RESIDUAL DIFFUSIVITY FOR EXPANDING BERNOULLI MAPS

WILLIAM COOPERMAN, GAUTAM IYER, AND JAMES NOLEN

Abstract. Consider a discrete time Markov process X^ε on $\mathbb{R}^d$ that makes a deterministic jump based on its current location, and then takes a small Gaussian step of variance ε^2. We study the behavior of the *asymptotic variance* as $\varepsilon \to 0$. In some situations (for instance if there were no jumps), then the asymptotic variance vanishes as $\varepsilon \to 0$. When the jumps are "chaotic", however, the asymptotic variance may be bounded from above and bounded away from 0, as $\varepsilon \to 0$. This phenomenon is known as *residual diffusivity*, and we prove this occurs when the jumps are determined by certain expanding Bernoulli maps.

1. Introduction

1.1. **Main Result.** Consider a Markov process $\{X_n^\varepsilon\}_{n \geqslant 0}$ that makes a deterministic jump based on its current location, followed by a Gaussian step of variance ε^2. Explicitly, X_{n+1}^ε is determined from X_n^ε by

$$X_{n+1}^\varepsilon = \varphi(X_n^\varepsilon) + \varepsilon \xi_{n+1} \,. \tag{1.1}$$

Here $\{\xi_n\}_{n \geqslant 1}$ is a family of independent standard Gaussian random variables, and $\varphi \colon \mathbb{R}^d \to \mathbb{R}^d$ is a Lebesgue measure preserving map with a periodic displacement (i.e. the function $x \mapsto \varphi(x) - x$ is $\mathbb{Z}^d$ periodic).

In this situation it is easy to see that for any $v \in \mathbb{R}^d$, the variance $\operatorname{var}(v \cdot X_n^\varepsilon)$ grows linearly with n as $n \to \infty$. Our interest is to study the growth rate of the variance asymptotically as $\varepsilon \to 0$. More precisely, we are interested in the behavior of the *asymptotic variance*

$$\lim_{n \to \infty} \frac{\operatorname{var}^{\mu_0}(v \cdot X_n^\varepsilon)}{n}$$

in the vanishing noise limit $\varepsilon \to 0$. Here μ_0 is a probability distribution on $\mathbb{R}^d$, and the notation $\operatorname{var}^{\mu_0}(v \cdot X_n^\varepsilon)$ denotes the variance of $v \cdot X_n^\varepsilon$ given $X_0^\varepsilon \sim \mu_0$.

If φ is "not too chaotic", then the asymptotic variance of $v \cdot X^\varepsilon$ may either vanish as $\varepsilon \to 0$ (for instance, if φ is the identity map) or diverge to $+\infty$ (for instance, if the map φ has a shear structure, with unbounded orbits). If, however, φ is "chaotic", then it may be possible for the asymptotic variance of $v \cdot X^\varepsilon$ to be bounded from above and bounded away from 0 as $\varepsilon \to 0$. This phenomenon is known as *residual diffusivity* and its study originated in fluid dynamics [Tay21, BCVV95, MCX+17]. The main result of this paper shows that the processes X^ε exhibits residual diffusivity when φ is obtained from a certain class of expanding Bernoulli maps (see Section 2,

2020 *Mathematics Subject Classification.* Primary: 60J05. Secondary: 37A25, 76M50,

Key words and phrases. homogenization, effective diffusivity, residual diffusivity, mixing.

This work has been partially supported by the National Science Foundation under grants DMS-2303355, DMS-2342349, DMS-2406853, and the Center for Nonlinear Analysis.

below). To the best of our knowledge, this is the only class of chaotic maps for which residual diffusion has been rigorously proved.

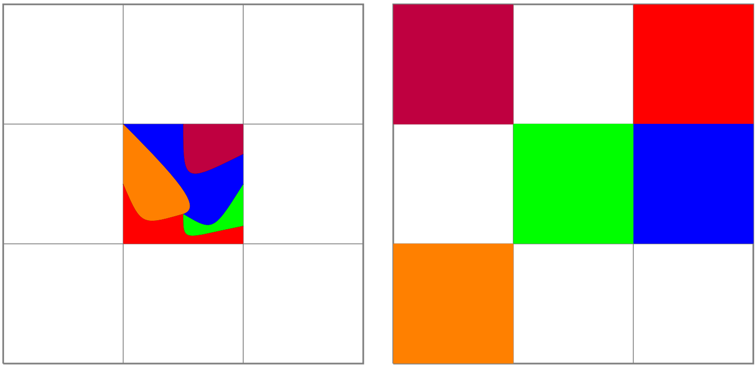

FIGURE 1. One example of an expanding Bernoulli map φ. The colored regions on the left are mapped to regions of the same color on the right.

Our main result is the following:

Theorem 1.1. *Suppose φ is obtained from an expanding Bernoulli map satisfying the conditions in Assumption 2.1, below. For all $v \in \mathbb{R}^d$, and all subgaussian initial distributions μ_0 we have*

$$\lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{\operatorname{var}^{\mu_0}(v \cdot X_n^\varepsilon)}{n} = \operatorname{var}^{\pi_0}(v \cdot \lfloor X_1^0 \rfloor) \,. \tag{1.2}$$

Here π_0 denotes the uniform distribution on the unit cube $Q_0 \stackrel{\text{def}}{=} [0,1)^d$, and the notation $\lfloor \cdot \rfloor$ denotes the bottom left vertex of the containing unit integer lattice cube. Explicitly, for $x \in \mathbb{R}^d$, the notation $\lfloor x \rfloor$ denotes the unique element in $\mathbb{Z}^d$ for which $x \in \lfloor x \rfloor + Q_0$.

Notice that the right side of (1.2) is non-zero if the image of Q_0 under φ intersects more than one cube. One example of an expanding Bernoulli map that satisfies the conditions of Theorem 1.1 is shown pictorially in Figure 1. Another example can be constructed from the one dimensional doubling map, and is defined by

$$\varphi(x) = 2x - \lfloor x \rfloor \,, \quad x \in \mathbb{R} \,. \tag{1.3}$$

Interestingly, if we introduce a shift by 1/2 into (1.3) then we obtain an example of a mixing map which does *not* exhibit residual diffusivity. Explicitly, we define the *shifted doubling map* by

$$\varphi_2(x) = 2x - \lfloor x \rfloor - \frac{1}{2} \,, \quad x \in \mathbb{R} \,.$$

The shifted doubling map satisfies all but one of the required hypothesis in Theorem 1.1 (specifically, for the shifted doubling map the cube centers o_i in Section 2, below, are not integers). The results of the numerical simulations in Figure 2 show that the doubling map exhibits residual diffusivity, while the shifted doubling map does not.

The main idea behind the proof of Theorem 1.1 is to use mixing to show that the asymptotic variance starting from any (subgaussian) initial distribution μ_0 is

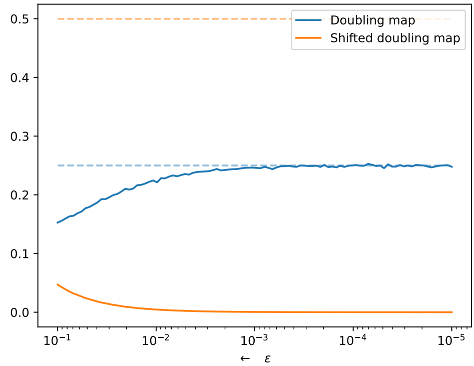

FIGURE 2. Asymptotic variance for the doubling map, and the shifted doubling map as ε varies from 10^{-1} to 10^{-5}. For the doubling map the asymptotic variance approaches the value on the right of (1.2) (blue dashed line) as $\varepsilon \to 0$. The shifted doubling map, however, doesn't exhibit residual diffusivity and the asymptotic variance approaches 0 instead of the value on the right of (1.2) (orange dashed line).

close to the asymptotic variance starting from the uniform distribution π_0 on Q_0. When $X_0^\varepsilon \sim \pi_0$, the Bernoulli structure of φ serendipitously makes increments of $\lfloor \varphi(X_n^\varepsilon) \rfloor - \lfloor X_n^\varepsilon \rfloor$ independent, allowing us to compute the asymptotic variance explicitly. Before delving into the technical details of the proof we briefly survey the literature and place Theorem 1.1 in the context of existing results.

1.2. **Motivation and Literature review.** Our interest in this problem stems from understanding the long time behavior of diffusions whose drift has "chaotic trajectories". Explicitly, consider the continuous time diffusion process X_t^ε defined by the SDE

$$dX_t^\varepsilon = u(X_t^\varepsilon)\, dt + \varepsilon\, dW_t \quad \text{on } \mathbb{R}^d \,, \tag{1.4}$$

where W is a d-dimensional Brownian motion, and u is a spatially periodic, divergence free vector field. One physical situation where this is relevant is in the study of diffusive tracer particles being advected by an incompressible fluid.

On small (i.e. $O(1)$) time scales, it the process of X^ε stays close to the deterministic trajectories of u, and a large deviations principle can be established (see for instance [FW12]). On intermediate (i.e. $O(|\ln \varepsilon|/\varepsilon^\alpha)$ for $\alpha \in [0, 2)$) time scales certain non-Markovian effects arise and lead to anomalous diffusion [You88, YPP89, Bak11, HKPG16, HIK+18]. On long (i.e. $O(1/\varepsilon^2)$) time scales, homogenization occurs and the net effect of the drift can be averaged and the process X^ε can be approximated by a Brownian motion with covariance matrix $D_{\text{eff}}^\varepsilon$ called the *effective diffusivity*. This was first studied in this setting by Freidlin [Fre64], and is now the subject of many standard books with several important applications [BLP78, PS08].

