# A Probabilistic Analysis of Enhanced Dissipation


by

Ethan Lu



A thesis submitted in partial fulfillment of the requirements for the Master of Science in Mathematical Sciences at Carnegie Mellon University.

April 26, 2022

Committee:

Gautam Iyer, Chair
Clinton Conley
Tomasz Tkocz

# Abstract

In this thesis, we present several results concerning *enhanced dissipation*, a physical phenomenon that manifests in a wide variety of contexts. In chapter 2, we review several results from the theory of PDE that concern decay of solutions to the *advection-diffusion equations*. In particular, we present an argument from [1, 2] that demonstrates how "mixing conditions" on the advection term can lead to an exponential improvement to naive bounds on the dissipation time obtained in the absence of advection. In chapter 3, we turn our attention to a discrete-time analog of the model above, focusing instead on bounding the mixing rate of Markov processes obtained by interleaving an exponentially mixing dynamical system with a random walk. Motivated by the results of [3], we approach such systems using tools from the *symbolic dynamics*, which have recently been shown to universally capture all properties of exponentially mixing dynamical systems. In section 3.6, we prove our main result by showing how coupling together with mild geometric conditions on such systems can be used in an incredibly elegant way to prove tight mixing time bounds in a fairly broad set of examples.

# Acknowledgements

Firstly, thanks to Gautam Iyer for advising this thesis and for his patience and guidance throughout, as well as Jim Nolen for many helpful discussions. I would also like to thank Clinton Conley and Tomasz Tkocz for serving on my committee and for their excellent classes and teaching during my time at CMU.

Thanks also to Keenan Crane and Ian Tice for the amazing research opportunities they have provided me with and their mentorship throughout.

Finally, thanks to my family for their continued support of my academic ventures.

# Contents

# Chapter 1

# Overview

## 1.1 Introduction

This thesis concerns itself with the phenomenon of **enhanced dissipation**, an effect that commonly manifests itself in a variety of physical contexts, ranging from tumor growth and bacterial movement to fluid flow and chemotaxis (through e.g., the Cahn-Hilliard and Keller-Segel equations). The basic setup can be thought of as follows. Suppose you're given a container of incompressible fluid and a drop of dye and are tasked with coloring all of the water as quickly as possible. While you could just drop in the dye and step away, perhaps the more natural thing to do would be to put the dye in and then to stir the fluid, thereby speeding up the process of convection. This simple observation captures exactly the idea of *enhanced dissipation*, which has become a recent subject of study within the PDE community (see [4] for a review). Although precise formulations of this phenomenon vary significantly from problem to problem, perhaps one of the simplest models to consider is that of the advection–diffusion equation, which reads

$$\partial_t \theta + u \cdot \nabla \theta - \kappa \Delta \theta = 0 \tag{1.1}$$

where $\theta$ is understood to represent the concentration of a dye, $u$ is a (possibly time-dependent and divergence free) velocity field advecting the fluid, and $\kappa$ is the diffusivity of the dye. Within this context, a standard measure of dissipation is to consider bounds on the *variance* of such solutions (e.g., bounds on $\left\|\theta - \fint \theta\right\|_{L^2}$). As we will see in section 2.2, with a simple energy estimate, we can obtain the bound

$$\|\theta(\cdot,t)\|_{L^2_0(\Omega)} \leq e^{-\lambda_1 \kappa t} \|\theta(\cdot,0)\|_{L^2_0(\Omega)}$$

for all mean 0 solutions to (1.1), where $\lambda_1$ denotes the smallest non-zero eigenvalue of the Laplacian. Although such a calculation already yields a useful insight into the nature of the problem, the energy estimate above is inherently limited because it fails to consider advection. Indeed, after integrating by parts, the effect of $u$ cancels, meaning that this bound has no dependence on the bulk motion of the fluid. The primary question to answer, then, is to see how different properties of $u$ can be parlayed into improved bounds on the exponent given above, also known as the *dissipation time* of the system. Towards doing so, the primary relation we will be interested in studying is that between the *ergodicity* of such a system and its dissipation time.

In continuous time, recent results from [1, 2] have shown that the $\mathcal{O}(\kappa^{-1})$ bound on the dissipation time above can be improved to $\mathcal{O}(\ln \kappa)^2$ given an "exponential mixing" assumption on the underlying flow $u$, the details of which are the focus of chapter 2. While already a substantial improvement to the naive bounds discussed above, however, there still remains the natural question of whether or not these bounds are optimal. Heuristically, the answer is no: in addition to intuitive arguments suggesting an optimal bound of order $\mathcal{O}(\ln \kappa)$, several classes of exponentially mixing flows for which this bound holds have been explicitly constructed, suggesting that it could hold in the general case (see [1]).

The focus of chapter 3, then, is to examine situations in which we can recover this tighter bound through a slight modification to the underlying model. More precisely, our approach in this chapter is to model solutions as Markov processes obtained by intertwining a mixing dynamical system with a random walk. The upshot of this approach is that a different set of tools from probability can be applied directly, rather than having to rely solely on PDE methods. In particular, we will find that tools from the theory of *mixing times* prove indispensable, allowing us a more diverse set of tools to tackle these problems. In doing so, we will discover an unexpectedly deep connection to the theory of *symbolic dynamics*, which will lead us to an entirely new approach for characterizing and studying the interactions between advection and diffusion in this setting. As a result, we'll then be able to prove our main result in section 3.6, which parlays mild geometric conditions on the symbolic dynamics perspective into a $\mathcal{O}(\ln \kappa)$ bound on the associated mixing time.

## 1.2 Main Results

We now briefly state the two main results of this thesis. As already mentioned, our first result is due to [1, 2], and provides a sufficient condition for an exponential improvement to the dissipation time of advection-diffusion type systems.

**Theorem 1** (Proved in Theorem 2.10)**.** *Suppose that $u$ is a smooth, divergence free, and exponentially mixing velocity field. Then for $\kappa \ll 1$, the dissipation time of the associated advection-diffusion equations is bounded by*

$$\tau_d \leq C\,|\ln \kappa|^2$$

*for some $C > 0$.*

Our second result is original, and leverages recently established connections between the theory of exponentially mixing dynamical systems and symbolic dynamics to provide a slightly tighter bound to the mixing time of Markov processes obtained by interleaving an exponentially mixing dynamical system and a homogeneous noise kernel. We take care to note that Theorem 2 below is not stated in the full generality of Theorem 3.23, and is instead a consequence of this theorem together with Proposition 3.22; for the sake of clarity and it's particularly notable applications, we choose to only highlight this special case here.

