Communications in
Mathematical Physics

Random Splitting of Fluid Models: Unique Ergodicity and Convergence

Andrea Agazzi[1], Jonathan C. Mattingly[1,2,3] , Omar Melikechi[1]

[1] Department of Mathematics, Duke University, Durham, NC, USA.
   E-mail: jonm@math.duke.edu
[2] Department of Statistical Science, Duke University, Durham, NC, USA.
[3] Institute for Advanced Study, Princeton, NJ, USA.

Abstract: We introduce a family of stochastic models motivated by the study of nonequilibrium steady states of fluid equations. These models decompose the deterministic dynamics of interest into fundamental building blocks, i.e., minimal vector fields preserving some fundamental aspects of the original dynamics. Randomness is injected by sequentially following each vector field for a random amount of time. We show under general conditions that these random dynamics possess a unique, ergodic invariant measure and converge almost surely to the original, deterministic model in the small noise limit. We apply our construction to the Lorenz-96 equations, often used in studies of chaos and data assimilation, and Galerkin approximations of the 2D Euler and Navier–Stokes equations. An interesting feature of the models developed is that they apply directly to the conservative dynamics and not just those with excitation and dissipation.

1. Introduction

This paper studies the long time dynamics of fluid-like equations that are kept out of equilibrium. Among the simplest examples of fluid models displaying interesting out-of-equilibrium behavior (such as fluxes across scales) are the two-dimensional Euler and incompressible Navier–Stokes equations. On the 2-dimensional torus $\mathbb{T}$, i.e., $\mathbb{T} := [0, 2\pi]^2$ with periodic boundary conditions, the Navier–Stokes equations, which model the flow of an incompressible fluid, are

$$\begin{cases} \partial_t u + (u \cdot \nabla)u = -\nabla p + F + \nu \Delta u \,, \\ \mathrm{div}(u) := \nabla \cdot u = 0 \,, \end{cases} \tag{1.1}$$

where $u : \mathbb{T} \times \mathbb{R} \to \mathbb{R}^2$ is the fluid velocity, $p : \mathbb{T} \times \mathbb{R} \to \mathbb{R}$ the fluid pressure,

$$(u \cdot \nabla)u = (u_1 \partial_1 u_1 + u_2 \partial_2 u_1, u_1 \partial_1 u_2 + u_2 \partial_2 u_2), \quad \text{and} \quad \Delta u = \partial_1^2 u_1 + \partial_2^2 u_2 \,.$$

Here $u = (u_1, u_2)$ and $\partial_j := \partial_{x_j}$. The viscosity $\nu > 0$ measures the strength of the dissipation introduced by the Laplacian Δ, and $F(x, t)$ is an external driving force whose role is to keep the system from relaxing to the trivial state $u \equiv 0$.

By balancing the dissipative effect of Δu, the forcing term allows the system to establish an out-of-equilibrium steady state. Such statistical equilibria often develop fluxes across scales, a phenomenon whose study is an active area of research. Often F is taken to live on only a few scales so that the flux out of those scales can be studied [18,24,29,40]. In practice, the forcing $F(x, t)$ is usually taken to be stochastic in space and time for some stationary distribution which is typically white in time [14,18,20,24]. A common choice in the literature is $F(x, t) = \sum \psi_k(x) \dot{W}_k(t)$ where each $\psi_k(x)$ is a fixed spatial forcing and $\{\dot{W}_k(t)\}$ are a collection of mutually independent white in time noise terms written here as the formal derivative of a Brownian motion. Stochastic forcing serves multiple purposes in these settings. On one hand, as already mentioned, it provides the energetic excitation which keeps the system out of equilibrium and allows for the establishment of a nontrivial statistical steady state. On the other hand, it provides local agitation which, modulo certain constraints, ensures the existence of a unique statistical steady state to which the system converges for most initial conditions. In other words, it guarantees the forcing is sufficiently varied and generic to ensure convergence to a single long time statistical behavior of the system, largely independent of the system's initial configuration.

This paper studies a class of processes, introduced in the next section, injecting randomness into the fluid models of interest while separating in a simple way the various roles served by noise in previous works. In particular, the randomness is used primarily to ensure that when the dynamics is sufficiently generic, unique ergodicity[1] holds for a broad class of initial conditions. This will free one to use a much less disruptive class of forcing to keep the system out of equilibrium. More specifically, the class of models introduced below have a number of desirable properties:

(1) They allow one to separate the effect of forcing, which keeps the system out of equilibrium, and stochastic agitation, which ensures the system has a unique long time statistical behavior.
(2) The stochastic agitation is strongly non-reversible since it is constructed from dynamics which only flow in the directions the original dynamics could already move.
(3) The stochastic agitation preserves the conserved quantities of the original dynamics. This allows the properties of the (stochastic) conservative dynamics to be studied directly rather than only as a limit of the forced-dissipated dynamics.
(4) The model dynamics will be constructed as the composition of simple dynamics, isolating particular nonlinear interactions which are relatively intuitive and can be explicitly analyzed.

By balancing between preservation of fundamental macroscopic properties of the original dynamics as in (3) and simplicity of the fundamental building blocks in our model dynamics as in (4), we expect the stochastic models introduced in this paper will provide meaningful physical and dynamical insight into nonequilibrium steady states of models such as (1.1).

Our decomposition into fundamental building blocks is partially motivated by the classical stylized models of dynamics studied in depth at the dawn of the theory of dynamical systems. Examples include the doubling map, quadratic maps, the Henon map, the Smale horseshoe, and extended systems like coupled map lattices (see [16,27]

---

[1] See Definition 1.1.

and references therein). The form of the decomposition is also motivated by the recent progress in proving ergodic properties of piecewise deterministic Markov processes (PDMPs) and their success as modeling and sampling tools. See for example [2–4,6,7, 9,13,15,17,33–35,38,44].

1.1. A class of stochastic models. We now introduce the general idea underlying the class of stochastic models, called *random splitting*, that we study in this paper. A more systematic definition of these models is deferred to Sect. 2. Consider an ordinary differential equation (ODE)

$$\dot{x} = V(x) = \sum_{k=1}^{n} V_k(x), \tag{1.2}$$

where $n \in \mathbb{N}$ and V and $\{V_k\}_{k=1}^{n}$ are vector fields on $\mathbb{R}^d$. In what follows, we choose the V_k so that the dynamics

$$\dot{x} = V_k(x) \tag{1.3}$$

are in some sense simpler than the dynamics corresponding to (1.2). We then approximate the solution $\Psi_t \colon x(0) \mapsto x(t)$ of (1.2) with compositions of the solution maps $\varphi_t^{(k)} \colon x(0) \mapsto x(t)$ of (1.3). This procedure is known as operator splitting in the numerical analysis literature and is often used in numerical simulations of various ordinary, partial, and stochastic differential equations [1,8,10,11,22,32,39,46,47]. Typically, the goal is to leverage the fact that each of the dynamics in (1.3) is more computationally tractable than (1.2) to construct an efficient and accurate numerical method. A variant of these models was also explored in the thesis [51].

  Here our goal is related but slightly different. Specifically, instead of evolving each $\varphi^{(k)}$ for a fixed time h as in traditional operator splitting methods, we evolve each of the $\varphi^{(k)}$ for a random time with mean h. Repeated composition then produces dynamics on $\mathcal{O}(1)$ times. The evolution times for each $\varphi^{(k)}$, and over each cycle, will be identically distributed and mutually independent, which implies our models are Markovian. As in the numerical analysis context, we hope to leverage the simplified nature of each $\varphi^{(k)}$, obtained from (1.3), to gain insight into the complex dynamics of the composition of maps. We will also see that as the mean evolution time $h \to 0$, the random splitting associated to (1.2) will almost surely converge to the deterministic dynamics Ψ_t on finite time intervals. However, we are most interested in studying the random splitting in its own right and not strictly as a approximation of (1.2). We will be particularly interested in its long time behavior and qualitative understanding of the stationary dynamics the random splitting produces when $h > 0$. More specifically, the property of the system we aim to establish is codified in the following standard definition from the theory of Markov processes; the supporting definition of *invariant measure* is given in the first paragraph of Sect. 3 after the transition kernel of random splitting is explicitly introduced.

Definition 1.1. A Markov process on a manifold $\mathcal{X}$ is uniquely ergodic on $\mathcal{X}$ if its transition kernel admits exactly one invariant probability measure on $\mathcal{X}$.

  We note that the definition of the set $\mathcal{X}$ where the above property holds can be quite delicate. While in general there might not exist a d-dimensional manifold $\mathcal{X}$ in $\mathbb{R}^d$ on which the random splitting is uniquely ergodic (see for example Remark 1.6), in the examples below we will identify a family of manifolds of lowest co-dimension where the above definition applies.

Remark 1.2. The set of invariant probability measures for a Markov transition kernel is convex, and the extremal points of this set are precisely the ergodic invariant measures [12,23]. In particular, if the transition kernel admits exactly one invariant measure, then it is necessarily extremal and therefore ergodic. This explains the use of the term ergodic in Definition 1.1.

1.2. Two motivating examples. In this paper, we consider two motivating examples: A conservative version of the Lorenz-96 model and Galerkin approximations of the vorticity formulation of the 2D Euler equations. We then use these analyses to study the full Lorenz-96 model and Galerkin approximations of the vorticity formulation of 2D Navier–Stokes.

Lorenz-96. Fix $n \geq 4$ and let $\{e_k\}_{k=1}^n$ denote the standard basis of $\mathbb{R}^n$. The Lorenz-96 model is

$$\dot{x} = \sum_{k=1}^n \big((x_{k+1} - x_{k-2})x_{k-1} - \nu x_k + F_k\big)e_k \tag{1.4}$$

for $x \in \mathbb{R}^n$, $\nu > 0$, and nonnegative constants F_k, where the indices are periodized via the identities $x_{-1} := x_{n-1}$, $x_0 := x_n$, and $x_{n+1} := x_1$. The $-\nu x_k$ term in (1.4) represents dissipation in the kth coordinate and F_k is a forcing constant. Initially, we study a variant of Lorenz-96, called *conservative Lorenz-96*, obtained by removing the dissipation and forcing terms from Lorenz-96. That is,

$$\dot{x} = V(x) := \sum_{k=1}^n (x_{k+1} - x_{k-2})x_{k-1}e_k \,. \tag{1.5}$$

We sometimes refer to the original Lorenz-96 model as the *forced Lorenz-96* model to emphasize the forcing (though the dissipation is equally important). For conservative Lorenz-96, we will decompose V into a collection of simple rotations by observing that

$$V(x) = \sum_{k=1}^n V_k(x) \tag{1.6}$$

where $V_k(x) := (x_{k+1}e_k - x_ke_{k+1})x_{k-1}$. The dynamics given by $\dot{x} = V_k(x)$ are easy to understand on their own; any complex behavior comes from interactions of the rotations. Importantly, each V_k is chosen to conserve, like V, the system's *energy*, which for Lorenz-96 is defined to be the square of the usual Euclidean norm, $\|x\|^2 := \sum_{k=1}^n x_k^2$.

2D Euler. Returning to (1.1), we begin by defining the scalar vorticity $q(x,t) = \operatorname{curl} u(x,t)$ of the velocity field $u(x,t)$. Initially, we will consider the Euler equations which are obtained from (1.1) by taking $\nu = F = 0$. Writing the equation for the jth Fourier mode $q_j \in \mathbb{C}$, defined by $q(x,t) = \sum_j q_j(t)e_j(x)$ for $e_j(x) := e^{ix \cdot j}$, and $j \in \{j \in \mathbb{Z}^2 : |j| < N, j \neq 0\}$, we have

$$\dot{q}_j = - \sum_{j+k+\ell=0} C_{k\ell} \bar{q}_k \bar{q}_\ell \tag{1.7}$$

for a constant $C_{k\ell}$ defined in Sect. 6.1. We will see that this system has two conserved quantities, the *enstrophy*, $\sum_j |q_j|^2$, and the *energy*, $\sum_j |j|^{-2}|q_j|^2$. Notice that the definition of energy differs between this equation and the Lorenz-96 model.

As in the Lorenz-96 model, we introduce the simpler dynamics $\dot{q} = V_{jk\ell}(q)$ where $V_{jk\ell}(q) = -C_{k\ell}\bar{q}_k\bar{q}_\ell e_j - C_{j\ell}\bar{q}_j\bar{q}_\ell e_k - C_{jk}\bar{q}_j\bar{q}_k e_\ell$ and observe that

$$V(q) = \sum_{j+k+\ell=0} V_{jk\ell}(q).$$

We will see in Sect. 6 that with this choice of splitting the dynamics $\dot{q} = V_{jk\ell}(q)$, like the original system $V(q)$, preserves the important physical quantities of enstrophy and energy.

Remark 1.3. In Sect. 6, we further simplify these complex-valued dynamics by projecting onto a real basis. The current choice is sufficient for an introductory discussion.

Remark 1.4. Our results do not focus on establishing minimal hypoellipticity assumptions for our models of Lorenz-96, 2D Euler and 2D Navier–Stokes; the stochastic agitation we use is more global than the minimal hypoellipticity forcing considered in [18,24]. We hope this will allow us to progress further than with previous models while preserving much of the physically interesting dynamics.

Remark 1.5. It is important to emphasize that, with regard to unique ergodicity, the main role of the forcing, when included, is only to destroy the fixed points and other low-dimensional invariant structures of the original flows and not to provide the stochastic mixing which ensures the existence of a unique, ergodic measure to which the system's statistics converge. The stochastic mixing is largely provided by the random splitting and is in contrast to the results in [5,18,19,24,29–31].

Remark 1.6. When considering conservative versions of our split dynamics (those without any explicit dissipation or body forcing), we cannot expect there to be a unique invariant measure for the system. In particular, since the dynamics will be constrained to level sets of the conserved quantities, there will be at least one invariant measure per level set. Furthermore, we will see that even on such constraint level sets there can be multiple ergodic invariant measures. Most will correspond to fixed points of the original dynamics and other lower-dimensional invariant structures. However we will see, in the two examples considered, that when our family of switched vector fields is sufficiently rich, there will be a unique ergodic invariant measure which is absolutely continuous with respect to the volume measure on the level set. This implies that in these examples, there is a unique ergodic invariant measure concentrated on a set of full measure inside each constraint level set. In this sense, we will demonstrate a form of uniqueness which aligns with the form of unique ergodicity often proven in the smooth deterministic dynamics setting, i.e., that there is only one invariant measure absolutely continuous with respect to the setting's natural Lebesgue measure.

Organization of paper. In Sect. 2, we introduce random splitting and its state spaces, called $\mathcal{V}$-*orbits*. In Sect. 3, we give conditions for random splitting to be uniquely ergodic on a $\mathcal{V}$-orbit. In Sect. 4, we show under general conditions that random splitting converges to its deterministic counterpart (1.2) on finite time intervals both in terms of its transition kernel and almost surely as the average time step h goes to zero. In Sects. 5 and 6, we construct random splittings of conservative Lorenz-96 and Galerkin approximations of

2D Euler and apply the preceding results to show these splittings are uniquely ergodic and converge on finite time intervals as $h \to 0$. In doing so we show each system has a unique invariant measure that is absolutely continuous (with respect to the volume measure) on the set defined by a given choice of the conserved quantities. In Sect. 7, we consider the Lorenz-96 and Euler models when fixed forcing and dissipation are added. When appropriate dissipation is chosen, the latter model corresponds to a random splitting of Galerkin approximations of 2D Navier–Stokes. We again construct random splittings of these models, prove convergence, and show that if the forcing is not aligned with the equations' invariant structures (such as fixed points) then both randomly split Lorenz-96 and Galerkin approximations of 2D Navier–Stokes have a unique invariant measure and the distribution starting from any initial condition converges exponentially to this measure.

2. Random Splitting in a General Setting

Let $\mathcal{V} := \{V_k\}_{k=1}^n$ be a family of complete[2], $\mathcal{C}^2$ vector fields[3] on $\mathbb{R}^d$ and set

$$V := \sum_{k=1}^n V_k. \tag{2.1}$$

Denote the flow of $\dot{x} = V(x)$ by Ψ and the flow of $\dot{x} = V_k(x)$ by $\varphi^{(k)}$. Ψ is the *true dynamics*. To construct a random dynamics approximating Ψ, fix $h > 0$, let $\tau = (\tau_k)_{k=1}^\infty$ be a sequence of independent exponential random variables with mean 1, and set $h\tau := (h\tau_k)_{k=1}^\infty$. The approximating dynamics, henceforth referred to as the *random splitting associated to* $\mathcal{V}$ or just *random splitting* for short, is the Markov chain $\{\Phi_{h\tau}^m\}_{m=0}^\infty$ defined by $\Phi_{h\tau}^0 := I$ and, for $m > 0$,

$$\Phi_{h\tau}^m := \varphi_{h\tau_{mn}}^{(n)} \circ \cdots \circ \varphi_{h\tau_{(m-1)n+1}}^{(1)} (\Phi_{h\tau}^{m-1}), \tag{2.2}$$

where I is the identity on $\mathbb{R}^d$, $\Phi := \varphi^{(n)} \circ \cdots \circ \varphi^{(1)}$, and Φ^m is the m-fold composition of Φ. Note that $h\tau_k \overset{iid}{\sim} \text{Exp}(1/h)$. Therefore, starting from the current step, the next step of the chain is obtained by following each V_k for $\text{Exp}(1/h)$ time in order from $k = 1$ to n. The chain is Markovian because the random times are independent. Its transition kernel P_h acts on measurable functions $f : \mathbb{R}^d \to \mathbb{R}$ via

$$P_h f(x) = \mathbb{E}\big(f(\Phi_{h\tau}(x))\big) = \int_{\mathbb{R}_+^n} f(\Phi_{ht}(x))e^{-\sum_{k=1}^n t_k} dt \tag{2.3}$$

where $\mathbb{R}_+ := (0, \infty)$, $t = (t_1, \ldots, t_n)$, and $dt = dt_1 \cdots dt_n$.

Remark 2.1. Throughout this paper the superscripts k in $\varphi^{(k)}$ and subscripts k in V_k are understood to be taken modulo n if $k \mod n \neq 0$ and to be n otherwise. For example, if $n = 3$,

$$\varphi^{(6)} \circ \varphi^{(5)} \circ \varphi^{(4)} \circ \varphi^{(3)} \circ \varphi^{(2)} \circ \varphi^{(1)} = \varphi^{(3)} \circ \varphi^{(2)} \circ \varphi^{(1)} \circ \varphi^{(3)} \circ \varphi^{(2)} \circ \varphi^{(1)}.$$

---

[2] A vector field is *complete* if its flow curve starting from any point exists for all time.

[3] We use calligraphic $\mathcal{C}^k$ for k-times continuously differentiable maps throughout to avoid confusion with constants which are often denoted by normal script C (for example, the constants C_{jk} in 2D Euler).

Also, the t in Φ_t^m is always a sequence $t = (t_1, \ldots, t_{mn})$ or, more generally, $t = (t_k)_{k=1}^{\infty}$, so that

$$\Phi_t^m(x) = \varphi_{t_{mn}}^{(n)} \circ \cdots \circ \varphi_{t_1}^{(1)}(x).$$

Note that the above is a composition of mn flows, as in (2.2).

Remark 2.2. All results in this paper remain true if at each step we randomly permute indices in the composition Φ. That is, given a current state x, the next step is $\varphi_{h\tau_n}^{(\sigma(n))} \circ \cdots \circ \varphi_{h\tau_1}^{(\sigma(1))}(x)$ where σ is a random permutation of $\{1, \ldots, n\}$. This yields both additional randomness and an avenue to higher order approximations of the true dynamics [10, 11,32,46,47]. We forgo this more general setting however to keep exposition more approachable and notationally light.

Remark 2.3. The times are assumed exponentially distributed for convenience. All results extend to any distribution on $[0, \infty)$ with positive density on $(0, \varepsilon)$ for some $\varepsilon > 0$ and exponential tails. The second condition, which is not sharp, guarantees sufficient concentration of averages of random flow times τ_i in Lemmas A.3 and A.4 and is required for the convergence results as $h \to 0$ in Sect. 4. The first condition is used in Sects. 5 and 6 to guarantee sufficient flexibility in the trajectories of the split systems of interest to establish the global irreducibility needed for ergodicity.

2.1. $\mathcal{V}$-Orbits. Throughout this paper we often restrict attention to certain subsets of $\mathbb{R}^d$ affiliated with the family of vector fields $\mathcal{V}$. Specifically, for each x in $\mathbb{R}^d$ define the *$\mathcal{V}$-orbit of x* by

$$\mathcal{X}(x) := \left\{ \Phi_t^m(x) : m \geq 0, t \in \mathbb{R}^{mn} \right\}. \tag{2.4}$$

This is the set of points in $\mathbb{R}^d$ that can be reached by the split dynamics starting from x in any finite number of steps and over arbitrary positive and negative times. $\mathcal{X}(x)$ is well-defined since the V_k are complete. Furthermore, since the time vectors t in (2.4) admit coordinates that are 0,

$$\mathcal{X}(x) = \left\{ \varphi_{t_{i_m}}^{(i_m)} \circ \cdots \circ \varphi_{t_{i_1}}^{(i_1)}(x) : m \in \mathbb{N}, 1 \leq i_j \leq n, t_{i_j} \in \mathbb{R} \right\}.$$

Hence (2.4) agrees with the definition of $\mathcal{V}$-orbits from control theory [26,50]. The collection $\{\mathcal{X}(x) : x \in \mathbb{R}^d\}$ partitions $\mathbb{R}^d$ and if the random splitting $\{\Phi_{h\tau}^m\}$ associated to $\mathcal{V}$ starts in $\mathcal{X}(x)$ then it stays in $\mathcal{X}(x)$ for all time. Therefore the $\{\Phi_{h\tau}^m\}$ previously defined on $\mathbb{R}^d$ also defines a Markov chain on $\mathcal{X}(x)$ whenever it starts in $\mathcal{X}(x)$, and its transition kernel P_h acts on measurable functions $f : \mathcal{X}(x) \to \mathbb{R}$ as in (2.3). When x is arbitrary or clear from context, we denote $\mathcal{X}(x)$ by $\mathcal{X}$. A classic result from geometric control theory, sometimes called *the orbit theorem*, says if every V_k in $\mathcal{V}$ is $\mathcal{C}^r$ for some $1 \leq r \leq \infty$ (respectively, analytic[4]), then every $\mathcal{X}$ is an immersed $\mathcal{C}^r$ (respectively, analytic) submanifold of $\mathbb{R}^d$ [26]. In particular, each $\mathcal{X}$ has a Riemannian structure induced by the Euclidean structure on $\mathbb{R}^d$ and an associated volume form, henceforth denoted λ, sometimes called *Hausdorff* or *Lebesgue measure on $\mathcal{X}$*, which serves as our reference measure on $\mathcal{X}$.

---

[4] Throughout this paper *analytic* means *real-analytic*.

Remark 2.4. The orbit theorem says every $\mathcal{V}$-orbit $\mathcal{X}$ is an immersed but not necessarily embedded submanifold of $\mathbb{R}^d$. For example, $\mathcal{X}$ can be a "figure-eight" curve in $\mathbb{R}^2$ [37, Example 5.19]. Nevertheless, every $\mathcal{X}$ is a manifold with a volume form induced by the ambient Euclidean structure, and every vector field in $\mathcal{V}$ restricts to a vector field on $\mathcal{X}$ by construction. In particular, $\{V_i(x) : V_i \in \mathcal{V}\}$ is a set of vectors in the tangent space $T_x\mathcal{X}$ for every x in $\mathcal{X}$. Throughout this paper submanifold will mean immersed submanifold without further qualification. See [36,37] for more on immersed and embedded submanifolds in general, and [26] for more on $\mathcal{V}$-orbits in particular.

3. Ergodicity

Let $\mathcal{V} := \{V_k\}_{k=1}^n$ be a family of complete, $\mathcal{C}^2$ vector fields on $\mathbb{R}^d$ as before and fix a p-dimensional $\mathcal{V}$-orbit $\mathcal{X}$. Also fix $h > 0$ and let P_h be the transition kernel of the associated random splitting on $\mathcal{X}$. A measure μ on $\mathcal{X}$ is P_h-*invariant* if $\mu P_h = \mu$ where μP_h is defined by

$$\mu P_h f := \int_{\mathcal{X}} P_h f(x) \mu(dx) \tag{3.1}$$

for all bounded, measurable functions $f : \mathcal{X} \to \mathbb{R}$. The main result of this section is

Theorem 3.1. *If there exists x_* in $\mathcal{X}$ such that for all x in $\mathcal{X}$ there is an m in $\mathbb{N}$ and t in $\mathbb{R}_+^{mn}$ with $\Phi^m(x, t) = x_*$ and $D_t\Phi^m(x, t) : T_t\mathbb{R}_+^{mn} \to T_{x_*}\mathcal{X}$ surjective, then P_h has at most one invariant measure on $\mathcal{X}$. Moreover, if such a measure exists, it is absolutely continuous with respect to the volume form on $\mathcal{X}$.*

Here $T_{x_*}\mathcal{X}$ is the tangent space of $\mathcal{X}$ at x_*. The proof of Theorem 3.1 follows from the classical minorization condition [25,41,43,48] given by the following result, which appears in [6, Lemma 6.3].

Lemma 3.2. *Let $p \leq m$ and let $F : \mathcal{X} \times U \to \mathcal{X}$ be $\mathcal{C}^1$, where U is an open subset of $\mathbb{R}^m$. Suppose τ is a U-valued random variable with continuous density ρ. If for some (x, t) in $\mathcal{X} \times U$ the map $D_t F(x, t)$ is surjective and ρ is bounded below by $c_0 > 0$ on a neighborhood of t, then there exists a constant $c > 0$ and neighborhoods U_x of x and U_* of $x_* := F(x, t)$ such that*

$$\mathbb{P}\big(F(y, \tau) \in B\big) \geq c\lambda(B \cap U_*) \tag{3.2}$$

for all y in U_x and B in the Borel σ-algebra $\mathcal{B}(\mathcal{X})$ of $\mathcal{X}$ (recall λ is the volume form on $\mathcal{X}$).

