Theory of Stochastic Processes
Vol. 29 (45), no. 2, 2025, pp. 79–87

KYRYLO KUCHYNSKYI

STOCHASTIC DIFFERENTIAL EQUATIONS WITH INTERACTION AND CONSTANT DIFFUSION

We analyze the evolution of probability measures associated with stochastic differential equations with interaction (SDEWI) featuring constant diffusion coefficients. Our main results concern the propagation of the logarithmic Sobolev inequality (LSI); we prove that if the initial measure satisfies an LSI, the solution preserves this property almost surely. Additionally, we provide exact growth estimates for the Wasserstein-2 distance between two solutions driven by the same Brownian motion, improving upon general stability bounds by utilizing the constant diffusion structure. Finally, we derive bounds for the expected displacement of the measure from its initial state, characterizing the short-time behavior of the system.

1. Introduction

We are interested in families $(\mu_t)_{t\geq 0}$ of measure-valued processes which are the solutions to the following equation on $\mathbb{R}^d$

$$\begin{cases} dX_t(u) &= \int_{\mathbb{R}^d} \varphi(X_t(u) - X_t(v))\mu_0(\mathrm{d}v)\mathrm{d}t + \sigma \mathrm{d}W_t, \\ X_0(u) &= u \in \mathbb{R}^d, \\ \mu_t &= \mu_0 \circ X_t^{-1}, \quad t \geq 0, \end{cases} \tag{1}$$

where $\sigma > 0$, $\varphi : \mathbb{R}^d \longrightarrow \mathbb{R}^d$, μ_0 is an initial probability distribution, and W_t is a standard d-dimensional Brownian motion.

This system describes the evolution of the law of a system of particles which "interact" with each other. Such equations are called stochastic differential equations with interactions and were proposed by Andrey A. Dorogovtsev in [1]. We consider a model with a constant diffusion term.

A key feature of our model is the constant diffusion term σ. Since the noise is additive and independent of the state, it acts uniformly on the system.

Throughout this paper, we rely on the following assumption regarding the interaction kernel φ:

Assumption 1 (A1)**.** *There exists a constant $L > 0$ such that for all $x, y \in \mathbb{R}^d$:*

$$|\varphi(x) - \varphi(y)| \leq L|x - y|.$$

We begin by proving that if μ_0 satisfies a logarithmic Sobolev inequality, then so does μ_t.

Definition 1.1 (Logarithmic Sobolev inequality)**.** Denoting by $\mathcal{C}_c^1(\mathbb{R}^d)$ the set of compactly supported smooth functions from $\mathbb{R}^d$ to $\mathbb{R}$, a probability measure μ on $\mathbb{R}^d$ is said

2020 *Mathematics Subject Classification.* 60B05 .
Key words and phrases. stochastic differential equations with interactions.
This work was supported by a grant from the Simons Foundation (SFI-PD-Ukraine-00014586, K.K.).

to satisfy a logarithmic Sobolev inequality with constant $C > 0$ if for every function $h \in \mathcal{C}^1_c(\mathbb{R}^d)$ satisfying the normalization condition

$$\int_{\mathbb{R}^d} h^2 \, d\mu = 1,$$

the following inequality holds:

$$\int_{\mathbb{R}^d} h^2 \ln(h^2) \, d\mu \le C \int_{\mathbb{R}^d} |\nabla h|^2 \, d\mu. \tag{2}$$

Equivalently, for any probability measure $\nu \in \mathcal{P}(\mathbb{R}^d)$ such that ν is absolutely continuous with respect to μ and $\sqrt{d\nu/d\mu} \in \mathcal{C}^1_c$, we have

$$\mathcal{H}(\nu \mid \mu) \le \frac{C}{4} \mathcal{I}(\nu \mid \mu),$$

where $\mathcal{H}, \mathcal{I}$ are the relative entropy and Fisher information defined respectively as follows:

$$\mathcal{H}(\nu \mid \mu) := \int_{\mathbb{R}^d} \ln \frac{d\nu}{d\mu} \, d\nu,$$

$$\mathcal{I}(\nu \mid \mu) := \int_{\mathbb{R}^d} \left| \nabla \ln \frac{d\nu}{d\mu} \right|^2 \, d\nu.$$