**Theorem 2** (Proved in Proposition 3.22 and Theorem 3.23)**.** *Let $M$ be a Riemannian manifold and $T : M \to M$ be a smooth exponentially mixing dynamical system. Suppose further that $K_\kappa$ is a homogenous Markov transition kernel on $M$ as in section 3.6 and proposition 3.22, and that there exists an exponentially fine Bernoulli partition associated to $T$. Then the mixing time of the Markov process associated to $T * K_\kappa$ satisfies the estimate*

$$\tau_{mix} \leq C \left|\ln \kappa\right| .$$

## 1.3 Layout of the Thesis

We conclude now with a brief overview of the layout of the rest of this document. As mentioned above, in chapter 2, our main goal is to build towards the recent results of [1, 2], which parlay an exponential mixing requirement on the advection term into a logarithmic dissipation time of solutions to the advection-diffusion equation in continuous time. Towards doing so, we begin in section 2.1 by establishing the basic notation and conventions used throughout. With these out of the way, in section 2.2 we then show how diffusion alone can already lead to mixing via the energy estimate mentioned above. After introducing an additional advection term, we then frontload many of the analogous estimates needed for the next section. The chapter then concludes with the proof of its main result in section 2.3.

Changing gears, in chapter 3, we then consider a slightly different model of advection-diffusion in discrete time. In section 3.1, we begin with a brief review of various results in the theory of *Markov processes*, which serve as the analogue of the process of diffusion in this model. As in the previous case, even in the presence of diffusion alone, various tools from probability are able to provide us with quantitative bounds on the mixing time of such systems. After establishing these results, we then review various results from the theory of *dynamical systems* in section 3.2, which serve as the analogue of advection, paying particular attention to those concerning ergodicity. In section 3.3, we provide several prototypical examples of the systems we are interested in studying, which will then motivate the results of sections 3.4 and 3.5, where we establish the connection between exponentially mixing dynamical systems and *Bernoulli systems*. We conclude in section 3.6 with our main result of this chapter, which leverages mild geometric conditions on the symbolic dynamics to prove a tighter bound on the mixing time of such systems.

# Chapter 2

# Dissipation Enhancement in Continuous Time

## 2.1 Norms, Function Spaces, and Setup

As mentioned above, the focus of this chapter is to investigate how interactions between large-scale bulk movement and small-scale diffusion can lead to quantitatively faster mixing rates within the context of the *advection-diffusion equation* in continuous time, with a particular interest paid to the speedup afforded by *exponentially mixing velocity fields*. The primary result of this chapter is Theorem 2.10, with our proof mirroring that of [1, 2].

For the sake of presentation, for the remainder of this chapter, we assume that our domain of interest $\Omega$ is given by the $d$-dimensional torus $\mathbb{T}^d := (\mathbb{R}/\mathbb{Z})^d$, though most of the results presented here also hold in the more general case when $\Omega$ is given to be e.g. a smooth Riemannian manifold. Throughout, we also fix a smooth incompressible (i.e. divergence free) velocity field $u : \mathbb{R} \times \Omega \to \mathbb{R}$ and consider solutions $\theta : \mathbb{R}_{\geq s} \times \Omega \to \mathbb{R}$ to the equations

$$\begin{cases} \partial_t \theta + u \cdot \nabla \theta - \kappa \Delta \theta = 0 & \text{for } t > s \\ \theta(t = s) = \theta_s \end{cases} \tag{2.1}$$

given $s \in \mathbb{R}$, initial data $\theta_s \in L^2_0(\Omega)$, and a diffusivity parameter $\kappa > 0$. Classical results then guarantee global well-posedness for this system, and hence we can make the following definition:

**Definition 2.1.** *For $t \geq s$, we define $\mathcal{S}_{s,t} : L^2_0(\Omega)$ to be the operator satisfying*

$$\mathcal{S}_{s,t}(\theta_s) = \theta(t),$$

*where $\theta$ is the unique solution to equation* (2.1) *with initial data $\theta_s$.*

We now take a moment to standardize the notation used in the remainder of this chapter.

**Definition 2.2** (Norms and function spaces)**.**

1. *We define $L^2(\Omega)$ (abbreviated as $L^2$) to be the space of all square-integrable mean-zero functions $f : \Omega \to \mathbb{R}$, equipped with the norm*

$$\|f\|_{L^2} := \left(\int_\Omega |f|^2\right)^{1/2}$$

*Unless otherwise specified, $\|\cdot\|$ will always denote this norm, and $\langle\cdot,\cdot\rangle$ the associated inner product.*
*We define $L_0^2(\Omega)$ (analogously abbreviated as $L_0^2$) as the space of all functions in $L^2$ with mean zero (i.e. $\int f = 0$).*

2. *For $s \in \mathbb{R}$, we define **the (inhomogenous) Sobolev space** $H^s$ to be the space of all distributions $f$ such that $\|f\|_{\dot{H}^s} < \infty$, where*

$$\|f\|_{H^s} := \left(\sum_{\eta\in\mathbb{Z}^d} \left(1+|\eta|^2\right)^s \hat{f}(\eta)\right)^{1/2}$$

*For $s \in \mathbb{N}$, this is precisely the space of distributions with all derivatives of order at most $\alpha$ in $L^2$.*

3. *For $\alpha \in \mathbb{R}$, we define **the homogenous Sobolev space of order** $\alpha$ to be the space of all locally integrable functions $f$ such that $\|f\|_{\dot{H}^\alpha} < \infty$, where*

$$\|f\|_{\dot{H}^\alpha} := \left(\sum_{\eta\in\mathbb{Z}^d\setminus\{0\}} |\eta|^{2\alpha} \hat{f}(\eta)\right)^{1/2} = \left\|(-\Delta)^{\alpha/2} f\right\|$$

*If $\alpha \in \mathbb{N}$, this is precisely the space of distributions with all derivatives of order $\alpha$ in $L^2$.*

4. *For $s \le t$, we let $\varphi_{s,t} : \Omega \to \Omega$ be the dynamical system generated by $u$; that is, $\varphi_{s,t}$ is the unique solution to*

$$\begin{cases} \partial_t \varphi_{s,t}(x) = u\left(\varphi_{s,t}(x), t\right) & t > s, x \in \Omega \\ \varphi_{s,s}(x) = x & x \in \Omega \end{cases}$$

With this notation established, we are now ready to define the primary objects we're interested in studying.

**Definition 2.3.** *The **dissipation time** associated to the flow generated by $u, \kappa$ is*

$$\tau_d := \inf\{t - s \mid \|\mathcal{S}_{s,t}(\theta_s)\| \le \frac{\|\theta_s\|}{e} \quad \forall \theta_s \in L_0^2, t \in \mathbb{R}\}.$$

Intuitively, the dissipation time is the minimum time needed for the variance of any solution to decrease by a constant fraction (chosen here to be $1/e$).
As mentioned previously, even in the absence of advection, it can be shown that $\tau_d \leq \mathcal{O}(1/\kappa)$, which, at least heuristically, matches with one's intuition; given a point source that diffuses at speed $\kappa$, one would expect after time $1/\kappa$ that the source has "covered the entire domain," and hence has become sufficiently diffused.
Upon introducing an advection or mixing term to the dynamics of this equation, however, one would naturally expect this timescale to become much faster. Although precise notions of mixing vary drastically in the literature (see e.g. [5, 4, 6, 3]), for our purposes, we will be mostly interested in **strongly exponentially mixing flows**, which we define now.