Remark 3.3. In our setting, $U = \mathbb{R}_+^{mn}$, $F = \Phi^m : \mathcal{X} \times \mathbb{R}_+^{mn} \to \mathcal{X}$, and $\tau = (\tau_1, \ldots, \tau_{mn})$ with the τ_k independent exponential random variables with mean h. In this case, if $x_* = \Phi^m(x, t)$ for some t with $D_t\Phi^m(x, t)$ surjective, then Lemma 3.2 guarantees the existence of a constant $c > 0$ and neighborhoods U_x of x and U_* of x_* such that, for all y in U_x and B in $\mathcal{B}(\mathcal{X})$,

$$P^m(y, B) \geq c\lambda(B \cap U_*) \,. \tag{3.3}$$

Proof of Theorem 3.1. The proof is by contradiction. Suppose μ_1 and μ_2 are distinct P_h-invariant probability measures. Assume without loss of generality both μ_i are ergodic and therefore mutually singular [12,28]. Then there exist disjoint measurable sets A_1 and A_2 partitioning $\mathcal{X}$ such that $\mu_i(B) = \mu_i(B \cap A_i)$ for all B in $\mathcal{B}(\mathcal{X})$. Fix x_i in the support of μ_i so, by definition, μ_i gives positive measure to every neighborhood of x_i. By hypothesis and Remark 3.3 there exist $c_i > 0$, $m_i \in \mathbb{N}$, and neighborhoods U_i of x_i and U_* of x_* such that $P_h^{m_i}(x, \cdot) \geq c_i \lambda(\cdot \cap U_*)$ for all x in U_i. So,

$$\mu_i(B) = \mu_i P_h^{m_i}(B) \geq \int_{U_i} P_h^{m_i}(x, B)\mu_i(dx) \geq c_i \lambda(B \cap U_*)\mu_i(U_i) \qquad (3.4)$$

for all B in $\mathcal{B}(\mathcal{X})$. In particular, $\mu_i(B) = 0$ implies $\lambda(B \cap U_*) = 0$ since c_i and $\mu_i(U_i)$ are strictly positive. But $\mu_1(A_2 \cap U_*) = \mu_2(A_1 \cap U_*) = 0$ and hence

$$0 < \lambda(U_*) = \lambda(A_1 \cap U_*) + \lambda(A_2 \cap U_*) = 0,$$

which is a contradiction. Absolute continuity of the P_h-invariant measure μ, provided it exists, follows from uniqueness together with the fact that the absolutely continuous part, μ_{ac}, and singular part, μ_s, of μ are P_h-invariant whenever μ is [6, Proposition 2.7]. Specifically, since μ_{ac} and μ_s are P_h-invariant and there can be at most one P_h-invariant probability measure, either μ_{ac} or μ_s is identically zero. Since μ_{ac} is nonzero by (3.4), it follows that $\mu_s = 0$ and therefore $\mu = \mu_{ac}$. □

Remark 3.4. The invariant measure μ, which we defined as a fixed point of the left action of the Markov semigroup P, is often called a stationary measure. This is since the sequence of random variables generated by the Markov process starting from an initial condition distributed according to μ will be stationary. This helps distinguish from the invariant measure of the skew flow $(x, \tau) \mapsto (\Psi_{h\tau}(x), \vartheta\tau)$ where the shift ϑ is defined by $\vartheta\tau : \tau = (\tau_1, \tau_2, \cdots) \mapsto (\tau_{n+1}, \tau_{n+2}, \cdots)$. The skew perspective captures more information about the dynamics and is preferred for many questions. However, we will not pursue it here as it complicates the simple picture we explore in this note.

3.1. The Lie bracket condition. Let $\mathfrak{X}(\mathcal{X})$ be the Lie algebra of smooth vector fields on $\mathcal{X}$ and assume throughout this subsection the vector fields in $\mathcal{V}$ are smooth. Then the smallest subalgebra $\mathrm{Lie}(\mathcal{V})$ of $\mathfrak{X}(\mathcal{X})$ containing $\mathcal{V}$ is well-defined, and for each x in $\mathcal{X}$ the collection $\mathrm{Lie}_x(\mathcal{V}) := \{V(x) : V \in \mathrm{Lie}(\mathcal{V})\}$ is a subspace of the tangent space $T_x\mathcal{X}$ at x.

Definition 3.5. The Lie bracket condition holds at x in $\mathcal{X}$ if $\mathrm{Lie}_x(\mathcal{V}) = T_x\mathcal{X}$.

The Lie bracket condition is called the *weak bracket condition* in [6] and *Condition B* in [2]. Both papers also consider a *strong bracket condition* (*Condition A*) which is used for results about continuous time Markov processes and is therefore not needed here. The Lie bracket condition has the following important consequence. Note $\mathbb{R}_+ := (0, \infty)$ throughout this paper.

Theorem 3.6. *If the Lie bracket condition holds at a point x_* in $\mathcal{X}$ then for every neighborhood U of x_* in $\mathcal{X}$ and every $T > 0$ there exists an x in U, an m, and a t in $\mathbb{R}_+^{mn}$ such that $\sum_{k=1}^{mn} t_k \leq T$ and $t \mapsto \Phi^m(x_*, t) = x$ is a submersion at t, i.e. $D_t\Phi^m(x_*, t) : T_t\mathbb{R}^{mn} \to T_x\mathcal{X}$ is surjective.*

A version of Theorem 3.6 appears as Theorem 3.1 in [26]; the equivalent version given here is better suited to random splitting and other classes of piecewise deterministic Markov processes. See Theorem 5 in [2] and Theorem 4.4 in [6] and their corresponding discussions for details. Intuitively, Theorem 3.6 says that if the Lie bracket condition holds at x_* then, as a consequence of surjectivity, the random splitting can move in any infinitesimal direction from x_* in arbitrarily small positive times. The next result is an immediate consequence of Theorems 3.1 and 3.6.

Corollary 3.7. *Suppose there is an x_* in $\mathcal{X}$ at which the Lie bracket condition holds and such that for every x in $\mathcal{X}$ there is an $m \in \mathbb{N}$ and a $t \in \mathbb{R}_+^{mn}$ satisfying $\Phi^m(x, t) = x_*$. Then P_h has at most one invariant measure on $\mathcal{X}$. Furthermore, if such a measure exists, it is absolutely continuous with respect to the volume form λ.*

One benefit of Corollary 3.7 is that it replaces the need to check the surjectivity assumption of Theorem 3.1, which can be challenging in practice, with the verification of the Lie bracket condition. The next result provides a further convenience in the analytic setting which will be used in the specific examples considered below. See [26,45] for further discussion and proof.

Theorem 3.8 *[Nagano]. Suppose the vector fields in $\mathcal{V}$ are analytic. If the Lie bracket condition holds at any point in $\mathcal{X}$, then it holds at every point in $\mathcal{X}$.*

Corollary 3.9. *Suppose the vector fields in $\mathcal{V}$ are analytic and there is a point x_* in $\mathcal{X}$ such that for every x in $\mathcal{X}$ there is an $m \in \mathbb{N}$ and a $t \in \mathbb{R}_+^{mn}$ satisfying $\Phi^m(x, t) = x_*$. If the Lie bracket condition holds at any point in $\mathcal{X}$, then P_h has at most one invariant measure on $\mathcal{X}$. Furthermore, if such a measure exists, it is absolutely continuous with respect to the volume form λ.*

Proof. Since the Lie bracket condition holds at one point in $\mathcal{X}$, it also holds at x_* by Nagano's theorem. The result follows by Corollary 3.7. □

4. Convergence as Mean Time Step Goes to Zero

A well-known result in the operator splitting literature is that the error incurred in approximating Ψ by the deterministic splitting scheme $\Phi_h = \varphi_h^{(n)} \circ \cdots \circ \varphi_h^{(1)}$ is $\mathcal{O}(h)$ [39]. That is, Φ_h converges to the true dynamics Ψ at worst linearly in h as $h \to 0$. In this section we give analogous results for random splitting; the pluralized "results" reflects that with randomness comes several different notions of convergence. Specifically, we give two main results. First, as in the deterministic case, the transition kernel P_h of random splitting converges to the true dynamics linearly in h as $h \to 0$. Second, random splitting converges almost-surely to the true dynamics as $h \to 0$. Each case requires a slightly different notion of $\mathcal{O}(h)$. These statements are made precise in Theorems 4.1 and 4.5, respectively, but to make sense of them we first introduce the appropriate setting.

The following assumption on $\mathcal{V}$-orbits is used throughout this section.

Assumption 1. $\mathcal{X}(x)$ is bounded for each x in $\mathbb{R}^d$.

Since the vector fields V_k are assumed $\mathcal{C}^2$, Assumption 1 implies the V_k are bounded with bounded first and second derivatives on every $\mathcal{X}$. In particular,

$$C_*(x_0) := \sup_{x \in \mathcal{X}(x_0)} \left\{ \|V_k(x)\|, \|DV_k(x)\|, \|D^2 V_k(x)\| : 1 \leq k \leq n \right\} < \infty, \quad (4.1)$$

where $\|V_k(x)\|$ is the usual Euclidean norm, $\|DV_k(x)\|$ is the operator norm of the linear map $DV_k(x) : \mathbb{R}^d \to \mathbb{R}^d$, and $\|D^2 V_k(x)\|$ is the operator norm of the bilinear map $D^2 V_k(x) : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$.

For a positive integer k let $\mathcal{C}^k(\mathcal{X})$ be the space of k-times continuously differentiable functions $f : \mathcal{X} \to \mathbb{R}$. For f in $\mathcal{C}^k(\mathcal{X})$ and $\ell \leq k$, the ℓth derivative $D^\ell f(x)$ of f at x is a multilinear operator from $\otimes_1^\ell T_x \mathcal{X}$ to $\mathbb{R}$. The operator norm of $D^\ell f(x)$ is then

$$\|D^\ell f(x)\| := \sup_{\|\eta\|=1} \left\{ |D^\ell f(x)\eta| \right\},$$

where $\eta \in \otimes_1^\ell T_x \mathcal{X}$. Defining $D^0 f(x) := f(x)$, this in turn induces a norm on $\mathcal{C}^k(\mathcal{X})$ given by

$$\|f\|_k := \sup_{x \in \mathcal{X}} \left\{ \|D^\ell f(x)\| : 0 \leq \ell \leq k \right\}.$$

The corresponding operator norm is denoted $\|\cdot\|_{k \to k}$. More generally, for any k and ℓ define a norm $\|\cdot\|_{k \to \ell}$ on the space of linear operators $L : \mathcal{C}^k(\mathcal{X}) \to \mathcal{C}^\ell(\mathcal{X})$ by

$$\|L\|_{k \to \ell} := \sup_{\|f\|_k=1} \|Lf\|_\ell.$$

We make frequent use of the *submultiplicity* of $\|\cdot\|_{k \to \ell}$. Namely, if A and B are bounded linear operators from $\mathcal{C}^j(\mathcal{X})$ to $\mathcal{C}^k(\mathcal{X})$ and from $\mathcal{C}^k(\mathcal{X})$ to $\mathcal{C}^\ell(\mathcal{X})$, respectively, then

$$\|BA\|_{j \to \ell} \leq \|B\|_{k \to \ell} \|A\|_{j \to k}.$$

The results below are stated in terms of semigroups of the flows Ψ and $\varphi^{(j)}$, which are $\mathcal{C}^2$ by assumption. Hence for all $k \leq 2$ the semigroup $\{S_t\}_{t \geq 0}$ corresponding to Ψ acts on $f \in \mathcal{C}^k(\mathcal{X})$ via

$$S_t f(x) = e^{tV} f(x) = f(\Psi_t(x)) \tag{4.2}$$

and, similarly, the semigroup $\{\widetilde{S}_t^{(j)}\}_{t \geq 0}$ corresponding to $\varphi^{(j)}$ is given by

$$\widetilde{S}_t^{(j)} f(x) = e^{tV_j} f(x) = f(\varphi_t^{(j)}(x)). \tag{4.3}$$

In particular, m steps of random splitting corresponds to $\widetilde{S}_{h\tau}^m := \widetilde{S}_{h\tau_1}^{(1)} \cdots \widetilde{S}_{h\tau_{mn}}^{(mn)}$ with superscripts taken as in Remark 2.2. The transition kernel P_h^m and semigroup composition $\widetilde{S}_{h\tau}^m$ are related via

$$P_h^m f = \mathbb{E}(f(\Phi_{h\tau}^m)) = \mathbb{E}(\widetilde{S}_{h\tau}^m f). \tag{4.4}$$

With the above notation we now present the two main results of this section, Theorems 4.1 and 4.5, which follow from Lemmas 4.2 and 4.6, respectively. The full proofs of both lemmas are given in the Appendix, but we discuss the general idea behind each at the end of this section.

Theorem 4.1. *Suppose Assumption 1 holds and fix $t > 0$. For all h sufficiently small and satisfying $mh = t$ for some $m \in \mathbb{N}$, there exists a constant $C(t)$ depending on t but not on h such that*

$$\|P_h^m - S_t\|_{2 \to 0} \leq C(t)h. \tag{4.5}$$

Lemma 4.2. *If Assumption 1 holds then there exists a constant C such that*

$$\|P_h - S_h\|_{2 \to 0} \le Ch^2 \tag{4.6}$$

for all h sufficiently small.

Recalling from (4.4) that $P_h = \mathbb{E}(\widetilde{S}_{h\tau}^1)$, informally Lemma 4.2 states that the average difference between one step of random splitting and the true dynamics is $\mathcal{O}(h^2)$ for sufficiently small h. For any finite time interval $[0, t]$ we can leverage this result to approximate S_t by successive steps of P_h. Specifically, choose h sufficiently small so that (4.6) holds and there exists an integer m with $mh = t$. Then the composition P_h^m corresponds to $\mathcal{O}(1/h)$ steps of P_h. Consequently, since the difference between P_h and S_h is $\mathcal{O}(h^2)$, the difference between P_h^m and S_t is $\mathcal{O}(h)$.

A possible interpretation of $\mathcal{O}(h^p)$ is given in Theorem 4.1 and Equation (4.5). This choice matched the particular results being proved. In Theorem 4.5 and Lemma 4.6 below, we chose to quantify the error in another fashion, though the same order of magnitude statements hold true. The same basic reasoning can be used to prove the following.

Remark 4.3. As we have made minimal assumptions on the splitting, we will only be able to deduce that $P_h - S_h = \mathcal{O}(h^2)$. In specific examples, it is often possible to arrange the splitting so that $P_h - S_h = \mathcal{O}(h^p)$ with $p > 2$. An example of a higher order splitting is Strang splitting [39]. Alternatively, higher order can also be obtained by fully randomizing the order [10] or randomly choosing between one ordering and its reverse [32,46,47].

Proof of Theorem 4.1. Let h be sufficiently small that (4.6) holds and such that $mh = t$ for some $m \in \mathbb{N}$. The quantity of interest can be written as the following telescoping sum:

$$P_h^m - S_t = \sum_{k=1}^{m} P_h^{k-1} (P_h - S_h) S_{h(m-k)} . \tag{4.7}$$

For any k and continuous function f with $\|f\|_0 = 1$,

$$\|P_h^k f\|_0 \le \mathbb{E}\left(\left\|f\big(\Phi_{h\tau}^k\big)\right\|_0\right) = 1 .$$

Hence $\|P_h^k\|_{0 \to 0} = 1$. Similarly, since $mh = t$ implies $h(m - k) \le t$ for $k \ge 0$ and $\mathcal{X}$ is bounded by Assumption 1 (so Ψ and its first and second derivatives are bounded on $\mathcal{X}$, uniformly on $[0, t]$),

$$\|S_{h(m-k)}\|_{2 \to 2} \le K(t)$$

for some $K(t)$ depending on t but not h. Hence, by submultiplicity, (4.7), and Lemma 4.2, we have

$$\|P_h^m - S_t\|_{2 \to 0} \le \sum_{k=1}^{m} \|P_h^{k-1}\|_{0 \to 0} \|P_h - S_h\|_{2 \to 0} \|S_{h(m-k)}\|_{2 \to 2} \le K(t) \sum_{k=1}^{m} Ch^2 = C(t)h ,$$

where $C(t) := K(t)C$, with C the constant from (4.6) in Lemma 4.2.   □

Remark 4.4. Theorem 4.1 had the relation $h = t/m$, while in the almost-sure results below we will take $h = t/m^2$ (note we explicitly write t/m^2, making no reference to the variable h). The reason, loosely speaking, is that the transition kernel depends only on the expectation of the randomness, while the almost-sure results additionally depend on fluctuations of the randomness about its mean. For example, Lemma 4.6 prepares for an application of the Borel-Cantelli lemma by establishing the summability of probabilities of "large" fluctuations over sets of $\mathcal{O}(m) = \mathcal{O}(1/\sqrt{h})$ cycles. This is discussed in more detail at the end of this section and worked out in full in the Appendix.

Theorem 4.5. *Suppose Assumption 1 holds and fix $t > 0$. Then for any $\varepsilon > 0$,*

$$\mathbb{P}\left(\limsup_{m\to\infty}\|\widetilde{S}^{m^2}_{t\tau/m^2} - S_t\|_{2\to 0} > \varepsilon\right) = 0.\tag{4.8}$$

Lemma 4.6. *Suppose Assumption 1 holds and fix $t > 0$. Then for any $\varepsilon > 0$,*

$$\sum_{m=1}^{\infty}\mathbb{P}\left(\|\widetilde{S}^{m}_{t\tau/m^2} - S_{t/m}\|_{2\to 0} > \tfrac{\varepsilon}{m}\right) < \infty.\tag{4.9}$$

Remark 4.7. There is a relationship between Theorems 4.1 and 4.5 and the averaging results from Wentzell-Freidlin theory, e.g., [21, Theorem 2.1, Chapter 7]. This theorem builds on local results like Lemmas 4.2 and 4.6. Since our averaging is that of a deterministic, cyclic process, the calculations can be more explicit and more precise. We are able to prove using simple calculations that the local error is $\mathcal{O}(h^2)$ which leads to $\mathcal{O}(h)$ error over order one times. Typical soft averaging results prove a local error of $o(h)$ [5] and then simply conclude that the order one error goes to 0. Of course, more careful calculations are possible in the averaging setting. However, the simple structure of our problems, where the only randomness is in the switching times and not the orderings, allows for the direct, straightforward proofs we have presented.

Proof of Theorem 4.5. By the Borel-Cantelli Lemma it suffices to show

$$\sum_{m=1}^{\infty}\mathbb{P}\left(\|\widetilde{S}^{m^2}_{t\tau/m^2} - S_t\|_{2\to 0} > \varepsilon\right) < \infty.$$

Consider the telescoping sum

$$\widetilde{S}^{m^2}_{t\tau/m^2} - S_t = \sum_{k=1}^{m}\widetilde{S}^{(k-1)}_{t\tau/m^2}\left(\widetilde{S}^{m}_{t\tau/m^2} - S_{t/m}\right)S_{(m-k)t/m}.\tag{4.10}$$

For any k and continuous function f with $\|f\|_0 = 1$,

$$\|\widetilde{S}^{k}_{t\tau/m^2}f\|_0 = \|f(\Phi^k_{h\tau})\|_0 = 1.$$

Hence $\|\widetilde{S}^{(k-1)}_{t\tau/m^2}\|_{0\to 0} = 1$. Similarly, since $(m-k)t/m \leq t$ for $k \geq 0$ and $\mathcal{X}$ is bounded by Assumption 1 (so Ψ and its first and second derivatives are bounded on $\mathcal{X}$, uniformly on $[0, t]$),

$$\|S_{(m-k)t/m}\|_{2\to 2} \leq K(t)$$

---

[5] $f(h)$ is $o(g(h))$ when $h \to 0$ if $\lim f(h)/g(h) = 0$ as $h \to 0$. $f(h)$ is $O(g(h))$ when $h \to 0$ if $\lim |f(h)/g(h)| \in (0, \infty)$.

for some $K(t)$ depending on t but not h. Hence, by submultiplicity, (4.10), and Lemma 4.6, we have

$$\left\|\widetilde{S}^{m^2}_{t\tau/m^2} - S_t\right\|_{2\to 0} \leq K(t) \sum_{k=1}^{m} \left\|\widetilde{S}^{m}_{t\tau/m^2} - S_{t/m}\right\|_{2\to 0} = K(t)m\left\|\widetilde{S}^{m}_{t\tau/m^2} - S_{t/m}\right\|_{2\to 0},$$

and hence by Lemma 4.6,

$$\sum_{m=1}^{\infty} \mathbb{P}\left(\left\|\widetilde{S}^{m^2}_{t\tau/m^2} - S_t\right\|_{2\to 0} > \varepsilon\right) \leq \sum_{m=1}^{\infty} \mathbb{P}\left(\left\|\widetilde{S}^{m}_{t\tau/m^2} - S_{t/m}\right\|_{2\to 0} > \frac{\varepsilon}{K(t)m}\right) < \infty.$$

$\square$

We conclude this section by sketching the proofs of Lemmas 4.2 and 4.6, which are inspired by ideas from [10,11] and given in full detail in the Appendix. In what follows we set $\widetilde{S}_{h\tau} := \widetilde{S}^1_{h\tau}$ and define $\widetilde{S}^{(i,j)}_{h\tau} := \widetilde{S}^{(i)}_{h\tau} \cdots \widetilde{S}^{(j)}_{h\tau}$. Consider first Lemma 4.2. Differentiating $\widetilde{S}_{h\tau}$ in h gives

$$\partial_h \widetilde{S}_{h\tau} = \sum_{k=1}^{n} \tau_k e^{h\tau_1} \cdots e^{h\tau_{k-1}} V_k e^{h\tau_k} \cdots e^{h\tau_n} = \sum_{k=1}^{n} \tau_k \widetilde{S}^{(1,k-1)}_{h\tau} V_k \widetilde{S}^{(k,n)}_{h\tau}.$$

Next, commute $\widetilde{S}^{(1,k-1)}_{h\tau}$ and V_k via the Lie bracket $[\widetilde{S}^{(1,k-1)}_{h\tau}, V_k] := \widetilde{S}^{(1,k-1)}_{h\tau} V_k - V_k \widetilde{S}^{(1,k-1)}_{h\tau}$ to get

$$\partial_h \widetilde{S}_{h\tau} = \sum_{k=1}^{n} \tau_k V_k \widetilde{S}_{h\tau} + \sum_{k=1}^{n} \tau_k [\widetilde{S}^{(1,k-1)}_{h\tau}, V_k]\widetilde{S}^{(k,n)}_{h\tau} = V\widetilde{S}_{h\tau} + (V_\tau - V)\widetilde{S}_{h\tau} + E_{h\tau}$$

where $V_\tau := \sum_{k=1}^{n} \tau_k V_k$ and $E_{h\tau} := \sum_{k=1}^{n} \tau_k [\widetilde{S}^{(1,k-1)}_{h\tau}, V_k]\widetilde{S}^{(k,n)}_{h\tau}$. So, by variation of constants,

$$\widetilde{S}_{h\tau} - S_h = \int_0^h S_{h-r}(V_\tau - V)\widetilde{S}_{r\tau} dr + \int_0^h S_{h-r} E_{r\tau} dr. \tag{4.11}$$

Loosely speaking, the first integrand is $\mathcal{O}(h)$ because

$$\mathbb{E}(V_\tau - V) = \sum_{k=1}^{n} \mathbb{E}(\tau_k - 1)V_k = 0 \tag{4.12}$$

cancels first order terms from the full expression, $S_{h-r}(V_\tau - V)\widetilde{S}_{r\tau}$. On the other hand the second integrand is $\mathcal{O}(h)$ because the bracket terms in $E_{h\tau}$ also cancel first order terms (most of the work of the proof in the Appendix is making these two statements precise). Thus, integrating these $\mathcal{O}(h)$ terms over the interval $(0, h)$, the difference on the right side of (4.11) is $\mathcal{O}(h^2)$ as claimed.

The proof of Lemma 4.6 is structurally similar to the one sketched above in that it again begins with an application of variation of constants. However, in this case our analysis aims to establish a concentration estimate and can therefore not rely solely on the vanishing first moment in as in (4.12). Instead, morally speaking, we expect the desired estimate to hold because of the averaging of iid flow times τ_i in the homologue

of (4.12). In order to capture such averaging, we cannot limit our analysis to one cycle, but have to consider a variation of constants estimate on $m \gg 1$ such cycles:

$$\widetilde{S}_{h\tau}^m - S_{mh} = \int_0^h S_{m(h-r)}(V_\tau - V)\widetilde{S}_{r\tau}^m \, dr + \int_0^h S_{m(h-r)} E_{r\tau}^{(m)} \, dr \qquad (4.13)$$

where now $E_{r\tau}^{(m)} := \sum_{k=1}^{mn} \tau_k [\widetilde{S}_{h\tau}^{(1,k-1)}, V_k] \widetilde{S}_{h\tau}^{(k,n)}$. Note that the second term contains $\mathcal{O}(m^2)$ commutators, each contributing $\mathcal{O}(h^2)$ as in the previous analysis. On the other hand, once integrated, the difference in the first integral, $\sum_{k=1}^{mn} (\tau_k - 1) V_k$, scales as $\mathcal{O}(\sqrt{m}h)$ by the central limit theorem for iid random variables. In order to have both terms decay faster than $\mathcal{O}(1/m)$ we choose $m \sim \mathcal{O}(1/\sqrt{h})$, whence the relation $h = t/m^2$.

5. Conservative Lorenz-96

In this section, we apply results of the previous sections to the conservative Lorenz-96 model introduced in Sect. 1.2. There we noted the vector field V in (1.5) splits as (1.6) where the flow of each V_k is a rotation. Specifically, each flow $\varphi^{(k)}$ of the splitting vector fields

$$V_k(x) = (x_{k+1}e_k - x_k e_{k+1})x_{k-1} \qquad (5.1)$$

is a rotation in the (x_k, x_{k+1})-plane with angular velocity x_{k-1} and therefore preserves Euclidean norm, which we refer to as the *energy* of the system. Throughout this section $\mathcal{V}$ denotes the family of splitting vector fields corresponding to (5.1). By the preceding remarks every $\mathcal{V}$-orbit lies on a sphere centered at the origin in $\mathbb{R}^n$. In particular, we have

Proposition 5.1. *All the finite time convergence results of Sect. 4 apply to the random splitting (1.6) of conservative Lorenz-96 starting from any initial condition.*

Proof. The splitting vector fields are smooth and Assumption 1 is satisfied since every $\mathcal{V}$-orbit lies on a sphere, so the conclusions of Theorems 4.1 and 4.5 both hold. □

5.1. Ergodicity. A complicating feature of the conservative Lorenz-96 dynamics is that it has fixed points. Specifically, a point x in $\mathbb{R}^n$ is a fixed point of (1.5) if and only if $\sum_{k=1}^n (x_k^2 + x_{k+1}^2) x_{k-1}^2 = 0$. For a 2-sphere embedded in $\mathbb{R}^3$ these are precisely the 6 points of intersection of the sphere with the standard coordinate axes. In higher dimensions, these fixed points lie on submanifolds that in general have dimension greater than 0 and in particular are no longer isolated. Nevertheless, nonfixed points cannot reach fixed points in finite time; in fact, the following result shows there is precisely one $\mathcal{V}$-orbit on each sphere that contains all the nonfixed points on that sphere.