We also use the entropy functional: for a probability measure μ and a measurable function $f \ge 0$ such that $\int f \, d\mu < \infty$, define

$$\mathrm{Ent}_\mu(f) := \int_{\mathbb{R}^d} f \ln f \, d\mu - \Big(\int_{\mathbb{R}^d} f \, d\mu \Big) \ln \Big(\int_{\mathbb{R}^d} f \, d\mu \Big), \tag{3}$$

with the convention $0 \ln 0 = 0$. In particular, if $\int g^2 \, d\mu = 1$, then

$$\mathrm{Ent}_\mu(g^2) = \int_{\mathbb{R}^d} g^2 \ln(g^2) \, d\mu.$$

The next result presents conditions under which the family $(\mu_t)_{t \ge 0}$ solving (1) satisfies an LSI.

Proposition 1.1 (Propagation of LSI)**.** *Assume that Assumption (A1) holds and μ_0 satisfies an LSI with constant $C_0 > 0$. Then, almost surely, the measure μ_t satisfies an LSI with constant $C_0 e^{2Lt}$.*

Remark 1.1. We note that Proposition 1.1 can alternatively be viewed as a consequence of the general stability of the LSI under Lipschitz pushforwards (see, e.g., [3, Proposition 2.1]). Specifically, if a map $T : \mathbb{R}^d \to \mathbb{R}^d$ is K-Lipschitz and μ_0 satisfies an LSI with constant C, then the pushforward measure $\mu_0 \circ T^{-1}$ satisfies an LSI with constant CK^2. In our context, the flow map $u \mapsto X_t(u)$ is almost surely e^{Lt}-Lipschitz. However, we provide the direct proof below to keep the exposition self-contained and to explicitly demonstrate how the interaction kernel's Lipschitz property controls the operator norm of the Jacobian ∇T_t.

For the next results, we consider the more generalized model with constant diffusion and drift $a(X_t(u), \mu_t)$

$$\begin{cases} dX_t(u) & = a(X_t(u), \mu_t) \, \mathrm{d}t + \sigma \, \mathrm{d}W_t, \\ X_0(u) & = u \in \mathbb{R}^d, \\ \mu_t & = \mu_0 \circ X_t^{-1}, \quad t \ge 0, \end{cases} \tag{4}$$

where $\sigma > 0$ is a constant diffusion coefficient, W_t is a standard d-dimensional Brownian motion, and μ_0 is the initial probability distribution. The drift coefficient is given by a function $a : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d$, where $\mathcal{P}_2(\mathbb{R}^d)$ denotes the space of probability measures

with finite second moments equipped with the Wasserstein-2 distance. For any $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$, this distance is defined as

$$W_2(\mu,\nu) := \left(\inf_{\pi \in \Pi(\mu,\nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x-y|^2 \, \pi(dx,dy) \right)^{1/2}, \tag{5}$$

where $\Pi(\mu,\nu)$ denotes the set of all joint probability measures on $\mathbb{R}^d \times \mathbb{R}^d$ with marginals μ and ν.

Assumption 2 (A2). *The drift coefficient $a(x,\mu)$ satisfies the following Lipschitz conditions:*

(i) *There exists a constant $L \in \mathbb{R}$ such that for all $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and $x, y \in \mathbb{R}^d$:*

$$|a(x,\mu) - a(y,\mu)| \leq L|x-y|.$$

(ii) *There exists a constant $K > 0$ such that for all $x \in \mathbb{R}^d$ and $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$:*

$$|a(x,\mu) - a(x,\nu)| \leq K W_2(\mu,\nu).$$

This system generalizes models of interacting particles where the drift depends on the distribution of the population. A straightforward example satisfying Assumption (A2) is the pairwise interaction model (1). It can be shown that for φ satisfying Assumption (A1), both constants in Assumption (A2) are exactly L.

