

SAMPLING AND OPTIMIZATION MEET ENHANCED FLOWS

YUAN GAO, SIMING HE, AND EITAN TADMOR

ABSTRACT. It is well known that the computational realization of Gibbs probability measures, $e^{-\mathbb{U}(\mathbf{x})}/Z$, plays a central role in sampling and optimization. In this paper, we introduce two types of dynamics that exhibit rapid convergence towards these Gibbs measures. The mechanism driving this rapid convergence is the enhanced dissipation associated with these transport-diffusion dynamics. Motivated by these enhanced dynamics, we design numerical algorithms for sampling from the target Gibbs measure. Finally, we provide the corresponding particle systems that may yield other effective numerical samplers.

CONTENTS

1. Introduction	1
2. The First-Order Enhanced Dynamics	10
3. The Second-Order Enhanced Dynamics	15
4. Numerical Simulations	22
Appendix A. Connections between the Models	24
Appendix B. Space-time Discretization and Unconditional Stability	25
Appendix C. The Particle System of the Sampling Dynamics	26
Appendix D. Enhanced Dissipation for Alternating Shear Flows	27
Appendix E. Mass-Searching Dynamics	34
References	34

1. INTRODUCTION

A classical problem in sampling theory and optimization is to design dynamical systems whose solutions converge rapidly towards a target

$$(1.1) \quad 0 \leq \Pi(\mathbf{x}) = \frac{e^{-\mathbb{U}(\mathbf{x})}}{Z}, \quad \mathbf{x} \in \mathbb{T}^d = (\mathbb{R}/2\pi\mathbb{Z})^d.$$

The normalization constant $Z := \int_{\mathbb{T}^d} e^{-\mathbb{U}(\mathbf{x})} d\mathbf{x}$ ensures that Π is a probability density. The efficient numerical realization of general densities Π is fundamental in two distinct but intertwined areas (see, e.g., [63]). In *sampling theory*, specifically in Bayesian inference, sampling from a fixed probability distribution Π is crucial for approximating statistically relevant quantities—such as expectations and variances—of high-dimensional observables; see, e.g., [53, 63]. On the other hand, in the context of *optimization*, approximating the target $\Pi = e^{-\mathbb{U}}/Z$ helps characterize the global optimum of the associated target function $\mathbb{U}$.

Although many methods have been developed, significant challenges remain in the vast literature on this problem. The first numerical challenge arises when one attempts to compute the value of the normalization constant Z —particularly in the high-dimensional setting, $d \gg 1$. In spite of

Date: August 7, 2026.

2020 Mathematics Subject Classification. Primary 35B40, 35Q84, 65C05; Secondary 35Q35, 37A25, 65C05.

Key words and phrases. Langevin dynamics, enhanced dissipation, mixing, sampling.

Acknowledgment. The research of YG is supported by NSF CAREER Award DMS-2440651. SH is supported by NSF grant DMS-2406293 and would like to thank Wuchen Li for helpful discussions. The research of ET was supported by ONR grant N00014-2412659 and NSF grant DMS-2508407.

the straightforward theoretical representation of the constant Z , numerical implementation of the corresponding high-dimensional quadrature is challenging; see, e.g., [35, 48, 60]. We mention direct estimators based on lattice methods [59], bridge and path sampling [28] and annealed importance sampling [46]. Another widely applied technique, which circumvents direct computation of the normalization constant Z , is the design of *Langevin sampling dynamics* whose long-time behavior converges toward the target density Π ; see, e.g., [50]. We recall the gradient flow associated with *overdamped Langevin sampling dynamics*, which is given by [3, 8, 29, 49, 50, 54],

$$(1.2) \quad \partial_t \rho = \nu \Delta_{\mathbf{x}} \rho + \nu \nabla_{\mathbf{x}} \cdot (\rho \nabla_{\mathbf{x}} \mathbb{U}), \quad \rho(t=0, \mathbf{x}) = \rho_0(\mathbf{x}), \quad \mathbf{x} \in \mathbb{T}^d.$$

Here, the solution $\rho \geq 0$ is a probability density, and the parameter $\nu^{-1} \geq 1$ encodes the intrinsic time scale of the Langevin dynamics. One readily checks that $Z^{-1}e^{-\mathbb{U}}$ is a stationary solution of (1.2), and hence the target measure Π can be approximated by the solution of (1.2) as $t \rightarrow \infty$. However, the convergence $\rho(t) \xrightarrow{t \rightarrow \infty} \Pi$ can be exceedingly slow; see, e.g., [12]. In this context, one may consider the following scenario, in which the function $\mathbb{U}$ has multiple “valleys” surrounding local minima, and the initial density is concentrated in one of them. Then, it can take the Langevin dynamics (1.2) an extremely long time to escape these “valleys” around local minima before successfully exploring other parts of the landscape of $\mathbb{U}$. However, one expects that the judicious introduction of an external drift can drive the density away from these locally trapped states and enable it to explore the landscape effectively. In conclusion, we aim to design new dynamics to address the following question:

Question: *By incorporating external drift, is it possible to design dynamics whose solution converges to a given smooth density $\Pi(\mathbf{x}) = e^{-\mathbb{U}(\mathbf{x})}/Z$ within a short time scale?*

Initial progress toward this goal was made in [12], where the authors considered the following modified Langevin dynamics (in PDE form):

$$(1.3) \quad \partial_t \rho + \nabla_{\mathbf{x}} \cdot (\mathbb{V} \rho) = \nu \Delta_{\mathbf{x}} \rho + \nu \nabla_{\mathbf{x}} \cdot (\rho \nabla_{\mathbf{x}} \mathbb{U}).$$

Here the external time-dependent drift $\mathbb{V}$ is chosen such that the *drift defect*,

$$(1.4) \quad \mathbb{D} := -e^{\mathbb{U}} \nabla_{\mathbf{x}} \cdot (e^{-\mathbb{U}} \mathbb{V}),$$

is identically zero, i.e., $\mathbb{D} \equiv 0$. The vanishing of the drift defect ensures that the dynamics preserves the invariant measure $\Pi = e^{-\mathbb{U}}/Z$. This can be verified by rewriting the dynamics in terms of the ratio $h = \rho/\Pi$ and observing that $h \equiv 1$ is a stationary solution. In [12], the authors identify delicate, rapidly alternating random drifts $\mathbb{V}$ that exhibit strong mixing properties [5, 9] while preserving a vanishing drift defect $\mathbb{D}$. As a consequence, rapid convergence of the dynamics towards Π is established theoretically. However, due to the randomness and rapid alternation involved in the construction, it is numerically challenging to simulate the dynamics for long times. Finally, the choice of $\mathbb{V}$ is constrained by the requirement that $\mathbb{D} \equiv 0$, which also complicates the numerical implementation.

Motivated by the aforementioned goal and the early development of [12], we propose two types of enhanced dynamics that exhibit rapid convergence towards the target distribution Π while reducing the complexity of the drift $\mathbb{V}$. We start by introducing a *first-order model* incorporating alternating shear flows, then we present a *second-order kinetic model* with a fixed drift. Both models are supplemented with their *simplified versions* and the quadrature-free *mass-searching dynamics* for numerical implementation.

1.1. The First-Order Model. Given the target distribution $\Pi = Z^{-1}e^{-\mathbb{U}}$ and a *divergence-free* ambient drift $\mathbb{V}$, we recall the drift defect $\mathbb{D}$ (1.4) associated with $\mathbb{U}$ and $\mathbb{V}$,

$$\mathbb{D} = -e^{\mathbb{U}} \nabla_{\mathbf{x}} \cdot (e^{-\mathbb{U}} \mathbb{V}) = \mathbb{V} \cdot \nabla_{\mathbf{x}} \mathbb{U},$$

and introduce the first-order *mass-preserving enhanced sampling* model,

$$(1.5) \quad \overbrace{\partial_t \rho + \nabla_{\mathbf{x}} \cdot (\mathbb{V} \rho)}^{\text{Transport}} = \overbrace{\nu \Delta_{\mathbf{x}} \rho + \nu \nabla_{\mathbf{x}} \cdot (\rho \nabla_{\mathbf{x}} \mathbb{U})}^{\text{Gradient Flow}} + \overbrace{\rho \int (\mathbb{D}(\mathbf{y}) - \mathbb{D}(\mathbf{x})) \rho(\mathbf{y}) d\mathbf{y}}^{\mathbb{D}\text{-alignment } \mathcal{Q}_{\mathbb{U}, \mathbb{V}}[\rho]},$$

$$\nabla_{\mathbf{x}} \cdot \mathbb{V} = 0, \quad \rho(t = 0, \mathbf{x}) = \rho_0(\mathbf{x}), \quad (t, \mathbf{x}) \in \mathbb{R}_+ \times \mathbb{T}^d.$$

The model describes the time evolution of the probability density $\rho \geq 0$. Here, the viscosity parameter $\nu > 0$ characterizes the strength of the classical Langevin dynamics (“gradient flow”). We focus on the parameter regime $0 < \nu \ll 1$, emphasizing the dominant roles of transport and alignment. Next, we observe that if the drift $\mathbb{V}$ is absent, the model (1.5) reduces to the classical equation (1.2). Thus, a suitably chosen divergence-free vector field $\mathbb{V}$ can be viewed as an external control to enhance the convergence of the Langevin dynamics. The key question to address is the *convergence rate*: below we measure the *relative error*, $\left\| \frac{\rho(t, \cdot)}{\Pi(\cdot)} - 1 \right\|_{L^2}$, proving that our enhanced model with properly chosen control $\mathbb{V}$ leads to a rapid exponential convergence in time. Finally, in order to preserve the invariant measure Π of the dynamics, the transport term $\nabla_{\mathbf{x}} \cdot (\mathbb{V} \rho)$ on the left-hand side of (1.5) is coupled on the right with the *defect alignment* term ($\mathbb{D}$ -alignment), which can be reformulated as follows

$$\mathcal{Q}_{\mathbb{U}, \mathbb{V}}[\rho](t, \mathbf{x}) = -\rho(t, \mathbf{x}) \mathbb{V}(t, \mathbf{x}) \cdot \nabla_{\mathbf{x}} \mathbb{U}(\mathbf{x}) + \rho(t, \mathbf{x}) \int \rho(t, \mathbf{y}) \mathbb{V}(t, \mathbf{y}) \cdot \nabla_{\mathbf{y}} \mathbb{U}(\mathbf{y}) d\mathbf{y}.$$

Furthermore, by replacing the divergence-free condition $\nabla_{\mathbf{x}} \cdot \mathbb{V} \equiv 0$ by the vanishing-defect constraint $\mathbb{D} \equiv 0$, one recovers the model in [12]. The commutator structure in the $\mathbb{D}$ -alignment $\mathcal{Q}_{\mathbb{U}, \mathbb{V}}[\rho]$ appears in various alignment/flocking dynamics [30, 31, 37, 45, 58]. The role of such an antisymmetric alignment mechanism is to drive the density ρ towards an “environmental average”, which in our case will be the desired distribution Π . Moreover, the antisymmetric form of the alignment term guarantees mass conservation, $\|\rho(\cdot, t)\|_{L^1_{\mathbf{x}}} = \|\rho_0\|_{L^1_{\mathbf{x}}} = 1$.

Our first main theorem illustrates that it is possible to design the dynamics (1.5) such that its solutions $\rho(t, \cdot)$ exhibit fast convergence to the target distribution Π .

Theorem 1.1. *Consider the enhanced dynamics (1.5) associated with the target probability density $0 < \Pi \in C^2(\mathbb{T}^d)$ with $2 \leq d \in 2\mathbb{N}$. Assume that the initial data is smooth and normalized: $0 \leq \rho_0 \in C^\infty(\mathbb{T}^d)$ and $\|\rho_0\|_{L^1} = 1$. There exist time-dependent velocity fields $\mathbb{V} \in C_{t, \mathbf{x}}^\infty$, constants $C_\dagger = C_\dagger(\rho_0, Z, \mathbb{U})$, $\delta = \mathcal{O}(1) > 0$ independent of d and ν , and a threshold $\nu_0(\rho_0, \mathbb{U})$ such that the following estimate holds for all $\nu \in (0, \nu_0]$:*

$$(1.6) \quad \left\| \frac{\rho(t)}{\Pi} - 1 \right\|_{L^2} \leq C_\dagger e^{-\delta \nu^{1/2} t}, \quad \forall t \geq 0.$$

Moreover, the velocity fields $\mathbb{V}$ have concrete expressions.

For the classical Langevin dynamics (1.2), the expected convergence time scale is $\mathcal{O}(\nu^{-1})$. By contrast, the convergence time scale for the model (1.5) is $\mathcal{O}(\nu^{-\frac{1}{2}})$, which is significantly shorter than the classical time scale in the transport-dominated regime $\nu \ll 1$. This is the so-called *enhanced-dissipation/relaxation-enhancement phenomenon* in the literature [6, 14], and we will refer to $\nu^{\frac{1}{2}}$ as the *enhanced sampling rate*. In the work [12], by incorporating fast-alternating randomized drift, the authors can achieve an even faster theoretical enhanced sampling rate $|\log \nu|^{-1}$. However, our model is simpler and more flexible in the sense that a large family of time-dependent vector fields $\mathbb{V}$ ensures the fast convergence stated in (1.6). In this paper, we focus on the time-periodic alternating sine-shear flows of period $\nu^{-1/2}$, see (1.18). Our argument extends to other vector fields $\mathbb{V}$ exhibiting the enhanced dissipation outlined in (1.22), e.g., the randomized Pierrehumbert flows analyzed in [9, 12] serve as natural alternatives.

Motivated by Theorem 1.1, one seeks effective discretizations of the dynamics (1.5) which would lead to fast sampling/optimization algorithms. One such discretization, based on a particle system, is proposed in Appendix C. Another approach, which will be our main focus, is to simulate the PDE dynamics directly. We emphasize that even though the $\mathbb{D}$ -alignment term on the right of (1.5) involves integration $\int \rho \mathbb{V} \cdot \nabla \mathbb{U} d\mathbf{y}$, *no quadrature is needed* to access the solution $\rho(t)$ of the PDE. Indeed, it can be verified, see Lemma A.1, that $\rho(t, \cdot) = \frac{\hat{\rho}(t, \cdot)}{\|\hat{\rho}(t, \cdot)\|_{L^1}}$, where $\hat{\rho}(t, \cdot)$ satisfies the following *simplified first-order dynamics*,

$$(1.7) \quad \begin{cases} \partial_t \hat{\rho} + \mathbb{V} \cdot \nabla_{\mathbf{x}} \hat{\rho} = \nu \Delta_{\mathbf{x}} \hat{\rho} + \nu \nabla_{\mathbf{x}} \cdot (\hat{\rho} \nabla_{\mathbf{x}} \mathbb{U}) - \hat{\rho} \mathbb{V} \cdot \nabla_{\mathbf{x}} \mathbb{U}, \\ \nabla_{\mathbf{x}} \cdot \mathbb{V}(t, \mathbf{x}) = 0, \quad \hat{\rho}(t=0, \mathbf{x}) = \rho_0(\mathbf{x}), \quad (t, \mathbf{x}) \in \mathbb{R}_+ \times \mathbb{T}^d. \end{cases}$$

Furthermore, a separate quadrature-free *mass-searching dynamics* will be introduced to compute the total mass $\|\hat{\rho}\|_{L^1}$ at any fixed time t (Section 1.4). Consequently, the simulation of the dynamics (1.5) can be decomposed into two steps:

- (1) Simulate the simplified dynamics (1.7) to obtain $\hat{\rho}(t)$.
- (2) Use the “mass-searching dynamics” to determine its mass $\|\hat{\rho}\|_{L^1}$ at a fixed time t .

By forming the quotient $\rho(t, \cdot) = \frac{\hat{\rho}(t, \cdot)}{\|\hat{\rho}(t, \cdot)\|_{L^1}}$, we end up with a quadrature-free scheme to simulate ρ , which in turn yields a rapidly convergent algorithm for sampling Π .

1.2. The Second-Order Model. One potential drawback of the first-order model is that the external drift controls $\mathbb{V}$ are chosen manually in our context and depend on time. Numerical realizations of such vector fields can be challenging, see, e.g., [12]. We next develop an alternative second-order sampling/optimization dynamics with a simpler, static external drift control. To set up the dynamics to sample $\Pi = Z^{-1}e^{-\mathbb{U}}$, we consider the kinetic state space $(\mathbf{x}, \mathbf{p})$ with position $\mathbf{x}$ and momentum $\mathbf{p}$. The assigned drift $\mathbf{v}$ is determined by the momentum, and the drift-defect $\mathbb{D}$ (1.4) is nontrivial unless the target Π is constant, since

$$\mathbb{D}(\mathbf{x}, \mathbf{p}) = -e^{\mathbb{U}(\mathbf{x})} \nabla_{\mathbf{x}} \cdot (e^{-\mathbb{U}(\mathbf{x})} \mathbf{v}(\mathbf{p})) = \mathbf{v}(\mathbf{p}) \cdot \nabla \mathbb{U}(\mathbf{x}).$$

With these ingredients introduced, we consider the following second-order kinetic *mass-preserving enhanced sampling* dynamics,

$$(1.8) \quad \begin{cases} \partial_t f + \overbrace{\mathbf{v} \cdot \nabla_{\mathbf{x}} f}^{\text{Transport}} = \overbrace{\nu \Delta_{\mathbf{p}} f}^{\text{Diffusion}} + \overbrace{\kappa \Delta_{\mathbf{x}} f + \kappa \nabla_{\mathbf{x}} \cdot (f \nabla_{\mathbf{x}} \mathbb{U})}^{\text{Langevin}} + \overbrace{\mathcal{Q}_{\mathbb{U}, \mathbf{v}}[f]}^{\mathbb{D}\text{-alignment}}, \\ \mathcal{Q}_{\mathbb{U}, \mathbf{v}}[f](t, \mathbf{x}, \mathbf{p}) = \iint (\mathbb{D}(\mathbf{y}, \mathbf{q}) - \mathbb{D}(\mathbf{x}, \mathbf{p})) f(t, \mathbf{x}, \mathbf{p}) f(t, \mathbf{y}, \mathbf{q}) d\mathbf{y} d\mathbf{q}, \\ f(t=0, \mathbf{x}, \mathbf{p}) = f_0(\mathbf{x}, \mathbf{p}), \quad (\mathbf{x}, \mathbf{p}) \in \mathbb{T}^d \times \mathbb{T}^d. \end{cases}$$

In this model, the probability density $f \geq 0$ describes the probability of finding agents at position $\mathbf{x}$ moving with velocity $\mathbf{v}(\mathbf{p}) := (\sin(p_1), \sin(p_2), \dots, \sin(p_d)) \in [-1, 1]^d$. The model is similar to *underdamped Langevin dynamics* (see, e.g., [1, 4, 10, 11, 20, 32, 38, 43, 47, 61, 66] and the references therein). We assume initial normalization, which will be preserved by the dynamics,

$$(1.9) \quad \|f_0\|_{L^1_{\mathbf{x}, \mathbf{p}}} = \|f(t)\|_{L^1_{\mathbf{x}, \mathbf{p}}} = 1, \quad \text{for all } t \geq 0.$$

The Langevin structure $\kappa \Delta_{\mathbf{x}} f + \kappa \nabla_{\mathbf{x}} \cdot (f \nabla_{\mathbf{x}} \mathbb{U})$ helps stabilize the numerical simulation. The $\mathbb{D}$ -alignment term $\mathcal{Q}_{\mathbb{U}, \mathbf{v}}[f]$ is essential to preserve the invariant measure $\Pi \otimes (|\mathbb{T}|^{-d} \text{Lebesgue})$, and can be expressed as follows

$$\mathcal{Q}_{\mathbb{U}, \mathbf{v}}[f](t, \mathbf{x}, \mathbf{p}) = f(t, \mathbf{x}, \mathbf{p}) \left(-\mathbf{v}(\mathbf{p}) \cdot \nabla_{\mathbf{x}} \mathbb{U}(\mathbf{x}) + \iint f(t, \mathbf{y}, \mathbf{q}) \mathbf{v}(\mathbf{q}) \cdot \nabla_{\mathbf{y}} \mathbb{U}(\mathbf{y}) d\mathbf{y} d\mathbf{q} \right).$$

As in the model (1.5), the $\mathbb{D}$ -alignment ensures mass conservation (1.9). In a broad sense, the dynamics (1.8) exhibit features similar to those found in second-order models in flocking dynamics (see, e.g., [19, 24, 33, 34, 44, 55–57, 64] and the references therein).

The motivation behind the second-order enhanced dynamics is that, by extending the problem's phase space, one can design a simple drift to speed up convergence. Once the solution is sufficiently close to the target, a hydrodynamic averaging operation yields an approximation of the target measure. Let us consider the following hydrodynamic density:

$$(1.10) \quad [f](t, \mathbf{x}) := \int_{\mathbb{T}^d} f(t, \mathbf{x}, \mathbf{p}) d\mathbf{p}.$$

Our second main theorem states that the hydrodynamic density converges to the target density Π at a fast rate.

Theorem 1.2. *Consider equation (1.8) with smooth, normalized initial data $0 \leq f_0 \in C^\infty(\mathbb{T}_{\mathbf{x}}^d \times \mathbb{T}_{\mathbf{p}}^d)$ satisfying $\|f_0\|_{L_{\mathbf{x}, \mathbf{p}}^1} = 1$, and with a positive target density $0 < \Pi \in C^\infty(\mathbb{T}^d)$ ($1 \leq d \in \mathbb{N}$). There exists a threshold $\nu_0(d, \mathbb{U}, f_0)$ such that for any $0 < \nu \leq \nu_0$, and any $\kappa \in (0, \nu]$, the following estimate holds*

$$(1.11) \quad \text{Sample : } \|[f](t) - \Pi\|_{L_{\mathbf{x}}^1} \leq C_1(d, f_0, \mathbb{U}) \|\Pi\|_{L_{\mathbf{x}}^\infty} \|f_0 e^{\mathbb{U}}\|_{L^2} e^{-\delta \nu^{1/2} t}, \quad \forall t \geq 0;$$

$$(1.12) \quad \text{Approximation : } \|[f](t) - \Pi\|_{L_{\mathbf{x}}^\infty} \leq C_2(d, f_0, \mathbb{U}, Z) e^{-\delta_2 \nu^{1/2} t}, \quad \forall t \geq 0.$$

The parameter δ is independent of the dimension d and diffusion coefficient ν , but the parameter δ_2 depends on d . Moreover, for an arbitrary $\varsigma > 0$ and any test function $\varphi \in H_{\mathbf{x}}^\varsigma$, the following mixing estimate holds for all $t \in (0, \delta^{-1} \nu^{-1/2}]$:

$$(1.13) \quad \left| \int \varphi(\mathbf{x}) \Pi(\mathbf{x}) d\mathbf{x} - \int \varphi(\mathbf{x}) [f](t, \mathbf{x}) d\mathbf{x} \right| \leq \frac{C_3(d, \varsigma, \Pi, f_0)}{t^{1/2}} \|\varphi\|_{H_{\mathbf{x}}^\varsigma}, \quad \sigma = \max\{1, d/2 + \varsigma\}.$$

Remark 1.1. *The integral $\int \varphi[f] d\mathbf{x}$ is the hydrodynamic observable that captures probabilistic information, e.g., means and variances. The mixing estimate (1.13) implies that the hydrodynamic observable converges to the target with a rate that is independent of the parameter ν .*

Remark 1.2. *Other second-order models give similar results. We mention a Vicsek-type model, which is adopted in our numerical test in Section 4:*

$$(1.14) \quad \begin{cases} \partial_t f + \mathbf{p} \cdot \nabla_{\mathbf{x}} f = \nu \Delta_{\mathbf{p}} f + \kappa \Delta_{\mathbf{x}} f + \kappa \nabla_{\mathbf{x}} \cdot (f \nabla_{\mathbf{x}} \mathbb{U}) + \mathcal{Q}_{\mathbb{U}, \mathbf{p}}[f], \\ \mathcal{Q}_{\mathbb{U}, \mathbf{p}}[f](t, \mathbf{x}, \mathbf{p}) = \iint (\mathbf{q} \cdot \nabla_{\mathbf{y}} \mathbb{U} - \mathbf{p} \cdot \nabla_{\mathbf{x}} \mathbb{U}) f(t, \mathbf{y}, \mathbf{q}) f(t, \mathbf{x}, \mathbf{p}) d\mathbf{y} dS_{\mathbf{q}}, \\ f(t=0, \mathbf{x}, \mathbf{p}) = f_0(\mathbf{x}, \mathbf{p}), \quad (\mathbf{x}, \mathbf{p}) \in \mathbb{T}^d \times \mathbb{S}^{d-1}. \end{cases}$$

For this model, one can consider the corresponding hydrodynamic density $[f] = \int_{\mathbb{S}^{d-1}} f dS_{\mathbf{p}}$ and apply ideas from this paper to derive the sampling and approximation estimates.

Similar to the first-order case, we consider the following second-order *simplified enhanced sampling* dynamics:

$$(1.15) \quad \begin{cases} \partial_t \hat{f} + \mathbf{v} \cdot \nabla_{\mathbf{x}} \hat{f} = \nu \Delta_{\mathbf{p}} \hat{f} + \kappa \Delta_{\mathbf{x}} \hat{f} + \kappa \nabla_{\mathbf{x}} \cdot (\hat{f} \nabla_{\mathbf{x}} \mathbb{U}) - \hat{f} \mathbf{v} \cdot \nabla_{\mathbf{x}} \mathbb{U}, \\ \hat{f}(t=0, \mathbf{x}, \mathbf{p}) = f_0(\mathbf{x}, \mathbf{p}), \quad (\mathbf{x}, \mathbf{p}) \in \mathbb{T}^d \times \mathbb{T}^d. \end{cases}$$

The solutions of (1.8) and (1.15) are related by the formula: $f = \frac{\hat{f}}{\|\hat{f}\|_{L^1}}$. Hence, one can use a similar scheme as in the first-order enhanced model (simplified dynamics (1.15) + “mass-searching dynamics”) to simulate the PDE dynamics (1.8).

1.3. Analysis of the Models. To motivate the mathematical tools applied in our analysis, we would like to consider the simplest possible target function,

$$\Pi_0(\mathbf{x}) \equiv \frac{1}{|\mathbb{T}^d|} = \text{constant}, \quad \mathbf{x} \in \mathbb{T}^d, \quad d \geq 2.$$

We want to address the following key question.

Question A: Fix arbitrary initial data that is sufficiently smooth. Is it possible to design a dynamical system whose solution converges to the constant state Π_0 within a short time?

We propose two solutions to address the above problem, which lead, respectively, to the first- and second-order models we design.