The effective diffusivity matrix $D_{\text{eff}}^\varepsilon$ is the unique symmetric matrix whose action on vectors $v \in \mathbb{R}^d$ is given by

$$v D_{\text{eff}}^\varepsilon v = \lim_{t \to \infty} \frac{\operatorname{var}(v \cdot X_t^\varepsilon)}{t} \,.$$

This, however, is hard to compute explicitly and authors usually characterize it in terms of a cell problem on $\mathbb{T}^d$, or a variational problem. In a few special situations (such as shear flows, or cellular flows) the asymptotic behavior of $D^\varepsilon_{\text{eff}}$ as $\varepsilon \to 0$ is known [Tay53, CS89, FP94, FP97, MK99, Hei03, Kor04, NPR05, RZ07].

The motivation for the present paper comes from thinking about the case when the deterministic flow of u is chaotic on the torus. In this case it has been conjectured that $D^\varepsilon_{\text{eff}}$ is $O(1)$ as $\varepsilon \to 0$, a phenomenon known as *residual diffusivity.* Study of this was initiated by Taylor [Tay21] over 100 years ago, and has since been extensively studied by many authors [ZSW93, BCVV95, MK99, Zas02, MCX$^+$17]. While this has been confirmed numerically and studied for elephant random walks [LXY17, LXY18, MCZ$^+$20, WXZ21, WXZ22, LWXZ22, KLX22], a rigorous proof is of this in even one example is still open.

The goal of this paper is to rigorously exhibit residual diffusivity in a simple setting. First we replace the notion of "chaotic" with the assumption that the flow of u on the torus $\mathbb{T}^d$ is *exponentially mixing* (see [SOW06]). In continuous time, however, examples of exponentially mixing flows are not easy to construct. The canonical example of an exponentially mixing flow is the geodesic flow on the unit sphere bundle of negatively curved manifolds [Dol98]. On the 3-torus, however, the existence of a divergence free, C^1, time independent, exponentially mixing velocity field is an open question. To the best of our knowledge, there are only examples of lower regularity [EZ19], and several time dependent examples [Pie94, BBPS19, MHSW22, BCZG23, ELM23, CFIN23].

On the other hand, there are several simple, well known, examples of exponentially mixing dynamical systems on the torus [SOW06]. Therefore, instead of studying a continuous-time system, we study the discrete time system (1.1). The system (1.1) can be viewed as a discretization of (1.4) where φ is the flow map of u after a fixed amount of time. The periodicity of u translates to the requirement that the displacement $x \mapsto \varphi(x) - x$ is periodic. Incompressibility of u translates to the requirement that φ is Lebesgue measure preserving. This leads us to study (1.1) in the general situation that φ is any Lebesgue measure preserving with a periodic displacement (and not necessarily a diffeomorphism obtained as the flow of an incompressible vector field). Such systems are interesting in their own right, and various aspects of them have been extensively studied [FW03, FNW04, TC03, FI19, OTD21, ILN24].

In this time-discrete setting, Theorem 1.1 exactly computes the vanishing noise limit of the effective diffusivity. Our proof, however, relies on the Bernoulli structure of φ and will not apply to general mixing maps. For general mixing maps φ we conjecture that (1.2) in Theorem 1.1 should be replaced with

$$\lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{\operatorname{var}^{\mu_0}(v \cdot X^\varepsilon_n)}{n} = \lim_{n \to \infty} \frac{\operatorname{var}^{\pi_0}(v \cdot X^0_n)}{n} \,.$$

Note, when $\varepsilon > 0$ it is natural to expect that the mixing effects of the noise eliminate the dependence of the asymptotic variance on the initial distribution. The same is true when $\varepsilon = 0$, provided the dynamics of φ are mixing and the initial distribution is regular. There is, however, no explicit formula for the asymptotic variance and no easy way to compute it in general. In our situation, the Bernoulli structure of φ allows us to compute the asymptotic variance when the initial distribution is π_0, and this is used in the proof of Theorem 1.1.

Plan of this paper. In Section 2 we fix our notation convention and precisely state the assumptions under which Theorem 1.1 is true. In Section 3 we explain the main idea behind the proof of Theorem 1.1, and carry out the details modulo the computation of the asymptotic variance when the initial distribution is uniform (Lemma 3.2, below). Finally in Section 4 we prove Lemma 3.2.

Acknowledgements. The authors would like to thank Jack Xin and Albert Fannjiang for helpful discussions.

2. Expanding Bernoulli maps

We begin by precisely describing the class of maps φ for which Theorem 1.1 holds. Partition $\mathbb{R}^d$ into unit cubes $\{Q_k \mid k \in \mathbb{Z}^d\}$, where $Q_k = k + [0,1)^d$. Let $M \geqslant 2$ and $E_1, \dots, E_M \subseteq Q_0$ be a Borel measurable partition of Q_0, with

$$|E_1| \leqslant |E_2| \cdots \leqslant |E_M| .$$

For each $i \in \{1, \dots, M\}$, let $o_i \in \mathbb{Z}^d$ and $\varphi_i \colon E_i \to Q_{o_i}$ be a Borel measurable bijection which pushes forward the normalized Lebesgue measure on E_i to the Lebesgue measure on Q_{o_i}.

Given $x \in \mathbb{R}^d$ we let $n = \lfloor x \rfloor$ denote the unique element in $\mathbb{Z}^d$ such that $x \in Q_n = n + Q_0$. Define $\varphi \colon \mathbb{R}^d \to \mathbb{R}^d$ by

$$\varphi(x) = n + \varphi_i(x - n) \quad \text{if } x - n \in E_i . \tag{2.1}$$

Clearly the function $x \mapsto \varphi(x) - x$ is $\mathbb{Z}^d$ periodic, and so φ projects to a well-defined map $\tilde{\varphi} \colon \mathbb{T}^d \to \mathbb{T}^d$ given by

$$\tilde{\varphi}(\tilde{x}) = \tilde{y} , \qquad \text{where} \quad y = \varphi(x) .$$

Here $x, y \in \mathbb{R}^d$ and $\tilde{x}, \tilde{y} \in \mathbb{T}^d$ denote the equivalence classes x and y modulo $\mathbb{Z}^d$ respectively.

We note the map $\tilde{\varphi}$ defined above is conjugate to a one-sided Bernoulli shift on sequences $\{1, \dots, M\}^{\mathbb{Z}}$. Moreover, since $\tilde{\varphi}$ expands each set $\tilde{E}_i$ to the torus $\mathbb{T}^d$, we can view $\tilde{\varphi}$ as an expanding Bernoulli map. In addition to the above structure, we require a mixing assumption $\tilde{\varphi}$ in order to prove Theorem 1.1. We now state this assumption precisely.

Assumption 2.1. *The map $\tilde{\varphi} \colon \mathbb{T}^d \to \mathbb{T}^d$, defined as above, is piecewise C^1 and exponentially mixing. That is, there exist constants $D < \infty$ and $\gamma > 0$ such that for every pair of test functions $\tilde{f}, \tilde{g} \in H^1(\mathbb{T}^d)$, we have*

$$\Big| \langle \tilde{f}, \tilde{g} \circ \tilde{\varphi}^n \rangle - \int_{\mathbb{T}^d} \tilde{f} \, dx \int_{\mathbb{T}^d} \tilde{g} \, dx \Big| \leqslant D e^{-\gamma n} \|\tilde{f}\|_{H^1} \|\tilde{g}\|_{H^1} . \tag{2.2}$$

We clarify that $\tilde{E}_i \subseteq \mathbb{T}^d$ is the projection of $E_i \subseteq \mathbb{R}^d$ to the torus $\mathbb{T}^d$. We also note that (2.1) and the fact that $\tilde{\varphi} \colon \mathbb{T}^d \to \mathbb{T}^d$ is Lebesgue measure preserving implies $\varphi \colon \mathbb{R}^d \to \mathbb{R}^d$ is also Lebesgue measure preserving. The shift invariance of φ implies the projected process $\tilde{X}^\varepsilon$ is a Markov process on the torus $\mathbb{T}^d$. Equation (1.1) and the fact that φ is Lebesgue measure preserving implies that the stationary distribution of $\tilde{X}^\varepsilon$ is $\tilde{\pi}$, the uniform measure on $\mathbb{T}^d$.

Remark 2.2. The exponential mixing requirement in Assumption 2.1 can be weakened. In the proof, this condition is only used to ensure the right hand side of (4.3) in

Proposition 4.2 (below) vanishes. Theorem 1.1 will still hold provided we replace Assumption 2.1 with the assumption that

$$\lim_{\varepsilon \to 0} \varepsilon \tilde{t}^{\varepsilon}_{\mathrm{mix}} = 0 \,, \tag{2.3}$$

where $\tilde{t}^{\varepsilon}_{\mathrm{mix}}$ is the 1/2 mixing time of $\tilde{X}^{\varepsilon}$ on $\mathbb{T}^d$. Available results (see for instance [FI19, ILN24, IZ23]) show even a quadratic decay in (2.2) implies (2.3).

Notation and convention. We now briefly clarify several convention that will be used throughout this paper.

