of p_b imply $(x-\bar{x})p_b(x)G_b'(x) \to 0$ at both infinities. Integrating the centered Poisson equation by parts gives

$$\pi_b(G_b') = -\int_{\mathbb{R}} (x-\bar{x})(p_b G_b')'(x)\,\mathrm{d}x = \frac{2}{\sigma^2}\int_{\mathbb{R}} (x-\bar{x})^2 p_b(x)\,\mathrm{d}x = \frac{2\,\mathrm{Var}_{\pi_b}(X)}{\sigma^2}.$$

This spatial average is at most $\|G_b'\|_\infty$, the limiting optimal coefficient over fixed spatial profiles. Equality can hold, as for an Ornstein–Uhlenbeck drift. For the drifts below, this norm equals $\int_0^\infty \kappa_b(t)\,\mathrm{d}t$ because the derivatives attain their maxima at zero.

Odd drifts and absorption. If b is odd, then $\bar{x} = 0$. The strong Markov cancellation in the proof of Lemma 3.1 uses only oddness and uniqueness: the expected contribution after a first hit of zero vanishes. For $x \geq 0$, $P_t^b \,\mathrm{id}(x) = \mathbb{E}_x[X_t \mathbf{1}_{\{t<\tau_0\}}]$. Tonelli and the integrability established above therefore give

$$G_b(x) = \mathbb{E}_x \int_0^{\tau_0} X_t \,\mathrm{d}t, \qquad x \geq 0. \tag{89}$$

Thus the centered whole-line potential restricts to the absorbed occupation potential, with $G_b(0) = 0$ supplied by symmetry. Its boundary derivative is

$$G_b'(0) = \frac{2}{\sigma^2} \int_0^\infty y \exp\left\{ \frac{2}{\sigma^2} \int_0^y b(u)\,\mathrm{d}u \right\} \mathrm{d}y = \frac{\mathbb{E}_{\pi_b}|X|}{\sigma^2 p_b(0)}.$$

For any primitive $\Psi' = 2b/\sigma^2$, the exponential is $e^{\Psi(y)-\Psi(0)}$, as in (86). At noise 2, the equation $(e^{\Psi} G_b')' = -(x/2)e^{\Psi}$ recovers the coefficient and zero-boundary normalization in Lemma B.3, which also supplies integrability. That lemma treats the affine comparison directly: its odd extension is discontinuous at zero for $\chi > 0$, outside the regular whole-line setting. The smooth limits are in Proposition E.1.

If in addition b' is nonincreasing on $[0, \infty)$, the common maximizer in (55) makes (88) an equality at zero. This proves (28) with its shape assumption.

E Smooth Approximation and Sharpness

For $L < 0$, synchronous coupling gives $\mathrm{Lip}_{W_1}(P_{s,t}) \leq e^{L(t-s)}$. We consider $L \geq 0$ as in Corollary 2.4. Use the switching point z_λ from Lemma B.2, with χ fixed. The clipped-linear drifts and their smooth approximations prove sharpness for fixed L; a sequence with $L_n \to \infty$ proves sharpness under the radial bound alone.

E.1 Evaluation of the time integral

The dimensionless total integral is

$$\mathcal{I}_{\chi,\lambda}^{\infty} = \begin{cases} 1 + \sqrt{\pi\chi} + \chi, & \lambda = 0, \\ \left(1 + \lambda^{-1}\right) e^{\lambda\chi/(1+\lambda)^2} - \lambda^{-1} + \sqrt{\pi\chi}\, e^{\lambda\chi/(1+\lambda)} \operatorname{erfc}\left(-\dfrac{\lambda\sqrt{\chi}}{1+\lambda} \right), & \lambda > 0. \end{cases} \tag{90}$$

To evaluate it, Lemma B.3 reduces the calculation to a single integral. With $c = \lambda\chi/(1+\lambda)^2$, its exponent is

$$\Psi_{\chi,\lambda}(z) = \frac{1}{2} \int_0^z \gamma_{\chi,\lambda}(u)\,\mathrm{d}u = \begin{cases} \lambda z^2/4, & 0 \leq z \leq z_\lambda, \\ \sqrt{\chi} z - z^2/4 - \chi/(1+\lambda), & z \geq z_\lambda. \end{cases} \tag{91}$$

The two values agree at z_λ, where both equal c. The inner contribution is

$$\frac{1}{2} \int_0^{z_\lambda} z e^{\Psi_{\chi,\lambda}(z)}\,\mathrm{d}z = \begin{cases} \chi, & \lambda = 0, \\ (e^c - 1)/\lambda, & \lambda > 0. \end{cases}$$

Integration by parts and completion of the square give the outer contribution:

$$\begin{aligned} \frac{1}{2} \int_{z_\lambda}^\infty z e^{\Psi_{\chi,\lambda}(z)}\,\mathrm{d}z &= e^c + \sqrt{\chi}\, e^{-\chi/(1+\lambda)} \int_{z_\lambda}^\infty e^{\sqrt{\chi} z - z^2/4}\,\mathrm{d}z \\ &= e^c + \sqrt{\pi\chi}\, e^{\chi\lambda/(1+\lambda)} \operatorname{erfc}\left(-\frac{\lambda\sqrt{\chi}}{1+\lambda} \right). \end{aligned}$$

Adding the contributions proves (90), including $\lambda = 0$ and $\chi = 0$.

For the affine and clipped-linear comparisons, completing the square gives the tail exponents $\chi - (z - 2\sqrt{\chi})^2/4$ and $\chi\lambda/(1+\lambda) - (z - 2\sqrt{\chi})^2/4$. For fixed $\lambda > 0$ as $\chi \to \infty$, the common peak at $z = 2\sqrt{\chi}$ lies increasingly far beyond z_λ, while its width stays constant. The factor z in the Poisson integral accounts for the $\sqrt{\chi}$ prefactor. At $\lambda = 0$, the potential is flat on $[0, 2\sqrt{\chi}]$, giving the inner contribution χ, while the tail contributes $1 + \sqrt{\pi\chi}$.

E.2 Clipped-linear drifts and smooth approximation

The clipped-linear drift b_L^0 in (56) is odd and globally Lipschitz and satisfies both drift bounds. The symmetric reflection coupling from Section 3.2, started at x and $-x$ with $x > 0$, has distance $R_u = 2X_u$ satisfying

$$\mathrm{d}R_u = \min\{LR_u, 2A - mR_u\}\,\mathrm{d}u + 2\sigma\,\mathrm{d}W_u \tag{92}$$

until absorption. Ordered transport gives (16), and $x \downarrow 0$ gives the coefficient $\kappa_L(t)$. To approximate these values within the regular class while preserving room for a perturbation, we use the smoothing and residual shrinkage from Appendix B.

Fix $m, \sigma > 0$ and $A \geq 0$. Use the smoothing family $b_{\varepsilon,\theta}$ from (57), with constant noise σ. Its drifts are regular and odd, with derivative even and nonincreasing on $[0, \infty)$. For a selected sequence, write $b_n = -m\,\mathrm{id} + u_n$ and $\kappa_n(t) = \mathrm{Lip}_{W_1}(P_t^{b_n})$ as in Section 3.2.