Approach #1: Introducing time-dependent flows. Incorporate the time-dependent divergence-free vector fields $\mathbb{V}(t, \mathbf{x})$, and keep track of the passive scalar solution,

$$(1.16) \quad \partial_t \rho + \nabla_{\mathbf{x}} \cdot (\mathbb{V} \rho) = \nu \Delta \rho, \quad \nabla_{\mathbf{x}} \cdot \mathbb{V} \equiv 0, \quad \rho(t=0) = \rho_0.$$

Define the average $\bar{\rho} := \frac{1}{|\mathbb{T}^d|} \int \rho d\mathbf{x}$. The standard L^2 -energy estimate shows that the norm $\|\rho - \bar{\rho}\|_{L^2}(t)$ decays to zero at a rate of $\mathcal{O}(\nu)$. It is a classical problem in the fluid mechanics community to identify suitable fluid vector fields $\mathbb{V}$ that improve the decay rate of the solution. These flows are commonly referred to as *relaxation enhancing* flows. We refer interested readers to the classical paper [14] for a precise definition. This phenomenon of flow-induced acceleration of dissipation is also called the *enhanced dissipation*. Later research also shows that the *enhanced dissipation* has a direct connection to the *mixing* phenomenon (see, e.g., [16, 22] and the expository article [15]). Specifically, smooth, exponentially mixing vector fields $\mathbb{V}$ yield a solution ρ of (1.16) with rapid decay

$$(1.17) \quad \|\rho(t) - \bar{\rho}\|_{L^2} \leq C \|\rho_0 - \bar{\rho}_0\|_{L^2} e^{-\delta |\log \nu|^{-2} t}.$$

We observe that if $0 < \nu \ll 1$, the decay rate here is much faster than ν . These exponentially mixing vector fields (such as the Pierrehumbert flow [51] and the stochastic Navier-Stokes solutions considered in [5]) have recently been identified using probabilistic methods (see, e.g., [9, 21]). One can design fast sampling algorithms by incorporating these random, exponentially mixing flows into the sampling dynamics. This is essentially the idea presented in the pioneering paper [12].

However, it has also been pointed out in [12] that applying exponentially mixing flows can lead to numerical challenges in sampling because the exponentially mixing flows constructed in the literature are random and vary drastically over time. For example, to generate the exponential decay in (1.17), the random vector fields $\mathbb{V}$ will need to switch orientation over time intervals of length $\mathcal{O}(1)$. When $\nu \ll 1$, these intervals are much shorter than the classical decay time scale of the heat equation, which is $\mathcal{O}(\nu^{-1})$.

Our proposal to resolve the problem above is to consider a slowly varying, *deterministic* alternating shear flow $\mathbb{V}$. Since the construction is straightforward for even dimensions, we assume that $d = 2\mathfrak{d}$ for some $\mathfrak{d} \in \mathbb{N} \setminus \{0\}$. The building blocks for this alternating shear are the following

$$(1.18) \quad \mathbf{u}_1(\{x_{\mathfrak{d}+j}\}_{j=1}^{\mathfrak{d}}) = \sum_{j=1}^{\mathfrak{d}} \mathcal{U}(x_{\mathfrak{d}+j}) \mathbf{e}_j, \quad \mathbf{u}_2(\{x_j\}_{j=1}^{\mathfrak{d}}) = \sum_{j=1}^{\mathfrak{d}} \mathcal{U}(x_j) \mathbf{e}_{\mathfrak{d}+j}.$$

Here, the set $\{\mathbf{e}_j\}_{j=1}^{\mathfrak{d}}$ is the collection of the canonical basis vectors of $\mathbb{R}^{\mathfrak{d}}$ and $\mathcal{U} \in C^\infty$. When we consider the passive scalar equation subject to one of these shear flows,

$$\partial_t \rho + \mathbf{u}_1 \cdot \nabla_{\mathbf{x}} \rho = \nu \Delta \rho, \quad 0 < \nu \leq 1.$$

There will be two distinct components of the solution. The first is the stream-line average in the flow direction $\langle \rho \rangle_1 := |\mathbb{T}|^{-\mathfrak{d}} \int \rho(t, \mathbf{x}) dx_1 \cdots dx_{\mathfrak{d}}$ and the other is the fluctuation $\rho - \langle \rho \rangle_1$. In two

dimensions ($d = 2$), it is well-established that the fluctuation experiences *enhanced dissipation* under suitable structural assumptions on the shear profile $\mathcal{U}$. The study of this shear flow-induced enhanced dissipation (and related hypoelliptic smoothing) can be traced back to the nineteenth-century work of Lord Kelvin [40], and to A. Kolmogorov [41] and L. Hörmander [39]. In recent years, the sharp enhanced dissipation estimate has been obtained for general shear flows. Assuming that $d = 2$, it is shown in the works [6, 65] that if the shear flow profile $\mathcal{U}$ has only finitely many non-degenerate critical points, the enhanced dissipation holds. In particular, if $\mathcal{U} = \sin$, the following decay holds for all $0 < \nu \leq 1$,

$$(1.19) \quad \|\rho(t) - \langle \rho \rangle_\ell(t)\|_{L^2} \leq C_{\sin} \|\rho_0 - \langle \rho_0 \rangle_\ell\|_{L^2} e^{-\delta_{\sin} \nu^{1/2} t}, \quad \mathcal{U}(\cdot) = \sin(\cdot), \quad \forall t \geq 0.$$

Here, the constants $C_{\sin} \geq 1$, $\delta_{\sin} \in (0, 1)$ are independent of the diffusion coefficient ν . The $\nu^{1/2}$ -rate has been proved to be sharp in [18]. We refer interested readers to the related works [2, 13, 23, 25, 27].

Finally, by alternately applying the shear flows $\mathbf{u}_1$ and $\mathbf{u}_2$ (1.18) (with profile $\mathcal{U} = \sin$) on time intervals of length $\mathcal{O}(\nu^{-1/2})$, one establishes enhanced dissipation on $\mathbb{T}^2$:

$$(1.20) \quad \|\rho(t) - \bar{\rho}\|_{L^2} \leq C \|\rho_0 - \bar{\rho}_0\|_{L^2} e^{-\delta \nu^{1/2} t}.$$

This result was first derived in [36]. In this paper, we derive the higher-dimensional analog.

Theorem 1.3 (Enhanced Dissipation). *Consider the passive scalar equation in even dimension,*

$$(1.21) \quad \partial_t \eta + \mathbb{V}(t) \cdot \nabla_{\mathbf{x}} \eta = \nu \Delta_{\mathbf{x}} \eta, \quad \eta(t=0) = \eta_0 \in L^2(\mathbb{T}^d), \quad d = 2d.$$

Let the time-dependent flow $\mathbb{V}$ be the linear combination of (1.18) $_{\mathcal{U}=\sin}$, $\mathbb{V}(t, \mathbf{x}) = \sum_{i=1}^2 \Phi^{(i)}(t) \mathbf{u}_i(\mathbf{x})$. There exists a family of coefficient functions $\{\Phi^{(i)}\}_{i=1}^2$ that vary on a time scale $\mathcal{O}(\nu^{-1/2})$ such that the following enhanced dissipation estimate holds for any $0 < \nu \leq 1$:

$$(1.22) \quad \|\eta(s+t, \cdot) - \bar{\eta}(s+t)\|_{L^2_{\mathbf{x}}(\mathbb{T}^d)} \leq C_{\text{ED}} \|\eta(s, \cdot) - \bar{\eta}(s)\|_{L^2_{\mathbf{x}}(\mathbb{T}^d)} e^{-\delta_{\text{ED}} \nu^{1/2} t}, \quad \forall s, t \in \mathbb{R}_+.$$

Here, $\delta_{\text{ED}} \in (0, 1)$ and $C_{\text{ED}} \geq 1$ are constants that are independent of the solution and d, ν .

Proof. The proof of Theorem 1.3 is detailed in Section D. □

Remark 1.3. The schematic picture of the coefficient functions $\{\Phi^{(i)}\}_{i=1}^2$ is shown in Figure 1 and the construction is detailed in Section D.

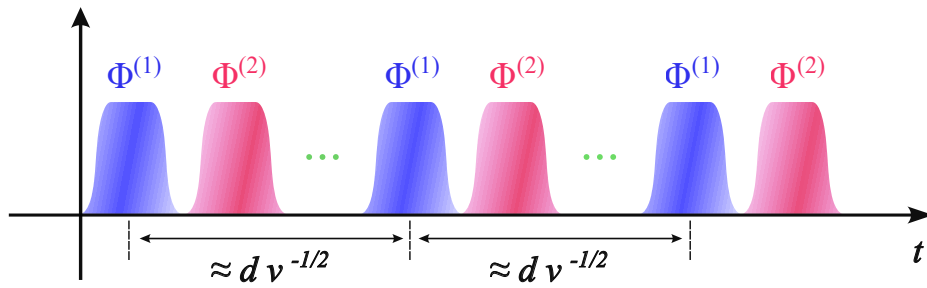


FIGURE 1. The coefficient functions $\Phi^{(i)}$

Remark 1.4. The parameter δ in Theorem 1.1 is closely connected to the parameters in Theorem 1.3:

$$(1.23) \quad \delta := \frac{\delta_{\text{ED}}}{2 + \log C_{\text{ED}}} < \delta_{\text{ED}}.$$

Remark 1.5. We remark that our construction is robust in the sense that replacing the shear profile $\mathcal{U}$ by another function with finitely many nondegenerate critical points leaves the result unchanged.

Remark 1.6. *One can extend the result to odd dimensions $d = 2d + 1$ by a small modification of the flow. We will detail the construction in Section D.*

In conclusion, we propose replacing the random exponentially mixing flow in the previous literature by slowly varying deterministic shear flows. Even though the enhanced dissipation rate (1.20) is slower than the rate in (1.17), it has the advantage that the flow fields do not change too drastically and hence have the potential to make the numerics simpler.

Approach #2: Extend the dimension of the phase space and use simpler flows.

The second approach is to introduce the momentum variable $\mathbf{p} \in \mathbb{T}^d$, and to consider dynamics

$$(1.24) \quad \partial_t f + \mathbf{v}(\mathbf{p}) \cdot \nabla_{\mathbf{x}} f = \nu \Delta_{\mathbf{p}} f, \quad f(t=0) = f_0, \quad (\mathbf{x}, \mathbf{p}) \in \mathbb{T}_{\mathbf{x}}^d \times \mathbb{T}_{\mathbf{p}}^d, \quad d = 1, 2, 3, \dots$$

Here, the transport vector field is $\mathbf{v}(\mathbf{p}) := \sum_{j=1}^d \sin(p_j) \mathbf{e}_{\mathbf{x},j}$, and $\{\mathbf{e}_{\mathbf{x},j}\}_{j=1}^d$ are the canonical basis vectors for $\mathbb{T}_{\mathbf{x}}^d$. Consider the average (in the shearing direction) $\langle f \rangle_{\mathbf{x}}$ and the remainder $\tilde{f}$,

$$(1.25) \quad \langle f \rangle_{\mathbf{x}}(t, \mathbf{p}) := \frac{1}{|\mathbb{T}|^d} \int_{\mathbb{T}^d} f(t, \mathbf{x}, \mathbf{p}) d\mathbf{x}, \quad \tilde{f}(t, \mathbf{x}, \mathbf{p}) = \tilde{\mathbb{P}}_{\mathbf{x}} f(t, \mathbf{x}, \mathbf{p}) := f(t, \mathbf{x}, \mathbf{p}) - \langle f \rangle_{\mathbf{x}}(t, \mathbf{p}).$$

Similar to the estimate (1.19), one can derive the estimate (Corollary D.1, Section D)

$$(1.26) \quad \|\tilde{f}(t)\|_{L_{\mathbf{x},\mathbf{p}}^2} \leq C \|f_0\|_{L_{\mathbf{x},\mathbf{p}}^2} e^{-\delta \nu^{1/2} t}, \quad \forall t \geq 0.$$

A direct consequence of the estimate (1.26) is that the hydrodynamic density (1.10) converges to the equilibrium with enhanced dissipation rate $\nu^{1/2}$: $\|[f](t) - [\bar{f}]\|_{L_{\mathbf{x}}^2} \leq C \|f_0\|_{L_{\mathbf{x},\mathbf{p}}^2} e^{-\delta \nu^{1/2} t}, \quad \forall t \geq 0$.

In conclusion, if the target density $\Pi_0 \equiv \frac{1}{|\mathbb{T}|^d}$, both the first-order dynamics (1.16) and the second-order dynamics (1.24) provide fast-converging solutions towards Π_0 . This settles question A. Next, we turn to address the following key question.

Question B: Fix arbitrary initial data that is sufficiently smooth. Is it possible to design a dynamical system whose solution converges to any given target density $\Pi(\mathbf{x}) = Z^{-1} e^{-\mathbb{U}(\mathbf{x})} \in L_+^1 \cap C^2(\mathbb{T}^d)$ within a short time?

As it turns out, we have already introduced all the essential ingredients to resolve question B. The missing piece to address is how to introduce nontrivial density information into the dynamics (1.16) and (1.24).

a) First-order model: First, we rewrite the dynamics (1.16) as follows:

$$\partial_t \rho - \nu \nabla_{\mathbf{x}} \cdot \left(\rho \nabla_{\mathbf{x}} \log \left(\frac{\rho}{\Pi_0} \right) \right) = -\Pi_0 \nabla_{\mathbf{x}} \cdot \left(\frac{\rho \mathbb{V}}{\Pi_0} \right), \quad \nabla_{\mathbf{x}} \cdot \mathbb{V} \equiv 0, \quad \rho(t=0) = \rho_0.$$

Naively, we could change the target density to $\Pi(\mathbf{x}) = e^{-\mathbb{U}(\mathbf{x})}/Z$. However, a direct replacement would break the mass-conservation structure in general. Motivated by the designs in [42, 62], we incorporate a “mass correction” term in the dynamics to ensure conservation of the total mass:

$$(1.27) \quad \begin{cases} \partial_t \rho - \nu \nabla_{\mathbf{x}} \cdot \left(\rho \nabla_{\mathbf{x}} \log \left(\frac{\rho}{\Pi} \right) \right) = -\Pi \nabla_{\mathbf{x}} \cdot \left(\frac{\rho \mathbb{V}}{\Pi} \right) - \rho \int \rho \mathbb{V} \cdot \nabla_{\mathbf{y}} \log \Pi d\mathbf{y}, \\ \nabla_{\mathbf{x}} \cdot \mathbb{V}(t, \mathbf{x}) = 0, \quad \rho(t=0, \mathbf{x}) = \rho_0(\mathbf{x}), \quad (t, \mathbf{x}) \in \mathbb{R}_+ \times \mathbb{T}^d. \end{cases}$$

Recalling the definition of drift defect (1.3) and applying the divergence-free condition of $\mathbb{V}$, we observe that the modified transport term naturally leads to the drift defect:

$$-\Pi \nabla_{\mathbf{x}} \cdot \left(\frac{\rho \mathbb{V}}{\Pi} \right) = -\nabla_{\mathbf{x}} \cdot (\rho \mathbb{V}) + \rho e^{\mathbb{U}} \nabla_{\mathbf{x}} \cdot (e^{-\mathbb{U}} \mathbb{V}) = -\nabla_{\mathbf{x}} \cdot (\rho \mathbb{V}) - \rho \mathbb{D}.$$

Hence, the equation (1.27), when rewritten properly, recovers the $\mathbb{D}$ -alignment $\mathcal{Q}_{\mathbb{U},\mathbb{V}}[\rho]$, yielding the enhanced model (1.5).

To conclude the discussion of the first-order model, we highlight that the equation (1.5) is equivalent to the following system satisfied by the quotient $h \stackrel{\text{def}}{=} \rho/\Pi$:

$$(1.28) \quad \begin{cases} \partial_t h + \mathbb{V} \cdot \nabla_{\mathbf{x}} h = \nu \Delta_{\mathbf{x}} h - \nu \nabla_{\mathbf{x}} \mathbb{U} \cdot \nabla_{\mathbf{x}} h - h \int h \mathbb{V} \cdot \nabla_{\mathbf{x}} \Pi d\mathbf{x}, \\ \nabla_{\mathbf{x}} \cdot \mathbb{V}(t, \mathbf{x}) = 0, \quad h(t=0, \mathbf{x}) = \frac{\rho_0(\mathbf{x})}{\Pi(\mathbf{x})}. \end{cases}$$

b) Second-order model A similar argument yields that the passive scalar dynamics (1.24) can be generalized to the following equation, which is equivalent to the equation (1.8):

$$\begin{cases} \partial_t f + \Pi \nabla_{\mathbf{x}} \cdot \left(\frac{\mathbf{v}(\mathbf{p}) f}{\Pi} \right) = \nu \Delta_{\mathbf{p}} f + \kappa \nabla_{\mathbf{x}} \cdot \left(f \nabla \log \left(\frac{f}{\Pi} \right) \right) - f \iint f \mathbf{v}(\mathbf{p}) \cdot \nabla_{\mathbf{x}} \log \Pi \, d\mathbf{x} d\mathbf{p} \\ f(t=0, \mathbf{x}, \mathbf{p}) = f_0(\mathbf{x}, \mathbf{p}), \quad \mathbf{x} \in \mathbb{T}^d, \mathbf{p} \in \mathbb{T}^d. \end{cases}$$

Now we can rewrite it with $h \stackrel{\text{def}}{=} f/\Pi$ as follows,

$$\begin{cases} \partial_t h + \mathbf{v}(\mathbf{p}) \cdot \nabla_{\mathbf{x}} h = \nu \Delta_{\mathbf{p}} h + \kappa \Delta_{\mathbf{x}} h - \kappa \nabla_{\mathbf{x}} \mathbb{U} \cdot \nabla_{\mathbf{x}} h - h \iint h \mathbf{v}(\mathbf{p}) \cdot \nabla_{\mathbf{x}} \Pi \, d\mathbf{x} d\mathbf{p}, \\ h(t=0, \mathbf{x}, \mathbf{p}) = \frac{f_0(\mathbf{x}, \mathbf{p})}{\Pi(\mathbf{x})}, \quad \mathbf{x} \in \mathbb{T}^d, \mathbf{p} \in \mathbb{T}^d. \end{cases}$$

In Sections 2–3, we develop enhanced dissipation estimates, which are in the spirit of (1.22) and (1.26) for the formulations presented in this section.

1.4. The Mass-searching Dynamics. As discussed in the previous sections, for an accurate simulation of the full enhanced dynamics (1.5) and (1.8), a mass-searching procedure is required to recover the exact solutions. We propose a dynamics to compute or approximate the mass of a given function $e^{-\mathbb{W}}$. This dynamics can be applied to find the normalizing constant Z in $\Pi(\mathbf{x}) = e^{-\mathbb{U}}/Z$ (1.1) or the mass of the solutions $\hat{\rho}(t, \mathbf{x})$, $[\hat{f}](t, \mathbf{x}) \geq 0$ to the simplified dynamics (1.7).

To compute the mass of a given function $F = e^{-\mathbb{W}} \in C_{\mathbf{x}}^3$, we propose the following first-order dynamics to identify $M := \|F\|_{L_{\mathbf{x}}^1}$:

$$(1.29) \quad \begin{cases} \partial_t \omega + \mathbb{V} \cdot \nabla_{\mathbf{x}} \omega - \omega \mathbb{V} \cdot \nabla_{\mathbf{x}} \mathbb{W} = \nu e^{\mathbb{W}} \Delta_{\mathbf{x}} (\omega e^{-\mathbb{W}}), \\ \nabla_{\mathbf{x}} \cdot \mathbb{V}(t, \mathbf{x}) = 0, \quad \omega(t=0, \mathbf{x}) \equiv |\mathbb{T}|^d, \quad (t, \mathbf{x}) \in \mathbb{R}_+ \times \mathbb{T}^d. \end{cases}$$

Here, the vector field $\mathbb{V}$ is defined in (1.18) _{$\mathcal{U}=\text{sin}$} and the parameter ν is chosen as in Theorem 1.3. The diffusion operator is reformulated as follows:

$$\nu e^{\mathbb{W}} \Delta_{\mathbf{x}} (\omega e^{-\mathbb{W}}) = \nu \Delta_{\mathbf{x}} \omega - 2\nu \nabla_{\mathbf{x}} \omega \cdot \nabla_{\mathbf{x}} \mathbb{W} + \nu \omega (-\Delta_{\mathbf{x}} \mathbb{W} + |\nabla_{\mathbf{x}} \mathbb{W}|^2), \quad \nabla_{\mathbf{x}} \mathbb{W} = -F^{-1} \nabla_{\mathbf{x}} F.$$

Hence, the gradient $\nabla \mathbb{W}$ is singular if F is close to zero. To avoid this, we assume that

$$(1.30) \quad \min_{\mathbf{x}} F(\mathbf{x}) \geq 1.$$

For functions $\tilde{F}$ that fail the condition (1.30), we can define the shifted function $F := \tilde{F} + 1$, apply the dynamics (1.29) to determine $\|F\|_{L^1}$, and then subtract the mass introduced by the shift to obtain the true mass, i.e., $\|\tilde{F}\|_{L_{\mathbf{x}}^1} = \|F\|_{L_{\mathbf{x}}^1} - |\mathbb{T}|^d$.

We emphasize that no quadrature is needed to simulate the dynamics (1.29), and the solutions converge to the mass $\|e^{-\mathbb{W}}\|_{L_{\mathbf{x}}^1}$ at an enhanced rate. This is the content of the next theorem.

Theorem 1.4. *Consider solutions to (1.29) subject to the vector field $\mathbb{V}$ in (1.18) _{$\mathcal{U}=\text{sin}$} . Assume that the parameter ν is chosen as in Theorem 1.3 and the target function F satisfies the condition (1.30). Then the following estimate holds*

$$\|\omega(t) e^{-\mathbb{W}} - M\|_{L^2} \leq C \| |\mathbb{T}|^d e^{-\mathbb{W}} - M \|_{L^2} e^{-\delta_{\text{ED}} \nu^{1/2} t}, \quad \forall t \geq 0.$$

Here, the mass $M := \|F\|_{L^1_{\mathbf{x}}}$.

The remainder of the paper is organized as follows: In Section 2, we present the analysis of the first-order model (1.5) and the proof of Theorem 1.1; in Section 3, we present the proof of Theorem 1.2, which characterizes the behavior of the second-order model (1.8); in Section 4, we present a numerical experiment to illustrate the fast convergence of our dynamics (see, e.g., Figure 3). In the appendix, we present several technical theorems needed in the main text and the particle dynamics associated with the systems (1.5) and (1.8).

Notation: Throughout the paper, the symbol C denotes a generic constant that may vary from line to line. The notation T_j ($j = 1, 2, \dots$) refers to intermediate terms appearing within a proof of a theorem or lemma. The T_j 's are local to the specific argument and can be redefined when we start the proof of a new theorem or lemma.

2. THE FIRST-ORDER ENHANCED DYNAMICS

2.1. Preliminaries. In this section, we present some preliminary facts about the system (1.28). Since ρ_0 and Π are regular, a standard argument yields that the equation (1.28) is locally well-posed. Hence, we only derive a priori estimates in the forthcoming text.

An important aspect of the proposed model is that it preserves the total mass of the density ρ . To see this, we compute the time derivative of the mass. Thanks to the maximum principle, the solution ρ stays positive as long as it stays regular. Hence, the mass $\|\rho(t)\|_{L^1}$ of the solution is the integral $\int \rho(t, \mathbf{x}) d\mathbf{x}$. We recall (1.27) and compute the time evolution of the mass as follows:

$$\frac{d}{dt} \int \rho d\mathbf{x} = \int_{\mathbb{T}^d} \nabla_{\mathbf{x}} \cdot (-\nabla \rho + \nu \nabla_{\mathbf{x}} \rho + \nu \rho \nabla_{\mathbf{x}} \mathbb{U}) d\mathbf{x} + \int \rho \left(-\nabla \cdot \nabla_{\mathbf{x}} \mathbb{U} + \int \rho \nabla \cdot \nabla_{\mathbf{x}} \mathbb{U} d\mathbf{x} \right) d\mathbf{x}.$$

Thanks to the divergence theorem, the first term is 0. For the second term, we rewrite it as follows:

$$\frac{d}{dt} \int \rho d\mathbf{x} = \left(\int \rho \nabla \cdot \nabla_{\mathbf{x}} \mathbb{U} d\mathbf{x} \right) \left(-1 + \int \rho d\mathbf{x} \right).$$

Since the first factor on the right-hand side is continuous in time, $\int \rho(t, \mathbf{x}) d\mathbf{x} \equiv 1$ is the unique solution to the above ODE with the prescribed initial mass. As a result, we conclude that the total mass is preserved by the dynamics: $\|\rho(t, \cdot)\|_{L^1_{\mathbf{x}}(\mathbb{T}^d)} \equiv \|\rho_0\|_{L^1_{\mathbf{x}}(\mathbb{T}^d)} = 1$, $\forall t \in \mathbb{R}_+$.

2.2. Proof of Theorem 1.1. In this section, we consider the first-order model (1.5) and derive the result (1.6). The proof is organized in steps.

Step #1: Key estimate. To distill the key estimate to establish, we observe that the following relations for all $t \geq 0$ imply estimate (1.6):

$$(2.1) \quad \left\| \left(\frac{\rho}{\Pi} \right)(t) \right\|_{L^2_{\mathbf{x}}} \leq C_1 \left\| \left(\frac{\rho_0}{\Pi} \right) \right\|_{L^2_{\mathbf{x}}} e^{-\delta \nu^{1/2} t}, \quad (\text{Key Estimate})$$

$$(2.2) \quad \left| \left(\frac{\rho}{\Pi} \right)(t) - 1 \right| \leq C_2 e^{-\delta \nu^{1/2} t}.$$

Here, $\bar{F} := \frac{1}{|\mathbb{T}|^d} \int_{\mathbb{T}^d} F d\mathbf{x}$ is the spatial average and $\tilde{F} := F - \bar{F}$ is the fluctuation. The constants C_1 (defined in (2.7)), C_2 (defined in (2.3)) depend only on the initial density ρ_0 , and the potential $\mathbb{U}$. We recall that the parameter δ (1.23) can be expressed in terms of C_{ED} , δ_{ED} in (1.22). To see the implication of the result (1.6), we apply the triangle inequality,

$$\left\| \frac{\rho}{\Pi}(t) - 1 \right\|_{L^2_{\mathbf{x}}} \leq \left(\left\| \left(\frac{\rho}{\Pi} \right)(t) \right\|_{L^2_{\mathbf{x}}} + |\mathbb{T}|^{d/2} \left| \left(\frac{\rho}{\Pi} \right)(t) - 1 \right| \right) \leq \left(C_1 \left\| \left(\frac{\rho_0}{\Pi} \right) \right\|_{L^2_{\mathbf{x}}} + |\mathbb{T}|^{d/2} C_2 \right) e^{-\delta \nu^{1/2} t}.$$

Defining $C_{\dagger} := C_1 \left\| \left(\frac{\rho_0}{\Pi} \right) \right\|_{L^2_{\mathbf{x}}} + |\mathbb{T}|^{d/2} C_2$ yields the result.

Next, we observe that the estimate (2.2) is a simple consequence of (2.1). Thanks to the mass conservation, the following equality holds

$$0 = \int \rho d\mathbf{x} - \int \Pi d\mathbf{x} = \int \left(\frac{\rho}{\Pi} - 1 \right) \Pi d\mathbf{x} = \int \left[\left(\frac{\rho}{\Pi} \right) + \left(\frac{\rho}{\Pi} \right) - 1 \right] \Pi d\mathbf{x} \implies \left(\frac{\rho}{\Pi} \right) - 1 = - \int \left(\frac{\rho}{\Pi} \right) \Pi d\mathbf{x}.$$

Hence,

$$(2.3) \quad \left| \left(\frac{\rho}{\Pi} \right)(t) - 1 \right| = \left| \int \left(\frac{\rho}{\Pi} \right) \Pi d\mathbf{x} \right| \leq \|\Pi\|_{L^2} \left\| \left(\frac{\rho}{\Pi} \right)(t) \right\|_{L^2} \\ \leq C_1 \|\Pi\|_{L^2} \left\| \left(\frac{\rho_0}{\Pi} \right) \right\|_{L^2} e^{-\delta\nu^{1/2}t} = C_1 \|e^{-\mathbb{U}}\|_{L^2} \left\| \left(\rho_0 e^{\mathbb{U}} \right) \right\|_{L^2} e^{-\delta\nu^{1/2}t} =: C_2(C_1, \rho_0, \mathbb{U}) e^{-\delta\nu^{1/2}t}.$$

This is (2.2). Hence, we conclude that the *key estimate* is the nonlinear enhanced dissipation (2.1), whose proof stands as the focus of the remainder of the section.