(1) A tilde is used to denote projections onto the torus. That is if $x \in \mathbb{R}^d$, then $\tilde{x} \in \mathbb{T}^d$ denotes the equivalence class $x \pmod{\mathbb{Z}^d}$. Conversely if $\tilde{y} \in \mathbb{T}^d$, we will implicitly use $y \in \mathbb{R}^d$ to denote any representative of the equivalence class $\tilde{y}$.
(2) If $f \colon \mathbb{R}^d \to \mathbb{R}$ is a function then $\tilde{f} \colon \mathbb{T}^d \to \mathbb{R}$ denotes its periodization

$$\tilde{f}(\tilde{x}) \stackrel{\text{def}}{=} \sum_{k \in \mathbb{Z}^d} f(x+k) \,.$$

Similarly, if μ is a measure on $\mathbb{R}^d$ then $\tilde{\mu}$ is the measure on $\mathbb{T}^d$ defined by $\tilde{\mu}(\tilde{E}) = \sum_{k \in \mathbb{Z}^d} \mu(E+k)$, where $E \subseteq \mathbb{R}^d$ is any Borel set such that $\tilde{E} = \{\tilde{x} \mid x \in E\}$.
(3) The expectation operator has lower precedence than multiplication. That is, if X, Y are random variables then $\boldsymbol{E} XY$ denotes the expectation of the product $\boldsymbol{E}(XY)$, and $\boldsymbol{E} X^2$ denotes the expectation of the square $\boldsymbol{E}(X^2)$.
(4) When X is a Markov process we will use $\boldsymbol{E}^{\mu} X_n$ to denote the expectation of X_n given $X_0 \sim \mu$. When $x \in \mathbb{R}^d$, we use $\boldsymbol{E}^{x} X_n$ to denote the expectation of X_n given $X_0 = x$. For random variables that are not associated to Markov processes, we will denote their expectation by $\boldsymbol{E}$, without any sub or superscripts.

3. Proof of Theorem 1.1

To prove Theorem 1.1 we first need to show that the process X^{ε} remains subgaussian, with norm controlled independently of ε. Recall [Ver18] the *subgaussian norm* of a random variable, denoted by $\|\cdot\|_{\psi_2}$, is defined by

$$\|Y\|_{\psi_2} \stackrel{\text{def}}{=} \inf\bigl\{c > 0 \bigm| \boldsymbol{E} e^{|Y|^2/c^2} \leqslant 2\bigr\} \,. \tag{3.1}$$

Lemma 3.1. *There exists a constant $\Lambda > 0$ such that for all $n \in \mathbb{N}$, and all $\varepsilon \in [0,1]$ we have*

$$\|X^{\varepsilon}_n\|_{\psi_2} \leqslant \frac{\Lambda}{2} \bigl(\|X_0\|_{\psi_2} + n\bigr) \,. \tag{3.2}$$

Note that even though Lemma 3.1 gives an upper bound on $\|X^{\varepsilon}_n\|_{\psi_2}$ that is ε-independent, it is more crude than Theorem 1.1. Indeed, the bound (3.2) implies the quadratic upper bound $\operatorname{var}(v \cdot X^{\varepsilon}_n) \leqslant C|v|^2 n^2$, whereas the conclusion of Theorem 1.1 involves bounds (both upper and lower) that are linear in n.

The next Lemma is a special case of Theorem 1.1 when the initial distribution is uniform on Q_0.

Lemma 3.2. *If $\pi_0 = \operatorname{unif}(Q_0)$ denotes the uniform distribution on Q_0 then*

$$\Bigl| \lim_{n \to \infty} \frac{\operatorname{var}^{\pi_0}(v \cdot X^{\varepsilon}_n)}{n} - \operatorname{var}^{\pi_0}\bigl(v \cdot \lfloor X^0_1 \rfloor\bigr) \Bigr| \leqslant C|v|^2 \sqrt{\varepsilon |\ln \varepsilon|^3} \,.$$

Lemma 3.2 is the only place in the proof of Theorem 1.1 which relies on the Bernoulli structure of φ. Even if $\tilde{\varphi}$ is not an expanding Bernoulli map then one can still show that the asymptotic variance of $v \cdot X^\varepsilon$ starting from any (subgaussian) initial distribution equals the asymptotic variance of $v \cdot X^\varepsilon$ starting form π_0. Computing this for $\varepsilon > 0$ (or even for $\varepsilon = 0$), however, is not easy in general. In our situation we compute it because Assumption 2.1 makes the increments $\lfloor \varphi(X_n^\varepsilon) \rfloor - \lfloor X_n^\varepsilon \rfloor$ independent (see Lemma 4.1, below).

Momentarily postponing the proof of Lemmas 3.1 and 3.2 we will now prove Theorem 1.1.

Proof of Theorem 1.1. Before delving into the technical details of the proof, we will first explain the main idea. Fix $\varepsilon \in (0,1)$, and let f_n denote the density of X_n^ε on $\mathbb{R}^d$. We will begin by choosing

$$m = \lfloor n^{1/4} \rfloor \,, \tag{3.3}$$

and finding a probability density function g_m such that

$$\tilde{g}_m = 1 \qquad \text{and} \qquad \|f_m - g_m\|_{L^1(\mathbb{R}^d)} = \|\tilde{f}_m - 1\|_{L^1(\mathbb{T}^d)} \,. \tag{3.4}$$

Once we have g_m, we define two new (coupled) Markov processes Y_k, Y_k', for $k \in \{m, \dots, n\}$, using the same evolution rule (1.1), the same noise, but different initial distributions. We specify the initial distributions for Y and Y' at time m by

$$Y_m \sim g_m \quad \text{and} \quad Y_m' \sim \pi_0 \,,$$

respectively. Since $\tilde{g}_m = 1$, Y_m and Y_m' differ by an element of the integer lattice $\mathbb{Z}^d$, and the periodic structure of φ will preserve this difference at all later times. This combined with Lemma 3.2 will allow us to show

$$\left| \operatorname{var}(v \cdot Y_n) - \operatorname{var}(v \cdot Y_n') \right| \leqslant |v|^2 o(n) \,. \tag{3.5}$$

The process $\tilde{X}^\varepsilon$ is a Markov process on the torus $\mathbb{T}^d$ with stationary distribution $\tilde{\pi}$, the uniform on $\mathbb{T}^d$. Additionally, available results [FI19, ILN24, IZ23] can be used to show that $\tilde{X}^\varepsilon$ is mixing on $\mathbb{T}^d$. In fact, as we will see shortly, the mixing time of $\tilde{X}^\varepsilon$ is at most $O(|\ln \varepsilon|^3)$. Combined with (3.3) and (3.4), this will show that $\|f_m - g_m\|_{L^1}$ is small when n is large.

Finally, using subgaussianity and the fact that any Markov evolution induces a contraction on the laws, we will show

$$|\operatorname{var}(v \cdot X_n^\varepsilon) - \operatorname{var}(v \cdot Y_n)| \xrightarrow{n \to \infty} 0 \,. \tag{3.6}$$

Combining (3.5), (3.6) and Lemma 3.2 will conclude the proof of Theorem 1.1.

We will now prove each of the above claims. Moreover, the proofs of (3.5) and (3.6) will be quantitative and we will obtain an explicit rate at which the right hand sides vanish as $n \to \infty$.

Step 1: Constructing the function g_m. Order the elements of $\mathbb{Z}^d$ as $\{0 = k_0, k_1, \dots\}$. Fix $x \in Q_0$, and define

$$\ell_0 = \ell_0(x) = \inf \Big\{ \ell \in \mathbb{N} \,\Big|\, \sum_{j=0}^{\ell} f(x + k_j) > 1 \Big\} \,.$$

If $\ell_0 = \infty$, then for every $\ell \in \mathbb{N}$ we define

$$g_m(x + k_\ell) = \begin{cases} f_m(x + k_\ell) & \ell > 0 \\ f_m(x + k_0) + 1 - \sum_{\ell' \in \mathbb{N}} f_m(x + k_{\ell'}) & \ell = 0 \,. \end{cases}$$

If $\ell_0 < \infty$, then we define

$$g_m(x + k_\ell) = \begin{cases} f_m(x + k_\ell) & \ell < \ell_0 \\ 1 - \sum_{\ell' < \ell_0} f_m(x + k_{\ell'}) & \ell = \ell_0 \\ 0 & \ell > \ell_0 \,. \end{cases}$$

Clearly, for every $x \in \mathbb{T}^d$ we have

$$\tilde{g}_m(x) = \sum_{\ell=0}^{\infty} g_m(x + k_\ell) = \sum_{k \in \mathbb{Z}^d} g_m(x + k) = 1 \,.$$

Moreover, when $\ell_0(x) = \infty$ we note $g_m(x + k) \geqslant f_m(x + k)$ for all $k \in \mathbb{Z}^d$. When $\ell_0(x) < \infty$, we note $0 \leqslant g_m(x + k) \leqslant f_m(x + k)$ for all $k \in \mathbb{Z}^d$. Thus, in both cases, the sign of $g_m(x+k) - f_m(x+k)$ does not change with k and we have

$$\begin{aligned} \sum_{k \in \mathbb{Z}^d} |g_m(x + k) - f_m(x + k)| &= \Big| \sum_{k \in \mathbb{Z}^d} (g_m(x + k) - f_m(x + k)) \Big| \\ &= \Big| 1 - \sum_{k \in \mathbb{Z}} f_m(x + k) \Big| = |1 - \tilde{f}_m(x)|. \end{aligned}$$

