

# Stationary measures for stochastic differential equations with degenerate damping

Jacob Bedrossian\*      Kyle Liss†

June 5, 2022

## Abstract

A variety of physical phenomena involve the nonlinear transfer of energy from weakly damped modes subjected to external forcing to other modes which are more heavily damped. In this work we explore this in (finite-dimensional) stochastic differential equations in  $\mathbb{R}^n$  with a quadratic, conservative nonlinearity  $B(x, x)$  and a linear damping term  $-Ax$  which is degenerate in the sense that  $\ker A \neq \emptyset$ .

We investigate sufficient conditions to deduce the existence of a stationary measure for the associated Markov semigroups. Existence of such measures is straightforward if  $A$  is full rank, but otherwise, energy could potentially accumulate in  $\ker A$  and lead to almost-surely unbounded trajectories, making the existence of stationary measures impossible. We give a relatively simple and general sufficient condition based on time-averaged coercivity estimates along trajectories in neighborhoods of  $\ker A$  and many examples where such estimates can be made.

2010 MSC. 60H10, 60H50, 37A25 37H30.

Keywords. Stationary measures, degenerate damping, nonlinear energy transfer

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Main results	4
1.2	Discussion and related work	8
<b>2</b>	<b>Time-averaged coercivity near <math>\ker A</math></b>	<b>9</b>
2.1	Time-averaged coercivity implies existence	9
2.2	Short-time coercivity near $\ker A$	11
<b>3</b>	<b>Conservative flow transverse to the kernel</b>	<b>14</b>
<b>4</b>	<b>Unstable equilibria in the kernel</b>	<b>18</b>
4.1	Spectrally unstable estimates	18
4.2	Jordan block unstable estimates	23
4.3	Concluding the proofs of Theorems 1.2 and 1.3	27
<b>5</b>	<b>Sabra and Galerkin Navier-Stokes</b>	<b>29</b>
5.1	Statement and proof of general result	29
5.2	Applications to the Sabra shell model and Galerkin Navier-Stokes	34

---

\*Department of Mathematics, University of Maryland, College Park, MD 20742, USA [jacob@math.umd.edu](mailto:jacob@math.umd.edu). J.B. was supported by NSF CAREER grant DMS-1552826 and NSF Award DMS-2108633.

†Department of Mathematics, Duke University, Durham, NC 27710, USA [kyle.liss@duke.edu](mailto:kyle.liss@duke.edu). K.L. was supported by NSF Award No. DMS-2038056 and DMS-1552826

<b>6</b>	<b>Lorenz-96 with a two-dimensional kernel</b>	<b>36</b>
<b>7</b>	<b>Stochastic triad model with non-trivial, invariant conservative dynamics in the kernel</b>	<b>45</b>
7.1	Local Lyapunov function	45
7.2	Growth of approximate solution	49
7.3	Justifying the approximation	50
<b>A</b>	<b>A basic energy estimate</b>	<b>53</b>
	<b>References</b>	<b>54</b>

## 1 Introduction

A variety of physical phenomena involve the nonlinear transfer of energy from weakly damped modes subjected to external forcing to other modes which are more heavily damped. In hydrodynamic turbulence for example, the forcing is considered to act at large scales whereas in the high Reynolds number limit, the viscous dissipation is only strong at very high frequencies. This leads to the phenomenon known as anomalous dissipation (see e.g. [6, 16]). A study of such phenomena in infinite-dimensional systems remains largely out of reach (with a few exceptions, for example some simplified shell models [15, 27] and Batchelor-regime passive scalar turbulence [3]). As suggested in e.g. [26], it is natural to first study the analogues in finite-dimensional systems. In this setting we will study systems with damping which only acts on a proper subset of the degrees of freedom and ask the question of whether or not a statistical equilibrium, i.e. a stationary measure, can still be shown to exist. If the undamped modes are directly forced at least, for this to be possible the nonlinearity must continually pump energy away from the modes without damping into modes with damping.

We study the following prototypical class of stochastic differential equations (SDEs) for  $x_t \in \mathbb{R}^n$

$$\begin{cases} dx_t = B(x_t, x_t)dt - Ax_t dt + \sigma dW_t \\ x_t|_{t=0} = x_0 \in \mathbb{R}^n. \end{cases} \quad (1.1)$$

Here,  $W_t = (W_t^{(1)}, \dots, W_t^{(n)})$  is an  $n$ -dimensional canonical Brownian motion on a complete probability space  $(\Omega, \mathcal{F}, \mathbf{P})$ ,  $A \in \mathbb{R}^{n \times n}$  is positive semi-definite (with  $\ker A \neq \emptyset$ ), and  $\sigma \in \mathbb{R}^{n \times n}$ . We will assume for simplicity throughout this introduction that  $\sigma$  is full rank, though, as discussed in the main body of the text, weaker conditions are possible for the examples we study. The nonlinear term  $B$  is bilinear such that the energy  $|x|^2$  is conserved:

$$x \cdot B(x, x) = 0. \quad (1.2)$$

Many of the specific examples we study also satisfy  $\nabla \cdot B = 0$ , but this is not required for our methods. This class of systems contains Galerkin truncations of both the 2d and 3d Navier-Stokes equations, as well as Lorenz-96 [23], and the classical shell models of hydrodynamic turbulence, GOY [17, 33] and Sabra [25]; see e.g. [26] for further discussions on the motivations for studying this class of SDEs. It is straightforward to show that the SDEs are globally well-posed and the associated Markov semigroups are well-behaved; see e.g. [Appendix A; 4]]. We will refer to the ODE

$$\frac{d}{dt} z_t = B(z_t, z_t)$$

as the *conservative dynamics*. This deterministic ODE plays a distinguished role, as it is the leading order dynamics at high energies, i.e. when  $|x| \gg 1$ .

Denote the generator

$$\mathcal{L} = \frac{1}{2} \sigma \sigma^T : \nabla^2 - Ax \cdot \nabla + B(x, x) \cdot \nabla \quad (1.3)$$

and the associated Markov semigroups  $\mathcal{P}_t = e^{t\mathcal{L}}$  and  $\mathcal{P}_t^* = e^{t\mathcal{L}^*}$ , the former acting on the space of bounded, Borel measurable observables  $B_b(\mathbb{R}^n; \mathbb{R})$  and the latter acting on Borel probability measures  $\mathcal{P}(\mathbb{R}^n)$ . When  $A$  is positive definite, it is not hard to prove that there always exists at least one stationary measure, i.e. a measure  $\mu \in \mathcal{P}(\mathbb{R}^n)$  such that  $\mathcal{P}_t^* \mu = \mu$ . This is proved by the Krylov-Bogoliubov procedure (see e.g. [10]) combined with the following energy balance obtained from Itô's lemma:

$$\frac{1}{2} \mathbf{E} |x_t|^2 + \mathbf{E} \int_0^t x_s \cdot Ax_s ds = \frac{t}{2} \sum_{i,j=1}^n \sigma_{ij}^2 + \frac{1}{2} \mathbf{E} |x_0|^2.$$

However, if  $\ker A \neq \emptyset$ , then there is the possibility that energy could accumulate in these degrees of freedom and the a priori estimate

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \mathbf{E} \int_0^t x_s \cdot Ax_s ds \leq \frac{1}{2} \sum_{i,j=1}^n \sigma_{ij}^2 \quad (1.4)$$

would not be sufficient to imply the compactness required for Krylov-Bogoliubov.