Proposition E.1 (Smooth approximation with strict drift margins)**.** *For $A > 0$ and every fixed $L \geq 0$, the smoothing family admits a sequence with $0 < \theta_n < 1$, $\theta_n \to 0$, and*

$$\|u_n\|_\infty \leq (1 - \theta_n)A, \qquad \mathrm{osLip}(b_n) \leq L - \theta_n(m + L). \tag{93}$$

For every finite T,

$$\begin{aligned} \sup_{x\in\mathbb{R},\,0\leq t\leq T} W_1(\delta_x P_t^{b_n}, \delta_x P_t^{0,L}) &\longrightarrow 0, \\ \kappa_n(t) &\longrightarrow \kappa_L(t) \quad \textit{pointwise and in } L^1([0,\infty)). \end{aligned} \tag{94}$$

There is also a separate sequence with $L_n \to \infty$, satisfying (93) with L_n in place of L, for which

$$\kappa_n(t) \longrightarrow \kappa(t) \quad \textit{for every } t \geq 0, \qquad \int_0^\infty |\kappa_n(t) - \kappa(t)|\,\mathrm{d}t \longrightarrow 0. \tag{95}$$

Along this sequence, for every fixed $r, t \geq 0$, $W_1(\delta_{r/2}P_t^{b_n}, \delta_{-r/2}P_t^{b_n}) \to \Phi_t(r)$.

By Lemma 3.1, each sequence has $\kappa_n(t) = \partial_x P_t^{b_n}\,\mathrm{id}(0)$. The L^1 limits give convergence of every bounded measurable time-weighted integral, including on $[0, \infty)$. Taking $L \to \infty$ after the fixed-L limits gives the same integrals involving κ by (61). At $A = 0$, the constant sequence $b_n(x) = -mx$ gives all propagation limits exactly.

Proof. For fixed L, use $b_{\varepsilon,\theta}$ from (57), with semigroup $P^{\varepsilon,\theta}$ and noise coefficient σ. The constraints (93) follow from Appendix B. Both $b_{\varepsilon,\theta}$ and b_L^0 are $(m + L)$-Lipschitz. Combining (58) with (31) yields

$$\sup_{x\in\mathbb{R},\,0\leq t\leq T} W_1(\delta_x P_t^{\varepsilon,\theta}, \delta_x P_t^{0,L}) \leq \big((m + L)\varepsilon + A\theta\big)Te^{(m+L)T}. \tag{96}$$

Thus any $\varepsilon_n, \theta_n \downarrow 0$ with $\theta_n > 0$ gives the fixed-L transition convergence, including $L = 0$. For every fixed $x > 0$, the symmetric transition pair converges to the pair for b_L^0. Equation (16) and the comparison upper bound (79) give

$$\frac{\Phi_t^L(2x)}{2x} \leq \liminf_{n\to\infty} \kappa_n(t) \leq \limsup_{n\to\infty} \kappa_n(t) \leq \kappa_L(t). \tag{97}$$

Taking $x \downarrow 0$ after the approximation limit proves pointwise convergence of κ_n; at $t = 0$ both coefficients equal one. Since $0 \leq \kappa_n(t) \leq \kappa_L(t)$ and the latter is integrable by Lemma B.3, dominated convergence proves the L^1 assertion in (94).

For the sequence with $L_n \to \infty$, take, for $n \geq 1$,

$$L_n = mn, \qquad \varepsilon_n = \frac{\sigma}{\sqrt{m}}e^{-n^3}, \qquad \theta_n = e^{-n^3}, \qquad b_n = b_{\varepsilon_n,\theta_n} \quad \text{with } L = L_n.$$

For each fixed T, (96) bounds the synchronous error by

$$\big(\sigma\sqrt{m}(1 + n) + A\big)e^{-n^3}Te^{m(1+n)T} \longrightarrow 0.$$

Let $\Delta_n(T)$ denote the preceding synchronous-error bound, and fix $r > 0$. For all sufficiently large n, $\sqrt{m}\, r/\sigma > z_{L_n/m}$. The triangle inequality for the two transition laws and (60) give

$$\sup_{0\le t\le T}\left|W_1(\delta_{r/2}P_t^{b_n}, \delta_{-r/2}P_t^{b_n}) - \Phi_t(r)\right| \le 2\Delta_n(T) + \frac{\sigma}{\sqrt{m}} M_\chi z_{L_n/m} \longrightarrow 0.$$

The bound is uniform in time, including zero, and proves convergence for every fixed separation; $r = 0$ is immediate. Then the sandwich in (97), with Φ and κ, gives the boundary limit. Since κ is integrable by Lemma B.3, dominated convergence proves (95).

□

Invariant-density and moment limits. For the same sequences, the invariant densities at zero and first absolute moments converge to those of b_L^0 at fixed L, and to those of the density proportional to $\exp((-mx^2 + 2A|x|)/\sigma^2)$ as $L_n \to \infty$. Write $V_n(x) = 2\int_0^x b_n(y)\,\mathrm{d}y/\sigma^2$ and $p_n(x) = e^{V_n(x)}/\int e^{V_n(y)}\,\mathrm{d}y$. The residual bound gives

$$e^{V_n(x)} \le \exp\left(\frac{-mx^2 + 2A|x|}{\sigma^2}\right).$$

For fixed L, the log-density exponents converge locally uniformly to $V_L(x) = 2\int_0^x b_L^0(y)\,\mathrm{d}y/\sigma^2$. For the sequence with $L_n \to \infty$, the residuals converge to $A\,\mathrm{sgn}(x)$ away from zero, and bounded convergence in their primitives gives $V_n(x) \to (-mx^2 + 2A|x|)/\sigma^2$. Gaussian domination gives convergence of the normalizing constants, of $p_n(0)$, and of $\int |x| p_n(x)\,\mathrm{d}x$; it also gives convergence of densities weighted by $1 + |x|$ in L^1. Each limiting density is positive at zero. Thus the invariant ratios in (28) converge as well. With $z = 2\sqrt{m}\, x/\sigma$, the fixed-L limiting ratio is

$$\frac{2}{\sigma^2}\int_0^\infty x e^{V_L(x)}\,\mathrm{d}x = \frac{1}{2m}\int_0^\infty z e^{\Psi_{\chi,\lambda}(z)}\,\mathrm{d}z = \frac{\mathcal{I}^\infty_{\chi,\lambda}}{m}.$$

The limiting exponent $(-mx^2 + 2A|x|)/\sigma^2$ gives $\mathcal{I}^\infty_\chi/m$ by the same substitution. At $A = 0$, all residuals vanish and the assertions reduce to the Ornstein–Uhlenbeck drift.

Asymptotics for a fixed one-sided Lipschitz bound. For fixed $\lambda > 0$, the two exponential rates in (90) and $\mathrm{erfc}(-x) \to 2$ give

$$\mathcal{I}^\infty_{\chi,\lambda} \sim 2\sqrt{\pi\chi}\, e^{\chi\lambda/(1+\lambda)} \qquad (\chi \to \infty). \tag{98}$$

At $\lambda = 0$, (90) gives $1 + \sqrt{\pi\chi} + \chi$. For fixed χ, monotonicity and convergence to $\mathcal{I}^\infty_\chi$ as $\lambda \to \infty$ follow from Lemma B.2.

E.3 Optimality for a prescribed initial distance

The bounds with and without a prescribed L are proved together in Section C. Let $\tau = t - s \ge 0$.

For $r > 0$, take $\mu = \delta_{r/2}$ and $\nu = \delta_{-r/2}$, whose distance is exactly r. Under the bound L, the clipped-linear drift attains $\Phi^L_\tau(r)$, and the fixed-L sequence in Proposition E.1 approaches this value within $\mathcal{D}^1_T(m, A, \sigma; L)$. The sequence with $L_n \to \infty$ approaches $\Phi_\tau(r)$. This proves class optimality at each prescribed distance; the cases $r = 0$ or $\tau = 0$ are immediate.