So integrating over $x \in \mathbb{T}^d$ shows $\|f_m - g_m\|_{L^1(\mathbb{R}^d)} = \|\tilde{f}_m - 1\|_{L^1(\mathbb{T}^d)}$. Thus g_m is a probability density function that satisfies both conditions in (3.4), as desired. Observe that this construction with the choice of $k_0 = 0$ guarantees that

$$g_m(x) \leqslant f_m(x), \quad \text{for all } x \notin Q_0.$$

Step 2: Proof of the bound (3.5). Once we have g_m, we define two new (coupled) Markov processes Y_k, Y'_k, for $k \in \{m, \dots, n\}$, as described above. These processes Y and Y' use the same evolution rule (1.1) as for X (for times $k \in \{m, \dots, n\}$). Moreover, Y and Y are couplied, using the same noise for both processes; they differ only in the initial condition (at time m). Specifically, let $Y_m \sim g_m$, and then define Y'_m by

$$Y'_m = Y_m - I_m,$$

where

$$I_m = \lfloor Y_m \rfloor = \sum_{k \in \mathbb{Z}^d} k \mathbf{1}_{\{Y_m \in Q_k\}} \,.$$

In particular, $Y'_m \sim \pi_0$, since $\tilde{g}_m = 1$. For $k \in \{m, \dots, n-1\}$, define $Y_{k+1} = \varphi(Y_k) + \varepsilon \xi_{k+1}$ and $Y'_{k+1} = \varphi(Y'_k) + \varepsilon \xi_{k+1}$. Because of this coupling and because I_m is integer valued, we have

$$Y_k = Y'_k + I_m \,, \quad \forall \;\; k \in \{m, \dots, n\}.$$

Thus

$$|\text{var}(v \cdot Y_n) - \text{var}(v \cdot (Y'_n))| \leqslant \text{var}(v \cdot I_m) + 2|\text{cov}(v \cdot Y'_n, v \cdot I_m)| \,. \tag{3.7}$$

By Lemma 3.1 we know

$$\text{var}(v \cdot I_m) \leqslant C|v|^2 m^2 \,. \tag{3.8}$$

Moreover, since $Y'_m \sim \operatorname{unif}(Q_0)$, Lemma 3.2 implies there exists $N_0 = N_0(\varepsilon)$ such that

$$\operatorname{var}(v \cdot Y'_n) \leqslant C|v|^2(n-m) \leqslant C|v|^2 n\,, \tag{3.9}$$

for all $n \geqslant N_0$. Using (3.8) and (3.9) in (3.7) immediately implies

$$|\operatorname{var}(v \cdot Y_n) - \operatorname{var}(v \cdot (Y'_n))| \leqslant C|v|^2(m^2 + m\sqrt{n}) \leqslant C|v|^2 n^{3/4}\,,$$

for all $n \geqslant N_0$. This proves (3.5) as desired.

Step 3: Proof of the bound (3.6)*.* At time m, the density of X_m is f_m, and the density of Y_m is g_m. For $k \in \{m, \dots, n\}$, both processes evolve according to the same transition probabilities, and we wish to compare their variances at time n. Let h_n be the density of Y_n, and note

$$\begin{aligned}|\operatorname{var}(v \cdot X_n) - \operatorname{var}(v \cdot Y_n)| \leqslant \int_{\mathbb{R}^d} (v\cdot x)^2 |f_n(x) - h_n(x)|\,dx \\ + \Bigl(\int_{\mathbb{R}^d} |v \cdot x||f_n(x) - h_n(x)|\,dx\Bigr)\Bigl(\boldsymbol{E}(|v\cdot X_n| + |v \cdot Y_n|)\Bigr)\,.\end{aligned} \tag{3.10}$$

We will bound each term on the right hand side by splitting the integral into two parts, one where $|x| < n^{3/2}$ and the other where $|x| \geqslant n^{3/2}$.

Step 3.1: First term in (3.10) *when $x \geqslant n^{3/2}$.* We first claim that there exists m_0 (depending only on the dimension), such that for all $m \geqslant m_0$ we have

$$\|Y_m\|_{\psi_2} \leqslant \Lambda(\|X_0\|_{\psi_2} + m)\,. \tag{3.11}$$

To see this, we note that the construction of g_m guarantees

$$g_m(x) \leqslant f_m(x) \quad \text{for all } x \notin Q_0\,,$$

and hence

$$\boldsymbol{P}(|Y_m| \geqslant t) \leqslant \boldsymbol{P}(|X_m| \geqslant t) \quad \text{for all } t \geqslant \sqrt{d}\,.$$

The definition of the subgaussian norm (3.1) implies

$$\boldsymbol{P}(|X_m| \geqslant t) \leqslant 2e^{-t^2/\|X_m\|_{\psi_2}^2}\,,$$

and hence for any $\alpha > \|X_m\|_{\psi_2}$ we have

$$\begin{aligned}\boldsymbol{E} e^{|Y_m|^2/\alpha^2} &= 1 + \int_0^\infty \frac{2t}{\alpha^2} e^{t^2/\alpha^2} \boldsymbol{P}(|Y_m| \geqslant t)\,dt \\ &\leqslant 1 + \int_0^{\sqrt{d}} \frac{2t}{\alpha^2} e^{t^2/\alpha^2}\,dt + \int_{\sqrt{d}}^\infty \frac{2t}{\alpha^2} e^{t^2/\alpha^2} \boldsymbol{P}(|X_m| \geqslant t)\,dt \\ &\leqslant e^{d/\alpha^2} + \int_0^\infty \frac{4t}{\alpha^2} \exp\Bigl(-t^2\Bigl(\frac{1}{\|X_m\|_{\psi_2}^2 - \frac{1}{\alpha^2}}\Bigr)\Bigr)\,dt \\ &\leqslant e^{d/\alpha^2} + \frac{2\|X_m\|_{\psi_2}^2}{\alpha^2 - \|X_m\|_{\psi_2}^2}\,.\end{aligned}$$

Choosing

$$\alpha = \Lambda(\|X_0\|_{\psi_2} + m)$$

and using Lemma 3.1 we see

$$\boldsymbol{E} e^{|Y_m|^2/\alpha^2} \leqslant 2\,, \quad \text{provided} \quad \Lambda^2(\|X_0\|_{\psi_2} + m)^2 \geqslant \frac{d}{\ln(4/3)}\,.$$

This proves (3.11), as claimed.

Now, since the processes Y_k and X_k evolve by the same transition probability for $k \in \{m, \dots, n\}$, Lemma 3.1 and (3.11) imply

$$\|Y_n\|_{\psi_2} \leqslant \frac{\Lambda}{2}(\|Y_m\|_{\psi_2} + (n-m)) \leqslant \frac{\Lambda^2}{2}(\|X_0\|_{\psi_2} + n) \leqslant \Lambda^2 n\,,$$

provided $n \geqslant \|X_0\|_{\psi_2}$. This implies

$$\boldsymbol{P}(|Y_n| > t) \leqslant 2e^{-t^2/(n^2\Lambda^4)} \quad \text{for all } t \geqslant 0\,.$$

Thus for any $R > 0$ we have

$$\begin{aligned}\boldsymbol{E}\big(\mathbf{1}_{\{|Y_n|>R\}}|Y_n|^2\big) &= \int_R^\infty 2t\boldsymbol{P}(|Y_n| > t)\,dt \leqslant \int_R^\infty 4te^{-t^2/(n^2\Lambda^4)}\,dt \\ &= 2n^2\Lambda^4 e^{-R^2/(n^2\Lambda^4)}\,.\end{aligned} \tag{3.12}$$

By the same argument and Lemma 3.1, we also obtain

$$\boldsymbol{E}\big(\mathbf{1}_{\{|X_n^\varepsilon|>R\}}|X_n^\varepsilon|^2\big) \leqslant 2n^2\Lambda^2 e^{-R^2/(n^2\Lambda^2)} \leqslant 2n^2\Lambda^4 e^{-R^2/(n^2\Lambda^4)}\,. \tag{3.13}$$

Choosing $R = n^{3/2}$ and combining (3.12) and (3.13) and shows

$$\int_{|x|\geqslant n^{3/2}} (v\cdot x)^2 |f_n(x) - h_n(x)|\,dx \leqslant |v|^2 4\Lambda^2 n^2 e^{-n/\Lambda^4} \tag{3.14}$$

for all $n \geqslant \|\mu_0\|_{\psi_2}$.