We address how the distance between two solutions μ_t and ν_t evolves relative to their initial distance $W_2(\mu_0,\nu_0)$. Lemma 2.4.1 in [1] establishes that the Wasserstein distance between solutions is bounded by the initial distance multiplied by some constant C,

$$\mathbb{E}\left[\sup_{t\in[0,T]} W_2^2(\mu_t,\nu_t) \right] \leq C W_2^2(\mu_0,\nu_0).$$

In contrast, our Proposition 1.2 provides a pathwise estimate that holds almost surely, rather than in expectation and we obtain the explicit growth rate $e^{(2L+2K)t}$ instead of a generic constant C

Proposition 1.2. *Assume that Assumption 2 holds. Let μ_t and ν_t be two solutions to (4) with initial data μ_0 and ν_0. Then, for all $t \geq 0$:*

$$W_2^2(\mu_t,\nu_t) \leq e^{(2L+2K)t} W_2^2(\mu_0,\nu_0) \quad \textit{almost surely.}$$

Remark 1.2. The inequality in Proposition 1.2 holds almost surely rather than in expectation because both particle systems are driven by the same realization of the Brownian motion W_t. As a result, the diffusion terms cancel out exactly in the difference equation, reducing the problem to a pathwise analysis of the drift.

Finally, we provide estimates on the growth of the Wasserstein distance from the initial measure, $\mathbb{E}[W_2^2(\mu_t,\mu_0)]$.

Proposition 1.3. *Assume that Assumption 2 holds and $|a(x,\mu)|$ grows at most linearly. Then, for any $T > 0$, there exists a constant C depending on L, K, T, and moments of μ_0 such that for all $t \in [0,T]$:*

$$\mathbb{E}\left[W_2^2(\mu_t,\mu_0)\right] \leq 2d\sigma^2 t + Ct^2.$$

2. Logarithmic Sobolev inequality

In this section, we prove Proposition 1.1, showing that the LSI is preserved by the flow with a constant that grows at most exponentially in time.

Proof of Proposition 1.1. The proof proceeds in two steps. First, we establish the estimate under the assumption that the interaction kernel φ is smooth $\mathcal{C}_c^\infty(\mathbb{R}^d)$. Second, we extend this result to general Lipschitz kernels via a regularization argument.
Step 1: The smooth case. Assume initially that $\varphi \in C_c^\infty(\mathbb{R}^d)$. Then, for each fixed (t,ω), the drift coefficient

$$a(x,\mu_t(\omega)) = \int_{\mathbb{R}^d} \varphi(x - X_t(v,\omega))\,\mu_0(dv)$$

is globally Lipschitz with constant $\leq L$ and belongs to $C_b^k(\mathbb{R}^d)$ for all $k \geq 1$, with all spatial derivatives bounded.

By Kunita's theorem on stochastic flows [2, Theorem 1.4.1], there exists a modification of the solution such that, for almost every ω and all $t > 0$, the map $T_t(\omega,\cdot) : u \mapsto X_t(u,\omega)$ is a global $\mathcal{C}^1$-diffeomorphism of $\mathbb{R}^d$. Fix such an ω.

Let $g \in \mathcal{C}_c^1(\mathbb{R}^d)$ satisfy the normalization

$$\int_{\mathbb{R}^d} g^2\,d\mu_t = 1.$$

Define $h := g \circ T_t$. Since $\mu_t = \mu_0 \circ T_t^{-1}$, the change-of-variables formula for pushforward measures gives

$$\int_{\mathbb{R}^d} h^2\,d\mu_0 = \int_{\mathbb{R}^d} (g\circ T_t)^2\,d\mu_0 = \int_{\mathbb{R}^d} g^2\,d\mu_t = 1.$$

