APPENDIX

A STATEMENTS OF REFERENCED THEOREMS

For a more comprehensive discussion and comparison, we provide the following related results from the literature.

Proposition A.1 (Caffarelli). *Let $\mu = \exp(-Q(x))dx$ and $\nu = \exp(-(Q(x) + V(x)))dx$ denote two Borel probability measures on Euclidean space $(\mathbb{R}^n, |\cdot|)$, where Q denotes a quadratic function, i.e.*

$$Q(x) = \langle Ax, x\rangle + \langle b, x\rangle + c,$$

with A positive-definite, and V is a convex function. Then the Brenier optimal-transport map $T = T_{\mathrm{opt}}$ pushing forward μ onto ν is a contraction:

$$\forall x, y \in \mathbb{R}^n, \quad |T(x) - T(y)| \leq |x - y|.$$

Proposition A.2 (Kim & Milman Thm.1.1). *Let $\mu = \exp(-U(x))dx$ and $\nu = \exp(-(U(x) + V(x)))dx$ denote two Borel probability measures on Euclidean space $(\mathbb{R}^n, |\cdot|)$. Assume that $U \in C^{3,\alpha}_{\mathrm{loc}}(\mathbb{R}^n)$ ($\alpha > 0$) is a convex function of the form:*

$$U(x) = Q(\mathrm{Proj}_{E_0} x) + \sum_{i=1}^{k} \rho_i(|\mathrm{Proj}_{E_i} x|), \quad \forall i = 1, \ldots, k, \ \rho_i''' \leq 0 \text{ on } \mathbb{R}_+,$$

where $Q : E_0 \to \mathbb{R}$ is a quadratic function, i.e.

$$Q(x) = \langle Ax, x\rangle + \langle b, x\rangle + c.$$

And that $V : \mathbb{R}^n \to \mathbb{R}$ is convex and satisfies our symmetry assumptions A.3. Then there exists a map $T : \mathbb{R}^n \to \mathbb{R}^n$ pushing forward μ onto ν and satisfying our symmetry assumptions, which is a contraction.

Definition A.3 (Kim & Milman symmetry assumptions). *We will say that a function $F : \mathbb{R}^n \to \mathbb{R}$ satisfies our symmetry assumptions if it is invariant under the action of the subgroup $O(E_1, \ldots, E_k) := 1 \times O(E_1) \times \ldots \times O(E_k)$ of the orthogonal group $O(n)$, or equivalently, if:*

$$\exists \Phi : \mathbb{R}^{\dim E_0 + k} \to \mathbb{R} \text{ so that } F(x) = \Phi(\mathrm{Proj}_{E_0} x, |\mathrm{Proj}_{E_1} x|, \ldots, |\mathrm{Proj}_{E_k} x|).$$

We will similarly say that a map $T : \mathbb{R}^n \to \mathbb{R}^n$ satisfies our symmetry assumptions if it commutes with the action of the latter subgroup.

Proposition A.4 (Kim & Milman Thm.1.3). *Proposition A.2 is also valid when replacing T with the Brenier optimal transport map T_{opt} pushing forward μ onto ν.*

Proposition A.5 (Colombo et al. Thm.1.1). *Let $V \in C^{1,1}_{loc}(\mathbb{R}^n)$ be such that $e^{-V(x)}dx \in \mathcal{P}(\mathbb{R}^n)$. Suppose that $V(0) = \inf_{\mathbb{R}^n} V$ and there exist constants $0 < \lambda, \Lambda < \infty$ for which $\lambda \mathrm{I_d} \leq D^2 V(x) \leq \Lambda \mathrm{I_d}$ for a.e. $x \in \mathbb{R}^n$. Moreover, let $R > 0$, $q \in C^0_c(B_R)$, and $c_q \in \mathbb{R}$ be such that $e^{-V(x)+c_q-q(x)}dx \in \mathcal{P}(\mathbb{R}^n)$. Assume that $-\lambda_q \mathrm{I_d} \leq D^2 q$ in the sense of distributions for some constant $\lambda_q \geq 0$. Then, there exists a constant $C_1 = C_1(R, \lambda, \Lambda, \lambda_q) > 0$, independent of n, such that the optimal transport map T that takes $e^{-V(x)}dx$ to $e^{-V(x)+c_q-q(x)}dx$ satisfies*

$$\|\nabla T\|_{L^\infty(\mathbb{R}^n)} \leq C_1.$$

Proposition A.6 (Neeman Thm.1.3). *Suppose that $d\mu = e^{-V(x)}d\gamma$ is a probability measure on $\mathbb{R}^n$, where $D^2 V \geq -\kappa$ (for $\kappa \geq 1$), and $\sup V - \inf V \leq c$. Then μ is an L-Lipschitz push-forward of γ, for*

$$L = 2(2\kappa)^{e^c}.$$

Proposition A.7 (Mikulincer & Shenfeld Thm.1). *Let μ be a κ-log-concave probability measure on $\mathbb{R}^d$, and set $D := diam(supp(\mu))$. Then, for the map $\varphi^{\mathrm{flow}} : \mathbb{R}^d \to \mathbb{R}^d$, which satisfies $\varphi^{\mathrm{flow}}_* \gamma_d = \mu$, the following holds:*

1. *If $\kappa > 0$ then*

$$\|\nabla \varphi^{\mathrm{flow}}(x)\|_{\mathrm{op}} \leq \frac{1}{\sqrt{\kappa}},$$

for μ-almost every x.

Table 2: Lip changes of variables via Heat flow

Target P_0	**Lip-constant**	**Result**
log-concave+sym.Ass.	1+sym.Ass.	Kim & Milman (2012) Prop.A.2
κ-log-concave+osc $\leq c$	$2(2\kappa)^{e^c}$	Neeman (2022) Prop. A.6
κ-log-concave	$e^{\frac{1-\kappa D^2}{2}}D$	Mikulincer & Shenfeld (2023) Prop. A.7
L-log-Lipschitz	$\frac{1}{\sqrt{\kappa}}e^{\left(\frac{L^2}{2\kappa}+2\frac{L}{\sqrt{\kappa}}\right)}$	Brigati & Pedrotti (2024) Prop. A.8
G-tail Ass. 3.7	$\frac{1}{\sqrt{\kappa}}e^{\left(\frac{L^2+L_1}{2\kappa}\right)}$	This work Cor. A.9

where $\kappa, c, D, L, L_1, K_1, K_2$ are dimension-independent constant.

2. *If* $\kappa D^2 < 1$ *then*

$$\|\nabla\varphi^{\text{flow}}(x)\|_{\text{op}} \leq e^{\frac{1-\kappa D^2}{2}}D,$$

for μ-almost every x.

Proposition A.8 (Brigati & Pedrotti Thm.1.4)**.** *Let $\mu = e^{-(W+H)} \in L^1_+(\mathbb{R}^d)$ be a probability density on $\mathbb{R}^d$ such that W is κ-convex for some $\kappa > 0$ and H is L-Lipschitz for some $L \geq 0$. Then, there exists a map $T^{\text{flow}} : \mathbb{R}^d \to \mathbb{R}^d$ such that $(T^{\text{flow}})_\#\gamma_d = \mu$ and T^{flow} is $\frac{1}{\sqrt{\kappa}}\exp\left(\frac{L^2}{2\kappa} + 2\frac{L}{\sqrt{\kappa}}\right)$-Lipschitz.*

Corollary A.9 (This work Cor. 3.11 with $A = (\kappa I_d)^{-1}$ and $C = I_d$)**.** *Let $\mu = e^{-(W(x)-H(x))}$ be a probability density on $\mathbb{R}^d$ such that $W(x) = \frac{\kappa|x|^2}{2}$ while $H(x)$ being L-Lipschitz and $\|\nabla^2 H\|_\infty < L_1$ for some $L, L_1 \geq 0$. Then, there exists a map $T^{\text{flow}} : \mathbb{R}^d \to \mathbb{R}^d$ such that $(T^{\text{flow}})_\#\gamma_d = \mu$ and T^{flow} is $\frac{1}{\sqrt{\kappa}}\exp\left(\frac{L^2+L_1}{2\kappa}\right)$-Lipschitz.*

Sketch of calculation. Applying Cor. 3.11 with $A = (\kappa I_d)^{-1}$ and $C = I_d$, Both A and C are diagonalizable, which leads to a more refined Lip upper bound:

$$\begin{aligned}
&\exp\left(\int_0^1 \frac{\|A\|^2(L^2+L_1)t}{(\|A\|t^2+\|C\|(1-t^2))^2}\,dt + \int_0^1 \frac{\|A-C\|t}{\|A\|t^2+\|C\|(1-t^2)}\,dt\right) \\
\leq &\exp\left(\frac{\|A\|(L^2+L_1)}{2\|C\|} + \frac{1}{2}\ln\frac{\|A\|}{\|C\|}\right) \\
\leq &\frac{1}{\sqrt{\kappa}}\exp\left(\frac{L^2+L_1}{2\kappa}\right).
\end{aligned}$$

□

Proposition A.10 (Albergo & Vanden-Eijnden Pro.3)**.** *Let $\rho_t(x)$ be the exact interpolant density and given a velocity field $\hat{v}_t(x)$, let us define $\hat{\rho}_t(x)$ as the solution of the initial value problem*

$$\partial_t\hat{\rho}_t + \nabla\cdot(\hat{v}_t\hat{\rho}_t) = 0, \quad \hat{\rho}_{t=0} = \rho_0.$$