Step 3.2: First term in (3.10) *when $x < n^{3/2}$.* We will show this term vanishes by showing $\|f_n - h_n\|_{L^1}$ vanishes exponentially as $n \to \infty$. Let $\tilde t^\varepsilon_{\text{mix}} = \tilde t^\varepsilon_{\text{mix}}(1/2)$ denote the mixing time of $\tilde X^\varepsilon$. That is, if $\tilde p^\varepsilon_k$ is the k-step transition density of $\tilde X^\varepsilon$, then $\tilde t^\varepsilon_{\text{mix}}$ is the smallest $k \in \mathbb{N}$ for which

$$\sup_{\tilde x \in \mathbb{T}^d} \frac{1}{2}\|\tilde p^\varepsilon_k(x,\cdot) - 1\|_{L^1(\mathbb{T}^d)} \leqslant \frac{1}{4}\,.$$

Using the fact that a Markov process induces an L^1 contraction on the density, and the fact that (see for instance Section 4.5 in [LP17])

$$\sup_{x\in\mathbb{T}^d} \frac{1}{2}\|\tilde p^\varepsilon_m(x,\cdot) - 1\|_{L^1(\mathbb{T}^d)} \leqslant 2^{-\lfloor m/\tilde t^\varepsilon_{\text{mix}}\rfloor}$$

we see

$$\begin{aligned}\int_{|x|\leqslant n^{3/2}} (v\cdot x)^2|f_n(x) - h_n(x)|\,dx &\leqslant C|v|^2 n^3\|f_n - h_n\|_{L^1} \leqslant C|v|^2 n^3 \|f_m - h_m\|_{L^1} \\ &= C|v|^2n^3\|f_m - g_m\|_{L^1} = C|v|^2 n^3\|\tilde f_m - 1\|_{L^1} \\ &\leqslant C|v|^2 n^3 2^{-\lfloor m/\tilde t^\varepsilon_{\text{mix}}\rfloor}\,.\end{aligned} \tag{3.15}$$

Step 3.3: Second integral in (3.10). Using the same argument as in the previous two steps we see

$$\begin{aligned}\Big|\int_{|x|\geqslant n^{3/2}} (x\cdot v)(f_n(x) - h_n(x))\,dx\Big| &\leqslant 4|v|\int_{n^{3/2}}^\infty e^{-t^2/(n^2\Lambda^4)}\,dt \\ &\leqslant 4|v|\Lambda^4\sqrt{n}e^{-n/\Lambda^4}\end{aligned} \tag{3.16}$$

and

$$\Big|\int_{|x|\leqslant n^{3/2}} (x\cdot v)(f_n(x) - h_n(x))\,dx\Big| \leqslant C|v|n^{3/2}2^{-\lfloor m/\tilde t^{\varepsilon}_{\mathrm{mix}}\rfloor}\,. \tag{3.17}$$

Moreover, by Lemma 3.1 and (3.11) we see

$$\boldsymbol{E}|X^\varepsilon_n| \leqslant \|X^\varepsilon_n\|_{\psi_2} \leqslant \frac{\Lambda}{2}\big(\|X_0\|_{\psi_2} + n\big)\,, \tag{3.18}$$

$$\text{and}\quad \boldsymbol{E}|Y_n| \leqslant \|Y_n\|_{\psi_2} \leqslant \Lambda\big(\|X_0\|_{\psi_2} + n\big)\,. \tag{3.19}$$

Using (3.14)–(3.19) in (3.10) and increasing N_0 if necessary, and recalling (3.3) we obtain (3.6), concluding Step 3.

Step 4: Combining (3.5) and (3.6) we obtain

$$\lim_{n\to\infty}\frac{\operatorname{var}(v\cdot X_n)}{n} = \lim_{n\to\infty}\frac{\operatorname{var}(v\cdot Y_n)}{n} = \lim_{n\to\infty}\frac{\operatorname{var}(v\cdot Y'_n)}{n}\,,$$

and using Lemma 3.2 concludes the proof. □

It remains to prove Lemmas 3.1 and 3.2. The proof of Lemma 3.1 follows quickly from Hoeffding's inequality, and we present it here. The proof of Lemma 3.2 is more involved and we present it in Section 4, below.

Proof of Lemma 3.1. Notice first

$$X^\varepsilon_{n+1} - X^\varepsilon_n = \varphi(X^\varepsilon_n) - X^\varepsilon_n + \varepsilon\xi_{n+1} \leqslant A + \varepsilon|\xi_{n+1}|\,,$$

where

$$A \stackrel{\text{def}}{=} \sup_{x\in\mathbb{R}^d}|\varphi(x) - x|\,.$$

Hence

$$|X^\varepsilon_n| \leqslant |X^\varepsilon_0| + An + \varepsilon\Xi_n\,,\quad \text{where}\quad \Xi_n \stackrel{\text{def}}{=} \sum_1^n|\xi_k|\,.$$

Inequality (3.2) now follows from Hoeffding's inequality (see for instance Theorem 2.6.2 in [Ver18]). □

4. Proof of Lemma 3.2

To prove Lemma 3.2 we write

$$X^\varepsilon_n = S_n + J_n + R_n \tag{4.1}$$

where

$$\begin{aligned} S_n &= S_{n-1} + \lfloor\varphi(X^\varepsilon_{n-1})\rfloor - \lfloor X^\varepsilon_{n-1}\rfloor\,,\\ J_n &= J_{n-1} + \lfloor X^\varepsilon_n\rfloor - \lfloor\varphi(X^\varepsilon_{n-1})\rfloor\,,\\ R_n &= X^\varepsilon_n - S_n - J_n\,, \end{aligned}$$

with $S_0 = 0$, $J_0 = 0$, $R_0 = X^\varepsilon_0$. We will prove Lemma 3.2 by obtaining a bound on the asymptotic variances of each of the processes S, J and R. The Bernoulli structure of φ implies the increments of S are independent, which allows us to compute the variance of S exactly.

Lemma 4.1. *For any $v\in\mathbb{R}^d$, $n\geqslant 0$ we have*

$$\operatorname{var}^{\pi_0}(v\cdot S_n) = n\operatorname{var}^{\pi_0}(v\cdot\lfloor X^0_1\rfloor)\,. \tag{4.2}$$

To bound the asymptotic variance of J, we rely on the fact that the mixing time of $\tilde{X}^\varepsilon$ is $o(1/\varepsilon)$, which only requires $\tilde{\varphi}$ to be sufficiently mixing, and does not rely on the Bernoulli structure.

Proposition 4.2. *If $\tilde{t}^\varepsilon_{\text{mix}} = \tilde{t}^\varepsilon_{\text{mix}}(1/2)$ denotes the mixing time of $\tilde{X}^\varepsilon$, then for all $v \in \mathbb{R}^d$ and all $\varepsilon \in (0,1)$ we have*

$$\operatorname{var}^{\pi_0}(v \cdot J_n) \leqslant C|v|^2 n\varepsilon \tilde{t}^\varepsilon_{\text{mix}} . \tag{4.3}$$

Momentarily postponing the proofs of Lemma 4.1 and Proposition 4.2, we now prove Lemma 3.2.

Proof of Lemma 3.2. For any $v \in \mathbb{R}^d$, (4.1) implies

$$\begin{aligned} |\operatorname{var}^{\pi_0}(v \cdot X^\varepsilon_n) - \operatorname{var}^{\pi_0}(v \cdot S_n)| \leqslant C\Big(&\operatorname{var}^{\pi_0}(v \cdot J_n) + \operatorname{var}^{\pi_0}(v \cdot R_n) \\ &+ \operatorname{std}^{\pi_0}(v \cdot S_n)\operatorname{std}^{\pi_0}(v \cdot J_n) + \operatorname{std}^{\pi_0}(v \cdot S_n)\operatorname{std}^{\pi_0}(v \cdot R_n)\Big) . \end{aligned} \tag{4.4}$$

Note

$$R_n - R_{n-1} = X_n - \lfloor X_n \rfloor - (X_{n-1} - \lfloor X_{n-1} \rfloor)$$

and hence

$$R_n - R_0 = X_n - \lfloor X_n \rfloor - (X_0 - \lfloor X_0 \rfloor) \in (-1,1)^d .$$

This implies $\operatorname{var}^{\pi_0}(R_n) \leqslant C$. Using this, Lemma 4.1 and Proposition 4.2 in (4.4) implies

$$\lim_{n\to\infty} \frac{|\operatorname{var}^{\pi_0}(v \cdot X^\varepsilon_n) - \operatorname{var}^{\pi_0}(v \cdot S_n)|}{n} \leqslant C|v|^2 \left(\varepsilon \tilde{t}^\varepsilon_{\text{mix}} + \sqrt{\varepsilon \tilde{t}^\varepsilon_{\text{mix}}} \right) . \tag{4.5}$$

To finish the proof we only a mixing time estimate which shows (2.3) holds. This can be done quickly from existing results. Indeed, the exponentially mixing condition in Assumption 2.1 allows us to use Corollary 2.5 in [FI19] to show that the dissipation time (aka the L^2 mixing time) of $\tilde{X}^\varepsilon$ is at most $O(|\ln \varepsilon|^2)$. Following this Proposition 1.3 in [ILN24] (or Proposition 2.2 in [IZ23]) implies

$$\tilde{t}^\varepsilon_{\text{mix}} \leqslant C|\ln \varepsilon|^3 .$$

This immediately implies (2.3) and using this and Lemma 4.1 in (4.5) finishes the proof. ◻

The remainder of this section is devoted to the proof of Lemma 4.1 and Proposition 4.2

4.1. **Variance of S (Lemma 4.1).** The proof of Lemma 4.1 relies on the Bernoulli structure of φ to show that the increments of S are independent.

Proof of Lemma 4.1. Define

$$D_n = S_{n+1} - S_n = \lfloor \varphi(X^\varepsilon_n) \rfloor - \lfloor X^\varepsilon_n \rfloor .$$

By Assumption 2.1, for every $\ell \in \mathbb{Z}^d$, the event $\{D_n = \ell\}$ can be partitioned into events of the form $\{\tilde{X}^\varepsilon_n \in \tilde{E}_i\}$. We will now show that

$$\boldsymbol{P}^{\pi_0}(D_n = k \mid \tilde{X}^\varepsilon_{n-1} \in \tilde{E}_i) = \boldsymbol{P}^{\pi_0}(D_n = k) = \boldsymbol{P}^{\pi_0}\big(\lfloor X^0_1 \rfloor = k\big) . \tag{4.6}$$

We will first show the last equality in (4.6). For this, we recall $\tilde{\pi}$ is the Lebesgue measure on $\mathbb{T}^d$, which is the stationary distribution for the Markov process $\tilde{X}^\varepsilon$.