**Definition 2.4.** *We say a smooth divergence free vector field $u$ is **exponentially mixing** if there exists $c_1, c_2 > 0$ such that*

$$\langle \varphi_{s,t} f, g \rangle \leq c_1 e^{-c_2(t-s)} \|f\|_{\dot{H}^1} \|g\|_{\dot{H}^1}$$

*for all $f, g \in L^2_0$.*

**Remark 2.5.** *As mentioned above, precise definitions of mixing vary widely throughout the literature, though most formulations end up being weaker than the one above. We take a moment to remark on some of these notions:*

1. *In the general case considered by [1], the authors define $\alpha, \beta$ flows that are mixing with rate function $h$ to be velocity fields $u$ that satisfy the condition*

$$\langle \varphi_{s,t} f, g \rangle \leq h(t-s) \|f\|_{\dot{H}^\alpha} \|g\|_{\dot{H}^\beta}$$

   *where $\alpha, \beta > 0$, $f, g$ are as above and $h : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is a generic decreasing function vanishing at infinity. In the special case that $h$ is of the form $h(x) := c_1 e^{-c_2 x}$ as above, it can be shown that our definition above with $\alpha = \beta = 1$ is actually equivalent to the same decay condition for any $\alpha, \beta > 0$ (though with a different choice of constants). Hence in the exponential case, our definition agrees exactly with that of this paper. Though the results proved in [1] continue to hold in the case of a generic rate function $h$, we choose to omit them for the sake of presentation.*

2. *In [2], a similar setup is considered where the velocity field is required to satisfy the estimate*

$$\|\varphi_{s,t} f\|_{\dot{H}^{-1}} \leq h(t-s) \|f\|_{\dot{H}^1}$$

   *Setting $\alpha = \beta = 1$ and dualizing, we see that this definition is then another special case of the previous item.*

3. *Finally, in the dynamical systems literature, the notion most commonly considered is that of being **weakly/strongly mixing**. Though we defer a precise definition of these terms to definition 3.11, we note that strong mixing turns out to be exactly equivalent to the assertion that*

$$\|\varphi_{s,t} f\|_{\dot{H}^{-1}} \overset{t\to\infty}{\to} 0$$

   *and hence is the weakest notion out of those considered here (see [5] for a proof of the fact above).*

## 2.2 Energy Estimates

With the notation of the previous section established, we turn now to the technical results and computations that we will need for our main result. Our first result in lemma 2.6 and corollary 2.7 is a simple calculation that allows us to explicitly calculate the rate of energy decay for solutions to equation (2.1). We then present a standard result that allows us to estimate the difference between the dynamics with and without diffusion for small timescales, and use it to prove lemma 2.9, which is the main technical result of this section.

**Lemma 2.6** (Energy Estimates)**.** *For $\theta$ as above, we have the equality*

$$\partial_t \|\theta(t,\cdot)\|_{L^2(\Omega)}^2 = -2\kappa \|\theta(t,\cdot)\|_{\dot{H}^1}^2$$

*which implies that*

$$\|\theta(t,\cdot)\|_{L^2(\Omega)} \leq e^{-\lambda_1\kappa(t-s)} \|\theta(s,\cdot)\|_{L^2(\Omega)}$$

*for all $s \leq t$, where $\lambda_1$ denotes the smallest non-zero eigenvalue of $\Delta$.*

*Proof.* Multiplying equation (2.1) by $\theta$ and integrating over $\Omega$, we have

$$\partial_t \|\theta(\cdot,t)\|_{L^2(\Omega)}^2 /2 = \int_\Omega \theta\partial_t\theta = \int_\Omega -u\theta\cdot(\nabla\theta) + \kappa\theta\Delta\theta = -\int \kappa \|\nabla\theta\|^2.$$

which is precisely the first equality. Applying Poincaré then yields the desired inequality. ∎

Already, the inequality above suffices to show that $\tau_d \leq 1/(\lambda_1\kappa)$. In the results that follow, however, we will also need the following corollary, which follows immediately by integrating the first equality.

**Corollary 2.7** (Energy Equality)**.** *For any $s \leq t$, we have that*

$$\|\theta(t)\|^2 = \|\theta(s)\|^2 \exp\left(-2\kappa\int_s^t \frac{\|\theta(r)\|_{\dot{H}^1}^2}{\|\theta(r)\|^2}dr\right).$$

As mentioned previously, we will also need the following result which allows us to estimate the difference between the inviscid and viscid dynamics.

**Lemma 2.8.** *Let $\phi : \mathbb{R}_{\geq s}\times\Omega \to \mathbb{R}$ be given by*

$$\phi(t) := \theta_s \circ \varphi_{s,t}$$

*where again $\varphi_{s,t}$ is as in definition 2.2. Then*

$$\|\theta(t)-\phi(t)\|^2 \leqslant 2\sqrt{2\kappa(t-s)}\,\|\theta_s\|\left(2\|\nabla u\|_{L^\infty}\int_s^t \|\theta\|_{H^1}^2 + \|\theta_s\|_{H^1}^2\right)^{1/2}. \tag{2.2}$$

*Proof.* For notational convenience we assume WLOG that $s = 0$. Multiplying equation (2.1) by $\Delta\theta(t)$ and integrating over space, we deduce that

$$\partial_t \|\theta\|_{H^1}^2 + 2\kappa \|\theta\|_{H^2}^2 \leq 2 \|\nabla u\|_{L^\infty} \|\theta\|_{H^1}^2 .$$

which, upon integrating in time, yields that

$$2\kappa \int_0^t \|\theta\|_{H^2}^2 \leq 2 \|\nabla u\|_{L^\infty} \left( \int_0^t \|\theta\|_{H^1}^2 \right) + \|\theta\|_{H^1}^2 \tag{2.3}$$

Also we note that

$$\partial_t \|\theta - \phi\|^2 = 2\kappa \langle \Delta\theta, \theta - \phi \rangle \leq 4\kappa \|\theta\|_{H^2} \|\theta_0\| ,$$

where the last inequality follows from corollary 2.7, which tells us that the energy is nondecreasing for both $\theta$ and $\phi$. Consequently,

$$\|\theta(t) - \phi(t)\|^2 \leq 4\kappa \|\theta_0\| \int_s^t \|\theta_s\|_{H^2} \leq 2\sqrt{2\kappa t} \|\theta_0\| \left( 2\kappa \int_s^t \|\theta_s\|_{H^2}^2 \right)^{1/2} .$$

Substituting in equation (2.3) then concludes. □

Using this estimate, we can now prove the main technical result of this section.