Finally, the absorbed mean identity in (52) and the survival bound by one give

$$e^{-m\tau} r \le \Phi_\tau(r) \le e^{-m\tau} r + \frac{2A}{m}(1 - e^{-m\tau}), \qquad \lim_{r\to\infty}\frac{\Phi_\tau(r)}{r} = e^{-m\tau}. \tag{99}$$

The same bounds and limiting ratio hold for Φ^L_τ: comparison with drift $-mr$ gives the lower bound, and comparison with Φ_τ gives the upper bound. The affine upper bound is also the synchronous radial estimate, obtained from $D' \le -mD + 2A$ before meeting.

E.4 Drift-error sharpness

For the $A > 0$, $w \not\equiv 0$ construction in Section 3.2, it remains to check initialization and admissibility at the selected amplitude. Proposition E.1 supplies the L^1 approximation and strict margins: (93) gives $\theta_n A > 0$ and $\theta_n(m+L) > 0$, including at $L = 0$; the sequence with $L_n \to \infty$ uses its selected finite L_n. The invariant law is symmetric, has positive density, and has Gaussian tails by Section D.3, so it meets the initialization hypotheses of Lemma 3.2. For the selected smooth even bump $h \geq 0$, normalized by $\pi_b h = 1$, choose $\eta > 0$ so that

$$\eta\|w\|_\infty\|h\|_\infty \leq A - \|b + m\,\mathrm{id}\,\|_\infty, \qquad \eta\|w\|_\infty\|h'\|_\infty \leq L - \mathrm{osLip}(b)$$

when L is imposed; otherwise only the first inequality is needed. The strict margins make the right-hand sides positive, including at $L = 0$. Once the normalized h is fixed, its norms are finite, so every sufficiently small $\eta > 0$ preserves the imposed drift constraints for $b + \eta w(t)h$.

If $w \equiv 0$, finite $D_w(\widehat{\mu})$ forces zero expected discrepancy and Duhamel gives zero error under common initialization. Exchanging the diffusions proves the same assertion for $D_w(\mu)$ when both drifts satisfy the constraints.

The $A = 0$ case of Theorem 2.3 and Corollary 2.4. For $A = 0$ and $w \not\equiv 0$, take $b(x) = -mx$ and $\widehat{b}_t(x) = -mx + \eta w(t)$, started at $\pi_b = \mathcal{N}(0, \sigma^2/(2m))$. Both drifts satisfy the radial bound and have one-sided Lipschitz coefficient $-m \leq L$. Their synchronous difference is the deterministic translation

$$\widehat{X}^\eta_T - X_T = \eta \int_0^T e^{-m(T-r)} w(r)\,\mathrm{d}r.$$

For $\eta > 0$, $D_w(\mu) = D_w(\widehat{\mu}) = \eta$. Since $k_0(q) = k_{0,\lambda}(q) = e^{-q}$, the constants are attained exactly. □

Sharpness in each dimension. The scalar examples extend to any fixed $d \geq 1$. Use

$$b^{(d)}(x) = \big(b(x_1), -mx_2, \ldots, -mx_d\big), \qquad h^{(d)}(x) = \big(h(x_1), 0, \ldots, 0\big),$$

with noise σI_d and independent Ornstein–Uhlenbeck coordinates after the first. If $r = \|x - y\|$, the scalar radial bound gives

$$\langle b^{(d)}(x) - b^{(d)}(y), x - y\rangle \leq -mr^2 + 2A|x_1 - y_1| \leq -mr^2 + 2Ar.$$

The same argument applies to the perturbed drift; an imposed one-sided bound is inherited because the other coordinates have derivative $-m \leq L$. For initial-error sharpness, start at $\pm(r/2)e_1$. For drift-error sharpness, use the product of the scalar invariant law and $\mathcal{N}(0, \sigma^2/(2m))^{\otimes(d-1)}$. At each time the two laws have the same independent remaining coordinates. Projection onto the first coordinate and coupling the remaining coordinates identically show that their Euclidean W_1 distance equals the scalar distance. Both drift-error amplitudes also equal their scalar counterparts. The scalar constructions therefore prove sharpness in every fixed dimension.

F Peak Amplification and Accumulation Time

Finite-time bounds. Use M_χ from Lemma B.1. Integrating (48) and the uniform bound in (49) gives, for $Q \geq 0$,

$$\mathcal{I}_\chi(Q) \leq \min\left\{Q + \frac{4}{3}\sqrt{\frac{2\chi}{\pi}}Q^{3/2} + \chi Q^2,\ M_\chi Q\right\}. \tag{100}$$

The same bound holds for $\mathcal{I}_{\chi,\lambda}$. Substitution of $Q = mT$ and $\chi = A^2/(m\sigma^2)$ into (20) gives, for every $T \geq 0$,

$$C^\star_T(1) \leq T + \frac{4}{3}\sqrt{\frac{2}{\pi}}\frac{A}{\sigma}T^{3/2} + \frac{A^2}{\sigma^2}T^2. \tag{101}$$

The alternative estimate is $M_\chi T$; both also bound $C^\star_{T,L}(1)$ uniformly in L.

For $\alpha \in (0, 1)$, define the accumulation times by

$$\mathcal{I}_\chi(Q_\alpha(\chi)) = \alpha\mathcal{I}^\infty_\chi, \qquad \mathcal{I}_{\chi,\lambda}(Q_\alpha(\chi,\lambda)) = \alpha\mathcal{I}^\infty_{\chi,\lambda}. \tag{102}$$

For either scalar model, abbreviate its finite and infinite integrals and accumulation time by I, I_∞, Q_α.

We prove Proposition 4.1 together with the fixed-window estimates and constrained variants used in Section 4.1.

Asymptotics on fixed time intervals. As $\chi \to \infty$,

$$\sup_{0\le q\le Q}\left|\frac{k_\chi(q)}{2\chi} - (1-e^{-q})\right| = O_Q(\chi^{-1/2}) \tag{103}$$

for every fixed $Q < \infty$. For every fixed $\lambda \ge 0$,

$$\sup_{0\le q\le Q}\left|k_{\chi,\lambda}(q) - e^{\lambda q}\right| \longrightarrow 0 \qquad (\chi\to\infty). \tag{104}$$

Accumulation time. For $\alpha \in (0,1)$,

$$\frac{\alpha I_\infty}{M_\chi} \le Q_\alpha \le I_\infty \log\frac{1}{1-\alpha}. \tag{105}$$

For each fixed $\alpha \in (0,1)$, as $\chi \to \infty$,

$$\begin{aligned} Q_\alpha(\chi,\lambda) &\sim \frac{\pi}{\sqrt{\lambda}} e^{\chi\lambda/(1+\lambda)} \log\frac{1}{1-\alpha} \quad (\lambda>0 \text{ fixed}),\\ Q_\alpha(\chi,0) &= \Theta_\alpha(\chi). \end{aligned} \tag{106}$$

Integrating the uniform fixed-window estimates gives $\mathcal{I}_\chi(Q)/(2\chi) \to Q - 1 + e^{-Q}$ and, for fixed $\lambda \ge 0$,

$$\mathcal{I}_{\chi,\lambda}(Q) \longrightarrow \int_0^Q e^{\lambda q}\,\mathrm{d}q = \begin{cases} (e^{\lambda Q}-1)/\lambda, & \lambda>0,\\ Q, & \lambda = 0. \end{cases}$$

Keeping m, σ, L, T fixed and sending $A \to \infty$, rescaling by $Q = mT$ gives

$$C^\star_{T,L}(1) \longrightarrow \begin{cases} (e^{LT}-1)/L, & L>0,\\ T, & L=0. \end{cases} \tag{107}$$

Thus moving the onset of contraction outward makes the one-sided Lipschitz estimate sharp on a fixed horizon. This limit keeps L fixed; the coefficient without a prescribed L optimizes over arbitrarily large local expansion.