It is well known that to prove the existence of a stationary measure it suffices to construct a *Lyapunov function*, i.e., a  $C^2$  function  $V : \mathbb{R}^n \rightarrow [0, \infty)$  satisfying  $\lim_{|x| \rightarrow \infty} V(x) = \infty$  and

$$\mathcal{L}V \leq -\alpha V^p + \beta \quad (1.5)$$

for some  $\alpha, \beta > 0$  and  $p \in (0, 1]$ . Indeed, this is a straightforward generalization of the argument recalled above using Itô's lemma and the Krylov-Bogoliubov procedure. Note that if the kernel of  $A$  is trivial, then  $V(x) = |x|^2$  is a Lyapunov function for (1.1), while if  $\ker A \neq \emptyset$  then (1.5) holds only in regions where  $|x| \lesssim |\Pi_{\ker A^\perp} x|$ . There are many works that have successfully constructed an invariant measure and/or obtained convergence rates to equilibrium for SDEs with partial dissipation or unstable deterministic dynamics by building a nontrivial Lyapunov function (see e.g. [1, 5, 14, 19, 20, 30, 32]). A general strategy for constructing a Lyapunov function is to patch together a sequence of local Lyapunov functions, each satisfying (1.5) in a different part of phase space. In regions where (1.5) is not obviously satisfied by some natural energy-type function, a common approach is to perform a scaling analysis and show (1.5) for a reduced generator, and then justify the full inequality by an approximation argument. For a discussion of scaling arguments and a meta-algorithm for constructing Lyapunov functions, see [1]. The Lyapunov functions obtained by such methods tend to be quite involved, even in low dimensional, relatively simple systems (see e.g. [1, 14, 19] and [section 2, [32]]), and require a careful gluing of separate local Lyapunov functions.

In this paper, we develop a framework for constructing invariant measures for partially damped systems based on returning to the simple a priori energy estimate (1.4). Rather than directly building a Lyapunov function, the idea is to recover compactness by proving that the time-averaged dissipation controls the average of some simple coercive function. More precisely, our strategy is to prove the following time-averaged coercivity estimate for some  $T \in (0, 2)$  and  $r \in (0, 1]$ ,

$$\frac{1}{T} \mathbf{E} \int_0^T \langle x_t \rangle^{2r} dt \lesssim 1 + \frac{1}{T} \mathbf{E} \int_0^T x_t \cdot Ax_t dt, \quad (1.6)$$

which we show is sufficient to imply existence in Lemma 2.1 by a straightforward iteration procedure. In Lemma 2.2, we reduce this to short-time coercivity estimates for trajectories starting in a relatively small neighborhood of  $\ker A$  at high energy. Specifically, we show that it suffices to prove (1.6) for initial conditions  $x_0 \in \mathbb{R}^n$  satisfying

$$|\Pi_{\ker A^\perp} x_0| \ll |\Pi_{\ker A} x_0|^r, \quad |x_0| \gg 1$$

and the time  $T$  depending on the initial energy  $|x_0|$ . The goal is thus to prove that at high energies, where the conservative dynamics dominate, solutions that start near  $\ker A$  must depart rapidly (on average) due to some kind of instability. Our strategy to prove the necessary time-averaged coercivity estimates is to use a suitable approximation of the solution when  $|\Pi_{\ker A^\perp} x_t| \ll |\Pi_{\ker A} x_t|^r$ , show that this approximate solution rapidly enters the region  $|\Pi_{\ker A^\perp} x| \gtrsim |\Pi_{\ker A} x|^r$ , and then argue that the approximation remains valid for as long as  $|\Pi_{\ker A^\perp} x_t| \lesssim |\Pi_{\ker A} x_t|^r$ .

The time-averaged coercivity framework is convenient in that it allows one to leverage in a natural way assumptions on the instability of  $\ker A$  under the dynamics to obtain existence of an invariant measure and an explicit convergence rate to equilibrium. Moreover, it avoids the need to carefully patch together separate local Lyapunov functions, which is required even if one uses a construction based on local exit times. We will showcase the flexibility of our methods by presenting a variety of examples to which they apply, in each case showing a different potential case that arises with degenerate damping. The examples below are chosen to show qualitatively distinct cases where the approximation procedure described above can be justified, although a different choice of approximate solution is used in each type of example.

## 1.1 Main results

We now discuss our main results and their connection to some of the existing literature on related SDEs. Below, denote the set of undamped configurations on the unit energy sphere by

$$\mathcal{U} = \ker A \cap \mathbb{S}^{n-1}.$$

The first theorem considers the case where  $\mathcal{U}$  contains no sets which are invariant under the conservative dynamics. This case is analogous to the settings considered by hypocoercivity, which usually studies nontrivial interplay between degenerate elliptic operators and conservative first order operators (such as transport) to obtain decay estimates, despite the lack of coercivity; see discussions in e.g. [18, 31]. Indeed, the results we are proving are quite similar to (sub-exponential) hypocoercivity results for the associated Markov semigroups (although here we use different, essentially probabilistic, methods). See [2, 8] for further discussion on the relationship between Harris' theorems and commonly used hypocoercivity methods. The intuition is clear: if  $\mathcal{U}$  contains no sets which are invariant under the conservative dynamics, then at high energies any trip to a small neighborhood of  $\ker A$  must necessarily be short lived.

**Theorem 1.1.** *Suppose that  $\exists J \in \mathbb{N}$  such that  $\forall x \in \mathcal{U}$ , if  $X_t$  solves the conservative dynamics*

$$\begin{cases} \frac{d}{dt} X_t = B(X_t, X_t) \\ X_0 = x, \end{cases} \quad (1.7)$$

then

$$\exists j \leq J, \Pi_{\ker A^\perp} \frac{d^j}{dt^j} X_t|_{t=0} \neq 0. \quad (1.8)$$

Then, there exists at least one stationary measure  $\mu$  and  $\langle x \rangle^p \in L^1(d\mu)$  for all  $p < \infty$ .

**Remark 1.** Condition (1.8) implies that solutions to (1.7) that start on  $\ker A$  instantly depart it (at least at a rate like  $\gtrsim (Kt)^J$  if  $|x| \approx K$ ; see Lemma 3.2). Note that the condition in (1.8) is purely algebraic, that is, in principle it could be investigated using methods from algebraic geometry, rather than being an abstract condition on trajectories. Related algebraic conditions describing the instability of a set under some conservative dynamics have appeared in [32].