Since $X_0 \sim \pi_0$ by assumption, and $\tilde\pi_0 = \tilde\pi$, we must have $\tilde X^\varepsilon_n \sim \tilde\pi$ for every $n \in \mathbb{N}$. Thus, the right hand side of (4.6) is given by

$$\begin{aligned}\boldsymbol{P}^{\pi_0}(D_n = k) &= \boldsymbol{P}\big(\lfloor \varphi(U)\rfloor - \lfloor U\rfloor = k \bigm| \tilde U \sim \tilde\pi\big)\\ &= \boldsymbol{P}^{\pi_0}\big(\lfloor X^0_1\rfloor = k\big)\,,\end{aligned}$$

proving the last equality in (4.6) as desired.

We now compute the left hand side of (4.6) and show it equals the right hand side of (4.6). Since $\tilde X^\varepsilon_n \sim \tilde\pi$, Assumption 2.1 implies that conditioned on the event $\tilde X^\varepsilon_{n-1} \in \tilde E_i$, the distribution of $\varphi(\tilde X^\varepsilon_{n-1})$ is also $\tilde\pi$. Since ξ_n is independent of X^ε_{n-1}, this in turn implies that conditioned on the event $\tilde X^\varepsilon_{n-1} \in \tilde E_i$, the distribution of $\tilde X^\varepsilon_n$ is the uniform distribution on $\mathbb{T}^d$. Hence

$$\boldsymbol{P}^{\pi_0}(D_n = k \mid \tilde X^\varepsilon_{n-1} \in \tilde E_i) = \boldsymbol{P}^{\pi_0}(D_n = k \mid \tilde X^\varepsilon_n \sim \tilde\pi) = \boldsymbol{P}^{\pi_0}\big(\lfloor X^0_1\rfloor = k\big)\,,$$

which concludes the proof of (4.6).

Now, (4.6) implies the increments of S are i.i.d. with distribution

$$\boldsymbol{P}^{\pi_0}(\Delta_n = k) = \boldsymbol{P}(\lfloor X^0_1\rfloor = k)\,.$$

This immediately implies (4.2), completing the proof. □

4.2. **Decorrelation bound for J (Proposition 4.2).** For any $n \in \mathbb{N}$ define

$$\Delta_n = J_{n+1} - J_n = \lfloor X^\varepsilon_{n+1}\rfloor - \lfloor \varphi(X^\varepsilon_n)\rfloor\,.$$

We first claim all moments of Δ_0 are of order ε.

Lemma 4.3. *For all $\varepsilon \leqslant 1$, and all $p \in [1,\infty)$ we have*

$$\boldsymbol{E}^{\pi_0}|\Delta_0|^p \leqslant C_p\varepsilon\,. \tag{4.7}$$

Next, we claim that the increments Δ_n decorrelate rapidly.

Lemma 4.4. *For any $v \in \mathbb{R}^d$, $m, n \in \mathbb{N}$ we have*

$$|\operatorname{cov}^{\pi_0}(v\cdot\Delta_m, v\cdot\Delta_{m+n+1})| \leqslant C\varepsilon|v|^2 \sup_{\tilde x \in \mathbb{T}^d} \|\tilde p^\varepsilon_n(\tilde x,\cdot) - 1\|_{L^1(\mathbb{T}^d)}\,, \tag{4.8}$$

where as before $\tilde p^\varepsilon_n$ is the n-step transition density of the process $\tilde X^\varepsilon_n$.

Momentarily postponing the proofs of Lemmas 4.3 and 4.4, we prove Proposition 4.2.

Proof of Proposition 4.2. Note

$$\begin{aligned}\operatorname{var}^{\pi_0}(v\cdot J_N) &= \operatorname{var}^{\pi_0}\Big(\sum_{n=0}^{N-1} v\cdot\Delta_n\Big)\\ &= \sum_{n=0}^{N-1}\operatorname{var}^{\pi_0}(v\cdot\Delta_n) + 2\sum_{m=0}^{N-1}\sum_{n=0}^{N-m-1}\operatorname{cov}^{\pi_0}(v\cdot\Delta_m, v\cdot\Delta_{m+n+1})\\ &\leqslant N|v|^2\boldsymbol{E}^{\pi_0}[|\Delta_0|^2 + |\Delta_0|]\Big(1 + C\sum_{n=1}^\infty \sup_{\tilde x\in\mathbb{T}^d}\|\tilde p_n(\tilde x,\cdot) - 1\|_{L^1(\mathbb{T}^d)}\Big)\,. \qquad (4.9)\end{aligned}$$

Now let $T = \tilde t^\varepsilon_{\text{mix}}$, and observe that for every $\tilde x \in \mathbb{T}^d$, $n \in \mathbb{N}$ and $j \in \{0, \dots, T-1\}$, we have

$$\|\tilde p_{nT+j}(\tilde x,\cdot) - 1\|_{L^1(\mathbb{T}^d)} \leqslant \frac{1}{2^n}\,.$$

Thus the series on the right hand side of (4.9) is bounded by $2T$. Combining this with Lemmas 4.3 and 4.4, we obtain (4.3) as desired. $\square$

It remains to prove Lemmas 4.3 and 4.4. We begin with the proof of Lemma 4.4.

Proof of Lemma 4.4. We first claim for any $n \in \mathbb{N}$,

$$\boldsymbol{E}^{\pi_0} \Delta_n = \boldsymbol{E}^{\pi_0} \Delta_0 \,. \tag{4.10}$$

Clearly, the Markov property implies

$$\boldsymbol{E}^{\pi_0} \Delta_n = \boldsymbol{E}^{\pi_0} \boldsymbol{E}^{X_n^\varepsilon} \Delta_0 = \boldsymbol{E}^{\pi_n^\varepsilon} \Delta_0 \,, \tag{4.11}$$

where $\pi_n^\varepsilon = \operatorname{dist}(X_n^\varepsilon)$. Now we note note that adding an element of $\mathbb{Z}^d$ to X_0^ε does not change Δ_n for any $n \in \mathbb{N}$. Thus, if μ_0, ν_0 are any two probability measures such that $\tilde{\mu}_0 = \tilde{\nu}_0$, then for any $n \in \mathbb{N}$,

$$\boldsymbol{E}^{\mu_0} \Delta_n = \boldsymbol{E}^{\nu_0} \Delta_n \,. \tag{4.12}$$

Since the Lebesgue measure $\tilde{\pi}$ is the stationary distribution of $\tilde{X}^\varepsilon$, we note $\tilde{\pi}_n^\varepsilon = \tilde{\pi}$. Hence (4.11) and (4.12) imply (4.10) as desired.

Now using the Markov property and (4.10) we compute

$$\begin{aligned}
\operatorname{cov}^{\pi_0}(v \cdot \Delta_m, v \cdot \Delta_{m+n+1}) &= \boldsymbol{E}^{\pi_0} v \cdot (\Delta_m - \boldsymbol{E}^{\pi_0} \Delta_0)\, v \cdot (\Delta_{m+n+1} - \boldsymbol{E}^{\pi_0} \Delta_0) \\
&= \boldsymbol{E}^{\pi_0} \boldsymbol{E}^{X_m^\varepsilon} v \cdot (\Delta_0 - \boldsymbol{E}^{\pi_0} \Delta_0)\, v \cdot (\Delta_{n+1} - \boldsymbol{E}^{\pi_0} \Delta_0) \\
&= \boldsymbol{E}^{\pi_m^\varepsilon} v \cdot (\Delta_0 - \boldsymbol{E}^{\pi_0} \Delta_0)\, v \cdot (\Delta_{n+1} - \boldsymbol{E}^{\pi_0} \Delta_0) \\
&= \boldsymbol{E}^{\pi_0} v \cdot (\Delta_0 - \boldsymbol{E}^{\pi_0} \Delta_0)\, v \cdot (\Delta_{n+1} - \boldsymbol{E}^{\pi_0} \Delta_0) \,,
\end{aligned} \tag{4.13}$$

where the last inequality followed because $\tilde{\pi}_m^\varepsilon = \tilde{\pi}$.

Next we define the function $f \colon \mathbb{R}^d \to \mathbb{R}$ by

$$f(x) \stackrel{\text{def}}{=} \boldsymbol{E}^x v \cdot (\Delta_0 - \boldsymbol{E}^{\pi_0} \Delta_0) = \boldsymbol{E}^x (v \cdot \big(\lfloor X_1^\varepsilon \rfloor - \lfloor \varphi(X_0^\varepsilon) \rfloor - \boldsymbol{E}^{\pi_0} \Delta_0\big) \,. \tag{4.14}$$

The Markov property and (4.13) imply

$$\operatorname{cov}^{\pi_0}(v \cdot \Delta_m, v \cdot \Delta_{m+n+1}) = \boldsymbol{E}^{\pi_0} \big[v \cdot (\Delta_0 - \boldsymbol{E}^{\pi_0} \Delta_0) P_n^\varepsilon f(X_1^\varepsilon) \big] \,. \tag{4.15}$$

Here P_n^ε is the n-step transition operator, whose action of functions is given by

$$P_n^\varepsilon g(x) = \boldsymbol{E}^x g(X_n^\varepsilon) = \int_{\mathbb{R}^d} p_n^\varepsilon(x, y) g(y) \, dy \,,$$

where p_n^ε is the n-step transition density of X^ε.