Assume that $\hat{v}_t(x)$ is continuously differentiable in (t,x) and Lipschitz in x uniformly on $(t,x) \in [0,1]\times\mathbb{R}^d$ with Lipschitz constant K. Then the square of the W_2 distance between ρ_1 and $\hat{\rho}_1$ is bounded by

$$W_2^2(\rho_1,\hat{\rho}_1) \leq e^{1+2K}H(\hat{v}) \tag{17}$$

where $H(\hat{v})$ is the objective function defined as

$$H(\hat{v}) = \int_0^1\int_{\mathbb{R}^d}|\hat{v}_t(x) - v_t(x)|^2\,\rho_t(x)\,dx\,dt.$$

Proposition A.11 (Albergo et al. Thm.2.23)**.** *Let ρ denote the solution of the Fokker-Planck equation (2.20) with $\epsilon(t) = \epsilon > 0$. Given two velocity fields $\hat{b}, \hat{s} \in C^0([0,1],(C^1(\mathbb{R}^d))^d)$, define*

$$\hat{b}_F(t,x) = \hat{b}(t,x) + \epsilon\hat{s}(t,x), \quad \hat{v}(t,x) = \hat{b}(t,x) + \gamma(t)\dot{\gamma}(t)\hat{s}(t,x),$$

where the function γ satisfies the properties listed in Definition 1. Let $\hat{\rho}$ denote the solution to the Fokker-Planck equation

$$\partial_t \hat{\rho} + \nabla \cdot (\hat{b}_F \hat{\rho}) = \epsilon \Delta \hat{\rho}, \quad \hat{\rho}(0) = \rho_0.$$

Then,

$$KL(\rho_1 \| \hat{\rho}(1)) \leq \frac{1}{2\epsilon} \left(\mathcal{L}_{\hat{b}}[\hat{b}] - \min_{\widetilde{b}} \mathcal{L}_{\hat{b}}[\widetilde{b}] \right) + \frac{\epsilon}{2} \left(\mathcal{L}_{\hat{s}}[\hat{s}] - \min_{\widetilde{s}} \mathcal{L}_{\hat{s}}[\widetilde{s}] \right),$$

where $\mathcal{L}_{\hat{b}}[\hat{b}]$ and $\mathcal{L}_{\hat{s}}[\hat{s}]$ are the objective functions defined in

$$\mathcal{L}_b[\hat{b}] = \int_0^1 \mathbb{E} \left(\frac{1}{2} |\hat{b}(t, x_t)|^2 - (\partial_t I(t, x_0, x_1) + \dot{\gamma}(t) z) \cdot \hat{b}(t, x_t) \right) dt$$

and

$$\mathcal{L}_s[\hat{s}] = \int_0^1 \mathbb{E} \left(\frac{1}{2} |\hat{s}(t, x_t)|^2 + \gamma^{-1}(t) z \cdot \hat{s}(t, x_t) \right) dt.$$

And

$$KL(\rho_1 \| \hat{\rho}(1)) \leq \frac{1}{2\epsilon} \left(\mathcal{L}_{\hat{v}}[\hat{v}] - \min_{\widetilde{v}} \mathcal{L}_{\hat{v}}[\widetilde{v}] \right) + \frac{\sup_{t \in [0,1]} (\gamma(t) \dot{\gamma}(t) - \epsilon)^2}{2\epsilon} \left(\mathcal{L}_{\hat{s}}[\hat{s}] - \min_{\widetilde{s}} \mathcal{L}_{\hat{s}}[\widetilde{s}] \right).$$

where $\mathcal{L}_{\hat{v}}[\hat{v}]$ is the objective function defined in

$$\mathcal{L}_v[\hat{v}] = \int_0^1 \mathbb{E} \left(\frac{1}{2} |\hat{v}(t, x_t)|^2 - \partial_t I(t, x_0, x_1) \cdot \hat{v}(t, x_t) \right) dt.$$

Proposition A.12 (Boffi et al. Prop.3.9: Lagrangian error bound). *Let $X_{s,t} : \mathbb{R}^d \to \mathbb{R}^d$ denote the flow map for a pre-trained stochastic interpolant or diffusion model, and let $\hat{X}_{s,t} : \mathbb{R}^d \to \mathbb{R}^d$ denote an approximate flow map. Given $x_0 \sim \rho_0$, let $\hat{\rho}_1$ be the density of $\hat{X}_{0,1}(x_0)$ and let ρ_1 be the target density of $X_{0,1}(x_0)$. Then,*

$$W_2^2(\rho_1, \hat{\rho}_1) \leq e^{1 + 2\int_0^1 |C_t| dt} \mathcal{L}_{LMD}(\hat{X}),$$

where C_t is the Lipschitz constant of the drift term.

Proposition A.13 (Boffi et al. Prop.3.10: Eulerian error bound). *Let $X_{s,t} : \mathbb{R}^d \to \mathbb{R}^d$ denote the flow map for a pre-trained stochastic interpolant or diffusion model, and let $\hat{X}_{s,t} : \mathbb{R}^d \to \mathbb{R}^d$ denote an approximate flow map. Given $x_0 \sim \rho_0$, let $\hat{\rho}_1$ be the density of $\hat{X}_{0,1}(x_0)$ and ρ_1 be the target density of $X_{0,1}(x_0)$. Then,*

$$W_2^2(\rho_1, \hat{\rho}_1) \leq e^1 \mathcal{L}_{EMD}(\hat{X}). \tag{18}$$

Proposition A.14 (Benton et al. Thm.4). *Suppose that the following Assumptions hold,*

- *The true and approximate drifts $v^X(\mathbf{x}, t)$ and $v_\theta(\mathbf{x}, t)$ satisfy $\int_0^1 \mathbb{E} \left[\left\| v_\theta(X_t, t) - v^X(X_t, t) \right\|^2 \right] dt \leq \epsilon^2$.*
- *For each $\mathbf{x} \in \mathbb{R}^d$ and $s \in [0,1]$ there exist unique flows $(Y_{s,t}^X)_{t \in [s,1]}$ and $(Z_{s,t}^X)_{t \in [s,1]}$ starting in $Y_{s,s}^X = \mathbf{x}$ and $Z_{s,s}^X = \mathbf{x}$ with velocity fields $v_\theta(\mathbf{x}, t)$ and $v^X(\mathbf{x}, t)$ respectively. Moreover, $Y_{s,t}^X$ and $Z_{s,t}^X$ are continuously differentiable in $\mathbf{x}$, s and t.*
- *The approximate flow $v_\theta(\mathbf{x}, t)$ is differentiable in both inputs. Also, for each $t \in (0,1)$ there is a constant L_t such that $v_\theta(\mathbf{x}, t)$ is L_t-Lipschitz in $\mathbf{x}$.*
- *For some $\lambda \geq 1$, $\alpha_t X_0 + \beta_t X_1$ is λ-regular for all $t \in [0,1]$.*

γ_t is a concave function on $[0,1]$ which determines the amount of Gaussian smoothing applied at time t, and that $\alpha_0 = \beta_1 = 1$ and $\gamma_0 = \gamma_1 = \gamma_{\min}$. Then, for any $v_\theta \in \mathcal{V}$, if Y is a flow starting in $\hat{\pi}_0$ with velocity field v_θ and $\hat{\pi}_1$ is the law of Y_1,

$$W_2(\hat{\pi}_1, \pi_1) \leq C^{\lambda^{1/2}} \varepsilon \left(\frac{\gamma_{\max}}{\gamma_{\min}} \right)^{2\lambda} + \sqrt{d} \gamma_{\min}$$

where $C = \exp\left\{ R \left(\int_0^1 (|\dot{\alpha}_t| / \gamma_t)\, dt + \int_0^1 (|\dot{\beta}_t| / \gamma_t)\, dt \right) \right\}$ with $\gamma_{\min} = \inf_{t \in [0,1]} \gamma_t$, $\gamma_{\max} = \sup_{t \in [0,1]} \gamma_t$.

Proposition A.15 (Cheng et al. Prop.5.4). *Suppose $q_N = \widetilde{q}_N = q \in \mathcal{P}_2^\gamma$, and the computed transport maps T_n and S_n satisfy the following assumptions:*

- *The learned transport T_{n+1} is non-degenerate and in $L^2(p_n)$; it is invertible on $\mathbb{R}^d$ and T_{n+1}^{-1} is also non-degenerate. In addition, for some $\varepsilon > 0$, $\exists \xi_{n+1} \in \partial_{W_2} F_{n+1}(p_{n+1})$ s.t. $\|\xi_{n+1}\|_{p_{n+1}} \leq \varepsilon$.*
- *For $n = N, \ldots, 1$, the computed reverse transport S_n is non-degenerate, in $L^2(\widetilde{q}_n)$, and satisfies that*
$$\|T_n \circ S_n - I_d\|_{\widetilde{q}_n} \leq \varepsilon_{inv}.$$
- *There is $K > 0$ s.t. T_n^{-1} is Lipschitz on $\mathbb{R}^d$ with Lipschitz constant $e^{\gamma K}$ for all $n = N, \ldots, 1$.*

Then all q_n and $\widetilde{q}_n$ are in $\mathcal{P}_2^\gamma$ and
$$W_2(\widetilde{q}_0, q_0) \leq \frac{\varepsilon_{inv}}{\gamma K} e^{\gamma K(N+1)},$$
where ε_{inv} denotes Inversion error and $e^{\gamma K}$ is the Lipschitz constant of inverse transport map.