Proposition 5.2. *If x is a nonfixed point of the conservative Lorenz-96 equations, then*

$$\mathcal{X}(x) = \mathcal{X} := \left\{ y \in \mathbb{R}^n : \|y\| = R \text{ and } \sum_{k=1}^n (y_k^2 + y_{k+1}^2) y_{k-1}^2 \neq 0 \right\}, \qquad (5.2)$$

where $R = \|x\|$. Furthermore the random splitting of conservative Lorenz-96 is uniquely ergodic on $\mathcal{X}$: for all $h > 0$ the volume form λ is the unique P_h-invariant probability measure on $\mathcal{X}$.

Corollary 5.3. *For all $h > 0$ the volume form λ on $S^{n-1}(R) := \{x \in \mathbb{R}^n : \|x\| = R\}$ is the unique ergodic P_h-invariant probability measure on $S^{n-1}(R)$ that is absolutely continuous with respect to λ.*

Proof. $\mathcal{X}$ in (5.2) is the complement of a closed, measure zero subset of $S^{n-1}(R)$. Thus λ on $\mathcal{X}$ agrees with the volume form, also denoted λ, on $S^{n-1}(R)$. In particular, λ is an ergodic invariant measure on $S^{n-1}(R)$ by Proposition 5.2. Since ergodic invariant measures are mutually singular, see e.g. [23], any other ergodic invariant measure on $S^{n-1}(R)$ must be singular with respect to λ.   □

Proof of Proposition 5.2. Let x be a nonfixed point with $\|x\| = R$. We first prove x can be mapped via the split dynamics to $x_* := (R/\sqrt{n}, \ldots, R/\sqrt{n})$. Since x is a nonfixed point, i.e. $\sum_{k=1}^{n}(x_k^2 + x_{k+1}^2)x_{k-1}^2 \neq 0$, there exists k such that $x_{k-1} \neq 0$ and x_k or x_{k+1} is nonzero. Now, since $\varphi^{(k)}$ is a rotation in the (x_k, x_{k+1})-plane with angular velocity x_{k-1}, there is a t_k such that both k and $k + 1$ coordinates of $\varphi^{(k)}(x, t_k)$ are nonzero. By the same argument there is a t_{k+1} such that the $k, k + 1$, and $k + 2$ coordinates of $x^{(k+1)} = \varphi^{(k+1)}(\varphi^{(k)}(x, t_k), t_{k+1})$ are nonzero. Continuing this way, we see x can be made to have nonzero coordinates in a finite number of steps.

Now since $\|x\| = R$, there exists an index k such that $|x_k| \geq R/\sqrt{n}$. If $k = n$, rotate in the $(n - 1, n)$-plane so that the nth coordinate of x becomes $R/\sqrt{n}$. If $k < n$, rotate in the $(k, k + 1)$-plane so that the $k + 1$ coordinate of x becomes $R/\sqrt{n}$, then rotate in the $(k + 1, k + 2)$-plane so that the $k + 2$ coordinate of x becomes $R/\sqrt{n}$, and so on until the nth-coordinate of x becomes $R/\sqrt{n}$. Such rotations are always possible because all coordinates of x are nonzero by the preceding argument. Thus, whether $k = n$ or $k < n$ we can evolve x via the split dynamics so that its last coordinate, x_n, is $R/\sqrt{n}$. In particular, there now must exist an index $k < n$ such that $|x_k| \geq R/\sqrt{n}$. By the same procedure, and without disturbing the last coordinate, we can use rotations to make the $n - 1$ coordinate of x equal $R/\sqrt{n}$. Iterating this process maps x to x_* in a finite number of steps. Since x was arbitrary it follows that every nonfixed point with norm R belongs to the same orbit, which is precisely the set $\mathcal{X}$ defined in (5.2).

Next we prove there is at most one P_h-invariant measure on $\mathcal{X}$. First note that since the split dynamics are all rotations, the above procedure mapping any arbitrary x in $\mathcal{X}$ to x_* can be done using strictly positive times. Furthermore, by direct observation, the matrix of splitting vector fields

$$
\begin{pmatrix} | & | & & | \\ V_1(x) & V_2(x) & \cdots & V_{n-1}(x) \\ | & | & & | \end{pmatrix} = \begin{pmatrix} x_2 x_n & 0 & \ldots & 0 \\ -x_1 x_n & x_3 x_1 & \ldots & 0 \\ 0 & -x_2 x_1 & \cdots & \vdots \\ \vdots & \vdots & \ddots & x_n x_{n-2} \\ 0 & 0 & & -x_{n-1} x_{n-2} \end{pmatrix}
$$

has rank $n - 1$ whenever all x_k are nonzero. In particular, since $\mathcal{X}$ is an open subset of the sphere of radius R and therefore itself an $n - 1$-dimensional manifold, the splitting vector fields V_k span $T_{x_*}\mathcal{X}$. Hence $\text{Lie}_{x_*}(\mathcal{V}) = T_{x_*}\mathcal{X}$. By Corollary 3.7, P_h has at most one invariant measure on $\mathcal{X}$.

We next show Lebesgue measure, Leb, in $\mathbb{R}^n$ is P_h-invariant. Let $S^{n-1}(R)$ denote the sphere of radius R in $\mathbb{R}^n$ and let $\text{Leb}_t^{(k)} := (\varphi_t^{(k)})_\# \text{Leb}$ be the pushforward of Leb by $\varphi_t^{(k)}$. Since the V_k in (5.1) are divergence free, the continuity equation, intended in

the weak sense,[6] becomes

$$0 = \partial_t \operatorname{Leb}_t^{(k)} + \operatorname{div}\left(V_k \operatorname{Leb}_t^{(k)}\right) = \partial_t \operatorname{Leb}_t^{(k)} + \nabla \operatorname{Leb}_t^{(k)} \cdot V_k. \tag{5.3}$$

The latter is a transport equation with constant initial condition $\operatorname{Leb}_0^{(k)} \equiv 1$ and hence $\operatorname{Leb}_t^{(k)} = \operatorname{Leb}$ for all t. Because the trajectories of all V_k conserve the energy $\|x\|$, we fiber $\mathbb{R}^n$ using spherical coordinates $(r, \vartheta) \in \mathbb{R}_+ \times S^{n-1}(R)$. In these coordinates, we have that $V_k(r, \vartheta) = 0\,\partial_r + r v_k(\vartheta)\nabla_\vartheta$ and by a change of coordinates of the divergence operator the stationarity equation becomes

$$0 = \operatorname{div}\left(V_k(x)\lambda(x)\right) = u(r)w(\vartheta)\operatorname{div}_\vartheta\left(\lambda(r, \vartheta)v_k(\vartheta)\right), \tag{5.4}$$

where $\operatorname{div}_\vartheta$ denotes the angular terms of the divergence in spherical coordinates, and $u(r)$, $w(\vartheta)$ result from the change of variables. Hence, we can factor the solution $\lambda(r, \vartheta) = \bar\lambda(\vartheta|r) \cdot \mu_R(dr) = \bar\lambda(\vartheta) \cdot \mu_R(dr)$, where $\bar\lambda(\vartheta|r)$ is the conditional density of Lebesgue measure on a fiber. The measure $\bar\lambda$ solves $w(\vartheta)\operatorname{div}_\vartheta(\bar\lambda(\vartheta)v_k(\vartheta)) = 0$ and is therefore invariant under the flows $\varphi_t^{(k)}$. By rotational symmetry of Leb, we must have that $\bar\lambda(\vartheta)$ is the volume form on $S^{n-1}(R)$. And since $\mathcal{X}$ is a full-measure open subset of $S^{n-1}(R)$, the volume form λ on $\mathcal{X}$ is just the restriction of $\bar\lambda$ to $\mathcal{X}$. Thus λ is also invariant under the flows and is therefore the unique P_h-invariant measure on $\mathcal{X}$.   □

6. Galerkin Approximations of 2D Euler

The 2D Euler equations on the torus $\mathbb{T}$ are obtained from the 2D Navier–Stokes equations (1.1) by dropping the dissipative and forcing terms:

$$\begin{cases} \partial_t u + (u \cdot \nabla)u = -\nabla p \\ \operatorname{div}(u) := \nabla \cdot u = 0 \end{cases} \tag{6.1}$$

where, as before, $u : \mathbb{T} \times \mathbb{R} \to \mathbb{R}^2$ is the fluid velocity, $p : \mathbb{T} \times \mathbb{R} \to \mathbb{R}$ the fluid pressure, and

$$(u \cdot \nabla)u = (u_1\partial_1 u_1 + u_2\partial_2 u_1, \, u_1\partial_1 u_2 + u_2\partial_2 u_2).$$

In this section we construct a convenient random splitting of (6.1). To do so we first write (6.1) in vorticity form and apply the Fourier transform. This yields an infinite system of ODEs which we truncate to systems of arbitrary finite size, referred to throughout as Galerkin approximations. Finally, we split these Galerkin approximations and apply the results of Sects. 3 and 4 to the associated random splitting.

---

[6] This equation should be interpreted as an equation on measures or, equivalently, as holding in the weak sense. In other words, the left and right side are equal when integrated against any compactly supported, smooth test function.

6.1. Constructing the splitting. The vorticity formulation of (6.1) is obtained by taking the curl of velocity. Specifically, setting $q := \mathrm{curl}(u) := \partial_2 u_1 - \partial_1 u_2$, equation (6.1) becomes

$$\begin{cases} \partial_t q + (\mathcal{K}q \cdot \nabla)q = 0, \\ \mathrm{div}(q) = 0, \end{cases} \tag{6.2}$$

where $\mathcal{K} := \nabla^\perp(-\Delta)^{-1}$ with $\nabla^\perp := (\partial_2, -\partial_1)$. To express (6.2) in Fourier space, set $\mathbb{Z}_\infty^2 := \mathbb{Z}^2 \setminus \{(0,0)\}$ and let $\{e_j\}_{j \in \mathbb{Z}_\infty^2}$ be the orthonormal basis of $L^2(\mathbb{T}, \mathbb{R})$ given by $e_j(x) := (2\pi)^{-1} \exp(ix \cdot j)$. Then $q(x,t) = \sum_{j \in \mathbb{Z}_\infty^2} q_j(t)e_j(x)$ where

$$q_j(t) := \langle q, e_j \rangle_{L^2} = \int_{\mathbb{T}} q(x,t)\overline{e}_j(x)dx$$

is the jth Fourier mode of q. Here $\langle \cdot, \cdot \rangle_{L^2}$ is the standard inner product on $L^2(\mathbb{T}, \mathbb{R})$ with $\overline{e}_j$ denoting the complex conjugate of e_j. The jth Fourier mode of $(\mathcal{K}q \cdot \nabla)q$ is

$$\langle (\mathcal{K}q \cdot \nabla)q, e_j \rangle_{L^2} = \sum_{k+\ell=j} C_{k\ell} q_k q_\ell$$

where

$$C_{k\ell} := \frac{\langle k, \ell^\perp \rangle}{4\pi} \left(\frac{1}{|k|^2} - \frac{1}{|\ell|^2} \right) \tag{6.3}$$

with $\langle \cdot, \cdot \rangle$ the standard inner product in $\mathbb{R}^2$, $\ell^\perp := (\ell_2, -\ell_1)$, and $|\ell|^2 := \ell_1^2 + \ell_2^2$. Therefore

$$\sum_j \dot{q}_j e_j = \partial_t q = -(\mathcal{K}q \cdot \nabla)q = -\sum_j \left(\sum_{k+\ell=j} C_{k\ell} q_k q_\ell \right) e_j$$

and hence $\dot{q}_j = -\sum_{k+\ell=j} C_{k\ell} q_k q_\ell$. Moreover, since q is real-valued,

$$\sum_j q_j e_j = q = \overline{q} = \sum_j \overline{q}_j e_{-j}$$

which gives $q_j = \overline{q}_{-j}$. In particular,

$$\dot{q}_j = \dot{\overline{q}}_{-j} = -\sum_{j+k+\ell=0} C_{k\ell} \overline{q}_k \overline{q}_\ell.$$

Writing each Fourier mode $q_j = a_j + ib_j$ in terms of real and imaginary parts then gives

$$\dot{a}_j + i\dot{b}_j = \dot{q}_j = -\sum_{j+k+\ell=0} C_{k\ell}(a_k - ib_k)(a_\ell - ib_\ell)$$

$$= \sum_{j+k+\ell=0} C_{k\ell}(b_k b_\ell - a_k a_\ell) + i \sum_{j+k+\ell=0} C_{k\ell}(a_k b_\ell + a_\ell b_k).$$

Thus the Fourier modes of solutions to the Euler equation in vorticity form satisfy

$$\begin{cases} \dot{a}_j = \sum_{j+k+\ell=0} C_{k\ell}(b_k b_\ell - a_k a_\ell) \\ \dot{b}_j = \sum_{j+k+\ell=0} C_{k\ell}(a_k b_\ell + a_\ell b_k) \end{cases} \tag{6.4}$$

for all $j \in \mathbb{Z}_\infty^2$. While (6.4) could be studied as is, notice the constraint $q_{-j} = \overline{q}_j$ implies $a_{-j} = a_j$ and $b_{-j} = -b_j$, which introduces redundancy in (6.4). Therefore we restrict to the subset

$$\mathbb{Z}_+^2 := \{j \in \mathbb{Z}^2 : j_2 > 0\} \cup \{j \in \mathbb{Z}^2 : j_2 = 0 \text{ and } j_1 > 0\}.$$

Specifically, by straightforward computation together with the identities $a_{-j} = a_j$, $b_{-j} = -b_j$, and $C_{k\ell} = C_{-k,-\ell} = -C_{-k,\ell} = -C_{k,-\ell}$, the system (6.4) can be re-expressed as

$$\begin{cases} \dot{a}_j = \sum_{j+k-\ell=0} C_{k\ell}(a_k a_\ell + b_k b_\ell) + \sum_{j-k-\ell=0} C_{k\ell}(b_k b_\ell - a_k a_\ell) \\ \dot{b}_j = \sum_{j+k-\ell=0} C_{k\ell}(a_k b_\ell - b_k a_\ell) - \sum_{j-k-\ell=0} C_{k\ell}(a_k b_\ell + b_k a_\ell) \end{cases} \tag{6.5}$$

for all $j \in \mathbb{Z}_+^2$ with each sum running over all pairs $k, \ell \in \mathbb{Z}_+^2$ satisfying the specified identity. To split (6.5) note that for any $j, k, \ell \in \mathbb{Z}_+^2$ satisfying $j + k - \ell = 0$ (and hence $\ell - j - k = 0$) we can isolate from the above sums exactly 6 equations involving only these indices:

$$\begin{aligned} \dot{a}_j &= C_{k\ell}(a_k a_\ell + b_k b_\ell), & \dot{a}_k &= C_{j\ell}(a_j a_\ell + b_j b_\ell), & \dot{a}_\ell &= C_{jk}(b_j b_k - a_j a_k), \\ \dot{b}_j &= C_{k\ell}(a_k b_\ell - b_k a_\ell), & \dot{b}_k &= C_{j\ell}(a_j b_\ell - b_j a_\ell), & \dot{b}_\ell &= -C_{jk}(a_j b_k + b_j a_k). \end{aligned} \tag{6.6}$$

For reasons to be made clear shortly, we recombine (6.6) into 4 groups of 3 equations:

$$\begin{cases} \dot{a}_j = C_{k\ell} a_k a_\ell \\ \dot{a}_k = C_{j\ell} a_j a_\ell \\ \dot{a}_\ell = -C_{jk} a_j a_k \end{cases} \begin{cases} \dot{a}_j = C_{k\ell} b_k b_\ell \\ \dot{b}_k = C_{j\ell} a_j b_\ell \\ \dot{b}_\ell = -C_{jk} a_j b_k \end{cases} \begin{cases} \dot{b}_j = C_{k\ell} a_k b_\ell \\ \dot{a}_k = C_{j\ell} b_j b_\ell \\ \dot{b}_\ell = -C_{jk} b_j a_k \end{cases} \begin{cases} \dot{b}_j = -C_{k\ell} b_k a_\ell \\ \dot{b}_k = -C_{j\ell} b_j a_\ell \\ \dot{a}_\ell = C_{jk} b_j b_k \end{cases}. \tag{6.7}$$

Let $V_{a_j a_k a_\ell}$, $V_{a_j b_k b_\ell}$, $V_{b_j a_k b_\ell}$, and $V_{b_j b_k a_\ell}$ be the vector fields associated to the equations of (6.7) from left to right. For example, $V_{a_j a_k a_\ell}$ is the vector field on $\mathbb{R}^\infty$ mapping the a_j coordinate to $-C_{k\ell} a_k a_\ell$, the a_k coordinate to $-C_{j\ell} a_j a_\ell$, the a_ℓ coordinate to $-C_{jk} a_j a_k$, and all other coordinates to 0. These are the *splitting vector fields*. Our sought-after splitting is

$$V = \sum_{j+k-\ell=0} V_{a_j a_k a_\ell} + V_{a_j b_k b_\ell} + V_{b_j a_k b_\ell} + V_{b_j b_k a_\ell}, \tag{6.8}$$

where V is the vector field associated to (6.5). As noted earlier, our focus will be on finite truncations of the infinite-dimensional system (6.5). Thus we fix an integer $N \geq 2$

and define the *N*th *Galerkin approximation* of (6.5) to be (6.5) with indices restricted
to the set

$$\mathbb{Z}_N^2 := \left\{ j \in \mathbb{Z}_+^2 : \max\{|j_1|, |j_2|\} \leq N \right\}.$$

The splitting (6.8) remains valid in this finite-dimensional setting, bearing in mind that
now all indices lie in $\mathbb{Z}_N^2$. By a slight abuse of notation, we denote the finite-dimensional
counterpart of V by V and similarly for the splitting vector fields. Thus our family of
splitting vector fields is

$$\mathcal{V} = \left\{ V_{a_j a_k a_\ell}, V_{a_j b_k b_\ell}, V_{b_j a_k b_\ell}, V_{b_j b_k a_\ell} : j, k, \ell \in \mathbb{Z}_N^2 \text{ and } j + k - \ell = 0 \right\}. \quad (6.9)$$

Since $\mathbb{Z}_N^2$ has cardinality $2N(N+1)$ and each index $j \in \mathbb{Z}_N^2$ has an associated a_j and
b_j coordinate, these are all vector fields on $\mathbb{R}^n$, where throughout this section we set
$n := 4N(N+1)$. We also abuse notation by conflating elements j in $\mathbb{Z}_N^2$ with elements
j in $\{1, \ldots, n/2\}$, which can be formalized via any bijection between the two sets.
Moreover, we denote elements of $\mathbb{R}^n$ by $q = (a_j, b_j)_{j=1}^{n/2}$. This reflects that the a_j and
b_j coordinates of q in $\mathbb{R}^n$ are in one-to-one correspondence with the real and imaginary
parts of the jth mode of q.

Remark 6.1. There are many possible splittings of a given equation. For the Euler equa-
tions, we made the particular choice we have so that both energy and enstrophy are
conserved but the dynamics of each splitting are still relatively easily understood. We
could have further decomposed the three-dimensional dynamics in the above splitting
into a number of two-dimensional dynamics, similar in spirit to the decomposition into
rotations used in Lorenz-96. However, that would have necessitated only conserving
either the energy or the enstrophy.

6.2. Conservation and convergence. The conservative Lorenz-96 dynamics discussed
in Sect. 5 conserves Euclidean norm (energy in that case) and therefore remains on
whichever sphere it starts on. So too do the flows of each of the splitting vector fields
(5.1). We now show a similar thing is true for Galerkin approximations of 2D Euler.
Define the *energy* and *enstrophy* of $q = (a_j, b_j)_{j=1}^{n/2}$ by

$$E(q) := \sum_{j \in \mathbb{Z}_N^2} \frac{a_j^2 + b_j^2}{|j|^2} \quad \text{and} \quad \mathcal{E}(q) := \sum_{j \in \mathbb{Z}_N^2} a_j^2 + b_j^2, \quad (6.10)$$

respectively (note the aforementioned conflation of j in $\mathbb{Z}_N^2$ and $j \in \{1, \ldots, n/2\}$ in
the summations). Straightforward computation shows that for all $j, k, \ell \in \mathbb{Z}_N^2$ with
$j + k - \ell = 0$,

$$C_{k\ell} + C_{j\ell} - C_{jk} = \frac{C_{k\ell}}{|j|^2} + \frac{C_{j\ell}}{|k|^2} - \frac{C_{jk}}{|\ell|^2} = 0,$$

which in turn implies that under the dynamics (6.5),

$$\partial_t E(q) = \partial_t \mathcal{E}(q) = 0$$

for all $q \in \mathbb{R}^n$. That is, both energy and enstrophy are conserved by the true dynamics and the set

$$\mathcal{Q}_0(E, \mathcal{E}) := \left\{ q \in \mathbb{R}^n : E(q) = E, \ \mathcal{E}(q) = \mathcal{E} \right\}. \tag{6.11}$$

is invariant under (6.5). This is a well-established property of the 2D Euler equations. Moreover, if we flow by $V_{a_j a_k a_\ell}$ starting from q for any $j, k, \ell \in \mathbb{Z}_N^2$ with $j + k - \ell = 0$, then

$$\tfrac{1}{2} \partial_t E(q) = \frac{a_j \dot{a}_j}{|j|^2} + \frac{a_k \dot{a}_k}{|k|^2} + \frac{a_\ell \dot{a}_\ell}{|\ell|^2} = \left(\frac{C_{k\ell}}{|j|^2} + \frac{C_{j\ell}}{|k|^2} - \frac{C_{jk}}{|\ell|^2} \right) a_j a_k a_\ell = 0,$$

and similarly $\partial_t \mathcal{E}(q) = 0$. The same computation shows energy and enstrophy are conserved by *all* of the splitting vector fields in $\mathcal{V}$, which provides the motivation for recombining (6.6) as (6.7) in the first place. In particular, we have

Proposition 6.2. *All of the finite time convergence results of Sect. 4 apply to the random splitting (6.8) of every Galerkin approximation of 2D Euler starting from any initial condition.*

Proof. The splitting vector fields are smooth and Assumption 1 is satisfied since every $\mathcal{V}$-orbit lies on a sphere, so the conclusions of Theorems 4.1 and 4.5 both hold. $\square$

6.3. Ergodicity. Fix energy and enstrophy values E and $\mathcal{E}$ and set $\mathcal{Q}_0 := \mathcal{Q}_0(E, \mathcal{E})$. $\mathcal{Q}_0$ is an $n - 2$-dimensional submanifold of $\mathbb{R}^n$ where, recall, $n := 4N(N + 1)$; denote its volume form by λ. As with conservative Lorenz-96, the Nth Galerkin approximation of 2D Euler has points q in $\mathcal{Q}_0$ whose $\mathcal{V}$-orbits are not dense in $\mathcal{Q}_0$. For example, any q with exactly one nonzero coordinate is a fixed point of (6.5) and of all the equations (6.7). In this subsection we characterize these points and prove there is exactly one $\mathcal{V}$-orbit $\mathcal{Q}$ on $\mathcal{Q}_0$ such that $\lambda(\mathcal{Q}) = 1$. By a slight abuse of notation we denote the restriction of λ to $\mathcal{Q}$ by λ as well. We then show there exists a unique P_h-invariant measure on $\mathcal{Q}$ – and hence on $\mathcal{Q}_0$ – that is absolutely continuous with respect to λ on $\mathcal{Q}_0$.

To make the above statements precise, we begin by enumerating the coordinates of $q \in \mathbb{R}^n$ by extending the indices $j \in \mathbb{Z}_N^2$ with an element $\chi \in \{+, -\}$ which denotes the real $(+)$ or imaginary $(-)$ part of the corresponding mode. Then, for $\mathbf{j} = (j, \chi) \in \mathbb{Z}_N^2 \times \{+, -\}$, we define the *type* of such coordinates via the function $\mathrm{T}(\mathbf{j}) = \chi$ so that $q_{\mathbf{j}}$ is identified with a_j if $\mathrm{T}(\mathbf{j}) = +$ and with b_j if $\mathrm{T}(\mathbf{j}) = -$. For $q \in \mathbb{R}^n$, denote by

$$\mathcal{A}(q) := \left\{ \mathbf{j} \in \mathbb{Z}_N^2 \times \{+, -\} \ : \ q_{\mathbf{j}} \neq 0 \right\}$$

the set of "active" coordinates. To streamline our analysis, we define the following operation to expand the set $\mathcal{A}$:

$$\mathcal{A} \oplus \ell := \begin{cases} \mathcal{A} \cup \{\ell\} & \text{if } \ell \in \{j + k, \ j - k\} \cap \mathbb{Z}_N^2 \text{ for } \mathbf{j}, \mathbf{k} \in \mathcal{A}, \, C_{jk} \neq 0, \, \mathrm{T}(\mathbf{j}) \cdot \mathrm{T}(\mathbf{k}) = \mathrm{T}(\ell), \\ \mathcal{A} & \text{else}, \end{cases} \tag{6.12}$$

where $\mathrm{T}(\mathbf{j}) \cdot \mathrm{T}(\mathbf{k})$ is $+$ if $\mathrm{T}(\mathbf{j}) = \mathrm{T}(\mathbf{k})$ and $-$ if $\mathrm{T}(\mathbf{j}) \neq \mathrm{T}(\mathbf{k})$. This operation corresponds to extending the nonzero coordinates of q from $\mathbf{j}, \mathbf{k}$ to ℓ by letting a triple $\iota = \mathbf{j k \ell}$ interact.

We assume that the initial condition is sufficiently nondegenerate, as stated in the following assumption similar to the one made in [24, Thm. 2.1].

Definition 6.3 [Nondegenerate point]. A point q in $\mathcal{Q}_0$ is nondegenerate if there exists $M \in \mathbb{N}$, $j^* \in \mathbb{Z}_N^2$ with $|j^*|^2 > 1$, and an ordered set of indices $(\ell_i)_{i=1}^M$ in $\mathbb{Z}_N^2 \times \{+, -\}$ such that

$$\left\{(1, 0, +), (0, 1, +), (j^*, -)\right\} \subseteq \left((\mathcal{A}(q) \oplus \ell_1) \oplus \ell_2\right) \cdots \oplus \ell_M. \qquad (6.13)$$

Definition 6.4 [Generic point]. A point in $\mathbb{R}^n$ is generic if all of its coordinates are nonzero.