Note that for $\tilde{x}, \tilde{y} \in \mathbb{T}^d$, the n-step transition density of $\tilde{X}^\varepsilon$ is given by

$$\tilde{p}_n^\varepsilon(\tilde{x}, \tilde{y}) = \sum_{k \in \mathbb{Z}^d} p_n^\varepsilon(x, y + k) \,.$$

Thus if $\tilde{P}_n^\varepsilon$ is the n-step transition operator of $\tilde{X}^\varepsilon$, then the action of $\tilde{P}_n^\varepsilon$ on functions $\tilde{g} \colon \mathbb{T}^d \to \mathbb{R}$ is given by

$$\begin{aligned}
\tilde{P}_n^\varepsilon \tilde{g}(\tilde{x}) &\stackrel{\text{def}}{=} \boldsymbol{E}^{\tilde{x}} \tilde{g}(\tilde{X}_n^\varepsilon) = \int_{\mathbb{T}^d} \tilde{p}_n^\varepsilon(\tilde{x}, \tilde{y}) \tilde{g}(\tilde{y}) \, d\tilde{y} \\
&= \int_{\mathbb{R}^d} p_n^\varepsilon(x, y + k) g(y) \, dy = P_n^\varepsilon g(x) \,.
\end{aligned} \tag{4.16}$$

Returning to (4.15), we note that shift invariance of φ implies the function f is $\mathbb{Z}^d$ periodic. Thus, defining $\tilde{f}\colon \mathbb{T}^d \to \mathbb{R}$ by $\tilde{f}(\tilde{x}) = f(x)$ and using (4.15) and (4.16) we see

$$\begin{aligned}
\operatorname{cov}^{\pi_0}(v\cdot\Delta_m, v\cdot\Delta_{m+n+1}) &= \boldsymbol{E}^{\pi_0}\big[v\cdot(\Delta_0 - \boldsymbol{E}^{\pi_0}\Delta_0)\tilde{P}^\varepsilon_n \tilde{f}(\tilde{X}^\varepsilon_1)\big] \\
&= \boldsymbol{E}^{\pi_0}\Big[v\cdot(\Delta_0 - \boldsymbol{E}^{\pi_0}\Delta_0)\Big(\tilde{P}^\varepsilon_n \tilde{f}(\tilde{X}^\varepsilon_1) - \int_{\mathbb{T}^d}\tilde{f}\,d\tilde{\pi}\Big)\Big] \\
&= \boldsymbol{E}^{\pi_0}\Big[v\cdot(\Delta_0 - \boldsymbol{E}^{\pi_0}\Delta_0)\Big(\int_{\mathbb{T}^d}\big(\tilde{p}^\varepsilon_n(\tilde{X}^\varepsilon_1, y) - 1\big)\tilde{f}(y)\,d\tilde{y}\Big)\Big].
\end{aligned}$$

Hence

$$\begin{aligned}
&\big|\operatorname{cov}^{\pi_0}(v\cdot\Delta_m, v\cdot\Delta_{m+n+1})\big| \\
&\qquad\leqslant \boldsymbol{E}^{\pi_0}\big|v\cdot(\Delta_0 - \boldsymbol{E}^{\pi_0}\Delta_0)\big| \sup_{\tilde{x}\in\mathbb{T}^d}\|\tilde{p}^\varepsilon_n(\tilde{X}^\varepsilon_1, y) - 1\|_{L^1(\mathbb{T}^d)}\|\tilde{f}\|_{L^\infty(\mathbb{T}^d)} \\
&\qquad= \|\tilde{f}\|_{L^1(\mathbb{T}^d)} \sup_{\tilde{x}\in\mathbb{T}^d}\|\tilde{p}^\varepsilon_n(\tilde{X}^\varepsilon_1, y) - 1\|_{L^1(\mathbb{T}^d)}\|\tilde{f}\|_{L^\infty(\mathbb{T}^d)}\,. \qquad (4.17)
\end{aligned}$$

It remains to estimate $\|\tilde{f}\|_{L^1}$ and $\|\tilde{f}\|_{L^\infty}$. Using (4.14) and Lemma 4.3 with $p=1$ we see

$$\|\tilde{f}\|_{L^1} \leqslant 2|v|\boldsymbol{E}^{\pi_0}|\Delta_0| \leqslant C\varepsilon\,, \tag{4.18}$$

and

$$\begin{aligned}
\|\tilde{f}\|_{L^\infty} \leqslant \boldsymbol{E} &= |v|\Big(|\boldsymbol{E}^{\pi_0}\Delta_0| + \sup_{x\in Q_0}\boldsymbol{E}\big|\lfloor\varphi(x)+\varepsilon\xi_1\rfloor - \lfloor\varphi(x)\rfloor)\big|\Big) \\
&\leqslant C(1+\varepsilon)|v| \leqslant C|v|\,. \qquad (4.19)
\end{aligned}$$

Using (4.18) and (4.19) in (4.17) implies (4.8), concluding the proof. □

Finally, we prove Lemma 4.3.

Proof of Lemma 4.3. For any $j\in\{1,\dots,d\}$ define

$$\Delta^j_n \stackrel{\text{def}}{=} \boldsymbol{e}_j\cdot\Delta_n\,,$$

and note

$$\boldsymbol{E}|\Delta^j_0|^p = \int_{y\in Q_0}\boldsymbol{E}|\boldsymbol{e}_j\cdot(\lfloor\varphi(y)+\varepsilon\xi_1\rfloor - \lfloor\varphi(y)\rfloor)|^p\,dy\,.$$

Clearly,

$$|\boldsymbol{e}_j\cdot(\lfloor\varphi(y)+\varepsilon\xi_1\rfloor - \lfloor\varphi(y)\rfloor)| \leqslant \mathbf{1}_{\{\boldsymbol{e}_j\cdot\varepsilon\xi\notin[-\psi_j,1-\psi_j)\}} + \varepsilon|\boldsymbol{e}_j\cdot\xi_1|\,,$$

where

$$\psi_j = \boldsymbol{e}_j\cdot(\varphi(y) - \lfloor\varphi(y)\rfloor)\,.$$

Hence,

$$\boldsymbol{E}|\boldsymbol{e}_j\cdot(\lfloor\varphi(y)+\varepsilon\xi_1\rfloor - \lfloor\varphi(y)\rfloor)|^p \leqslant C_p\big(\boldsymbol{P}\big(\boldsymbol{e}_j\cdot\varepsilon\xi\notin[-\psi_j,1-\psi_j)\big) + \varepsilon^p\big) \tag{4.20}$$

When $\psi_j\leqslant 1/2$ a standard Gaussian tail bound implies

$$\boldsymbol{P}\big(\boldsymbol{e}_j\cdot\varepsilon\xi_1\notin[-\psi_j,1-\psi_j)\big) \leqslant \boldsymbol{P}\big(\boldsymbol{e}_j\cdot\varepsilon\xi_1\leqslant-\psi_j\big) \leqslant 2e^{-\psi_j^2/(2\varepsilon^2)}\,.$$

Similarly, when $\psi_j\geqslant 1/2$, we note

$$\boldsymbol{P}\big(\boldsymbol{e}_j\cdot\varepsilon\xi_1\notin[-\psi_j,1-\psi_j)\big) \leqslant \boldsymbol{P}\big(\boldsymbol{e}_j\cdot\varepsilon\xi_1\geqslant 1-\psi_j\big) \leqslant 2e^{-(1-\psi_j)^2/(2\varepsilon^2)}\,.$$

Using these two inequalities and integrating (4.20) in y immediately implies (4.7) as desired. $\square$

References

[Bak11] Y. Bakhtin. Noisy heteroclinic networks. *Probab. Theory Related Fields*, 150(1-2):1–42, 2011. doi:10.1007/s00440-010-0264-0.

[BBPS19] J. Bedrossian, A. Blumenthal, and S. Punshon-Smith. Almost-sure exponential mixing of passive scalars by the stochastic navier-stokes equations, 2019, 1905.03869.

[BCVV95] L. Biferale, A. Crisanti, M. Vergassola, and A. Vulpiani. Eddy diffusivities in scalar transport. *Physics of Fluids*, 7(11):2725–2734, 1995. doi:10.1063/1.868651.

[BCZG23] 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. doi:10.1214/23-aop1627.

[BLP78] A. Bensoussan, J.-L. Lions, and G. Papanicolaou. *Asymptotic analysis for periodic structures*, volume 5 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam-New York, 1978.

[CFIN23] A. Christie, Y. Feng, G. Iyer, and A. Novikov. Speeding up Langevin dynamics by mixing, 2023, 2303.18168.

[CS89] S. Childress and A. M. Soward. Scalar transport and alpha-effect for a family of cat's-eye flows. *J. Fluid Mech.*, 205:99–133, 1989. doi:10.1017/S0022112089001965.

[Dol98] D. Dolgopyat. On decay of correlations in Anosov flows. *Ann. of Math. (2)*, 147(2):357–390, 1998. doi:10.2307/121012.

[ELM23] T. M. Elgindi, K. Liss, and J. C. Mattingly. Optimal enhanced dissipation and mixing for a time-periodic, lipschitz velocity field on $\mathbb{T}^2$, 2023, 2304.05374.