Therefore, by inequality (2) applied to μ_0,

$$\int_{\mathbb{R}^d} h^2 \ln(h^2)\,d\mu_0 \leq C_0 \int_{\mathbb{R}^d} |\nabla h|^2\,d\mu_0.$$

Using again the pushforward relation $\mu_t = \mu_0 \circ T_t^{-1}$, the left-hand side equals

$$\int_{\mathbb{R}^d} (g\circ T_t)^2 \ln\big((g\circ T_t)^2\big)\,d\mu_0 = \int_{\mathbb{R}^d} g^2 \ln(g^2)\,d\mu_t.$$

For the right-hand side, we apply the chain rule: for any $u \in \mathbb{R}^d$,

$$\nabla h(u) = \nabla(g\circ T_t)(u) = (\nabla T_t(u))^\top\,\nabla g(T_t(u)).$$

Taking norms and using the definition of operator norm yields

$$|\nabla h(u)|^2 \leq \|\nabla T_t(u)\|_{\mathrm{op}}^2\,|\nabla g(T_t(u))|^2.$$

Integrating over μ_0 and applying the pushforward relation once more gives

$$\int_{\mathbb{R}^d} |\nabla h|^2\,d\mu_0 \leq \Big(\sup_{u\in\mathbb{R}^d} \|\nabla T_t(u)\|_{\mathrm{op}}^2\Big) \int_{\mathbb{R}^d} |\nabla g|^2\,d\mu_t.$$

Combining these estimates, we obtain

$$\int_{\mathbb{R}^d} g^2 \ln(g^2)\,d\mu_t \leq C_0\Big(\sup_{u} \|\nabla T_t(u)\|_{\mathrm{op}}^2\Big) \int_{\mathbb{R}^d} |\nabla g|^2\,d\mu_t.$$

In our context, $\nabla T_t(u)$ is the Jacobian matrix of the flow map T_t at point u, and the operator norm quantifies the maximum stretching factor of the flow, so we need to evaluate it to get the constant for the logarithmic Sobolev inequality.

For any $u, v \in \mathbb{R}^d$, let $\Delta_t = X_t(u) - X_t(v)$. We have:

$$\frac{d}{dt}|\Delta_t|^2 \leq 2|\Delta_t| \int |\varphi(X_t(u) - X_t(w)) - \varphi(X_t(v) - X_t(w))|\,\mu_0(dw).$$

Using the Lipschitz property of φ:

$$\frac{d}{dt}|\Delta_t|^2 \le 2L|\Delta_t|^2.$$

By Grönwall's inequality, $|\Delta_t| \le e^{Lt}|\Delta_0|$, which implies:

$$\sup_{u\in\mathbb{R}^d} \|\nabla T_t(u)\|_{\mathrm{op}}^2 \le e^{2Lt}.$$

Substituting this back into the entropy estimate:

$$\mathrm{Ent}_{\mu_t}(g^2) \le C_0 e^{2Lt} \int_{\mathbb{R}^d} |\nabla g(T_t(u))|^2 \,\mu_0(du) = C_0 e^{2Lt} \int_{\mathbb{R}^d} |\nabla g(y)|^2 \,\mu_t(dy). \tag{6}$$

Thus, the result holds when the kernel φ is in $\mathcal{C}_b^\infty(\mathbb{R}^d)$.
Step 2: Regularization. Now consider a general interaction kernel φ satisfying Assumption 1. Let ρ_ε be a standard mollifier in $\mathbb{R}^d$ and define $\varphi_\varepsilon = \varphi * \rho_\varepsilon$, such that $\|\varphi - \varphi_\varepsilon\|_\infty \le L\varepsilon$ and that φ_ε preserves the Lipschitz constant L.