Proposition A.16 (Gao et al. Thm.1.2). *Suppose that the target distribution has a bounded support, or is strongly log-concave, or is a mixture of Gaussians. Let n be the sample size and $0 < \underline{t} \ll 1$ satisfying $\underline{t} \approx n^{-1/(d+5)}$. By properly setting the deep ReLU network structure and the forward Euler discretization step sizes, the distribution estimation error of the CNFs learned with linear interpolation and flow matching is evaluated by*
$$\mathbb{E}W_2(\hat{p}_{1-\underline{t}}, p_1) = \mathcal{O}(n^{-\frac{1}{d+5}}),$$
where the expectation is taken with respect to all random samples, $\hat{p}_{1-\underline{t}}$ is the law of generated data, p is the law of target data, $W_2(\cdot,\cdot)$ is the Wasserstein-2 distance, and a polylogarithmic prefactor in n is omitted.

Definition A.17 (Cattiaux & Guillin). *A probability measure $\mu(dx) = \exp(-U(x))dx$ is κ-semi-log-concave for some $\kappa \in \mathbb{R}$ if its support $\Omega \subseteq \mathbb{R}^d$ is convex and $U \in C^2(\Omega)$ satisfies*
$$\nabla^2 U(x) \succeq \kappa I_d, \quad \forall x \in \Omega.$$

Definition A.18 (Eldan & Lee). *A probability measure $\mu(dx) = \exp(-U(x))dx$ is β-semi-log-convex for some $\beta > 0$ if its support $\Omega \subseteq \mathbb{R}^d$ is convex and $U \in C^2(\Omega)$ satisfies*
$$\nabla^2 U(x) \preceq \beta I_d, \quad \forall x \in \Omega.$$

Assumption A.19 (Dai et al. Semi-log-convexity). *ν is β-semi-log-convex for some $\beta > 0$.*

Assumption A.20 (Dai et al. Structural condition). *Set $D := \frac{1}{\sqrt{2}}\operatorname{diam}(\operatorname{supp}(\nu))$. One of the following holds:*

1. *ν is κ-semi-log-concave for some $\kappa > 0$ with $D \in (0, \infty]$;*
2. *ν is κ-semi-log-concave for some $\kappa \leq 0$ with $D \in (0, \infty)$;*
3. *$\nu = \mathcal{N}(0, \sigma^2 I_d) * \rho$ with a probability ρ supported on a ball of radius R in R^d.*

B Proofs

In this part, we provide the detailed proofs of the theories in this paper.

B.1 Proof of Theorem 3.5

Proof. Recall Monge's Optimal Transport (OT) problem (Monge, 1781), which seeks a map T pushing μ to ν that minimizes $\displaystyle\int c(x, T(x))\, d\mu(x)$ subject to $T_\sharp \mu = \nu$, inducing the p-Wasserstein distance
$$\mathcal{W}_p(\mu, \nu) := \Big(\inf_{\gamma \in \Gamma(\mu,\nu)} \mathbb{E}_{(x,y)\sim\gamma}[\|x - y\|^p]\Big)^{1/p},$$

where $\Gamma(\mu,\nu)$ denotes the set of couplings with marginals μ and ν.

Applying this to the push-forward measures $\overleftarrow{P}_{t_{n+1}}$ and $\overleftarrow{Q}_{t_{n+1}}$ (cf. definition (4)) at $[t_n, t_{n+1}]$ gives

$$
\begin{aligned}
&\mathcal{W}_2(\overleftarrow{P}_{t_{n+1}}, \overleftarrow{Q}_{t_{n+1}}) \\
\leq &\sqrt{\mathbb{E}_{(\overleftarrow{Y}_{t_n}, \overleftarrow{X}_{t_n}) \sim \Gamma(\overleftarrow{Q}_{t_n}, \overleftarrow{P}_{t_n})} |\widetilde{T}_n(\overleftarrow{Y}_{t_n}) - T_n(\overleftarrow{X}_{t_n})|^2} \\
\leq &\sqrt{\mathbb{E}_{(\overleftarrow{Y}_{t_n}, \overleftarrow{X}_{t_n}) \sim \Gamma(\overleftarrow{Q}_{t_n}, \overleftarrow{P}_{t_n})} |\widetilde{T}_n(\overleftarrow{Y}_{t_n}) - \widetilde{T}_n(\overleftarrow{X}_{t_n})|^2} + \sqrt{\mathbb{E}_{\overleftarrow{X}_{t_n} \sim \overleftarrow{P}_{t_n}} |\widetilde{T}_n(\overleftarrow{X}_{t_n}) - T_n(\overleftarrow{X}_{t_n})|^2} \\
\leq &\mathrm{Lip}(\widetilde{T}_n)\sqrt{\mathbb{E}_{(\overleftarrow{Y}_{t_{n-1}}, \overleftarrow{X}_{t_{n-1}}) \sim \Gamma(\overleftarrow{Q}_{t_{n-1}}, \overleftarrow{P}_{t_{n-1}})} |\widetilde{T}_{n-1}(\overleftarrow{Y}_{t_{n-1}}) - \widetilde{T}_{n-1}(\overleftarrow{X}_{t_{n-1}})|^2} \\
&+ \left(\left(\overline{K}\sqrt{M_0} + \frac{\overline{K_1}}{\sqrt{1-t_n^2}} + \overline{K_2} \right) h + \epsilon \right),
\end{aligned}
$$

where the last inequality uses Assumption 3.3 and Assumption 3.4. Then applying the discrete Grönwall inequality yields (10). □

B.2 Proof of Theorem 3.8

Proof. For simplicity, denote

$$
G(t,x,y) := \exp\left(-\frac{|K(t)(\sqrt{C})^{-1}x - y|^2_{B(t)}}{2} \right),\ g(t,x,y) := \partial_t \left(-\frac{|K(t)(\sqrt{C})^{-1}x - y|^2_{B(t)}}{2} \right).
$$

Under the Gaussian tail Assumption 3.7, the score function of Föllmer flow can be calculated by

$$
S(t,x) := C\nabla \log p_t = C\nabla \log \int_{\mathbb{R}^d} (2\pi \det(B(t)))^{-\frac{d}{2}}\, G(t,x,y) \cdot \exp\left(-\frac{|x|^2_{\overline{A}_t}}{2} \right) \exp\left(h(\sqrt{C}y) \right) \mathrm{d}y,
$$

where $\overline{A}_t = At^2 + C(1-t^2)$, $K(t) = (A\overline{A}_t^{-1})t$, $B(t) = (A\overline{A}_t^{-1})(1-t^2)$.

First, we consider the modified score function over the time interval $(0, 1]$,

$$
\begin{aligned}
\widetilde{S}(t,x) &= S(t,x) + C\overline{A}_t^{-1}x \\
&= C\nabla \log \int_{\mathbb{R}^d} (2\pi B(t))^{-\frac{d}{2}}\, G(t,x,y) \exp\left(h(\sqrt{C}y) \right) \mathrm{d}y \\
&= C\frac{\int_{\mathbb{R}^d} \nabla_x G(t,x,y) \exp\left(h(\sqrt{C}y) \right) \mathrm{d}y}{\int_{\mathbb{R}^d} G(t,x,y) \exp\left(h(\sqrt{C}y) \right) \mathrm{d}y} \\
&= -\frac{K(t)\sqrt{C}\int_{\mathbb{R}^d} \nabla_y G(t,x,y) \exp\left(h(\sqrt{C}y) \right) \mathrm{d}y}{\int_{\mathbb{R}^d} G(t,x,y) \exp\left(h(\sqrt{C}y) \right) \mathrm{d}y} \\
&= \frac{K(t)\sqrt{C}\int_{\mathbb{R}^d} G(t,x,y) \nabla_y h(\sqrt{C}y) \exp\left(h(\sqrt{C}y) \right) \mathrm{d}y}{\int_{\mathbb{R}^d} G(t,x,y) \exp\left(h(\sqrt{C}y) \right) \mathrm{d}y},
\end{aligned}
\tag{19}
$$

Here, the last equal sign is derived from integration by parts.

Since $G(t,x,y)\exp\left(h(\sqrt{C}y) \right) \geq 0$,

$$
|\widetilde{S}(t,\cdot)| \leq |K(t)\sqrt{C}\nabla_y h(\sqrt{C}y)|.
$$

Let $K = \sup_{0\leq t\leq 1} |\frac{1}{t}K(t)| = \sup_{0\leq t\leq 1} |A\overline{A}_t^{-1}| \leq \max\{1, \|AC^{-1}\|\}$, we have

$$|\widetilde{S}(t,\cdot)| \leq K\|C\|^{1/2}|\sqrt{C}\nabla h|_\infty t = K_0 t.$$

Taking the derivative twice along that direction and using the same method as above, we get :

$$\|\nabla\widetilde{S}(t,\cdot)\|_\infty \leq K(t)^2(\|C\nabla^2 h\|_\infty + |\sqrt{C}\nabla h|_\infty^2) = K_1 t^2.$$

where $K_1 := K^2(\|C\nabla^2 h\|_\infty + |\sqrt{C}\nabla h|_\infty^2)$.

Define $K_2 := \sup_{0\leq t\leq 1} \|\frac{1}{t^2}(I - C\overline{A}_t^{-1})\| = \sup_{0\leq t\leq 1} \|(A-C)(At^2 + C(1-t^2))^{-1}\|$, then $\|(I - C\overline{A}_t^{-1})\| \leq K_2 t^2$.

Recall definition of $V(t,x)$ in (8), we have

$$|V(t,x)| = \left|\frac{x + S(t,\cdot)}{t}\right| = \left|\frac{\widetilde{S}(t,\cdot) + (I - C\overline{A}_t^{-1})x}{t}\right| \leq K_0 + K_2 t|x|,$$

which implies the velocity field $|V(t,x)|$ remains locally uniformly bounded and grows at most linearly.