F.1 Peak height and fixed-window asymptotics

For the shifted Ornstein–Uhlenbeck process, Giorgini et al. (2020) analyze the short-time boundary layer and long survival tail by matched asymptotics. To control the boundary slope, the bounded subsolution below gives a survival inequality at every positive separation before we send that separation to zero.

Proof of the peak bound in Proposition 4.1 and (103). Fix $\chi > 1$, put $c_\chi = \sqrt{\chi} - 1 > 0$, and set

$$f(z) = 1 - e^{-c_\chi z}, \qquad \delta_\chi = \frac{1}{e(e^{2c_\chi}-1)}, \qquad \mathcal{L}_\chi = 2\partial_{zz} + (2\sqrt{\chi} - z)\partial_z.$$

Direct differentiation gives $\mathcal{L}_\chi f = c_\chi(2-z)e^{-c_\chi z}$. This is nonnegative for $0 \le z \le 2$. For $z \ge 2$, let $u = c_\chi(z-2) \ge 0$; then

$$-\mathcal{L}_\chi f = u e^{-2c_\chi - u} \le e^{-2c_\chi - 1}, \qquad f(z) \ge 1 - e^{-2c_\chi}.$$

Thus $\mathcal{L}_\chi f \ge -\delta_\chi f$ on $z > 0$. Apply Itô to $e^{\delta_\chi(s\wedge\tau_0)} f(Z_{s\wedge\tau_0})$ and stop also at an upper radius. The functions f, f' and $\mathcal{L}_\chi f$ are bounded; nonexplosion and bounded convergence remove the upper stop, and the stochastic integral is square-integrable on a fixed horizon. Since $f(0) = 0$, the result is

$$\mathbb{P}_z(\tau_0 > s) \ge \mathbb{E}_z f(Z_s) \ge e^{-\delta_\chi s}(1 - e^{-c_\chi z}). \tag{108}$$

Insert this bound in (52), divide by z, and let $z \downarrow 0$. The factor $(1 - e^{-c_\chi z})/z$ tends to c_χ, so

$$\begin{aligned} k_\chi(q) &\geq e^{-q} + 2\sqrt{\chi}\, c_\chi \int_0^q e^{-(q-s)} e^{-\delta_\chi s}\, \mathrm{d}s \\ &\geq e^{-q} + 2(\chi - \sqrt{\chi}) e^{-\delta_\chi q}(1 - e^{-q}). \end{aligned} \tag{109}$$

The bound also holds at $q = 0$.

Combining this lower bound with (49), and using $\int_0^q e^{-(q-s)} s^{-1/2}\, \mathrm{d}s \leq 2$, gives, for every $q \geq 0$,

$$\begin{aligned} e^{-q} + 2(\chi - \sqrt{\chi}) e^{-\delta_\chi q}(1 - e^{-q}) &\leq k_\chi(q) \\ &\leq e^{-q} + 2\chi(1 - e^{-q}) + 2\sqrt{2\chi/\pi}. \end{aligned} \tag{110}$$

On a fixed window $[0, Q]$, divide (110) by 2χ. Since $e^{-\delta_\chi q} = 1 + O_Q(\delta_\chi)$ and $\delta_\chi = O(e^{-2\sqrt{\chi}})$, both deviations from $1 - e^{-q}$ are $O_Q(\chi^{-1/2})$, proving the uniform rate and, by integration, (23).

For the global peak, evaluate (109) at $q = \log \chi$. The facts $1 - e^{-q} = 1 - \chi^{-1}$ and $\chi \delta_\chi \log \chi = o(1)$ show

$$\sup_q k_\chi(q) \geq 2\chi - 2\sqrt{\chi} - O(1).$$

The upper bound $M_\chi = 2\chi + O(\sqrt{\chi})$ proves the peak assertion in (21). The total asymptotic follows from (69). □

Proof of (104). Fix $Q > 0$ and $\lambda \geq 0$. Since $\gamma_{\chi,\lambda}(z) \leq \lambda z$, stopped Gronwall gives $\mathbb{E}_z Z_q^{\chi,\lambda} \leq e^{\lambda q} z$, hence $k_{\chi,\lambda}(q) \leq e^{\lambda q}$. Let U solve $\mathrm{d}U_q = \lambda U_q\, \mathrm{d}q + 2\, \mathrm{d}W_q$ from $U_0 = 1$, and let $\widetilde{U}$ be its absorption at zero. The absorbed mean is still $\mathbb{E}\widetilde{U}_q = e^{\lambda q}$: its stopped noise has mean zero, while its drift is $\lambda \widetilde{U}_q$.

Put $R = 2\sqrt{\chi}/(1 + \lambda)$ and take χ large enough that $R > 1$. Couple $\widetilde{U}$ and $Z^{\chi,\lambda}$ from 1 using the same Brownian motion. They coincide until hitting zero or R, and scalar comparison gives $0 \leq Z_q^{\chi,\lambda} \leq \widetilde{U}_q$. For $M_Q = \sup_{q \leq Q} \widetilde{U}_q$, it follows pathwise that

$$\sup_{q \leq Q} |\widetilde{U}_q - Z_q^{\chi,\lambda}| \leq M_Q \mathbf{1}_{\{M_Q \geq R\}}.$$

The explicit solution $U_q = e^{\lambda q}(1 + 2\int_0^q e^{-\lambda s}\, \mathrm{d}W_s)$ and Doob's inequality give $\mathbb{E}M_Q \leq e^{\lambda Q}(1 + 4\sqrt{Q}) < \infty$. Since $k_{\chi,\lambda}(q) \geq v_{\chi,\lambda}(q, 1)$,

$$0 \leq \sup_{q \leq Q} \big(e^{\lambda q} - k_{\chi,\lambda}(q)\big) \leq \mathbb{E}[M_Q \mathbf{1}_{\{M_Q \geq R\}}] \longrightarrow 0.$$

This proves (104), including $\lambda = 0$; the case $Q = 0$ is immediate. □

F.2 Accumulation time

Proof of the finite-parameter bounds (105) *and the* $L = 0$ *order.* Fix either scalar model and abbreviate its absorbed mean, coefficient, and integrals by v, k, I, I_∞. Both scalar drifts dominate $-z$, so $k(q) \geq e^{-q}$. The function I is therefore continuous and strictly increasing to the finite positive I_∞. Measurability of k follows by taking its supremum over positive rational initial points. Consequently each Q_α in (102) exists and is unique.