Remark 6.5. Every point with all coordinates nonzero is a nonfixed point of conservative Lorenz-96; similarly, every generic point in $\mathcal{Q}_0$ is nondegenerate. However, comparing (6.13) with (5.2), we see the conditions defining nondegenerate points in $\mathcal{Q}_0$ are more complicated than the easily characterized nonfixed points of conservative Lorenz-96. The difference is that, unlike spheres in conservative Lorenz-96, there are proper subspaces of $\mathcal{Q}_0$ which are invariant for our splitting of the Euler dynamics but are not fixed points. One such subspace is the collection of purely real points; another is the purely imaginary points.

The following analogs of Proposition 5.2 and Corollary 5.3 are the main results of this subsection.

Proposition 6.6. *Every nondegenerate point in $\mathcal{Q}_0$ belongs to the same $\mathcal{V}$-orbit, $\mathcal{Q}$, and for all $h > 0$ there exists a unique P_h-invariant probability measure on $\mathcal{Q}$. Furthermore, this unique invariant measure is absolutely continuous with respect to the volume form on $\mathcal{Q}$.*

Proof. By Proposition 6.9 there is a q^* in $\mathcal{Q}_0$ such that every nondegenerate point in $\mathcal{Q}_0$ belongs to the $\mathcal{V}$-orbit $\mathcal{Q} := \mathcal{Q}(q^*)$, and for every q in $\mathcal{Q}$ there is an $m \in \mathbb{N}$ and a $t \in \mathbb{R}_+^{mn}$ satisfying $\Phi^m(q, t) = q^*$. By Lemma 6.15 the splitting vector fields span the tangent space of $\mathcal{Q}$ at generic points; in particular, the Lie bracket condition holds at every generic point. Thus, since the vector fields in $\mathcal{V}$ are analytic, Corollary 3.9 implies P_h has at most one invariant probability measure on $\mathcal{Q}$, which is necessarily the one identified by Lemma 6.14.   □

Corollary 6.7. *For all $h > 0$ the measure from Proposition 6.6 is the unique P_h-invariant ergodic probability measure on $\mathcal{Q}_0$ that is absolutely continuous with respect to the volume form on $\mathcal{Q}_0$.*

Proof. Let λ denote volume form on $\mathcal{Q}_0$. Since $\mathcal{Q}$ contains all generic points in $\mathcal{Q}_0$, it is an open subset of $\mathcal{Q}_0$ satisfying $\lambda(\mathcal{Q}) = 1$. In particular, the unique invariant measure on $\mathcal{Q}$ from Proposition 6.6 is an ergodic invariant measure on $\mathcal{Q}_0$. Since ergodic invariant measures are mutually singular, see e.g. [23], any other ergodic invariant measure on $\mathcal{Q}_0$ must be singular with respect to λ.   □

Remark 6.8. Continuing in the spirit of Remark 6.1, we observe the splitting in (6.7) splits q_j into its real and imaginary parts. We could have chosen another basis of $\mathbb{C}$ and even randomized over this choice for each evolution of an interacting triple (j, k, ℓ). More explicitly, if we define $e(\vartheta) = cos(\vartheta) + i \sin(\vartheta)$ then $e(\vartheta)$ and $e(\vartheta + \frac{\pi}{2})$ form an orthonormal basis of $\mathbb{C}$ for any ϑ. Then we can drive a system analogous to (6.7) by setting $q_\ell = a_\ell^\vartheta e(\vartheta) + b_\ell^\vartheta e(\vartheta + \frac{\pi}{2})$. As the form is similar to (6.7), the results of the paper extend to this system. In particular, by randomizing the choice of ϑ for each such triple (j, k, ℓ), we can relax the characterization of nondegenerate points in Definition 6.3 by destroying some of the invariant structures discussed in Remark 6.5 which obstruct controllability starting from some initial conditions.

6.3.1. Controllability. In this section, we prove controllability of the dynamics (6.7). By conservation of energy and enstrophy, the $\mathcal{V}$-orbit of an initial condition $q^{(0)}$ in $\mathcal{Q}_0$ is contained in $\mathcal{Q}_0$. Recalling the definition of extended indices in Sect. 6.3, we define the set of *interacting coordinate triples*

$$
\mathcal{I} := \big\{ (\mathbf{j}, \mathbf{k}, \boldsymbol{\ell}) \in (\mathbb{Z}_N^2 \times \{+, -\})^3 \; : \; j + k = \ell, \; (C_{jk}, C_{j\ell}, C_{k\ell})
$$
$$
\neq (0, 0, 0), \; \mathrm{T}(\mathbf{j}) \cdot \mathrm{T}(\mathbf{k}) = \mathrm{T}(\boldsymbol{\ell}) \big\}.
$$

Then, for any such triple of interacting indices $\iota \in \mathcal{I}$ we denote by $\varphi_t^\iota \; : \; \mathcal{Q}_0 \to \mathcal{Q}_0$ the flow of the ODEs (6.7) evolving the corresponding coordinates. The dynamics we consider is then obtained by cycling through the set $\mathcal{I}$ in a fixed or random order. For any $\iota \in \mathcal{I}$ we denote by $\Phi_t^\iota \; : \; \mathcal{Q}_0 \to \mathcal{Q}_0$ the flow of (6.7) after one such full cycle where the flow times are chosen as

$$
\tau^\xi = \begin{cases} t & \text{if } \xi = \iota \,, \\ 0 & \text{else} \,, \end{cases} \tag{6.14}
$$

so that for any $q \in \mathcal{Q}_0$, $\Phi_t^\iota(q) = \varphi_t^\iota(q)$.

Let $q^* = (a_j^*, b_j^*)_{j=1}^{n/2}$ be the point in $\mathcal{Q}_0$ defined as follows:

$$
q_{(1,0)}^* = q_{(0,1)}^* = (a^*, 0) \,, \qquad q_{(N,N)}^* = (0, b^*) \,, \tag{6.15}
$$

for $a^*, b^* \geq 0$ and $q_j^* = (0, 0)$ for all other $j \in \mathbb{Z}_N^2$. We show below that for any nondegenerate initial condition $q^{(0)} \in \mathcal{Q}_0$ the system can be driven to this unique point q^*.

Proposition 6.9. *For any nondegenerate point $q^{(0)} = (a_j^{(0)}, b_j^{(0)})_{j=1}^{n/2}$ in $\mathcal{Q}_0$ there exists $M \in \mathbb{N}$ and a joint sequence of transition times and coordinate triples $\{(\iota(m), \tau(m))\}_{m=1}^M$ such that*

$$
\Phi_{\tau(M)}^{\iota(M)} \circ \cdots \circ \Phi_{\tau(1)}^{\iota(1)}(q^{(0)}) = q^* \,. \tag{6.16}
$$

Thus every nondegenerate point belongs to the same orbit, $\mathcal{Q} := \mathcal{Q}(q^*)$. Furthermore, for every q in $\mathcal{Q}$ there is an $m \in \mathbb{N}$ and a $t \in \mathbb{R}_+^{mn}$ such that $\Phi^m(q, t) = q^*$.

Recall from Remark 2.3 that the only property of the exponential distribution used in this proof is the fact that it has a density around 0, allowing to choose the flow of some of the split vector fields to be the identity as, e.g., in (6.14). This comment also applies to the proof of Proposition 5.2 in the previous section. We further note that, since the trajectories of each of the $\varphi^{\iota(m)}$ in the above theorem are periodic (see Lemmas 6.11 and 6.12), each of these transformations can be inverted by choosing complementary transition times to $\tau(m)$. Inverting the order of the transformations yields the converse statement:

Corollary 6.10. *For any nondegenerate point $q^{(0)} = (a_j^{(0)}, b_j^{(0)})_{j=1}^{n/2}$ in $\mathcal{Q}_0$ there exists $M \in \mathbb{N}$ and a joint sequence of transition times and coordinate triples $\{(\tilde{\iota}(m), \tilde{\tau}_0(m))\}_{m=1}^M$ such that*

$$
\Phi_{\tilde{\tau}(M)}^{\tilde{\iota}(M)} \circ \cdots \circ \Phi_{\tilde{\tau}(1)}^{\tilde{\iota}(1)}(q^*) = q^{(0)} \,.
$$

While the Corollary 6.10 will not be used in the remainder of the paper, it offers an alternative to Theorem 3.8 in proving that, when applying Corollary 3.7, it is sufficient to verify that Lie bracket condition holds at *any* point in $\mathcal{Q}$, not necessarily at q^*.

Proof of Proposition 6.9. We prove the first statement by first evolving the initial condition $q^{(0)}$ into a sufficiently nondegenerate state $q^{(1)}$, and then by sequentially shrinking the set of active components of the coordinate vector q to the ones listed in (6.15). We realize this program by following, in order, the sequence of steps described below, represented schematically in Fig. 1:

(0) If it is not the case at initialization, Lemma B.1 shows that we can "prepare" our state by evolving $q^{(0)}$ into $q^{(1)}$ such that

$$a_{(1,0)}^{(1)}, b_{(1,0)}^{(1)}, a_{(0,1)}^{(1)}, b_{(0,1)}^{(1)}, a_{(1,1)}^{(1)}, b_{(1,1)}^{(1)} \neq 0, \tag{6.17}$$

as represented in Fig. 1a.
(1) As shown in Lemma B.2, we can then transform $q^{(1)}$ into $q^{(2)}$ with the property

$$q_j^{(2)} = (0,0) \quad \text{for all } j \in \mathbb{Z}_N^2 \setminus \{(0,1),(1,0),(1,1),(N,N),(-N,N)\}, \tag{6.18}$$

as represented in Fig. 1b, and

$$a_{(1,0)}^{(2)}, b_{(1,0)}^{(2)}, a_{(0,1)}^{(2)}, b_{(0,1)}^{(2)}, a_{(1,1)}^{(2)}, b_{(1,1)}^{(2)} \neq 0. \tag{6.19}$$

(2) Lemma B.3 shows that we can then "transfer" the amplitude from modes $a_{(-N,N)}$, $b_{(-N,N)}$, $a_{(N,N)}$ to mode $b_{(N,N)}$ i.e., we can reach a state $q^{(3)}$ that satisfies

$$q_j^{(3)} = (0,0) \qquad \text{for all } j \in \mathbb{Z}_N^2 \setminus \{(0,1),(1,0),(1,1),(N,N)\}, \tag{6.20}$$

$$q_{(N,N)}^{(3)} = (0, b_{(N,N)}^{(3)}) \quad \text{with } b_{(N,N)}^{(3)} \geq 0. \tag{6.21}$$

This state is represented in Fig. 1c.
(3) Finally, Lemma B.5 shows that we can "transfer" the amplitude from modes $a_{(1,1)}$, $b_{(1,1)}$, $b_{(0,1)}$ and $b_{(1,0)}$ to modes $a_{(0,1)}, a_{(1,0)}, b_{(N,N)}$ so that, after the transfer, $a_{(0,1)} = a_{(1,0)}$ and $a_{(0,1)}, a_{(1,0)}, b_{(N,N)} > 0$ i.e., we reach the unique state q^* from (6.15) (represented in Fig. 1d).

This proves the first part of Proposition 6.9, which immediately implies nondegenerate points in $\mathcal{Q}_0$ belong to $\mathcal{Q} = \mathcal{Q}(q^*)$. Let q be *any* point in $\mathcal{Q}$. By definition there exist m and t in $\mathbb{R}^{mn}$ such that

$$\Phi^m(q,t) = \varphi_{t_{mn}}^{(n)} \circ \cdots \varphi_{t_1}^{(1)}(q) = q^*.$$

Note that the times t_i may be negative; however, by Lemma 6.11 each $\varphi^{(i)}$ is periodic. Thus for every $t_i \leq 0$ there exists a $t_i' > 0$ such that $\varphi_{t_i}^{(i)}(q') = \varphi_{t_i'}^{(i)}(q')$ for all q' in $\mathcal{Q}$. Let t' be t with all $t_i \leq 0$ replaced by t_i'. Then t' is in $\mathbb{R}_+^{mn}$ and $\Phi^m(q,t') = \Phi^m(q,t) = q^*$.
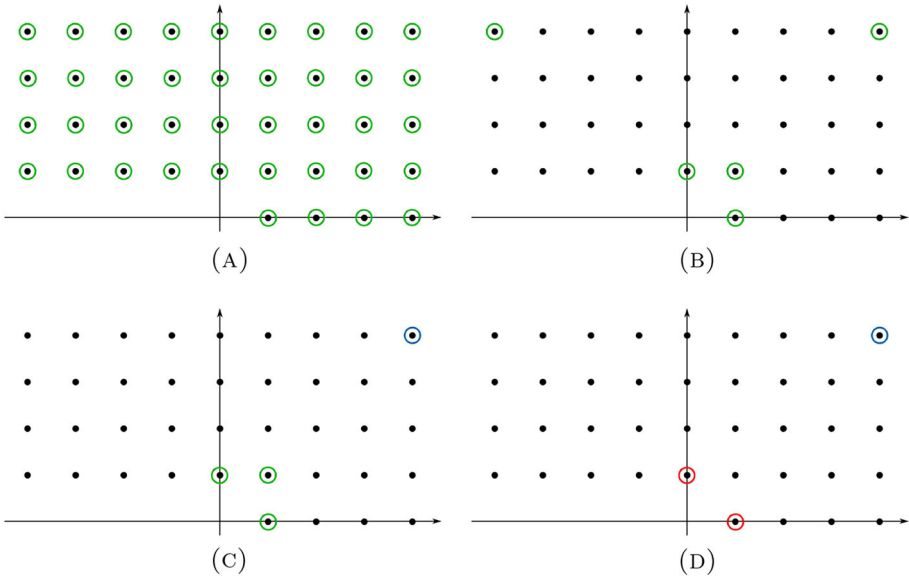$\square$

Fig. 1. Representation of the state of the network in a generic initial state (a), after step 1 of the procedure in the proof of Proposition 6.9 (b), and after step 2 (c) and after step 3 (d) of the same procedure. In the above pictures, each point corresponds to a mode, i.e., an element of $\mathbb{Z}_N^2$ while the color of each circle represents the real/complex value of the corresponding mode: zero (white, no circle), purely imaginary (red), purely real (blue) or having both nonvanishing real and imaginary parts (green)

Defining similarly to (6.22) the operation of removing a coordinate from the set $\mathcal{A}$

$$\mathcal{A} \ominus \ell = \begin{cases} \mathcal{A} \backslash \{\ell\} & \text{if } \ell \in \{j+k, j-k\} \cap \mathbb{Z}_N^2 \text{ for } \mathbf{j}, \mathbf{k} \in \mathcal{A}, C_{jk} \neq 0, \mathrm{T}(\mathbf{j}) \cdot \mathrm{T}(\mathbf{k}) = \mathrm{T}(\ell), \\ \mathcal{A} & \text{else}, \end{cases}$$

(6.22)

we now proceed to *construct* (sequences of) times τ and interacting triples ι such that the transformations $\Phi_\tau^{(\iota)}$ of q implement the operations $\oplus, \ominus$ from (6.22), (6.22) through the flow of (6.7), i.e., such that $\mathcal{A}(q) \oplus \ell = \mathcal{A}(\Phi_\tau^\iota(q))$ or $\mathcal{A}(q) \ominus \ell = \mathcal{A}(\Phi_\tau^\iota(q))$ respectively. To do so we separate the possible interactions between the modes in two types:

$$\begin{aligned} a) \quad & \iota = \mathbf{jk\ell} \in \mathcal{I} \; : \; |j| \neq |k| \neq |\ell|, \\ b) \quad & \iota = \mathbf{jk\ell} \in \mathcal{I} \; : \; |j| = |k| \neq |\ell|. \end{aligned}$$

(6.23)

Note that these two types of interactions are exhaustive, since if $|j| = |k| = |\ell|$, $C_{j\ell} = C_{jk} = C_{k\ell} = 0$.

The following preparatory lemmas describe the properties of these two types of interactions that we will leverage throughout our proof. The first one shows that for interactions of type a), ordering the indices so that $|j| < |k| < |\ell|$, it is always possible to activate all modes $\mathbf{j}, \mathbf{k}, \boldsymbol{\ell}$ or to distribute the amplitude of the k-mode to the j and ℓ-modes reaching, in finite time, a state with $q_k = 0$. As we show in the proof below, while such a point with $q_k = 0$ always exists on the orbits of (6.7), this point is reachable
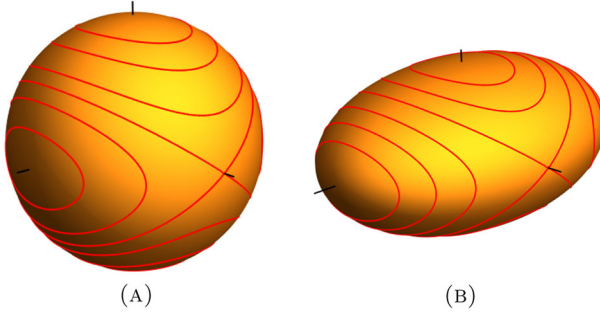
Fig. 2. Orbits $\mathcal{Q}_\iota$ of (6.7) (in red) corresponding in (A) to various values of the energy $\mathcal{E}_\iota(q)$ on the sphere of constant enstrophy $E_\iota(q)$ and in (B) to various values of the enstrophy $E_\iota(q)$ on the ellipsoid of constant energy $\mathcal{E}_\iota(q)$. The axes are, sequentially, q_k, q_j, q_ℓ. The orbit with a degenerate point at the pole of the sphere or ellipsoid corresponds to values of $E_\iota, \mathcal{E}_\iota$ violating (6.24)

in finite time for $\iota = jk\ell \in \mathcal{I}$ with $|j| < |k| < |\ell|$ only if

$$E_\iota(q) \neq |k|^2 \mathcal{E}_\iota(q), \tag{6.24}$$

where $E_\iota(q)$ and $\mathcal{E}_\iota(q)$ denote the energy and enstrophy of the coordinates in $\iota \in \mathcal{I}$:

$$E_\iota(q) := \sum_{\ell \in \iota} |q_\ell|^2, \qquad \mathcal{E}_\iota(q) := \sum_{\ell \in \iota} \frac{|q_\ell|^2}{|\ell|^2}. \tag{6.25}$$

In the following lemma and throughout the section, we abuse notation slightly by defining $\text{sign}(x) = +1$ for $x \in [0, \infty)$ and -1 otherwise.

Lemma 6.11. *Fix $\iota = jk\ell \in \mathcal{I}$ with $|j| < |k| < |\ell|$. Let q be a nondegenerate point in $\mathcal{Q}_0$ satisfying (6.24) and let $q_l = 0$ for at most an index $l \in \{j, k, \ell\}$. Then the orbit of V_ι is periodic and there exist $\tau_-^\iota, \tau_+^\iota \geq 0$ such that*

(a) $\varphi_{\tau_-^\iota}^\iota(q) = q'$ with $q_k' = 0$, $\text{sign}(q_j) = \text{sign}(q_j')$ and $\text{sign}(q_\ell) = \text{sign}(q_\ell')$,
(b) $\varphi_{\tau_+^\iota}^\iota(q) = q''$ with $q_j'', q_k'', q_\ell'' \neq 0$, $\text{sign}(q_j) = \text{sign}(q_j'')$ and $\text{sign}(q_\ell) = \text{sign}(q_\ell'')$.

Furthermore, if $|j|^2 \mathcal{E}_\iota(q) < E_\iota(q) < |k|^2 \mathcal{E}_\iota(q)$, there exists $\tau_=^\iota \geq 0$ such that

(c) $\varphi_{\tau_=^\iota}^\iota(q) = q'''$ with $q_\ell''' = 0$, $\text{sign}(q_j) = \text{sign}(q_j''')$ and $\text{sign}(q_k) = \text{sign}(q_k''')$.

Proof. We consider the intersection between the sphere and the ellipse corresponding to the enstrophy and the energy in the coordinates $\iota = jk\ell \in \mathcal{I}$ of interest, resulting in the set

$$\mathcal{Q}_\iota := \left\{ (q_j', q_k', q_\ell') \in \mathbb{R}^3 \; : \; |q_j'|^2 + |q_k'|^2 + |q_\ell'|^2 = E_\iota(q), \frac{|q_j'|^2}{|j|^2} + \frac{|q_k'|^2}{|k|^2} + \frac{|q_\ell'|^2}{|\ell|^2} = \mathcal{E}_\iota(q) \right\}. \tag{6.26}$$

This set is represented in Fig. 2. We observe that this set has exactly 2 disjoint simply connected components when $|j|^2 \mathcal{E}_\iota(q) < E_\iota(q) < |k|^2 \mathcal{E}_\iota(q)$ and $|k|^2 \mathcal{E}_\iota(q) < E_\iota(q) < |\ell|^2 \mathcal{E}_\iota(q)$. These components are diffeomorphic to S^1. By continuity the dynamics are limited to one such component of $\mathcal{Q}_\iota$. Furthermore, $|\dot{q}|^2$ is uniformly bounded away from

0 on each such component: the fixed points of (6.7) must have at least two coordinates vanishing, which cannot be realized on the curves of interest. Therefore the dynamics on these sets are periodic.

We start by proving part (b) of the lemma. If $q_j, q_k, q_\ell \neq 0$ the result follows by choosing $\tau_+^\iota = 0$. Else, if $q_l = 0$ for $l \in \iota$ the result follows immediately choosing τ_+^ι small enough by combining the continuity of the flow Φ_t^ι and the fact that $\dot{q}_l = C_{l'l''}q_{l'}q_{l''} \neq 0$ for $\{l', l''\} = \iota \setminus \{l\}$.

To prove part (a) we consider the cases where $|j|^2 \mathcal{E}_\iota(q) < E_\iota(q) < |k|^2 \mathcal{E}_\iota(q)$ and $|k|^2 \mathcal{E}_\iota(q) < E_\iota(q) < |\ell|^2 \mathcal{E}_\iota(q)$ separately. In the first case, we see that there is no point $q \in \mathcal{Q}_\iota$ with $q_j = 0$: if that were the case we would have

$$E_\iota(q) = q_k^2 + q_\ell^2 = |k|^2 \left(\frac{q_k^2}{|k|^2} + \frac{q_\ell^2}{|k|^2} \right) > |k|^2 \mathcal{E}_\iota(q), \tag{6.27}$$

contradicting our assumption. Consequently the points $(p_j, 0, p_\ell), (p_j, 0, -p_\ell)$ with $p_\ell > 0$, $\text{sign}(p_j) = \text{sign}(q_j)$ and

$$p_j^2 + p_\ell^2 = E_\iota(q), \qquad \frac{p_j^2}{|j|^2} + \frac{p_\ell^2}{|\ell|^2} = \mathcal{E}_\iota(q), \tag{6.28}$$

belong to the same connected component as q and by the lower bound on the velocity on this connected component both these points are reachable in finite time from q. This also proves part (c) by continuity of the dynamics. The second case where $|k|^2 \mathcal{E}_\iota(q) < E_\iota(q) < |\ell|^2 \mathcal{E}_\iota(q)$ can be handled analogously: in this case we have $\mathcal{Q}_\iota \cap \{q_\ell = 0\} = \emptyset$ and we can reach $(p_j, 0, p_\ell), (-p_j, 0, p_\ell)$ with $p_j > 0$, $\text{sign}(p_\ell) = \text{sign}(q_\ell)$ in finite time. $\square$

The following lemma considers interactions of type b) in (6.23). Recalling the definition $j^\perp := (j_2, -j_1)$ we show that interactions with $|j| = |k| \neq |\ell|$ leave component ℓ fixed and move j, k in a circle at constant angular speed.