**Remark 2.** As to be expected, Theorem 1.1 requires no assumptions on  $\sigma$ .

In [14], the authors consider the stochastically driven Lorenz-63 model, a classical three dimensional model introduced in [22]. This model does not take exactly the form of (1.1) due to the presence of a non-dissipative linear term, but the setting is essentially the same since there still exists a natural energy function that yields an invariant measure when  $\ker A = \emptyset$ . The authors consider the case where  $\ker A = \text{span}\{e_k\}$  for some canonical unit vector  $e_k$  and consists of conservative equilibria that to leading order at high energies exhibit a Jordan block instability. They prove using a Lyapunov function approach that if the noise directly excites the instability, then there always exists a stationary measure. The next theorem is a similar kind of result but generalized to higher dimensional systems in which  $\ker A = \text{span}\{e_k\}$  for  $e_k$  a general unstable equilibrium point of the conservative dynamics. Unlike in the setting of Theorem 1.1, in this case we cannot depend purely on the conservative dynamics to simply transport the  $x_t$  away from  $\ker A$ . Instead, we must rely on the noise to push the dynamics off of the equilibrium and its stable manifold so that  $x_t$  is repelled quickly from neighborhoods of  $\ker A$  at high energy. We denote the (instantaneous) linearization of the conservative nonlinearity around any fixed  $x$  as

$$L_x v = B(x, v) + B(v, x) \quad (1.9)$$

and for the restriction to  $\ker A^\perp$  we write

$$L_x^\perp v = \Pi_{\ker A^\perp} L_x \Pi_{\ker A^\perp}.$$

Recall that for simplicity we assume for now unless otherwise stated that  $\text{rank}(\sigma) = n$ .

**Theorem 1.2.** *Suppose that  $\mathcal{U} = \{x_0, -x_0\}$  for some unit vector  $x_0$  and that for each  $x \in \mathcal{U}$  there holds*

$$B(x, x) = 0, \quad \lim_{t \rightarrow \infty} \left\| e^{tL_x^\perp} \right\| = \infty.$$

*Then, there exists at least one stationary measure  $\mu$  and  $\langle x \rangle^p \in L^1(d\mu)$  for all  $p < 1/3$ .*

**Remark 3.** If  $x_0$  and  $-x_0$  are spectrally unstable, i.e.  $L_{\pm x_0}$  has an eigenvalue  $\lambda$  with  $\text{Re}\lambda > 0$ , then the stationary measure satisfies  $\langle x \rangle^p \in L^1(d\mu)$  for all  $p < \infty$ . Notice however, that in general we do *not* require that  $x_0$  is spectrally unstable, that is, it is sufficient for the equilibria to have an  $O(t)$  growth coming from a non-trivial Jordan block. We did not take care in this paper to optimize the moment bounds on the stationary measures that we construct and in general they are probably far from sharp. For example, it is likely that  $\mu$  has exponential moments in many cases. In fact, the existence of an invariant measure with exponential moments was proven for a 3d model satisfying the conditions of Theorem 1.2 in [32].

**Remark 4.** The condition  $\text{rank}(\sigma) = n$  is not necessary. What is used in the proof is essentially that the range of  $\sigma$  contains at least one eigenvector or generalized eigenvector associated with the fastest instability of  $L_x^\perp$ . For the precise statement of Theorem 1.2 with weaker assumptions on  $\sigma$ , see Theorem 4.5. In fact, none of the theorems we prove require the forcing to act on all variables. We expect that all of the theorems that rely on unstable equilibria hold only under the assumption that the forcing is hypoelliptic if all of the instabilities are spectral, however, we did not pursue this direction here. Similarly, we expect variations of these results to be valid with multiplicative stochastic forcing under suitable assumptions.

We can also treat cases with  $\dim(\ker A) > 1$  provided that  $\mathcal{U}$  consists either entirely of spectrally unstable equilibria or Jordan block unstable equilibria. In the latter case we require an additional cancellation condition due to the slower timescale of the instability.

**Theorem 1.3.** *Suppose that  $B(x, x) = 0$  for every  $x \in \mathcal{U}$  and that there exists a constant  $C > 0$  so that*

$$\sup_{x \in \mathcal{U}} (\|P_x\| + \|P_x^{-1}\|) \leq C, \quad (1.10)$$

*where  $J_x^\perp = P_x^{-1} L_x^\perp P_x$  is the Jordan canonical form of  $L_x^\perp$ . Then, we have the following results.*

- If for every  $x \in \mathcal{U}$  there is an eigenvalue of  $L_x^\perp$  with positive real part, then there exists at least one stationary measure  $\mu$  and  $\langle x \rangle^p \in L^1(d\mu)$  for every  $p < 2/3$ .
- Assume that

$$\Pi_{\ker A} (B(\Pi_{\ker A} x, \Pi_{\ker A^\perp} x) + B(\Pi_{\ker A^\perp} x, \Pi_{\ker A} x)) = 0$$

for every  $x \in \mathbb{R}^n$ . If for every  $x \in \mathcal{U}$  there exists  $J \in \{1, 2, \dots, n-2\}$  so that there holds

$$t^J \lesssim_x \|e^{tL_x^\perp}\| \lesssim_x (1 + t^J)$$

for all  $t \geq 0$ , then there exists at least one stationary measure  $\mu$  and  $\langle x \rangle^p \in L^1(d\mu)$  for every  $p < 1/3$ .

**Remark 5.** Analogous criteria to Theorems 1.2, 1.3, and Theorem 1.1 can be found for much more general nonlinearities, i.e. systems of the form  $dx_t = F(x_t)dt - Ax_t + \sigma dW_t$  with  $x \cdot F(x) = 0$ , however, the lack of scaling invariance requires slightly more care.

A first natural question is whether or not Theorem 1.1 and Theorems 1.2, 1.3 can be combined into one. We do not know how to do this in reasonable generality due to difficulties in dealing with transitions between “transverse” regions as in Theorem 1.1 and unstable equilibria as in Theorems 1.2 and 1.3. However, in Section 5 we prove Theorem 5.1, which provides at least one general setting where this is possible. Specifically, we consider systems for which  $\ker A = V_1 \oplus V_2$  for subspaces  $V_1, V_2 \subseteq \mathbb{R}^n$  consisting of spectrally unstable equilibria and such that the region where  $\Pi_{V_1} x$  and  $\Pi_{V_2} x$  are both sufficiently large can be treated as a transverse zone. Note that in this setting the instability of  $\Pi_{V_j} x$  need not cause growth of the damped modes directly, but could instead cause the solution to enter a transverse region, where it is then subsequently expelled from  $\ker A$  in a manner similar to Theorem 1.1. While we require some additional structural assumptions to justify the approximations, Theorem 5.1 applies to several well-known examples, for example the Sabra model with  $\ker A$  given by the first two frequency shells (which means  $\dim(\ker A) = 4$ ) and the 2d Galerkin-Navier-Stokes equations with  $\ker A$  consisting of a four-dimensional subspace of suitably chosen shear flows. We will state here our result on the Navier-Stokes equations, and defer the general result and application to Sabra to Section 5.