The absorbed Markov property and $v(r, 0) = 0$ give

$$v(q + r, z) = \mathbb{E}_z v(r, Z_q) \leq k(r) v(q, z).$$

Division by z and a supremum prove submultiplicativity. For $F(Q) = \int_Q^\infty k(q)\, \mathrm{d}q$, integration yields $F(Q) \leq k(Q) I_\infty$. Since F is absolutely continuous and $F' = -k$ almost everywhere, the integrating-factor argument gives

$$k(q + r) \leq k(q)k(r), \qquad I_\infty - I(Q) \leq I_\infty e^{-Q/I_\infty}. \tag{111}$$

At $Q = Q_\alpha$, the peak bound $I(Q) \le M_\chi Q$ gives the lower estimate in (105); the tail estimate gives the upper one. At $\lambda = 0$, the bound $k_{\chi,\lambda}(q) \le e^{\lambda q}$ proved above gives $k_{\chi,0}(q) \le 1$. The peak argument therefore improves to

$$\alpha \mathcal{I}^\infty_{\chi,0} \le Q_\alpha(\chi, 0) \le \mathcal{I}^\infty_{\chi,0} \log \frac{1}{1-\alpha}.$$

Since (90) gives $\mathcal{I}^\infty_{\chi,0} = 1 + \sqrt{\pi\chi} + \chi \sim \chi$, this proves the $L = 0$ assertion in (106) for fixed α. In original time, $T_\alpha = Q_\alpha/m$. □

Accumulation as an absorption probability. For either drift γ in (44), let P_q denote the semigroup killed at zero, so $P_q 1(z) = \mathbb{P}_z(\tau_0 > q)$ and $P_q\,\mathrm{id}(z) = v(q, z)$. Use Ψ_γ from (64) and set

$$\begin{aligned} \rho(z) = e^{\Psi_\gamma(z)}, \qquad N_j = \int_0^\infty z^j \rho(z)\,\mathrm{d}z \quad (j = 0, 1), \\ \pi(\,\mathrm{d}z) = \frac{\rho(z)\,\mathrm{d}z}{N_0}, \qquad \eta(\,\mathrm{d}z) = \frac{z\rho(z)\,\mathrm{d}z}{N_1}. \end{aligned} \tag{112}$$

All moments of these weights are finite. The probability π symmetrizes the killed semigroup and is invariant for the reflected diffusion. Its size-biased law η initializes the auxiliary absorbed process.

Lemma F.1 (Integrated boundary flux)**.** *For either scalar comparison, all $\chi \ge 0$ (and $\lambda \ge 0$ with the one-sided Lipschitz bound), and every $Q \ge 0$,*

$$\frac{I(Q)}{I_\infty} = \mathbb{P}_\eta(\tau_0 \le Q). \tag{113}$$

Proof. The generator is $G = 2\partial_{zz} + \gamma\partial_z = 2\rho^{-1}(\rho\partial_z)'$. In $L^2(\pi)$, let $\mathcal{H}_D$ be the closure of $C_c^\infty(0, \infty)$ for the norm whose square is

$$\|f\|^2_{L^2(\pi)} + 2\int_0^\infty |f'(z)|^2 \pi(\,\mathrm{d}z).$$

The closed form $\mathcal{E}(f, g) = 2\int f'g'\,\mathrm{d}\pi$ defines the nonnegative self-adjoint operator $-G$ with zero trace at zero. There is no extra boundary condition at infinity. On $(0, R)$, integration by parts and the Dirichlet backward equation identify the form semigroup with the diffusion killed at both endpoints. As $R \to \infty$, the form resolvents converge by energy minimization over an increasing union dense in $\mathcal{H}_D$, and the stopped expectations converge by nonexplosion. Thus the half-line semigroup is P_q and is self-adjoint. Here each drift is globally Lipschitz, and (47) supplies nonexplosion.

For $f(z) = z \in L^2(\pi)$, define the Bochner integral $u_Q = \int_0^Q P_q f\,\mathrm{d}q$. Strong continuity gives

$$\frac{P_h u_Q - u_Q}{h} = \frac{1}{h}\left(\int_Q^{Q+h} P_q f\,\mathrm{d}q - \int_0^h P_q f\,\mathrm{d}q\right) \longrightarrow P_Q f - f \quad \text{in } L^2(\pi).$$

Thus u_Q lies in the generator domain and satisfies

$$G u_Q = P_Q f - f, \qquad 2(\rho u_Q')' = \rho(P_Q f - f).$$

The latter equality first holds weakly and gives a locally absolutely continuous flux. By concavity and (63), $u_Q(0) = 0$, $u_Q'(0+) = I(Q)$, and $0 \le u_Q'(z) \le I(Q)$. Thus $\rho(z)u_Q'(z) \to 0$ at infinity. Integrating the flux identity and using $\rho(0) = 1$ yields

$$2I(Q) = \int_0^\infty \rho(z)\{f(z) - P_Q f(z)\}\,\mathrm{d}z.$$

Both integrals are finite by Cauchy–Schwarz. The functions f and 1 lie in $L^2(\pi)$; self-adjointness of the semigroup requires no Dirichlet trace condition on either test. Thus

$$\int \rho P_Q f = \int \rho f P_Q 1 = \int_0^\infty z\rho(z)\mathbb{P}_z(\tau_0 > Q)\,\mathrm{d}z.$$

The Poisson identity (66) gives $2I_\infty = N_1$, proving (113). □

The exponential accumulation profile. In the affine case, $(a_\chi - Z_q)/2$, with $a_\chi = 2\sqrt{\chi}$, is a unit-noise Ornstein–Uhlenbeck process with absorption level $\sqrt{\chi}$. Its hitting-time transform solves the classical backward boundary problem (Alili et al., 2005). In the normalization of Giorgini et al. (2020), the coordinate is $(a_\chi - Z_q)/\sqrt{2}$ and the level is $\sqrt{2\chi}$; their outer survival rate is T_χ^{-1}. The following proof obtains the cumulative limit under the law in (112), and treats a fixed positive one-sided Lipschitz bound by the same argument.

Proposition F.2 (Exponential accumulation)**.** *Let T_χ be defined in (22). As $\chi \to \infty$,*

$$\sup_{s\ge 0}\left|\frac{\mathcal{I}_\chi(T_\chi s)}{\mathcal{I}_\chi^\infty} - (1-e^{-s})\right| \longrightarrow 0. \tag{114}$$

For every fixed $\lambda > 0$, define

$$T_{\chi,\lambda} := \frac{\pi}{\sqrt{\lambda}} e^{\chi\lambda/(1+\lambda)}. \tag{115}$$

Then (114) also holds with $(\mathcal{I}_{\chi,\lambda}, \mathcal{I}_{\chi,\lambda}^\infty, T_{\chi,\lambda})$ in place of $(\mathcal{I}_\chi, \mathcal{I}_\chi^\infty, T_\chi)$. Consequently (22) and the $\lambda > 0$ case of (106) hold uniformly for α in any compact subset of $(0,1)$. This limit keeps λ fixed.

Proof. Use (112) for each member of either family. Recall $a_\chi = 2\sqrt{\chi}$ and choose $0 < c = c_\chi < a_\chi$ as specified below. Set

$$S(c) = \int_0^c \rho(z)^{-1}\,\mathrm{d}z, \qquad \widehat{T} = \frac{N_0 S(c)}{2}. \tag{116}$$

This is the reciprocal of a finite-interval capacity. Keeping π normalized on the whole half-line, restrict only its energy to $(0,c)$:

$$\inf_{\substack{f\in H^1(0,c)\\ f(0)=0,\ f(c)=1}} 2\int_0^c |f'(z)|^2 \pi(\,\mathrm{d}z) = \frac{2}{N_0 S(c)} = \widehat{T}^{-1}. \tag{117}$$

The endpoint values are Sobolev traces. Since ρ is continuous and positive on $[0,c]$, Cauchy–Schwarz gives $1 = (\int_0^c f')^2 \le S(c)\int_0^c |f'|^2\rho$. Equality holds for the unique minimizer $f(z) = S(c)^{-1}\int_0^z \rho(u)^{-1}\,\mathrm{d}u$. The mass N_0 and boundary-to-c integral $S(c)$ thus determine this variational quantity exactly; its role as an asymptotic absorption-time scale follows from the concentration estimates below.