Lemma 6.12. *Fix an unordered interacting triple $\iota = jk\ell$ with $|k| = |j|$ and $q_\ell \neq 0$. For all ϑ in $[0, 2\pi)$ there exists $t \geq 0$ such that $\varphi_t^\iota(q) = q'$ with $(q_j', q_k') = \sqrt{q_j^2 + q_k^2}(\cos(\vartheta), \sin(\vartheta))$ and $q_\ell' = q_\ell$.*

Corollary 6.13. *Fix an (unordered) interacting triple $\iota = jk\ell \in \mathcal{I}$ with $|k| = |j|$ and let $q_\ell, q_k \neq 0$. Then there exist $\tau_+^\iota, \tau_-^\iota \geq 0$ such that $(\varphi_{\tau_+^\iota}^\iota(q))_j > 0$ and $(\varphi_{\tau_-^\iota}^\iota(q))_j = 0$.*

Proof of Lemma 6.12. Recall from (6.3) that if $|j| = |k| \neq |\ell|$ we have $C_{jk} = 0$. This implies that, by our choice of $|k| = |j|$, $\dot{q}_\ell = 0$ and $q_\ell' = q_\ell$. Again by (6.3) and since to have an interacting triple $\ell = j + k$ we must have

$$\langle k^\perp, \ell \rangle = \langle k^\perp, k + j \rangle = \langle k^\perp, j \rangle = \langle (k + j)^\perp - j^\perp, j \rangle = \langle \ell^\perp, j \rangle = -\langle j^\perp, \ell \rangle, \tag{6.29}$$

so that

$$C_{k\ell} = \frac{\langle k, \ell^\perp \rangle}{4\pi} \left(\frac{1}{|k|^2} - \frac{1}{|\ell|^2} \right) = -\frac{\langle j, \ell^\perp \rangle}{4\pi} \left(\frac{1}{|j|^2} - \frac{1}{|\ell|^2} \right) = -C_{j\ell}. \tag{6.30}$$

This implies that the dynamics of the vector $\tilde{q} := (q_j, q_k)$ can be written as $\dot{\tilde{q}} = \tilde{C}\tilde{q}^\perp$ for $\tilde{C} := C_{j\ell}q_\ell \neq 0$, proving the claim. $\square$

6.3.2. Existence of invariant measure. As with conservative Lorenz-96, each vector field of the 2D Euler splitting is divergence free and so Lebesgue measure in $\mathbb{R}^n$ is invariant. Consequently, we have

Lemma 6.14. *Let λ denote the Lebesgue measure on $\mathbb{R}^n$. The measure obtained by conditioning λ to lie on $\mathcal{Q} \subset Q_0(E, \mathcal{E})$, (or equivalently conditioned to lie on $Q_0(E, \mathcal{E})$) is P_h-invariant.*

Proof. As in the proof of Proposition 5.2 we have that Lebesgue measure in $\mathbb{R}^n$ is P_h-invariant. Since the vector fields V_k defined in (6.9) are divergence free, the continuity equation[7] reads

$$\partial_t \lambda + \text{div}\,(V_k \lambda) = \partial_t \lambda + \nabla \lambda \cdot V_k = 0. \tag{6.31}$$

Because each flow $\varphi^{(k)}$ conserves energy E and enstrophy $\mathcal{E}$, we locally fiber $\mathbb{R}^n$ using coordinates $(E, \mathcal{E}, \vartheta) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^{n-2}$. In these coordinates, we have $V_k(E, \mathcal{E}, \vartheta) = 0\,\partial_E + 0\,\partial_{\mathcal{E}} + v_k(E, \mathcal{E}, \vartheta)\nabla_\vartheta$ so by a change of coordinates of the divergence operator the stationary equation becomes

$$0 = \text{div}\,(V_k(x)\lambda(x)) = u(E, \mathcal{E}, \vartheta)\,\text{div}_\vartheta\,(\lambda(E, \mathcal{E}, \vartheta)v_k(E, \mathcal{E}, \vartheta)), \tag{6.32}$$

where div_ϑ denotes the "angular" terms of the divergence in $(E, \mathcal{E}, \vartheta)$-coordinates, and $u(E, \mathcal{E}, \vartheta)$ result from the change of variables. Hence, we can factor the solution $\lambda(E, \mathcal{E}, \vartheta) = \bar{\lambda}(\vartheta | E, \mathcal{E}) \cdot \lambda^\perp(E, \mathcal{E})$, where $\bar{\lambda}(\vartheta | E, \mathcal{E})$ is the conditional density of Lebesgue measure on a fiber, solving $u(E, \mathcal{E}, \vartheta)\,\text{div}_\vartheta\,(\bar{\lambda}(\vartheta | E, \mathcal{E})v_k(E, \mathcal{E}, \vartheta)) = 0$ for any choice of $E/(2N^2) < \mathcal{E} < E$. This proves the invariance of $\bar{\lambda}(\vartheta | E, \mathcal{E})$ under the flow map for any value of the flow times τ. The stationarity of $\bar{\lambda}(\vartheta)$ under P_h follows immediately as in Proposition 5.2 □

6.3.3. Spanning. For $j, k, \ell \in \mathbb{Z}_N^2$ with $j + k - \ell = 0$ define $M_{jk\ell}$ to be the matrix

$$
\begin{aligned}
M_{jk\ell} &:= \begin{pmatrix} | & | & | & | \\ V_{a_j a_k a_\ell} & V_{a_j b_k b_\ell} & V_{b_j a_k b_\ell} & V_{b_j b_k a_\ell} \\ | & | & | & | \end{pmatrix} \\
&= \begin{pmatrix} C_{k\ell a_k a_\ell} & C_{k\ell b_k b_\ell} & 0 & 0 \\ 0 & 0 & C_{k\ell a_k b_\ell} & -C_{k\ell b_k a_\ell} \\ C_{j\ell a_j a_\ell} & 0 & C_{j\ell b_j b_\ell} & 0 \\ 0 & C_{j\ell a_j b_\ell} & 0 & -C_{j\ell b_j a_\ell} \\ -C_{jk a_j a_k} & 0 & 0 & C_{jk b_j b_k} \\ 0 & -C_{jk a_j b_k} & -C_{jk b_j a_k} & 0 \end{pmatrix}
\end{aligned} \tag{6.33}
$$

and let $M'_{jk\ell}$ and $M''_{jk\ell}$ be the 4-by-4 and 2-by-4 matrices consisting of the bottom four and bottom two rows of $M_{jk\ell}$, respectively. Straightforward Gaussian elimination shows that M, M', and M'' have ranks 4, 3, and 2 whenever C_{jk}, $C_{j\ell}$, $C_{k\ell}$, a_j, b_j, a_k, b_k, a_ℓ, and b_ℓ are nonzero.

Recalling that a point $q \in \mathbb{R}^n$ is *generic* if all its coordinates are nonzero, we have

Lemma 6.15. *The family of vector fields*

$$\mathcal{V} := \left\{ V_{a_j a_k a_\ell}, V_{a_j b_k b_\ell}, V_{b_j a_k b_\ell}, V_{b_j b_k a_\ell} : j, k, \ell \in \mathbb{Z}_N^2 \text{ and } j + k - \ell = 0 \right\}$$

span $T_q \mathcal{Q}$ at every generic point q in $\mathcal{Q}$.

---

[7] As in the proof of Proposition 5.2, the continuity equation is intended here in the weak sense.

Proof. Fix a generic point q in $\mathcal{Q}$. The main idea of the proof is to choose an enumeration of $\mathbb{Z}_N^2$ and a subset of vector fields from $\mathcal{V}$ so that the matrix made up of these vector fields evaluated at q is in a convenient form whose rank is readily deduced. Formally, the enumeration is the bijection $F : \mathbb{Z}_N^2 \to \{1, \ldots, 2N(N+1)\}$ given by

$$
F(j) := \begin{cases}
1 & j = (1, 0), \\
5 + N & j = (2, 0), \\
j_1 + N(2N+1) & j = (j_1, 0) \text{ with } j_1 > 2, \\
j_1 + 2 + N & j = (j_1, 1) \text{ with } j_1 < 3, \\
j_1 + 3 + N & j = (j_1, 1) \text{ with } j_1 \geq 3, \\
j_1 + 2 - N + (2N+1)j_2 & j = (j_1, j_2) \text{ with } j_2 > 1.
\end{cases}
$$

Figure 2 gives this enumeration in the case $N = 4$. Informally, F starts at $(1, 0)$, then counts lattice points from left to right along the horizontal line $y = 1$ until the point $(2, 1)$, which corresponds to $4 + N$. It then assigns $5 + N$ to $(2, 0)$ and continues counting along the line $y = 1$. From there it moves up to the lines $y = 2$, $y = 3$, and so on, counting from left to right along each. Finally, it goes back down to the line $y = 0$ and counts the remaining indices from left to right.

The motivation for F is that all horizontally-adjacent indices (j_1, j_2) and $(j_1 + 1, j_2)$ form an interacting triple together with $(1, 0)$. Fix for the moment an integer $y > 1$ and consider the yth horizontal line of $\mathbb{Z}_N^2$; that is, the points with second coordinate y. These are outlined by red blocks in Fig. 3. By the preceding remarks we can choose the vector fields corresponding to the horizontally-adjacent indices and concatenate them column-wise to get the block matrix

$$
B_y := \begin{pmatrix}
\widetilde{M}_y & * & * & * \\
0 & M''_{y,-N+2} & * & * \\
0 & 0 & \ddots & * \\
0 & 0 & 0 & M''_{y,N}
\end{pmatrix}.
$$

Here, slightly abusing notation, each $M''_{y,i}$ is the 2-by-4 matrix consisting of the bottom two rows of (6.33) for the indices $j = (1, 0)$, $k = (i - 1, y)$, $\ell = (j, y)$ and

$$
\widetilde{M}_y := \begin{pmatrix}
C_{j\ell}a_j a_\ell & 0 & C_{j\ell}b_j b_\ell & 0 & 0 & 0 \\
0 & C_{j\ell}a_j b_\ell & 0 & -C_{j\ell}b_j a_\ell & 0 & 0 \\
-C_{jk}a_j a_k & 0 & 0 & C_{jk}b_j b_k & -C_{j'k'}a_{j'} a_{k'} & 0 \\
0 & -C_{jk}a_j b_k & -C_{jk}b_j a_k & 0 & 0 & -C_{j'k'}a_{j'} b_{k'}
\end{pmatrix}
$$

where $j = (1, 0)$, $k = (-N, y)$, $\ell = (-N + 1, y)$ and $j' = (0, 1)$ and $k' = (-N + 1, y - 1)$. This is M' with two columns from the interacting triple $(0, 1)$, $(-N + 1, y - 1)$, $(-N + 1, y)$ adjoined to the end. Note that these adjoined columns contribute entries in the coordinates corresponding to $(0, 1)$ and $(-N + 1, y - 1)$, but these come before all indices in the yth row for our ordering. By adding the latter two columns, $\widetilde{M}_y$ has rank 4 at any generic point. Further, since each $M''_{y,j}$ has rank 2, each B_y has rank $4 + 2(2N - 1) = 4N + 2$. This establishes spanning of the red blocks in Fig. 3.

For the blue block we perform a similar procedure to the one above to get

$$
B_1 := \begin{pmatrix}
M_{123} & * & * & * & * & * \\
0 & M''_{1,-N+2} & * & * & * & * \\
\hline
0 & 0 & \ddots & * & * & * \\
0 & 0 & 0 & \widehat{M} & * & * \\
\hline
0 & 0 & 0 & 0 & \ddots & * \\
0 & 0 & 0 & 0 & 0 & M''_{1,N}
\end{pmatrix}
$$

where M_{123} is the matrix from (6.33) for the interacting triple $(1,0), (-N,1), (-N+1,1)$, each M'' is as before, and $\widehat{M}$ is the 6-by-8 matrix

$$
\widehat{M} := \begin{pmatrix}
& & & & & 0 & 0 & 0 & 0 \\
& M'_{1,N+3,N+4} & & & & 0 & 0 & 0 & 0 \\
& & & & & \text{---} & \text{---} & \text{---} & \text{---} \\
\text{---} & \text{---} & \text{---} & \text{---} & & & & & \\
0 & 0 & 0 & 0 & & & M'_{N+2,N+4,N+5} & & \\
0 & 0 & 0 & 0 & & & & &
\end{pmatrix}
$$

located at the rows corresponding to $N+3$, $N+4$, and $N+5$. The reason for $\widehat{M}$, and for considering the blue block separately, is that $C_{jk} = 0$ when $j = (1,0)$ and $k = (0,1)$. The matrix M has rank 6 at a generic point. Since M_{123} has rank 4, $\widehat{M}$ has rank 6, and each of the $2N-3$ remaining M'' blocks has rank 2, the matrix B_y has rank $4 + 6 + 2(2N-3) = 4N + 4$.

Finally, none of the indices of the green block interact with $(1,0)$ since the C_{jk} are all 0 in this case. However, by an entirely similar procedure to above, we can use the interactions between $(0,1), (x,0)$, and $(x,1)$ for $x > 1$ to get a rank $2(N-2)$ block matrix for the last $N-2$ coordinates of the form

$$
B_{N+1} := \begin{pmatrix}
\tilde{M}''_{0,2} & * & * & * \\
0 & \tilde{M}''_{0,3} & * & * \\
0 & 0 & \ddots & * \\
0 & 0 & 0 & \tilde{M}''_{0,N}
\end{pmatrix}
$$

where $\tilde{M}''_{0,x} = M''_{(0,1),(x,0),(x,1)}$ for $M''_{jk\ell}$ consisting of the two bottom rows of (6.33). Combining the above results we observe that there is an ordering of indices and vector fields such that the matrix whose columns consist of these vector fields has the form

$$
B := \begin{pmatrix}
B_1 & * & * & * \\
0 & B_2 & * & * \\
0 & 0 & \ddots & * \\
0 & 0 & 0 & B_{N+1}
\end{pmatrix}.
$$

Moreover, B has rank

$$
\text{rank}(B) = \text{rank}(B_1) + \text{rank}(B_{N+1}) + \sum_{y=2}^{N} \text{rank}(B_y) = 4N(N+1) - 2 = n - 2
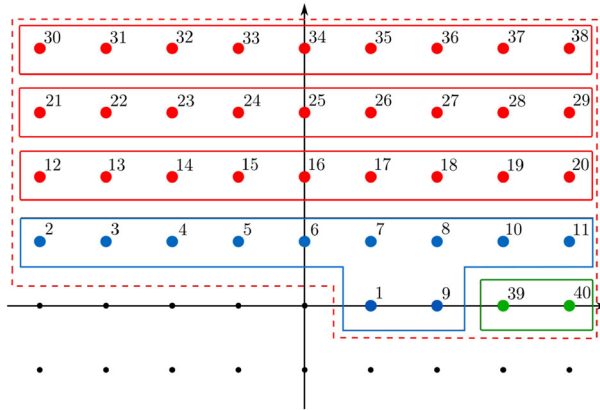$$

Fig. 3. Ordering of $\mathbb{Z}_N^2$ when $N = 4$

at every generic point in $\mathcal{Q}$. Now since the dynamics conserve energy and enstrophy, every tangent vector to $\mathcal{Q}$ is perpendicular to the normal vectors for these two quantities which are linearly independent at every generic point. Therefore the maximum dimension of $T_q\mathcal{Q}$ is $n - 2$, and by the above argument we have shown the vector fields $\mathcal{V}$ span $T_q\mathcal{Q}$ at q. □

7. Adding Forcing and Dissipation: Lorenz-96 and 2D Navier–Stokes

In this section we add dissipation and fixed body forcing to both conservative Lorenz-96 and Galerkin approximations of 2D Euler by introducing a new vector field

$$V_0(x) = -\nu\Lambda x + F \tag{7.1}$$

to the splittings constructed in Sects. 5 and 6, where $\nu > 0$ is an arbitrary constant, F a fixed nonzero vector with nonnegative entries, and Λ a linear operator satisfying

$$\Lambda x \cdot x \geq \alpha\|x\|^2 \tag{7.2}$$

for some $\alpha > 0$. For the remainder of this section we consider random splittings associated to families of complete, smooth vector fields $\mathcal{V} = \{V_k\}_{k=0}^n$ on $\mathbb{R}^d$ satisfying.

Assumption 2. V_0 is as in (7.1) and the flows of the other V_k conserve Euclidean norm.

Fix $h > 0$ and let P_h be the transition kernel of a random splitting satisfying Assumption 2. When Λ is the identity matrix, the addition of V_0 to the splitting of conservative Lorenz-96 gives a splitting of the full Lorenz-96 model, (1.4), while for 2D Euler the resulting V_0 corresponds to a friction or drag term sometimes called *Ekman damping*. When Λ is diagonal with diagonal entry $|k|^2$ in the spots associated to[8] a_k and b_k, which corresponds to a Laplacian written in Fourier space, the addition of V_0 to the splitting of 2D Euler gives a splitting of 2D Navier–Stokes, (1.1).

Note that the dissipative part of V_0 in (7.1) depends linearly on x whereas the forcing is constant. Thus dissipation dominates forcing for sufficiently large x and, since the

---

[8] Recall that for each index $k \in \mathbb{Z}_N^2$, we have two real coordinates a_k and b_k.

remaining vector fields are conservative, the splitting dynamics cannot grow too large. Specifically, letting $\Phi_{h\tau}$ be as in (2.2) but with the solution $\varphi^{(0)}$ of $\dot{x} = V_0(x)$ appended to the beginning of each cycle, we have

Lemma 7.1. *Under Assumption 2 for any initial x and $m > 0$,*

$$\|\Phi_{h\tau}^m(x)\|^2 \leq \|x\|^2 e^{-\nu\alpha h \sum_{k=0}^m \tau_{k(n+1)}} + \frac{1}{\nu^2\alpha^2}\|F\|^2 \left(1 - e^{-\nu\alpha h \sum_{k=0}^m \tau_{k(n+1)}}\right). \quad (7.3)$$

Proof. Letting $\varphi = \varphi^{(0)}$, we have

$$\partial_t \|\varphi_t\|^2 = 2\langle F, \varphi_t\rangle - 2\nu\langle \Lambda\varphi_t, \varphi_t\rangle$$

$$\leq \frac{1}{\nu\alpha}\|F\|^2 + \nu\alpha\|\varphi_t\|^2 - 2\nu\alpha\|\varphi_t\|^2 = \frac{1}{\nu\alpha}\|F\|^2 - \nu\alpha\|\varphi_t\|^2,$$

where the inequality follows from (7.2) and $2\langle F, \varphi_t\rangle \leq (\nu\alpha)^{-1}\|F\|^2 + \nu\alpha\|\varphi_t\|^2$. Solving

$$\dot{y} = \frac{1}{\nu\alpha}\|F\|^2 - \nu\alpha y$$

from $y(0) = \|x\|$ together with the comparison theorem for ODEs [42] then gives

$$\|\varphi_t(x)\|^2 \leq \|x\|^2 e^{-\nu\alpha t} + \frac{1}{\nu^2\alpha^2}\|F\|^2 \left(1 - e^{-\nu\alpha t}\right)$$

for all time. Furthermore, since $\varphi^{(k)}$ conserves norm for $1 \leq k \leq n$, the above implies

$$\|\Phi_{h\tau}(x)\|^2 = \|\varphi_{h\tau_n}^{(n)} \circ \cdots \circ \varphi_{h\tau_0}^{(0)}(x)\|^2$$

$$= \|\varphi_{h\tau_0}^{(0)}(x)\|^2 \leq \|x\|^2 e^{-\nu\alpha\tau_0} + \frac{1}{\nu^2\alpha^2}\|F\|^2 \left(1 - e^{-\nu\alpha\tau_0}\right).$$

The result then follows by straightforward induction on the number of cycles, namely m. □

Remark 7.2. The convergence results of Sect. 4 do not directly apply to Lorenz-96 and Galerkin approximations of 2D Navier–Stokes since $\mathcal{V}$-orbits are generally unbounded in both models. However, Lemma 7.1 implies that any splitting starting from x whose vector fields satisfy Assumption 2 will lie inside the ball of radius $\|x\|^2 + (\nu\alpha)^{-2}\|F\|^2$ centered at the origin for all nonnegative times. In particular, since the splitting vector fields are smooth, a bound analogous to (4.1) holds for all x in the ball $B_r(0)$ of radius r centered at the origin in the ambient Euclidean space. Thus all convergence results of Sect. 4 hold for these random splittings when $\mathcal{C}^k(\mathcal{X})$ is replaced by $\mathcal{C}_r^k(\mathcal{X})$, the space of k-times continuously differentiable functions that vanish outside $B_r(0)$. Intuitively, this says that for any initial condition x, the trajectories of a random splitting satisfying Assumption 2 will converge on average and almost surely as $h \to 0$ to the trajectory of the true dynamics starting from x.

Corollary 7.3. *The Euclidean norm is a Lyapunov function for P_h. That is, there exist constants $K \geq 0$ and $\gamma \in (0, 1)$ such that for all $x \in \mathbb{R}^d$,*

$$(P_h\|\cdot\|)(x) \leq \gamma\|x\| + K.$$

Proof. By Lemma 7.1, specifically $\|\Phi_{ht}(x)\| \leq \|x\|e^{-\frac{1}{2}\nu\alpha t_0} + (\nu\alpha)^{-1}\|F\|$, we have

$$(P_h\|\cdot\|)(x) = \int_{\mathbb{R}_+^{n+1}} \|\Phi_{ht}(x)\|e^{-\sum t_k} dt \leq \frac{1}{1 + \frac{1}{2}\nu\alpha h}\|x\| + \frac{1}{\nu\alpha}\|F\|$$

for any x. The result follows with $K = (\nu\alpha)^{-1}\|F\|$ and $\gamma = (1 + \frac{1}{2}\nu\alpha h)^{-1}$. □

7.1. Ergodicity. We now present a variation of Theorem 3.1, namely Theorem 7.4, which simplifies verification of ergodicity in the present setting. Recall from Sects. 5 and 6 that one of the difficulties in verifying Theorem 3.1 was proving controllability, i.e., the existence of a distinguished point x_* that could be reached by the splitting dynamics in finite time from any other point. With the addition of dissipation, the fixed point $\nu^{-1}\Lambda^{-1}F$ of $\dot{x} = V_0(x)$ is a natural candidate for x_* and, as we will see, the fact that it is globally attracting obviates several technicalities associated with controllability in the conservative cases discussed above.

Theorem 7.4. *Suppose Assumption 2 holds and set $x_* = \nu^{-1}\Lambda^{-1}F$. If there exist $m \geq 0$ and t in $\mathbb{R}_+^{mn}$ such that the Lie bracket condition holds at $\widetilde{x} := \Phi_{ht}^m(x_*)$, then P_h has a unique invariant measure μ for all $h > 0$. Furthermore, there exist $C > 0$ and γ in $(0, 1)$ such that for all x in $\mathbb{R}^d$,*

$$\|P_h^m(x, \cdot) - \mu\| \leq C\gamma^m \tag{7.4}$$

where $\|\cdot\|$ is the norm on probability measures induced by the weighted supremum norm $\|f\| := \sup_x |f(x)|/(1 + \|x\|)$ on bounded measurable functions $f : \mathbb{R}^d \to \mathbb{R}$.

The proof of Theorem 7.4 uses the following lemmas. The first, due to Krylov-Bogolubov, is a standard result from the theory of Markov processes [23]. The second, which follows from Lemma 3.2 and Theorem 3.6, is from [6, Theorem 4.4]. For the statement of Lemma 7.5, recall a transition kernel P on $\mathbb{R}^d$ is *Feller* if Pf is continuous whenever $f : \mathbb{R}^d \to \mathbb{R}$ is continuous and bounded. Also, a sequence of probability measures $\{\mu_m\}$ on $\mathbb{R}^d$ is *tight* if for every $\varepsilon > 0$ there exists a compact subset K of $\mathbb{R}^d$ such that $\mu_m(K) \geq 1 - \varepsilon$ for all m.

Lemma 7.5. *Let P be a Feller probability transition kernel on $\mathbb{R}^d$. If there exists x in $\mathbb{R}^d$ such that $\{P^m(x, \cdot)\}_{m=0}^\infty$ is tight, then P has an invariant probability measure.*

Lemma 7.6. *Suppose $\Phi_{ht}^m(x) = \widetilde{x}$ and the Lie bracket condition holds at $\widetilde{x}$. Then there exists a $c > 0$, an $\widetilde{m}$, and neighborhoods U_x of x and $\widetilde{U}$ of $\widetilde{x}$ such that for all y in U_x and B in $\mathcal{B}(\mathcal{X})$,*

$$P_h^{\widetilde{m}}(y, B) \geq c\lambda\left(B \cap \widetilde{U}\right).$$

The following proof is another instance of the rather classical idea, dating at least back to the split chains of Nummelin [49] and work of Meyn and Tweedie [43], that the existence of a globally accessible point at which the dynamics is continuous in the right sense implies the transition densities converge to a unique equilibrium measure. If the return to the globally accessible point has finite expectation, then mixing is exponential. The same basic structure of the SDE version of our system was leveraged in [18] to prove exponential mixing (see also [41]). In the closely related PDMP setting, analogous results are found in [38] in a specific example and [7] in a more general context.

Proof of Theorem 7.4. We first prove existence. Continuity of Φ_{ht} immediately implies P_h is Feller. Furthermore, Lemma 7.1 implies that random splitting starting from any x is constrained to lie in a compact subset of $\mathbb{R}^d$, namely the closed ball of radius $\|x\|^2 + (\nu\alpha)^{-2}\|F\|^2$ centered at the origin. Thus, for any x, the sequence $\{P_h^m(x, \cdot)\}_{m=0}^\infty$ is tight and existence follows from Lemma 7.5.

Next we prove uniqueness. The hypothesis and Lemma 7.6 together imply the existence of $c > 0$, $\widetilde{m}$, and neighborhoods U_* of x_* and $\widetilde{U}$ of $\widetilde{x}$ such that

$$P_h^{\widetilde{m}}(x, B) \geq c\lambda\left(B \cap \widetilde{U}\right) \tag{7.5}$$

for all $x \in U_*$ and Borel sets B. Also, positive-definiteness of Λ implies

$$\|\varphi_t^{(0)}(x) - x_*\| \le e^{-\alpha t} \|x - x_*\|$$

for any $x \in \mathbb{R}^d$ and $t \ge 0$. In particular, for any open ball B_r of radius r centered at the origin, there exists $T_0 > 0$ such that $\varphi_{ht}^{(0)}(B_r)$ is properly contained in U_* whenever $ht > T_0$. And since $\varphi_{ht}^{(0)}(B_r)$ is properly contained in U_* and the $\varphi^{(k)}$ are continuous, there exist $T_k > 0$ such that $\Phi_{ht} = \varphi_{ht_n}^{(n)} \circ \cdots \circ \varphi_{ht_0}^{(0)}(x) \in U_*$ for all $x \in B_r$ and $ht_k \in (0, T_k)$. So, for any $x \in B_r$,

$$P_h(x, U_*) \ge \int_0^{T_n} \cdots \int_0^{T_1} \int_{T_0}^{\infty} \mathbb{1}_{U_*}(\Phi_{ht}(x)) e^{-\sum t_k} dt = \frac{1}{T_0} \prod_{k=1}^{n} \left(1 - e^{-T_k}\right) > 0$$

and hence $\inf_{x \in B_r} P_h(x, U_*) > 0$.

As in the proof of Theorem 3.1, suppose toward a contradiction that μ_1 and μ_2 are distinct P_h-ergodic probability measures and that A_1 and A_2 are disjoint measurable sets partitioning $\mathbb{R}^d$ with $\mu_i(B) = \mu_i(B \cap A_i)$ for all Borel sets B. Fix x_i in the support of μ_i, let r be sufficiently large that $x_1, x_2 \in B_r$, and set $\kappa := \inf_{x \in B_r} P_h(x, U_*) > 0$. Then by (7.5) for any Borel set B,

$$\mu_i(B) = \mu_i P_h^{\widetilde{m}+1}(B) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} P_h^{\widetilde{m}}(y, B) P_h(x, dy) \mu_i(dx)$$

$$\ge \int_{B_r} \int_{U_*} P_h^{\widetilde{m}}(y, B) P_h(x, dy) \mu_i(dx) \ge \kappa c \lambda \left(B \cap \widetilde{U}\right) \mu_i(B_r). \quad (7.6)$$

In particular, $\mu_i(B) = 0$ implies $\lambda(B \cap \widetilde{U}) = 0$ since c, κ, and $\mu_i(B_r)$ are all strictly positive (the latter because B_r is an open set containing both x_1 and x_2 which were chosen to be in the supports of μ_1 and μ_2, respectively). But $\mu_1(A_2 \cap \widetilde{U}) = \mu_2(A_1 \cap \widetilde{U}) = 0$ and so we obtain the contradiction

$$0 < \lambda\left(\widetilde{U}\right) = \lambda\left(A_1 \cap \widetilde{U}\right) + \lambda\left(A_2 \cap \widetilde{U}\right) = 0,$$

which concludes the proof of uniqueness.