**Lemma 2.9.** *Let $H(\kappa) \in \mathbb{R}$ satisfy*

$$\sqrt{H(\kappa)}(\ln c_1 + \ln(H(\kappa)) + \ln 2) = \frac{c_2}{64\sqrt{\kappa \|\nabla u\|_{L^\infty}}}$$

*and suppose that $\lambda_N$ is the largest eigenvalue of $\Delta$ in $[0, H(\kappa)]$. Then if $\|\theta_s\|_{\dot{H}^1}^2 < \lambda_N \|\theta_s\|^2$ we have the estimate*

$$\|\theta(s + t_0)\|^2 \leq \exp\left( -\frac{\kappa H(\kappa) t_0}{8} \right) \|\theta_s\|^2$$

*where*

$$t_0 := 2 (\ln c_1 + \ln(2\lambda_N)) / c_2.$$

*Proof.* Again for the sake of clarity we assume that $s = 0$. In light of corollary 2.7 and the inequality $1 - x \leq e^{-x}$, it suffices to show that

$$\int_0^{t_0} \|\theta(t)\|_{\dot{H}^1}^2 \, dt \geq \frac{\lambda_N t_0 \|\theta_0\|^2}{8}$$

since for $\kappa \ll 1, \lambda_N \geq H(\kappa)/2$ by Weyl's lemma, and hence this will imply

$$\|\theta(t_0)\|^2 \leq \left( 1 - \frac{2\kappa\lambda_N t_0}{8} \right) \|\theta_0\|^2 \leq \exp\left( -\frac{\kappa H(\kappa) t_0}{8} \right) \|\theta_0\|^2 .$$

To do so, suppose towards a contradiction that the opposite inequality holds, e.g. that

$$\int_0^{t_0} \|\theta(t)\|_{\dot{H}^1}^2 \, dt < \frac{\lambda_N t_0 \|\theta_0\|^2}{8}$$

Then

$$\begin{aligned}
\int_0^{t_0} \|\theta(t)\|_{\dot H^1}^2 \, dt &\geq \lambda_N \int_{t_0/2}^{t_0} \|(I - P_N)\,\theta(t)\|^2 \, dt \\
&\geq \frac{\lambda_N}{2} \int_{t_0/2}^{t_0} \|(I - P_N)\,\phi(t)\|^2 \, dt - \lambda_N \int_{t_0/2}^{t_0} \|(I - P_N)\,(\theta(t) - \phi(t))\|^2 \, dt \\
&\geq \frac{\lambda_N t_0}{4} \|\theta_0\|^2 - \frac{\lambda_N}{2} \int_{t_0/2}^{t_0} \|P_N \phi(t)\|^2 \, dt - \lambda_N \int_0^{t_0} \|\theta(t) - \phi(t)\|^2 \, dt
\end{aligned}$$

where $P_N$ denotes the projection operator onto the first $N$ eigenvalues of $\Delta$. Using the fact that $u$ is exponentially mixing, we can control the second term in the final expression by

$$\begin{aligned}
\int_{t_0/2}^{t_0} \|P_N \phi(t)\|^2 \, dt &\leq \lambda_N \int_{t_0/2}^{t_0} \|\phi(t)\|_{\dot H^{-1}}^2 \, dt \leq \lambda_N \int_{t_0/2}^{t_0} (c_1 e^{-c_2 t})^2 \|\theta_0\|_{\dot H^1}^2 \, dt \\
&\leq \frac{t_0}{2} \lambda_N \left(c_1 e^{-c_2 t_0}\right)^2 \|\theta_0\|_{\dot H^1}^2 \leq \frac{t_0}{2} \lambda_N^2 \left(c_1 e^{-c_2 t_0}\right)^2 \|\theta_0\|^2 .
\end{aligned}$$

Finally, using lemma 2.8, we can control the last term by

$$\begin{aligned}
\int_0^{t_0} \|\theta(t) - \phi(t)\|^2 \, dt &\leq \int_0^{t_0} 2\sqrt{2\kappa t}\, \|\theta_0\| \left(2 \|\nabla u\|_{L^\infty} \int_0^t \|\theta\|_{H^1}^2 + \|\theta_0\|_{H^1}^2\right)^{1/2} dt \\
&\leq 2\sqrt{2\kappa} t_0^{3/2} \|\theta_0\| \left(2 \|\nabla u\|_{L^\infty} \int_0^{t_0} \|\theta\|_{H^1}^2 \, dt + \|\theta_0\|_{H^1}^2\right)^{1/2} \\
&\leq 2\sqrt{2\kappa} t_0^{3/2} \|\theta_0\|^2 \left(\frac{\|\nabla u\|_{L^\infty} \lambda_N t_0}{4} + \lambda_N\right)^{1/2} .
\end{aligned}$$

Plugging in these estimates, our choice of $t_0$, and dividing by $\lambda_N t_0 \|\theta_0\|^2$ yields

$$\frac{1}{8} > \frac{1}{4} - \frac{1}{16} - 2\sqrt{\|\nabla u\|_{L^\infty \kappa} \lambda_N} t_0$$

which contradicts our choice of $t_0$, finishing the proof. □

## 2.3 Exponentially Mixing Flows

With the previous calculations in hand, we are now ready to prove the main theorem of this chapter.

**Theorem 2.10.** *Suppose that $u$ is a smooth, divergence free, and exponentially mixing velocity field. Then for $\kappa \ll 1$, the dissipation time of the associated advection-diffusion equations is bounded by*

$$\tau_d \leq \frac{18}{\kappa H(\kappa)} \leq C (\ln \kappa)^2$$

*where $H(\kappa)$ is as in lemma 2.9.*

*Proof.* First note that $H(\kappa) \leq \mathcal{O}\left(\frac{1}{\kappa\left|\ln\kappa\right|^2}\right)$ for $\kappa \ll 1$, so it suffices to prove the first inequality.
Towards doing so, let $\lambda_N$ be as in lemma 2.9, and note that if $\left\|\theta(t)\right\|_{\dot{H}^1}^2 \geq \lambda_N \left\|\theta(t)\right\|^2$ for $t \in [s, t_0]$, then

$$\left\|\theta(t)\right\|^2 \leq \exp(-2\kappa\lambda_N(t-s)) \left\|\theta_s\right\|^2$$

for all $s \leq t \leq t_0$.
Using this observation together with lemma 2.9 allows us to find $t_1' < t_2' < \cdots$ increasing to infinity such that

$$\left\|\theta\left(t_k'\right)\right\|^2 \leq \exp\left(-\frac{\kappa H(\kappa)\left(t_k'-s\right)}{8}\right)\left\|\theta_s\right\|^2, \text{ and } t_{k+1}' - t_k' \leq t_0.$$

This immediately implies

$$\tau_d \leq \frac{16}{\kappa H(\kappa)} + t_)$$

By our choice of $\lambda_N$ and $t_0$, we know that $t_0 \leq 1/\left(\kappa\lambda_N\right) \leq 2/\left(\kappa H(\kappa)\right)$ for $\kappa \ll 1$, which shows the first inequality as desired. $\square$

# Chapter 3

# Dissipation Enhancement in Discrete Time

We turn our attention now to the phenomenon of dissipation enhancement in *discrete time*, which we aim to understand through the lens of probability and mixing times rather than the more dynamical approach of chapter 2. In doing so, we will adopt a slightly different model of the underlying processes, separately splitting the "advection" and "diffusion" terms of the last chapter into the Markov processes covered in section 3.1 and the dynamical systems defined in section 3.2. After reviewing several classical results in these fields, we will then provide several motivating examples for the types of systems we are interested in studying, which will then lead us to our discussion of the symbolic dynamics and its connection to these systems in sections 3.4 and 3.5. We conclude in section 3.6 with the proof of our main result in this chapter.