Let $X_t^\varepsilon(u)$ denote the solution associated with φ_ε, driven by the same Brownian path $W_t(\omega)$ as the original flow $X_t(u)$. Since φ_ε is smooth, the estimate (6) holds:

$$\mathrm{Ent}_{\mu_t^\varepsilon}(g^2) \le C_0 e^{2Lt} \int_{\mathbb{R}^d} |\nabla g|^2 \, d\mu_t^\varepsilon, \quad \forall g \in \mathcal{C}_c^1(\mathbb{R}^d). \tag{7}$$

Consider the quantity

$$D_t := \sup_{u\in\mathbb{R}^d} |X_t(u) - X_t^\varepsilon(u)|.$$

Now

$$X_t(u) - X_t^\varepsilon(u) = \int_0^t \int_{\mathbb{R}^d} [\varphi(X_s(u) - X_s(v)) - \varphi_\varepsilon(X_s^\varepsilon(u) - X_s^\varepsilon(v))] \,\mu_0(dv)\, ds.$$

Using the triangle inequality:

$$\begin{aligned}
&|\varphi(X_s(u) - X_s(v)) - \varphi_\varepsilon(X_s^\varepsilon(u) - X_s^\varepsilon(v))| \le \\
&\le \|\varphi - \varphi_\varepsilon\|_\infty + L|(X_s(u) - X_s(v)) - (X_s^\varepsilon(u) - X_s^\varepsilon(v))| \\
&\le L\varepsilon + L|X_s(u) - X_s^\varepsilon(u)| + L|X_s(v) - X_s^\varepsilon(v)|.
\end{aligned}$$

Thus

$$D_t \le \int_0^t (L\varepsilon + 2LD_s)\, ds.$$

By Grönwall's lemma:

$$D_t \le L\varepsilon \int_0^t e^{2L(t-s)}\, ds = \frac{\varepsilon}{2}\big(e^{2Lt} - 1\big),$$

which implies pointwise convergence for all u.

Using the representation $\mu_t^\varepsilon = \mu_0 \circ (X_t^\varepsilon)^{-1}$ and the entropy definition (3), the left-hand side in (7) can be rewritten in terms of the initial measure. Define

$$\begin{aligned}
h_0(x) &:= g(x)^2, \\
h_1(x) &:= g(x)^2 \ln(g(x)^2) \quad (0 \ln 0 := 0), \\
h_2(x) &:= |\nabla g(x)|^2,
\end{aligned}$$

and set $m_\varepsilon := \int_{\mathbb{R}^d} h_0(X_t^\varepsilon(u))\,\mu_0(du) = \int h_0 \, d\mu_t^\varepsilon$. Then

$$\mathrm{Ent}_{\mu_t^\varepsilon}(g^2) = \int_{\mathbb{R}^d} h_1(X_t^\varepsilon(u))\,\mu_0(du) - m_\varepsilon \ln(m_\varepsilon),$$

and similarly

$$\int_{\mathbb{R}^d} |\nabla g|^2 \, d\mu_t^\varepsilon = \int_{\mathbb{R}^d} h_2(X_t^\varepsilon(u)) \, \mu_0(du).$$

Thus (7) becomes

$$\int_{\mathbb{R}^d} h_1(X_t^\varepsilon(u)) \, \mu_0(du) - m_\varepsilon \ln(m_\varepsilon) \le C_0 e^{2Lt} \int_{\mathbb{R}^d} h_2(X_t^\varepsilon(u)) \, \mu_0(du).$$

Since $g \in \mathcal{C}_c^1(\mathbb{R}^d)$, the functions h_0, h_1, h_2 are continuous and compactly supported, hence bounded. Let $M = \max(\|h_0\|_\infty, \|h_1\|_\infty, \|h_2\|_\infty)$.