Finally, for the exponential convergence statement (7.4), we have from (7.6) that for any $r > 0$,

$$\inf_{x \in B_r} P_h^{\widetilde{m}+1}(x, B) \ge \kappa c \lambda \left(B \cap \widetilde{U}\right)$$

for all Borel sets B. That is, the transition probabilities $P_h^{\widetilde{m}+1}(x, \cdot)$ are minorized uniformly over B_r by the probability measure $\overline{\lambda} := \lambda(\widetilde{U})^{-1} \lambda(\cdot \cap \widetilde{U})$. Exponential convergence then follows from Corollary 7.3 upon taking $r > 2K/(1 - \gamma)$. See for example Theorem 1.2 in [25]. $\qquad\square$

Corollary 7.7. *Consider the random splitting of Lorenz-96 associated to the vector fields $\{V_k\}_{k=0}^n$, where $V_0(x) = -\nu x + F$ and $\{V_k\}_{k=1}^n$ are the splitting vector fields of conservative Lorenz-96 from Sect. 5. If $x_* := -\nu F$ is not a fixed point of conservative Lorenz-96, i.e., $\nu^2 \sum_{k=1}^n (F_k^2 + F_{k+1}^2) F_{k-1}^2 \ne 0$, then the random splitting has a unique, and hence ergodic, invariant measure on $\mathbb{R}^n$ and the dynamics converge to this measure at an exponential rate in the sense of (7.4).*

Proof. The determinant of the n-by-n matrix

$$
\begin{pmatrix} | & | & & | \\ V_0(x) & V_1(x) & \cdots & V_{n-1}(x) \\ | & | & & | \end{pmatrix} = \begin{pmatrix} -\nu x_1 + F_1 & x_2 x_n & 0 & \cdots & 0 \\ -\nu x_2 + F_2 & -x_1 x_n & x_3 x_1 & \cdots & 0 \\ -\nu x_3 + F_3 & 0 & -x_2 x_1 & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & x_n x_{n-2} \\ -\nu x_n + F_n & 0 & 0 & & -x_{n-1} x_{n-2} \end{pmatrix}
$$

is

$$
x_1 x_{n-1} x_n \left(\prod_{x=2}^{n-2} x_k^2 \right) \left(\nu^2 \|x\|^2 - \langle F, x \rangle \right).
$$

So the $\{V_k\}_{k=0}^n$ span $\mathbb{R}^n$ at every x with nonzero coordinates and satisfying $\nu^2 \|x\|^2 \neq \langle F, x \rangle$. In particular, since x_* is not a fixed point of conservative Lorenz-96, we showed in the proof of Proposition 5.2 that x_* can be moved via the splitting dynamics to some $\widetilde{x}$ with nonzero coordinates. Finally, by rotating slightly more on the last step if necessary, we can also guarantee $\nu^2 \|\widetilde{x}\|^2 \neq \langle F, \widetilde{x} \rangle$. Thus the Lie bracket condition holds at $\widetilde{x}$ and the result follows by Theorem 7.4. $\square$

Corollary 7.8. *Fix $N \geq 2$ and set $n = 4N(N + 1)$. Consider the random splitting of the Nth Galerkin approximation of 2D Navier–Stokes associated to $\{V_k\}_{k=0}^n$, where $V_0(x) = -\nu \Lambda x + F$ with Λ the n-by-n diagonal matrix corresponding to the Laplacian discussed at the beginning of this section, and $\{V_k\}_{k=1}^n$ the splitting vector fields of 2D Euler from Sect. 6. If F is nondegenerate in the sense of Definition 6.3, then the random splitting has a unique, and hence ergodic, invariant measure and the dynamics converge to this measure at an exponential rate in the sense of (7.4).*

Proof. Recall in this case $V_0(x) = -\nu \Lambda x + F$ where Λ is the diagonal matrix with diagonal entry $|k|^2$ in the slots corresponding to the coordinates a_k and b_k. Fix $j, k, \ell \in \mathbb{Z}_N^2$ with $j + k - \ell = 0$ and let W be one of the vector fields $V_{a_j a_k a_\ell}, V_{a_j b_k b_\ell}, V_{b_j a_k b_\ell}$, or $V_{b_j b_k a_\ell}$. Letting e.g. $(x_j, x_k, x_\ell) = (a_j, a_k, a_\ell)$ when $W = V_{a_j a_k a_\ell}$ and similarly for the other cases, direct computation yields

$$
[V_0, W]_j(x) = C_{k\ell} \left(F_k x_\ell + F_\ell x_k + \nu(|j|^2 - |k|^2 - |\ell|^2) x_k x_\ell \right),
$$

$$
[V_0, W]_k(x) = C_{j\ell} \left(F_j x_\ell + F_\ell x_j + \nu(|k|^2 - |j|^2 - |\ell|^2) x_j x_\ell \right),
$$

$$
[V_0, W]_\ell(x) = -C_{jk} \left(F_j x_k + F_k x_j + \nu(|\ell|^2 - |j|^2 - |k|^2) x_j x_k \right), \qquad (7.7)
$$

where $[V_0, W]_j(x)$ is the component of $[V_0, W]$ corresponding to the component x_j of x, and similarly for $[V_0, W]_k$ and $[V_0, W]_\ell$. As in the 2D Euler case, Gaussian elimination shows that the 6-by-6 matrix (see (6.33) for an explicit form of the middle 4 columns)

$$
\begin{pmatrix} | & | & | & | & | & | \\ V_0 & V_{a_j a_k a_\ell} & V_{a_j b_k b_\ell} & V_{b_j a_k b_\ell} & V_{b_j b_k a_\ell} & [V_0, W] \\ | & | & | & | & | & | \end{pmatrix} \qquad (7.8)
$$

is rank 6 at every generic[9] point q in $\mathbb{R}^n$. Thus V_0 and $[V_0, W]$ add two new directions to the splitting vector fields of 2D Euler and by an entirely similar argument to the spanning

---

[9] Recall a *generic point* is one with all coordinates nonzero; see Definition 6.4.

argument in Sect. 6.3.3 we have that the Lie bracket condition holds at every such q. Furthermore, since F is nondegenerate the controllability argument of Sect. 6.3.1 implies x_* can be evolved via the split dynamics to a generic point. The result then follows by Theorem 7.4. $\square$

Remark 7.9. A very similar argument to the one above proves unique ergodicity for Ekman damping as well, i.e., when Λ is the identity matrix on $\mathbb{R}^n$. In this case (7.1) becomes

$$[V_0, W]_j(x) = C_{k\ell}\left(F_k x_\ell + F_\ell x_k - \nu x_k x_\ell\right),$$
$$[V_0, W]_k(x) = C_{j\ell}\left(F_j x_\ell + F_\ell x_j - \nu x_j x_\ell\right),$$
$$[V_0, W]_\ell(x) = -C_{jk}\left(F_j x_k + F_k x_j - \nu x_j x_k\right),$$

and the rest of the argument goes through unchanged.

Declarations

Appendix A. Convergence Lemmas

A.1. Semigroups, norms, and bounds. In this subsection we elaborate on the semigroup framework of Sect. 4. The notation and results are used extensively in the proofs of Lemmas 4.2 and 4.6, which are given in Sects. A.2 and A.4, respectively.

Fix a $\mathcal{V}$-orbit $\mathcal{X}$. The $\mathcal{C}^2$ assumption implies the V_k, which act on functions f via $V_k f(x) = Df(x)V_k(x)$, are linear operators from $\mathcal{C}^2(\mathcal{X})$ to $\mathcal{C}^1(\mathcal{X})$ and from $\mathcal{C}^1(\mathcal{X})$ to $\mathcal{C}(\mathcal{X})$. It also implies the semigroups $\{S_t\}_{t\geq 0}$ and $\{\widetilde{S}_t^{(k)}\}_{t\geq 0}$ defined in (4.2) and (4.3) are linear operators on $\mathcal{C}^k(\mathcal{X})$ for $k \leq 2$. Our aim now is to obtain bounds on norms of compositions of these random semigroups. For $i \leq j$ define $\Phi_{h\tau}^{(i,j)} := \varphi_{h\tau_j}^{(j)} \circ \cdots \circ \varphi_{h\tau_i}^{(i)}$ and $\widetilde{S}_{h\tau}^{(i,j)} := \widetilde{S}_{h\tau}^{(i)} \cdots \widetilde{S}_{h\tau}^{(j)}$. Note $\widetilde{S}_{h\tau}^{(i,j)}$ acts on functions f via

$$\widetilde{S}_{h\tau}^{(i,j)} f(x) = f\left(\Phi_{h\tau}^{(i,j)}(x)\right) = f\left(\varphi_{h\tau_j}^{(j)} \circ \cdots \circ \varphi_{h\tau_i}^{(i)}(x)\right).$$

So for any $f \in \mathcal{C}(\mathcal{X})$ with $\|f\|_\infty = 1$, we have

$$\|\widetilde{S}_{h\tau}^{(i,j)} f\|_\infty = \|f(\Phi_{h\tau}^{(i,j)})\|_\infty = 1$$

and hence $\|\widetilde{S}_{h\tau}^m\|_{0\to 0} = 1$. Next, let $\varphi = \varphi^{(k)}$ for arbitrary k. Then

$$\varphi_t(x) = x + \int_0^t V(\varphi_s(x))ds$$

and so

$$D\varphi_t(x) = I + \int_0^t DV(\varphi_s(x))D\varphi_s(x)ds$$

and

$$D^2\varphi_t(x) = \int_0^t D^2V(\varphi_s(x))(D\varphi_s(x), D\varphi_s(x)) + DV(\varphi_s(x))D^2\varphi_s(x)ds.$$

In particular, $\|D\varphi_t(x)\| \leq 1 + C_* \int_0^t \|D\varphi_s(x)\|ds$ for all x in $\mathcal{X}$ and Grönwall's inequality implies

$$\sup_{x\in\mathcal{X}} \|D\varphi_t(x)\| \leq e^{C_*t}, \tag{A.1}$$

where here and throughout C_* is the constant from (4.1) corresponding to $\mathcal{X}$. Similarly, since $\|D^2V(D\varphi, D\varphi)\| \leq \|D^2V\|\|D\varphi\|^2 \leq C_*\|D\varphi\|^2$,

$$\|D^2\varphi_t(x)\| \leq C_* \int_0^t \|D\varphi_s(x)\|^2 + \|D^2\varphi_s(x)\|ds \leq C_*te^{2C_*t} + C_* \int_0^t \|D^2\varphi_s(x)\|ds$$

and Grönwall implies

$$\sup_{x\in\mathcal{X}} \|D^2\varphi_t(x)\| \leq C_*te^{3C_*t}. \tag{A.2}$$

Note (A.1) and (A.2) hold uniformly over all $\varphi^{(k)}$. Thus, for $f \in C^1(\mathcal{X})$ with $\|f\|_1 = 1$,

$$\left\| D\left(\widetilde{S}_{h\tau}^{(i,j)}f\right) \right\| = \left\| Df\left(\Phi_{h\tau}^{(i,j)}\right)D\Phi_{h\tau}^{(i,j)} \right\| \leq \prod_{k=i}^j \|D\varphi_{h\tau_k}^{(k)}\| \leq e^{C_*h\sum_{k=i}^j \tau_k},$$

where the first inequality follows from submultiplicity and the second from (A.1). Similarly,

$$D^2\Phi_{h\tau}^{(i,j)} = \sum_{k=i}^j D\varphi_{h\tau_j}^{(j)} \cdots D\varphi_{h\tau_{k+1}}^{(k+1)} D^2\varphi_{h\tau_k}^{(k)} \left(D\Phi_{h\tau}^{(i,k-1)}, D\Phi_{h\tau}^{(i,k-1)}\right)$$

together with (A.1) and (A.2) gives

$$\left\| D^2\Phi_{h\tau}^{(i,j)} \right\| \leq \sum_{k=i}^j \left\| D\varphi_{h\tau_j}^{(j)} \right\| \cdots \left\| D\varphi_{h\tau_{k+1}}^{(k+1)} \right\| \left\| D^2\varphi^{(k)} \right\| \left\| D\Phi_{h\tau}^{(i,k-1)} \right\|^2$$

$$\leq C_* \sum_{k=i}^j h\tau_k e^{C_*h\sum_{k+1}^j \tau_\ell} e^{3C_*h\tau_k} e^{2C_*h\sum_1^{k-1}\tau_\ell} \leq C_*he^{3C_*h\sum_{k=i}^j \tau_k} \sum_{k=i}^j \tau_k.$$

Therefore

$$
\left\| D^2 \left(\widetilde{S}_{h\tau}^{(i,j)} f \right) \right\| = \left\| D^2 f \left(\Phi_{h\tau}^{(i,j)} \right) \left(D\Phi_{h\tau}^{(i,j)}, D\Phi_{h\tau}^{(i,j)} \right) + Df \left(\Phi_{h\tau}^{(i,j)} \right) D^2 \Phi_{h\tau}^{(i,j)} \right\|
$$

$$
\leq \left\| D\Phi_{h\tau}^{(i,j)} \right\|^2 + \left\| D^2 \Phi_{h\tau}^{(i,j)} \right\| \leq e^{2C_* h \sum_{k=i}^{j} \tau_k} + \left\| D^2 \Phi_{h\tau}^{(i,j)} \right\|
$$

$$
\leq e^{2C_* h \sum_{k=i}^{j} \tau_k} + C_* h e^{3C_* h \sum_{k=i}^{j} \tau_k} \sum_{k=i}^{j} \tau_k
$$

$$
\leq \left(1 + C_* h \sum_{k=i}^{j} \tau_k \right) e^{3C_* h \sum_{k=i}^{j} \tau_k}.
$$

The above computations prove

Lemma A.1. *For any $h > 0$ and $i \leq j$, we have $\|\widetilde{S}_{h\tau}^{(i,j)}\|_{0 \to 0} = 1$ as well as*

$$
\left\| \widetilde{S}_{h\tau}^{(i,j)} \right\|_{1 \to 1} \leq e^{C_* h \sum_{k=i}^{j} \tau_k} \quad and \quad \left\| \widetilde{S}_{h\tau}^{(i,j)} \right\|_{2 \to 2} \leq \left(1 + C_* h \sum_{k=i}^{j} \tau_k \right) e^{3C_* h \sum_{k=i}^{j} \tau_k}.
$$

In particular, $\left\| \widetilde{S}_{h\tau}^{(i,j)} \right\|_{\ell \to \ell} \leq \left(1 + C_* h \sum_{k=i}^{j} \tau_k \right) e^{3C_* h \sum_{k=i}^{j} \tau_k}$ for all $\ell \leq 2$.

Note that under the $\mathcal{C}^2$ assumption $\widetilde{S}_{h\tau}^{(i,j)}$ can also be regarded as a linear operator from $\mathcal{C}^2(\mathcal{X})$ to $\mathcal{C}^1(\mathcal{X})$. So since $\{f \in \mathcal{C}^2(\mathcal{X}) : \|f\|_2 = 1\}$ is a subset of $\{f \in \mathcal{C}^1(\mathcal{X}) : \|f\|_1 = 1\}$, we have

$$
\left\| \widetilde{S}_{h\tau}^{(i,j)} \right\|_{2 \to 1} = \sup_{\|f\|_2 = 1} \left\| \widetilde{S}_{h\tau}^{(i,j)} f \right\|_1 \leq \sup_{\|f\|_1 = 1} \left\| \widetilde{S}_{h\tau}^{(i,j)} f \right\|_1 = \left\| \widetilde{S}_{h\tau}^{(i,j)} \right\|_{1 \to 1} \leq e^{C_* h \sum_{k=i}^{j} \tau_k}.
$$

$$(A.3)$$

We also have the following corollary of Lemma A.1.

Corollary A.2. *Fix $i \leq j$ and set $m := j - i + 1$. For all $\ell \leq 2$ and polynomial $p : \mathbb{R}_+^m \to \mathbb{R}$ there exists $h_* > 0$ such that for all $h < h_*$,*

$$
\mathbb{E} \| p(\tau_i, \ldots, \tau_j) \widetilde{S}_{h\tau}^{(i,j)} \|_{k \to k} < \infty.
$$

$$(A.4)$$

Proof. Writing $t = (t_i, \ldots, t_j)$ and $dt = dt_i \cdots dt_j$, we have

$$
\mathbb{E} \| p(\tau_i, \ldots, \tau_j) \widetilde{S}_{h\tau}^{(i,j)} \|_{\ell \to \ell} = \int_{\mathbb{R}_+^m} |p(t)| \left\| \widetilde{S}_{ht}^{(i,j)} \right\|_{\ell \to \ell} e^{-\sum t_k} dt
$$

$$
\leq \int_{\mathbb{R}_+^m} |p(t)| \left(1 + C_* h \sum_{k=i}^{j} t_k \right) e^{(3C_* h - 1) \sum_{k=i}^{j} t_k} dt
$$

which is finite for all $h < h_* := (3C_*)^{-1}$. $\qquad \square$

A.2. Proof of Lemma 4.2. We highlight the steps of the proof with italicized font.
Variation of constants. We begin by differentiating $\widetilde{S}_{h\tau}$ in h:

$$\partial_h \widetilde{S}_{h\tau} = \sum_{k=1}^{n} \tau_k e^{h\tau_1} \cdots e^{h\tau_{k-1}} V_k e^{h\tau_k} \cdots e^{h\tau_n} = \sum_{k=1}^{n} \tau_k \widetilde{S}_{h\tau}^{(1,k-1)} V_k \widetilde{S}_{h\tau}^{(k,n)}.$$

Next, commute $\widetilde{S}_{h\tau}^{(1,k-1)}$ and V_k via $[\widetilde{S}_{h\tau}^{(1,k-1)}, V_k] := \widetilde{S}_{h\tau}^{(1,k-1)} V_k - V_k \widetilde{S}_{h\tau}^{(1,k-1)}$ to get

$$\partial_h \widetilde{S}_{h\tau} = \sum_{k=1}^{n} \tau_k V_k \widetilde{S}_{h\tau} + \sum_{k=1}^{n} \tau_k [\widetilde{S}_{h\tau}^{(1,k-1)}, V_k] \widetilde{S}_{h\tau}^{(k,n)} = V \widetilde{S}_{h\tau} + (V_\tau - V) \widetilde{S}_{h\tau} + E_{h\tau}$$

where $V_\tau := \sum_{k=1}^{n} \tau_k V_k$ and $E_{h\tau} := \sum_{k=1}^{n} \tau_k [\widetilde{S}_{h\tau}^{(1,k-1)}, V_k] \widetilde{S}_{h\tau}^{(k,n)}$. So, by variation of constants,

$$\widetilde{S}_{h\tau} - S_h = \int_0^h S_{h-r}(V_\tau - V) \widetilde{S}_{r\tau} dr + \int_0^h S_{h-r} E_{r\tau} dr. \tag{A.5}$$

Call $S_{h-r}(V_\tau - V)\widetilde{S}_{r\tau}$ *error term 1* and $S_{h-r} E_{r\tau}$ *error term 2*. These terms will be treated
separately in what follows. First however, we invoke variation of constants again to get
an expression for $[\widetilde{S}_{r\tau}^{(1,k-1)}, V_k]$ that will be used to control error term 2. Differentiating
in r gives

$$\partial_r [\widetilde{S}_{r\tau}^{(1,k-1)}, V_k] = \sum_{j=1}^{k-1} \tau_j [\widetilde{S}_{r\tau}^{(1,j-1)} V_j \widetilde{S}_{r\tau}^{(j,k-1)}, V_k]$$

$$= \sum_{j=1}^{k-1} \tau_j \left([V_j \widetilde{S}_{r\tau}^{(1,k-1)}, V_k] + [[\widetilde{S}_{r\tau}^{(1,j-1)}, V_j] \widetilde{S}_{r\tau}^{(j,k-1)}, V_k] \right)$$

$$= \sum_{j=1}^{k-1} \tau_j V_j [\widetilde{S}_{r\tau}^{(1,k-1)}, V_k]$$

$$+ \sum_{j=1}^{k-1} \tau_j \left([V_j, V_k] \widetilde{S}_{r\tau}^{(1,k-1)} + [[\widetilde{S}_{r\tau}^{(1,j-1)}, V_j] \widetilde{S}_{r\tau}^{(j,k-1)}, V_k] \right).$$

The second equality follows from commuting $\widetilde{S}_{h\tau}^{(1,j-1)}$ and V_j as before, and the third
follows from the identity $[XY, Z] = X[Y, Z] + [X, Z]Y$. So, by variation of constants,

$$[\widetilde{S}_{r\tau}^{(1,k-1)}, V_k] = \sum_{j=1}^{k-1} \int_0^r \tau_j e^{(r-s) \sum_{j=1}^{k-1} \tau_j V_j} [V_j, V_k] \widetilde{S}_{s\tau}^{(1,k-1)} ds$$

$$+ \sum_{j=1}^{k-1} \int_0^r \tau_j e^{(r-s) \sum_{j=1}^{k-1} \tau_j V_j} [[\widetilde{S}_{s\tau}^{(1,j-1)}, V_j] \widetilde{S}_{s\tau}^{(j,k-1)}, V_k] ds. \tag{A.6}$$

Note $\|e^{(r-s) \sum_{j=1}^{k-1} \tau_j V_j}\|_{0 \to 0} = 1$. So, by Corollary A.2 the integrands above satisfy

$$\mathbb{E} \|\tau_j e^{(r-s) \sum_{j=1}^{k-1} \tau_j V_j} [V_j, V_k] \widetilde{S}_{s\tau}^{(1,k-1)}\|_{2 \to 0}$$

$$\leq \|[V_j, V_k]\|_{2\to 0}\mathbb{E}\|\tau_j \widetilde{S}_{s\tau}^{(1,k-1)}\|_{2\to 2} < C$$

and

$$\mathbb{E}\big\|\tau_j e^{(r-s)\sum_{j=1}^{k-1}\tau_j V_j}\big[[\widetilde{S}_{s\tau}^{(1,j-1)}, V_j]\widetilde{S}_{s\tau}^{(j,k-1)}, V_k]\big\|_{2\to 0}$$
$$\leq \mathbb{E}\big\|\tau_j[[\widetilde{S}_{s\tau}^{(1,j-1)}, V_j]\widetilde{S}_{s\tau}^{(j,k-1)}, V_k]\big\|_{2\to 0} < C$$

for some C. Therefore

$$\mathbb{E}\|[\widetilde{S}_{r\tau}^{(1,k-1)}, V_k]\|_{2\to 0} \leq 2\sum_{j=1}^{k-1}\int_0^r C\,ds \leq Cr \tag{A.7}$$

for some new constant C (we will often absorb arbitrary constants into existing ones).

Error term 1. Rewrite error term 1 as

$$S_{h-r}(V_\tau - V)\widetilde{S}_{r\tau} = \sum_{k=1}^n (\tau_k - 1)S_{h-r}V_k\widetilde{S}_{r\tau}$$

$$= \sum_{k=1}^n (\tau_k - 1)S_{h-r}V_k\widetilde{S}_{r\tau}^{(1,k-1)}\widetilde{S}_{r\tau}^{(k+1,n)}$$

$$+ \sum_{k=1}^n (\tau_k - 1)S_{h-r}V_k\widetilde{S}_{r\tau}^{(1,k-1)}(e^{r\tau_k V_k} - I)\widetilde{S}_{r\tau}^{(k+1,n)}$$

$$=: \mathcal{A}_1 + \mathcal{A}_2 \tag{A.8}$$

where $\mathcal{A}_1$ and $\mathcal{A}_2$ are the first and second sums in the preceding expression. The second equality is obtained by adding and subtracting the identity I as follows:

$$\widetilde{S}_{r\tau} = \widetilde{S}_{r\tau}^{(1,k-1)}(e^{r\tau_k V_k} - I + I)\widetilde{S}_{r\tau}^{(k+1,n)}$$
$$= \widetilde{S}_{r\tau}^{(1,k-1)}\widetilde{S}_{r\tau}^{(k+1,n)} + \widetilde{S}_{r\tau}^{(1,k-1)}(e^{r\tau_k V_k} - I)\widetilde{S}_{r\tau}^{(k+1,n)}.$$

Notice $\widetilde{S}_{r\tau}^{(1,k-1)}\widetilde{S}_{r\tau}^{(k+1,n)}$ does not depend on τ_k. So, since the τ_i are independent with mean 1,

$$\mathbb{E}(\mathcal{A}_1) = \sum_{k=1}^n S_{t-r}V_k\mathbb{E}(\tau_k - 1)\mathbb{E}\big(\widetilde{S}_{r\tau}^{(1,k-1)}\widetilde{S}_{r\tau}^{(k+1,n)}\big) = 0. \tag{A.9}$$

For the second sum, Taylor expanding $r \mapsto e^{r\tau_k V_k}$ about $r = 0$ with remainder gives

$$e^{r\tau_k V_k} - I = r\tau_k V_k e^{r_* \tau_k V_k}$$

for some $r_* \in [0, r]$. Therefore

$$\mathcal{A}_2 = r\sum_{k=1}^n \tau_k(\tau_k - 1)S_{h-r}V_k\widetilde{S}_{r\tau}^{(1,k-1)}V_k e^{r_* \tau_k V_k}\widetilde{S}_{r\tau}^{(k+1,n)}$$

and by Lemma A.1 and Corollary A.2,

$$\|\mathbb{E}(\mathcal{A}_2)\|_{2\to 0} \le Cr \sum_{k=1}^{n} \mathbb{E}\|\widetilde{S}_{r\tau}^{(1,k-1)}\|_{1\to 1}\mathbb{E}\|\tau_k(\tau_k-1)\widetilde{S}_{r\tau}^{(k,n)}\|_{2\to 2} \le Cr \qquad (\text{A}.10)$$

for some $C > 0$. Combining Equations (A.8), (A.9), and (A.10) gives

$$\|\mathbb{E}(S_{h-r}(V_\tau - V)\widetilde{S}_{r\tau})\|_{2\to 0} \le Cr. \qquad (\text{A}.11)$$

Error term 2. Recall error term 2 is $S_{h-r}E_{r\tau} := \sum_{k=1}^{n}\tau_k S_{h-r}[\widetilde{S}_{r\tau}^{(1,k-1)}, V_k]\widetilde{S}_{r\tau}^{(k,n)}$. So, we have that

$$\|S_{h-r}E_{r\tau}\|_{2\to 0} \le \sum_{k=1}^{n}\tau_k\|S_{h-r}\|_{0\to 0}\|[\widetilde{S}_{r\tau}^{(1,k-1)}, V_k]\|_{2\to 0}\|\tau_k\widetilde{S}_{r\tau}^{(k,n)}\|_{2\to 2}.$$

Note $[\widetilde{S}_{r\tau}^{(1,k-1)}, V_k]$ is independent of τ_k. So, by (A.7) Corollary A.2,

$$\|\mathbb{E}(S_{h-r}E_{r\tau})\|_{2\to 0} \le Cr \qquad (\text{A}.12)$$

for some $C > 0$.

Final step. Combining (A.5), (A.11), and (A.12) and absorbing constants into C, we have

$$\begin{aligned}
\|P_h - S_h\|_{2\to 0} &= \|\mathbb{E}(\widetilde{S}_{h\tau} - S_h)\|_{2\to 0} \\
&\le \int_0^h \|\mathbb{E}(S_{h-r}(V_\tau - V)\widetilde{S}_{r\tau})\|_{2\to 0}dr + \int_0^h \|\mathbb{E}(S_{h-r}E_{r\tau})\|_{2\to 0}dr \\
&\le C\int_0^h r\,dr = \tfrac{1}{2}Ch^2.
\end{aligned}$$

$\square$

A.3. Concentration of the sum of exponential random variables. The proof of Lemma 4.6 will itself use two lemmas.