Step #2: Bootstrap. To develop the key estimate (2.1), we recall the quantity $h = \rho\Pi^{-1}$ and the equation (1.28) that it satisfies,

$$(2.4) \quad \partial_t h + \mathbb{V} \cdot \nabla_{\mathbf{x}} h = \nu \Delta_{\mathbf{x}} h - \nu \nabla_{\mathbf{x}} \mathbb{U} \cdot \nabla_{\mathbf{x}} h - Qh, \quad Q(t) := \int_{\mathbb{T}^d} h \mathbb{V} \cdot \nabla_{\mathbf{x}} \Pi d\mathbf{x}, \quad h \geq 0.$$

The challenge in analyzing this system stems from the reaction term Qh . Unlike other terms on the right-hand side of (2.4), it does not have a small parameter ν in front. Therefore, to control this term, we need to estimate the time integral of the quantity Q . This observation motivates our bootstrap setup. To prove the key estimate (2.1), we apply a bootstrap argument. Assume that $[0, T_*)$ is the maximal interval on $[0, \infty)$ such that the following two *hypotheses* are satisfied:

a) Control on $Q(t)$: for all $t \in [0, T_*)$ there holds

$$(2.5a) \quad \left| \int_0^t Q(s) ds \right| \leq 2\mathfrak{B}_1(\rho_0, \mathbb{U}) := 2 \left(1 + \left| \log \left(\frac{2\|e^{-\mathbb{U}}\|_{L^\infty_{\mathbf{x}}}^{-1}}{3\|\rho_0 e^{\mathbb{U}}\|_{L^1_{\mathbf{x}}}} \right) \right| + \left| \log \left(\frac{2\|e^{\mathbb{U}}\|_{L^\infty_{\mathbf{x}}}}{\|\rho_0 e^{\mathbb{U}}\|_{L^1_{\mathbf{x}}}} \right) \right| \right).$$

b) Enhanced Dissipation

$$(2.5b) \quad \|\tilde{h}(t)\|_{L^2} \leq 2\mathfrak{B}_2(\mathfrak{B}_1) \|\tilde{h}_0\|_{L^2} \exp \left\{ -\delta\nu^{1/2}t \right\}, \quad \mathfrak{B}_2 = e^{2\mathfrak{B}_1+2}.$$

Here, the parameter δ is defined in (1.23).

To establish the estimate (2.1), it is sufficient to prove the following stronger *conclusion*: If the threshold $0 < \nu_0(\rho_0, \mathbb{U}, \delta^{-1})$ is small enough, then for all $0 < \nu \leq \nu_0$, the following estimates hold

a) Control on $Q(t)$

$$(2.6a) \quad \left| \int_0^t Q(s) ds \right| \leq \mathfrak{B}_1(\rho_0, \mathbb{U}), \quad t \in [0, T_*);$$

b) Enhanced Dissipation

$$(2.6b) \quad \|\tilde{h}(t)\|_{L^2} \leq \mathfrak{B}_2(\rho_0, \mathbb{U}) \|\tilde{h}_0\|_{L^2} \exp \left\{ -\delta\nu^{1/2}t \right\}, \quad t \in [0, T_*).$$

Remark 2.1 (The choice of ν_0). *The ν_0 satisfies all constraints (2.9), (2.12) and (2.15).*

Once the estimates (2.6a) and (2.6b) are established, we can see that $T_* = \infty$ and the following estimate holds

$$\left\| \left(\frac{\rho}{\Pi} \right)(t) \right\|_{L^2_{\mathbf{x}}} = \|\tilde{h}(t)\|_{L^2_{\mathbf{x}}} \leq e^{2\mathfrak{B}_1+2} \|\tilde{h}_0\|_{L^2_{\mathbf{x}}} e^{-\delta\nu^{1/2}t}, \quad \forall t \geq 0.$$

This is (2.1) with

$$(2.7) \quad C_1(\rho_0, \mathbb{U}) := \exp \left\{ 4 + 2 \left| \log \left(\frac{2\|e^{-\mathbb{U}}\|_{L_{\mathbf{x}}^{\infty}}^{-1}}{3\|\rho_0 e^{\mathbb{U}}\|_{L_{\mathbf{x}}^1}} \right) \right| + 2 \left| \log \left(\frac{2\|e^{\mathbb{U}}\|_{L_{\mathbf{x}}^{\infty}}}{\|\rho_0 e^{\mathbb{U}}\|_{L_{\mathbf{x}}^1}} \right) \right| \right\}.$$

This will conclude the proof of Theorem 1.1.

Step #3: Proof of conclusion (2.6a). To capture the quantity Q , we observe that it varies in time and is constant in space. Hence, it mainly causes fluctuations in the total mass of the solution h . Therefore, one can apply an a priori estimate of the L^1 -norm of h to recover Q . We compute, $\frac{d}{dt}\|h\|_{L_{\mathbf{x}}^1} = \nu \int_{\mathbb{T}^d} \tilde{h} \Delta_{\mathbf{x}} \mathbb{U} d\mathbf{x} - \|h\|_{L_{\mathbf{x}}^1} Q$, which implies

$$(2.8) \quad \begin{aligned} \|h(t)\|_{L_{\mathbf{x}}^1} &= \|h_0\|_{L_{\mathbf{x}}^1} \exp \left\{ - \int_0^t Q(s) ds \right\} + \nu \int_0^t \exp \left\{ - \int_s^t Q(\tau) d\tau \right\} \int_{\mathbb{T}^d} \tilde{h}(s, \mathbf{x}) \Delta_{\mathbf{x}} \mathbb{U}(\mathbf{x}) d\mathbf{x} ds \\ &=: T_1 + T_2, \quad \forall t \in [0, T_*]. \end{aligned}$$

We first observe that thanks to the bootstrap assumptions (2.5a) and (2.5b), the second term T_2 can be controlled as follows:

$$\begin{aligned} |T_2| &\leq \nu \int_0^t \exp\{4\mathfrak{B}_1\} \|\tilde{h}(s)\|_{L_{\mathbf{x}}^2} \|\Delta_{\mathbf{x}} \mathbb{U}\|_{L_{\mathbf{x}}^2} ds \leq \nu \int_0^t 2\mathfrak{B}_2 \exp\{4\mathfrak{B}_1\} \|\tilde{h}(0)\|_{L_{\mathbf{x}}^2} e^{-\delta\nu^{1/2}s} \|\Delta_{\mathbf{x}} \mathbb{U}\|_{L_{\mathbf{x}}^2} ds \\ &\leq \nu^{1/2} \frac{1}{\delta} 2\mathfrak{B}_2 \exp\{4\mathfrak{B}_1\} Z \left\| \widetilde{(\rho_0 e^{\mathbb{U}})} \right\|_{L_{\mathbf{x}}^2} \|\Delta_{\mathbf{x}} \mathbb{U}\|_{L_{\mathbf{x}}^2}. \end{aligned}$$

Here, Z is the normalizing factor. Hence, for any

$$(2.9) \quad 0 < \nu \leq \nu_0(\rho_0, \mathbb{U}, \delta^{-1}) \leq \left(\frac{\delta \|\rho_0 e^{\mathbb{U}}\|_{L_{\mathbf{x}}^1}}{4(1 + \left\| \widetilde{(\rho_0 e^{\mathbb{U}})} \right\|_{L_{\mathbf{x}}^2})(1 + \|\Delta_{\mathbf{x}} \mathbb{U}\|_{L_{\mathbf{x}}^2}) \mathfrak{B}_2 \exp\{6\mathfrak{B}_1\}} \right)^2,$$

the following relation holds

$$|T_2| \leq \frac{Z}{2} \|\rho_0 e^{\mathbb{U}}\|_{L_{\mathbf{x}}^1} \exp\{-2\mathfrak{B}_1\} = \frac{1}{2} \|h_0\|_{L_{\mathbf{x}}^1} \exp\{-2\mathfrak{B}_1\} \leq \frac{1}{2} \|h_0\|_{L_{\mathbf{x}}^1} \exp \left\{ - \int_0^t Q(s) ds \right\}.$$

As a consequence of this estimate, we obtain the following

$$\|h(t)\|_{L_{\mathbf{x}}^1} \in \left[\frac{1}{2} \|h_0\|_{L_{\mathbf{x}}^1} e^{-\int_0^t Q(s) ds}, \frac{3}{2} \|h_0\|_{L_{\mathbf{x}}^1} e^{-\int_0^t Q(s) ds} \right].$$

On the other hand, we also have the following observation

$$\|h\|_{L_{\mathbf{x}}^1} = \|\rho/\Pi\|_{L_{\mathbf{x}}^1} \leq \|\rho\|_{L_{\mathbf{x}}^1} \|\Pi^{-1}\|_{L_{\mathbf{x}}^{\infty}} \quad \text{and} \quad \|\rho\|_{L_{\mathbf{x}}^1} = \|h\Pi\|_{L_{\mathbf{x}}^1} \leq \|h\|_{L_{\mathbf{x}}^1} \|\Pi\|_{L_{\mathbf{x}}^{\infty}}.$$

Hence, $\|\Pi\|_{L_{\mathbf{x}}^{\infty}}^{-1} \leq \|h(t)\|_{L_{\mathbf{x}}^1} \leq \|\Pi^{-1}\|_{L_{\mathbf{x}}^{\infty}}, \forall t \in [0, \infty)$. This further yields the bound

$$\frac{2}{3} \|\Pi\|_{L_{\mathbf{x}}^{\infty}}^{-1} \leq \|h_0\|_{L_{\mathbf{x}}^1} \exp \left\{ - \int_0^t Q(s) ds \right\} \leq 2 \|\Pi^{-1}\|_{L_{\mathbf{x}}^{\infty}}.$$

Rewriting this bound gives

$$- \int_0^t Q(s) ds \in \left[\log \left(\frac{2\|\Pi\|_{L_{\mathbf{x}}^{\infty}}^{-1}}{3\|\rho_0/\Pi\|_{L_{\mathbf{x}}^1}} \right), \log \left(\frac{2\|\Pi^{-1}\|_{L_{\mathbf{x}}^{\infty}}}{\|\rho_0/\Pi\|_{L_{\mathbf{x}}^1}} \right) \right].$$

Hence, we obtain the following estimate which concludes the proof of (2.6a),

$$\left| \int_0^t Q(s) ds \right| \leq \left| \log \left(\frac{2\|e^{-\mathbb{U}}\|_{L_{\mathbf{x}}^{\infty}}^{-1}}{3\|\rho_0 e^{\mathbb{U}}\|_{L_{\mathbf{x}}^1}} \right) \right| + \left| \log \left(\frac{2\|e^{\mathbb{U}}\|_{L_{\mathbf{x}}^{\infty}}}{\|\rho_0 e^{\mathbb{U}}\|_{L_{\mathbf{x}}^1}} \right) \right| \leq \mathfrak{B}_1(\mathbb{U}, \rho_0), \quad \forall t \in [0, T_*].$$

Step #4: The proof of conclusion (2.6b) proceeds in two steps.

Step #4a: Reformulation. We define the augmented quantity $H := h \exp \left\{ \int_0^t Q(s) ds \right\}$. Thanks to hypothesis (2.5a), H and h are comparable:

$$(2.10) \quad H e^{-2\mathfrak{B}_1} \leq H e^{-\int_0^t Q(s) ds} = h \leq H e^{2\mathfrak{B}_1}.$$

The equation (2.4) can be rephrased as follows

$$\partial_t H + \mathbb{V} \cdot \nabla_{\mathbf{x}} H = \nu \Delta_{\mathbf{x}} H - \nu \nabla_{\mathbf{x}} \mathbb{U} \cdot \nabla_{\mathbf{x}} H, \quad H \geq 0.$$

We can further decompose the solution,

$$\begin{aligned} \partial_t \bar{H} &= \frac{\nu}{|\mathbb{T}|^d} \int \tilde{H} \Delta_{\mathbf{x}} \mathbb{U} d\mathbf{x}, \quad \bar{H}(t=0) = \left(\frac{\rho_0}{\Pi} \right); \\ \partial_t \tilde{H} + \mathbb{V} \cdot \nabla_{\mathbf{x}} \tilde{H} &= \nu \Delta \tilde{H} - \nu \tilde{\mathbb{P}}_{\mathbf{x}} \left(\nabla_{\mathbf{x}} \tilde{\mathbb{U}} \cdot \nabla_{\mathbf{x}} \tilde{H} \right), \quad \tilde{H}(t=0) = \left(\frac{\rho_0}{\Pi} \right). \end{aligned}$$

Here, we recall the definition of $\tilde{\mathbb{P}}_{\mathbf{x}}$ (1.25).

Step #4b: Enhanced Dissipation Estimate. We recall the linear enhanced dissipation (1.22) and the definition of δ (1.23) to obtain that

$$\delta = \frac{\delta_{\text{ED}}}{2 + \log C_{\text{ED}}} < \delta_{\text{ED}} \rightsquigarrow C_{\text{ED}} \exp\{-\delta_{\text{ED}} \delta^{-1}\} = e^{-2}.$$

Now we consider a decomposition of the time horizon $[0, \infty)$ as

$$(2.11) \quad \mathbb{R}_+ = \bigcup_{i=0}^{\infty} [T_i, T_{i+1}), \quad T_i := i \delta^{-1} \nu^{-1/2}; \quad \bigcup_{i=0}^{i_*} [T_i, T_{i+1}) \subset [0, T_*) \subset \bigcup_{i=0}^{i_*+1} [T_i, T_{i+1}).$$

Let $\tilde{\eta}$ solve the reference equation

$$\begin{cases} \partial_t \tilde{\eta}^{(i)} + \mathbb{V} \cdot \nabla_{\mathbf{x}} \tilde{\eta}^{(i)} = \nu \Delta_{\mathbf{x}} \tilde{\eta}^{(i)}, \\ \tilde{\eta}^{(i)}(T_i, \mathbf{x}) = \tilde{H}(T_i, \mathbf{x}), \quad \forall t \in [T_i, T_{i+1}]. \end{cases}$$

We compute the time evolution of the squared L^2 -norm of the solution $\tilde{H}$. Since $\int \bar{H} \tilde{g} d\mathbf{x} \equiv 0$ for any function $g \in L^1$, we have

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|\tilde{H}\|_{L^2}^2 &= -\nu \int |\nabla_{\mathbf{x}} \tilde{H}|^2 d\mathbf{x} + \int \nabla_{\mathbf{x}} \cdot \mathbb{V} \left(\frac{\tilde{H}^2}{2} \right) d\mathbf{x} + \nu \int \Delta_{\mathbf{x}} \tilde{\mathbb{U}} \left(\frac{\tilde{H}^2}{2} \right) d\mathbf{x} \\ &= -\nu \|\nabla_{\mathbf{x}} \tilde{H}\|_{L_{\mathbf{x}}^2}^2 + \frac{\nu}{2} \int (\Delta_{\mathbf{x}} \mathbb{U}) \tilde{H}^2 d\mathbf{x} =: -\mathfrak{D} + \mathfrak{S}. \end{aligned}$$

Here, we have the diffusion contribution $\mathfrak{D}$ and the source term $\mathfrak{S}$. Note that the term $\mathfrak{S}$ can be bounded as follows

$$\mathfrak{S} \leq \frac{\nu}{2} \|(\Delta_{\mathbf{x}} \mathbb{U})_+\|_{L_{\mathbf{x}}^{\infty}} \|\tilde{H}\|_{L_{\mathbf{x}}^2}^2.$$

Hence, the squared L^2 -norm of $\tilde{H}$ satisfies the differential inequality,

$$\begin{aligned} \frac{d}{dt} \|\tilde{H}\|_{L_{\mathbf{x}}^2}^2 &\leq \nu \|(\Delta_{\mathbf{x}} \mathbb{U})_+\|_{L_{\mathbf{x}}^{\infty}} \|\tilde{H}\|_{L_{\mathbf{x}}^2}^2 \rightsquigarrow \\ \|\tilde{H}(T_i + \tau)\|_{L_{\mathbf{x}}^2}^2 &\leq \|\tilde{H}(T_i)\|_{L_{\mathbf{x}}^2}^2 \exp\{\nu \|(\Delta_{\mathbf{x}} \mathbb{U})_+\|_{L_{\mathbf{x}}^{\infty}} \tau\}, \quad \forall \tau \in [0, \delta^{-1} \nu^{-1/2}]. \end{aligned}$$

If we choose ν_0 such that

$$(2.12) \quad \nu_0^{1/2} \delta^{-1} \|\Delta_{\mathbf{x}} \mathbb{U}\|_{L_{\mathbf{x}}^{\infty}} \leq 2,$$

then for any $\nu \in (0, \nu_0]$, we obtain that

$$(2.13) \quad \|\tilde{H}(T_i + \tau)\|_{L_{\mathbf{x}}^2} \leq e \|\tilde{H}(T_i)\|_{L_{\mathbf{x}}^2}, \quad \forall \tau \in [0, \delta^{-1} \nu^{-1/2}].$$

Next, we derive the deviation estimate. Consider the time evolution of the difference $\tilde{H} - \tilde{\eta}^{(i)}$. Standard L^2 -energy estimate yields that

$$\begin{aligned}
 (2.14) \quad & \frac{1}{2} \frac{d}{dt} \|\tilde{H} - \tilde{\eta}^{(i)}\|_{L^2}^2 \\
 &= -\nu \|\nabla(\tilde{H} - \tilde{\eta}^{(i)})\|_{L^2}^2 + \nu \int \nabla_{\mathbf{x}}(\tilde{H} - \tilde{\eta}^{(i)}) \cdot \nabla_{\mathbf{x}} \tilde{\mathbb{U}} \tilde{H} d\mathbf{x} + \nu \int (\tilde{H} - \tilde{\eta}^{(i)}) \Delta_{\mathbf{x}} \tilde{\mathbb{U}} \tilde{H} d\mathbf{x} \\
 &=: -\mathfrak{D}_* + \mathfrak{S}_{*,1} + \mathfrak{S}_{*,2}.
 \end{aligned}$$

To estimate the $\mathfrak{S}_{*,1}$ -term in (2.14), we apply the Hölder inequality, Young's inequality and the estimate (2.13) to obtain the following

$$\mathfrak{S}_{*,1} \leq \nu \|\nabla_{\mathbf{x}}(\tilde{H} - \tilde{\eta}^{(i)})\|_{L_{\mathbf{x}}^2} \|\nabla_{\mathbf{x}} \tilde{\mathbb{U}}\|_{L_{\mathbf{x}}^\infty} \|\tilde{H}\|_{L_{\mathbf{x}}^2} \leq \frac{1}{4} \mathfrak{D}_* + e^2 \nu \|\nabla_{\mathbf{x}} \tilde{\mathbb{U}}\|_{L_{\mathbf{x}}^\infty}^2 \|\tilde{H}(T_i)\|_{L_{\mathbf{x}}^2}^2.$$

Similarly, we apply the Poincaré inequality $\|\tilde{H} - \tilde{\eta}^{(i)}\|_{L_{\mathbf{x}}^2} \leq \|\nabla_{\mathbf{x}}(\tilde{H} - \tilde{\eta}^{(i)})\|_{L_{\mathbf{x}}^2}$ and a similar argument to obtain that

$$\mathfrak{S}_{*,2} \leq \nu \|\tilde{H} - \tilde{\eta}^{(i)}\|_{L_{\mathbf{x}}^2} \|\Delta_{\mathbf{x}} \tilde{\mathbb{U}}\|_{L_{\mathbf{x}}^\infty} \|\tilde{H}\|_{L_{\mathbf{x}}^2} \leq \frac{1}{4} \mathfrak{D}_* + e^2 \nu \|\Delta_{\mathbf{x}} \tilde{\mathbb{U}}\|_{L_{\mathbf{x}}^\infty}^2 \|\tilde{H}(T_i)\|_{L_{\mathbf{x}}^2}^2.$$

Combining the estimates above yields that

$$\frac{1}{2} \frac{d}{dt} \|\tilde{H} - \tilde{\eta}^{(i)}\|_{L^2}^2 \leq -\frac{1}{2} \mathfrak{D}_* + e^2 \nu \left(\|\nabla_{\mathbf{x}} \tilde{\mathbb{U}}\|_{L_{\mathbf{x}}^\infty}^2 + \|\Delta_{\mathbf{x}} \tilde{\mathbb{U}}\|_{L_{\mathbf{x}}^\infty}^2 \right) \|\tilde{H}(T_i)\|_{L^2}^2.$$

Integrating in time yields

$$\|\tilde{H} - \tilde{\eta}^{(i)}\|_{L^2}(T_i + \tau) \leq \frac{\sqrt{2} e \nu^{1/4}}{\delta^{1/2}} \left(\|\nabla_{\mathbf{x}} \tilde{\mathbb{U}}\|_{L_{\mathbf{x}}^\infty}^2 + \|\Delta_{\mathbf{x}} \tilde{\mathbb{U}}\|_{L_{\mathbf{x}}^\infty}^2 \right)^{1/2} \|\tilde{H}(T_i)\|_{L^2}.$$

As long as ν_0 is chosen small enough, i.e.,

$$(2.15) \quad 0 < \nu \leq \nu_0(\rho_0, \mathbb{U}, \delta^{-1}) \leq \left(\frac{\delta}{8e^4(1 + \|\nabla_{\mathbf{x}} \tilde{\mathbb{U}}\|_{L_{\mathbf{x}}^\infty}^2 + \|\Delta_{\mathbf{x}} \tilde{\mathbb{U}}\|_{L_{\mathbf{x}}^\infty}^2)} \right)^2,$$

we have that $\|\tilde{H} - \tilde{\eta}^{(i)}\|_{L^2}(T_i + \tau) \leq \frac{1}{2e} \|\tilde{H}(T_i)\|_{L^2}$, $\forall \tau \in [0, \delta^{-1} \nu^{-1/2}]$. As a consequence of the above discussion, if $\nu \leq \nu_0$ is chosen small such that all conditions (2.9), (2.12) and (2.15) are fulfilled,

$$\|\tilde{H}(T_i + \tau)\|_{L^2}^2 \leq e^2 \|\tilde{H}(T_i)\|_{L^2}^2, \quad \|\tilde{H} - \tilde{\eta}^{(i)}\|_{L^2}(T_i + \tau) \leq \frac{1}{2e} \|\tilde{H}(T_i)\|_{L^2}, \quad \forall \tau \in [0, \delta^{-1} \nu^{-1/2}].$$

Thanks to the choice of δ (1.23), we have that $\|\tilde{\eta}^{(i)}(T_i + \delta^{-1} \nu^{-1/2})\|_{L^2} \leq \frac{1}{2e} \|\tilde{H}(T_i)\|_{L^2}$. Combining the estimates above, we have that

$$(2.16) \quad \|\tilde{H}(T_{i+1})\|_{L^2} \leq \|\tilde{H}(T_{i+1}) - \tilde{\eta}^{(i)}(T_{i+1})\|_{L^2} + \|\tilde{\eta}^{(i)}(T_{i+1})\|_{L^2} \leq \frac{1}{e} \|\tilde{H}(T_i)\|_{L^2}.$$

The remaining argument for deriving the enhanced dissipation estimate (2.6b) is standard. If $t \in \delta^{-1} \nu^{-1/2} \mathbb{N}$, then we have that by (2.16),

$$\|\tilde{H}(t)\|_{L^2} \leq \|\tilde{H}(0)\|_{L^2} \exp \left\{ -\frac{t}{\delta^{-1} \nu^{-1/2}} \right\} = \|\tilde{h}_0\|_{L^2} \exp \left\{ -\delta \nu^{1/2} t \right\}.$$

On the other hand, if $t \notin \delta^{-1} \nu^{-1/2} \mathbb{N}$, we choose the largest integer N such that $\delta^{-1} \nu^{-1/2} N \leq t$. Hence, we have the relation $t \in \delta^{-1} \nu^{-1/2} [N, N+1]$. Then, we combine the estimates (2.13) and (2.16) to obtain that

$$(2.17) \quad \|\tilde{H}(t)\|_{L^2} \leq e \|\tilde{H}(T_N)\|_{L^2} \leq e^2 \|\tilde{h}_0\|_{L^2} e^{-(N+1)} \leq e^2 \|\tilde{h}_0\|_{L^2} e^{-\delta \nu^{1/2} t}, \quad \forall t \in [0, T_*].$$

We recall the relation (2.10),

$$\|\tilde{h}(t)\|_{L^2} \leq e^{2\mathfrak{B}_1} \|\tilde{H}(t)\|_{L^2} \leq e^{2\mathfrak{B}_1+2} \|\tilde{h}_0\|_{L^2} e^{-\delta \nu^{1/2} t}, \quad \forall t \in [0, T_*].$$

This proves the nonlinear enhanced dissipation (2.6b) and concludes the proof of Theorem 1.1.

3. THE SECOND-ORDER ENHANCED DYNAMICS

In this section, we consider the second-order model (1.8) and prove Theorem 1.2. Part of the analysis is similar to that for the first-order model, and we provide a sketch of the argument in Section 3.1. In Section 3.2, we provide the proof of Theorem 1.2.