Furthermore,

$$\|\nabla V(t,x)\|_\infty = \left\|\nabla\left(\frac{\widetilde{S}(t,x) + (I - C\overline{A}_t^{-1})x}{t}\right)\right\|_\infty \leq (K_1 + K_2)t,$$

which yields a Lipschitz constant that is uniform over space and independent of the dimension, ensuring uniform equicontinuity.

Next, we give the properties of $V(t,x)$ with respect to time over $(0,1)$. By taking (19), we have

$$\begin{aligned}
\partial_t V(t,x) =& \partial_t\left(\frac{\widetilde{S}(t,\cdot) + (I - C\overline{A}_t^{-1})x}{t}\right) \\
=& \partial_t\left(\frac{K(t)\sqrt{C}\int_{\mathbb{R}^d} G(t,x,y)\nabla_y h(\sqrt{C}y)\exp\left(h(\sqrt{C}y)\right)\mathrm{d}y}{t\int_{\mathbb{R}^d} G(t,x,y)\exp\left(h(\sqrt{C}y)\right)\mathrm{d}y}\right) + \partial_t\left(\frac{(I - C\overline{A}_t^{-1})x}{t}\right) \\
:=& A_1 + A_2
\end{aligned} \tag{20}$$

We begin by calculating the first part,

$$
\begin{aligned}
A_1 =& \partial_t \left(\frac{(A\bar{A}_t^{-1})\sqrt{C} \int_{\mathbb{R}^d} G(t,x,y) \nabla_y h(\sqrt{C}y) \exp\Big(h(\sqrt{C}y)\Big) \,\mathrm{d}y}{\int_{\mathbb{R}^d} G(t,x,y) \exp\Big(h(\sqrt{C}y)\Big) \,\mathrm{d}y} \right) \\
=& \frac{\partial_t (A\bar{A}_t^{-1})\sqrt{C} \int_{\mathbb{R}^d} G(t,x,y) \nabla_y h(\sqrt{C}y) \exp\Big(h(\sqrt{C}y)\Big) \,\mathrm{d}y}{\int_{\mathbb{R}^d} G(t,x,y) \exp\Big(h(\sqrt{C}y)\Big) \,\mathrm{d}y} \\
&+ \frac{A\bar{A}_t^{-1}\sqrt{C} \int_{\mathbb{R}^d} G(t,x,y) g(t,x,y) \exp\Big(h(\sqrt{C}y)\Big) \nabla_y h(\sqrt{C}y) dy}{\int_{\mathbb{R}^d} G(t,x,y) \exp\Big(h(\sqrt{C}y)\Big) \,\mathrm{d}y} \\
&- \frac{A\bar{A}_t^{-1}\sqrt{C} \int_{\mathbb{R}^d} G(t,x,y) \exp\Big(h(\sqrt{C}y)\Big) \nabla_y h(\sqrt{C}y) dy}{\int_{\mathbb{R}^d} G(t,x,y) \exp\Big(h(\sqrt{C}y)\Big) \,\mathrm{d}y} \\
&\cdot \frac{\int_{\mathbb{R}^d} G(t,x,y) g(t,x,y) \exp\Big(h(\sqrt{C}y)\Big) dy}{\int_{\mathbb{R}^d} G(t,x,y) \exp\Big(h(\sqrt{C}y)\Big) \,\mathrm{d}y} \\
\le& \frac{\partial_t (A\bar{A}_t^{-1})\sqrt{C} \int_{\mathbb{R}^d} G(t,x,y) \nabla_y h(\sqrt{C}y) \exp\Big(h(\sqrt{C}y)\Big) \,\mathrm{d}y}{\int_{\mathbb{R}^d} G(t,x,y) \exp\Big(h(\sqrt{C}y)\Big) \,\mathrm{d}y} \\
&+ A\bar{A}_t^{-1}\sqrt{C}\, \mathrm{Cov}_{p_t(y|x)} \Big(g(t,x,y), \nabla_y h(\sqrt{C}y) \Big) \\
\le& \frac{\partial_t (A\bar{A}_t^{-1})\sqrt{C} \int_{\mathbb{R}^d} G(t,x,y) \nabla_y h(\sqrt{C}y) \exp\Big(h(\sqrt{C}y)\Big) \,\mathrm{d}y}{\int_{\mathbb{R}^d} G(t,x,y) \exp\Big(h(\sqrt{C}y)\Big) \,\mathrm{d}y} \\
&+ A\bar{A}_t^{-1}\sqrt{C} \sqrt{\mathrm{Var}\left(g(t,x,y)\right)} \sqrt{\mathrm{Var}\Big(\nabla_y h(\sqrt{C}y)\Big)} \\
:=& I_1 + I_2,
\end{aligned}
\tag{21}
$$

where

$$
p_t(y|x) = \frac{G(t,x,y) \exp\Big(h(\sqrt{C}y)\Big)}{\int_{\mathbb{R}^d} G(t,x,y) \exp\Big(h(\sqrt{C}y)\Big) \,\mathrm{d}y}.
$$

Define $K_3 := \sup_{0 \le t \le 1} \| \frac{A\partial_t(\bar{A}_t^{-1})}{t} \| = \sup_{0 \le t \le 1} \| 2A(A-C)((At^2 + C(1-t^2))^2)^{-1} \|$, we obtain the first term of A_1,

$$
\begin{aligned}
|I_1| &\le \left\| 2A(A-C)((At^2 + C(1-t^2))^2)^{-1} t \right\| |\sqrt{C}| |\sqrt{C}\nabla h|_\infty \\
&\le |\sqrt{C}| |\sqrt{C}\nabla h|_\infty K_3 t \\
&\le K_0 K_3 K^{-1}.
\end{aligned}
\tag{22}
$$

For the second term I_2 in (21), the analysis is carried out separately for the following two cases:

Case I **-** $\|B(t)\| \le \frac{1}{2\|C\nabla^2 h\|_\infty}$**:** Then $p_t(y|x)$ is a log-concave measure as $-\nabla^2 \log p_t(y|x) = B(t)^{-1} - C\nabla^2 h \succ 0$. Then, by Brascamp-Lieb inequality, we have

$$
\sqrt{\mathrm{Var}\left(g(t,x,y)\right)} \le \sqrt{\mathbb{E}_{p_t(y|x)} |\nabla^T g(t,x,y) \Big(-\nabla^2 \log p_t(y|x) \Big)^{-1} \nabla g(t,x,y)|}, \tag{23}
$$

in which

$$\nabla g(t,x,y) = \partial_t K(t)B(t)^{-1}(\sqrt{C})^{-1}x - \partial_t B(t)B(t)^{-2}(K(t)(\sqrt{C})^{-1}x - y).$$

We have the following estimate for the right-hand side of (23) via integration by parts,

$$\begin{aligned}
&\text{RHS of (23)}\\
\leq &\int_{\mathbb{R}^d} \left(\partial_t K(t)B(t)^{-1}(\sqrt{C})^{-1}x - \partial_t B(t)B(t)^{-2}(K(t)(\sqrt{C})^{-1}x - y)\right)^T (B(t)^{-1} - C\nabla^2 h)^{-1}\\
&\quad \cdot \left(\partial_t K(t)B(t)^{-1}(\sqrt{C})^{-1}x - \partial_t B(t)B(t)^{-2}(K(t)(\sqrt{C})^{-1}x - y)\right)p_t(y|x)\mathrm{d}y\\
\leq &\sqrt{|\partial_t K(t)x|^2\|C\|^{-1} + 2\|\partial_t K(t)\|\|\partial_t B(t)\||\nabla h|_\infty|x| + \|\partial_t B(t)\|^2(\|C\nabla^2 h\|_\infty + |\sqrt{C}\nabla h|_\infty^2)}\\
&\quad \cdot \sqrt{(\|B(t)\|\left(1 - \|B(t)\|\|C\nabla^2 h\|_\infty)\right)^{-1}} + \|\partial_t B(t)\|\sqrt{\left(\|B(t)\|^2(1 - \|B(t)\|\|C\nabla^2 h\|_\infty)\right)^{-1}}.
\end{aligned}$$

Combining with

$$\sqrt{\mathrm{Var}\left(\nabla_y h(\sqrt{C}y)\right)} \leq \|C\nabla^2 h\|_\infty \|\sqrt{B(t)}\|\sqrt{(1 - \|B(t)\|\|C\nabla^2 h\|_\infty)^{-1}},$$

we derive that

$$\begin{aligned}
|I_2| \leq &K\|C\nabla^2 h\|_\infty (1 - \|B(t)\|\|C\nabla^2 h\|_\infty)^{-1}\Big(|\partial_t K(t)x| + \sqrt{2\|\partial_t K(t)\|\|\partial_t B(t)\||C\nabla h|_\infty |x|}\\
&+ \|\partial_t B(t)\|\sqrt{(\|C\nabla^2 h\|_\infty + |\sqrt{C}\nabla h|_\infty^2)}\Big)\\
&+ \|\sqrt{B(t)^{-1}}\|K\|\sqrt{C}\|\|\partial_t B(t)\|\|C\nabla^2 h\|_\infty (1 - \|B(t)\|\|C\nabla^2 h\|_\infty)^{-1}\\
\leq &2K\|C\nabla^2 h\|_\infty \big(|\partial_t K(t)x| + \sqrt{2\|\partial_t K(t)\|\|\partial_t B(t)\||\nabla h|_\infty |x|}\\
&+ \|\partial_t B(t)\|\sqrt{(\|C\nabla^2 h\|_\infty + |\sqrt{C}\nabla h|_\infty^2}\big) + 2K\|\sqrt{C}\|\|\partial_t B(t)\|\|C\nabla^2 h\|_\infty \|\sqrt{B(t)^{-1}}\|\\
\leq &2K_1 C_1 |x| + 2\sqrt{2}K_1\sqrt{K_0}\sqrt{C_1}\sqrt{C_2}\sqrt{|x|}\\
&+ 2K_1^{\frac{3}{2}}C_2 + 2K_1\|\sqrt{C}\|C_2\|\sqrt{B(t)^{-1}}\|\\
\leq &3K_1 C_1 |x| + \frac{2K^{-1}K_1\|\sqrt{C}\|C_2}{\sqrt{1-t^2}} + C_3,
\end{aligned}$$

where the second inequality use the fact $(1 - \|B(t)\|\|C\nabla^2 h\|_\infty)^{-1} \leq 2$ under Case I, while the last inequality is obtained using Young's inequality, C_1, C_2, C_3 are dimension-free constants defined in table 3.