Lemma A.3. *Let $\{\tau_k\}_{k=1}^{\infty}$ be iid exponential with mean 1. For any $m \in \mathbb{N}$, $K > 0$ and $\beta > 1$,*

$$\mathbb{P}\left(\sum_{k=1}^{m}\tau_k > Km^\beta\right) \le 2^m e^{-\frac{1}{2}Km^\beta}. \qquad (\text{A}.13)$$

Proof. Note if $\tau \sim \text{Exp}(1)$ then $\mathbb{E}(e^{\tau/2}) = 2$. So, by Markov's inequality and independence,

$$\begin{aligned}
\mathbb{P}\left(\sum_{k=1}^{m}\tau_k > Km^\beta\right) &= \mathbb{P}\left(e^{\frac{1}{2}\sum_{k=1}^{m}\tau_k} > e^{\frac{1}{2}Km^\beta}\right) \\
&\le e^{-\frac{1}{2}Km^\beta}\left(\mathbb{E}\left[e^{\frac{1}{2}\tau}\right]\right)^m = 2^m e^{-\frac{1}{2}Km^\beta}.
\end{aligned}$$

$\square$

Lemma A.4. *Let* $\{\tau_k\}_{k=1}^{\infty}$ *be iid exponential with mean* 1. *For any* $m \in \mathbb{N}$ *and* $K \in (0, 1)$,

$$\mathbb{P}\left(\left|\sum_{k=1}^{m} \tau_k - 1\right| > Km\right) < 2e^{-\frac{1}{2}K^2 m}. \tag{A.14}$$

Proof. Fix m. For any $\gamma \in (0, 1)$,

$$\mathbb{P}\left(\left|\sum_{k=1}^{m} \tau_k - 1\right| > Km\right) = \mathbb{P}\left(\sum_{k=1}^{m} \tau_k > (1 + K)m\right) + \mathbb{P}\left(-\sum_{k=1}^{m} \tau_k > -(1 - K)m\right)$$

$$= \mathbb{P}\left(e^{\gamma \sum_{k=1}^{m} \tau_k} > e^{(1+K)\gamma m}\right) + \mathbb{P}\left(e^{-\gamma \sum_{k=1}^{m} \tau_k} > e^{-(1-K)\gamma m}\right)$$

$$\leq e^{-(1+K)\gamma m} \left(\mathbb{E}\left[e^{\gamma \tau}\right]\right)^m + e^{(1-K)\gamma m} \left(\mathbb{E}\left[e^{-\gamma \tau}\right]\right)^m$$

$$= e^{-(1+K)\gamma m} (1 - \gamma)^{-m} + e^{(1-K)\gamma m} (1 + \gamma)^{-m}$$

$$= \exp\left(-\gamma m\left[1 + K + \frac{\log(1 - \gamma)}{\gamma}\right]\right)$$

$$+ \exp\left(\gamma m\left[1 - K - \frac{\log(1 + \gamma)}{\gamma}\right]\right).$$

The inequality is Markov's inequality and the equality immediately after the inequality follows from independence together with $\mathbb{E}[\exp(\alpha \tau)] = (1 - \alpha)^{-1}$ for any $\alpha \in (-1, 1)$. The other steps are all algebraic manipulations. By Taylor's theorem with remainder there exists $\gamma_1 \in (-\gamma, 0)$ such that

$$\frac{1}{\gamma} \log(1 - \gamma) = -1 - \frac{\gamma}{2(1 - \gamma_1)^2} > -1 - \frac{\gamma}{2},$$

where the inequality follows since $\gamma_1 < 0$. Therefore

$$\exp\left(-\gamma m\left[1 + K + \frac{\log(1 - \gamma)}{\gamma}\right]\right) \leq \exp\left(-\gamma m\left[K - \frac{\gamma}{2}\right]\right).$$

Similarly,

$$\exp\left(\gamma m\left[1 - K - \frac{\log(1 + \gamma)}{\gamma}\right]\right) \leq \exp\left(-\gamma m\left[K - \frac{\gamma}{2}\right]\right).$$

So combining with the first computation of this proof and taking $\gamma = K$ gives

$$\mathbb{P}\left(\left|\sum_{k=1}^{m} \tau_k - 1\right| > Km\right) \leq 2\exp\left(-\gamma m\left[K - \frac{\gamma}{2}\right]\right) = 2e^{-\frac{1}{2}K^2 m}.$$

$\square$

A.4. Proof of Lemma 4.6. Fix $t > 0$. The argument is similar to that of Lemma 4.2.
Variation of constants. Fix $m \in \mathbb{N}$. Since $\widetilde{S}^m_{h\tau} = \exp(h\tau_1 V_1) \cdots \exp(h\tau_{mn} V_{mn})$,

$$\partial_h \widetilde{S}^m_{h\tau} = \sum_{k=1}^{mn} \tau_k \widetilde{S}^{(1,k-1)}_{h\tau} V_k \widetilde{S}^{(k,mn)}_{h\tau} = \sum_{k=1}^{mn} \tau_k V_k \widetilde{S}^m_{h\tau} + \tau_k [\widetilde{S}^{(1,k-1)} h\tau, V_k] \widetilde{S}^{(k,mn)}_{h\tau}$$

$$= m V \widetilde{S}^m_{h\tau} + \sum_{k=1}^{mn} (\tau_k - 1) V_k \widetilde{S}^m_{h\tau} + \sum_{k=1}^{mn} \tau_k [\widetilde{S}^{(1,k-1)}_{h\tau}, V_k] \widetilde{S}^{(k,mn)}_{h\tau},$$

where the second equality is obtained by commuting $\widetilde{S}^{(1,k-1)}_{h\tau}$ and V_k, and the third by replacing τ_k with $\tau_k - 1 + 1$. So, setting $E^{(m)}_{h\tau} := \sum_{k=1}^{mn} \tau_k [\widetilde{S}^{(1,k-1)}_{h\tau}, V_k] \widetilde{S}^{(k,mn)}_{h\tau}$, variation of constants implies

$$\widetilde{S}^m_{h\tau} - S_{hm} = \int_0^h S_{m(h-r)} \left(\sum_{k=1}^{mn} (\tau_k - 1) V_k \right) \widetilde{S}^m_{r\tau} dr + \int_0^h S_{m(h-r)} E^{(m)}_{r\tau} dr.$$

Therefore, since $\|S_{m(h-r)}\|_{0\to 0} = 1$,

$$\|\widetilde{S}^m_{h\tau} - S_{hm}\|_{2\to 0} \leq \int_0^h \left\| \sum_{k=1}^{mn} (\tau_k - 1) V_k \right\|_{1\to 0} \|\widetilde{S}^m_{r\tau}\|_{2\to 1} dr + \int_0^h \|E^{(m)}_{r\tau}\|_{2\to 0} dr.$$

Let $I_1(h)$ and $I_2(h)$ denote the first and second integrals, respectively. Then for any $\varepsilon > 0$,

$$\mathbb{P}\left(\|\widetilde{S}^m_{h\tau} - S_{hm}\|_{2\to 0} > \frac{\varepsilon}{m} \right) \leq \mathbb{P}\left(I_1(h) > \frac{\varepsilon}{2m} \right) + \mathbb{P}\left(I_2(h) > \frac{\varepsilon}{2m} \right). \qquad \text{(A.15)}$$

We consider the two probabilities on the right, called the *first* and *second probabilities*, separately.

First probability. Note $\sum_{k=1}^{mn} (\tau_k - 1) V_k = \sum_{k=1}^{n} \sum_{j=1}^{m} (\tau^{(k)}_j - 1) V_k$ where $\tau^{(k)}_j := \tau_{(j-1)n+k}$. So

$$\left\| \sum_{k=1}^{mn} (\tau_k - 1) V_k \right\|_{1\to 0} \leq C_* \sum_{k=1}^{n} \left| \sum_{j=1}^{m} \tau^{(k)}_j - 1 \right|,$$

and together with Lemma A.1 and Equation (A.3),

$$I_1(h) \leq C_* \sum_{k=1}^{n} \left| \sum_{j=1}^{m} \tau^{(k)}_j - 1 \right| \int_0^h \prod_{k=1}^{n} e^{C_* r \sum_{j=1}^{m} \tau^{(k)}_j} dr.$$

Therefore

$$\mathbb{P}\left(I_1(h) > \frac{\varepsilon}{2m} \right) \leq \mathbb{P}\left(C_* \sum_{k=1}^{n} \left| \sum_{j=1}^{m} \tau^{(k)}_j - 1 \right| \int_0^h \prod_{k=1}^{n} e^{C_* r \sum_{j=1}^{m} \tau^{(k)}_j} dr > \frac{\varepsilon}{2m} \right)$$

$$\leq \sum_{k=1}^{n} \mathbb{P}\left(\left| \sum_{j=1}^{m} \tau^{(k)}_j - 1 \right| \int_0^h \prod_{k=1}^{n} e^{C_* r \sum_{j=1}^{m} \tau^{(k)}_j} dr > \frac{\varepsilon}{2 C_* mn} \right).$$

The second inequality follows from a union bound together with the fact that for any nonnegative random variables X_k and constant c, $\{\sum_{k=1}^n X_k > c\} \subseteq \cup_{k=1}^n \{X_k > c/n\}$. Set

$$A(h) := \bigcap_{k=1}^n \left\{ h \sum_{j=1}^m \tau_j^{(k)} \leq \alpha \right\} \quad \text{and}$$

$$B_k(h) := \left\{ \left| \sum_{j=1}^m \tau_j^{(k)} - 1 \right| \int_0^h \prod_{k=1}^n e^{C_* r \sum_{j=1}^m \tau_j^{(k)}} dr > \frac{\varepsilon}{2 C_* mn} \right\}$$

for arbitrary $\alpha > 0$ and note that

$$A(h) \cap B_k(h) \subseteq \left\{ \left| \sum_{j=1}^m \tau_j^{(k)} - 1 \right| h e^{C_* n\alpha} > \frac{\varepsilon}{2 C_* mn} \right\} =: B(h).$$

Therefore

$$\mathbb{P}\left(I_1(h) > \frac{\varepsilon}{m} \right) \leq \sum_{k=1}^n \mathbb{P}\left(B_k(h) \cap A(h) \right) + \mathbb{P}\left(B_k(h) \cap A(h)^c \right) \leq n \big[\mathbb{P}\left(B(h) \right) + \mathbb{P}\left(A(h)^c \right) \big].$$

Set $h = t/m^2$. By Lemma A.4 for all $\varepsilon > 0$ such that $K := \varepsilon (2 C_* tn)^{-1} e^{-C_* n\alpha} < 1$,

$$\mathbb{P}\left(B(h) \right) = \mathbb{P}\left(\left| \sum_{j=1}^m \tau_j^{(k)} - 1 \right| > \frac{\varepsilon m}{2 C_* tn e^{C_* n\alpha}} \right) \leq 2 e^{-\frac{1}{2} K^2 m}.$$

And by Lemma A.3,

$$\mathbb{P}\left(A(h)^c \right) = \mathbb{P}\left(\bigcup_{k=1}^n \left\{ \sum_{j=1}^m \tau_j^{(k)} > \frac{\alpha}{h} \right\} \right) \leq n \mathbb{P}\left(\sum_{j=1}^m \tau_j > \frac{\alpha m^2}{t} \right) \leq n 2^m e^{-\frac{1}{2} K' m^2}$$

where $K' := \alpha/t$. Therefore

$$\mathbb{P}\left(I_1(h) > \frac{\varepsilon}{2m} \right) \leq 2 e^{-\frac{1}{2} K^2 m} + 2^m n e^{-\frac{1}{2} K' m^2} \leq 2^m C e^{-\frac{1}{2} C m^2} \tag{A.16}$$

for some positive constant C independent of m.

Second probability. Recall $E_{r\tau}^{(m)} := \sum_{k=1}^{mn} \tau_k [\widetilde{S}_{r\tau}^{(1,k-1)}, V_k] \widetilde{S}_{r\tau}^{(k,mn)}$. Also, from Equation (A.6),

$$[\widetilde{S}_{r\tau}^{(1,k-1)}, V_k] \widetilde{S}_{r\tau}^{(k,mn)} = \sum_{j=1}^{k-1} \int_0^r \tau_j e^{(r-s) \sum_{j=1}^{k-1} \tau_j V_j} [V_j, V_k] \widetilde{S}_{s\tau}^{(1,k-1)} \widetilde{S}_{r\tau}^{(k,mn)} ds$$

$$+ \sum_{j=1}^{k-1} \int_0^r \tau_j e^{(r-s) \sum_{j=1}^{k-1} \tau_j V_j} [[\widetilde{S}_{s\tau}^{(1,j-1)}, V_j] \widetilde{S}_{s\tau}^{(j,k-1)}, V_k] \widetilde{S}_{r\tau}^{(k,mn)} ds.$$

Lemma A.1 together with $\|[V_j, V_k]\|_{2\to 0} \leq \|V_j\|_{1\to 0}\|V_k\|_{2\to 1} + \|V_k\|_{1\to 0}\|V_j\|_{2\to 1} \leq 2C_*^2$ give

$$\left\|[V_j, V_k]\widetilde{S}_{s\tau}^{(1,k-1)}\widetilde{S}_{r\tau}^{(k,mn)}\right\|_{2\to 0} \leq 2C_*^2\left(1 + C_*r\sum_{j=1}^{mn}\tau_j\right)e^{3C_*r\sum_1^{mn}\tau_j}.$$

Also,

$$[[\widetilde{S}_{s\tau}^{(1,j-1)}, V_j]\widetilde{S}_{s\tau}^{(j,k-1)}, V_k] = \widetilde{S}_{s\tau}^{(1,j-1)}V_j\widetilde{S}_{s\tau}^{(j,k-1)}V_k - V_k\widetilde{S}_{s\tau}^{(1,j-1)}V_j\widetilde{S}_{s\tau}^{(j,k-1)}$$
$$- V_j\widetilde{S}_{s\tau}^{(1,k-1)}V_k + V_kV_j\widetilde{S}_{s\tau}^{(1,k-1)}$$

together with Lemma A.1 gives

$$\left\|[[\widetilde{S}_{s\tau}^{(1,j-1)}, V_j]\widetilde{S}_{s\tau}^{(j,k-1)}, V_k]\widetilde{S}_{r\tau}^{(k,mn)}\right\|_{2\to 0} \leq 4C_*^2\left(1 + C_*r\sum_{j=1}^{mn}\tau_j\right)e^{3C_*r\sum_1^{mn}\tau_j}.$$

Therefore for any $0 \leq r \leq h$,

$$\left\|E_{r\tau}^{(m)}\right\|_{2\to 0} \leq \sum_{k=1}^{mn}\sum_{j=1}^{k-1}\tau_k\tau_j\int_0^r\left\|[V_j, V_k]\widetilde{S}_{s\tau}^{(1,k-1)}\widetilde{S}_{r\tau}^{(k,mn)}\right\|_{2\to 0}$$
$$+ \left\|[[\widetilde{S}_{s\tau}^{(1,j-1)}, V_j]\widetilde{S}_{s\tau}^{(j,k-1)}, V_k]\widetilde{S}_{r\tau}^{(k,mn)}\right\|_{2\to 0}ds$$
$$\leq 6C_*^2r\left(1 + C_*r\sum_{\ell=1}^{mn}\tau_\ell\right)e^{3C_*r\sum_1^{mn}\tau_\ell}\sum_{k=1}^{mn}\sum_{j=1}^{k-1}\tau_k\tau_j$$
$$\leq Ch\left(1 + Ch\sum_{\ell=1}^{mn}\tau_\ell\right)e^{Ch\sum_1^{mn}\tau_\ell}\left(\sum_{k=1}^{mn}\tau_k\right)^2$$

for some $C > 0$. So, we have that

$$I_2(h) = \int_0^h\|E_{r\tau}^{(m)}\|_{2\to 0}dr \leq Ch^2\left(1 + Ch\sum_{\ell=1}^{mn}\tau_\ell\right)e^{Ch\sum_1^{mn}\tau_\ell}\left(\sum_{k=1}^{mn}\tau_k\right)^2.$$

For arbitrary $\alpha > 0$, set

$$A(h) := \left\{h\sum_{k=1}^{mn}\tau_k \leq \alpha\right\}\quad\text{and}$$
$$B(h) := \left\{Ch^2\left(1 + Ch\sum_{\ell=1}^{mn}\tau_\ell\right)e^{Ch\sum_1^{mn}\tau_\ell}\left(\sum_{k=1}^{mn}\tau_k\right)^2 > \frac{\varepsilon}{2m}\right\}.$$

Then taking $h = t/m^2$ as before,

$$
\mathbb{P}\left(I_2(h) > \frac{\varepsilon}{2m}\right) = \mathbb{P}\left(A(h) \cap B(h)\right) + \mathbb{P}\left(A(h)^c \cap B(h)\right)
$$

$$
\leq \mathbb{P}\left(Ch^2\left(1 + C\alpha\right)e^{C\alpha}\left(\sum_{k=1}^{mn}\tau_k\right)^2 > \frac{\varepsilon}{2m}\right) + \mathbb{P}\left(h\sum_{k=1}^{mn}\tau_k > \alpha\right)
$$

$$
= \mathbb{P}\left(\sum_{k=1}^{mn}\tau_k > Km^{\frac{3}{2}}\right) + \mathbb{P}\left(\sum_{k=1}^{mn}\tau_k > \frac{\alpha m^2}{t}\right) \qquad \text{(A.17)}
$$

$$
\leq n\left[\mathbb{P}\left(\sum_{k=1}^{m}\tau_k > K'm^{\frac{3}{2}}\right) + \mathbb{P}\left(\sum_{k=1}^{m}\tau_k > \frac{\alpha m^2}{nt}\right)\right]
$$

$$
\leq n\left(2^m e^{-\frac{1}{2}K'm^{3/2}} + 2^m e^{-\frac{1}{2}K''m^2}\right) \leq 2^m C' e^{-\frac{1}{2}C'm^{3/2}}
$$

for some $C' > 0$ where $K = (\varepsilon(2t^2 C(1 + C\alpha)e^{C\alpha})^{-1})^{1/2}$, $K' = Kn^{-1}$, $K'' = \alpha(nt)^{-1}$, and the second-to-last last inequality follows from Lemma A.3. Combining (A.15), (A.16), and (A.17) and taking $h = t/m^2$ we therefore have that for all ε sufficiently small,

$$
\mathbb{P}\left(\|\widetilde{S}^m_{t\tau/m^2} - S_{t/m}\|_{2\to 0} > \frac{\varepsilon}{m}\right) \leq 2^m C'' e^{-\frac{1}{2}C''m^{3/2}}
$$

for some constant $C'' > 0$ independent of m. So, we have that

$$
\sum_{m=1}^{\infty}\mathbb{P}\left(\|\widetilde{S}^m_{t\tau/m^2} - S_{t/m}\|_{2\to 0} > \frac{\varepsilon}{m}\right) \leq \sum_{m=1}^{\infty} 2^m C'' e^{-\frac{1}{2}C''m^{3/2}} < \infty.
$$

$\square$

Appendix B. Controllability Lemmas

Combining the partial results obtained above we show the existence of transformations implementing the steps listed at the beginning of the section:

Lemma B.1. *If $q^{(0)}$ in $\mathcal{Q}_0$ is nondegenerate, then there exists M_1 and a sequence of transition times and interaction triples $\{\iota(m), \tau(m)\}_{m=1}^{M_1}$ such that $\Phi^{\iota(M_1)}_{\tau(M_1)} \circ \cdots \circ \Phi^{\iota(1)}_{\tau(1)}(q^{(0)}) = q^{(1)}$ as in (6.17).*

Proof. If (6.17) is satisfied by $q^{(0)}$ we simply set $M_1 = 0$, $q^{(1)} = q^{(0)}$. If not, by nondegeneracy there exists a sequence of triples $\{\iota(m)\}_{m=1}^{M}$ with $\iota(m) = j(m)k(m)\ell(m)$ such that $\mathcal{A}_0 := \mathcal{A}(q^{(0)})$ and $\mathcal{A}_m = \mathcal{A}_{m-1} \oplus \ell(m)$ with $\{(0, 1, +), (1, 0, +), (j^*, -)\} \subset \mathcal{A}_M$. We notice that all steps of this procedure satisfy, upon possibly reordering the indices within each triple, either the conditions of Lemma 6.11 (b) or of Lemma 6.12, so we sequentially choose $\tau(m) = \tau^{\iota(m)}_+$ from those lemmas.

To activate coordinate $(1, 1, -)$ – if this was not already done in the previous procedure – we start with component $b_{j^*} \neq 0$ for $|j^*| \neq 1$ and consider a nearest neighbors path $\{\ell(n)\}_{n=1}^{M'}$ in $\mathbb{Z}_N^2$ connecting j^* to $(1, 1)$ without performing any step on the axes. It is easy to see that such path can be realized through repeated application of Lemma 6.11

(b) by choosing for the n-th step the triples $\iota(n) = (0, 1, +)(\ell(n), -)(\ell(n)\pm(0, 1), -)$ or $\iota(n) = (1, 0, +)(\ell(n), -)(\ell(n)\pm(1, 0), -)$ for vertical and horizontal steps respectively.

Finally, coordinates $(1, 0, -)$ and $(0, 1, -)$ can be activated by applying Lemma 6.13 to the triples $(1, 0, -)(0, 1, +)(1, 1, -)$ and $(1, 0, +)(0, 1, -)(1, 1, -)$ respectively, while $(1, 1, +)$ is activated by (b) by interchanging the type of modes $(1, 1, -)$ and $(1, 0, +)$ (or $(0, 1, +)$) in $\iota(M')$ from the previous paragraph to $(1, 1, +)$ and $(1, 0, -)$ (or $(0, 1, -)$).□

Lemma B.2. *Let $q^{(1)}$ be a nondegenerate point in $\mathcal{Q}_0$ satisfying (6.17). Then there exists M_2 and a sequence of interacting triples and transition times $\{\iota(m), \tau(m)\}_{m=1}^{M_2}$ such that $\Phi_{\tau(M_2)}^{\iota(M_2)} \circ \cdots \circ \Phi_{\tau(1)}^{\iota(1)}(q^{(1)}) = q^{(2)}$ is a nondegenerate point in $\mathcal{Q}_0$ satisfying (6.18) and (6.19).*

Proof. In this part of the proof, we only consider interactions involving triples $\iota(m)$ of the form

$$\left\{ (0, 1)(l, h)(l, h \pm 1) \text{ or } (1, 0)(l, h)(l \pm 1, h) : |l|, |h| \leq N, \ |(l, h)| \neq 1 \right\}. \quad \text{(B.1)}$$

By Lemma 6.11 (a), if $|j| < |k| < |\ell|$ and $(0, 1), (l, h) \in \mathcal{A}(q)$ there exists $\tau(m) = \tau_-^{\iota(m)}$ such that defining $\mathcal{A}_m = \mathcal{A}(\varphi_{\tau(m)}^{\iota(m)}(q))$ we have $(l, h) \notin \mathcal{A}_m$ and $(0, 1) \in \mathcal{A}_m$ (and similarly for $(1, 0)$).[10] Note that while a triple as above satisfies by assumption that $|j| < |k| < |\ell|$ and at least two of its coordinates are nonvanishing, it does not, in general, satisfy (6.24). However, assuming that q does not satisfy (6.24), by Lemma 6.12 and setting $\iota' = (1, 0)(0, 1)(1, 1)$, there exists $\tau^{\iota'}$ such that $|q_{(1,0)}| \neq |(\Phi_{\tau^{\iota'}}^{\iota'}(q))_{(1,0)}| > 0$. Since none of the coordinates in $\mathbb{Z}_N^2 \setminus \{(1, 0)(0, 1)(1, 1)\}$ are affected by this operation, $(\Phi_{\tau^{\iota'}}^{\iota'}(q))$ satisfies (6.24) and Lemma 6.11 can be applied to this state.

To conclude the proof we identify a sequence of triples $\iota(m) = (j(m), k(m), \ell(m)) \in \mathcal{I}$ of the form (B.1) such that for $\mathcal{A}_0 = \mathcal{A}(q^{(1)}) \subseteq \mathbb{Z}_N^2 \times \{+, -\}$

$$\begin{aligned}(((\mathcal{A}_0 \ominus k(1)) \ominus k(2)) \ominus \dots) \ominus k(M_2) \\ = \{(1, 0, \chi), (0, 1, \chi), (1, 1\chi), (N, N\chi), (-N, N\chi), \chi \in \{+, -\}\}.\end{aligned}$$

A possible such sequence is given by triples of the form

$$\left\{ (1, 0, +)(l, h, \chi)(l + 1, h, \chi) \ : \ (l, h) \in \{(0, 2), \dots, (0, N)\}, \chi \in \{+, -\} \right\}$$

to remove the vertical column of $\mathbb{Z}_N^2$ (which cannot interact with $(0, 1)$), followed by

$$\left\{ \left((0, 1, +)(l, h, \chi)(l, h + 1, \chi) \ : \ (l, h) \in \{(l, 0), \dots, (l, N) : |l| \in (1, \dots, N - 1)\} \setminus \{(1, 1)\} \right), \chi \in \{+, -\} \right\},$$

where importantly the set of transitions for each l is ordered. The above transformation zeroes all coefficients except those in the set $\{(1, 1), (0, 1), (1, 0)\} \cup \{(l, N) \ : \ l \in (-N, \dots, N)\}$. We further remove the coefficients from $\{(l, N) \ : \ l \in (-N+1, \dots, N-1)\}$ by sequentially applying Lemma 6.11 to the ordered sequence of interacting triples

$$\left((1, 0, +)(l, h, \chi)(l + 1, h, \chi) \ : \ (l, h) \in \{(0, N), \dots, (N - 1, N)\}, \chi \in \{+, -\} \right),$$

---

[10] Note that the same result can trivially be obtained if $(l, h) \notin \mathcal{A}(q)$ setting $\tau_-^{\iota(m)} = 0$.

and then

$$\left((1, 0, +)(l, h, \chi)(l - 1, h, \chi) \ : \ (l, h) \in \{(-1, N), \dots, (-N + 1, N)\}, \chi \in \{+, -\}\right).$$

It is easy to check that each transition in the above construction sequentially satisfies the assumptions of Lemma 6.11 (a), and that once a mode has been removed from $\mathcal{A}$ it will not interact again in this procedure. The fact that (6.19) holds follows from (6.17) and that in an interacting triple $\iota = jk\ell$ with $|j| < |k| < |l|$ both modes j and ℓ are in $\mathcal{A}$ at the end of the interaction by τ_-^ι. □

Lemma B.3. *Let $q^{(2)}$ be a nondegenerate point in $\mathcal{Q}_0$ satisfying (6.18) and (6.19). Then there exists M_3 and a sequence of interacting triples and transition times $\{\iota(m), \tau(m)\}_{m=1}^{M_3}$ such that $\Phi_{\tau(M_3)}^{\iota(M_3)} \circ \cdots \circ \Phi_{\tau(1)}^{\iota(1)}(q^{(2)}) = q^{(3)}$ is a nondegenerate point in $\mathcal{Q}_0$ satisfying (6.20) and (6.21).*

Since it may not be possible to "transfer" the content of *e.g.*, mode $(-N, N)$ to $(-N + 1, N)$ through one single interaction with mode $(1, 0)$ – and therefore it won't be possible to transfer the amplitude of mode $(-N, N)$ to (N, N) in one single "pass" – we proceed to prove that, through a sequence of interactions, we can transfer a *finite* and $q_{(-N,N)}$-independent amount of energy from mode $(-N, N)$ to (N, N). Therefore, the transfer of amplitude from mode $(-N, N)$ to (N, N) may be accomplished by repeating this sequence of interactions sufficiently many times.