We first prove the limit on scale $\widehat{T}$ from the following estimates, then verify them and evaluate $\widehat{T}$ for each drift:

$$\pi([0,c]) \longrightarrow 0, \qquad \frac{1}{S(c)}\int_c^{a_\chi+K} \rho(z)^{-1}\,\mathrm{d}z \longrightarrow 0 \quad \text{for every fixed } K>0, \tag{118}$$
$$\lim_{K\to\infty}\limsup_{\chi\to\infty} \pi((a_\chi+K,\infty)) = 0, \qquad \left\| \frac{z}{\int u\pi(\,\mathrm{d}u)} - 1 \right\|_{L^2(\pi)} \longrightarrow 0.$$

Laplace-transform convergence. Fix $s > 0$. The transform $h(z) = \mathbb{E}_z e^{-s\tau_0/\widehat{T}}$ is bounded between zero and one, has $h(0) = 1$, and solves $Gh = (s/\widehat{T})h$. This follows from the stopped backward equation on bounded intervals and exhaustion; boundedness and local ODE regularity pass the equation to the half-line. In divergence form, $(\rho h')' = s\rho h/(2\widehat{T})$. Its right side is integrable, so $\rho h'$ has a finite limit at infinity. A nonzero limit would contradict boundedness of h, since $\int^\infty \rho^{-1} = \infty$. Hence

$$h'(u) = -\frac{s}{2\widehat{T}\rho(u)}\int_u^\infty \rho(y)h(y)\,\mathrm{d}y. \tag{119}$$

Write $\bar{h} = \int h\,\mathrm{d}\pi$. Integrating from zero to c, and using $N_0 S(c)/(2\widehat{T}) = 1$, gives

$$h(c) = 1 - s\bar{h} + \varepsilon_\chi, \qquad 0 \le \varepsilon_\chi \le s\pi([0,c]). \tag{120}$$

Indeed, subtract $\int_0^u \rho h$ from $N_0\bar{h}$ inside the inner integral of (119); the resulting error is at most $s\pi([0,c])$. The same identity gives

$$\sup_{c\le z\le a_\chi+K} |h(z) - h(c)| \le \frac{s}{S(c)}\int_c^{a_\chi+K} \rho(u)^{-1}\,\mathrm{d}u. \tag{121}$$

Split $\bar{h} - h(c)$ over $[0, c]$, $[c, a_\chi + K]$, and $(a_\chi + K, \infty)$. By boundedness and (118)–(121),

$$\limsup_{\chi\to\infty} |\bar{h} - h(c)| \le \limsup_{\chi\to\infty} \pi((a_\chi + K, \infty)).$$

Taking $\chi \to \infty$ first and then $K \to \infty$ shows $\bar{h} - h(c) \to 0$. Equation (120) now gives $(1+s)\bar{h} \to 1$. Finally,

$$\left| \int h \, \mathrm{d}\eta - \bar{h} \right| \le \left\| \frac{z}{\int u\pi(\,\mathrm{d}u)} - 1 \right\|_{L^2(\pi)} \longrightarrow 0.$$

Thus $\mathbb{E}_\eta e^{-s\tau_0/\widehat{T}} \to (1+s)^{-1}$ for every $s > 0$.

Affine drift. Here $\rho(z) = e^\chi e^{-(z-a_\chi)^2/4}$; take $c = \sqrt{\chi}$. Gaussian integration gives $N_0 \sim 2\sqrt{\pi}e^\chi$. Substituting $u = \sqrt{\chi}z$ yields

$$\sqrt{\chi}S(c) = \int_0^\chi e^{-u+u^2/(4\chi)} \, \mathrm{d}u \longrightarrow 1.$$

The integrand, extended by zero outside $[0, \chi]$, is bounded by $e^{-3u/4}$, which justifies dominated convergence. For each fixed $K > 0$,

$$\pi([0, c]) = O(\sqrt{\chi}e^{-\chi/4}), \qquad \int_c^{a_\chi+K} \rho^{-1} = O_K(\sqrt{\chi}e^{-3\chi/4}).$$

For the second bound, the exponent $\chi - (z - a_\chi)^2/4$ is at least $3\chi/4$ on this interval for all sufficiently large χ. Under π, $z - a_\chi$ has a centered Gaussian law of variance 2 truncated below $-a_\chi$; its first two moments therefore converge to $0, 2$. This proves both the upper-tail condition and the L^2 condition in (118), the latter with squared norm $O(\chi^{-1})$. Hence all its conditions hold, and $\widehat{T} \sim T_\chi$.

Fixed $\lambda > 0$. Fix $\lambda > 0$, put $z_\lambda = a_\chi/(1+\lambda)$, and take $c = (z_\lambda + a_\chi)/2$. Direct integration of $\gamma_{\chi,\lambda}$ gives

$$\rho(z) = \begin{cases} e^{\lambda z^2/4}, & 0 \le z \le z_\lambda, \\ e^{\chi\lambda/(1+\lambda)} e^{-(z-a_\chi)^2/4}, & z \ge z_\lambda. \end{cases} \tag{122}$$

The mass below z_λ is at most $z_\lambda e^{\lambda\chi/(1+\lambda)^2}$, exponentially smaller than $e^{\chi\lambda/(1+\lambda)}$. The Gaussian part starts a distance $a_\chi\lambda/(1+\lambda) \to \infty$ to the left of its center. Consequently

$$N_0 \sim 2\sqrt{\pi}e^{\chi\lambda/(1+\lambda)}, \qquad S(c) = \int_0^{z_\lambda} e^{-\lambda z^2/4} \, \mathrm{d}z + O_\lambda(\sqrt{\chi}e^{-\lambda\chi/(1+\lambda)^2}) \longrightarrow \frac{\sqrt{\pi}}{\sqrt{\lambda}}.$$

The error bound uses monotonicity of ρ on $[z_\lambda, c]$. Monotonicity also gives, for each fixed $K > 0$,

$$\pi([0, c]) = O_\lambda\left(\sqrt{\chi}e^{-\chi\lambda^2/[4(1+\lambda)^2]}\right),$$
$$\int_c^{a_\chi+K} \rho^{-1} = O_{\lambda,K}\left(\sqrt{\chi}e^{-\chi\lambda(3\lambda+4)/[4(1+\lambda)^2]}\right).$$

For the second estimate, the exponent of (122) is minimized at c for large χ. The normalized contribution below z_λ to $\int (z - a_\chi)^2 \, \mathrm{d}\pi$ is $O_\lambda(\chi^{3/2}e^{-\chi\lambda^2/(1+\lambda)^2})$ and tends to zero. Thus $z - a_\chi$ again converges to the centered Gaussian law of variance 2, together with its first two moments. Its Gaussian upper tail and the moment estimate verify the remaining conditions in (118). We obtain $\widehat{T} \sim T_{\chi,\lambda}$.

Cumulative limits and quantiles. The continuity theorem for Laplace transforms gives $\tau_0/\widehat{T} \Rightarrow \mathrm{Exp}(1)$ under η. Since the limit distribution function is continuous, the distribution functions converge uniformly: use a finite grid with small limiting increments on a compact time interval, monotonicity between grid points, and then a small upper tail. Replacing $\widehat{T}$ by the asymptotically equivalent T_χ or $T_{\chi,\lambda}$ preserves uniformity, since $\sup_{s\ge0} |e^{-cs} - e^{-s}| \to 0$ as $c \to 1$. Lemma F.1 proves the asserted cumulative limits.