Case II **-** $\|B(t)\| > \frac{1}{2\|C\nabla^2 h\|}$**:** According to the definition of variance

$$\begin{aligned}
\mathrm{Var}(g(t,x,y)) &= \int_{\mathbb{R}^d} (g(t,x,y))^2 p_t(y|x)\mathrm{d}y - \left(\int_{\mathbb{R}^d} g(t,x,y)p_t(y|x)\mathrm{d}y\right)^2\\
&\leq \int_{\mathbb{R}^d} (g(t,x,y))^2 p_t(y|x)\mathrm{d}y,
\end{aligned}$$

We have the following estimate, analogous to the right-hand side of (23):

$$
\begin{aligned}
&\sqrt{\mathrm{Var}\left(g(t,x,y)\right)}\\
\le&\frac{\|\sqrt{B(t)^{-1}}\|}{2}\Big(2\|\sqrt{C}^{-1}\||\partial_t K(t)x| + 2\sqrt{3}\sqrt{\|\partial_t K(t)\|\|\partial_t B(t)\||\nabla h|_\infty |x|}\\
&+\sqrt{6}\|\partial_t B(t)\|\sqrt{\|C\nabla^2 h\|_\infty + |\sqrt{C}\nabla h|_\infty^2}\Big)\\
&+|\partial_t K(t)x|\sqrt{\|\nabla^2 h\|_\infty + |\nabla h|_\infty^2} + \sqrt{\|\partial_t K(t)\|\|\partial_t B(t)\||\nabla h|_\infty(\|C\nabla^2 h\|_\infty + |\sqrt{C}\nabla h|_\infty^2)|x|}\\
&+\frac{1}{2}\|\partial_t B(t)\|\sqrt{|\sqrt{C}\nabla h|_\infty(\|C\nabla^2 h\|_\infty + |\sqrt{C}\nabla h|_\infty^2)} + \frac{\sqrt{3}}{2}\|\partial_t B(t)\|\|B(t)^{-1}\|\\
\le&\sqrt{\frac{\|\nabla^2 h\|_\infty}{2}}\Big(2|\partial_t K(t)x| + 2\sqrt{3}\sqrt{\|\partial_t K(t)\|\|\partial_t B(t)\||C\nabla h|_\infty |x|}\\
&+\sqrt{6}\|\sqrt{C}\|\|\partial_t B(t)\|\sqrt{\|C\nabla^2 h\|_\infty + |\sqrt{C}\nabla h|_\infty^2}\Big)\\
&+|\partial_t K(t)x|\sqrt{(\|\nabla^2 h\|_\infty + |\nabla h|_\infty^2)} + \sqrt{\|\partial_t K(t)\|\|\partial_t B(t)\||\nabla h|_\infty(\|C\nabla^2 h\|_\infty + |\sqrt{C}\nabla h|_\infty^2)|x|}\\
&+\frac{1}{2}\|\partial_t B(t)\|\sqrt{|\sqrt{C}\nabla h|_\infty(\|C\nabla^2 h\|_\infty + |\sqrt{C}\nabla h|_\infty^2)} + \sqrt{3}\|\partial_t B(t)\|\|C\nabla^2 h\|_\infty.
\end{aligned}
$$

Together with

$$\sqrt{\mathrm{Var}\left(\nabla_y h(\sqrt{C}y)\right)} \le |\sqrt{C}\nabla h|_\infty,$$

we obtain

$$
\begin{aligned}
|I_2| \le& K|\sqrt{C}\nabla h|_\infty\Bigg(|\partial_t K(t)x|\big(\sqrt{2\|C\nabla^2 h\|_\infty} + \sqrt{(\|C\nabla^2 h\|_\infty + |\sqrt{C}\nabla h|_\infty^2)}\big)\\
&\quad+\sqrt{6}\sqrt{\|C\nabla^2 h\|_\infty}\sqrt{\|\partial_t K(t)\|\|\partial_t B(t)\||C\nabla h|_\infty |x|}\\
&\quad+\sqrt{\|\partial_t K(t)\|\|\partial_t B(t)\|\sqrt{C}\nabla h|_\infty(\|C\nabla^2 h\|_\infty + |\sqrt{C}\nabla h|_\infty^2)|x|}\\
&\quad+\sqrt{3}\|\sqrt{C}\|\sqrt{\|C\nabla^2 h\|_\infty}\|\|\partial_t B(t)\|\|\sqrt{\|C\nabla^2 h\|_\infty + |\sqrt{C}\nabla h|_\infty^2}\\
&\quad+\frac{1}{2}\|\sqrt{C}\|\|\partial_t B(t)\|\sqrt{|\sqrt{C}\nabla h|_\infty(\|C\nabla^2 h\|_\infty + |\sqrt{C}\nabla h|_\infty^2)} + \sqrt{3}\|\sqrt{C}\|\|\partial_t B(t)\|\|C\nabla^2 h\|_\infty\Bigg)\\
\le&(1+\sqrt{2})K_0\sqrt{K_1}C_1|x| + (1+\sqrt{6})K_0^{\frac{3}{2}}\sqrt{K_1}\sqrt{C_1}\sqrt{C_2}|\sqrt{x}|\\
&\quad+K_0C_2(2\sqrt{3}K_1 + \frac{1}{2}\sqrt{K_0}\sqrt{K_1})\\
\le&2(1+\sqrt{2})K_0\sqrt{K_1}C_1|x| + C_4,
\end{aligned}
$$

where C_4 defined in table 3 is also constant indepent of dimension.

Combining the above two cases, we obtain

$$|I_2| \le \max\{3K_1C_1, 2(1+\sqrt{2})K_0\sqrt{K_1}C_1\}|x| + \frac{2K^{-1}K_1\|\sqrt{C}\|C_2}{\sqrt{1-t^2}} + \max\{C_3, C_4\}. \tag{24}$$

Using (21), (22) and (24), we derive

$$
\begin{aligned}
|A_1| \le& |I_1| + |I_2|\\
\le& K_0K_3K^{-1} + \max\{3K_1C_1, 2(1+\sqrt{2})K_0\sqrt{K_1}C_1\}|x| + \frac{2K_1K^{-1}\|\sqrt{C}\|C_2}{\sqrt{1-t^2}} + \max\{C_3, C_4\}.
\end{aligned}
$$

The next step is to calculate the absolute value of the second term of (20), i.e.

$$
\begin{aligned}
|A_2| &= \left|\frac{-C\partial_t(\bar{A}_t^{-1})xt + (I - C\bar{A}_t^{-1})x}{t^2}\right|\\
&\le \|2C(A-C)((At^2 + C(1-t^2))^2)^{-1}\||x| + K_2|x|\\
&\le (K_4 + K_2)|x|.
\end{aligned}
$$

where $K_4 := \sup_{0\le t\le 1} \|\frac{C\partial_t(\bar{A}_t^{-1})}{t}\| = \sup_{0\le t\le 1} \|2C(A-C)((At^2 + C(1-t^2))^2)^{-1}\|$.

It then follows from (20) that

$$|\partial_t V(t,x)| \le |A_1| + |A_2| \le K_5|x| + \frac{K_6}{\sqrt{1-t^2}} + K_7.$$

where $K_5 = \max\{3K_1C_1, 2(1+\sqrt{2})K_0\sqrt{K_1}C_1\} + K_2 + K_4$, $K_6 = 2K_1K^{-1}\|\sqrt{C}\|C_2$, $K_7 = \max\{C_3, C_4\} + K_0K_3K^{-1}$.

When $t = 0$, by (12) in the subsequently stated well-posedness Lemma 3.10, we have

$$|V(0,x)| = \left|\sqrt{C}\mathbb{E}_{\bar{p}_0}[X]\right| \lesssim K\|C\|^{1/2}\left|\sqrt{C}\nabla h\right|_\infty = K_0.$$

Similarly, $\|\nabla V(0,\cdot)\|_\infty$ is also bounded by $K_1 + K_2$. Then we conclude that the velocity field $V(t,x)$ satisfied the condition of Theorem 3.8 for all $t \in [0,1]$.

For clarity, the coefficients are summarized in Table 3. In corollary 3.18, we take $C = I_d$, $A =$

Table 3: Explicit for coefficients in Thm. 3.8 and Thm. 3.15.