3.1. Preliminaries. We recall the model (1.8),

$$\begin{cases} \partial_t f + \mathbf{v} \cdot \nabla_{\mathbf{x}} f = \nu \Delta_{\mathbf{p}} f + \kappa \Delta_{\mathbf{x}} f + \kappa \nabla_{\mathbf{x}} \cdot (f \nabla_{\mathbf{x}} \mathbb{U}) + f \left(-\mathbf{v} \cdot \nabla_{\mathbf{x}} \mathbb{U} + \iint f \mathbf{v} \cdot \nabla_{\mathbf{x}} \mathbb{U} \, d\mathbf{x} d\mathbf{p} \right) \\ \mathbf{v}(\mathbf{p}) := \sum_{j=1}^d \sin(p_j) \mathbf{e}_{\mathbf{x},j}, \quad \Pi = Z^{-1} e^{-\mathbb{U}}, \quad f(t=0, \mathbf{x}, \mathbf{p}) = f_0(\mathbf{x}, \mathbf{p}), \quad \mathbf{x} \in \mathbb{T}^d, \mathbf{p} \in \mathbb{T}^d. \end{cases}$$

Assume the following initial normalization:

$$\|f_0\|_{L^1_{\mathbf{x},\mathbf{p}}} = 1, \quad f_0 \geq 0.$$

Standard arguments yield that the solution is locally well-posed, and we will focus on the a priori estimates in the forthcoming steps. The model preserves mass. Thanks to the parabolic maximum principle, $f \geq 0$, and we can compute the time derivative of the mass:

$$(3.1) \quad \frac{d}{dt} \|f\|_{L^1_{\mathbf{x},\mathbf{p}}} = \iint f \mathbf{v} \cdot \nabla_{\mathbf{x}} \log \Pi \, d\mathbf{x} d\mathbf{p} - \|f\|_{L^1_{\mathbf{x},\mathbf{p}}} \iint f \mathbf{v} \cdot \nabla_{\mathbf{x}} \log \Pi \, d\mathbf{x} d\mathbf{p},$$

Since $\|f_0\|_{L^1} = 1$, the scalar ODE (3.1) implies

$$(3.2) \quad \|f(t)\|_{L^1_{\mathbf{x},\mathbf{p}}} \equiv 1 = \|\Pi\|_{L^1_{\mathbf{x}}}.$$

Now we consider the equation for $h(t, \mathbf{x}, \mathbf{p}) := \frac{f(t, \mathbf{x}, \mathbf{p})}{\Pi(\mathbf{x})}$, which solves the following equation:

$$\begin{cases} \partial_t h + \mathbf{v} \cdot \nabla_{\mathbf{x}} h = \nu \Delta_{\mathbf{p}} h + \kappa \Delta_{\mathbf{x}} h - \kappa \nabla_{\mathbf{x}} \mathbb{U} \cdot \nabla_{\mathbf{x}} h - h \iint h \mathbf{v} \cdot \nabla_{\mathbf{x}} \Pi \, d\mathbf{x} d\mathbf{p}, \\ \mathbf{v}(\mathbf{p}) := \sum_{j=1}^d \sin(p_j) \mathbf{e}_{\mathbf{x},j}, \quad \Pi = Z^{-1} e^{-\mathbb{U}}, \quad h(t=0, \mathbf{x}, \mathbf{p}) = \frac{f_0(\mathbf{x}, \mathbf{p})}{\Pi(\mathbf{x})}. \end{cases}$$

As in the first-order case, we define the renormalization factor

$$G(t) := \iint h(t, \mathbf{x}, \mathbf{p}) \mathbf{v}(\mathbf{p}) \cdot \nabla_{\mathbf{x}} \Pi(\mathbf{x}) \, d\mathbf{x} d\mathbf{p},$$

and define the renormalized solution

$$(3.3) \quad H(t, \mathbf{x}, \mathbf{p}) := h(t, \mathbf{x}, \mathbf{p}) \exp \left\{ \int_0^t G(s) ds \right\}, \quad \forall t \geq 0.$$

The equation satisfied by the renormalized solution is as follows:

$$(3.4) \quad \begin{cases} \partial_t H + \mathbf{v} \cdot \nabla_{\mathbf{x}} H = \nu \Delta_{\mathbf{p}} H + \kappa \Delta_{\mathbf{x}} H - \kappa \nabla_{\mathbf{x}} \mathbb{U} \cdot \nabla_{\mathbf{x}} H, \\ H(t=0, \mathbf{x}, \mathbf{p}) = \frac{f_0(\mathbf{x}, \mathbf{p})}{\Pi(\mathbf{x})}. \end{cases}$$

To analyze this equation, we further decompose the solution H into the average $\langle H \rangle_{\mathbf{x}}$ and the fluctuation $\tilde{H} = H - \langle H \rangle_{\mathbf{x}}$:

$$(3.5) \quad \partial_t \langle H \rangle_{\mathbf{x}} = \nu \Delta_{\mathbf{p}} \langle H \rangle_{\mathbf{x}} + \kappa \langle \tilde{H} \Delta_{\mathbf{x}} \mathbb{U} \rangle_{\mathbf{x}}, \quad \langle H \rangle_{\mathbf{x}} \geq 0;$$

$$(3.6) \quad \partial_t \tilde{H} = \nu \Delta_{\mathbf{p}} \tilde{H} + \kappa \Delta_{\mathbf{x}} \tilde{H} - \kappa \tilde{\mathbb{P}}_{\mathbf{x}} (\nabla_{\mathbf{x}} \mathbb{U} \cdot \nabla_{\mathbf{x}} \tilde{H}) - \mathbf{v} \cdot \nabla_{\mathbf{x}} \tilde{H}.$$

Here, we recall the definition of $\tilde{\mathbb{P}}_{\mathbf{x}}$ in (1.25).

Finally, we fix the parameter δ in Theorem 1.2 as follows

$$(3.7) \quad \delta = \frac{\delta_0}{5} \quad (\text{Theorem 1.2}).$$

The constant δ_0 is defined in Theorem D.1 and is independent of the diffusion coefficient ν , the dimension d , and the solution.

The proof of the following nonlinear enhanced dissipation estimates is analogous to that of (1.6).

Theorem 3.1. *Under the assumptions of Theorem 1.2, the following claims hold:*

- *There exists a threshold $\nu_0(f_0, \delta^{-1}, \|\mathbb{U}\|_{C_x^2})$ such that for all $0 < \nu \leq \nu_0$, the nonlinear estimates hold:*

$$(3.8) \quad \left| \int_0^t G(s) ds \right| \leq 1 + \left| \log \left(\frac{2\|e^{-\mathbb{U}}\|_{L_x^\infty}^{-1}}{3\|f_0 e^{\mathbb{U}}\|_{L_{x,p}^1}} \right) \right| + \left| \log \left(\frac{2\|e^{\mathbb{U}}\|_{L_x^\infty}}{\|f_0 e^{\mathbb{U}}\|_{L_{x,p}^1}} \right) \right|,$$

$$(3.9) \quad \|\tilde{h}(t)\|_{L_{x,p}^2} \leq C_1^* \|\tilde{h}_0\|_{L_{x,p}^2} e^{-\delta\nu^{1/2}t},$$

$$C_1^* := \exp \left\{ 4 + 2 \left| \log \left(\frac{2\|e^{-\mathbb{U}}\|_{L_x^\infty}^{-1}}{3\|f_0 e^{\mathbb{U}}\|_{L_{x,p}^1}} \right) \right| + 2 \left| \log \left(\frac{2\|e^{\mathbb{U}}\|_{L_x^\infty}}{\|f_0 e^{\mathbb{U}}\|_{L_{x,p}^1}} \right) \right| \right\}, \quad \forall t \in [0, \infty).$$

- *Consider the passive scalar solution $\tilde{\zeta}$,*

$$(3.10) \quad \partial_t \tilde{\zeta} + \mathbf{v} \cdot \nabla_x \tilde{\zeta} = \nu \Delta_p \tilde{\zeta} + \kappa \Delta_x \tilde{\zeta}, \quad \tilde{\zeta}(0, \mathbf{x}, \mathbf{p}) = \tilde{H}(0, \mathbf{x}, \mathbf{p}).$$

Then the following estimate holds for all $\nu \leq 1$,

$$(3.11) \quad \|\tilde{H} - \tilde{\zeta}\|_{L_{x,p}^2}(t) \leq C(\delta^{-1}, \|\mathbb{U}\|_{C_x^2}) \nu^{\frac{1}{4}} \|\tilde{H}(0)\|_{L_{x,p}^2}, \quad \forall t \in [0, \delta^{-1}\nu^{-1/2}].$$

Proof. We apply a bootstrap argument and decompose the proof into three steps.

Step #1: Setup. Assume that $[0, T_*)$ is the maximal interval on $[0, \infty)$ such that the following two hypotheses are satisfied for all $t \in [0, T_*)$:

a) Control on $G(t)$:

$$(3.12a) \quad \left| \int_0^t G(s) ds \right| \leq 2 \left(1 + \left| \log \left(\frac{2\|e^{-\mathbb{U}}\|_{L_x^\infty}^{-1}}{3\|f_0 e^{\mathbb{U}}\|_{L_{x,p}^1}} \right) \right| + \left| \log \left(\frac{2\|e^{\mathbb{U}}\|_{L_x^\infty}}{\|f_0 e^{\mathbb{U}}\|_{L_{x,p}^1}} \right) \right| \right) =: 2\mathfrak{C}_1(f_0, \mathbb{U}).$$

b) Enhanced Dissipation:

$$(3.12b) \quad \|\tilde{h}(t)\|_{L^2} \leq 2\mathfrak{C}_2 \|\tilde{h}_0\|_{L^2} e^{-\delta\nu^{1/2}t}, \quad \mathfrak{C}_2 := e^{2\mathfrak{C}_1+2}.$$

Here, the parameter δ is chosen in (3.7). Moreover, the constants $\mathfrak{C}_1, \mathfrak{C}_2$ will only be used in this proof.

The *goal for the proof* is to show that on the same time interval, the estimates (3.12a) and (3.12b) can be improved provided the diffusion coefficient $0 < \nu < \nu_0(f_0, \mathbb{U}, \delta^{-1})$ is chosen small enough. The concrete improved conclusions are stated below:

$$(3.13a) \quad \left| \int_0^t G(s) ds \right| \leq \mathfrak{C}_1(f_0, \mathbb{U});$$

$$(3.13b) \quad \|\tilde{h}(t)\|_{L^2} \leq \mathfrak{C}_2(f_0, \mathbb{U}) \|\tilde{h}_0\|_{L^2} e^{-\delta\nu^{1/2}t}.$$

The proof of (3.13a) will be presented in Step # 2, and the proof of (3.13b) will be presented in Step # 3. Finally, we develop (3.11) in Step # 4.

Step #2: Proof of (3.13a). We can compute the time evolution of the function h for all $t \in [0, \infty)$:

$$\frac{d}{dt} \|h\|_{L_{x,p}^1} = \kappa \iint \tilde{h} \Delta_x \mathbb{U} d\mathbf{x} d\mathbf{p} - \|h\|_{L_{x,p}^1} G(t) \rightsquigarrow$$

$$\|h(t)\|_{L^1_{\mathbf{x},\mathbf{p}}} = \|h_0\|_{L^1_{\mathbf{x},\mathbf{p}}} e^{-\int_0^t G(s)ds} + \kappa \int_0^t \iint e^{-\int_s^t G(\tau)d\tau} \tilde{h}(s, \mathbf{x}, \mathbf{p}) \Delta_{\mathbf{x}} \mathbb{U}(\mathbf{x}) d\mathbf{x} d\mathbf{p} ds.$$

The structure of this equation mirrors that of (2.8). In particular, the second term only involves the $\tilde{h}$ remainder of the solution. Hence, thanks to the bootstrap hypothesis (3.12b) and the choice of $\kappa \leq \nu$, we have that the time integral of this term is small. The remaining argument of this step is similar to the proof of (2.6a), and hence we omit it for the sake of brevity.

Step #3: Proof of (3.13b). The main difference in the argument comes from the analysis of the enhanced dissipation phenomenon. Here, the linear enhanced dissipation effect (D.5) is weaker because it only acts in the $\mathbf{x}$ -direction.

First of all, we observe that the renormalized solution H and the original solution h are comparable under the hypothesis (3.12a):

$$(3.14) \quad H(t, \mathbf{x}, \mathbf{p}) e^{-2\mathfrak{C}_1} \leq H(t, \mathbf{x}, \mathbf{p}) e^{-\int_0^t G(s)ds} = h(t, \mathbf{x}, \mathbf{p}) \leq H(t, \mathbf{x}, \mathbf{p}) e^{2\mathfrak{C}_1}, \quad \forall t \in [0, T_*].$$

Hence, a suitable estimate on $\tilde{H}$ leads to an estimate of $\tilde{h}$ as stated in (3.13b). Therefore, we will mainly focus on $\tilde{H}$.

Now the *goal* is to show that the $\tilde{H}$ has enhanced dissipation of the form

$$(3.15) \quad \|\tilde{H}(t)\|_{L^2} \leq e^2 \|\tilde{h}(0)\|_{L^2} e^{-\delta \nu^{1/2} t}.$$

Here, the δ is chosen as in (3.7). As a result of the comparison (3.14), the estimate (3.15) implies the conclusion (3.13b).

To prove (3.15), we implement the same time-discretization as in (2.11). On each time interval $[T_i, T_{i+1}]$, $T_i = i\delta^{-1}\nu^{-1/2}$, let $\tilde{\eta}^{(i)}$ solve the reference equation

$$\begin{cases} \partial_t \tilde{\eta}^{(i)} + \mathbf{v} \cdot \nabla_{\mathbf{x}} \tilde{\eta}^{(i)} = \nu \Delta_{\mathbf{p}} \tilde{\eta}^{(i)} + \kappa \Delta_{\mathbf{x}} \tilde{\eta}^{(i)}, \\ \tilde{\eta}^{(i)}(T_i, \mathbf{x}, \mathbf{p}) = \tilde{H}(T_i, \mathbf{x}, \mathbf{p}), \quad \forall t \in [T_i, T_{i+1}]. \end{cases}$$

We carry out energy estimates similar to those used above,

$$(3.16) \quad \frac{1}{2} \frac{d}{dt} \|\tilde{H}\|_{L^2}^2 = -(\nu \|\nabla_{\mathbf{p}} \tilde{H}\|_{L^2}^2 + \kappa \|\nabla_{\mathbf{x}} \tilde{H}\|_{L^2}^2) - \kappa \iint \tilde{H} \nabla_{\mathbf{x}} \mathbb{U} \cdot \nabla_{\mathbf{x}} \tilde{H} d\mathbf{x} d\mathbf{p} =: -\mathcal{D} + \mathcal{S};$$

$$(3.17) \quad \begin{aligned} \frac{1}{2} \frac{d}{dt} \|\tilde{H} - \tilde{\eta}^{(i)}\|_{L^2}^2 &= -(\nu \|\nabla_{\mathbf{p}} (\tilde{H} - \tilde{\eta}^{(i)})\|_{L^2}^2 + \kappa \|\nabla_{\mathbf{x}} (\tilde{H} - \tilde{\eta}^{(i)})\|_{L^2}^2) \\ &\quad - \kappa \iint (\tilde{H} - \tilde{\eta}^{(i)}) \nabla_{\mathbf{x}} \mathbb{U} \cdot \nabla_{\mathbf{x}} \tilde{H} d\mathbf{x} d\mathbf{p} =: -\mathcal{D}_* + \mathcal{S}_*. \end{aligned}$$

Now we estimate the terms in (3.17)

$$\begin{aligned} |\mathcal{S}_*| &\leq \frac{1}{4} \kappa \|\nabla_{\mathbf{x}} (\tilde{H} - \tilde{\eta}^{(i)})\|_{L^2}^2 + C \kappa (\|\nabla_{\mathbf{x}} \mathbb{U}\|_{L^\infty}^2 + \|\Delta_{\mathbf{x}} \mathbb{U}\|_{L^\infty}^2) \|\tilde{H}\|_{L^2}^2 \\ &\leq \frac{1}{4} \kappa \|\nabla_{\mathbf{x}} (\tilde{H} - \tilde{\eta}^{(i)})\|_{L^2}^2 + C (\|\mathbb{U}\|_{C_x^2}) \nu \|\tilde{H}\|_{L^2}^2, \end{aligned}$$

Hence,

$$\frac{d}{dt} \|\tilde{H} - \tilde{\eta}^{(i)}\|_{L^2}^2 \leq -\frac{1}{2} (\nu \|\nabla_{\mathbf{p}} (\tilde{H} - \tilde{\eta}^{(i)})\|_{L^2}^2 + \kappa \|\nabla_{\mathbf{x}} (\tilde{H} - \tilde{\eta}^{(i)})\|_{L^2}^2) + C (\|\mathbb{U}\|_{C^2}) \nu \|\tilde{H}\|_{L^2}^2.$$

Through similar estimates on the equation (3.16), we have $\frac{d}{dt} \|\tilde{H}\|_{L^2}^2 \leq C (\|\mathbb{U}\|_{C^2}) \nu \|\tilde{H}\|_{L^2}^2$. As a consequence of the above differential inequalities, by the Grönwall inequality, for all $t \in [T_i, T_{i+1} = T_i + \delta^{-1}\nu^{-1/2}]$,

$$(3.18) \quad \begin{aligned} \|\tilde{H}(t)\|_{L^2}^2 &\leq e^{C(\|\mathbb{U}\|_{C^2})\delta^{-1}\nu^{\frac{1}{2}}} \|\tilde{H}(T_i)\|_{L^2}^2, \\ \|\tilde{H} - \tilde{\eta}^{(i)}\|_{L^2}^2(t) &\leq \nu^{1/2} C(\|\mathbb{U}\|_{C^2}) \delta^{-1} e^{C(\|\mathbb{U}\|_{C^2})\delta^{-1}\nu^{\frac{1}{2}}} \|\tilde{H}(T_i)\|_{L^2}^2. \end{aligned}$$

Hence, if $\nu_0 = \nu_0(\delta^{-1}, \|\mathbb{U}\|_{C^2})$ is chosen small enough,

$$\begin{aligned} \|\tilde{H}(t)\|_{L^2} &\leq e\|\tilde{H}(T_i)\|_{L^2}, \\ \|\tilde{H} - \tilde{\eta}^{(i)}\|_{L^2}(t) &\leq \frac{1}{4e}\|\tilde{H}(T_i)\|_{L^2}, \quad \forall t \in [T_i, T_i + \delta^{-1}\nu^{-1/2}]. \end{aligned}$$

Thanks to the choice of $\delta = \frac{\delta_0}{5}$, we have that by the linear enhanced dissipation (D.5),

$$\|\tilde{\eta}^{(i)}(T_i + \delta^{-1}\nu^{-1/2})\|_{L^2} \leq e^{\frac{1}{2}(1-\delta_0\nu^{\frac{1}{2}}5\delta_0^{-1}\nu^{-\frac{1}{2}})}\|\tilde{H}(T_i)\|_{L^2} \leq \frac{1}{e^2}\|\tilde{H}(T_i)\|_{L^2}.$$

Combining the estimates above, we have that

$$\|\tilde{H}(T_{i+1})\|_{L^2} \leq \|\tilde{H}(T_{i+1}) - \tilde{\eta}^{(i)}(T_{i+1})\|_{L^2} + \|\tilde{\eta}^{(i)}(T_{i+1})\|_{L^2} \leq \frac{1}{e}\|\tilde{H}(T_i)\|_{L^2}.$$

The remaining argument to derive the enhanced dissipation estimate (3.15) is identical to the argument in the first-order case (2.17). Hence, we omit the details.

Step #4: Proof of (3.11). The estimate is a consequence of (3.18). We choose $\tilde{\zeta} = \tilde{\eta}^{(0)}$, and observe that $\nu \leq 1$ is sufficient to derive (3.18), which implies (3.11). This concludes the proof. $\square$

3.2. Proof of Theorem 1.2. We divide the proof into steps.

Step #1: Convergence of the hydrodynamic density $[f]$. In this step, we establish the convergence (1.11): $\int_{\mathbb{T}^d} f(t, \mathbf{x}, \mathbf{p}) d\mathbf{p} \xrightarrow{t \rightarrow \infty} \Pi(\mathbf{x})$. To this end, the key is to establish the following convergence $|\mathbb{T}|^{-d}\|h\|_{L_{\mathbf{x}, \mathbf{p}}^1} \xrightarrow{t \rightarrow \infty} 1$. To see this convergence, we observe that by mass conservation (3.2) ($\|f(t)\|_{L_{\mathbf{x}, \mathbf{p}}^1} \equiv 1 = \|\Pi\|_{L_{\mathbf{x}}^1}$), we have

$$(3.19) \quad 0 = \iint f d\mathbf{x} d\mathbf{p} - \int \Pi d\mathbf{x} = \iint \left(h - \frac{1}{|\mathbb{T}|^d} \right) \Pi d\mathbf{x} d\mathbf{p} = \iint \left(\langle h \rangle_{\mathbf{x}} + \tilde{h} - \frac{1}{|\mathbb{T}|^d} \right) \Pi d\mathbf{x} d\mathbf{p}.$$

This, together with the enhanced dissipation (3.9), leads to the relation

$$\begin{aligned} (3.20) \quad & \left| |\mathbb{T}|^{-d}\|h\|_{L_{\mathbf{x}, \mathbf{p}}^1} - 1 \right| = \left| \|\langle h \rangle_{\mathbf{x}}\|_{L_{\mathbf{p}}^1} - 1 \right| = \left| \int \left(\langle h \rangle_{\mathbf{x}} - \frac{1}{|\mathbb{T}|^d} \right) d\mathbf{p} \right| = \left| \iint \tilde{h} \Pi d\mathbf{x} d\mathbf{p} \right| \\ & \leq |\mathbb{T}|^{d/2} \|\tilde{h}\|_{L_{\mathbf{x}, \mathbf{p}}^2} \|\Pi\|_{L_{\mathbf{x}}^2} \leq |\mathbb{T}|^{d/2} C_1^*(f_0, \mathbb{U}) \|\tilde{h}_0\|_{L_{\mathbf{x}, \mathbf{p}}^2} \|\Pi\|_{L_{\mathbf{x}}^2} e^{-\delta\nu^{1/2}t} \\ & = |\mathbb{T}|^{d/2} C_1^*(f_0, \mathbb{U}) \|\tilde{\mathbb{P}}_{\mathbf{x}}(f_0 e^{\mathbb{U}})\|_{L_{\mathbf{x}, \mathbf{p}}^2} \|e^{-\mathbb{U}}\|_{L_{\mathbf{x}}^2} e^{-\delta\nu^{1/2}t}. \end{aligned}$$

Here, we recall the definition of $\tilde{\mathbb{P}}_{\mathbf{x}}$ in (1.25). Now we make the decomposition:

$$\begin{aligned} & \left\| \int f(t, \mathbf{x}, \mathbf{p}) d\mathbf{p} - \Pi \right\|_{L_{\mathbf{x}}^1} = \left\| \int f(t, \mathbf{x}, \mathbf{p}) d\mathbf{p} - \frac{\int h(t, \mathbf{x}, \mathbf{p}) \Pi d\mathbf{p}}{|\mathbb{T}|^{-d}\|h\|_{L_{\mathbf{x}, \mathbf{p}}^1}} + \frac{\int h(t, \mathbf{x}, \mathbf{p}) \Pi d\mathbf{p}}{|\mathbb{T}|^{-d}\|h\|_{L_{\mathbf{x}, \mathbf{p}}^1}} - \Pi \right\|_{L_{\mathbf{x}}^1} \\ & \leq \left\| \int f(t, \mathbf{x}, \mathbf{p}) d\mathbf{p} - \frac{\int h(t, \mathbf{x}, \mathbf{p}) \Pi d\mathbf{p}}{|\mathbb{T}|^{-d}\|h\|_{L_{\mathbf{x}, \mathbf{p}}^1}} \right\|_{L_{\mathbf{x}}^1} + \left\| \frac{\int h(t, \mathbf{x}, \mathbf{p}) \Pi d\mathbf{p}}{\int \langle h \rangle_{\mathbf{x}} d\mathbf{p}} - \Pi \right\|_{L_{\mathbf{x}}^1} =: T_1 + T_2. \end{aligned}$$

Here we have used the fact that $h \geq 0$ and $\|\langle h \rangle_{\mathbf{x}}\|_{L_{\mathbf{p}}^1} = |\mathbb{T}|^{-d}\|h\|_{L_{\mathbf{x}, \mathbf{p}}^1}$ (recall that $\langle h \rangle_{\mathbf{x}} := |\mathbb{T}|^{-d} \int h(t, \mathbf{x}, \mathbf{p}) d\mathbf{p}$ (1.25)). We further note that $\int f(t, \mathbf{x}, \mathbf{p}) d\mathbf{p} = \int h(t, \mathbf{x}, \mathbf{p}) d\mathbf{p} \Pi$. Thanks to the estimate $\|h\|_{L_{\mathbf{x}, \mathbf{p}}^1} \geq \|f\|_{L_{\mathbf{x}, \mathbf{p}}^1} \|\Pi\|_{L_{\mathbf{x}}^\infty}^{-1} = \|\Pi\|_{L_{\mathbf{x}}^\infty}^{-1}$ and the relation (3.20), we have that,

$$T_1 = \left\| \int f d\mathbf{p} \left(1 - \frac{|\mathbb{T}|^d}{\|h\|_{L_{\mathbf{x}, \mathbf{p}}^1}} \right) \right\|_{L_{\mathbf{x}}^1} = \left| \frac{\|h\|_{L_{\mathbf{x}, \mathbf{p}}^1} - |\mathbb{T}|^d}{\|h\|_{L_{\mathbf{x}, \mathbf{p}}^1}} \right| \leq C(f_0, \mathbb{U}) \|\Pi\|_{L_{\mathbf{x}}^\infty} \|\tilde{\mathbb{P}}_{\mathbf{x}}(f_0 e^{\mathbb{U}})\|_{L_{\mathbf{x}, \mathbf{p}}^2} e^{-\delta\nu^{1/2}t}.$$

On the other hand, we invoke the enhanced dissipation to obtain that

$$\begin{aligned} T_2 &= \left\| \frac{\int \Pi(\mathbf{x}) \left(\langle h \rangle(t, \mathbf{p}) + \tilde{h}(t, \mathbf{x}, \mathbf{p}) \right) d\mathbf{p}}{\int \langle h \rangle(t, \mathbf{p}) d\mathbf{p}} - \Pi(\mathbf{x}) \right\|_{L^1_{\mathbf{x}}} \\ &\leq \frac{\|\Pi \tilde{h}(t)\|_{L^1_{\mathbf{x}, \mathbf{p}}}}{\|\langle h \rangle_{\mathbf{x}}(t)\|_{L^1_{\mathbf{x}, \mathbf{p}}}} \leq \frac{|\mathbb{T}|^{d/2} \|\Pi\|_{L^2_{\mathbf{x}}} \|\tilde{h}(t)\|_{L^2_{\mathbf{x}, \mathbf{p}}}}{\|h(t)\|_{L^1_{\mathbf{x}, \mathbf{p}}} |\mathbb{T}|^{-d}} \leq \|\Pi\|_{L^\infty_{\mathbf{x}}} |\mathbb{T}|^{3d/2} C(f_0, \mathbb{U}) \|e^{-\mathbb{U}}\|_{L^2} \|\tilde{\mathbb{P}}_{\mathbf{x}}(f_0 e^{\mathbb{U}})\|_{L^2} e^{-\delta \nu^{1/2} t}. \end{aligned}$$

Here, we recall the notation $\tilde{\mathbb{P}}_{\mathbf{x}}$ from (1.25). Hence, we have obtained the bound (1.11).

Step #2: Approximation. In this step, we discuss $L^\infty_{\mathbf{x}}$ -convergence of the dynamics to the target Π . To this end, we consider higher-regularity estimates for the solution h . We adopt the d -dimensional multi-index notation $D_{\mathbf{x}}^\alpha = \prod_{j=1}^d \partial_{x_j}^{\alpha_j}$, $|\alpha| = \sum_{j=1}^d \alpha_j$, $\alpha' \leq \alpha$ if and only if $\alpha'_j \leq \alpha_j$ for every j , and $\alpha' < \alpha$ if and only if $\alpha' \leq \alpha$ and $\alpha' \neq \alpha$. We then consider the renormalized solution H (3.4) and derive higher-regularity estimates.