Recall the 2d Navier-Stokes equations in vorticity form on a square torus  $\mathbb{T}^2$  subjected to stochastic forcing:

$$dw + (u \cdot \nabla w - \Delta w)dt = \sum_{k \in \mathbb{Z}^2: k \neq 0} \sigma_k^{(1)} \cos(k \cdot x) dW_t^{(k;1)} + \sigma_k^{(2)} \sin(k \cdot x) dW_t^{(k;2)}$$

$$u = \begin{pmatrix} -\partial_{x_2} \\ \partial_{x_1} \end{pmatrix} (-\Delta)^{-1} w.$$

Let  $\Pi_{\leq N}$  be the projection to the modes such that  $\max(|k_1|, |k_2|) =: |k|_\infty \leq N$  (any choice of  $\ell^p$  works). Then the Galerkin Navier-Stokes equations are given by the SDE defined for mean-zero  $w \in \text{Im} \Pi_{\leq N}$  by

$$dw + (\Pi_{\leq N}(u \cdot \nabla w) + Aw)dt = \sum_{0 < |k|_\infty \leq N} \sigma_k^{(1)} \cos(k \cdot x) dW_t^{(k;1)} + \sigma_k^{(2)} \sin(k \cdot x) dW_t^{(k;2)}$$

$$u = \begin{pmatrix} -\partial_{x_2} \\ \partial_{x_1} \end{pmatrix} (-\Delta)^{-1} w,$$

where we have replaced the matrix  $-\Pi_{\leq N} \Delta \Pi_{\leq N}$  with a general positive semi-definite matrix  $A$ .

**Theorem 1.4.** Let  $N \geq 3$  be arbitrary and define the two subspaces of  $\text{Im} \Pi_{\leq N}$

$$V_1 \oplus V_2 = \text{span}(\cos \ell x_1, \sin \ell x_1) \oplus \text{span}(\cos k x_2, \sin k x_2),$$

for two arbitrary integers  $\ell, k \geq 2$  such that  $\ell \neq k$  and  $\max(\ell, k) \leq N$ . Suppose further that the forcing coefficients  $\sigma_p^{(j)}$  are all non-zero. If  $\ker A = V_1 \oplus V_2$  then there exists a (unique) invariant measure  $\mu_*$  of the Galerkin Navier-Stokes equations with truncation  $N$  and for all  $p < 2/3$  there holds

$$\int_{\text{Im}\Pi_{\leq N}} |w|^p d\mu_* < \infty.$$

As an additional example in a setting similar to Theorem 5.1 described above, we consider the Lorenz-96 model, put forward by Lorenz in [23], for  $n$  real-valued unknowns  $u_1, \dots, u_n$  in a periodic ensemble  $u_{i+kn} = u_i$ :

$$du_m = (u_{m+1} - u_{m-2})u_{m-1}dt - (Au)_m dt + q_m dW_t^{(m)}. \quad (1.11)$$

Here,  $\{W_t^{(m)}\}$  are independent Brownian motions and  $\{q_m\}$  are fixed parameters. This model has been studied as a prototypical high dimensional chaotic system (see e.g. [21, 24, 26]). We consider (1.11) with

$$\ker A = \{u_1 = u_2 = 0\}.$$

Similar to the general setting of Theorem 5.1, this example contains a mixture of all of Theorems 1.1, 1.2, and 1.3 in the sense that  $\mathcal{U}$  contains both unstable equilibria and a region in which the conservative dynamics expel from  $\ker A$  as in Theorem 1.1. However, the equilibria are only Jordan block unstable, so Theorem 5.1 (the proof of which relies crucially on the exponential instability of the equilibria in  $V_j$ ) does not apply. The linear instability of the equilibria defined by  $u = \alpha e_2$  (i.e. only supported in the second mode) causes growth of the  $e_1$  direction, rather than a mode in  $\ker A^\perp$ . In this region of  $\ker A$ , a careful (and somewhat nonlinear) argument is used to show that the linear instability moves the dynamics into a region where the nonlinearity can then transport the dynamics out of  $\ker A$ . Despite the lack of unstable eigenvalues, using the precise structure of (1.11) we can justify the approximations needed to apply our methods and construct an invariant measure.

**Theorem 1.5.** *Let  $6 \leq n < \infty$  and suppose that  $q_{n-1}, q_n$  are both non-zero. Suppose that  $\ker A = \{u_1 = u_2 = 0\}$ . Then, (1.11) admits at least one stationary measure  $\mu$  and  $\langle x \rangle^p \in L^1(d\mu)$  for all  $p < 1/3$ .*

**Remark 6.** After completion of this work, we have been made aware of a similar result for Lorenz-96 in the upcoming thesis [7], which considers the case where  $n = 4$ ,  $\ker A$  consists of two modes, and the forcing acts only on the two modes in  $\ker A^\perp$ .

The above theorems do not contain all of the interesting possible relationships between  $\ker A$  and the dynamics of  $B$ . In particular, none of the above examples consider a case in which  $\mathcal{U}$  contains a non-equilibrium invariant set for the conservative dynamics. We give one such example where our methods apply, based on the following simple ‘‘stochastic triad’’ model [26] defined by the nonlinearity

$$B(x, x) = \begin{pmatrix} x_2 x_3 \\ x_1 x_3 \\ -2x_1 x_2 \end{pmatrix}. \quad (1.12)$$

The  $x_3$ -axis contains unstable equilibria and so Theorem 1.2 shows that if  $\ker A = \text{span}\{e_3\}$ , then there exists a stationary measure (this result was already proven in [32]). To contrast, the plane defined by  $\{x : x_1 = x_2\}$  consists of heteroclinic connections between the unstable equilibria with  $x_3 > 0$  and those with  $x_3 < 0$ , and so neither Theorem 1.2 nor Theorem 1.3 apply to the case that  $\ker A = \{x : x_1 = x_2\}$ . Nevertheless, we are able to adapt our methods to cover this case since we can precisely describe the conservative dynamics restricted to  $\ker A$ .

**Theorem 1.6.** *Consider the stochastic triad model defined by (1.12) in  $\mathbb{R}^3$  and suppose  $\ker A = \{x : x_1 = x_2\}$ . Then, there exists at least one stationary measure  $\mu$  and  $\langle x \rangle^p \in L^1(d\mu)$  for all  $p < 2/3$ .*

In all of the above examples, existing results give uniqueness and regularity of the stationary measure once existence is proved; see e.g. the Doob-Khasminskii theorem [10]. Moreover, the proof yields a sub-geometric Lyapunov function and one can apply a suitable variation of Harris' theorem to obtain explicit convergence estimates on the Markov semigroups in the total variation norm [13].