Coefficient	**Explicit expressions**
$\bar{A}_t$	$At^2 + C(1-t^2)$
$K(t)$	$(A\bar{A}_t^{-1})t$
$B(t)$	$(A\bar{A}_t^{-1})(1-t^2)$
K	$\sup_{0\le t\le 1} \vert A\bar{A}_t^{-1}\vert \le \max\{1, \Vert AC^{-1}\Vert\}$
K_0	$K\Vert C\Vert^{1/2}\vert\sqrt{C}\nabla h\vert_\infty$
K_1	$K^2(\Vert C\nabla^2 h\Vert_\infty + \vert\sqrt{C}\nabla h\vert_\infty^2)$
K_2	$\sup_{0\le t\le 1} \Vert\frac{1}{t^2}(I - C\bar{A}_t^{-1})\Vert = \sup_{0\le t\le 1} \Vert(A-C)(At^2 + C(1-t^2))^{-1}\Vert$
K_3	$\sup_{0\le t\le 1} \Vert\frac{A\partial_t(\bar{A}_t^{-1})}{t}\Vert = \sup_{0\le t\le 1} \Vert 2A(A-C)((At^2 + C(1-t^2))^2)^{-1}\Vert$
K_4	$\sup_{0\le t\le 1} \Vert\frac{C\partial_t(\bar{A}_t^{-1})}{t}\Vert = \sup_{0\le t\le 1} \Vert 2C(A-C)((At^2 + C(1-t^2))^2)^{-1}\Vert$
K_5	$\max\{3K_1C_1, 2(1+\sqrt{2})K_0\sqrt{K_1}C_1\} + K_2 + K_4$
K_6	$2K_1K^{-1}\Vert C\Vert^{\frac{1}{2}}C_2$
K_7	$\max\{C_3, C_4\} + K_0K_3K^{-1}$
K_9	$\frac{1}{4}K_6\pi + K_7$
C_1	$\max_{t\in[\sqrt{1-\frac{1}{2\Vert C\Vert\Vert\nabla^2 h\Vert_\infty}},1)}\{\partial_t K(t)\} = \max\{\frac{\Vert A\Vert(\Vert(A-C)(1-\frac{1}{2\Vert C\Vert\Vert\nabla^2 h\Vert_\infty})-C\Vert)}{(\Vert C\Vert+\Vert(A-C)\Vert(1-\frac{1}{2\Vert C\Vert\Vert\nabla^2 h\Vert_\infty}))^2}, \frac{\Vert A-2C\Vert}{\Vert A\Vert}, \frac{\Vert A\Vert}{8\Vert C\Vert}\}$
C_2	$\max_{t\in[\sqrt{1-\frac{1}{2\Vert C\Vert\Vert\nabla^2 h\Vert_\infty}},1)}\{\partial_t B(t)\} = \max\{2, \frac{2\Vert A\Vert^2\sqrt{1-\frac{1}{2\Vert C\Vert\Vert\nabla^2 h\Vert_\infty}}}{(\Vert C\Vert+\Vert(A-C)\Vert(1-\frac{1}{2\Vert C\Vert\Vert\nabla^2 h\Vert_\infty}))^2}, \frac{9A^2}{8\Vert C\Vert^2}\sqrt{\frac{\Vert C\Vert}{3\Vert A-C\Vert}}\}$
C_3	$2K_1C_2(K_1^{\frac{1}{2}} + K_0)$
C_4	$2\sqrt{3}K_0K_1C_2 + \frac{1}{2}K_0^{\frac{3}{2}}K_1^{\frac{1}{2}}C_2 + \frac{(1+\sqrt{6})^2}{4(1+\sqrt{2})}K_0^2K_1^{\frac{1}{2}}C_2$

$(1-(1-\delta)^2)I_d$, $\bar{A}_t = (1-t^2)I_d$. According to (16), we deduce the explicit for coefficients $K_0^*, K_1^*, K_2^*, K_3^*, K_4^*, K_5^*, K_6^*, K_7^*, K_9^*$ in Table 4. □

B.3 Proof of Lemma 3.10

Proof. First, we prove the velocity field $V(t,x)$ is well-defined at $t = 0$ ((12) in Lemma 3.10), i.e.

$$V(0,x) := \lim_{t\to 0} V(t,x) = \lim_{t\to 0} \frac{x + S(t,x)}{t} = \sqrt{C}\mathbb{E}_{\bar{p}_0}[X].$$

Let $t \to 0$, then it yields

$$\lim_{t\to 0} V(t,x) = \lim_{t\to 0} \partial_t S(t,x) = \lim_{t\to 0} \left\{\frac{C\nabla(\partial_t p_t(x))}{p_t(x)} - \frac{\partial_t p_t(x)}{p_t(x)} S(t,x)\right\}.$$

Table 4: Explicit for coefficients in Cor. 3.18, and Thm. 3.19.

Coefficient	**Explicit expressions**
$\bar{A}_t$	$(1-t^2)I_d$
K_0^*	$\frac{R}{1-(1-\delta)^2}$
K_1^*	$3\left(\frac{R}{1-(1-\delta)^2}\right)^2$
K_2^*	$\sup_{t\le 1-\delta}\Vert\frac{1}{t^2}(I-\bar{A}_t^{-1})\Vert = \frac{1}{1-(1-\delta)^2}$
K_3^*	$\frac{2}{1-(1-\delta)^2}$
K_4^*	$\frac{2}{(1-(1-\delta)^2)^2}$
K_5^*	$\frac{9C_1^* R^2}{(1-(1-\delta)^2)^2} + \frac{2}{(1-(1-\delta)^2)^2} + \frac{1}{1-(1-\delta)^2}$
K_6^*	$\frac{6C_2^* R^2}{(1-(1-\delta)^2)^2}$
K_7^*	$6(\sqrt{3}+1)\left(\frac{R}{1-(1-\delta)^2}\right)^3 C_2^* + \frac{2R}{(1-(1-\delta)^2)^2}$
K_9^*	$\frac{3\pi}{2}\,\frac{C_2^* R^2}{(1-(1-\delta)^2)^2} + 6(\sqrt{3}+1)\,\frac{C_2^* R^3}{(1-(1-\delta)^2)^3} + \frac{2R}{(1-(1-\delta)^2)^2}.$
C_1^*	$\max\{\frac{(1-(1-\delta)^2)\Vert(1-\delta)^2(1-\frac{1}{2\Vert\nabla^2 h\Vert_\infty})-1\Vert}{\Vert 1+(1-\delta)^2(1-\frac{1}{2\Vert\nabla^2 h\Vert_\infty})\Vert^2}, \frac{\Vert -(1-\delta)^2-1\Vert}{\Vert 1-(1-\delta)^2\Vert}, \frac{\Vert 1-(1-\delta)^2\Vert}{8}\},$
C_2^*	$\max\{2, \frac{2(1-(1-\delta)^2)^2(1-\frac{1}{2\Vert\nabla^2 h\Vert_\infty})^{\frac{1}{2}}}{\Vert 1+(1-\delta)^2(1-\frac{1}{2\Vert\nabla^2 h\Vert_\infty})\Vert^2}, \frac{9(1-(1-\delta)^2)^2}{8\sqrt{3}(1-\delta)},$

By simple calculation, it holds that

$$
\begin{aligned}
&\frac{\nabla(\partial_t p_t(x))}{p_t(x)} \\
=&-\partial_t\left(\det B(t)\right)(2\det B(t))^{-1} d\left(-K(t)^T K(t)(CB(t))^{-1}-\bar{A}_t^{-1}\right)(x^T x)x \\
&-\partial_t B(t)(2B(t))^{-1}K(t)(\sqrt{C}B(t))^{-1}\int_{\mathbb{R}^d} y p(1,y|t,x)\mathrm{d}y \\
&+\left(-K(t)\partial_t K(t)(CB(t))^{-1}-K(t)^T K(t)\partial_t(B(t)^{-1})(2CB(t))^{-1}-\frac{1}{2}\partial_t(\bar{A}_t^{-1})\right) \\
&\quad\cdot\Bigg(2x+(-K(t)^T K(t)(CB(t))^{-1}-\bar{A}_t^{-1})(x^T x)x \\
&\qquad+2K(t)(\sqrt{C}B(t))^{-1}\int_{\mathbb{R}^d}(x^T y)xp(1,y|t,x)\mathrm{d}y\Bigg) \\
&+\left(\partial_t K(t)(\sqrt{C}B(t))^{-1}+K(t)\partial_t(B(t)^{-1})(\sqrt{C})^{-1}\right) \\
&\quad\cdot\Bigg(\int_{\mathbb{R}^d}\left(y+K(t)(\sqrt{C}B(t))^{-1}(x^T y)y\right)p(1,y|t,x)\mathrm{d}y \\
&\qquad+\int_{\mathbb{R}^d}\left(-K(t)^T K(t)(CB(t))^{-1}-\bar{A}_t^{-1}\right)(x^T y)xp(1,y|t,x)\mathrm{d}y\Bigg) \\
&-\frac{1}{2}\partial_t(B(t)^{-1})\int_{\mathbb{R}^d}\left(\left(-K(t)^T K(t)(CB(t))^{-1}-\bar{A}_t^{-1}\right)(y^T y)x\right)p(1,y|t,x)\mathrm{d}y \\
&-\frac{1}{2}\partial_t(B(t)^{-1})\int_{\mathbb{R}^d}K(t)(\sqrt{C}B(t))^{-1}(y^T y)yp(1,y|t,x)\mathrm{d}y,
\end{aligned}
$$

while

$$\begin{aligned}\frac{\partial_t p_t(x)}{p_t(x)} =& -\partial_t\left(\det B(t)\right)(2\det B(t))^{-1}d \\ &+ x^T\left(-K(t)\partial_t K(t)(CB(t))^{-1} - (2C)^{-1}K(t)K(t)^T\partial_t(B(t)^{-1}) - \frac{1}{2}\partial_t(\bar{A}_t^{-1})\right)x \\ &+ \int_{\mathbb{R}^d} x^T\left(\partial_t K(t)(\sqrt{C}B(t))^{-1} + K(t)(\sqrt{C})^{-1}\partial_t(B(t)^{-1})\right)yp(1,y|t,x)\,\mathrm{d}y \\ &- \int_{\mathbb{R}^d} y^T\frac{1}{2}\partial_t(B(t)^{-1})yp(1,y|t,x)\,\mathrm{d}y.\end{aligned}$$