First of all, we observe that the $\mathbf{x}$ -derivatives of the solution are straightforward. For any fixed $M \in \mathbb{N}$, we have that

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \sum_{|\alpha| \leq M} \|D_{\mathbf{x}}^\alpha H\|_{L^2}^2 &= -\kappa \sum_{|\alpha| \leq M} \|\nabla_{\mathbf{x}} D_{\mathbf{x}}^\alpha H\|_{L^2}^2 - \nu \sum_{|\alpha| \leq M} \|\nabla_{\mathbf{p}} D_{\mathbf{x}}^\alpha H\|_{L^2}^2 \\ &\quad - \sum_{|\alpha| \leq M} \sum_{\alpha' \leq \alpha} C_{\alpha, \alpha'} \kappa \iint \nabla_{\mathbf{x}} D_{\mathbf{x}}^{\alpha'} \mathbb{U} \cdot \nabla_{\mathbf{x}} D_{\mathbf{x}}^{\alpha - \alpha'} H D_{\mathbf{x}}^\alpha H d\mathbf{x} d\mathbf{p} \\ &\leq C_M \kappa \|\mathbb{U}\|_{W_{\mathbf{x}}^{M+1, \infty}}^2 \sum_{|\alpha| \leq M} \|D_{\mathbf{x}}^\alpha H\|_{L^2}^2. \end{aligned}$$

As a consequence, we have that $\|H(t)\|_{H_{\mathbf{x}}^M} \leq \|H(0)\|_{H_{\mathbf{x}}^M} e^{C_*(M) \kappa \|\mathbb{U}\|_{W_{\mathbf{x}}^{M+1, \infty}}^2 t}$.

Next, we observe that $f = h\Pi = H\Pi e^{\int_0^t G(s) ds}$ (3.3) and the integral factor is bounded by (3.8). As a consequence, $\|f(t)\|_{H_{\mathbf{x}}^M} \leq C(M, f_0, \mathbb{U}, Z) \exp\{C_*(M) \kappa \|\mathbb{U}\|_{W_{\mathbf{x}}^{M+1, \infty}}^2 t\}$, $\kappa \leq \nu$.

Since the L^1 -difference between $\int f(t, \mathbf{x}, \mathbf{p}) d\mathbf{p}$ and Π decays to zero at an enhanced rate $\mathcal{O}(\nu^{1/2})$, an application of the Nash interpolation inequality (with $M > d/2$) yields that the $\int f(t, \mathbf{x}, \mathbf{p}) d\mathbf{p}$ converges to Π in the space $L^\infty_{\mathbf{x}}$ as $t \rightarrow \infty$ (1.11),

$$\begin{aligned} \left\| \int_{\mathbb{T}^d} f d\mathbf{p} - \Pi \right\|_{L^\infty_{\mathbf{x}}} &\leq C(d, M) \left\| \int_{\mathbb{T}^d} f d\mathbf{p} - \Pi \right\|_{L^1_{\mathbf{x}}}^{\gamma(d, M)} \left\| \int_{\mathbb{T}^d} (f - |\mathbb{T}|^{-d} \Pi) d\mathbf{p} \right\|_{H_{\mathbf{x}}^M}^{1-\gamma(d, M)} \\ &\leq C(d, f_0, \mathbb{U}, Z) e^{-\delta \gamma \nu^{1/2} t} \sum_{|\alpha| \leq M} \left(\int \left(\int_{\mathbb{T}^d} D_{\mathbf{x}}^\alpha (f - |\mathbb{T}|^{-d} \Pi) d\mathbf{p} \right)^2 d\mathbf{x} \right)^{(1-\gamma)/2} \\ &\leq C(d, f_0, \mathbb{U}, Z) e^{-\delta \gamma \nu^{1/2} t} (\|f(t)\|_{H_{\mathbf{x}}^M L^2_{\mathbf{p}}} + \|\Pi\|_{H_{\mathbf{x}}^M})^{1-\gamma} \\ &\leq C(d, f_0, \mathbb{U}, Z) \exp \left\{ C_*(M) (1-\gamma) \nu \|\mathbb{U}\|_{W_{\mathbf{x}}^{M+1, \infty}}^2 t - \delta \gamma \nu^{1/2} t \right\} \\ &\leq C(d, f_0, \mathbb{U}, Z) e^{-\delta \gamma \nu^{1/2} t/2}. \end{aligned}$$

Here, we have chosen ν to be small enough in the last line. This completes the proof of the approximation estimate (1.12).

Step #3: Nonlinear Mixing. Now we focus on the initial time layer $t \in [0, \delta^{-1} \nu^{-1/2}]$ to derive the nonlinear mixing estimate (1.13). We start by proving the following.

Lemma 3.1. *Assume the conditions in Theorem 1.2. There exists a constant $C_{\text{ID}} = C_{\text{ID}}(\varsigma, f_0, \mathbb{U}, \delta^{-1}, d)$ such that the following estimate holds for all $F \in H_{\mathbf{x}, \mathbf{p}}^{1 \vee \{d/2 + \varsigma\}}$:*

$$(3.21) \quad \left| \iint \tilde{h}(t) F d\mathbf{x} d\mathbf{p} \right| \leq \frac{C_{\text{ID}}(\varsigma, f_0, \mathbb{U}, \delta^{-1}, d)}{t^{1/2}} \|\tilde{h}_0\|_{H_{\mathbf{x}, \mathbf{p}}^{1 \vee \{d/2 + \varsigma\}}} \|F\|_{H_{\mathbf{x}, \mathbf{p}}^{1 \vee \{d/2 + \varsigma\}}}, \quad \forall t \in (0, \delta^{-1} \nu^{-1/2}].$$

Proof. Throughout this proof, we define $\sigma := 1 \vee \{d/2 + \varsigma\}$. We organize the proof in three steps.
Step # 1: Setup and error estimate. Instead of considering the coupling between $\tilde{h}$ and F directly, we consider the coupling $\left| \iint \tilde{H}(t) F d\mathbf{x} d\mathbf{p} \right|$. Here $\tilde{H}$ is the solution to (3.6) and is the renormalized solution associated with $\tilde{h}$ through the formula (3.3). We further consider the passive scalar solution ζ (3.10) initialized with the same initial data $\tilde{H}_0$ on the time interval $[0, \delta^{-1} \nu^{-1/2}]$. Combining all the considerations above, we have that

$$(3.22) \quad \begin{aligned} \left| \iint \tilde{h}(t) F d\mathbf{x} d\mathbf{p} \right| &= e^{-\int_0^t G(s) ds} \left| \iint \tilde{H}(t) F d\mathbf{x} d\mathbf{p} \right| \\ &\leq e^{-\int_0^t G(s) ds} \left| \iint (\tilde{H}(t) - \zeta(t)) F d\mathbf{x} d\mathbf{p} \right| + e^{-\int_0^t G(s) ds} \left| \iint \zeta(t) F d\mathbf{x} d\mathbf{p} \right| =: T_1 + T_2. \end{aligned}$$

Thanks to (3.11) and the relation (3.3), we have that

$$\|\tilde{H} - \zeta\|_{L^2}(t) \leq C(\delta^{-1}, \mathbb{U}) \nu^{1/4} \|\tilde{H}_0\|_{L^2} = C(\delta^{-1}, \mathbb{U}) \nu^{1/4} \|\tilde{h}_0\|_{L^2}, \quad t \in [0, \delta^{-1} \nu^{-1/2}].$$

As a consequence, we have that the first term in (3.22) can be estimated as follows for all $t \in [0, \delta^{-1} \nu^{-1/2}]$,

$$(3.23) \quad T_1 = e^{-\int_0^t G(s) ds} \left| \iint (\tilde{H}(t) - \zeta(t)) F d\mathbf{x} d\mathbf{p} \right| \leq C(\delta^{-1}, \mathbb{U}) \nu^{1/4} e^{-\int_0^t G(s) ds} \|\tilde{h}_0\|_{L_{\mathbf{x}, \mathbf{p}}^2} \|F\|_{L_{\mathbf{x}, \mathbf{p}}^2}.$$

Step # 2: Mixing of the passive scalar. Next, we observe that the $\mathbf{x}$ -Fourier transform of ζ solves the following equation:

$$\begin{aligned} \partial_t \hat{\zeta}(t, \mathbf{k}, \mathbf{p}) + i\mathbf{v}(\mathbf{p}) \cdot \mathbf{k} \hat{\zeta}(t, \mathbf{k}, \mathbf{p}) &= \nu \Delta_{\mathbf{p}} \hat{\zeta}(t, \mathbf{k}, \mathbf{p}) - \kappa |\mathbf{k}|^2 \hat{\zeta}(t, \mathbf{k}, \mathbf{p}), \\ \hat{\zeta}(t=0, \mathbf{k}, \mathbf{p}) &= \hat{H}_0(\mathbf{k}, \mathbf{p}), \quad |\mathbf{k}| \neq 0. \end{aligned}$$

One can reformulate the problem as follows:

$$\begin{aligned} \partial_t \left(\hat{\zeta}(t, \mathbf{k}, \mathbf{p}) e^{\kappa |\mathbf{k}|^2 t} \right) + i\mathbf{v}(\mathbf{p}) \cdot \mathbf{k} \left(\hat{\zeta}(t, \mathbf{k}, \mathbf{p}) e^{\kappa |\mathbf{k}|^2 t} \right) &= \nu \Delta_{\mathbf{p}} \left(\hat{\zeta}(t, \mathbf{k}, \mathbf{p}) e^{\kappa |\mathbf{k}|^2 t} \right), \\ \hat{\zeta}(t=0, \mathbf{k}, \mathbf{p}) &= \hat{h}_0(\mathbf{k}, \mathbf{p}), \quad |\mathbf{k}| \neq 0. \end{aligned}$$

We observe that the equation is identical to the equation (D.2) with slightly adjusted notation ($\mathbf{p}$ is the same as $\mathbf{y}$ and $\mathbf{v}(\mathbf{p})$ is the same as $u_1(\mathbf{y})$). Hence, the linear inviscid damping estimate in Theorem D.2 and the relation (D.13) ensure the following for all $t \geq 0$,

$$(3.24) \quad \begin{aligned} \left| \iint \zeta(t) F d\mathbf{x} d\mathbf{p} \right| &= \left| \sum_{|\mathbf{k}| \neq 0} \int e^{\kappa |\mathbf{k}|^2 t} \hat{\zeta}_{\mathbf{k}}(t, \mathbf{p}) \overline{\hat{F}_{\mathbf{k}}(\mathbf{p})} d\mathbf{p} \right| \\ &\leq C_{d, \varsigma} \min \left\{ \frac{\nu^{1/4}}{\min\{1, \nu^{1/4} t^{1/2}\}}, e^{-\delta_0 \nu^{1/2} t} \right\} \|\tilde{h}_0\|_{H_{\mathbf{x}, \mathbf{p}}^{\sigma}} \|e^{\kappa t \Delta_{\mathbf{x}}} F\|_{L_t^{\infty} H_{\mathbf{x}, \mathbf{p}}^{\sigma}} \\ &\leq C_{d, \varsigma} \min \left\{ \frac{\nu^{1/4}}{\min\{1, \nu^{1/4} t^{1/2}\}}, e^{-\delta_0 \nu^{1/2} t} \right\} \|\tilde{h}_0\|_{H_{\mathbf{x}, \mathbf{p}}^{\sigma}} \|F\|_{H_{\mathbf{x}, \mathbf{p}}^{\sigma}}. \end{aligned}$$

Here, in the last line, we have used the fact that $e^{\kappa t \Delta_{\mathbf{x}}}$ is a contraction in $H_{\mathbf{x}, \mathbf{p}}^{\sigma}$.

Step # 3: Conclusion. Finally, we observe that for $t \leq \delta^{-1}\nu^{-1/2}$, the relation $\nu^{1/4} \leq C(\delta^{-1})t^{-1/2}$ holds. Combining the decomposition (3.22) with the estimates (3.23) and (3.24), we obtain, for all $t \in (0, \delta^{-1}\nu^{-1/2}]$:

$$\begin{aligned} \left| \iint \tilde{h}(t) F d\mathbf{x} d\mathbf{p} \right| &\leq C(\delta^{-1}, \mathbb{U}, d, \varsigma) e^{-\int_0^t G(s) ds} \left(\nu^{1/4} + \min \left\{ \frac{\nu^{1/4}}{\min\{1, \nu^{1/4} t^{1/2}\}}, e^{-\delta_0 \nu^{1/2} t} \right\} \right) \|\tilde{h}_0\|_{H_{\mathbf{x}, \mathbf{p}}^{\sigma}} \|F\|_{H_{\mathbf{x}, \mathbf{p}}^{\sigma}} \\ &\leq C(\delta^{-1}, \mathbb{U}, d, \varsigma) \frac{e^{-\int_0^t G(s) ds}}{t^{1/2}} \|\tilde{h}_0\|_{H^{\sigma}} \|F\|_{H_{\mathbf{x}, \mathbf{p}}^{\sigma}} \leq \frac{C(\delta^{-1}, f_0, \mathbb{U}, d, \varsigma)}{t^{1/2}} \|\tilde{h}_0\|_{H^{\sigma}} \|F\|_{H_{\mathbf{x}, \mathbf{p}}^{\sigma}}. \end{aligned}$$

Here, in the last line, we used the estimate (3.8). This concludes the proof of (3.21). $\square$

Proof of Theorem 1.2. Using the lemma above, we are finally ready to derive the last estimate (1.13) in Theorem 1.2. We assume that $\varphi \in H_{\mathbf{x}}^{\sigma}$ and recall $[f](t, \mathbf{x}) := \int_{\mathbb{T}^d} f(t, \mathbf{x}, \mathbf{p}) d\mathbf{p}$ (1.10), and consider the difference,

$$\mathcal{T} := \left| \int \varphi(\mathbf{x}) \Pi(\mathbf{x}) d\mathbf{x} - \int \varphi(\mathbf{x}) [f](t, \mathbf{x}) d\mathbf{x} \right| = \left| \int \varphi(\mathbf{x}) \Pi(\mathbf{x}) d\mathbf{x} - \iint f(t, \mathbf{x}, \mathbf{p}) \varphi(\mathbf{x}) d\mathbf{x} d\mathbf{p} \right|.$$

We observe that by the relations $h\Pi = f$ and $\|\langle h \rangle\|_{L_{\mathbf{p}}^1} = |\mathbb{T}|^{-d} \|h\|_{L_{\mathbf{x}, \mathbf{p}}^1}$, the following decomposition holds

$$\begin{aligned} (3.25) \quad \left| \iint f(t, \mathbf{x}, \mathbf{p}) \varphi(\mathbf{x}) d\mathbf{p} d\mathbf{x} - \int \varphi(\mathbf{x}) \Pi(\mathbf{x}) d\mathbf{x} \right| &\leq \left| \iint f(t, \mathbf{x}, \mathbf{p}) \varphi(\mathbf{x}) d\mathbf{p} d\mathbf{x} \left(1 - \frac{1}{\|\langle h \rangle\|_{L_{\mathbf{p}}^1}} \right) \right| \\ &\quad + \left| \int \frac{\int h(t, \mathbf{x}, \mathbf{p}) \Pi(\mathbf{x}) \varphi(\mathbf{x}) d\mathbf{p}}{\|\langle h \rangle\|_{L_{\mathbf{p}}^1}} - \Pi \varphi d\mathbf{x} \right| =: T_1 + T_2. \end{aligned}$$

We start by considering the term T_1 with a focus on estimating the factor $1 - \frac{1}{|\mathbb{T}|^{-d} \|h\|_{L^1}}$. This quantity can be estimated by a suitable combination of $\tilde{h}$ and Π . Hence, we first invoke the nonlinear mixing estimate (3.21) to derive the following bound

$$\begin{aligned} \left| \iint \tilde{h}(t, \mathbf{x}, \mathbf{p}) \Pi(\mathbf{x}) d\mathbf{p} d\mathbf{x} \right| &= \left| \iint \tilde{h}(t, \mathbf{x}, \mathbf{p}) (\Pi(\mathbf{x}) - \langle \Pi \rangle) d\mathbf{p} d\mathbf{x} \right| \\ &\leq \frac{C(\varsigma, f_0, \mathbb{U}, \delta^{-1}, d)}{t^{1/2}} \|\tilde{h}_0\|_{H_{\mathbf{x}, \mathbf{p}}^{\sigma}} \|\Pi - |\mathbb{T}|^{-d}\|_{H_{\mathbf{x}}^{\sigma}}. \end{aligned}$$

Here, we have used that $\langle \Pi \rangle_{\mathbf{x}}(\mathbf{p}) \equiv |\mathbb{T}|^{-d}$ and the fact that the function $\Pi - |\mathbb{T}|^{-d}$ does not depend on $\mathbf{p}$. Now the relations $\|\langle h \rangle_{\mathbf{x}}\|_{L_{\mathbf{p}}^1} = |\mathbb{T}|^{-d} \|h\|_{L_{\mathbf{x}, \mathbf{p}}^1}$ and $0 = \iint (\langle h \rangle_{\mathbf{x}} + \tilde{h} - \frac{1}{|\mathbb{T}|^d}) \Pi$ (3.19), together with the estimate above, yield the following:

$$\begin{aligned} ||\mathbb{T}|^{-d} \|h\|_{L_{\mathbf{x}, \mathbf{p}}^1} - 1| &= ||\langle h \rangle_{\mathbf{x}}\|_{L_{\mathbf{p}}^1} - 1| = \left| \int \left(\langle h \rangle_{\mathbf{x}} - \frac{1}{|\mathbb{T}|^d} \right) d\mathbf{p} \int \Pi(\mathbf{x}) d\mathbf{x} \right| \\ &= \left| \iint \tilde{h} \Pi d\mathbf{x} d\mathbf{p} \right| = \left| \iint \tilde{h} (\Pi - |\mathbb{T}|^{-d}) d\mathbf{x} d\mathbf{p} \right| \leq \frac{C(\varsigma, f_0, \mathbb{U}, \delta^{-1}, d)}{t^{1/2}} \|\tilde{h}_0\|_{H_{\mathbf{x}, \mathbf{p}}^{\sigma}} \|\Pi - |\mathbb{T}|^{-d}\|_{H_{\mathbf{x}, \mathbf{p}}^{\sigma}}. \end{aligned}$$

Hence, thanks to the estimate $\|h\|_{L_{\mathbf{x}, \mathbf{p}}^1} \geq \|\Pi\|_{L_{\mathbf{x}}^{\infty}}^{-1}$, the fact that $\|f(t)\|_{L^1} = 1$ and the Sobolev embedding,

$$\begin{aligned} T_1 &\leq \left| \iint f \varphi d\mathbf{x} d\mathbf{p} \right| \left| \frac{\|h\|_{L_{\mathbf{x}, \mathbf{p}}^1} - |\mathbb{T}|^d}{\|h\|_{L_{\mathbf{x}, \mathbf{p}}^1}} \right| \\ &\leq C(d) \|f\|_{L_{\mathbf{x}, \mathbf{p}}^1} \|\varphi\|_{L_{\mathbf{x}}^{\infty}} \|\Pi\|_{L_{\mathbf{x}}^{\infty}} \frac{C(\varsigma, f_0, \mathbb{U}, \delta^{-1}, d)}{t^{1/2}} \|\tilde{h}_0\|_{H_{\mathbf{x}, \mathbf{p}}^{\sigma}} \|\Pi - |\mathbb{T}|^{-d}\|_{H_{\mathbf{x}}^{\sigma}} \leq \frac{C(d, \varsigma, f_0, \mathbb{U}, \Pi)}{t^{1/2}} \|\varphi\|_{H_{\mathbf{x}}^{\sigma}}. \end{aligned}$$

For the T_2 -term in (3.25), we observe a cancellation of the $\langle h \rangle_{\mathbf{x}}$ -information,

$$\begin{aligned} T_2 &= \frac{1}{\|\langle h \rangle_{\mathbf{x}}\|_{L^1_{\mathbf{p}}}} \left| \iint \left(\langle h \rangle_{\mathbf{x}}(t, \mathbf{p}) + \tilde{h}(t, \mathbf{x}, \mathbf{p}) \right) \varphi(\mathbf{x}) \Pi(\mathbf{x}) d\mathbf{p} d\mathbf{x} - \int \langle h \rangle_{\mathbf{x}}(t, \mathbf{p}) d\mathbf{p} \int \varphi(\mathbf{x}) \Pi(\mathbf{x}) d\mathbf{x} \right| \\ &= \frac{1}{\|\langle h \rangle_{\mathbf{x}}\|_{L^1_{\mathbf{p}}}} \left| \iint \tilde{h}(t, \mathbf{x}, \mathbf{p}) \varphi(\mathbf{x}) \Pi(\mathbf{x}) d\mathbf{p} d\mathbf{x} \right|. \end{aligned}$$

Now we invoke duality, the lower bound $\|\langle h \rangle_{\mathbf{x}}\|_{L^1_{\mathbf{p}}} \geq |\mathbb{T}|^{-d} \|\Pi\|_{L^\infty_{\mathbf{x}}}^{-1}$ and the mixing estimate (3.21), to obtain the bound

$$T_2 \leq \frac{C(d, \varsigma, f_0, \mathbb{U})}{t^{1/2}} \|\Pi\|_{L^\infty_{\mathbf{x}}} \|\Pi\varphi - \langle \Pi\varphi \rangle_{\mathbf{x}}\|_{H^s_{\mathbf{x}}} \|\tilde{h}_0\|_{H^s_{\mathbf{x}, \mathbf{p}}} \leq \frac{C(d, \varsigma, f_0, \mathbb{U}, \Pi)}{t^{1/2}} \|\varphi\|_{H^s_{\mathbf{x}}}.$$

In the last line, we have used the product estimate for Sobolev functions and the observation that $\langle \Pi\varphi \rangle_{\mathbf{x}}$ is a constant. Hence, we obtain $\mathcal{T} \leq \frac{C(d, \varsigma, \Pi, \mathbb{U}, f_0)}{t^{1/2}} \|\varphi\|_{H^s_{\mathbf{x}}}$, $\forall t \in (0, \delta^{-1} \nu^{-1/2}]$. This concludes the proof of (1.13) and therefore the proof of Theorem 1.2. $\square$

4. NUMERICAL SIMULATIONS

4.1. Numerical Setup. In this section, we present a numerical experiment for the flow-accelerated Fokker–Planck equation (1.8) to illustrate the accelerated convergence from arbitrary initial data to the target invariant measure Π , as proved in Theorem 1.2. The design is based on an unconditionally stable finite-volume scheme.

For a cleaner presentation, we choose the Vicsek-type equation in (1.14) with $d = 2$ and $\kappa = \nu$:

$$(4.1) \quad \begin{cases} \partial_t f = \nu \Delta_\theta f + \nu \nabla_{\mathbf{x}} \cdot \left(f \nabla_{\mathbf{x}} \log \left(\frac{f}{\Pi} \right) \right) - \Pi \nabla_{\mathbf{x}} \cdot \left(\frac{\mathbf{p} f}{\Pi} \right) - f \iint f \mathbf{p} \cdot \nabla_{\mathbf{x}} \log \Pi \, d\mathbf{x} d\theta \\ f(t = 0, \mathbf{x}, \theta) = f_0(\mathbf{x}, \theta), \quad \mathbf{x} \in \mathbb{T}^2, \theta \in \mathbb{S}, \end{cases}$$

where $\mathbf{p}(\theta) = (\cos \theta, \sin \theta)^T$. All the enhanced estimates developed in the previous sections can be adapted to this setting. Similar to the renormalization procedure in (3.3), we observe that if $\tilde{f}$ solves (4.1), then $\tilde{f} = f \exp \left\{ \int_0^t \iint f \mathbf{p} \cdot \nabla_{\mathbf{x}} \log \Pi \, d\mathbf{x} d\theta ds \right\}$ solves the equation (after dropping $\tilde{\cdot}$)

$$(4.2) \quad \partial_t f = \nu \Delta_\theta f + \nu \nabla_{\mathbf{x}} \cdot \left(f \nabla_{\mathbf{x}} \log \left(\frac{f}{\Pi} \right) \right) - \Pi \nabla_{\mathbf{x}} \cdot \left(\frac{\mathbf{p} f}{\Pi} \right), \quad \mathbf{x} = (x_1, x_2) \in \mathbb{T}^2, \theta \in \mathbb{S}.$$

Thus, we only need to design schemes to compute (4.2); the renormalized solution $\frac{f(t, \mathbf{x}, \theta)}{\int f(t, \mathbf{x}, \theta) d\mathbf{x} d\theta}$ then solves the original equation (4.1). The corresponding h -equation (after dropping tilde) in terms of $h = \frac{f}{\Pi}$ is $\partial_t h = \nu \Delta_\theta h + \nu \nabla_{\mathbf{x}} (\log \Pi) \cdot \nabla_{\mathbf{x}} h - \left(\frac{\cos \theta}{\sin \theta} \right) \cdot \nabla_{\mathbf{x}} h$.

We first recast the Fokker–Planck equation (4.2) as follows

$$(4.3) \quad \begin{aligned} \partial_t f &= \nu \nabla_\theta \cdot \left(\Pi \nabla_\theta \frac{f}{\Pi} \right) + \nu \nabla_{\mathbf{x}} \cdot \left(\Pi \nabla_{\mathbf{x}} \frac{f}{\Pi} \right) - \Pi \nabla_{\mathbf{x}} \cdot \left(\mathbf{p} \frac{f}{\Pi} \right) \\ &= \nu \nabla \cdot \left(\Pi \nabla \frac{f}{\Pi} \right) - \Pi \nabla_{\mathbf{x}} \cdot \left(\mathbf{p} \frac{f}{\Pi} \right), \quad \mathbf{x} = (x_1, x_2) \in \mathbb{T}^2, \theta \in \mathbb{T}. \end{aligned}$$

Here $\mathbf{p} = (\cos \theta, \sin \theta)^\top$ so we have $\nabla_{\mathbf{x}} \cdot \mathbf{p} = 0$. Expressed in terms of the drift $\mathbf{b} = (\cos \theta, \sin \theta, 0)^\top$, the last term on the right of (4.3) reads $-\Pi \nabla_{\mathbf{x}} \cdot \left(\mathbf{p} \frac{f}{\Pi} \right) = -\Pi \nabla \cdot \left(\mathbf{b} \frac{f}{\Pi} \right)$.

To design a finite volume scheme based on a general Voronoi tessellation in $\mathbb{T}^d$, we let x_i , $i = 1, \dots, n$ denote the cell (or volume) centers and define the Voronoi cells

$$C_i := \{x \in \mathbb{T}^d; d(x, x_i) \leq d(x, x_j) \text{ for all } x_j\} \quad \text{with volume } |C_i| = \mathcal{H}^d(C_i).$$

Here $\mathcal{H}^d(C_i)$ is the d -dimensional Hausdorff measure of cell C_i . Then $\mathbb{T}^d = \bigcup_{i=1}^n C_i$ is a Voronoi tessellation of $\mathbb{T}^d$. Let $VF(i) := \{j; \Gamma_{ij} \neq \emptyset\}$ denote the non-empty faces adjacent to cell C_i ,

$$\Gamma_{ij} := C_i \cap C_j \text{ with its area } |\Gamma_{ij}| = \mathcal{H}^{d-1}(\Gamma_{ij}), \quad j = 1, \dots, n.$$

In each cell C_i , the discrete densities $f_i|C_i|$, $\Pi_i|C_i|$ are computed through flux interface, using the following finite-volume scheme. Let $\mathbf{n} \in \mathbb{R}^d$ denote the outward unit normal vector field on ∂C_i (pointing from i to its adjacent j), and let $(\mathbf{b} \cdot \mathbf{n})_{ij}^\pm \geq 0$ denote the positive and respectively negative parts of $\mathbf{b} \cdot \mathbf{n}$ at the intersection point of the geodesic from x_i to x_j . Integrating (4.3) over C_i ,

$$\frac{d}{dt} f_i|C_i| \approx \frac{d}{dt} \int_{C_i} f d\mathcal{H}^d \approx \sum_{j \in VF(i)} \int_{\Gamma_{ij}} \mathbf{n} \cdot \left(\nu \Pi \nabla \frac{f}{\Pi} \right) d\mathcal{H}^{d-1} + \Pi_i \sum_{j \in VF(i)} \int_{\Gamma_{ij}} \mathbf{n} \cdot \left(-\mathbf{b} \frac{f}{\Pi} \right) d\mathcal{H}^{d-1},$$

where we use the constant cell probabilities $f_i|C_i|$ and $\Pi_i|C_i|$ to approximate the exact solution on each cell C_i . We end up with the following finite volume scheme for cells $i = 1, \dots, n$,

$$(4.4) \quad \frac{d}{dt} f_i|C_i| = \sum_{j \in VF(i)} |\Gamma_{ij}| \left(\nu \frac{\Pi_i + \Pi_j}{2|x_j - x_i|} \left(\frac{f_j}{\Pi_j} - \frac{f_i}{\Pi_i} \right) + [(\mathbf{b} \cdot \mathbf{n})_{ij}^- \frac{f_j}{\Pi_j} - (\mathbf{b} \cdot \mathbf{n})_{ij}^+ \frac{f_i}{\Pi_i}] \Pi_i \right).$$

A detailed 3D time-space discretization and stability analysis of (4.4) are given in Appendix B.