## 3.1 Markov Processes and Mixing

We begin with a brief review of some basic properties of *Markov processes*, which serve as the analogue of the diffusion processes seen in chapter 2. Although we will be primarily interested in studying the functional dynamics of such operators (e.g. their characteristics when viewed as linear maps between $L^p$ spaces), we include for completeness the standard probabilistic definitions of these processes.
Below, $M$ is always taken to be a complete separable metric space, with all associated measures being associated with the Borel sigma-algebra $\mathcal{B}(M)$.

**Definition 3.1.** *We denote by $\mathcal{M}_b(M)$ the vector space of all finite signed measures on $M$, equipped with the total variation norm*

$$\|\mu\|_{TV} := |\mu|(M).$$

*We define $\mathcal{M}_p(M) \subseteq \mathcal{M}_b(M)$ to be the space of all probability measures on $M$.*

**Definition 3.2.** *A* Markov transition kernel $K$ *on $M$ is a function $K : M \times \mathcal{B}(M) \to [0,1]$ such that*

- *For any fixed $A \in \mathcal{B}(M)$, $K(\cdot, A)$ is measurable,*
- *and for any fixed $x \in M, K(x, \cdot)$ is a probability measure on $M$.*

*Given any such kernel, we can naturally extend it to a continuous linear map $K : \mathcal{M}_b(M) \to \mathcal{M}_b(M)$ via the action*

$$K\mu(A) := \int_M K(x, A)\mu(dx)$$

*and to a linear map between measurable functions via*

$$Kf(x) := \int_M f(y)K(x, dy).$$

*Given two kernels $K, L$, we define their product $K \star L$ (often abbreviated $KL$) via*

$$K \star L(x, A) = KL(x, A) := \int K(x, dy)L(y, A)$$

*and similarly, we define exponentiation of a kernel $K$ as*

$$K^n := \underbrace{K \star K \star \cdots \star K}_{n \text{ times}}.$$

**Definition 3.3.** *Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ a sequence of $M$-valued adapted random variables $\{X_i\}_{i\in\mathbb{N}}$ is a **Markov Process** with transition kernel $K$ if there exists a sequence of i.i.d. uniform $[0, 1]$ random variables $\{U_i\}_{i\in\mathbb{N}}$ and a measurable function $f : M \times [0, 1] \to M$ such that*

- *$f(x, U_1)$ is equal in law to $K(x, \cdot)$ for all $x \in M$,*
- *and we have the identity $X_{i+1} = f(X_i, U_i)$ for all $i \in \mathbb{N}$.*

Note as a consequence of the definition above that we will always have

$$\mu_{X_{i+1}} = K\mu_{X_i}$$

where $\mu_X$ denotes the law of $X$, and that if $\{X_i\}$ has transition kernel $KL$, then

$$\mu_{X_{i+1}} = KL\mu_{X_i} = K(L\mu_{X_i}).$$

Although general Markov processes can behave arbitrarily poorly, for our purposes, we will typically require that $K$ is sufficiently regular to admit a unique stationary measure, e.g., that there exists a unique probability measure $\pi$ satisfying $K\pi = \pi$.
In particular, we will typically posit that $K(x, \cdot)$ can be identified with a function in $L^1(\pi) \cap L^\infty(\pi)$ via it's Radon-Nikodym derivative, which, under some mild symmetry assumptions, already guarantees existence, and that the operator $K$ restricted to $L^2(\pi)$ admits a spectral gap, which will guarantee uniqueness.
Although we will not make extensive use of it here, the following fact is often helpful when dealing with Markov operators from the functional analytic perspective:

**Theorem 3.4.** *Let $K$ be a Markov transition kernel, and $\pi$ be an invariant measure. Then $K$ restricts to a contraction of $L^p(\pi)$ for all $1 \leq p \leq \infty$, e.g. $\|Kf\|_{L^p(\pi)} \leq \|f\|_{L^p(\pi)}$ for all $f \in L^p(\pi)$.*

*Proof.* Applying Jensen's inequality for $1 \leq p < \infty$, we have that

$$\begin{aligned} \|Tf\|_{L^p(\pi)}^p &= \int_M \left| \int_M f(y)K(x,dy) \right|^p d\pi(x) \leq \int_M \int_M |f(y)|^p K(x,dy)d\pi(x) \\ &= \int_M |f(y)|^p d(K\pi)(y) = \int_M |f(y)|^p d\pi(y) = \|f\|_p^p \end{aligned}$$

The result for $p = \infty$ is obvious. □

For completeness, however, we briefly record two more general criteria that also guarantee existence and uniqueness of a stationary measure:

**Theorem 3.5** (Convergence to Stationary)**.** *Suppose that $K$ is a kernel and that there exists a probability measure $\varphi$ such that either*

- ***[Strong $\varphi$ recurrence]****: For all $x \in X$ and all $A \in \mathcal{B}(M)$ with $\varphi'(A) > 0$, there exists $N_{x,A}$ such that such that $K^n(x, A) > 0$ for all $n \geq N_{x,A}$.*
- ***[Lower bounded transition probabilities]*** *or that there exists $\varepsilon > 0$, such that*

$$K(x, A) \geq \varepsilon\varphi(A) \quad \forall x \in M, A \in \mathcal{B}(M).$$

*Then there exists a unique stationary measure $\pi$ for $K$ such that*

$$\|K^n\mu - \pi\|_{TV} \to 0 \text{ as } n \to \infty$$

*for all $\mu \in \mathcal{M}_p$.*

*Proof.* This is precisely the content of theorems 1 and 2 in [7]. □

In light of the convergence criterion identified above, we can now define the mixing time of a Markov kernel, which is exactly the analogue of the dissipation time in the previous chapter.

**Definition 3.6.** *Given a Markov kernel $K$ with unique stationary measure $\pi$, we define the* ***mixing time*** *$\tau_{mix}$ of $K$ to be*

$$\tau_{mix} := \sup_{\mu \in \mathcal{M}_p} \inf_{n \in \mathbb{N}} \|K^n\mu - \pi\|_{TV} \leq 1/e.$$

As in the continuous time case, the mixing time is the minimum time needed for the deviation of any initial probability distribution from stationary to decrease by a constant fraction (chosen here to be $1/e$).

Though more commonly used in the analysis of discrete space Markov chains, we will also find techniques from the theory of coupling extremely helpful, which we present now.

**Definition 3.7.** *Given two probability distributions $\mu, \nu$ on $M$, a* ***coupling*** *of $\mu, \nu$, is a pair of random variables $(X, Y)$ defined on a common probability space such that their marginal distributions are precisely $\mu$ and $\nu$.*

The following proposition then establishes a natural connection between the total variation norm and coupling.

**Proposition 3.8.**

$$\|\mu - \nu\|_{TV} = \min_{X,Y \text{ a coupling of } \mu,\nu} \mathbb{P}(X \neq Y)$$

Using this, we can establish a link between mixing times and Markovian couplings, which we define now.