For either explicit scale T, let $\delta_\chi = \sup_{s\ge0} |I(Ts)/I_\infty - (1 - e^{-s})|$. Whenever $\delta_\chi < \min\{\alpha, 1-\alpha\}$, the defining equality for Q_α gives

$$-\log(1 - \alpha + \delta_\chi) \le \frac{Q_\alpha}{T} \le -\log(1 - \alpha - \delta_\chi).$$

As $\delta_\chi \to 0$, these bounds prove uniform convergence of Q_α/T to $-\log(1-\alpha)$ on compact subsets of $(0,1)$. The limit is bounded away from zero there, so the ratio equivalents are uniform as well. Together with (98), this also proves (26). □

G Reverse Diffusion Examples

Gaussian smoothing and the radial bound. Let $\|X\| \le a$ almost surely, with $a \ge 0$, and let $Z \sim \mathcal{N}(0, I_d)$ be independent. For $v > 0$, write φ_v for the density of $\mathcal{N}(0, vI_d)$. Then $q_v(x) = \mathbb{E}\varphi_v(x - X) > 0$, and differentiation under the expectation gives

$$\nabla q_v(x) = \frac{1}{v}\mathbb{E}[(X-x)\varphi_v(x-X)], \qquad \nabla \log q_v(x) = -\frac{x}{v} + \frac{1}{v}\mathbb{E}[X \mid X + \sqrt{v}Z = x].$$

The posterior mean has norm at most a. Further differentiation yields

$$\nabla^2 \log q_v(x) = -\frac{I_d}{v} + \frac{\mathrm{Cov}(X \mid X + \sqrt{v}Z = x)}{v^2}.$$

For any unit vector e, the conditional variance of $e \cdot X$ is at most its conditional second moment, hence at most a^2. Thus $0 \preccurlyeq \mathrm{Cov}(X \mid X + \sqrt{v}Z = x) \preccurlyeq a^2 I_d$, giving the upper bound in Section 2. On $0 < v_{\min} \le v \le v_{\max} < \infty$, for example,

$$L = \max\left\{0, \frac{a^2}{v_{\min}^2} - \frac{1}{v_{\max}}\right\}$$

is a valid common one-sided Lipschitz bound on the interval. See also Beyler & Bach (2025).

The third derivatives of $\log q_v$ are the conditional third centered moments divided by v^3. Boundedness of X therefore gives uniformly bounded first two spatial derivatives of the score on this interval. The score and these derivatives are jointly continuous, and the score has linear growth; see also Beyler & Bach (2025) for the conditional-moment formulas.

For $T = v_{\max} - v_{\min} > 0$, the Brownian heat flow reversed from $q_{v_{\max}}$ has unit noise and drift $b_t = \nabla \log q_{v_{\max}-t}$ on $[0, T]$ (Beyler & Bach, 2025). In reverse time t, its generator is $b_t \cdot \nabla + \frac{1}{2}\Delta$, and its terminal law is $q_{v_{\min}}$. The decomposition above gives (2) with $m = 1/v_{\max}$ and $A = a/v_{\min}$, and the derivative bounds give Assumption 2.1. Choosing $X = \pm a$ equiprobably in dimension one and $v = 1 + s$ gives the mixture below. Its posterior mean is $a\tanh(ax/v)$.

Instantaneous noise and forward smoothing. The lower bound on σ_t concerns the SDE noise coefficient. For example, in forward time s, the variance-preserving (VP) SDE has drift $-\beta(s)x/2$ and noise coefficient $\sqrt{\beta(s)}$. For continuous $\beta \ge \beta_{\min} > 0$, this coefficient is bounded below, whereas the conditional Gaussian variance is $1 - \exp\{-\int_0^s \beta(u)\,\mathrm{d}u\}$ and vanishes at the data endpoint $s = 0$ (Song et al., 2021). The exact reverse SDE has the same instantaneous noise coefficient under time reversal. Thus a positive lower bound on the noise coefficient is distinct from imposing a nonzero forward smoothing variance at the terminal target.

For the sub-VP SDE, the noise coefficient is $\sqrt{\beta(s)(1 - \exp\{-2\int_0^s \beta(u)\,\mathrm{d}u\})}$ and vanishes at $s = 0$. The probability-flow ODE has zero noise coefficient (Song et al., 2021). These endpoint coefficients do not satisfy the positive noise floor on the full closed interval. For a VP reverse diffusion, the radial and regularity assumptions must still be checked separately.

G.1 Transition coefficients and terminal error

Let $a > \sqrt{2}$ and

$$p_s = \tfrac{1}{2}\mathcal{N}(-a, 1+s) + \tfrac{1}{2}\mathcal{N}(a, 1+s), \qquad 0 \le s \le 1. \tag{123}$$

The unit-noise reverse drift and its constant perturbation, both started at p_1, are

$$b_t(x) = -\frac{x}{2-t} + \frac{a}{2-t}\tanh\left(\frac{ax}{2-t}\right), \qquad \widehat{b}_t(x) = b_t(x) + \varepsilon, \qquad \varepsilon > 0. \tag{124}$$

Write $P_{r,t}$ for the reference transitions, $\widehat{\mu}_t^{\varepsilon}$ for the perturbed law, and $W_a(\varepsilon) = W_1(\widehat{\mu}_1^{\varepsilon}, p_0)$. The derivative is

$$b_t'(x) = -\frac{1}{2-t} + \frac{a^2}{(2-t)^2}\operatorname{sech}^2\left(\frac{ax}{2-t}\right).$$

Its spatial maximum is at zero. Since $v \mapsto a^2/v^2 - 1/v$ decreases on $[1,2]$ for $a > \sqrt{2}$, $\operatorname{osLip}(b) = a^2 - 1$, unchanged by the constant perturbation. Here L varies with separation; (107) keeps L fixed.

Proposition G.1 (Stability and terminal error for the mixture)**.** *For $0 \le r \le t \le 1$, set $v = 2 - r$ and $u = 2 - t$. Then*

$$\operatorname{Lip}_{W_1}(P_{r,t}) = \frac{u}{v} + \frac{a^2(v-u)}{v^2}. \tag{125}$$

For every $\varepsilon > 0$, the terminal error satisfies

$$\log 2 \le \frac{W_a(\varepsilon)}{\varepsilon} \le \min\left\{1 + \frac{8a}{3\sqrt{2\pi}} + a^2,\ \log 2 + a^2(\log 2 - \tfrac{1}{2}),\ \sqrt{1+a^2}\right\}. \tag{126}$$

For each fixed $\varepsilon > 0$,

$$\frac{W_a(\varepsilon)}{\varepsilon} \longrightarrow \log 2 \qquad (a \to \infty). \tag{127}$$

Proof of the transition coefficient and upper bounds. Both drifts in (124) are regular for each fixed a, ε. Their pairwise radial drift is at most $-r/2 + 2a$: the linear part contracts at rate at least $1/2$, the nonlinear residual is bounded by a, and a constant translation cancels from pair differences. Thus $m = 1/2, A = a, \sigma = 1, \chi = 2a^2$ suffice, and (101) gives the first upper bound in (126). The discrepancy is exactly ε under every probability law.