**Corollary 1.7.** *If  $\text{rank}(\sigma) = n$  and if  $x_t$  is irreducible, that is, if  $\forall x \in \mathbb{R}^n$ , open sets  $\mathcal{O} \subset \mathbb{R}^n$  and  $\forall t > 0$ ,*

$$\mathbf{P}(x_t \in \mathcal{O}) > 0,$$

*then in any of the above examples, there is a unique stationary measure  $\mu$ , and this stationary measure is  $C^\infty$ . Moreover, if  $V(x) = \langle x \rangle^2$ , then for  $T$  and  $r$  as in (1.6),*

$$\tilde{V}(x) = \frac{1}{T} \int_0^T \mathcal{P}_t V(x) dt$$

*satisfies for  $r$  as above and some constants  $c, C > 0$ ,*

$$\mathcal{L}\tilde{V} \leq -c\tilde{V}^r + C,$$

*and hence by results in [13], for any  $x \in \mathbb{R}^n$ , there holds (with the convention that if  $r = 1$ , then the decay is exponential)*

$$\|\mathcal{P}_t(x, \cdot) - \mu\|_{TV} \lesssim \langle t \rangle^{-\frac{r}{1-r}} \tilde{V}(x).$$

**Remark 7.** A simple energy estimate (Lemma A.1) shows that necessarily  $\tilde{V}(x) \gtrsim \langle x \rangle^2$ .

## 1.2 Discussion and related work

As alluded to above, the proofs of Theorems 1.1–1.6 are all about ruling out the possibility that energy accumulates into  $\ker A$  which is be done by demonstrating a time-averaged coercivity estimate of the form (1.6). For (1.6) to hold, we see that it would suffice to show that the solution does not spend a significant percentage of its time near  $\ker A$ . In fact, at higher energies, we show that the dynamics are expelled from neighborhoods of  $\ker A$  faster. This has a clear analogy with variations of hypocoercivity that emphasize this aspect (see discussions in [31]), however, these previously existing works are all essentially in the case of Theorem 1.1.

Section 2 provides two important lemmas: Lemma 2.1 shows that (1.6) suffices to prove the existence of stationary measures and Lemma 2.2 reduces this to short-time coercivity estimates in a small region of  $\ker A$  (see Assumption 1). Moreover, (1.6) implies the existence of a sub-geometric drift condition as pointed out in Corollary 1.7.

In order to prove (1.6) (via Lemma 2.2), it makes sense to proceed by contradiction. When (1.6) fails at high energy, it is necessary for the majority of the energy to be concentrated in a small region around  $\ker A$ , which could allow a perturbative treatment for as long as the dynamics remain in the small region. Theorem 1.1 simply uses the pure conservative dynamics as the approximate solution, whereas Theorems 1.2 and 1.3 use the linearization around  $\Pi_{\ker A} x_0$  (frozen in time) to justify the expulsion. Notice that the noise remains important here near the stable manifold of the equilibria. Theorems 1.4, 1.5, and 1.6 use more careful approximations based on what region of  $\ker A$  the solution is close to. For example, to prove Theorem 1.6, if one is near an equilibrium  $(0, 0, x_3)$  with  $x_3 > 0$ , then we first prove that with high probability the solution is rapidly transported away along the heteroclinic connections that run through  $\{x : x_1 = x_2\}$ , and then



show that it is likely to be expelled from  $\ker A$  along the unstable manifold of the corresponding equilibrium at  $(0, 0, -x_3)$ .

We remark here that many aspects of our work are not specific to the system (1.1) and could easily be adapted to various other regimes. However, systems with an underlying conservative dynamics which is a homogeneous polynomial (and hence a scale invariance is available), linear damping, and additive noise seem to be the simplest case to consider.

There are several works in the literature related to ours. The works that consider settings most similar to what we study here are [14] and [32]. In addition to the existence result discussed above around Theorem 1.2, it is proven in [14] that if the noise does not excite the instability and  $e_k$  is directly forced, then no stationary measure exists. The work [32] considers (1.1) with an additional structural assumption on  $B$  motivated by the nonlinearity in the Navier-Stokes equations. When the deterministic invariant subset of  $\ker A$ , denoted by  $\mathcal{N}$ , consists only of spectrally unstable equilibria, existence of an invariant measure is proven under an algebraic assumption that describes growth of the damped modes for initial conditions near  $\mathcal{N}$ . In the context of our work, this main result of [32] seems closely related to Theorem 5.1 and can be viewed essentially as sufficiently strict assumptions under which the combination of Theorems 1.1 and 1.3 is possible.

A set of works with close links to ours considers noise-induced stabilization for systems with deterministic dynamics that contain finite-time blow-up solutions; see e.g. [1, 19, 28, 30]. In these works, despite the finite-time blow-up of certain deterministic trajectories, depending on the noise or whether the blow-ups are unstable, one can nevertheless obtain almost-sure global well-posedness and prove the existence of stationary measures. The works using additive noise proceed by a Lyapunov function approach and so are closely related to [14]. Another related work is that of Coti Zelati and Hairer [9], which considers the Lorenz-63 system with  $\ker A = \emptyset$ , but where the forcing *only* acts on  $\text{span}\{e_3\}$ . This makes  $\text{span}\{e_3\}$  an almost-surely invariant set for the stochastically forced system, in which case, an argument based on transverse Lyapunov exponents can be made, providing another method for dynamically driving solutions away.

## 2 Time-averaged coercivity near $\ker A$

The purpose of this section is to prove a useful general result that will be applied to construct an invariant measure in each of the examples discussed in Section 1. The main abstract condition for the existence of invariant measures is stated below as Assumption 1. Intuitively, the condition requires that if the process enters the vicinity of  $\ker A$  at high energies, then it is quickly ejected and subsequently stays away from  $\ker A$  for some amount of time.

### 2.1 Time-averaged coercivity implies existence

In what follows denote

$$D(x) := x \cdot Ax.$$

Notice that

$$|\Pi_{\ker A^\perp} x|^2 \lesssim D(x) \lesssim |\Pi_{\ker A^\perp} x|^2.$$

We begin with a preliminary lemma which reduces the existence of a stationary measure to the kind of time-averaged coercivity alluded to in (1.6). The proof follows in a straightforward way from the Krylov-Bogoliubov procedure and the energy conservation property of  $B$ , however, we include it for the sake of completeness.

**Lemma 2.1.** *Let  $V(x) = \langle x \rangle^2$  and let  $\mathcal{P}_t$  be the Feller Markov semigroup on  $\mathbb{R}^n$  generated by  $\mathcal{L}$  (defined in (1.3)). If there exists  $r \in (0, 1]$ ,  $1 \leq p < (1 - r)^{-1}$ ,  $C > 0$  and  $T \in (0, 2)$  such that*

$$\frac{1}{T} \int_0^T \mathcal{P}_t V^{rp}(x) dt \leq C \left( 1 + \frac{1}{T} \int_0^T \mathcal{P}_t D^p(x) dt \right) \quad \forall x \in \mathbb{R}^n, \quad (2.1)$$

then there exists at least one stationary measure  $\mu_*$  of  $\mathcal{P}_t$  such that  $V^{rp} \in L^1(d\mu_*)$ . Moreover, there exist  $\alpha, \beta > 0$  such that the function

$$\tilde{V}(x) = \frac{1}{T} \int_0^T \mathcal{P}_t V^p(x) dt$$

satisfies

$$\mathcal{L}\tilde{V} \leq -\alpha\tilde{V}^r + \beta. \quad (2.2)$$

That is,  $\tilde{V}$  is a sub-geometric (when  $r < 1$ ) Lyapunov function.