From observing that

$$\lim_{t\to 0}\partial_t\left(\det B(t)\right) = 0,\ \lim_{t\to 0}\partial_t B(t) = 0,\ \lim_{t\to 0}K(t) = 0,\ \lim_{t\to 0}\partial_t(\bar{A}_t^{-1}) = 0,$$

and Assumption 3.6, which ensures

$$\lim_{t\to 0}\int_{\mathbb{R}^d}|y|^3 p(1,y|t,x)\mathrm{d}y = \int_{\mathbb{R}^d}|y|^3\lim_{t\to 0}p(1,y|t,x)\mathrm{d}y = \mathbb{E}_{\bar{p}_0}[|X|^3] < +\infty,$$

we have

$$\lim_{t\to 0}\frac{\partial_t p_t(x)}{p_t(x)}S(t,x) = -\frac{x^\top x}{\sqrt{C}}\mathbb{E}_{\bar{p}_0}[X],\quad \lim_{t\to 0}\frac{C\nabla(\partial_t p_t(x))}{p_t(x)} = \sqrt{C}\mathbb{E}_{\bar{p}_0}[X] - \frac{x^\top x}{\sqrt{C}}\mathbb{E}_{\bar{p}_0}[X].$$

Therefore, it yields $\lim_{t\to 0}V(t,x) = \sqrt{C}\mathbb{E}_{\bar{p}_0}[X]$, which completes the proof of (12).

Next, together with the regularity of the velocity field (Theorem 3.8), the Arzelà–Ascoli theorem (Arzela, 1895) ensures the existence of a subsequence $\{V(t,x)_{n_k}\}_{k\in N}$ that converges locally uniformly to $V(0,x)$, thereby guaranteeing the well-posedness of the ODE (8) on the entire time interval $t\in[0,1]$. □

B.4 Proof of Corollary 3.11

Proof. Recall the Föllmer flow (8) with $\|\nabla V(t,\cdot)\|_\infty \le (K_1+K_2)t$ in Theorem 3.8, by following the Proposition A.7 (Mikulincer & Shenfeld, 2023), we arrive at the following result,

$$\mathrm{Lip}(\overleftarrow{X}_1(x)) \le \|\nabla\overleftarrow{X}_1(x)\|_{op} \le \exp\left(\int_0^1 (K_1+K_2)sds\right) \le \exp\left(\frac{K_1+K_2}{2}\right).$$

□

B.5 Proof of Corollary 3.14

Proof. For any $x, y\in\mathbb{R}^d$, $t\in[t_n, t_{n+1}]$ with $k = 0,1,\dots,N-1$, Itô's formula gives

$$\frac{\mathrm{d}|\overleftarrow{Y}_t(x) - \overleftarrow{Y}_t(y)|^2}{dt} = 2\langle\overleftarrow{Y}_t(x) - \overleftarrow{Y}_t(y), \widetilde{V}\big(t_n,\overleftarrow{Y}_{t_n}(x)\big) - \widetilde{V}\big(t_n,\overleftarrow{Y}_{t_n}(y)\big)\rangle.$$

Using the Cauchy-Schwarz inequality, we obtain

$$\frac{\mathrm{d}|\overleftarrow{Y}_t(x) - \overleftarrow{Y}_t(y)|^2}{dt} \le 2\sqrt{|\overleftarrow{Y}_t(x) - \overleftarrow{Y}_t(y)|^2}\sqrt{|\widetilde{V}\big(t_n,\overleftarrow{Y}_{t_n}(x)\big) - \widetilde{V}\big(t_n,\overleftarrow{Y}_{t_n}(y)\big)|^2}.$$

Therefore,

$$\begin{aligned}\frac{d|\overleftarrow{Y}_t(x) - \overleftarrow{Y}_t(y)|}{dt} &\le |\widetilde{V}\big(t_n,\overleftarrow{Y}_{t_n}(x)\big) - \widetilde{V}\big(t_n,\overleftarrow{Y}_{t_n}(y)\big)| \\ &\le \nabla\widetilde{V}|\overleftarrow{Y}_{t_n}(x) - \overleftarrow{Y}_{t_n}(y)| \\ &\le (K_1+K_2+K_8)t_n|\overleftarrow{Y}_{t_n}(x) - \overleftarrow{Y}_{t_n}(y)|,\end{aligned}$$

where the last inequality uses Lipschitzness of $\widetilde{V}$ in Assumption 3.13. Integration over time yields

$$
\begin{aligned}
|\overleftarrow{Y}_{t_{n+1}}(x) - \overleftarrow{Y}_{t_{n+1}}(y)| &\leq |\overleftarrow{Y}_{t_n}(x) - \overleftarrow{Y}_{t_n}(y)| + (t_{n+1} - t_n)(K_1 + K_2 + K_8)t_n|\overleftarrow{Y}_{t_n}(x) - \overleftarrow{Y}_{t_n}(y)| \\
&= \mathrm{Lip}(\widetilde{T}_n)|\overleftarrow{Y}_{t_n}(x) - \overleftarrow{Y}_{t_n}(y)|,
\end{aligned}
$$

where $\mathrm{Lip}(\widetilde{T}_n) = 1+(t_{n+1}-t_n)(K_1+K_2+K_8)t_n$. Iterating this bound over all $n = 0, 1, \ldots, N-1$, we obtain the following estimate over the full interval $[0, 1]$,

$$
\begin{aligned}
|\overleftarrow{Y}_1(x) - \overleftarrow{Y}_1(y)| &\leq \left(\prod_{n=0}^{N-1} \mathrm{Lip}(\overleftarrow{Y}_n)\right) |\overleftarrow{Y}_{t_0}(x) - \overleftarrow{Y}_{t_0}(y)| \\
&\leq \exp\left(\frac{K_1 + K_2 + K_8}{2}\right) |x - y|.
\end{aligned}
$$

Then we complete the proof. □

B.6 Proof of Theorem 3.15

Proof. Recall that Assumption 3.6, 3.7 ensure the well-podeness and Lipschitzness of Föllmer flow $(\overleftarrow{X}_t)_{t\in[0,1]}$ in (8), with

$$
\mathrm{Lip}(T_n) = \exp\left(\int_{t_n}^{t_{n+1}} (K_1 + K_2)tdt\right) \quad \text{and} \quad \prod_{j=0}^{N-1} \mathrm{Lip}(T_j) = \exp\left(\frac{K_1 + K_2}{2}\right),
$$

as established in Lemma 3.10 and Corollary 3.11.

Furthermore, Assumption 3.13 guarantees the Lipschitzness of learned discret Föllmer flow $(\overleftarrow{Y}_t)_{t\in[0,1]}$ in (9), with

$$
\mathrm{Lip}(\widetilde{T}_n) = 1 + (t_{n+1} - t_n)(K_1 + K_2 + K_8)t_n \quad \text{and} \quad \prod_{j=0}^{N-1} \mathrm{Lip}(\widetilde{T}_j) = \exp\left(\frac{K_1 + K_2 + K_8}{2}\right),
$$

as shown in Corollary 3.14.

Therefore, it only remains to verify that the stepwise approximation error satisfies **Assumption 3.4**. To analyze the discretization error at each step, we recall the expression in (3):

$$
T_n(\overleftarrow{X}_{t_n}) = \overleftarrow{X}_{t_{n+1}}.
$$

Applying the vector-valued Taylor expansion of $\overleftarrow{X}_{t_{n+1}}$ over $[t_n, t_{n+1}]$, the remainder is defined by

$$
R(t) := T_n(\overleftarrow{X}_{t_n}) - \overleftarrow{X}_{t_n} - hV(t_n, \overleftarrow{X}_{t_n}).
$$

Under Assumption 3.6, we can derive the second moment bound by the forward diffusion process (6)

$$
\begin{aligned}
\mathbb{E}_{\bar{p}_t}|\overrightarrow{X}_t|^2 &= \mathbb{E}_{\bar{p}_t}|\overrightarrow{X}_t - (1-t)\overrightarrow{X}_0|^2 + \mathbb{E}_{\bar{p}_t}|(1-t)\overrightarrow{X}_0|^2 \\
&\leq t(2-t)Tr(C) + (1-t)^2 M_2 \\
&\leq M_0.
\end{aligned}
$$