Represent the forward heat flow as $Y_s = a\xi + G_0 + B_s$, where ξ is a symmetric sign and G_0, B are independent standard Gaussian and Brownian variables. With φ_s the centered Gaussian density of variance s, its reverse transition is

$$P_{r,t}(x, \mathrm{d}y) = \frac{p_{1-t}(y)\varphi_{t-r}(x-y)}{p_{1-r}(x)}\,\mathrm{d}y, \qquad r < t. \tag{128}$$

Bayes' rule verifies normalization and composition. Differentiating this kernel at $t = r$ gives generator $b_r\partial_x + \frac{1}{2}\partial_{xx}$, and regularity gives uniqueness; in particular the reference marginals are p_{1-t}. This is the reverse-diffusion formula of Anderson (1982); Haussmann & Pardoux (1986) in the present heat-flow setting.

For $v = 2 - r$ and $u = 2 - t$, Gaussian conditioning within each component and Bayes' rule give

$$\mathbb{E}[Y_{u-1} \mid Y_{v-1} = x, \xi] = \frac{u}{v}x + \left(1 - \frac{u}{v}\right)a\xi, \qquad \mathbb{E}[a\xi \mid Y_{v-1} = x] = a\tanh(ax/v).$$

Consequently the transition mean and derivative are

$$\begin{aligned} M_{u,v}(x) &:= P_{r,t}\,\mathrm{id}(x) = \frac{u}{v}x + \left(1 - \frac{u}{v}\right)a\tanh(ax/v), \\ M_{u,v}'(x) &= \frac{u}{v} + \frac{a^2(v-u)}{v^2}\operatorname{sech}^2(ax/v). \end{aligned} \tag{129}$$

The log-density Hessian $\partial_x b_t$ and its midpoint maximum are explicit in Saremi et al. (2024); see also Beyler & Bach (2025). Synchronous scalar solutions preserve initial order, so (46) identifies the Wasserstein distance with the difference of transition means. Their secants are bounded by $M_{u,v}'(0)$ and approach it as both initial points approach zero, proving (125); at $r = t$ both sides equal one. Duhamel with reference transitions and drift error averaged under $\widehat{\mu}_r^{\varepsilon}$ therefore yields the second upper bound, since

$$\int_0^1 \operatorname{Lip}_{W_1}(P_{r,1})\,\mathrm{d}r = \int_1^2 \left(\frac{1}{v} + a^2\frac{v-1}{v^2}\right)\mathrm{d}v = \log 2 + a^2(\log 2 - \tfrac{1}{2}).$$

Let $\mathbb{P}$ and $\mathbb{Q}^{\varepsilon}$ be the reference and perturbed path laws, with coordinate process X and filtration $\mathcal{F}_t = \sigma(X_s : 0 \le s \le t)$. Under $\mathbb{P}$, $B_t = X_t - X_0 - \int_0^t b_s(X_s)\,\mathrm{d}s$ is Brownian motion independent of X_0. The exponential martingale for the constant drift change gives

$$\left.\frac{\mathrm{d}\mathbb{Q}^{\varepsilon}}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{F}_t} = L_t := \exp\left(\varepsilon B_t - \frac{\varepsilon^2 t}{2}\right), \qquad \mathbb{E}_{\mathbb{P}}L_t^2 = e^{\varepsilon^2 t}. \tag{130}$$

Indeed, $\mathbb{E}_{\mathbb{P}}[L_1 \mid \mathcal{F}_0] = 1$, so weighting by L_1 preserves the initial law p_1. Girsanov's theorem (Girsanov, 1960) makes $B_t - \varepsilon t$ Brownian under the new law, and X then solves the equation with drift $b_t + \varepsilon$. Uniqueness identifies this law with $\mathbb{Q}^\varepsilon$. The Gaussian moment formula gives the second identity in (130) and $\mathrm{KL}(\mathbb{Q}^\varepsilon \| \mathbb{P}) = \mathbb{E}_{\mathbb{Q}^\varepsilon} \log L_1 = \varepsilon^2/2$.

For the entropy bound, write a sample from p_0 as $a\xi + G$. Conditional Gaussian concentration for a 1-Lipschitz f has variance proxy one, and its two conditional means differ by at most $2a$. The sign average uses $\cosh x \le e^{x^2/2}$, giving

$$\log p_0 e^{\theta(f - p_0 f)} \le \tfrac{1}{2}(1 + a^2)\theta^2.$$

The conditional Gaussian estimate follows from the heat-martingale representation with integrand bounded by one; truncation handles unbounded f. Entropy duality and optimization in θ give the transport–entropy estimate (Bobkov & Götze, 1999)

$$W_1(\rho, p_0) \le \sqrt{2(1 + a^2)\,\mathrm{KL}(\rho \| p_0)}.$$

The path entropy above and terminal data processing prove $W_a(\varepsilon) \le \varepsilon\sqrt{1 + a^2}$. □

G.2 Constant forcing at large separation

Proof of the lower bound and limit in Proposition G.1. For fixed a, ε, apply (43) with $h_t = w(t) = 1$ and amplitude ε. The drifts are regular and $p_1 \in \mathcal{P}_2$. With X^ε denoting the perturbed process and $M_t = P_{t,1}\,\mathrm{id}$, (129) at $u = 1$ gives

$$\frac{W_a(\varepsilon)}{\varepsilon} = \int_0^1 \mathbb{E} M_t'(X_t^\varepsilon)\,\mathrm{d}t = \log 2 + a^2 \int_0^1 \frac{1 - t}{(2 - t)^2} \mathbb{E}\,\mathrm{sech}^2\left(\frac{a X_t^\varepsilon}{2 - t}\right)\,\mathrm{d}t. \tag{131}$$

The nonnegative integrand proves the lower bound $\log 2$.

Put $v_t = 2 - t$ and $\phi_t(x) = \mathrm{sech}^2(ax/v_t)$. Under $\mathbb{P}$, X_t has law $\frac{1}{2}\mathcal{N}(-a, v_t) + \frac{1}{2}\mathcal{N}(a, v_t)$, with $1 \le v_t \le 2$. A one-sided Gaussian tail bound in each component gives

$$\mathbb{P}(|X_t| < a/2) \le e^{-a^2/(8v_t)} \le e^{-a^2/16}.$$

On $\{|X_t| \ge a/2\}$, $\phi_t(X_t) \le 4e^{-a^2/v_t}$. Since $0 \le \phi_t \le 1$, splitting over these two events yields

$$\mathbb{E}_{\mathbb{P}}\phi_t(X_t)^2 \le e^{-a^2/16} + 4e^{-a^2/2} \le 5e^{-a^2/16}.$$

By (130) and Cauchy–Schwarz,

$$\widehat{\mu}_t^\varepsilon(\phi_t) = \mathbb{E}_{\mathbb{P}}[L_t \phi_t(X_t)] \le \sqrt{5}\exp\left(\frac{\varepsilon^2 t}{2} - \frac{a^2}{32}\right) \le \sqrt{5}\exp\left(\frac{\varepsilon^2}{2} - \frac{a^2}{32}\right) \quad (0 \le t \le 1). \tag{132}$$

Inserting this uniform bound in (131) and using $\int_0^1 (1 - t)/(2 - t)^2\,\mathrm{d}t = \log 2 - \frac{1}{2}$ gives, for every $a > \sqrt{2}$ and $\varepsilon > 0$,

$$0 \le \frac{W_a(\varepsilon)}{\varepsilon} - \log 2 \le \sqrt{5}(\log 2 - \tfrac{1}{2})a^2 \exp\left(\frac{\varepsilon^2}{2} - \frac{a^2}{32}\right). \tag{133}$$

For each fixed $\varepsilon > 0$ the right side tends to zero, proving (127). □

The likelihood estimate uses the pointwise constant drift discrepancy. The expected L^1 condition in Theorem 2.3 also permits errors concentrated on sets of small probability, with large second moments.