4.2. Numerical Results. Based on the upwind finite-volume scheme (4.4), we compute the evolution of the dynamic density f from the prescribed initial data toward a smile-shaped target invariant measure.

Take a smile-shaped target invariant measure

$$(4.5) \quad \begin{aligned} \Pi(x, y) \propto & e^{-20 \left[\left(x - \frac{6}{5}\right)^2 + \left(y - \frac{6}{5}\right)^2 - \frac{1}{2} \right]^2 + \log(e^{-10(y-2)^2})} + e^{-20 \left[\left(x + \frac{6}{5}\right)^2 + \left(y - \frac{6}{5}\right)^2 - \frac{1}{2} \right]^2 + \log(e^{-10(y-2)^2})} \\ & + e^{-20(x^2 + y^2 - 2)^2 + \log(e^{-10(y+1)^2})} + 0.1; \end{aligned}$$

see the sixth subfigure in the bottom right corner of Fig. 2.

Take a Gaussian mixture

$$(4.6) \quad f_0(x, y) \propto e^{-16(x+3)^2 - 4y^2} + e^{-16(x-3)^2 - 4y^2} + e^{-4x^2 - 16(y+3)^2} + e^{-4x^2 - 16(y-3)^2} + 0.1$$

as an initial density $f_0(x, y, \theta) = f_0(x, y)$; see the first subfigure in the top left corner of Fig. 2.

The time evolution of the hydrodynamic density $\int f(t, x, y, \theta) d\theta$ is shown in Fig. 2 with the time step $\Delta t = 0.002$, the diffusion constant $\kappa = \nu = 0.25$ and the total computational time $T = 6$. The grid sizes are listed in detail in Appendix B. One can see the time-dependent density f converges rapidly to the invariant measure Π . In particular, the density f at $t = 0.2$ and $T = 6$ (shown at the bottom of Fig. 2) already looks very similar to the invariant measure Π .

The L^1 and relative L^2 errors are computed via $\| \int_0^{2\pi} f(t, x, y, \theta) d\theta - \Pi \|_{L^1_{x,y}}$ and $\left\| \frac{2\pi f}{\Pi} - 1 \right\|_{L^2_{x,y,\theta}}$, respectively. The decay of these errors with respect to time is shown in Fig. 3 with a red dashed line and a black dashed line, respectively.

For comparison, we also show the convergence rate of the dynamic solutions to the invariant measure for both the original linear Fokker-Planck equation in the gradient flow form and the flow-accelerated Fokker-Planck equation (4.1). For the original linear Fokker-Planck equation in the gradient flow form, i.e., the no-flow case $\mathbf{b} = 0$, the scheme (4.4) reduces to a classical finite volume scheme for

$$(4.7) \quad \partial_t f = \nu \nabla_{\mathbf{x}} \cdot \left(f \nabla_{\mathbf{x}} \log \left(\frac{f}{\Pi} \right) \right) = \nu \nabla_{\mathbf{x}} \cdot \left(\Pi \nabla_{\mathbf{x}} \frac{f}{\Pi} \right).$$

We refer to [26] for the convergence analysis of the finite volume scheme for the reduced Fokker-Planck equation (4.7) on a compact manifold. The decay rates of the L^1 and relative L^2 errors for

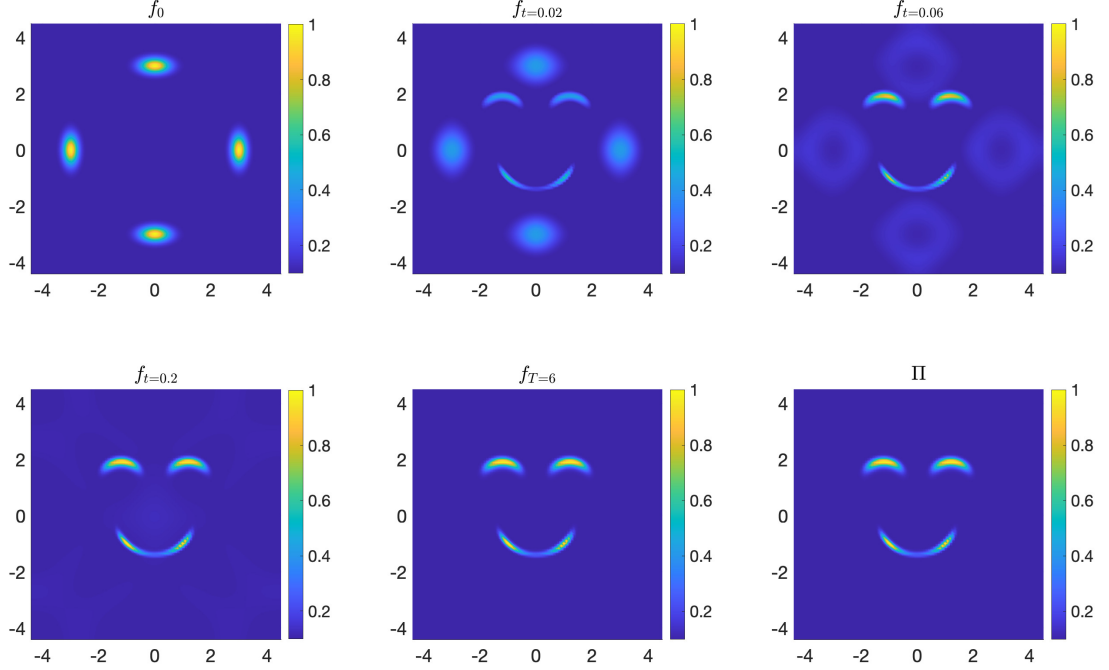


FIGURE 2. The time evolution of density f computed from scheme (4.4). The initial density f_0 is taken as (4.6) and the target invariant measure is taken as (4.5). Snapshots of f_t are shown at $t = 0.02, 0.05, 0.2, 6$.

the original linear Fokker-Planck equation in the gradient flow form are shown in Fig. 3 with a red solid line and a black solid line, respectively.

APPENDIX A. CONNECTIONS BETWEEN THE MODELS

Lemma A.1 (Normalization). *There is a one-to-one correspondence between the regular solution $\hat{\rho}$ to the simplified dynamics (1.7) and the mass-preserving dynamics (1.5). Specifically, given a solution $\hat{\rho}$ to (1.7), the mass-normalized solution $\hat{\rho}/\|\hat{\rho}\|_{L^1_{\mathbf{x}}}$ solves (1.5). Similarly, there is a one-to-one correspondence between the regular solution $\hat{f}$ to the simplified dynamics (1.15) and the mass-preserving dynamics (1.8).*

Proof. We focus on the first-order model because the argument for the second-order model is similar. As in the renormalization procedure in (3.3), denote $g(t) := \int \rho \nabla \cdot \nabla_{\mathbf{x}} \mathbb{U} \, d\mathbf{x} = \int \hat{\rho} \nabla_{\mathbf{x}} \cdot (\nabla \mathbb{U}) \, d\mathbf{x}$. One can see that if ρ solves the mass-preserving dynamics (1.5), then $\hat{\rho} = \rho \exp\{\int_0^t g(s) ds\}$ solves the following equation

$$(A.1) \quad \partial_t \hat{\rho} + \nabla \cdot \nabla_{\mathbf{x}} \hat{\rho} = \nu \Delta_{\mathbf{x}} \hat{\rho} + \nu \nabla_{\mathbf{x}} \cdot (\hat{\rho} \nabla_{\mathbf{x}} \mathbb{U}) - \hat{\rho} \nabla \cdot \nabla \mathbb{U}.$$

On the other hand, if $\hat{\rho} \geq 0$ solves the equation (A.1), its mass solves the equation: $\frac{d}{dt} \|\hat{\rho}\|_{L^1_{\mathbf{x}}} = - \int \hat{\rho} \nabla \cdot \nabla \mathbb{U} \, d\mathbf{x}$. Combining this relation and (A.1) yields (1.5). This establishes the relation between regular solutions to (1.5) and (A.1). $\square$

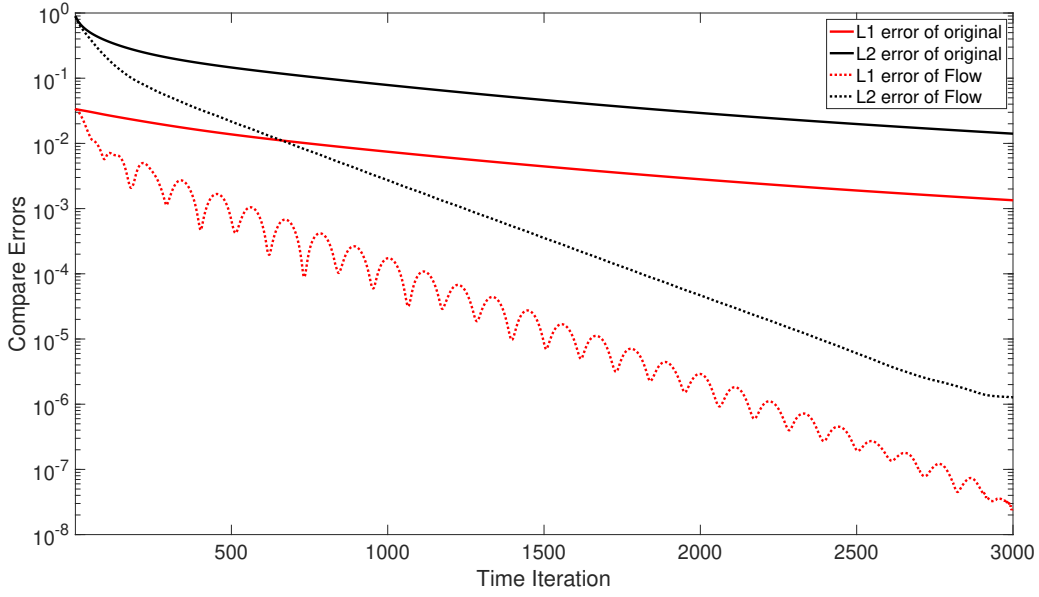


FIGURE 3. The decays of L^1 and relative L^2 errors between the dynamic density and the invariant measure are shown for both the flow-accelerated Fokker-Planck equation (4.1) and the original Fokker-Planck equation in the gradient flow form (4.7). The solid red and black lines represent the decay of L^1 and L^2 errors for (4.7). The dashed red and black lines represent the decay of L^1 and L^2 errors for (4.1) after flow acceleration.

APPENDIX B. SPACE-TIME DISCRETIZATION AND UNCONDITIONAL STABILITY

We outline the numerical scheme used for the computations in section 4.2, based in the 3D domain box $\Omega := [a, b] \times [c, d] \times [e, f]$ with $[a, b] = [-4.5, 4.5]$, $[c, d] = [-4.5, 4.5]$ and $[e, f] = [0, 2\pi]$. We use an equi-partition of Ω into small cells of size $\Delta x = \frac{b-a}{I}$, $\Delta y = \frac{d-c}{J}$ and $\Delta z = \frac{f-e}{K}$ with $I = 100$, $J = 100$, $K = 50$. Let (x_i, y_j, z_k) denote the cell centers

$$x_i = a + (i - \frac{1}{2})\Delta x, \quad y_j = c + (j - \frac{1}{2})\Delta y, \quad z_k = e + (k - \frac{1}{2})\Delta z, \quad 1 \leq i \leq I, \quad 1 \leq j \leq J, \quad 1 \leq k \leq K.$$

We use $f_\ell(t)$ to approximate the value of $f(t, x_i, y_j, z_k)$ and set, $g_\ell := \frac{f_\ell}{\pi_\ell}$ with $\pi_\ell = \pi(x_i, y_j, z_k)$, where ℓ stand for the multi-index (i, j, k) . The finite volume discretization (4.4) reads (we use the usual notation of $\mu^\pm$ and $\delta^\pm$ to denote forward and backward averaging and differencing¹, and abbreviate $c_k^\pm = \cos^\pm(z_k)$, $s_k^\pm = \sin^\pm(z_k)$)

$$\begin{aligned} \dot{f}_\ell = & \frac{1}{\Delta x^2} \left((\nu \mu_x^+ \pi_\ell + \Delta x c_k^- \pi_\ell) \delta_x^+ g_\ell - (\nu \mu_x^- \pi_\ell + \Delta x c_k^+ \pi_\ell) \delta_x^- g_\ell \right) \\ & + \frac{1}{\Delta y^2} \left((\nu \mu_y^+ \pi_\ell + \Delta y s_k^- \pi_\ell) \delta_y^+ g_\ell - (\nu \mu_y^- \pi_\ell + \Delta y s_k^+ \pi_\ell) \delta_y^- g_\ell \right) + \frac{1}{\Delta z^2} \left((\nu \mu_z^+ \pi_\ell) \delta_z^+ g_\ell - (\nu \mu_z^- \pi_\ell) \delta_z^- g_\ell \right). \end{aligned}$$

This is augmented with periodic boundary conditions for f_ℓ and π_ℓ . The expression on the right involves a seven point stencil: the value at the cell center, $g_{i,j,k}$, and its six neighboring values

¹Thus, for example, $\mu_y^+ f_{i,j,k} = \frac{1}{2}(f_{i,j+1,k} + f_{i,j,k})$ and $\delta_x^- f_{i,j,k} = f_{i,j,k} - f_{i-1,j,k}$.

$\{g_\ell : |\ell - (i, j, k)| = 1\}$; we use the abbreviated writing $f_\ell = RHS(g_\ell)$. We then use semi-implicit time discretization. Denoting $f_\ell^n = f(t^n, x_i, y_j, z_k)$ as the value of f at $t^n = n\Delta t$, we compute

$$(B.1) \quad \frac{f_{i,j,k}^{n+1} - f_{i,j,k}^n}{\Delta t} = RHS(g_{i,j,k}^{n+1}, g_{\ell:|\ell-(i,j,k)|=1}^n)$$

This enables us to recover an explicit form of f_ℓ^{n+1} . Define

$$\begin{aligned} \lambda_\ell := & \frac{1}{\Delta x^2} \left(\frac{\nu}{\pi_\ell} (\mu_x^+ \pi_\ell + \mu_x^- \pi_\ell) + \Delta x (c_k^+ + c_k^-) \right) \\ & + \frac{1}{\Delta y^2} \left(\frac{\nu}{\pi_\ell} (\mu_y^+ \pi_\ell + \mu_y^- \pi_\ell) + \Delta y (s_k^+ + s_k^-) \right) + \frac{1}{\Delta z^2} \left(\frac{\nu}{\pi_k} (\mu_z^+ \pi_\ell + \mu_z^- \pi_\ell) \right), \end{aligned}$$

then (B.1) reads

$$(B.2) \quad \begin{aligned} (1 + \Delta t \lambda_\ell) g_\ell^{n+1} = & g_\ell^n + \frac{\Delta t}{\Delta x^2} \left(\frac{\nu}{\pi_\ell} \mu_x^+ \pi_\ell + \Delta x c_k^- \right) g_{i+1,j,k}^n + \frac{\Delta t}{\Delta x^2} \left(\frac{\nu}{\pi_\ell} \mu_x^- \pi_\ell + \Delta x c_k^+ \right) g_{i-1,j,k}^n \\ & + \frac{\Delta t}{\Delta y^2} \left(\frac{\nu}{\pi_\ell} \mu_y^+ \pi_\ell + \Delta y s_k^- \right) g_{i,j+1,k}^n + \frac{\Delta t}{\Delta y^2} \left(\frac{\nu}{\pi_\ell} \mu_y^- \pi_\ell + \Delta y s_k^+ \right) g_{i,j-1,k}^n \\ & + \frac{\Delta t}{\Delta z^2} \left(\frac{\nu}{\pi_\ell} \mu_z^+ \pi_\ell \right) g_{i,j,k+1}^n + \frac{\Delta t}{\Delta z^2} \left(\frac{\nu}{\pi_\ell} \mu_z^- \pi_\ell \right) g_{i,j,k-1}^n. \end{aligned}$$

Observe that the coefficients of the seven g^n -terms corresponding to seven-point stencil on the right of (B.2) are non-negative and sum up to $1 + \Delta t \lambda_\ell$. Thus, $g_{i,j,k}^{n+1}$ is given as a convex combination of $g_{i,j,k}^n$ and its six neighboring values that participate in the stencil, $g_{\ell:|\ell-(i,j,k)|=1}^n$. In particular, we conclude that $\frac{f}{\pi}$ and likewise, $\frac{f}{\pi} - 1$, satisfy the maximum principle.

Lemma B.1. *If $\pi > 0$ and initial data $f_0 > 0$, then the scheme (B.2) satisfies*

- (1) *positivity-preserving property* $f_{i,j,k}^n > 0$ for any i, j, k, n ;
- (2) *unconditional maximum principle*: $\max_{i,j,k} \frac{f_{i,j,k}^{n+1}}{\pi_{i,j,k}} \leq \max_{i,j,k} \frac{f_{i,j,k}^n}{\pi_{i,j,k}}$;
- (3) ℓ^∞ *stability*: $\max_{i,j,k} \left| \frac{f_{i,j,k}^{n+1}}{\pi_{i,j,k}} - 1 \right| \leq \max_{i,j,k} \left| \frac{f_{i,j,k}^n}{\pi_{i,j,k}} - 1 \right|$.

APPENDIX C. THE PARTICLE SYSTEM OF THE SAMPLING DYNAMICS

In this section, we propose two particle models which resemble the dynamics of (1.5) and (1.8). As a consequence, the effective simulation of these particle dynamics also provides a fast sampling algorithm.

Motivated by [7, 52], we consider the empirical measure $\mu_N(t) := \frac{1}{N} \sum_{i=1}^N m_i(t) \delta_{\mathbf{x}_i(t)}$. Let $\mathbf{m}_N := (m_1, \dots, m_N)$, where $m_i(t=0) = 1$ with time-invariant total mass $\sum_{i=1}^N m_i = N$. Then the dynamics is

$$\begin{cases} d\mathbf{x}_i(t) = \mathbb{V}(t, \mathbf{x}_i(t))dt - \nu \nabla_{\mathbf{x}} \mathbb{U}(\mathbf{x}_i(t))dt + \sqrt{2\nu} dB_i(t), \\ dm_i(t) = \frac{m_i(t)}{N} \sum_{j=1}^N m_j(t) \left(\mathcal{A}(t, \mathbf{x}_j(t)) - \mathcal{A}(t, \mathbf{x}_i(t)) \right) dt, \\ \mathcal{A}(t, \mathbf{x}) := \mathbb{V}(t, \mathbf{x}) \cdot \nabla \mathbb{U}(\mathbf{x}), \\ \mathbf{x}_i(t=0) = \mathbf{x}_{i,0}, \quad m_i(t=0) = m_{i,0} = 1, \quad i \in \{1, \dots, N\}. \end{cases}$$

Next we consider the second-order sampling dynamics (1.8). Let the empirical measure be $\mu_N := \frac{1}{N} \sum_{i=1}^N m_i \delta_{(\mathbf{x}_i(t), \mathbf{p}_i(t))}$, Let $\mathbf{m}_N := (m_1, \dots, m_N)$, with time-invariant total mass $\sum_{i=1}^N m_i = N$. Then the dynamics is

$$\begin{cases} d\mathbf{x}_i(t) = \mathbf{v}(\mathbf{p}_i(t))dt - \kappa \nabla_{\mathbf{x}} \mathbb{U}(\mathbf{x}_i(t))dt + \sqrt{2\kappa} d\mathbb{B}_t^i, \\ d\mathbf{p}_i(t) = \sqrt{2\nu} d\mathbb{B}_t^i, \\ dm_i(t) = \frac{1}{N} \sum_{j=1}^N m_i(t)m_j(t) \left(\mathbf{v}(\mathbf{p}_j(t)) \cdot \nabla \mathbb{U}(\mathbf{x}_j(t)) - \mathbf{v}(\mathbf{p}_i(t)) \cdot \nabla \mathbb{U}(\mathbf{x}_i(t)) \right) dt, \\ \mathbf{x}_i(t=0) = \mathbf{x}_{i;0}, \quad \mathbf{p}_i(t=0) = \mathbf{p}_{i;0}, \quad m_i(t=0) = m_{i;0}. \end{cases}$$

Here, the Brownian motions $\{B_t^i\}_{i=1}^N, \{\mathbb{B}_t^i\}_{i=1}^N$ are independent and identically distributed.

APPENDIX D. ENHANCED DISSIPATION FOR ALTERNATING SHEAR FLOWS

D.1. General Setup. In this section, we establish the enhanced-dissipation result in Theorem 1.3. We rephrase the problem (1.21) on $\mathbb{T}^{2d}$ as follows

$$\partial_t \eta + \mathbf{u} \cdot \nabla_{\mathbf{x}, \mathbf{y}} \eta = \nu \Delta \eta, \quad \eta(t=0, \mathbf{x}, \mathbf{y}) = \eta_0(\mathbf{x}, \mathbf{y}).$$

Here, $(\mathbf{x}, \mathbf{y}) \in \mathbb{T}_{\mathbf{x}}^d \times \mathbb{T}_{\mathbf{y}}^d$ where $\mathbb{T}_{\mathbf{x}} = \mathbb{T}_{\mathbf{y}} = \mathbb{R}/2\pi\mathbb{Z}$. Without loss of generality, we assume that $\overline{\eta_0} = \frac{1}{|\mathbb{T}^{2d}|} \iint \eta_0 d\mathbf{x} d\mathbf{y} = 0$. Let $\{\mathbf{e}_j^{\mathbf{x}}\}_{j=1}^d$ be the canonical basis vectors for $\mathbb{T}_{\mathbf{x}}^d$ and $\{\mathbf{e}_j^{\mathbf{y}}\}_{j=1}^d$ be the canonical basis vectors for $\mathbb{T}_{\mathbf{y}}^d$. The time-dependent velocity field is defined as follows:

$$(D.1) \quad \mathbf{u}(t, \mathbf{x}, \mathbf{y}) = \Phi_1(t) \left[\sum_{j=1}^d \sin(y_j) \mathbf{e}_j^{\mathbf{x}} \right] + \Phi_2(t) \left[\sum_{j=1}^d \sin(x_j) \mathbf{e}_j^{\mathbf{y}} \right] = \Phi_1(t) u_1(\mathbf{y}) + \Phi_2(t) u_2(\mathbf{x}).$$

Here, Φ_1 and Φ_2 are two smooth temporal cutoff functions. The cutoff $\Phi_1(t)$ is supported in the open set $\cup_{k \in \mathbb{N}} \delta_*^{-1} \nu^{-\frac{1}{2}} (6k, 3+6k)$. It is identically one for $t \in \delta_*^{-1} \nu^{-\frac{1}{2}} [1+6k, 2+6k]$ and is smooth and monotone in $\delta_*^{-1} \nu^{-\frac{1}{2}} [6k, 1+6k]$ and $\delta_*^{-1} \nu^{-\frac{1}{2}} [2+6k, 3+6k]$. Similarly, the temporal cutoff $\Phi_2(t)$ is supported in the open set $\cup_{k \in \mathbb{N}} \delta_*^{-1} \nu^{-\frac{1}{2}} (3+6k, 6(k+1))$ and is equal to 1 on the interval $\delta_*^{-1} \nu^{-1/2} [4+6k, 5+6k]$. Here, the δ_* is a constant to be determined in Section D.5. The Fourier variables are

$$\mathbf{x} = (x_1, x_2, \dots, x_d) \implies \mathbf{k} = (k_1, k_2, \dots, k_d); \quad \mathbf{y} = (y_1, y_2, \dots, y_d) \implies \mathbf{l} = (\ell_1, \ell_2, \dots, \ell_d).$$

First of all, we develop the enhanced-dissipation estimate for the following system

$$\partial_t \eta + u_1(\mathbf{y}) \cdot \nabla_{\mathbf{x}} \eta = \nu \Delta_{\mathbf{x}, \mathbf{y}} \eta, \quad \eta(t=0, \mathbf{x}, \mathbf{y}) = \eta_0(\mathbf{x}, \mathbf{y}).$$

We take the Fourier transform in the $\mathbf{x}$ variables to obtain

$$\partial_t \hat{\eta}_{\mathbf{k}}(t, \mathbf{y}) + \underbrace{\left(\sum_{j=1}^d \sin(y_j) i k_j \right)}_{=: i u_1 \cdot \mathbf{k}} \hat{\eta}_{\mathbf{k}}(t, \mathbf{y}) = -\nu |\mathbf{k}|^2 \hat{\eta}_{\mathbf{k}}(t, \mathbf{y}) + \nu \Delta_{\mathbf{y}} \hat{\eta}_{\mathbf{k}}(t, \mathbf{y}).$$

Now, if we define $f_{\mathbf{k}}(t, \mathbf{y}) := \hat{\eta}_{\mathbf{k}}(t, \mathbf{y}) e^{\nu |\mathbf{k}|^2 t}$, we have the following equation

$$(D.2) \quad \partial_t f_{\mathbf{k}}(t, \mathbf{y}) + i u_1 \cdot \mathbf{k} f_{\mathbf{k}}(t, \mathbf{y}) = \nu \Delta_{\mathbf{y}} f_{\mathbf{k}}(t, \mathbf{y}).$$

Next, we would like to show that if $|\mathbf{k}| \neq 0$, the solution experiences the enhanced dissipation at a rate $\mathcal{O}(\nu^{\frac{1}{2}})$. As long as $|\mathbf{k}| \neq 0$, we have that $\max\{|k_j|\}_{j=1}^d \neq 0$ and we define

$$m \in \operatorname{argmax}_{1 \leq j \leq d} |k_j|, \quad k_m \neq 0, \quad \epsilon := \nu |k_m|^{-1}.$$

We consider the following Hypocoercivity functional:

$$(D.3) \quad \begin{aligned} \mathcal{G}[f_{\mathbf{k}}] := & \|f_{\mathbf{k}}\|_2^2 + \alpha\phi\epsilon^{1/2}\|\partial_{y_m}f_{\mathbf{k}}\|_2^2 \\ & + \beta\phi^2\Re\langle i\operatorname{sign}(k_m)\cos(y_m)f_{\mathbf{k}}, \partial_{y_m}f_{\mathbf{k}} \rangle + \gamma\phi^3\epsilon^{-1/2}\|\cos(y_m)f_{\mathbf{k}}\|_2^2. \end{aligned}$$

Here, we consider only the maximizing k_m -direction and define $\epsilon := \frac{\nu}{|k_m|}$ and $\phi := \min\{1, \sqrt{\nu}|k_m|^{1/2}t\}$.