[EZ19] T. M. Elgindi and A. Zlatoš. Universal mixers in all dimensions. *Adv. Math.*, 356:106807, 33, 2019. doi:10.1016/j.aim.2019.106807.

[FI19] Y. Feng and G. Iyer. Dissipation enhancement by mixing. *Nonlinearity*, 32(5):1810–1851, 2019. doi:10.1088/1361-6544/ab0e56.

[FNW04] A. Fannjiang, S. Nonnenmacher, and L. Wołowski. Dissipation time and decay of correlations. *Nonlinearity*, 17(4):1481–1508, 2004. doi:10.1088/0951-7715/17/4/018.

[FP94] A. Fannjiang and G. Papanicolaou. Convection enhanced diffusion for periodic flows. *SIAM J. Appl. Math.*, 54(2):333–408, 1994. doi:10.1137/S0036139992236785.

[FP97] A. Fannjiang and G. Papanicolaou. Convection-enhanced diffusion for random flows. *J. Statist. Phys.*, 88(5-6):1033–1076, 1997. doi:10.1007/BF02732425.

[Fre64] M. I. Freĭdlin. The Dirichlet problem for an equation with periodic coefficients depending on a small parameter. *Teor. Verojatnost. i Primenen.*, 9:133–139, 1964.

[FW03] A. Fannjiang and L. Wołowski. Noise induced dissipation in Lebesgue-measure preserving maps on d-dimensional torus. *J. Statist. Phys.*, 113(1-2):335–378, 2003. doi:10.1023/A:1025787124437.

[FW12] M. I. Freidlin and A. D. Wentzell. *Random perturbations of dynamical systems*, volume 260 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Heidelberg, third edition, 2012. doi:10.1007/978-3-642-25847-3. Translated from the 1979 Russian original by Joseph Szücs.

[Hei03] S. Heinze. Diffusion-advection in cellular flows with large Peclet numbers. *Arch. Ration. Mech. Anal.*, 168(4):329–342, 2003. doi:10.1007/s00205-003-0256-7.

[HIK+18] M. Hairer, G. Iyer, L. Koralov, A. Novikov, and Z. Pajor-Gyulai. A fractional kinetic process describing the intermediate time behaviour of cellular flows. *Ann. Probab.*, 46(2):897–955, 2018. doi:10.1214/17-AOP1196.

[HKPG16] M. Hairer, L. Koralov, and Z. Pajor-Gyulai. From averaging to homogenization in cellular flows—an exact description of the transition. *Ann. Inst. Henri Poincaré Probab. Stat.*, 52(4):1592–1613, 2016. doi:10.1214/15-AIHP690.

[ILN24] G. Iyer, E. Lu, and J. Nolen. Using Bernoulli maps to accelerate mixing of a random walk on the torus. *Quart. Appl. Math.*, 82(2):359–390, 2024. doi:10.1090/qam/1668.

[IZ23] G. Iyer and H. Zhou. Quantifying the dissipation enhancement of cellular flows. *SIAM J. Math. Anal.*, 55(6):6496–6516, 2023. doi:10.1137/22M1524576.

[KLX22] C. Kao, Y.-Y. Liu, and J. Xin. A semi-Lagrangian computation of front speeds of G-equation in ABC and Kolmogorov flows with estimation via ballistic orbits. *Multiscale Model. Simul.*, 20(1):107–117, 2022. doi:10.1137/20M1387699.

[Kor04] L. Koralov. Random perturbations of 2-dimensional Hamiltonian flows. *Probab. Theory Related Fields*, 129(1):37–62, 2004. doi:10.1007/s00440-003-0320-0.

[LP17] D. A. Levin and Y. Peres. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2017. doi:10.1090/mbk/107.

[LWXZ22] J. Lyu, Z. Wang, J. Xin, and Z. Zhang. A convergent interacting particle method and computation of KPP front speeds in chaotic flows. *SIAM J. Numer. Anal.*, 60(3):1136–1167, 2022. doi:10.1137/21M1410786.

[LXY17] J. Lyu, J. Xin, and Y. Yu. Computing residual diffusivity by adaptive basis learning via spectral method. *Numer. Math. Theory Methods Appl.*, 10(2):351–372, 2017. doi:10.4208/nmtma.2017.s08.

[LXY18] J. Lyu, J. Xin, and Y. Yu. Residual diffusivity in elephant random walk models with stops. *Commun. Math. Sci.*, 16(7):2033–2045, 2018. doi:10.4310/CMS.2018.v16.n7.a12.

[MCX+17] N. B. Murphy, E. Cherkaev, J. Xin, J. Zhu, and K. M. Golden. Spectral analysis and computation of effective diffusivities in space-time periodic incompressible flows. *Ann. Math. Sci. Appl.*, 2(1):3–66, 2017. doi:10.4310/AMSA.2017.v2.n1.a1.

[MCZ+20] N. B. Murphy, E. Cherkaev, J. Zhu, J. Xin, and K. M. Golden. Spectral analysis and computation for homogenization of advection diffusion processes in steady flows. *J. Math. Phys.*, 61(1):013102, 34, 2020. doi:10.1063/1.5127457.

[MHSW22] J. Myers Hill, R. Sturman, and M. C. Wilson. Exponential mixing by orthogonal non-monotonic shears. *Physica D: Nonlinear Phenomena*, 434:133224, 2022. doi:10.1016/j.physd.2022.133224.

[MK99] A. J. Majda and P. R. Kramer. Simplified models for turbulent diffusion: Theory, numerical modelling, and physical phenomena. *Phys. Rep.*, 314(4-5):237–574, 1999. doi:10.1016/S0370-1573(98)00083-0.

[NPR05] A. Novikov, G. Papanicolaou, and L. Ryzhik. Boundary layers for cellular flows at high Péclet numbers. *Comm. Pure Appl. Math.*, 58(7):867–922, 2005. doi:10.1002/cpa.20058.

[OTD21] B. W. Oakley, J.-L. Thiffeault, and C. R. Doering. On mix-norms and the rate of decay of correlations. *Nonlinearity*, 34(6):3762–3782, 2021. doi:10.1088/1361-6544/abdbbd.

[Pie94] R. T. Pierrehumbert. Tracer microstructure in the large-eddy dominated regime. *Chaos, Solitons & Fractals*, 4(6):1091–1110, 1994. doi:10.1016/0960-0779(94)90139-2.

[PS08] G. A. Pavliotis and A. M. Stuart. *Multiscale methods – Averaging and homogenization*, volume 53 of *Texts in Applied Mathematics*. Springer, New York, 2008.

[RZ07] L. Ryzhik and A. Zlatoš. KPP pulsating front speed-up by flows. *Commun. Math. Sci.*, 5(3):575–593, 2007. URL `http://projecteuclid.org/euclid.cms/1188405669`.

[SOW06] R. Sturman, J. M. Ottino, and S. Wiggins. *The mathematical foundations of mixing*, volume 22 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, Cambridge, 2006. doi:10.1017/CBO9780511618116.

[Tay21] G. I. Taylor. Diffusion by Continuous Movements. *Proc. London Math. Soc. (2)*, 20(3):196–212, 1921. doi:10.1112/plms/s2-20.1.196.

[Tay53] G. Taylor. Dispersion of soluble matter in solvent flowing slowly through a tube. *Proc. R. Soc. Lond. A*, 219(1137):186–203, 1953. doi:10.1098/rspa.1953.0139.

[TC03] J.-L. Thiffeault and S. Childress. Chaotic mixing in a torus map. *Chaos*, 13(2):502–507, 2003. doi:10.1063/1.1568833.

[Ver18] R. Vershynin. *High-dimensional probability*, volume 47 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 2018. doi:10.1017/9781108231596. An introduction with applications in data science, With a foreword by Sara van de Geer.

[WXZ21] Z. Wang, J. Xin, and Z. Zhang. Sharp error estimates on a stochastic structure-preserving scheme in computing effective diffusivity of 3D chaotic flows. *Multiscale Model. Simul.*, 19(3):1167–1189, 2021. doi:10.1137/19M1275516.

[WXZ22] Z. Wang, J. Xin, and Z. Zhang. Computing effective diffusivities in 3D time-dependent chaotic flows with a convergent Lagrangian numerical method. *ESAIM Math. Model. Numer. Anal.*, 56(5):1521–1544, 2022. doi:10.1051/m2an/2022049.

[You88] W. R. Young. Arrested shear dispersion and other models of anomalous diffusion. *J. Fluid Mech.*, 193:129–149, Aug 1988. doi:10.1017/S0022112088002083.

[YPP89] W. Young, A. Pumir, and Y. Pomeau. Anomalous diffusion of tracer in convection rolls. *Phys. Fluids A*, 1(3):462–469, 1989. doi:10.1063/1.857415.

[Zas02] G. M. Zaslavsky. Chaos, fractional kinetics, and anomalous transport. *Phys. Rep.*, 371(6):461–580, 2002. doi:10.1016/S0370-1573(02)00331-9.

[ZSW93] G. M. Zaslavsky, D. Stevens, and H. Weitzner. Self-similar transport in incomplete chaos. *Phys. Rev. E*, 48:1683–1694, Sep 1993. doi:10.1103/PhysRevE.48.1683.

Department of Mathematics, ETH Zürich, Switzerland
Email address: `bill@cprmn.org`

Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213.
Email address: `gautam@math.cmu.edu`

Department of Mathematics, Duke University, Durham, NC 27708.
Email address: `james.nolen@duke.edu`