**Definition 3.9.** *Given two probability distributions $\mu, \nu$ on $M$ and a Markov transition kernel $K$, a* ***Markovian coupling*** *of the process $K$ is a sequence of random variables $\{(X_i, Y_i)\}_{i\in\mathbb{N}}$ defined on a common probability space such*

1. *$(X_0, Y_0)$ is a coupling of $\mu, \nu$,*
2. *$\{X_i\}$ and $\{Y_i\}$ are Markov processes with transition kernel $K$,*
3. *and $\{X_i = Y_i\} \subseteq \{X_{i+1} = Y_{i+1}\}$.*

*Although most authors only insist on (1) and (2) in their definitions of a Markovian coupling, in practice nearly all such couplings will satisfy (3), so we include it here.*
*Under this assumption, we can also define*

$$\tau_{couple} := \inf\{t \in \mathbb{N} \mid X_t = Y_t\}.$$

**Theorem 3.10.** *For a Markov transition kernel $K$, the associated mixing time satisfies*

$$\tau_{mix} \leq 8 \sup_{\mu,\nu \in \mathcal{M}_p} \inf \mathbb{E}(\tau_{couple})$$

*where the inf is taken over all Markovian couplings of $\mu, \nu$.*

## 3.2 Dynamical Systems

We now review some basic definitions from the field of **dynamical systems**, which play the role of advection in our new setting.
Recall that a **dynamical system** is a measure space $(M, \mathcal{F}, \pi)$ equipped with a measure-preserving $T : X \to X$.
Note also that each dynamical system $T$ naturally induces a Markov transition kernel $K_T$ on $M$ via

$$K_T(x, \cdot) := \delta_{T(x)}$$

Abusing notation slightly, we will use $T$ and $K_T$ interchangeably in the remainder of this document. As in the previous chapter, we first begin by introducing the relevant notions of ergodicity that we'll be interested in throughout.

**Definition 3.11.** *A dynamical system $T$ is said to be* ***strongly mixing*** *if, for all $A, B \in \mathcal{F}$ that*

$$\lim_{n\to\infty} \mu(T^{-n}(A) \cap B) = \mu(A)\mu(B).$$

As mentioned in remark 2.5, this notion turns out to be exactly equivalent to decay of the $H^{-1}$ norm from the previous chapter.
While this condition already suffices to generate a number of interesting dynamical properties, again, we will primarily be interested in more quantitative restrictions on ergodicity, such as the one below.

**Definition 3.12.** *A dynamical system $T$ is* ***exponentially mixing*** *if, there exists $c_1, c_2 > 0$ and $k \in \mathbb{N}$ such that*

$$\left| \int fg \circ T^n - \int f \int g \right| \leq c_1 e^{-c_2 n} \|f\|_{C^k} \|g\|_{C^k}$$

*for all $f, g \in C^k$.*

## 3.3 Examples

We now consider some particularly nice examples of the systems we are interested in studying, which henceforth will be given by interleaving a mixing dynamical system and a Markov process. Though the examples given here are chosen to admit particularly elegant and elementary proofs, it should be noted that the methods used are extremely sensitive to the exact dynamics being used, and break almost immediately if e.g. the noise kernel is slightly asymmetrized.
The first example considered here is almost an exact analogue of the advection-diffusion system considered in the last chapter, motivated by the fact that the heat equation admits a fundamental solution:

**Proposition 3.13** (Doubling Map)**.** *Let $\mathbb{T} := \mathbb{R}/\mathbb{Z}$, and $T : \mathbb{T} \to \mathbb{T}$ be given by $x \mapsto 2x$. For $\kappa > 0$, let $K_\kappa$ be the transition kernel corresponding to the operator $e^{-\kappa\Delta}$, and consider the Markov transition kernel given by $TK_\kappa$. Then the Lebesgue measure is the unique invariant measure of this system with $\tau_{mix} \leq C |\ln \kappa|$.*

*Proof.* For the sake of consistency, denote by $\pi$ the Lebesgue measure, which is clearly invariant. Now observe that

$$\widehat{Tf}(k) = \hat{f}(k/2), \widehat{K_\kappa f}(k) = e^{-\kappa k^2} \hat{f}(k) \quad \forall f \in L^2$$

Where we understand $\hat{f}(k/2)$ to be 0 when $k$ is odd. Using Parseval, this immediately implies

$$\left\| (TK_\kappa)^{n+1}(f-1) \right\|_{L^2} \leq e^{-\kappa 2^n} \|f-1\|_{L^2}$$

for all $f$ with mean 1.
Now note that

$$K_\kappa : \mathcal{M}_p \to \{f d\pi \mid \{\|f\|_{L^1} = 1\}$$

and also that

$$\|K_\kappa\|_{L^1\to L^2} \leq C/\kappa$$

by Young's inequality. This implies that

$$\left\|(KP_\kappa)^{n+2}(\mu-\pi)\right\|_{TV} \leq Ce^{-\kappa 2^n}/\kappa$$

which suffices. □

In the previous example, our general approach was to use purely analytic techniques to bound the dissipation time, then to use a key $L^1$ to $L^2$ estimate to convert this into a mixing time bound, establishing a connection between the two.
Although we won't discuss it further in this document, it should be noted that similar techniques can also be used to bound the dissipation time in terms of the mixing time, though such bounds naturally remain suboptimal.

**Proposition 3.14** (Doubling Map with Uniform Noise)**.** *Let $\mathbb{T}, T$ be as in the previous theorem, and for $n \in \mathbb{N}$, let $\kappa_n := 2^{-n-1}$ and $K_n$ be the transition kernel given by*

$$K_n(x, A) = \pi(A \cap [x-\kappa, x+\kappa])/2\kappa$$

*where again $\pi$ denotes the Lebesgue measure. Then the Markov transition kernel given by $TK_\kappa$ has the Lebesgue measure as its the unique invariant measure with $\tau_{mix} \leq 2n = \mathcal{O}(|\ln\kappa|)$.*

*Proof.* A direct calculation yields that $(KP_n)^{2n}\mu = \pi$ for any $\mu \in \mathcal{M}_p$, which clearly suffices. □

As already noted above, though these two bounds admit particularly nice proofs, the techniques used to obtain them are rather brittle, and fail if e.g. the noise is not chosen to be homogeneous over space. Reframing these examples in the framework of sections 3.4 and 3.5, in section 3.6, we will return to these examples and show how our main result can be used to obtain analogous bound on asymmetrized analogues of these systems.

**Proposition 3.15.** *Let $M := \mathbb{T}^2$ and $T : \mathbb{T}^2 \to \mathbb{T}^2$ be the dynamical system given by*

$$(0.x_1x_2\cdots, 0.y_1y_2\cdots) \mapsto (0.x_2\cdots, 0.x_1y_1y_2\cdots),$$

*where $0.x_1x_2\cdots$ denotes the binary expansion of $x \in [0,1)$.*
*For $N \in \mathbb{N}$, let $K_N$ be the transition kernel associated to the process*

$$X_{i+1} = X_i + 2^{-N}\,\mathrm{Unif}[-1,1]$$

*Then the Lebesgue measure is the unique invariant measure of the kernel $TP_N$ with associated mixing time $\tau_{mix} \leq CN$.*

We defer a proof of this particular bound to section 3.6.