The following corollary of Lemma 6.12 will be instrumental for the proof of Lemma B.3:

Corollary B.4. *Let $q_{(1,1)}, b_{(1,1)} \neq 0$ then for any q, q' with $q_j = q'_j$ for all $|j| > 1$ there exist a sequence $\{\iota(m), \tau(m)\}_{m=1}^4$ such that $\Phi_{\tau(4)}^{\iota(4)} \circ \cdots \circ \Phi_{\tau(1)}^{\iota(1)}(q) = q'$.*

Proof of Lemma B.3. The desired result follows upon showing that for any $i \in \{-N, \dots, N\}$, setting $\ell = (-i, N, \chi)$, $\ell' = (i, N, \chi')$ for $\chi, \chi' \in \{-, +\}$ there exists $M_{\ell,\ell'}$ and a sequence of triples and interaction times $\{\iota(m), \tau(m)\}_{m=1}^{M_{\ell,\ell'}}$ such that for any q satisfying $\bigcup_{|i'|<i}\{(i', N, +), (i', N, -)\} \cap \mathcal{A}(q) = \emptyset$ and $q' = \Phi_{\tau(M_{\ell,\ell'})}^{\iota(M_{\ell,\ell'})} \circ \cdots \circ \Phi_{\tau(1)}^{\iota(1)}(q)$ we have

$$q'_j = \begin{cases} q_j & \text{for } j \in \mathbb{Z}_N^2 \setminus \{\ell, \ell'\}, \\ 0 & \text{for } j = \ell \text{ if } \ell \neq \ell', \end{cases} \tag{B.2}$$

and for $k \in \{\ell, \ell'\}$, $\text{sign}(q_k) = \text{sign}(q'_k)$ holds if $q'_k \neq 0$ (recalling our choice of notation $\text{sign}(0) = +1$). Indeed, if $\text{sign}(b_{(N,N)}) \geq 0$ we sequentially apply the above result to the pairs

$$(\ell, \ell') = ((N, N, +), (-N, N, +)), ((-N, N, +), (N, N, -)), ((-N, N, -), (N, N, -)).$$

Otherwise, when $\text{sign}(b_{(N,N)}) = -1$ we first apply the above result to $\ell = (N, N, -)$, $\ell' = (-N, N, -)$ and then proceed as in the previous case.

We prove the result above by induction on $i \in \{0, \dots, N\}$. The proof for $i \leq 0$ is analogous.

Base case ($i = 0 : (0, N, \chi) \rightarrow (0, N, \chi')$): If $\ell = \ell'$ there is nothing to show. We proceed to consider the case $\ell = (0, N, +)$, $\ell' = (0, N, -)$, as the converse follows by analogous arguments. In this case, for a sufficiently small $\varepsilon > 0$ we consider

the interactions $\iota = (1, 0, +)(0, N, +)(1, N, +)$ and $\iota' = (1, 0, -)(0, N, -)(1, N, +)$, running the corresponding flow maps by a small amount of time $\tau(\varepsilon)$, $\tau'(\varepsilon)$ such that $(\Phi^{\iota'}_{\tau'(\varepsilon)} \circ \Phi^{\iota}_{\tau(\varepsilon)}(q)_{(0,N,-)})^2 = b^2_{(0,N)} + \varepsilon$. We then apply Corollary B.4 to the coordinates $(1, 0, +)$, $(1, 0, -)$ to return them in the initial configuration. Note that the existence of a uniform $\varepsilon > 0$ such that the transitions above can be performed in a single pair of interactions (and therefore the finiteness of the total number of interactions required to perform the desired transformation) follows from the fact that $b_{(0,N)}$ is nondecreasing and the continuity of the dynamics together with Lemma 6.11.

Induction step ($i > 0 : (-i, N, \chi) \to (i, N, \chi')$): We consider two possibilities for q: a) there exists q'' with $|a''_{(1,0)}| \in [|a_{(1,0)}|/2, |a_{(1,0)}|]$, $q''_{(-i,N,\chi)} = 0$ and for $\iota'' = (1, 0, +)(-i+1, N, \chi)(-i, N, \chi)$

$$E_{\iota''}(q) = E_{\iota''}(q''), \quad \mathcal{E}_{\iota''}(q) = \mathcal{E}_{\iota''}(q''),$$

or b) such q'' does not exist.

In case a) the state q'' can be reached by letting $\iota = (1, 0, +)(-i+1, N, \chi)(-i, N, \chi)$ interact for a finite amount of time τ from Lemma 6.11 (c). Then, by the induction assumption there is a sequence of triples and interaction times allowing to reach a state q''' with $q'''_{(-i+1,N,\chi)} = 0$, $q'''_{(i-1,N,\chi')} = q''_{(-i+1,N,\chi)}$ and $q'''_j = q''_j$ for all other $j \in \mathbb{Z}_N^2$. The desired state can then be reached by application of Lemma 6.11 (a) to the triple $\iota = (1, 0, +)(i - 1, N, \chi')(i, N, \chi')$. We proceed to check that the final state satisfies (B.2). Because modes $j \notin \{(-i, N), \ldots, (i, N), (1, 0)\}$ did not interact in the procedure above for such j we must have that $q_j = q'_j$. The fact that for $j \in \{(-i, N), \ldots, (i - 1, N)\}$ $q'_j = 0$ follows by construction and the induction assumption. It remains to check that $|a'_{(1,0)}| = |a_{(1,0)}|$. Since the only modes affected by the above transformation are $(-i, N, \chi)$, $(i, N\chi')$, $(1, 0, +)$, this follows directly by conservation of energy and enstrophy:

$$(q_{(-i,N,\chi)})^2 + (q_{(i,N,\chi')})^2 + (q_{(1,0,+)})^2 = (q'_{(i,N,\chi')})^2 + (q'_{(1,0,+)})^2,$$

$$\frac{(q_{(-i,N,\chi)})^2}{N^2 + i^2} + \frac{(q_{(i,N,\chi')})^2}{N^2 + i^2} + (q_{(1,0,+)})^2 = \frac{(q'_{(-i,N,\chi)})^2}{N^2 + i^2} + (q'_{(1,0,+)})^2.$$

In case b) we proceed to show that case a) can be reached with a finite number of interactions. More specifically if condition a) is not satisfied we let the triple $\iota'' = (-i, N, \chi)(-i+1, N, \chi)(1, 0, +)$ for $\chi \in \{+, -\}$ interact as described by Lemma 6.11 for a time τ'' to reach a nondegenerate point q'' in $\mathcal{Q}_0$ with $q''_j = q_j$ for $j \notin \{(-i, N, \chi), (-i+1, N, \chi), (1, 0, +)\}$, $a''_{(1,0)} = a_{(1,0)}/2$ and $q''_{(-i,N,\chi)}$, $q''_{(-i+1,N,\chi)}$ satisfying the conservation laws

$$(q_{(-i,N,\chi)})^2 + (q_{(1,0,+)})^2 = (q''_{(-i,N,\chi)})^2 + (q''_{(-i+1,N,\chi)})^2 + (q_{(1,0,+)}/2)^2,$$

$$\frac{(q_{(-i,N,\chi)})^2}{N^2 + i^2} + (q_{(1,0,+)})^2 = \frac{(q''_{(-i,N,\chi)})^2}{N^2 + i^2} + \frac{(q''_{(-i+1,N,\chi)})^2}{N^2 + (i-1)^2} + (q_{(1,0,+)}/2)^2,$$

so that $(q''_{(-i,N,\chi)})^2 = (q_{(-i,N,\chi)})^2 - C_{N,i}(q_{(1,0)})^2$ for $C_{N,i} = \frac{3}{4} \frac{N^2+i^2}{i^2-(i-1)^2}(N^2+(i-1)^2-1)$. We see that a positive, $q_{(1,0,+)}$-dependent amplitude is removed from $(q_{(-i,N,\chi)})^2$. Again applying the induction step and Lemma 6.11 (a) to transfer, respectively, the amplitude from $(-i + 1, N, \chi)$ to $(i - 1, N, \chi')$ and from $(i - 1, N, \chi')$ to (i, N, χ')

we reach the state q' with $q_j = q'_j$ for modes $j \notin \{(-i, N), \ldots, (i, N), (1, 0)\}$ (since these modes either vanish in both cases or they did not interact). Further, by conservation of energy and enstrophy, we have that

$$(q_{(-i,N,\chi)})^2 + (q_{(i,N,\chi')})^2 + (q_{(1,0,+)})^2 = (q''_{-i,N,\chi})^2 + (q''_{(i,N,\chi')})^2 + (q''_{(1,0,+)})^2 ,$$

$$\frac{(q_{(-i,N,\chi)})^2}{N^2 + i^2} + \frac{(q_{(i,N,\chi')})^2}{N^2 + i^2} + (q_{(1,0,+)})^2 = \frac{(q''_{-i,N,\chi})^2}{N^2 + i^2} + \frac{(q''_{(i,N,\chi')})^2}{N^2 + i^2} + (q''_{(1,0,+)})^2 ,$$

so that $|q''_{(1,0,+)}| = |q_{(1,0,+)}|$. This shows that the amplitude $C_{N,i}(q_{(1,0,+)})^2$ subtracted to $q_{(-i,N,\chi)}$ is constant at each cycle, showing by boundedness of $q_{(-i,N,\chi)}$ that with a finite number of iterations as the one described above we can reach state a), concluding the proof.   □

Lemma B.5. *Let $q^{(3)}$ be a nondegenerate point in $\mathcal{Q}_0$ satisfying* (6.20) *and* (6.21). *Then there exists M_4 and a sequence of interacting triples and transition times $\{\iota(m), \tau(m)\}_{m=1}^{M_4}$ such that $\Phi_{\tau(M_4)}^{\iota(M_4)} \circ \cdots \circ \Phi_{\tau(1)}^{\iota(1)}(q^{(3)}) = q^*$ is a nondegenerate point in $\mathcal{Q}_0$ satisfying* (6.15).

Proof. We start the proof by applying Corollary B.4 to transform the state $q^{(3)}$ into $q = \Phi_{\tau(1)}(q^{(3)})$ satisfying $q_j^{(3)} = q_j$ for all $|j| > 1$ and $a_{(0,1)} = b_{(0,1)} = b_{(1,0)} = a_{(1,0)} > 0$. Throughout this proof, we refer to states q such that $q_{(i,i',\chi)} = q_{(i',i,\chi)}$ for all $i, i' \in (0, \ldots, N)$, $\chi \in \{+, -\}$ as *symmetric*.

We then proceed to transfer the amplitude from $a_{(1,1)}$ to $b_{(2,1)}, b_{(1,2)}$ by transforming q into another symmetric state q' with $(2, 1, -), (1, 2, -) \in \mathcal{A}(q')$ and $(1, 1, +) \notin \mathcal{A}(q')$. This can be done by letting triples $\iota(2) = (1, 0, -)(1, 1, +)(2, 1, -) \in \mathcal{I}$ and $\iota(3) = (0, 1, -)(1, 1, +)(1, 2, -) \in \mathcal{I}$ interact, and choosing the interaction times $\tau, \tau'(\tau)$ such that $\Phi_{\tau'(\tau)}^{\iota(3)} \circ \Phi_{\tau}^{\iota(2)}(q)_{(1,1,+)} = 0$. Further, we note that the difference $b'_{(1,2)} - b'_{(2,1)}$ is negative for $\tau = 0$, positive for $\tau'(\tau) = 0$ and is continuous in τ, so there must exist τ^* such that $b'_{(1,2)} = b'_{(2,1)}$. To show that q' is symmetric it only remains to show that $b'_{(1,0)} = b'_{(0,1)}$. This follows from the conservation laws:

$$B_{(1,0)(1,1)} \left((b'_{(1,0)})^2 - (b_{(1,0)})^2\right) = B_{(2,1)(1,1)}(b'_{(2,1)})^2 = B_{(1,2)(1,1)}(b'_{(1,2)})^2$$
$$= B_{(0,1)(1,1)} \left((b'_{(0,1)})^2 - (b_{(0,1)})^2\right)$$

where

$$B_{jk} := \frac{1}{|j|^2} - \frac{1}{|k|^2} .$$

Next, we let the triples $\iota(4) = (1, 0, -)(0, 1, +)(1, 1, -)$ and $\iota(5) = (0, 1, -)(1, 0, +)(1, 1, -)$ interact. By Lemma 6.12 there exists an interaction time such that the initial state q' is mapped to q'' with $b''_{(1,0)} = b''_{(0,1)} = 0$ and $a''_{(1,0)} = a''_{(0,1)} > 0$, so that $(1, 0, -), (0, 1, -) \notin \mathcal{A}(q'')$.

We then proceed to transfer the amplitude from modes $(1, 2, -)$ and $(2, 1, -)$ to $(2, 2, -)$. This is done letting triples $\iota(6) = (1, 0, +)(1, 2, -)(2, 2, -)$ and $\iota(7) = (0, 1, +)(2, 1, -)(2, 2, -)$ interact until the modes $(2, 1, -), (1, 2, -)$ are depleted, as proved in Lemma 6.11. The symmetry of the final state q''' is again a consequence of the conservation laws:

$$B_{(1,0)(2,2)} \left((a'''_{(1,0)})^2 - (a''_{(1,0)})^2\right) = B_{(2,1)(2,2)}(b''_{(2,1)})^2 = B_{(1,2)(2,2)}(b''_{(1,2)})^2$$

$$= B_{(0,1)(2,2)} \left((a'''_{(0,1)})^2 - (a''_{(0,1)})^2 \right).$$

Summarizing, we have reached a symmetric state $q''' = \Phi^{\iota(7)}_{\tau(7)} \circ \cdots \circ \Phi^{\iota(2)}_{\tau(2)}(q)$ with

$$\mathcal{A}(q''') = \{(1, 0, +), (0, 1, +), (2, 2, -), (1, 1, -), (N, N, -)\}.$$

The desired result then follows immediately if we can show that we can transfer the amplitude of mode $(i - 1, i - 1, -)$ to $(i, i, -)$ for $i \in (2, \ldots, N)$ while preserving the fact that $a'_{(1,0)} = a'_{(0,1)}$. We show this by considering, sequentially, the interaction triples

$$\begin{aligned}
\iota(4i) &= (1, 0, +)(i - 1, i - 1, -)(i, i - 1, -), \\
\iota(4i + 1) &= (0, 1, +)(i - 1, i - 1, -)(i - 1, i, -), \\
\iota(4i + 2) &= (0, 1, +)(i, i - 1, -)(i, i, -), \\
\iota(4i + 3) &= (1, 0, +)(i - 1, i, -)(i, i, -).
\end{aligned}$$

More specifically, we consider the family of endpoints

$$q''(t) = \Phi^{\iota(4i+3)}_{\tau_-^{\iota(4i+3)}} \circ \Phi^{\iota(4i+2)}_{\tau_-^{\iota(4i+2)}} \circ \Phi^{\iota(4i+1)}_{\tau_-^{\iota(4i+1)}} \circ \Phi^{\iota(4i)}_{t}(q'),$$

where τ_-^ι is defined in Lemma 6.11 (a). By construction, this sequence implies that $a''_{(i-1,i-1)} = a''_{(i-1,i)} = a''_{(i,i-1)} = 0$ and $a''_{(i,i)} \neq 0$. It remains to prove that $a''_{(1,0)} = a''_{(0,1)}$. As a composition of continuous functions, $q''(t)$ is continuous in t and therefore so is $\Delta q(t) = a''_{(1,0)}(t) - a''_{(0,1)}(t)$. Further, since by symmetry $a''_{(1,0)}(0) = a''_{(0,1)}(\tau_-^{\iota(4i)})$, we must have $\mathrm{sign}(\Delta q(0)) = -\mathrm{sign}(\Delta q(\tau_-^{\iota_1}))$. This implies the existence of $\tau(4i) \in [0, \tau_-^{\iota_1}]$ with $\Delta q(0) = 0$, concluding the proof. $\qquad\square$

References

1. Arnold, A., Ringhofer, C.: An operator splitting method for the Wigner–Poisson problem. SIAM J. Numer. Anal. **33**(4), 1622–1643 (1996)
2. Bakhtin, Y., Hurth, T.: Invariant densities for dynamical systems with random switching. Nonlinearity **25**, 2937–2952 (2012)
3. Bakhtin, Y., Hurth, T., Lawley, S.D., Mattingly, J.C.: Smooth invariant densities for random switching on the torus. Nonlinearity **31**(4), 1331–1350 (2018)
4. Bakhtin, Y., Hurth, T., Mattingly, J.C.: Regularity of invariant densities for 1D systems with random switching. Nonlinearity **28**(11), 3755–3787 (2015)
5. Bedrossian, J., Blumenthal, A., Punshon-Smith, S.: Almost-sure enhanced dissipation and uniform-in-diffusivity exponential mixing for advection–diffusion by stochastic Navier–Stokes. Probab. Theory Relat. Fields **179**(3–4), 777–834 (2021)
6. Benaïm, M., Borgne, S.L., Malrieu, F., Zitt, P.-A.: Qualitative properties of certain piecewise deterministic Markov processes. Ann. Inst. Henri Poincaré Probab. Stat. **51**(3), 1040–1075 (2015)
7. Benaïm, M., Hurth, T., Strickler, E.: A user-friendly condition for exponential ergodicity in randomly switched environments. Electron. Commun. Probab. **23**, 1–12 (2018)
8. Bermúdez, B., Nikolás, A., Sánchez, F.J.: On operator splitting methods with upwinding for the unsteady Navier–Stokes equations. East-West J. Numer. Math. **4**(2), 83–98 (1996)
9. Bierkens, J., Roberts, G.O., Zitt, P.-A.: Ergodicity of the zigzag process. Ann. Appl. Probab. **29**(4), 2266–2301 (2019)
10. Childs, A.M., Ostrander, A., Su, Y.: Faster quantum simulation by randomization. Quantum **3**, 182 (2019)
11. Childs, A.M., Su, Y., Tran, M.C., Wiebe, N., Zhu, S.: Theory of trotter error with commutator scaling. Phys. Rev. X **11**(1), 011020 (2021)

12. Cornfeld, I.P., Fomin, S.V., Sinaĭ, Y.G.: Ergodic Theory, Volume 245 of Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer, New York (1982). (**Translated from the Russian by A. B. Sosinskiĭ**)
13. Costa, O.L.V., Dufour, F.: Singular perturbation for the discounted continuous control of piecewise deterministic Markov processes. Appl. Math. Optim. **63**(3), 357–384 (2011)
14. Da Prato, G., Debussche, A.: Two-dimensional Navier–Stokes equations driven by a space-time white noise. J. Funct. Anal. **196**(1), 180–210 (2002)
15. Debussche, A., Nguepedja Nankep, M.J.: A piecewise deterministic limit for a multiscale stochastic spatial gene network. Appl. Math. Optim. **84**(suppl. 2), S1731–S1767 (2021)
16. Devaney, R.L.: An Introduction To Chaotic Dynamical Systems, 3rd edn. CRC Press, Boca Raton (2022)
17. Durmus, A., Guillin, A., Monmarché, P.: Piecewise deterministic Markov processes and their invariant measures. Ann. Inst. Henri Poincaré Probab. Stat. **57**(3), 1442–1475 (2021)
18. E, W., Mattingly, J.C.: Ergodicity for the Navier–Stokes equation with degenerate random forcing: finite-dimensional approximation. Commun. Pure Appl. Math. **54**(11), 1386–1402 (2001)
19. E, W., Mattingly, J.C., Sinai, Y.: Gibbsian dynamics and ergodicity for the stochastically forced Navier–Stokes equation. Commun. Math. Phys. **224**(1), 83–106 (2001). (**Dedicated to Joel L. Lebowitz**)
20. Flandoli, F., Maslowski, B.: Ergodicity of the 2-D Navier–Stokes equation under random perturbations. Commun. Math. Phys. **172**(1), 119–141 (1995)
21. Freidlin, M.I., Wentzell, A.D.: Random Perturbations of Dynamical Systems. Volume 260 of Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences, 2nd edn. Springer, New York (1998). (**Translated from the 1979 Russian original by Joseph Szücs**)
22. Goldman, D., Kaper, T.J.: Nth-order operator splitting schemes and nonreversible systems. SIAM J. Numer. Anal. **33**(1), 349–367 (1996)
23. Hairer, M.: Convergence of Markov processes. Lecture notes (2010)
24. Hairer, M., Mattingly, J.C.: Ergodicity of the 2d Navier–Stokes equations with degenerate stochastic forcing. Ann. Math. **164**, 993–1032 (2006)
25. Hairer, M., Mattingly, J.C.: Yet another look at Harris' ergodic theorem for Markov chains. In: Seminar on Stochastic Analysis, Random Fields and Applications VI, Volume 63 of Progress in Probability, pp. 109–117. Birkhäuser, Basel (2011)
26. Jurdjevic, V.: Geometric Control Theory. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge (1996)
27. Katok, A., Hasselblatt, B.: Introduction to the Modern Theory of Dynamical Systems, Volume 54 of Encyclopedia of Mathematics and Its Applications. Cambridge University Press, Cambridge (1995). (**With a supplementary chapter by Katok and Leonardo Mendoza**)
28. Kifer, Y.: Ergodic Theory of Random Transformations. Progress in Probability and Statistics, vol. 10. Birkhäuser Boston Inc, Boston, MA (1986)
29. Kuksin, S., Nersesyan, V., Shirikyan, A.: Exponential mixing for a class of dissipative PDEs with bounded degenerate noise. Geom. Funct. Anal. **30**(1), 126–187 (2020)
30. Kuksin, S., Shirikyan, A.: Stochastic dissipative PDEs and Gibbs measures. Commun. Math. Phys. **213**(2), 291–330 (2000)
31. Kuksin, S., Shirikyan, A.: Some limiting properties of randomly forced two-dimensional Navier–Stokes equations. Proc. R. Soc. Edinb. Sect. A **133**(4), 875–891 (2003)
32. Kusuoka, S.: Approximation of Expectation of Diffusion Processes Based on Lie Algebra and Malliavin Calculus, Volume 6 of Advanced Mathematical Economics, pp. 69–83. Springer, Berlin (2004)
33. Lawley, S.D.: Extreme first passage times of piecewise deterministic Markov processes. Nonlinearity **34**(5), 2750–2780 (2021)
34. Lawley, S.D., Mattingly, J.C., Reed, M.C.: Sensitivity to switching rates in stochastically switched ODEs. Commun. Math. Sci. **12**(7), 1343–1352 (2014)
35. Lawley, S.D., Mattingly, J.C., Reed, M.C.: Stochastic switching in infinite dimensions with applications to random parabolic PDE. SIAM J. Math. Anal. **47**(4), 3035–3063 (2015)
36. Lee, J.M.: Riemannian Manifolds: An Introduction to Curvature. Springer, New York (1997)
37. Lee, J.M.: Introduction to Smooth Manifolds. Springer, New York (2000)
38. Li, D., Liu, S., Cui, J.: Threshold dynamics and ergodicity of an sirs epidemic model with Markovian switching. J. Differ. Equ. **263**(12), 8873–8915 (2017)
39. MacNamara, S., Strang, G.: Operator splitting. In: Splitting Methods in Communication, Imaging, Science, and Engineering. Springer, Cham, pp. 95–114 (2016)
40. Mattingly, J.C.: On recent progress for the stochastic Navier Stokes equations. In: Journées "Équations aux Dérivées Partielles". University of Nantes, Nantes, pp. Exp. No. XI, 52 (2003)
41. Mattingly, J.C., Stuart, A.M., Higham, D.J.: Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. Stoch. Process. Appl. **101**(2), 185–232 (2002)
42. McNabb, A.: Comparison theorems for differential equations. J. Math. Anal. Appl. **119**(1–2), 417–428 (1986)

43. Meyn, S., Tweedie, R.L.: Markov Chains and Stochastic Stability, 2nd edn. Cambridge University Press, Cambridge (2009). (**With a prologue by Peter W. Glynn**)
44. Monmarché, P.: On $\mathcal{H}^1$ and entropic convergence for contractive PDMP. Electron. J. Probab. **20**(:Paper No. 128), 30 (2015)
45. Nagano, T.: Linear differential systems with singularities and an application to transitive Lie algebras. J. Math. Soc. Jpn. **18**(4), 398–404 (1966)
46. Ninomiya, M., Ninomiya, S.: A new higher-order weak approximation scheme for stochastic differential equations and the Runge–Kutta method. Finance Stoch. **13**(3), 415–443 (2009)
47. Ninomiya, S., Victoir, N.: Weak approximation of stochastic differential equations and application to derivative pricing. Appl. Math. Finance **15**(2), 107–121 (2008)
48. Nummelin, E.: The discrete skeleton method and a total variation limit theorem for continuous-time Markov processes. Math. Scand. **42**(1), 150–160 (1978)
49. Nummelin, E.: General Irreducible Markov Chains and Nonnegative Operators. Cambridge Tracts in Mathematics, vol. 83. Cambridge University Press, Cambridge (1984)
50. Sussmann, H.J.: Orbits of families of vector fields and integrability of distributions. Trans. Am. Math. Soc. **180**, 171–188 (1973)
51. Williamson, B.: On SDEs with partial damping inspired by the Navier–Stokes equations. Ph.D. thesis, Duke University (2019). https://hdl.handle.net/10161/18773

Communicated by M. Hairer