*Proof.* Let  $\mu$  be any Borel measure on  $\mathbb{R}^n$  with  $\int V^p(x)\mu(dx) < \infty$  and let  $T$  be given as in the assumption. We first claim that for every  $n \in \mathbb{N}$  there holds

$$\int_0^{nT} \int V^{rp}(x) \mathcal{P}_s^* \mu(dx) ds \leq C \left( nT + \int_0^{nT} \int D^p(x) \mathcal{P}_s^* \mu(dx) ds \right). \quad (2.3)$$

By the assumption (2.1) and Fubini's theorem, we have

$$\int_0^T \int V^{rp}(x) \mathcal{P}_s^* \mu(dx) ds = \int \left( \int_0^T \mathcal{P}_s V^{rp}(x) ds \right) \mu(dx) \leq C \left( T + \int \left( \int_0^T \mathcal{P}_s D^p(x) ds \right) \mu(dx) \right),$$

and hence

$$\int_0^T \int V^{rp}(x) \mathcal{P}_s^* \mu(dx) ds \leq C \left( T + \int_0^T \int D^p(x) \mathcal{P}_s^* \mu(dx) ds \right). \quad (2.4)$$

By the semigroup property and (2.4), for any  $m \in \mathbb{N}$  we have

$$\begin{aligned} \int_{mT}^{(m+1)T} \int V^{rp}(x) \mathcal{P}_s^* \mu(dx) ds &\leq C \left( T + \int_0^T \int D^p(x) \mathcal{P}_s^* (\mathcal{P}_{mT}^* \mu)(dx) ds \right) \\ &= C \left( T + \int_{mT}^{(m+1)T} \int D^p(x) \mathcal{P}_s^* \mu(dx) ds \right). \end{aligned} \quad (2.5)$$

Summing (2.5) over  $0 \leq m < n$  yields (2.3).

Next, notice a direct computation using  $B(x, x) \cdot x = 0$  shows that

$$\mathcal{L}V^p(x) \leq C_1 V(x)^{p-1} - C_2 D^p(x) \quad (2.6)$$

for some constants  $C_1, C_2 > 0$ . Thus,

$$\frac{d}{dt} \int V^p(x) \mathcal{P}_t^* \mu(dx) \leq C_1 \int V(x)^{p-1} \mathcal{P}_t^* \mu(dx) - C_2 \int D^p(x) \mathcal{P}_t^* \mu(dx). \quad (2.7)$$

Let  $n \in \mathbb{N}$ . Integrating the previous inequality over  $0 \leq t \leq nT$  we see that

$$\int_0^{nT} \int D^p(x) \mathcal{P}_s^* \mu(dx) ds \lesssim 1 + \int_0^{nT} \int V(x)^{p-1} \mathcal{P}_s^* \mu(dx) ds, \quad (2.8)$$

where the implicit constant depends on  $\int V^p(x)\mu(dx)$ . Applying (2.3) and (2.8) gives

$$\int_0^{nT} \int V^{rp}(x) \mathcal{P}_s^* \mu(dx) ds \leq C(1 + nT) + C \int_0^{nT} \int V(x)^{p-1} \mathcal{P}_s^* \mu(dx) ds. \quad (2.9)$$

The choice  $p < (1 - r)^{-1}$  ensures that  $p - 1 < rp$ , and so for every  $\epsilon > 0$  there is  $C_\epsilon$  such that

$$V^{p-1} \leq \epsilon V^{rp} + C_\epsilon.$$

Hence, the integral on the right-hand side of (2.9) can be absorbed into the left-hand side, yielding

$$\int_0^{nT} \int V^{rp}(x) \mathcal{P}_s^* \mu(dx) ds \lesssim (1 + nT).$$

Therefore

$$\sup_{n \in \mathbb{N}} \frac{1}{nT} \int_0^{nT} \int V^{rp}(x) \mathcal{P}_s^* \mu(dx) ds < \infty. \quad (2.10)$$

Using the tightness implied by (2.10), the existence of a stationary measure  $\mu_*$  with  $V^{rp} \in L^1(d\mu_*)$  follows by the usual Krylov-Bogoliubov method (see e.g. [11]).

It remain to prove (2.2). First, by (2.6), (2.1), and  $p < (1 - r)^{-1}$  there exist constants  $c, C > 0$  such that

$$\mathcal{L}\tilde{V} \leq -\frac{c}{T} \int_0^T \mathcal{P}_t V^{rp} dt + C. \quad (2.11)$$

Next, one can show using Grönwall's lemma that for any  $q \geq 1$  there exists  $C_q \geq 1$  such that

$$\langle x \rangle^{2q} - C_q t (\langle x \rangle^{2q} + 1) \leq \mathcal{P}_t V^q(x) \leq C_q (\langle x \rangle^{2q} + 1) \quad \forall t \in [0, 2].$$

It follows that for all  $x$  with  $|x|$  sufficiently large and  $T \in (0, 2)$  there holds

$$\frac{1}{T} \int_0^T \mathcal{P}_t V^q(x) dt \approx_q \langle x \rangle^{2q}. \quad (2.12)$$

Using (2.12) in (2.11) completes the proof.  $\square$

## 2.2 Short-time coercivity near $\ker A$

Next, we formulate a sufficient condition for (2.1) based on short-time (time-averaged) coercivity of solutions near  $\ker A$ . Intuitively, this is similar to estimating average exit times from the vicinity of  $\ker A$ , but not quite the same.

**Assumption 1.** Let  $\mathcal{P}_t$  be the Feller Markov semigroup on  $\mathbb{R}^n$  generated by  $\mathcal{L}$  defined in (1.3). We say that  $\mathcal{P}_t$  satisfies Assumption 1 if there exist  $r \in (0, 1]$ ,  $K_* \geq 1$ ,  $c_* > 0$ ,  $\delta \in (0, 1)$ , and a finite collection of times  $\eta_j : [0, \infty) \rightarrow (0, 1]$ ,  $1 \leq j \leq m$ , such that:

- $\lim_{K \rightarrow \infty} \sup_{1 \leq j \leq m} \eta_j(K) = 0$ ;
- for every  $K \geq K_*$  the set

$$B_K = \{x \in \mathbb{R}^n : |\Pi_{\ker A^\perp} x|^2 \leq \delta |\Pi_{\ker A} x|^{2r} \text{ and } (1 - \delta)K^2 \leq |x|^2 \leq (1 + \delta)K^2\}$$

admits a decomposition

$$B_K = \bigcup_{j=1}^m B_{K,j}$$

for which  $B_{K,j} \subseteq \mathbb{R}^n$  is such that  $x \in B_{K,j}$  implies

$$\frac{1}{\eta_j(K)} \int_0^{\eta_j(K)} \mathcal{P}_t D(x) dt \geq c_* K^{2r}. \quad (2.13)$$

**Remark 8.** By the Hölder and Jensen inequalities, (2.13) implies that for any  $p \geq 1$  and  $x \in B_{K,j}$  there holds

$$\frac{1}{\eta_j(K)} \int_0^{\eta_j(K)} \mathcal{P}_t D^p(x) dt \geq c_*^p K^{2rp}. \quad (2.14)$$

The main result of this section is the following lemma, which shows that Assumption 1 implies (2.1), which is a sufficient condition for the existence of an invariant measure due to Lemma 2.1.