## 3.4 Symbolic Dynamics

We now take a brief digression to investigate properties of a particular class of dynamical systems known as *Bernoulli systems*.

Let $M \in \mathbb{N}$, and $p_1, \ldots, p_M \in [0,1]$ be such that $\sum p_i = 1$. We then set $\Omega = \{1, \ldots, M\}^{\mathbb{Z}}$ to be our sample space, and given $n \leq m \in \mathbb{Z}$ and $x \in \Omega$, we define the cylinder set

$$C_{n,m,\omega} := \{\eta \in \Omega \mid \eta_i = \omega_i \text{ for all } i \in [n,m]\} .$$

Now let $\mathcal{B}$ be the $\sigma$-algebra generated by all cylinder sets, and let $\mu$ be the associated *Bernoulli measure* defined by

$$\mu(C_{n,m,\omega}) = \prod_{i=n}^{m} p_{\omega_i} .$$

**Definition 3.16.** *The* ***Bernoulli shift*** *on* $(\Omega, \mathcal{B}, \mu)$ *is the dynamical system defined by*

$$(T\omega)_n = \omega_{n+1} .$$

**Lemma 3.17** (Mixing)**.** *The Bernoulli shift is strongly mixing.*

*Proof.* Our aim is to show that for any two measurable sets $A, B$ we have

$$\lim_{n\to\infty} \mu(T^{-n}A \cap B) = \mu(A)\mu(B) .$$

Approximating by finite unions of cylinder sets, it is enough to show this when $A, B$ are both cylinder sets, and so we now assume $A = C_{n,m,\omega}$ and $B = C_{k,l,\eta}$. In this case, when $n > 10 \max\{|n|, |m|, |k|, |l|\}$ we have

$$T^{-n}C_{n,m,\omega} \cap C_{k,l,\eta} = \{\zeta \in \Omega \mid \zeta_i = \eta_i \text{ for } i \in [k,l] \text{, and } \zeta_{j+n} = \omega_j \text{, for } j \in [n,m]\} .$$

and hence

$$\mu(T^{-n}C_{n,m,\omega} \cap C_{k,l,\eta}) = \mu(C_{n,m,\omega})\mu(C_{k,l,\eta}) , \qquad \text{for all } n > k .$$

□

## 3.5 Systems isomorphic to Bernoulli shifts

We now let $M$ be a Riemannian manifold, and $\varphi\colon M \to M$ be a volume preserving dynamical system.

**Definition 3.18.** *We say* $\varphi$ *is* Bernoulli *if there exists a Bernoulli shift* $(\Omega, \mathcal{B}, \mu)$ *and a bijective, measure preserving* $\pi\colon M \to \Omega$ *such that*

$$\pi \circ \varphi = T \circ \pi ,$$

*where* $T$ *is the shift operator on* $\Omega$*. In other words, we have that the following diagram commutes*

$$\begin{array}{ccc} M & \xrightarrow{\varphi} & M \\ \downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle \pi} \\ \Omega & \xrightarrow{S} & \Omega \end{array}$$

If $\varphi$ is Bernoulli, then let

$$A_i = \pi^{-1}(C_{1,1,i})\,.$$

where $i$ denotes the constant sequence. The sets $A_1, \ldots, A_N$ form a partition of $M$ and are called a *Bernoulli partition.* Note that for any $x \in M$ we have

$$\varphi^n(x) \in A_i \iff \pi(x)_n = i\,.$$

Thus, one can think of the map $\pi$ as identifying every point $x$ with its orbital history $(a_0, a_1, \ldots)$, where $a_i = j$ if and only if $\varphi^i(x) \in A_j$. For convenience, define

$$A_{n,m,x} = \pi^{-1} C_{n,m,\pi(x)} \subseteq M\,,$$

to be the inverse image of the cylinder sets.
In light of lemma 3.17, it's clear that any Bernoulli system is mixing. Surprisingly, a partial converse to this statement holds, which is precisely the content of our next theorem.

**Theorem 3.19** (Proved in [3])**.** *If $T$ is an exponentially mixing $C^{1+\alpha}$ diffeomorphism, then it is isomorphic to a Bernoulli system.*

Given the universality result quoted above, then, a natural next question to ask is whether or not information from the perspective of symbolic dynamics can be parlayed into improved mixing bounds on the original dynamical system. Perhaps unsurprisingly, the answer is yes, though we will require some additional geometric assumptions on the Bernoulli partition, which we define now.

**Definition 3.20.** *Given a manifold $M$, a map $\pi : M \to \Omega$ as above, and a decreasing function $\delta : \mathbb{N} \to \mathbb{R}$, we say that the induced Bernoulli partition is $\delta$-**fine** if*

$$\sup_{x \in M} \operatorname{diam}(A_{-N,N,x}) \le \delta(N)$$

*for all $N \in \mathbb{N}$. In the case $\delta$ can be chosen to be of the form $N \mapsto c_1 e^{-c_2 N}$ for $c_1, c_2 > 0$, we say that the induced partition is **exponentially fine**.*

Henceforth, we will be assuming that all the systems we work with are $\delta$-fine for some unspecified $\delta$.

## 3.6 Geometric Coupling

We are now ready to proceed to our main result. Although we defer a precise statement of the conditions we need to section 3.6, we provide here a heuristic motivation for the definitions that follow.
In chapter 2 and the previous examples, our primary interest was in the asymptotic behavior of $\tau_d$ and $\tau_{mix}$ as a function of the diffusivity parameter $\kappa$, which roughly captured the "length scale" of our noise kernel. In order to to effectively transfer this information to the side of the symbolic dynamics, using our assumption of $\delta$-fineness, we will want to choose a parameter $N$ such that for every $x \in M$, there exists a finite collection of cylinder sets $A_{-N,N,x_i}$ with the

associated uniform measure "well approximating" the kernel $K(x,\cdot)$, which will then allow us to couple in expected $\mathcal{O}(N)$ time.