Theorem D.1. *For a C^2 solution of (D.2), the Hypocoercivity functional has the following estimate for $0 < \nu \leq |k_m|$,*

$$(D.4) \quad \|f_{\mathbf{k}}(t)\|_{L^2}^2 \leq \mathcal{G}[f_{\mathbf{k}}](t) \leq e\|f_{0;\mathbf{k}}\|_{L^2}^2 e^{-2\delta_0\nu^{\frac{1}{2}}|k_m|^{\frac{1}{2}}t}, \quad \forall t \geq 0.$$

Here, the constant $\delta_0 \in (0, 1)$ is universal.

As a direct corollary of the above theorem, we have the following estimate.

Corollary D.1. *Consider the C^2 -solution to the hypoelliptic passive scalar equation with mean-zero initial data $\langle f_0 \rangle_{\mathbf{x}} \equiv 0$,*

$$\partial_t f + u_1(\mathbf{y}) \cdot \nabla_{\mathbf{x}} f = \nu \Delta_{\mathbf{y}} f, \quad f(t=0) = f_0.$$

Then, the following estimate holds for all $0 < \nu \leq 1$,

$$(D.5) \quad \|f(t)\|_{L^2_{\mathbf{x},\mathbf{y}}}^2 \leq e\|f_0\|_{L^2_{\mathbf{x},\mathbf{y}}}^2 e^{-2\delta_0\nu^{\frac{1}{2}}t}, \quad \forall t \geq 0.$$

Here, the constant $\delta_0 \in (0, 1)$ is universal.

Proof. Since the solution preserves the mean-zero constraint, we have that $f(t, \mathbf{x}, \mathbf{y}) = C \sum_{|\mathbf{k}| \neq 0} f_{\mathbf{k}}(t, \mathbf{y}) e^{i\mathbf{k} \cdot \mathbf{x}}$.

A combination of the estimate (D.4) and the Plancherel equality yields the result

$$\|f(t)\|_{L^2_{\mathbf{x},\mathbf{y}}}^2 = C_d \sum_{|\mathbf{k}| \neq 0} \|f_{\mathbf{k}}(t)\|_{L^2_{\mathbf{y}}}^2 \leq C_d e \sum_{|\mathbf{k}| \neq 0} \|f_{\mathbf{k}}(0)\|_{L^2_{\mathbf{y}}}^2 e^{-2\delta_0\nu^{\frac{1}{2}}t} = e\|f_0\|_{L^2_{\mathbf{x},\mathbf{y}}}^2 e^{-2\delta_0\nu^{\frac{1}{2}}t}. \quad \square$$

D.2. Preliminary Lemmas. The proof relies on a spectral inequality and a comparison lemma. We collect them in this section.

Lemma D.1. *There exists a constant $\mathfrak{C}_{\text{Spec}}$ such that the following estimate holds for every $j \in \{1, \dots, d\}$, any $k_j \neq 0$, and $\nu \leq |k_j|$:*

$$(D.6) \quad \left(\frac{\nu}{|k_j|} \right)^{1/2} \|f_{\mathbf{k}}\|_{L^2(\mathbb{T}^d)}^2 \leq \frac{\nu}{|k_j|} \|\partial_{y_j} f_{\mathbf{k}}\|_{L^2(\mathbb{T}^d)}^2 + \mathfrak{C}_{\text{Spec}} \|\cos(y_j) f_{\mathbf{k}}\|_{L^2(\mathbb{T}^d)}^2, \quad \mathfrak{C}_{\text{Spec}} = \frac{3\pi^2}{2} + \frac{1024}{\pi^2}.$$

Remark D.1. *This estimate is a consequence of the uncertainty principle in $\mathbb{R}^d$. The key point in the estimate is that the bound only depends on the regularity in the y_j -direction.*

Proof. The proof is standard, we refer the readers to [13] for details. $\square$

Lemma D.2. *Assume the relation $\beta^2 \leq \alpha\gamma$. Then, the following equivalence estimate concerning the functional $\mathcal{G}$ (D.3) holds*

$$\begin{aligned} \|f\|_2^2 + \frac{1}{2} \left(\alpha\phi\epsilon^{1/2}\|\partial_{y_m}f\|_2^2 + \gamma\phi^3\epsilon^{-1/2}\|\cos(y_m)f\|_2^2 \right) \\ \leq \mathcal{G}[f] \leq \|f\|_2^2 + \frac{3}{2} \left(\alpha\phi\epsilon^{1/2}\|\partial_{y_m}f\|_2^2 + \gamma\phi^3\epsilon^{-1/2}\|\cos(y_m)f\|_2^2 \right). \end{aligned}$$

Proof. The proof is standard, we refer the readers to [33, Lemma A.1] for details. $\square$

D.3. Energy Estimates. By taking the time derivative of the hypocoercivity functional, (D.3), we end up with the following decomposition:

$$\begin{aligned} \frac{d}{dt} \mathcal{G}[f_{\mathbf{k}}(t)] &= \frac{d}{dt} \|f_{\mathbf{k}}\|_2^2 + \alpha \epsilon^{1/2} \frac{d}{dt} (\phi \|\partial_{y_m} f_{\mathbf{k}}\|_2^2) + \beta \frac{d}{dt} (\phi^2 \Re \langle i \cos(y_m) f_{\mathbf{k}}, \partial_{y_m} f_{\mathbf{k}} \rangle) \\ &\quad + \gamma \epsilon^{-1/2} \frac{d}{dt} (\phi^3 \|\cos(y_m) f_{\mathbf{k}}\|_2^2) =: T_{L^2} + T_{\alpha} + T_{\beta} + T_{\gamma}. \end{aligned}$$

Standard energy estimates yield that

$$(D.7) \quad T_{L^2} = -2\nu \|\nabla_{\mathbf{y}} f_{\mathbf{k}}\|_{L^2}^2.$$

The estimates for the T_{α} , T_{β} , and T_{γ} terms are collected in the following technical lemma.

Lemma D.3. *For any constant $B > 0$, the following estimates hold:*

$$T_{\alpha} \leq \alpha \nu \mathbb{1}_{t \leq (\nu |k_m|)^{-1/2}} \|\partial_{y_m} f\|_2^2 - 2\alpha \phi \nu^{\frac{3}{2}} |k_m|^{-\frac{1}{2}} \|\nabla_{\mathbf{y}} \partial_{y_m} f\|_2^2 + \frac{\beta \phi^2}{B} |k_m| \|\cos(y_m) f_{\mathbf{k}}\|_{L^2}^2 + \frac{B \alpha^2}{\beta} \nu \|\partial_{y_m} f\|_{L^2}^2,$$

$$T_{\beta} \leq 2\beta \nu \|\nabla_{\mathbf{y}} f_{\mathbf{k}}\|_2^2 + \frac{7\alpha}{4} \phi \epsilon^{1/2} \nu \|\nabla_{\mathbf{y}} \partial_{y_m} f_{\mathbf{k}}\|_2^2 + \left(\frac{3\beta^2}{4\alpha\gamma} \right) \gamma \phi^3 \epsilon^{-1/2} \nu \|\cos(y_m) \nabla_{\mathbf{y}} f\|_2^2$$

$$+ \frac{\beta^2}{\alpha} \phi^3 \nu^{\frac{1}{2}} |k_m|^{\frac{1}{2}} \|f_{\mathbf{k}}\|_{L^2}^2 - \frac{3}{4} \beta |k_m| \phi^2 \|\cos(y_m) f_{\mathbf{k}}\|_{L^2}^2,$$

$$T_{\gamma} \leq 3\gamma \phi^2 \mathbb{1}_{t \leq (\nu |k_m|)^{-\frac{1}{2}}} |k_m| \|\cos(y_m) f_{\mathbf{k}}\|_2^2 + 4\gamma \phi^3 \nu^{\frac{1}{2}} |k_m|^{\frac{1}{2}} \|f_{\mathbf{k}}\|_2^2 - \gamma \phi^3 \nu^{\frac{1}{2}} |k_m|^{\frac{1}{2}} \|\cos(y_m) \nabla_{\mathbf{y}} f_{\mathbf{k}}\|_2^2.$$

Proof. The proof of these lemmas are similar to the arguments in [33, (A.11)–(A.15)]. Hence we omit them for the sake of brevity. $\square$

These estimates allow us to prove Theorem D.1.

Proof of Theorem D.1. If $T \leq 2\nu^{-1/2} |k_m|^{-1/2}$, then a standard L^2 -energy estimate yields (D.4). Hence, we assume $T > 2\nu^{-1/2} |k_m|^{-1/2}$ without loss of generality. We distinguish between two time intervals, i.e., $\mathcal{I}_1 = [0, \nu^{-1/2} |k_m|^{-1/2}]$, $\mathcal{I}_2 = [\nu^{-1/2} |k_m|^{-1/2}, T]$. We organize the proof into three steps. In step # 1, we choose the α, β, γ parameters and derive the energy dissipation relation. In step # 2, we estimate the functional $\mathcal{G}$ in the time interval $\mathcal{I}_1$. In step # 3, we estimate the functional $\mathcal{G}$ in the time interval $\mathcal{I}_2$ and conclude the proof.

Step # 1: Energy bounds. Combining the estimate (D.7) and Lemma D.3, we pick $B = 4$ and use the fact that $\phi^3 \leq \phi^2$ to obtain that

$$\begin{aligned} \frac{d}{dt} \mathcal{G}[f(t)] &\leq - \left(2 - \alpha - \frac{4\alpha^2}{\beta} - 2\beta \right) \nu \|\nabla_{\mathbf{y}} f_{\mathbf{k}}\|_2^2 - \frac{\alpha}{4} \phi \nu^{\frac{3}{2}} |k_m|^{-\frac{1}{2}} \|\nabla_{\mathbf{y}} \partial_{y_m} f\|_2^2 \\ &\quad - \left(\frac{1}{2} - 3\frac{\gamma}{\beta} \right) \beta \phi^2 |k_m| \|\cos(y_m) f_{\mathbf{k}}\|_{L^2}^2 + \left(\frac{\beta}{\alpha} + 4\frac{\gamma}{\beta} \right) \beta \nu^{1/2} \phi^3 |k_m|^{\frac{1}{2}} \|f_{\mathbf{k}}\|_2^2 \\ &\quad - \left(1 - \frac{3\beta^2}{4\alpha\gamma} \right) \gamma \phi^3 \nu^{\frac{1}{2}} |k_m|^{\frac{1}{2}} \|\cos(y_m) \nabla_{\mathbf{y}} f\|_2^2. \end{aligned}$$

We choose α, γ in terms of β as follows: $\alpha = \frac{\beta^{1/2}}{4}$, $\gamma = 4\beta^{3/2}$. Since we assume that $\nu \leq 1$, we can invoke the spectral inequality (D.6) to obtain that

$$\frac{d}{dt} \mathcal{G}[f(t)] \leq - \left(1 - \sqrt{\beta} - 2\beta - 20\beta^{\frac{3}{2}} \right) \nu \|\nabla_{\mathbf{y}} f\|_2^2 - \left(\frac{1}{2} - 12\sqrt{\beta} - \mathfrak{C}_{\text{Spec}} 20\sqrt{\beta} \right) \beta |k_m| \phi^2 \|\cos(y_m) f_{\mathbf{k}}\|_2^2$$

$$-\frac{\gamma}{4}\phi^3\nu^{1/2}|k_m|^{\frac{1}{2}}\|\cos(y_m)\nabla_{\mathbf{y}}f_{\mathbf{k}}\|_2^2.$$

Hence, we can choose $\beta = \beta(\mathfrak{C}_{\text{Spec}}) < 1$, sufficiently small that

$$\begin{aligned} \frac{d}{dt}\mathcal{G}[f(t)] &\leq -\frac{1}{2}\nu\|\nabla_{\mathbf{y}}f_{\mathbf{k}}\|_2^2 - \frac{\beta}{4}\phi^2|k_m|\|\cos(y_m)f_{\mathbf{k}}\|_2^2 \\ (D.8) \quad &\leq -\frac{\beta\phi^2}{4\mathfrak{C}_{\text{Spec}}}\nu^{1/2}|k_m|^{1/2}\|f_{\mathbf{k}}\|_2^2 - \frac{1}{4}\nu^{1/2}|k_m|^{1/2}\epsilon^{1/2}\phi\|\partial_{y_m}f_{\mathbf{k}}\|_2^2 - \frac{\beta}{8}\nu^{\frac{1}{2}}|k_m|^{\frac{1}{2}}\phi^2\epsilon^{-\frac{1}{2}}\|\cos(y_m)f_{\mathbf{k}}\|_2^2 \\ &\leq -2\delta_0(\beta, \mathfrak{C}_{\text{Spec}}^{-1})\nu^{1/2}|k_m|^{1/2}\phi^2\mathcal{G}[f], \quad 0 < \delta_0(\beta, \mathfrak{C}_{\text{Spec}}^{-1}) \leq \frac{1}{8}. \end{aligned}$$

Here, in the last line, we have invoked the spectral inequality (D.6) and the comparison lemma D.2. Finally, we observe that the parameter δ_0 is universal.

Step # 2: Initial time layer estimate. Thanks to the energy dissipation relation (D.8), we obtain that $\mathcal{G}[f_{\mathbf{k}}(t)] \leq \|f_{0;\mathbf{k}}\|_{L^2}^2 \leq \sqrt{e}\|f_{0;\mathbf{k}}\|_{L^2}^2 e^{-2\delta_0\nu^{1/2}|k_m|^{1/2}t}$, $\forall t \in [0, \nu^{-1/2}|k_m|^{-1/2}]$.

Step # 3: Long-time estimate. Assume $t \geq \nu^{-1/2}|k_m|^{-1/2}$. Thanks to the energy dissipation relation (D.8), we obtain $\frac{d}{dt}\mathcal{G}[f] \leq -2\delta_0\nu^{1/2}|k_m|^{1/2}\mathcal{G}[f]$. Hence, we obtain that

$$\mathcal{G}[f(t)] \leq \mathcal{G}[f(\nu^{-1/2}|k_m|^{-1/2})]e^{-2\delta_0\nu^{1/2}|k_m|^{1/2}(t-\nu^{-1/2}|k_m|^{-1/2})} \leq \sqrt{e}\mathcal{G}[f(\nu^{-1/2}|k_m|^{-1/2})]e^{-2\delta_0\nu^{1/2}|k_m|^{1/2}t}.$$

Now, the results from Steps 2 and 3 yield exponential decay with a universal coefficient. This concludes the proof of (D.4). $\square$

D.4. The Mixing Estimates. In this section, we derive the mixing estimates. The key is to establish bounds for the following vector fields:

$$\begin{aligned} \Gamma_m^+ f_{\mathbf{k}} &:= \partial_{y_m} f_{\mathbf{k}} + \mathcal{A}^+(t, \nu, k_m) \cos(y_m) i k_m f_{\mathbf{k}}, \quad \mathcal{A}^+(t, \nu, k_m) := \frac{\tanh((1-i)\sqrt{\nu|k_m|}t)}{(1-i)\sqrt{\nu|k_m|}}, \\ \Gamma_m^- f_{\mathbf{k}} &:= \partial_{y_m} f_{\mathbf{k}} + \mathcal{A}^-(t, \nu, k_m) \cos(y_m) i k_m f_{\mathbf{k}}, \quad \mathcal{A}^-(t, \nu, k_m) := \frac{\tanh((1+i)\sqrt{\nu|k_m|}t)}{(1+i)\sqrt{\nu|k_m|}}. \end{aligned}$$

These vector fields mimic the behavior of the vector field $\partial_y + t \cos(y_m) i k_m$, which commutes with the inviscid dynamics (D.2) $_{\nu=0}$. For the sake of simplicity, we further define the quantities

$$\begin{aligned} \mathcal{H}_+(t, \nu, k_m) &:= -2i\nu|k_m|\mathcal{A}^+(t, \nu, k_m) = (1-i)\sqrt{\nu|k_m|}\tanh((1-i)\sqrt{\nu|k_m|}t), \\ \mathcal{H}_-(t, \nu, k_m) &:= 2i\nu|k_m|\mathcal{A}^-(t, \nu, k_m) = (1+i)\sqrt{\nu|k_m|}\tanh((1+i)\sqrt{\nu|k_m|}t). \end{aligned}$$

Since in the sequel, we will mostly work with $\mathcal{H}_+$, we use the simplified notation $\mathcal{H} = \mathcal{H}_+$. Here, we observe that

$$(D.9) \quad |\mathcal{A}^\pm(t)| = \frac{|\mathcal{H}_\pm(t)|}{2\nu|k_m|} \approx \min\left\{t, \frac{1}{\nu^{\frac{1}{2}}|k_m|^{\frac{1}{2}}}\right\} = \nu^{-\frac{1}{2}}|k_m|^{-\frac{1}{2}}\phi(t).$$

Here, the $\phi(t)$ function is defined in (D.3). Moreover, we can check that

$$\Re \mathcal{H}_\pm(t) = \sqrt{\nu|k_m|} \frac{\sinh(2\sqrt{\nu|k_m|}t) - \sin(2\sqrt{\nu|k_m|}t)}{\cosh(2\sqrt{\nu|k_m|}t) + \cos(2\sqrt{\nu|k_m|}t)} \geq 0.$$

Computation yields that

$$(\partial_t + i u_1(\mathbf{y}) \cdot \mathbf{k} - \nu \Delta_{\mathbf{y}} + \mathcal{H}_\pm) \Gamma_m^\pm f_{\mathbf{k}} = \mathcal{H}_\pm \partial_{y_m} ((1 \mp \sin(y_m) \text{sign}(k_m)) f_{\mathbf{k}}) \pm \frac{1}{2} \text{sign}(k_m) \mathcal{H}_\pm \cos(y_m) f_{\mathbf{k}}.$$

Now we define two smooth cutoff functions,

$$\chi_+ := \chi^{(0)} + \chi^{(1)}, \quad \chi_- := \chi^{(0)} + \chi^{(2)}.$$

Thus, χ_+ (respectively, χ_-) vanishes in an $\mathcal{O}(1)$ -neighborhood of $-\pi/2$ (respectively, $\pi/2$).

First we have the following energy estimate for the weighted L^2 -norm $\|\chi_+\Gamma_m^+ f\|_{L^2}^2$. The proof is similar to the arguments in [33, Lemma A.3]. and is omitted for the sake of brevity.

Lemma D.4. *There exists a constant C such that, for $k_m > 0$, the following estimate holds:*

$$(D.10) \quad \|\chi_+\Gamma_m^+ f_{\mathbf{k}}(t)\|_{L^2}^2 + \|\chi_-\Gamma_m^- f_{\mathbf{k}}(t)\|_{L^2}^2 \leq C \|f_{0,\mathbf{k}}\|_{H_y^1}^2.$$

For $k_m < 0$, we have the corresponding estimate, $\|\chi_+\Gamma_m^- f_{\mathbf{k}}(t)\|_{L^2}^2 + \|\chi_-\Gamma_m^+ f_{\mathbf{k}}(t)\|_{L^2}^2 \leq C \|f_{0,\mathbf{k}}\|_{H_y^1}^2$.

With the vector-field estimates established, we are ready to present the following mixing theorem.

Theorem D.2 (Linear Mixing). *Consider (regular) solutions η to the following equation associated with (D.2),*

$$(D.11) \quad \begin{aligned} \partial_t \eta + u_1(\mathbf{y}) \cdot \nabla_{\mathbf{x}} \eta &= \nu \Delta_{\mathbf{y}} \eta, \quad \eta(t=0) = \eta_0, \quad \nu \in (0, 1], \\ u_1(\mathbf{y}) &= \sum_{j=1}^{\mathfrak{d}} \sin(y_j) \mathbf{e}_j^{\mathbf{x}}, \quad \langle \eta_0 \rangle_{\mathbf{x}} \equiv 0. \end{aligned}$$

Then the following mixing estimate holds for all $t > 0$ and all $F \in L_t^\infty H_{\mathbf{x},\mathbf{y}}^\sigma$, $\sigma := \max\{1, \mathfrak{d}/2 + \varsigma\}$:

$$(D.12) \quad \left| \iint \eta(t) \overline{F(t)} d\mathbf{x} d\mathbf{y} \right| \leq C_{\mathfrak{d},\varsigma} \min \left\{ \frac{\nu^{1/4}}{\min\{1, \nu^{1/4} t^{1/2}\}}, e^{-\delta_0 \nu^{1/2} t} \right\} \|\eta_0\|_{H_{\mathbf{x},\mathbf{y}}^\sigma} \|F\|_{L_t^\infty H_{\mathbf{x},\mathbf{y}}^\sigma}.$$

Here, $C_{\mathfrak{d},\varsigma} \geq 1$ depends only on $\mathfrak{d}$ and ς , $\varsigma > 0$ is an arbitrarily small positive number, $f \vee g := \max\{f, g\}$, and δ_0 is defined in Theorem D.1.

Proof. we begin by observing that the solutions to (D.2) and (D.11) are related by $\eta(t, \mathbf{x}, \mathbf{y}) = \sum_{|\mathbf{k}| \neq 0} \eta_{\mathbf{k}}(t, \mathbf{y}) e^{i\mathbf{k} \cdot \mathbf{x}}$. Hence, by the Plancherel equality, we have the following relation

$$(D.13) \quad \left| \iint \eta(t, \mathbf{x}, \mathbf{y}) \overline{F(t, \mathbf{x}, \mathbf{y})} d\mathbf{x} d\mathbf{y} \right| = C \left| \sum_{|\mathbf{k}| \neq 0} \int \eta_{\mathbf{k}}(t, \mathbf{y}) \overline{F_{\mathbf{k}}(t, \mathbf{y})} d\mathbf{y} \right|.$$

Therefore, it is sufficient to establish the estimate for the right-hand side. Furthermore, we note that for $t \leq 1$, the estimate (D.12) is a direct consequence of the dissipative nature of the L^2 -norm of η :

$$\left| \iint \eta(t) F d\mathbf{x} d\mathbf{y} \right| \leq \|\eta(t)\|_{L^2} \|F\|_{L_t^\infty L^2} \leq \|\eta_0\|_{L^2} \|F\|_{L_t^\infty L^2} \leq e^{\delta_0} e^{-\delta_0 \nu^{1/2} t} \|\eta_0\|_{L^2} \|F\|_{L_t^\infty H^\sigma}, \quad \forall t \leq 1.$$

Hence, without loss of generality, we assume that $t > 1$ in the remainder of the proof. We follow the idea of the proof of Proposition 1.7 in the paper [17]. For a general test function $F \in H_{\mathbf{x},\mathbf{p}}^\sigma(\mathbb{T}^{\mathfrak{d}})$, we decompose the right-hand side of (D.13) as

$$\sum_{|\mathbf{k}| \neq 0} \int \eta_{\mathbf{k}} \overline{F_{\mathbf{k}}} d\mathbf{y} = \sum_{|\mathbf{k}| \neq 0} \left(\int \eta_{\mathbf{k}} \overline{F_{\mathbf{k}}} \chi_+ d\mathbf{y} + \int \eta_{\mathbf{k}} \overline{F_{\mathbf{k}}} (1 - \chi_+) d\mathbf{y} \right) =: \sum_{|\mathbf{k}| \neq 0} (I_{\mathbf{k};+} + I_{\mathbf{k};-}).$$

As explained in the paper [17], a symmetry consideration yields that it is enough to consider the first part of the expression under the assumption that $k_m > 0$. For the second part, one can use the vector field Γ_m^- and associated cutoffs to derive similar estimates. The $k_m < 0$ case can be treated in a similar manner, and we omit it for the sake of brevity. One can introduce another cutoff function $\chi_\zeta(y_m)$ such that it is 1 in a ζ -neighborhood of $\pi/2$ (with $\zeta \in (0, 1)$ to be chosen later). Moreover, $\|\partial_{y_m} \chi_\zeta\|_{L_y^\infty} \leq C\zeta^{-1}$. With this cutoff, we further decompose the $I_{\mathbf{k};+}$ as follows

$$I_{\mathbf{k};+} = \int_{\mathbb{T}^{\mathfrak{d}}} \eta_{\mathbf{k}} \overline{F_{\mathbf{k}}} \chi_\zeta \chi_+ d\mathbf{y} + \int_{\mathbb{T}^{\mathfrak{d}}} \eta_{\mathbf{k}} \overline{F_{\mathbf{k}}} (1 - \chi_\zeta) \chi_+ d\mathbf{y} =: I_{\mathbf{k};+}^{(1)} + I_{\mathbf{k};+}^{(2)}.$$

For the $I_{\mathbf{k};+}^{(1)}$ -term, we estimate it using the length of the interval, the observation that $\|\eta_{\mathbf{k}}(t)\|_{L^\infty} \leq \|\eta_{\mathbf{k}}(0)\|_{L^\infty}$ and the Sobolev embedding for $\mathbb{T}^d$:

$$|I_{\mathbf{k};+}^{(1)}| \leq C\zeta \|F_{\mathbf{k}}(t)\|_{L_y^\infty} \|\eta_{\mathbf{k}}(0)\|_{L^\infty} \leq C_\zeta \zeta \|F_{\mathbf{k}}(t)\|_{H_y^\sigma} \|\eta_{\mathbf{k}}(0)\|_{H_y^\sigma}.$$

Next we estimate the $I_{\mathbf{k};+}^{(2)}$ term using the observation that $|\cos(y_m)| > 0$ on this interval and $(\nu|k_m|^{-1})^{\frac{1}{2}}(\Gamma_m^+ \eta_{\mathbf{k}} - \partial_{y_m} \eta_{\mathbf{k}}) = \frac{i-1}{2} \tanh((1-i)\sqrt{\nu|k_m|}t) \cos(y_m) \eta_{\mathbf{k}}$:

$$\begin{aligned} |I_{\mathbf{k};+}^{(2)}| &\leq \nu^{1/2} |k_m|^{-1/2} \left| \int \cos(y_m) \frac{\Gamma_m^+ \eta_{\mathbf{k}} - \partial_{y_m} \eta_{\mathbf{k}}}{\frac{i-1}{2} \tanh((1-i)\sqrt{\nu|k_m|}t)} \frac{\overline{F_{\mathbf{k}}}}{\cos^2(y_m)} (1 - \chi_\zeta) \chi_+ d\mathbf{y} \right| \\ &\leq \int \nu^{1/2} |k_m|^{-1/2} |\cos(y_m)| \left| \frac{\Gamma_m^+ \eta_{\mathbf{k}}}{\frac{i-1}{2} \tanh((1-i)\sqrt{\nu|k_m|}t)} \right| \left| \frac{\overline{F_{\mathbf{k}}}}{\cos^2(y_m)} \right| (1 - \chi_\zeta) \chi_+ d\mathbf{y} \\ &\quad + \left| \frac{\nu^{1/2} |k_m|^{-1/2}}{\frac{i-1}{2} \tanh((1-i)\sqrt{\nu|k_m|}t)} \right| \left| \int \cos(y_m) \partial_{y_m} \eta_{\mathbf{k}} \frac{\overline{F_{\mathbf{k}}}}{\cos^2(y_m)} (1 - \chi_\zeta) \chi_+ d\mathbf{y} \right| =: T_4 + T_5. \end{aligned}$$