For every u, the pointwise convergence $X_t^\varepsilon(u) \to X_t(u)$ implies:

$$h_i(X_t^\varepsilon(u)) \to h_i(X_t(u)), \quad i = 0, 1, 2.$$

Furthermore, the integrands are dominated by the constant M, which is integrable with respect to the probability measure μ_0:

$$|h_i(X_t^\varepsilon(u))| \le M, \quad i = 0, 1, 2.$$

Therefore, by the dominated convergence theorem,

$$\int h_i(X_t^\varepsilon(u)) \, \mu_0(du) \to \int h_i(X_t(u)) \, \mu_0(du), \quad i = 0, 1, 2.$$

In particular $m_\varepsilon \to m := \int h_0(X_t(u)) \, \mu_0(du) = \int g^2 \, d\mu_t$, and since $x \mapsto x \ln x$ is continuous on $[0, \infty)$ we also have $m_\varepsilon \ln(m_\varepsilon) \to m \ln(m)$. Letting $\varepsilon \to 0$ in the inequality above yields

$$\mathrm{Ent}_{\mu_t}(g^2) \le C_0 e^{2Lt} \int_{\mathbb{R}^d} |\nabla g|^2 \, d\mu_t,$$

which completes the proof. □

3. Wasserstein Stability

3.1. **Stability with respect to Initial Data.** We prove Proposition 1.2. Consider two initial measures $\mu_0, \nu_0 \in \mathcal{P}_2(\mathbb{R}^d)$. Let $X_t(u)$ and $Y_t(v)$ be the flows starting from u and v respectively, driven by the *same* Brownian motion W_t. The associated measures are $\mu_t = \mu_0 \circ X_t^{-1}$ and $\nu_t = \nu_0 \circ Y_t^{-1}$.

In this section, we aim to bound the Wasserstein distance between two evolving measures, μ_t and ν_t, in terms of their initial measures μ_0 and ν_0. By employing a synchronous coupling—using the same Brownian drivers for both processes—we can isolate the effects of the drift components, allowing us to derive a pathwise stability estimate that depends on the Lipschitz constants of the interaction kernel almost surely. We work with the generalized model described in (4):

$$dX_t(u) = a(X_t(u), \mu_t) \, \mathrm{d}t + \sigma \, \mathrm{d}W_t.$$

The dynamics of the difference $Z_t = X_t(u) - Y_t(v)$ are given by:

$$dZ_t = (a(X_t(u), \mu_t) - a(Y_t(v), \nu_t)) \, dt.$$

Note that the diffusion terms σdW_t cancel out exactly.

Proof of Proposition 1.2. We estimate the time derivative of the squared Euclidean distance:

$$\frac{1}{2} \frac{d}{dt} |X_t(u) - Y_t(v)|^2 = \langle X_t(u) - Y_t(v), a(X_t(u), \mu_t) - a(Y_t(v), \nu_t) \rangle.$$

We decompose the drift difference by introducing a mixed term $a(Y_t(v), \mu_t)$:

$$a(X_t, \mu_t) - a(Y_t, \nu_t) = \underbrace{a(X_t, \mu_t) - a(Y_t, \mu_t)}_{\text{Spatial difference}} + \underbrace{a(Y_t, \mu_t) - a(Y_t, \nu_t)}_{\text{Measure difference}}.$$

Using Assumption 2:

- Spatial term: $\langle X_t - Y_t, a(X_t, \mu_t) - a(Y_t, \mu_t)\rangle \leq L|X_t - Y_t|^2$.
- Measure term: $|a(Y_t, \mu_t) - a(Y_t, \nu_t)| \leq KW_2(\mu_t, \nu_t)$.

Thus,

$$\frac{1}{2}\frac{d}{dt}|X_t - Y_t|^2 \leq L|X_t - Y_t|^2 + |X_t - Y_t| \cdot KW_2(\mu_t, \nu_t).$$

Using the inequality $xy \leq \frac{1}{2}(x^2 + y^2)$, we bound $|X_t - Y_t|W_2(\mu_t, \nu_t) \leq \frac{1}{2}|X_t - Y_t|^2 + \frac{1}{2}W_2^2(\mu_t, \nu_t)$.