**Lemma 2.2.** *Suppose that Assumption 1 holds for some  $r \in (0, 1]$ . Then, there exists at least one stationary measure  $\mu_*$  of  $\mathcal{P}_t$  and  $\langle x \rangle^q \in L^1(d\mu_*)$  for every  $q < 2r/(1-r)$ . Moreover, for every  $1 \leq p < (1-r)^{-1}$  the function  $\tilde{V}$  defined in Lemma 2.1 satisfies (2.2).*

Before proceeding to the details of the proof we give a few remarks on the intuition behind Lemma 2.2. As we will see at the beginning of the proof below, the lemma reduces to showing that there are  $T \in (0, 2)$  and  $C > 0$  so that for every  $x \in \mathbb{R}^n$  with  $|x| = K \gg 1$  there holds

$$K^{2rp} \leq \frac{C}{T} \int_0^T \mathcal{P}_t D^p(x) dt. \quad (2.15)$$

The idea behind proving (2.15) is to first note that if the process ever enters the set

$$\tilde{B} = \{z \in \mathbb{R}^n : |\Pi_{\ker A^\perp} z|^2 \leq \delta |\Pi_{\ker A} z|^{2r}\}, \quad (2.16)$$

then, provided  $T$  is small, by Lemma A.1 it is in  $B_K$  with high probability. Hence one can essentially assume that  $x_t(\omega) \in B_K$  whenever  $x_t(\omega) \in \tilde{B}$ . Now, the time average of  $D(x_t)^p$  controls  $K^{2rp}$  when the process is not in  $\tilde{B}$  and, by the discussion above and Assumption 1, with high probability it controls  $K^{2rp}$  on some short time interval if the process ever does enter  $\tilde{B}$ . By tracking the return times of the process to  $\tilde{B}$ , (2.15) follows by a suitable iteration of Assumption 1.

We now give the proof of Lemma 2.2.

*Proof of Lemma 2.2.* First notice that by Lemma 2.1 and (2.12) we just need to show that for every  $1 \leq p < (1-r)^{-1}$  there exists  $C > 0$  and  $T \in (0, 2)$  so that for every  $x \in \mathbb{R}^n$

$$\langle x \rangle^{2rp} = V^{rp}(x) \leq C \left( 1 + \frac{1}{T} \int_0^T \mathcal{P}_t D^p(x) dt \right). \quad (2.17)$$

We now set out to prove (2.17). Let  $\delta$ ,  $c_*$ , and  $K_*$  be as in Assumption 1 and let  $x_t$  denote the solution to (1.1) with  $x_0 = x$ . The bound (2.17) is trivial if  $|x| < K_*$  because we can take  $C$  sufficiently large depending on  $K_*$ , and so we need only consider when  $|x| := K \geq K_*$ . In this case we make precise the intuition described directly after the statement of the lemma. Let  $\tilde{B}$  be as in (2.16) and define the sequence of stopping times  $\tau_0(\omega) = 0$ ,  $\tau_1(\omega) = \inf\{t \geq 0 : x_t(\omega) \in \tilde{B}\} \wedge T$ , and for  $n \geq 1$

$$\tau_{n+1}(\omega) = \begin{cases} \tau_n(\omega) & \text{if } x_{\tau_n(\omega)}(\omega) \notin B_K \\ \inf\{t \geq \tau_n(\omega) + \eta_1(K) : x_t(\omega) \in \tilde{B}\} \wedge T & \text{if } x_{\tau_n(\omega)}(\omega) \in B_{K,1} \\ \vdots & \\ \inf\{t \geq \tau_n(\omega) + \eta_m(K) : x_t(\omega) \in \tilde{B}\} \wedge T & \text{if } x_{\tau_n(\omega)}(\omega) \in B_{K,m}. \end{cases}$$

Moreover, define

$$\bar{\tau}(\omega) = \inf\{t \geq 0 : ||x_t(\omega)|^2 - K^2| \geq \delta K^2\}.$$

Due to Lemma A.1 applied with  $\epsilon = \delta$ , by taking  $T$  sufficiently small and  $K_*$  sufficiently large (both depending only on  $\delta$ ) we may assume that

$$\mathbf{P}(\bar{\tau} \geq T) \geq 1/2. \quad (2.18)$$

With  $T$  fixed, choosing  $K_*$  perhaps even larger and recalling  $\lim_{K' \rightarrow \infty} \sup_{1 \leq j \leq m} \eta_j(K') = 0$  implies that we may assume

$$\sup_{1 \leq j \leq m} \eta_j(K) \leq T/2. \quad (2.19)$$

Let now

$$A_n = \{\omega \in \Omega : \tau_n(\omega) \leq T/2\}$$

and

$$A_{n,j} = \{\omega \in \Omega : x_{\tau_n(\omega)}(\omega) \in B_{K,j}\}.$$

Using that  $\tau_n$  is an increasing sequence with  $\lim_{n \rightarrow \infty} \tau_n(\omega) \leq T$  we have

$$\begin{aligned} \int_0^T \mathcal{P}_t D^p(x) dt &\geq \int_{\Omega} \sum_{n=0}^{\infty} \int_{\tau_n(\omega)}^{\tau_{n+1}(\omega)} D^p(x_t(\omega)) dt d\mathbf{P} \\ &= \int_{\Omega} \sum_{n=0}^{\infty} \int_0^{\tau_{n+1}(\omega) - \tau_n(\omega)} D^p(x_{\tau_n+t}(\omega)) dt d\mathbf{P} \\ &\geq \sum_{n=0}^{\infty} \sum_{j=1}^m \int_{A_{n,j} \cap A_n} \int_0^{\tau_{n+1}(\omega) - \tau_n(\omega)} D^p(x_{\tau_n+t}(\omega)) dt d\mathbf{P}. \end{aligned} \quad (2.20)$$

Now, if  $x_{\tau_n(\omega)}(\omega) \in B_{K,j}$  and  $\tau_n(\omega) \leq T/2$ , then  $\tau_{n+1}(\omega) - \tau_n(\omega) \geq \eta_j(K)$  due to (2.19) and the definition of  $\tau_{n+1}$ . For  $1 \leq j \leq m$  we thus have

$$\begin{aligned} &\int_{A_{n,j} \cap A_n} \int_0^{\tau_{n+1}(\omega) - \tau_n(\omega)} D^p(x_{\tau_n+t}(\omega)) dt d\mathbf{P} \\ &= \int_{A_{n,j} \cap A_n} \int_0^{\eta_j(K)} D^p(x_{\tau_n+t}(\omega)) dt d\mathbf{P} + \int_{A_{n,j} \cap A_n} \int_{\eta_j(K)}^{\tau_{n+1}(\omega) - \tau_n(\omega)} D^p(x_{\tau_n+t}(\omega)) dt d\mathbf{P}. \end{aligned} \quad (2.21)$$