For the T_4 -term, we estimate it with the Sobolev embedding, the equivalence $\phi(t) \approx |\tanh((1 \pm i)\sqrt{\nu|k_m|}t)|$, and (D.9) as follows:

$$\begin{aligned} T_4 &\leq \frac{C\nu^{1/2}}{|k_m|^{1/2}\phi(t)} \|F_{\mathbf{k}}(t)\|_{L_y^\infty} \int_{|y_m - \pi/2| \geq \zeta} \chi_+ \frac{|\Gamma_m^+ \eta_{\mathbf{k}}|}{|\cos(y_m)|} d\mathbf{y} \\ &\leq \frac{C\nu^{1/2} \|F_{\mathbf{k}}(t)\|_{H_y^\sigma} \|\chi_+ \Gamma_m^+ \eta_{\mathbf{k}}\|_{L^2} \zeta^{-\frac{1}{2}}}{|k_m|^{1/2} \min\{\nu^{1/2}|k_m|^{1/2}t, 1\}} \leq \frac{C\nu^{1/2} \|F_{\mathbf{k}}(t)\|_{H_y^\sigma} \|\chi_+ \Gamma_m^+ \eta_{\mathbf{k}}\|_{L^2} \zeta^{-\frac{1}{2}}}{|k_m|^{1/2} \min\{\nu^{1/2}t, 1\}}, \quad \forall t > 0. \end{aligned}$$

To estimate the second term T_5 , we recall the relation (D.9), integrate by parts, and estimate each resulting term as follows:

$$\begin{aligned} T_5 &\leq C \left| \frac{\nu^{1/2} |k_m|^{-1/2}}{\frac{i-1}{2} \tanh((1-i)\sqrt{\nu|k_m|}t)} \right| \left| \int \eta_{\mathbf{k}} \partial_{y_m} \left(\frac{\overline{F_{\mathbf{k}}}}{\cos(y_m)} (1 - \chi_\zeta) \chi_+ \right) d\mathbf{y} \right| \\ &\leq \frac{C\nu^{\frac{1}{2}} |k_m|^{-\frac{1}{2}}}{\min\{\nu^{1/2}t, 1\}} \|\eta_{\mathbf{k}}\|_{L^\infty} \left(\|F_{\mathbf{k}}(t)\|_{H_y^1} \left(\int \frac{(1 - \chi_\zeta)^2 \chi_+^2}{|\cos(y_m)|^2} d\mathbf{y} \right)^{1/2} + \|F_{\mathbf{k}}(t)\|_{L_y^\infty} \int \frac{|\sin(y_m)| (1 - \chi_\zeta) \chi_+}{|\cos(y_m)|^2} d\mathbf{y} \right. \\ &\quad \left. + \|F_{\mathbf{k}}(t)\|_{L_y^\infty} \int \frac{|\partial_{y_m} \chi_\zeta| \chi_+}{|\cos(y_m)|} d\mathbf{y} + \|F_{\mathbf{k}}(t)\|_{L_y^\infty} \int \frac{|\partial_{y_m} \chi_+| (1 - \chi_\zeta)}{|\cos(y_m)|} d\mathbf{y} \right) \\ &\leq \frac{C\nu^{\frac{1}{2}} |k_m|^{-\frac{1}{2}}}{\min\{\nu^{1/2}t, 1\}} \|\eta_{0;\mathbf{k}}\|_{L^\infty} \|F_{\mathbf{k}}(t)\|_{H_y^\sigma} \left(\zeta^{-\frac{1}{2}} + \zeta^{-1} + |\log \zeta| \right) \leq \frac{C\nu^{1/2} \|\eta_{0;\mathbf{k}}\|_{H^\sigma} \|F_{\mathbf{k}}(t)\|_{H_y^\sigma}}{|k_m|^{1/2} \min\{\nu^{1/2}t, 1\}} \frac{1}{\zeta}. \end{aligned}$$

Hence, we observe that if we set $\zeta = \left(\frac{\nu^{1/2}}{\min\{\nu^{1/2}t, 1\}} \right)^{1/2}$ and invoke the vector-field bound (D.10), the following estimate holds

$$|I_{\mathbf{k};+}| \leq \frac{C\nu^{1/4}}{\min\{\nu^{1/4}t^{1/2}, 1\}} \|\eta_{\mathbf{k}}(0)\|_{H^\sigma} \|F_{\mathbf{k}}(t)\|_{H_y^\sigma}, \quad \forall t \geq 0.$$

Summing all the $\mathbf{k}$ -components, we obtain that

$$(D.14) \quad \left| \sum_{|\mathbf{k}| \neq 0} \int \eta_{\mathbf{k}}(t, \mathbf{y}) \overline{F_{\mathbf{k}}(t, \mathbf{y})} d\mathbf{y} \right| \leq \sum_{|\mathbf{k}| \neq 0} \frac{C\nu^{1/4}}{\min\{\nu^{1/4}t^{1/2}, 1\}} \|\eta_{\mathbf{k}}(0)\|_{H_{\mathbf{y}}^g} \|F_{\mathbf{k}}(t)\|_{H_{\mathbf{y}}^g} \\ \leq \frac{C\nu^{1/4}}{\min\{\nu^{1/4}t^{1/2}, 1\}} \|\eta_0\|_{H_{\mathbf{x}, \mathbf{y}}^g} \|F(t)\|_{H_{\mathbf{x}, \mathbf{y}}^g} \leq \frac{C\nu^{1/4}}{\min\{\nu^{1/4}t^{1/2}, 1\}} \|\eta_0\|_{H_{\mathbf{x}, \mathbf{y}}^g} \|F\|_{L_t^\infty H_{\mathbf{x}, \mathbf{y}}^g}.$$

Moreover, thanks to the enhanced dissipation estimate in Theorem D.1, we have the following estimate for all times

$$(D.15) \quad \left| \sum_{|\mathbf{k}| \neq 0} \int \eta_{\mathbf{k}}(t, \mathbf{y}) \overline{F_{\mathbf{k}}(t, \mathbf{y})} d\mathbf{y} \right| \leq C \|\eta(0)\|_{L_{\mathbf{x}, \mathbf{y}}^2} e^{-\delta_0 \nu^{1/2} t} \|F\|_{L_t^\infty L_{\mathbf{x}, \mathbf{y}}^2}.$$

Combining the estimates (D.14) and (D.15), we obtain the result. $\square$

D.5. The Alternating Scheme. In this section, we consider the alternating shear flow scenario. We recall that $d = 2d$. Recall the smooth cutoff functions $\Phi_1(t), \Phi_2(t)$ and the alternating shear $\mathbf{u}$ (D.1). We set $\delta_* := 5^{-1}\delta_0$ where δ_0 is defined in (D.4) and (D.8). Without loss of generality, we assume that $\bar{\eta}_0 = 0$, which is a property that is preserved by the dynamics.

The goal is to show that at time $t = 2\mathcal{T} := 30\delta_0^{-1}\nu^{-\frac{1}{2}}$, the L^2 -norm $\eta - \bar{\eta}$ decays to e^{-1} times its original value. On the support of the first temporal cutoff $\Phi_1(t)$, there are two distinct regions: a) the effective region $t \in I_1^{(E)} := [1, 2] \times 5\delta_0^{-1}\nu^{-\frac{1}{2}}$; b) the warm-up/cool-down region $\text{supp } \Phi_1 \setminus I_1^{(E)}$. For $t \in \text{supp } \Phi_1(t)$, we observe that the streamline average $\langle \eta \rangle_{\mathbf{x}}(t, \mathbf{y}) := \frac{1}{|\mathbb{T}|^d} \int \eta(t, \mathbf{x}, \mathbf{y}) d\mathbf{x}$ solves the heat equation

$$\partial_t \langle \eta \rangle_{\mathbf{x}}(t, \mathbf{y}) + \Phi_1(t) \langle u_1(\mathbf{y}) \cdot \nabla_{\mathbf{x}} \eta \rangle_{\mathbf{x}} = \nu \Delta_{\mathbf{y}} \langle \eta \rangle_{\mathbf{x}}(t, \mathbf{y}) \implies \partial_t \langle \eta \rangle_{\mathbf{x}}(t, \mathbf{y}) = \nu \Delta_{\mathbf{y}} \langle \eta \rangle_{\mathbf{x}}(t, \mathbf{y}).$$

Hence,

$$(D.16) \quad \|\langle \eta \rangle_{\mathbf{x}}(t)\|_{L^2} \leq \|\langle \eta \rangle_{\mathbf{x}}(0)\|_{L^2}, \quad \forall t \in \text{supp } \Phi_1.$$

Next we consider the fluctuation $\eta - \langle \eta \rangle_{\mathbf{x}}$. The fluctuation solves the passive scalar equation, which dissipates the L^2 -norm. Moreover, on the effective region $I_1^{(E)}$, the fluctuation experiences strong enhanced dissipation (D.4). Hence, it is not surprising to derive the following estimate at the transition time $\mathcal{T}$: $\|\eta - \langle \eta \rangle_{\mathbf{x}}\|_{L^2}^2(\mathcal{T}) \leq \|\eta - \langle \eta \rangle_{\mathbf{x}}\|_{L^2}^2(10\delta_0^{-1}\nu^{-\frac{1}{2}}) \underbrace{\leq}_{\text{E.D.}} e^{-4} \|\eta - \langle \eta \rangle_{\mathbf{x}}\|_{L^2}^2(5\delta_0^{-1}\nu^{-\frac{1}{2}}) \leq$

$\frac{1}{e^4} \|\eta_0 - \langle \eta_0 \rangle_{\mathbf{x}}\|_{L^2}^2$. This concludes the argument in the first region.

For $t \in \text{supp } \Phi_2$, an argument similar to that in (D.16) shows that the L^2 -norm of the $\langle \eta \rangle_{\mathbf{y}}$ is non-expansive on the support of $\Phi_2(t)$. Moreover, we observe that at the transition time $\mathcal{T}$

$$\|\langle \eta \rangle_{\mathbf{y}}\|_{L_{\mathbf{x}, \mathbf{y}}^2}(\mathcal{T}) \leq \|\eta - \langle \eta \rangle_{\mathbf{x}}\|_{L_{\mathbf{x}, \mathbf{y}}^2}(\mathcal{T}) \leq \frac{1}{e^2} \|\eta_0 - \langle \eta_0 \rangle_{\mathbf{x}}\|_{L_{\mathbf{x}, \mathbf{y}}^2} \leq \frac{1}{e^2} \|\eta_0\|_{L_{\mathbf{x}, \mathbf{y}}^2}.$$

As a consequence, $\|\langle \eta \rangle_{\mathbf{y}}(t)\|_{L_{\mathbf{x}, \mathbf{y}}^2} \leq \frac{1}{e^2} \|\eta_0\|_{L_{\mathbf{x}, \mathbf{y}}^2}, \forall t \in \text{supp } \Phi_2$. Now by Theorem D.1, we have that at time $2\mathcal{T}$, $\|\eta - \langle \eta \rangle_{\mathbf{y}}\|_{L_{\mathbf{x}, \mathbf{y}}^2}(2\mathcal{T}) \leq \frac{1}{e^2} \|\eta - \langle \eta \rangle_{\mathbf{y}}\|_{L_{\mathbf{x}, \mathbf{y}}^2}(\mathcal{T}) \leq \frac{1}{e^2} \|\eta_0\|_{L_{\mathbf{x}, \mathbf{y}}^2}$. Hence, $\|\eta(2\mathcal{T})\|_{L_{\mathbf{x}, \mathbf{y}}^2} \leq \|\langle \eta \rangle_{\mathbf{y}}\|_{L_{\mathbf{x}, \mathbf{y}}^2}(2\mathcal{T}) + \|\eta - \langle \eta \rangle_{\mathbf{y}}\|_{L_{\mathbf{x}, \mathbf{y}}^2}(2\mathcal{T}) \leq \frac{2}{e^2} \|\eta_0 - \bar{\eta}_0\|_{L_{\mathbf{x}, \mathbf{y}}^2}$. By the same argument, we see that for $s, t \in 2\mathcal{T}\mathbb{N}$, the following estimate holds:

$$(D.17) \quad \|\eta(s+t) - \bar{\eta}(s+t)\|_{L_{\mathbf{x}, \mathbf{y}}^2} \leq \left(\frac{2}{e^2} \right)^{\frac{t}{2\mathcal{T}}} \|\eta(s) - \bar{\eta}(s)\|_{L_{\mathbf{x}, \mathbf{y}}^2}.$$

For general $s, t \geq 0$, we find the smallest integer N and largest integer M so that $(2\mathcal{T})N \geq s$, $(2\mathcal{T})M \leq s+t$, $M, N \in \mathbb{N}$. Note that if $t \leq 2\mathcal{T}$, then the estimate (1.22) is direct:

$$\|\eta(s+t) - \bar{\eta}(s+t)\|_2 \leq \|\eta(s) - \bar{\eta}(s)\|_2 \leq e^2 \|\eta(s) - \bar{\eta}(s)\|_2 e^{-\frac{1}{2\mathcal{T}}t}, \quad 0 \leq t \leq 2\mathcal{T}.$$

Hence we assume $t > 2\mathcal{T}$ and observe that $2\mathcal{T}(M - N) \geq t - 4\mathcal{T}$. Now we apply the estimate (D.17) with $s, t \in 2\mathcal{T}\mathbb{N}$, and the non-increasing nature of L^2 -norm of the solutions to derive that

$$\begin{aligned} \|\eta - \bar{\eta}\|_2(s + t) &\leq \|\eta - \bar{\eta}\|_2(M2\mathcal{T}) \leq \|\eta - \bar{\eta}\|_2(N2\mathcal{T}) \left(\frac{2}{e^2}\right)^{M-N} \leq \|\eta - \bar{\eta}\|_2(s) e^{-\frac{1}{2\mathcal{T}}(2\mathcal{T})(M-N)} \\ &\leq \|\eta - \bar{\eta}\|_2(s) e^{-\frac{1}{2\mathcal{T}}(t-4\mathcal{T})} = e^2 \|\eta - \bar{\eta}\|_2(s) e^{-\frac{1}{2\mathcal{T}}t}, \quad \forall s, t \geq 0. \end{aligned}$$

This concludes the proof of (1.22) in the general case.

APPENDIX E. MASS-SEARCHING DYNAMICS

Proof of Theorem 1.4. Consider the ratio $v := \omega e^{-\mathbb{W}}$, which solves the equation $\partial_t v + \mathbb{V} \cdot \nabla_{\mathbf{x}} v = \nu \Delta_{\mathbf{x}} v$. Direct computation yields that the spatial average of v satisfies the relation $\bar{v}(t) = \bar{v}(t = 0) = M$ for all $t \geq 0$. As a consequence of Theorem 1.3, we obtain that $\|\omega e^{-\mathbb{W}} - M\|_{L^2}(t) = \|v - \bar{v}\|_{L^2}(t) \leq C \|v(t = 0) - \bar{v}\|_{L^2} e^{-\delta_{\text{ED}} \nu^{1/2} t} \leq C \|\mathbb{T}\|^d e^{-\mathbb{W}} - M\|_{L^2} e^{-\delta_{\text{ED}} \nu^{1/2} t}$. Hence, the solution converges to the normalization constant M in L^2 at the enhanced exponential rate $\mathcal{O}(\nu^{1/2})$. $\square$

REFERENCES

- [1] D. Albritton, S. Armstrong, J.-C. Mourrat, and M. Novack. Variational methods for the kinetic fokker-planck equation. *Analysis & PDE*, 17(6):1953–2010, 2024.
- [2] D. Albritton, R. Beekie, and M. Novack. Enhanced dissipation and Hörmander’s hypoellipticity. *J. Funct. Anal.*, 283(3):Paper No. 109522, 38, 2022.
- [3] J. M. Altschuler and K. Talwar. Resolving the mixing time of the langevin algorithm to its stationary distribution for log-concave sampling. *arXiv preprint arXiv:2210.08448*, 2022.
- [4] F. Baudoin. Bakry-émery meet villani. *Journal of functional analysis*, 273(7):2275–2291, 2017.
- [5] J. Bedrossian, A. Blumenthal, and S. Punshon-Smith. Almost-sure enhanced dissipation and uniform-in-diffusivity exponential mixing for advection-diffusion by stochastic Navier-Stokes. *arXiv:1911.01561*.
- [6] J. Bedrossian and M. Coti Zelati. Enhanced dissipation, hypoellipticity, and anomalous small noise inviscid limits in shear flows. *Arch. Ration. Mech. Anal.*, 224(3):1161–1204, 2017.
- [7] I. Ben-Porat, J. A. Carrillo, and P.-E. Jabin. Singular flows with time-varying weights. *arXiv preprint arXiv:2503.02276*, 2025.
- [8] R. N. Bhattacharya. Criteria for recurrence and existence of invariant measures for multidimensional diffusions. *The Annals of Probability*, pages 541–553, 1978.
- [9] A. Blumenthal, M. Coti Zelati, and R. S. Gvalani. Exponential mixing for random dynamical systems and an example of Pierrehumbert. *Ann. Probab.*, 51(4):1559–1601, 2023.
- [10] Y. Cao, J. Lu, and L. Wang. On explicit L^2 -convergence rate estimate for underdamped Langevin dynamics. *Arch. Ration. Mech. Anal.*, 247(5):Paper No. 90, 34, 2023.
- [11] P. Cattiaux, A. Guillin, P. Monmarché, and C. Zhang. Entropic multipliers method for langevin diffusion and weighted log sobolev inequalities. *Journal of Functional Analysis*, 277(11):108288, 2019.
- [12] A. Christie, Y. Feng, G. Iyer, and A. Novikov. Speeding up langevin dynamics by mixing. *Multiscale Modeling & Simulation*, 23(4):1696–1743, 2025.
- [13] D. Coble and S. He. A note on enhanced dissipation and taylor dispersion of time-dependent shear flows. *arXiv:2309.15738*, 2023.
- [14] P. Constantin, A. Kiselev, L. Ryzhik, and A. Zlatoš. Diffusion and mixing in fluid flow. *Ann. of Math. (2)*, 168(2):643–674, 2008.
- [15] M. Coti Zelati, G. Crippa, G. Iyer, and A. L. Mazzucato. Mixing in incompressible flows: transport, dissipation, and their interplay. *Notice of AMS*, 71(5), 2024.
- [16] M. Coti Zelati, M. G. Delgadino, and T. M. Elgindi. On the relation between enhanced dissipation timescales and mixing rates. *Comm. Pure Appl. Math.*, 73(6):1205–1244, 2020.
- [17] M. Coti Zelati, H. Dietert, and D. Gérard-Varet. Orientation mixing in active suspensions. *Ann. PDE*, 9(20), 2023.
- [18] M. Coti-Zelati and T. D. Drivas. A stochastic approach to enhanced diffusion. *arXiv:1911.09995v1*, 2019.
- [19] F. Cucker and S. Smale. On the mathematics of emergence. *Japanese Journal of Mathematics*, 2(1):197–227, 2007.
- [20] J. Dolbeault, C. Mouhot, and C. Schmeiser. Hypocoercivity for linear kinetic equations conserving mass. *Transactions of the American Mathematical Society*, 367(6):3807–3828, 2015.

- [21] T. M. Elgindi, K. Liss, and J. C. Mattingly. Optimal enhanced dissipation and mixing for a time-periodic, lipschitz velocity field on $\mathbb{T}^2$. *arXiv preprint arXiv:2304.05374*, 2023.
- [22] Y. Feng and G. Iyer. Dissipation enhancement by mixing. *Nonlinearity*, 32(5):1810–1851, 2019.
- [23] Y. Feng, A. L. Mazzucato, and C. Nobili. Enhanced dissipation by circularly symmetric and parallel pipe flows. *Physica D: Nonlinear Phenomena*, 445:133640, 2023.
- [24] A. Frouvelle and J.-G. Liu. Dynamics in a kinetic model of oriented particles with phase transition. *SIAM Journal on Mathematical Analysis*, 44(2):791–826, 2012.
- [25] T. Gallay and M. C. Zelati. Enhanced dissipation and taylor dispersion in higher-dimensional parallel shear flows. *arXiv preprint arXiv:2108.11192*, 2021.
- [26] Y. Gao, J.-G. Liu, and N. Wu. Data-driven efficient solvers for langevin dynamics on manifold in high dimensions. *Applied and Computational Harmonic Analysis*, 62:261–309, 2023.
- [27] V. Gardner, K. L. Liss, and J. C. Mattingly. A pathwise approach to the enhanced dissipation of passive scalars advected by shear flows. *arXiv preprint arXiv:2410.05657*, 2024.
- [28] A. Gelman and X.-L. Meng. Simulating normalizing constants: From importance sampling to bridge sampling to path sampling. *Statistical Science*, 13(2):163–185, 1998.
- [29] M. Girolami and B. Calderhead. Riemann manifold langevin and hamiltonian monte carlo methods. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 73(2):123–214, 2011.
- [30] C. Griffin, S. Rajtmajer, A. Squicciarini, and A. Belmonte. Consensus and information cascades in game-theoretic imitation dynamics with static and dynamic network topologies. *SIAM Journal on Applied Dynamical Systems*, 18(2):597–628, 2019.
- [31] P. Grindrod. Models of individual aggregation or clustering in single and multi-species communities. *J. Math. Biol.* 26:651–660, 1988.
- [32] M. Grothaus and P. Stilgenbauer. Geometric langevin equations on submanifolds and applications to the stochastic melt-spinning process of nonwovens and biology. *Stochastics and Dynamics*, 13(04):1350001, 2013.
- [33] M. Gu and S. He. Mixing, enhanced dissipation and phase transition in the kinetic vicsek model. *arXiv:2409.19859*, 2024.
- [34] S.-Y. Ha and E. Tadmor. From particle to kinetic and hydrodynamic descriptions of flocking. *Kinetic and Related Models*, 1(3):415–435, 2008.
- [35] S. Haber. Numerical evaluation of multiple integrals. *SIAM Review*, 12(4):481–526, 1970.
- [36] S. He and A. Kiselev. Stirring speeds up chemical reaction. *Nonlinearity*, 35(8):4599, 2022.
- [37] S. He and E. Tadmor. A game of alignment: collective behavior of multi-species. *Ann. Inst. H. Poincaré C Anal. Non Linéaire*, 38(4):1031–1053, 2021.
- [38] F. Hérau and F. Nier. Isotropic hypoellipticity and trend to equilibrium for the Fokker-Planck equation with a high-degree potential. *Arch. Ration. Mech. Anal.*, 171(2):151–218, 2004.
- [39] L. Hörmander. Hypoelliptic second order differential equations. *Acta Math.*, 119:147–171, 1967.
- [40] L. Kelvin. Stability of fluid motion-rectilinear motion of viscous fluid between two parallel plates. *Phil. Mag.*, (24):188, 1887.
- [41] A. Kolmogoroff. Zufällige Bewegungen (zur Theorie der Brownschen Bewegung). *Ann. of Math. (2)*, 35(1):116–117, 1934.
- [42] Y. Lu, J. Lu, and J. Nolen. Accelerating langevin sampling with birth-death. *arXiv preprint arXiv:1905.09863*, 2019.
- [43] J. C. Mattingly, A. M. Stuart, and D. J. Higham. Ergodicity for sdes and approximations: locally lipschitz vector fields and degenerate noise. *Stochastic processes and their applications*, 101(2):185–232, 2002.
- [44] S. Motsch and E. Tadmor. A new model for self-organized dynamics and its flocking behavior. *Journal of Statistical Physics* 144(5), 2011.
- [45] S. Motsch and E. Tadmor. Heterophilious dynamics enhances consensus. *SIAM Review* 56(4), 2014.
- [46] R. M. Neal. Annealed importance sampling. *Statistics and Computing*, 11(2):125–139, 2001.
- [47] F. Nier and B. Helffer. *Hypoelliptic estimates and spectral theory for Fokker-Planck operators and Witten Laplacians*. Springer, 2005.
- [48] E. Novak. Some results on the complexity of numerical integration. In R. Cools and D. Nuyens, editors, *Monte Carlo and Quasi-Monte Carlo Methods: MCQMC, Leuven, Belgium, April 2014*, volume 163 of *Springer Proceedings in Mathematics & Statistics*, pages 161–183. Springer, Cham, 2016.
- [49] G. Parisi. Correlation functions and computer simulations. *Nuclear Physics B*, 180(3):378–384, 1981.
- [50] G. A. Pavliotis. Stochastic processes and applications. *Texts in applied mathematics*, 60, 2014.
- [51] R. Pierrehumbert. Tracer microstructure in the large-eddy dominated regime. *Chaos, Solitons & Fractals*, 4(6):1091–1110, 1994. Special Issue: Chaos Applied to Fluid Mixing.
- [52] N. Pouradieu Duteil. Mean-field limit of collective dynamics with time-varying weights. *Netw. Heterog. Media*, 17(2):129–161, 2022.
- [53] C. P. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer, New York, second edition, 2004.

- [54] G. O. Roberts and R. L. Tweedie. Exponential convergence of langevin distributions and their discrete approximations. *Bernoulli*, 2(4):341–363, 1996.
- [55] R. Shvydkoy. Global hypocoercivity of kinetic Fokker-Planck-alignment equations. *Kinet. Relat. Models*, 15(2):213–237, 2022.
- [56] R. Shvydkoy. Environmental averaging. *EMS Surv. Math. Sci.*, 11(2):277–413, 2024.
- [57] R. Shvydkoy et al. *Dynamics and analysis of alignment models of collective behavior*. Springer, 2021.
- [58] R. Shvydkoy and E. Tadmor. Eulerian dynamics with a commutator forcing. *Transactions of Mathematics and its Applications*, 1(1):1–26, 2017.
- [59] I. H. Sloan and S. Joe. *Lattice Methods for Multiple Integration*. Clarendon Press, Oxford University Press, Oxford, 1994.
- [60] I. H. Sloan and H. Woźniakowski. An intractability result for multiple integration. *Mathematics of Computation*, 66(219):1119–1124, July 1997.
- [61] D. Talay. Stochastic hamiltonian systems: exponential convergence to the invariant measure, and discretization by the implicit euler scheme. *Markov Process. Related Fields*, 8(2):163–198, 2002.
- [62] L. Tan and J. Lu. Accelerate langevin sampling with birth-death process and exploration component. *arXiv preprint arXiv:2305.05529*, 2023.
- [63] N. G. Trillos and D. Sanz-Alonso. From optimization to sampling through gradient flows. *arXiv preprint arXiv:2302.10037*, 2023.
- [64] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet. Novel type of phase transition in a system of self-driven particles. *Phys. Rev. Lett.*, 75(6):1226–1229, 1995.
- [65] D. Wei. Diffusion and mixing in fluid flow via the resolvent estimate. *Science China Mathematics*, pages 1–12, 2019.
- [66] L. Wu. Large and moderate deviations and exponential convergence for stochastic damping hamiltonian systems. *Stochastic processes and their applications*, 91(2):205–238, 2001.

DEPARTMENT OF MATHEMATICS, PURDUE UNIVERSITY, WEST LAFAYETTE, IN, USA
 Email address: gao662@purdue.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF SOUTH CAROLINA, COLUMBIA, SC, USA
 Email address: siming@mailbox.sc.edu

DEPARTMENT OF MATHEMATICS AND IPST, UNIVERSITY OF MARYLAND, COLLEGE PARK, MD, USA
 Email address: tadmor@umd.edu