$$\frac{d}{dt}|X_t - Y_t|^2 \leq (2L + K)|X_t - Y_t|^2 + KW_2^2(\mu_t, \nu_t).$$

Now, integrate this inequality with respect to an optimal coupling π_0 of the initial measures (μ_0, ν_0). Let

$$E_t = \int |X_t(u) - Y_t(v)|^2 \pi_0(du, dv),$$

with $E_0 = W_2^2(\mu_0, \nu_0)$. Since μ_t and ν_t are the push-forwards of μ_0 and ν_0 by the flows, the pair (X_t, Y_t) constitutes a valid coupling for (μ_t, ν_t). By definition of the Wasserstein distance:

$$W_2^2(\mu_t, \nu_t) \leq E_t.$$

Substituting this into the differential inequality:

$$\frac{d}{dt}E_t \leq (2L + K)E_t + KE_t = (2L + 2K)E_t.$$

Using Grönwall's inequality, we obtain:

$$E_t \leq E_0 e^{(2L+2K)t} = W_2^2(\mu_0, \nu_0)e^{(2L+2K)t}.$$

Since $W_2^2(\mu_t, \nu_t) \leq E_t$, the result follows. □

3.2. **Expected Displacement.**

Remark 3.1 (Linear Growth). Assumption 2 implies that the drift $a(x, \mu)$ satisfies a linear growth condition. Indeed, fixing δ_0 as the Dirac measure at zero, we have:

$$\begin{aligned}|a(x, \mu)| &\leq |a(x, \mu) - a(x, \delta_0)| + |a(x, \delta_0) - a(0, \delta_0)| + |a(0, \delta_0)| \\ &\leq KW_2(\mu, \delta_0) + L|x| + |a(0, \delta_0)|.\end{aligned}$$

Noting that $W_2^2(\mu, \delta_0) = \int |y|^2 \mu(dy)$, this indicates that the drift is controlled by the second moment of the measure.

This observation allows us to prove the following a priori moment bound.

Lemma 3.1 (Moment Bounds)**.** *Assume that Assumption 2 holds and μ_0 has the second moment. Then for any $T > 0$, there exists a constant $C_T > 0$ (depending on T, L, K, σ and the second moment of μ_0) such that:*

$$\sup_{t\in[0,T]} \mathbb{E}\left[|X_t(u)|^2\right] \leq C_T(1 + |u|^2).$$

Proof. By applying Itô's formula to the function $f(x) = |x|^2$, we obtain:

$$|X_t(u)|^2 = |u|^2 + \int_0^t \left(2\langle X_s(u), a(X_s(u), \mu_s)\rangle + d\sigma^2\right) ds + 2\sigma \int_0^t X_s(u) \cdot dW_s.$$

Taking expectations eliminates the martingale term, and using $2\langle x, y\rangle \leq |x|^2 + |y|^2$ we obtain:

$$\mathbb{E}|X_t(u)|^2 \leq |u|^2 + \int_0^t \left(\mathbb{E}|X_s(u)|^2 + \mathbb{E}|a(X_s(u), \mu_s)|^2 + d\sigma^2\right) ds.$$

By Remark 3.1 and $(x+y+z)^2 \le 3(x^2+y^2+z^2)$, there exists a constant $C_1 > 0$ (depending on L and K) such that for all $x \in \mathbb{R}^d$ and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$,

$$|a(x,\mu)|^2 \le C_1 \left(1 + |x|^2 + W_2^2(\mu, \delta_0)\right) = C_1 \left(1 + |x|^2 + \int_{\mathbb{R}^d} |y|^2 \, \mu(dy)\right).$$

To make the dependence on the measure explicit in the integral, we introduce the moment quantity:

$$M_s := \int_{\mathbb{R}^d} \mathbb{E}\big[|X_s(v)|^2\big] \, \mu_0(dv).$$

Since $\mu_s = \mu_0 \circ X_s^{-1}$, we have almost surely

$$\int_{\mathbb{R}^d} |y|^2 \, \mu_s(dy) = \int_{\mathbb{R}^d} |X_s(v)|^2 \, \mu_0(dv).$$