With this in mind, we can now precisely state our necessary assumptions. Recall also that for a signed measure $\mu$, we write $\mu^+,\mu^-$ to denote the positive and negative parts of $\mu$ respectively, and for positive measures $\mu,\nu$, we set

$$\mu \wedge \nu := \mu - (\mu-\nu)^+$$

**Definition 3.21.** *Let $K$ be a Markov transition kernel on $(\Omega,\mathcal{B})$ and fix $N\in\mathbb{N}$. We say that $K$ has **resolution** $N$ if there exists $c_1,c_2>0$ such that all of the following hold for arbitrary $\omega,\eta\in\Omega$:*

1. *[**High frequency coupling**] There exists a bijective $H_{\omega,\eta}:\Omega\to\Omega$ such that*

$$(H_{\omega,\eta}\star K(\omega,\cdot)\wedge K(\eta,\cdot))(\Omega)\geq c_1$$

   *and $H_{\omega,\eta}(\gamma)_i=\gamma_i$ for all $i\geq N,\gamma\in\Omega$.*

2. *[**Closeness**] If $\omega_i=\eta_i$ for all $i\geq -N$, then*

$$(K(\omega,\cdot)\wedge K(\eta,\cdot))(\Omega)\geq c_2.$$

3. *[**Homogenity**] For all $-N\leq m\leq N$, there exists a bijective $M_{m,\omega,\eta}:\Omega\to\Omega$ such that if $\omega_i=\eta_i$ for all $i\geq m$,*

$$M_{m,\omega,\eta}\star K(\omega,\cdot)=K(\eta,\cdot)$$

   *and $M(\gamma)_i=\gamma_i$ for all $i\geq m,\gamma\in\Omega$.*

Roughly speaking, the first condition above imposes a "scale condition" on the noise that forces it to be approximately uniform when restricting it's action to only those coordinates with index $\geq N$, which will be key in allowing us to couple these higher-order bits. The third homogeneity condition As mentioned previously, the conditions are intentionally stated in as abstract a form as possible; in practice, conditions like the one below are typically easier to check.

**Proposition 3.22.** *If on $M$, the kernel $K$ satisfies*

$$K(x,\cdot)\geq\alpha\frac{1}{\mu B(x,\kappa)}\chi_{B(x,\kappa)}d\mu$$

*for some $\alpha>0$, then it satisfies the first two assumptions above for any $N$ with $\delta(N)\leq\kappa/2$.*

*Proof.* We show that the first item holds, fixing $x,y\in M$ and $N$ as above. First note that $A_{-N,N,x}\subseteq B(x,\kappa), A_{-N,N,y}\subseteq B(y,\kappa)$, hence we can choose $H_{\pi(x),\pi(y)}$ to simply be the map sending $\pi(x)\to\pi(y)$ that acts trivially on all coordinates greater than $N$ in absolute value. On the other hand, to show that the second item holds, we can note that $\pi(x)_i=\pi(y)_i$ for all $i\geq -N$ implies that $d(x,y)<\kappa/2$, upon which the conclusion is immediate. □

With these definitions out of the way we are now ready to prove our main result.

**Theorem 3.23.** *If $K$ is as in section 3.6, then the Markov transition kernel given by $K \star P$ satisfies*

$$\tau_{mix} \le 16N/(c_1 c_2).$$

*Proof.* In light of theorem 3.10, it suffices to construct a coupling of any two initial states $\omega_0, \eta_0$ that couples in expected $2n/(c_1c_2)$ time. For convenience, we set

$$\mathcal{H}_i := (H_{\omega_i,\eta_i} \star P(\omega_i,\cdot) \wedge P(\eta_i,\cdot)) \in \mathcal{M}_b(\Omega).$$

Our construction is as follows, viewing the following as an iterative process to construct the sequences $\{\omega_i\}_{i\in\mathbb{N}}, \{\eta_i\}_{i\in\mathbb{N}}$:

1. Draw $U \sim \mathrm{Unif}[0,1]$. If $U \le \mathcal{H}_i(\Omega)$, sample $\eta$ proportionally to $\mathcal{H}_i$, and set

   $$\eta_{i+1} := T\eta, \omega_{i+1} = TS^{-1}_{\omega_i,\eta_i}(\eta_{i+1}).$$

   Otherwise, draw $\omega_{i+1}, \eta_{i+1}$ from $T(P(\omega_i) - S^{-1}_{\omega_i,\eta_i} \star P(\eta_i,\cdot))^+$ and $T(S_{\omega_i,\eta_i} \star P(\omega_i,\cdot) - P(\eta_i,\cdot))^+$ respectively. Repeat until the former case occurs.

2. For the next $N$ steps, sample $\omega$ from $K(\omega_i,\cdot)$, and set

   $$\omega_{i+1} = T(\omega)\eta i + 1 = M_{m,\omega_i,\eta_i}(\omega).$$

3. Draw $U \sim \mathrm{Unif}[0,1]$. If $U \le (K(\omega_i,\cdot) \wedge K(\eta_i,\cdot))(\Omega)$, sample $\eta_{i+1} = \omega_{i+1}$ proportional to the measure $K(\omega_i,\cdot) \wedge K(\eta_i,\cdot)$, and henceforth set $\omega, \eta$ to be equal.
   Otherwise, draw $\omega_{i+1}, \eta_{i+1}$ from $(K(\omega_i,\cdot) - K(\eta_i,\cdot))^+$ and $(K(\omega_i,\cdot) - K(\eta_i,\cdot))^+$ respectively, and return to the first step.

Now note that if the former case occurs in step three, the two processes have coupled, and that we reach this step in at most expected $2n/(c_1c_2)$ time as desired. □

We now prove proposition 3.15.

*Proof of proposition 3.15.* □

Upon making the identification of each $(a,b) \in \mathbb{T}^2$ with the bi-infinite sequence $\cdots b_2 b_1 a_1 a_2 \cdots$, it's easy to see that $\varphi$ is Bernoulli with Bernoulli partition given by dyadic rectangles.
It therefore suffices to check the conditions in . By Proposition 3.22, we know that $K$ satisfies items 1 and 2, so it suffices to check homogeneity, which we can do by simply setting

$$M_{m,x_0,y_0}(x) := x + (y_0 - x_0)$$

which concludes.

**Theorem 3.24** (Shifted Bernoulli Shift)**.** *Let $T$ be as above and $K_N$ be the transition kernel associated to the process*

$$X_{i+1} = X_i + 2^{-N}\,\mathrm{Unif}[-1,1] + \varepsilon$$

*Then the Lebesgue measure is the unique invariant measure of the kernel $TP_N$ with associated mixing time $\tau_{mix} \le CN$.*

*Proof.* The same construction as above works. □

# Bibliography


[1] Yuanyuan Feng and Gautam Iyer. Dissipation enhancement by mixing. *Nonlinearity*, 32(5):1810–1851, Apr 2019.

[2] Michele Coti Zelati, Matias G. Delgadino, and Tarek M. Elgindi. On the relation between enhanced dissipation timescales and mixing rates. *Communications on Pure and Applied Mathematics*, 73(6):1205–1244, Jun 2020.

[3] Dmitry Dolgopyat, Adam Kanigowski, and Federico Rodriguez-Hertz. Exponential mixing implies Bernoulli. 2021.

[4] Jean-Luc Thiffeault. Using multiscale norms to quantify mixing and transport. *Nonlinearity*, 25(2):R1–R44, Jan 2012.

[5] ZHI LIN, JEAN-LUC THIFFEAULT, and CHARLES R. DOERING. Optimal stirring strategies for passive scalar mixing. *Journal of Fluid Mechanics*, 675:465–476, 2011.

[6] Ravi Montenegro and Prasad Tetali. Mathematical aspects of mixing times in Markov chains. *Foundations and Trends in Theoretical Computer Science*, 1(3):237–354, 2006.

[7] Persi Diaconis and David Freedman. On Markov chains with continuous state space. 1997.