Then the expectation of the $R(t)$ is controlled by

$$
\begin{aligned}
\mathbb{E}|R(t)|^2 &= \mathbb{E}_{\overleftarrow{X}_{t_n} \sim \overleftarrow{P}_{t_n}} |T_n(\overleftarrow{X}_{t_n}) - \overleftarrow{X}_{t_n} - hV(t_n, \overleftarrow{X}_{t_n})|^2 \\
&\leq \frac{h^4}{4} \sup_{\tau\in(t_n,t_{n+1})} \mathbb{E}_{\overleftarrow{X}_\tau \sim \overleftarrow{P}_\tau} |\partial_\tau V(\tau, \overleftarrow{X}_\tau)|^2 \\
&\leq \frac{3h^4}{4}\left(K_5^2 M_0 + \frac{K_6^2}{1-t^2} + K_7^2\right), \quad \forall t \in [0, 1),
\end{aligned}
$$

where the last inequality follows from the bound on $|\partial_t V(t,x)|$ in Theorem 3.8, which gives

$$\mathbb{E}_{\overleftarrow{X}_\tau \sim \overleftarrow{P}_\tau} |\partial_\tau V(\tau, \overleftarrow{X}_\tau))|^2 \le 3\left(K_5^2 M_0 + \frac{K_6^2}{1-\tau^2} + K_7^2\right).$$

Consequently, the local truncation error is bounded by

$$\begin{aligned}
&\sqrt{\mathbb{E}_{\overleftarrow{X}_{t_n} \sim \overleftarrow{P}_{t_n}} |T_n(\overleftarrow{X}_{t_n}) - \widetilde{T}_n(\overleftarrow{X}_{t_n})|^2} \\
=&\sqrt{\mathbb{E}_{\overleftarrow{X}_{t_n} \sim \overleftarrow{P}_{t_n}} |\overleftarrow{X}_{t_n} + hV(t_n, \overleftarrow{X}_{t_n}) + R(t_n) - \overleftarrow{X}_{t_n} - h\widetilde{V}(t_n, \overleftarrow{X}_{t_n})|^2} \\
\le&\sqrt{h^2 \mathbb{E}_{\overleftarrow{X}_{t_n} \sim \overleftarrow{P}_{t_n}} |V(t_n, \overleftarrow{X}_{t_n}) - \widetilde{V}(t_n, \overleftarrow{X}_{t_n})|^2} + \sqrt{\mathbb{E}|R(t_n)|^2} \\
\le&h\left(\frac{\sqrt{3}}{2}\Big(K_5\sqrt{M_0} + \frac{K_6}{\sqrt{1-t_n^2}}h + K_7\Big)h + \epsilon\right).
\end{aligned}$$

The second inequality holds by the error between $\widetilde{V}(t_n,x)$ and $V(t_n,x)$ stated in Assumption 3.12. This completes the verification of Assumption 3.4.

Now in Theorem 3.5, we employ coupling between $\overleftarrow{X}_0 \sim \overrightarrow{P}_1 = \gamma_C$ and $\overleftarrow{Y}_0 \sim \overleftarrow{Q}_0 = \gamma_C$,

$$\begin{aligned}
\mathcal{W}_2(\overleftarrow{P}_1, \overleftarrow{Q}_1) \le& \exp\left(\frac{K_1+K_2+K_8}{2}\right)\mathcal{W}_2(\overleftarrow{P}_0, \overleftarrow{Q}_0) \\
&+ \frac{\exp\left(\frac{K_1+K_2+K_8}{2}\right)-1}{((K_1+K_2+K_8)h)/2}\cdot h\left(\frac{\sqrt{3}}{2}\left(K_5\sqrt{M_0}+K_7\right)h+\epsilon\right) \\
&+ \exp\left(\frac{K_1+K_2+K_8}{2}\right)\frac{\sqrt{3}K_6\pi}{4}h \\
\le& \exp\big(\frac{K_1+K_2+K_8}{2}\big)\left(\sqrt{3}\left(K_5\sqrt{M_0}+K_9\right)h+2\epsilon\right).
\end{aligned}$$

where a straightforward calculation shows $\sum_{n=0}^{N-1}\frac{K_6}{\sqrt{1-t_n^2}} = \frac{K_6\pi}{2}$; Accordingly, set $K_9 := \frac{K_6\pi}{4} + K_7$. Noting that $\overleftarrow{P}_1 = \overrightarrow{P}_0$, we obtain the conclusion in (14). □

C Convergence in the Bayesian Inverse problems

We are aware of several posterior analyses, such as Bayesian inverse problems (van de Schoot et al., 2021), used in uncertainty quantification to infer model parameters x from observations $y \in \mathbb{R}^m$. The posterior typically takes the form of

$$\bar{p}_0(x) = D_0 \exp\left(-\frac{|x|_C^2}{2}\right)\exp\left(-\frac{|G(x)-y|_\Sigma^2}{2}\right), \tag{25}$$

where D_0 is a normalizing constant, C denotes the covariance matrix of the Gaussian prior, Σ represents the covariance of the observational noise and $G \in C_2^b(\mathbb{R}^d, \mathbb{R}^m)$ is a nonlinear forward operator. In our training framework, we adopt the Gaussian prior with covariance C from (25) as the invariant measure of the forward diffusion process (6). The conditioned score (Batzolis et al., 2021) in the score matching is trained by minimizing

$$\mathbb{E}_{\bar{p}_t(x;y)}\left|\widetilde{s}(1-t,x;y) - C\nabla_x \log \bar{p}_t(x;y)\right|^2,$$

where $\bar{p}_t$ denotes the joint law of (X_t, Y) with $Y = G(X_0) + \mathcal{N}(0,\Sigma)$. For ODE-based generation of the posterior distribution with observation y, we impose the following assumption on the approximation error of the velocity field $V(t,x;y)$ given in (8).

Assumption C.1. *Fixing observation y, for each time discretization point t_n,*

$$\mathbb{E}_{\overrightarrow{P}_{1-t_n};y}|V(t_n,x;y)-\widetilde{V}(t_n,x;y)|^2 \le \epsilon^2.$$

Theorem C.2. *Suppose third moment Assumption 3.6, accuracy Assumption C.1 and regularity Assumption 3.13 hold. Using the Euler method to the Föllmer flow with uniform step size $h = t_{n+1}-t_n \le 1$ yields,*

$$\mathcal{W}_2(\overrightarrow{P}_0(\cdot,y),\overleftarrow{Q}_1(\cdot,y)) \le \exp\big(\frac{\widetilde{K_1}+K_8}{2}\big)\left(\frac{\sqrt{3}}{N}\left(\widetilde{K_5}\sqrt{M_0}+\widetilde{K_9}\right)+2\epsilon\right). \tag{26}$$

where $\widetilde{K_1},\widetilde{K_5},\widetilde{K_9}$ are dimension-free constants depending on $(\|C\|,\|\Sigma\|,G,y)$, see Table 5, and the constant K_8 is defined in Assumption 3.13.

Proof. Take $A = C$, and $h(x) = -\frac{|G(x)-y|^2_\Sigma}{2}$, then $h(x)$ satisfies,

$$\begin{aligned}
|\sqrt{C}\nabla h(x)| &= |\sqrt{C}\nabla G(x)\Sigma^{-1}(G(x)-y)| \le \|C\|^{\frac{1}{2}}\left(|G|_\infty+|y|\right)\|\Sigma^{-1}\|\|\nabla G\|_\infty, \\
\|C\nabla^2 h(x)\| &= \|C\nabla^2 G(x)\Sigma^{-1}(G(x)-y)+C\nabla G(x)\Sigma^{-1}\nabla G(x)^T\| \\
&\le \|C\|\|\Sigma\|^{-1}\|\nabla^2 G\|_\infty(|G|_\infty+|y|\|\nabla G\|^2_\infty).
\end{aligned}$$

Then by Theorem 3.15, we obtain the bound (26) with the constants replaced as specified in Table 5. $\square$

Table 5: Explicit for coefficients in Thm. C.2.

Coefficient	**Explicit expressions**
$\widetilde{K_0}$	$\|C\|\|\Sigma^{-1}\|\,\|\nabla G\|_\infty\left(\|G\|_\infty+\|y\|\right)$
$\widetilde{K_1}$	$\|C\|\Big(\|\Sigma^{-1}\|\big(\|\nabla G\|^2_\infty+(\|G\|_\infty+\|y\|)\|\nabla^2 G\|_\infty\big)+\|\Sigma^{-2}\|\,\|\nabla G\|^2_\infty\left(\|G\|_\infty+\|y\|\right)^2\Big)$
$\widetilde{K_5}$	$\max\{3\widetilde{K_1},2(1+\sqrt{2})\widetilde{K_0}\widetilde{K_1}^{\frac{1}{2}}\}$
$\widetilde{K_6}$	$4\|C\|^{\frac{1}{2}}\widetilde{K_1}$
$\widetilde{K_7}$	$\max\{\widetilde{C_3},\widetilde{C_4}\}$
$\widetilde{K_9}$	$\|C\|^{\frac{1}{2}}\widetilde{K_1}\pi+\widetilde{K_7}$
$\widetilde{C_3}$	$4\widetilde{K_1}(\widetilde{K_1}^{\frac{1}{2}}+\widetilde{K_0})$
$\widetilde{C_4}$	$4\sqrt{3}\widetilde{K_0}\widetilde{K_1}+\widetilde{K_0}^{\frac{3}{2}}\widetilde{K_1}^{\frac{1}{2}}+\frac{(1+\sqrt{6})^2}{2(1+\sqrt{2})}\widetilde{K_0}^2\widetilde{K_1}^{\frac{1}{2}}$

Remark C.3. *With fixed $\widetilde{K}_1$, $\widetilde{K}_5$ and $\widetilde{K}_9$, for ϵ_0 accuracy in W_2 distance for 26, one requires,*

$$N = \mathcal{O}\left(\frac{\sqrt{M_0}}{\epsilon_0}\right), \quad \epsilon = \mathcal{O}(\epsilon_0).$$