Taking expectations, we see that $\mathbb{E}[\int |y|^2 \mu_s(dy)] = M_s$. Therefore

$$\mathbb{E}|a(X_s(u), \mu_s)|^2 \le C_1 \left(1 + \mathbb{E}|X_s(u)|^2 + M_s\right).$$

Substituting this into our main inequality, there exists a constant $C > 0$ such that for all $t \in [0,T]$:

$$\mathbb{E}|X_t(u)|^2 \le |u|^2 + C \int_0^t \left(1 + \mathbb{E}|X_s(u)|^2 + M_s\right) ds.$$

Integrating this inequality with respect to $\mu_0(du)$ gives

$$M_t \le \int_{\mathbb{R}^d} |u|^2 \, \mu_0(du) + C \int_0^t (1 + 2M_s) \, ds.$$

By Grönwall's inequality: $\sup_{t\in[0,T]} M_t < \infty$. Substituting this bound back into the estimate for $\mathbb{E}|X_t(u)|^2$ and applying Grönwall's inequality once more yields the claimed bound. □

With the moment bound established, we can now rigorously derive the estimate for the expected displacement.

Proof of Proposition 1.3. We estimate the Wasserstein distance using the coupling induced by the flow $u \mapsto X_t(u)$:

$$\mathbb{E}\left[W_2^2(\mu_t, \mu_0)\right] \le \int_{\mathbb{R}^d} \mathbb{E}|X_t(u) - u|^2 \mu_0(du).$$

Using the integral formulation of the SDE $X_t(u) - u = \int_0^t a(X_s(u), \mu_s) \, ds + \sigma W_t$, we square and take expectations:

$$\mathbb{E}|X_t(u) - u|^2 \le 2\mathbb{E}\left|\int_0^t a(X_s(u), \mu_s) \, ds\right|^2 + 2d\sigma^2 t.$$

Applying Hölder's inequality to the drift term:

$$\mathbb{E}\left|\int_0^t a(X_s(u), \mu_s) \, ds\right|^2 \le t \int_0^t \mathbb{E}|a(X_s(u), \mu_s)|^2 \, ds.$$

We now bound the integrand using the linear growth condition and Lemma 3.1. Using the constant C_1 from the proof of Lemma 3.1:

$$\mathbb{E}|a(X_s(u), \mu_s)|^2 \le C_1 \left(1 + \mathbb{E}|X_s(u)|^2 + \int \mathbb{E}|X_s(v)|^2 \mu_0(dv)\right).$$

Substituting the bound from Lemma 3.1 ($\mathbb{E}|X_s(u)|^2 \le C_T(1 + |u|^2)$):

$$\mathbb{E}|a(X_s(u), \mu_s)|^2 \le C_1 \left(1 + C_T(1 + |u|^2) + \int C_T(1 + |v|^2)\mu_0(dv)\right).$$

Integrating this bound with respect to $\mu_0(u)$ and using the finiteness of the second moment of μ_0, we obtain

$$\int_{\mathbb{R}^d} \mathbb{E}|X_t(u) - u|^2 \mu_0(du) \le 2t^2\tilde{C} + 2d\sigma^2 t,$$

where $\tilde{C}$ depends on T, L, K and the second moment of μ_0. For small t, the linear diffusion term dominates. □

References

1. Andrey A. Dorogovtsev, *Stochastic flows with interaction and measure-valued processes*, International Journal of Mathematics and Mathematical Sciences **2003** (2003), no. 63, 3963–3977. https://doi.org/10.1155/S0161171203301073
2. Hiroshi Kunita, *Lectures on stochastic flows and applications*, Tata Institute of Fundamental Research, Bombay, 1986.
3. Dirk Zimmermann, *Elementary proof of logarithmic Sobolev inequalities for Gaussian convolutions on R*, Annales mathématiques Blaise Pascal **23** (2016), no. 1, 129–140. https://doi.org/10.5802/ambp.360

Ukraine, Kyiv-4, 3, Tereschenkivska st.
E-mail address: kkuchynskyi@imath.kiev.ua