For the first piece, observe that  $A_{n,j} \cap A_n$  is measurable with respect to  $\mathcal{F}_{\tau_n}$ , the  $\sigma$ -algebra of events determined prior to the stopping time  $\tau_n$  (this agrees with the  $\sigma$ -algebra generated by  $\{W_{s \wedge \tau_n}; s \geq 0\}$ ). Thus,

$$\int_{A_{n,j} \cap A_n} \int_0^{\eta_j(K)} D^p(x_{\tau_n+t}(\omega)) dt d\mathbf{P} = \int_0^{\eta_j(K)} \int_{A_{n,j} \cap A_n} \mathbf{E}(D^p(x_{\tau_n+t}) | \mathcal{F}_{\tau_n}) d\mathbf{P} dt.$$

It then follows from the strong Markov property and Assumption 1 in the form (2.14) that

$$\int_{A_{n,j} \cap A_n} \int_0^{\eta_j(K)} D^p(x_{\tau_n+t}(\omega)) dt d\mathbf{P} = \int_{A_{n,j} \cap A_n} \int_0^{\eta_j(K)} \mathcal{P}_t D^p(x_{\tau_n}) dt d\mathbf{P} \geq \eta_j(K) c_*^p K^{2rp} \mathbf{P}(A_{n,j} \cap A_n). \quad (2.22)$$

For the second piece, first note that

$$\int_{A_{n,j} \cap A_n} \int_{\eta_j(K)}^{\tau_{n+1}(\omega) - \tau_n(\omega)} D^p(x_{\tau_n+t}(\omega)) dt d\mathbf{P} \geq \int_{A_{n,j} \cap A_n \cap \{\bar{\tau} \geq T\}} \int_{\eta_j(K)}^{\tau_{n+1}(\omega) - \tau_n(\omega)} D^p(x_{\tau_n+t}(\omega)) dt d\mathbf{P}.$$

Now, by construction, if  $\omega \in A_{n,j} \cap A_n \cap \{\bar{\tau} \geq T\}$  then for each  $t \in (\eta_j(K), \tau_{n+1}(\omega) - \tau_n(\omega))$  one has

$$|\Pi_{\ker A^\perp} x_{\tau_n+t}(\omega)|^2 \geq \delta |\Pi_{\ker A} x_{\tau_n+t}(\omega)|^{2r} \quad \text{and} \quad (1 - \delta) K^2 \leq |x_{\tau_n+t}(\omega)|^2 \leq (1 + \delta) K^2.$$

Thus, there is a constant  $c_0 \in (0, 1)$  so that, for any  $c \in (0, c_0)$ , over the same time interval there holds

$$D^p(x_{\tau_n+t}(\omega)) \geq c \delta^p K^{2rp}.$$

Consequently,

$$\begin{aligned} & \int_{A_{n,j} \cap A_n} \int_{\eta_j(K)}^{\tau_{n+1}(\omega) - \tau_n(\omega)} D^p(x_{\tau_n+t}(\omega)) dt d\mathbf{P} \\ & \geq c\delta^p K^{2rp} \int_{A_{n,j} \cap A_n \cap \{\bar{\tau} \geq T\}} (\tau_{n+1}(\omega) - \tau_n(\omega)) d\mathbf{P} - c\eta_j(K) \delta^p K^{2rp} \mathbf{P}(A_{n,j} \cap A_n). \end{aligned} \quad (2.23)$$

Choosing  $c < \min(c_0, c_*^p)$  and then putting (2.22) and (2.23) into (2.21) we find

$$\int_{A_{n,j} \cap A_n} \int_0^{\tau_{n+1}(\omega) - \tau_n(\omega)} D^p(x_{\tau_n+t}(\omega)) dt d\mathbf{P} \geq c\delta^p K^{2rp} \int_{A_{n,j} \cap A_n \cap \{\bar{\tau} \geq T\}} (\tau_{n+1}(\omega) - \tau_n(\omega)) d\mathbf{P}.$$

Using this bound in (2.20) and noting that if  $\bar{\tau}(\omega) \geq T$  and  $\tau_n(\omega) \leq T/2$  then  $x_{\tau_n(\omega)} \in B_K$  gives

$$\begin{aligned} \int_0^T \mathcal{P}_t D^p(x) dt & \geq c\delta^p K^{2rp} \sum_{n=0}^{\infty} \sum_{j=1}^m \int_{A_{n,j} \cap A_n \cap \{\bar{\tau} \geq T\}} (\tau_{n+1}(\omega) - \tau_n(\omega)) d\mathbf{P} \\ & = c\delta^p K^{2rp} \sum_{n=0}^{\infty} \int_{\{\bar{\tau} \geq T\}} \mathbf{1}_{\{\tau_n \leq T/2\}} (\tau_{n+1}(\omega) - \tau_n(\omega)) d\mathbf{P}. \end{aligned}$$

By the telescoping summation and the definition of  $\tau_n$ , we have

$$\sum_{n=0}^{\infty} \mathbf{1}_{\{\tau_n \leq T/2\}} (\tau_{n+1}(\omega) - \tau_n(\omega)) \geq T/2$$

whenever  $\bar{\tau}(\omega) \geq T$ , and hence we conclude

$$\frac{1}{T} \int_0^T \mathcal{P}_t D^p(x) dt \geq \frac{c}{2} \delta^p K^{2rp} \mathbf{P}(\bar{\tau} \geq T) \geq \frac{c}{4} \delta^p K^{2rp},$$

which completes the proof.  $\square$

### 3 Conservative flow transverse to the kernel

Theorem 1.1 is an immediate consequence of Lemma 2.2 together with the following proposition. The proof consists of two main steps: the first is to deduce growth of the damped modes for a suitable approximate solution (in this case, the deterministic, conservative dynamics) and the second is to justify the approximation on a long enough time-scale to verify Assumption 1 for the true solution.

**Proposition 3.1.** *Let  $B$  and  $A$  satisfy the conditions of Theorem 1.1. Then, Assumption 1 holds for  $r = 1$ .*

*Proof.* Let  $x_0 \in \mathbb{R}^n$  with  $K/2 \leq |x_0| \leq 2K$  and  $|\Pi_{\ker A^\perp} x_0| \leq \delta |\Pi_{\ker A} x_0|$  for some  $K \geq 1$  and  $\delta \in (0, 1)$ . Note that the assumptions on  $x_0$  imply that

$$|\Pi_{\ker A} x_0| \geq \frac{K}{\sqrt{8}}. \quad (3.1)$$

Let  $X_t$  solve

$$\begin{cases} \frac{d}{dt} X_t = B(X_t, X_t) \\ X_t|_{t=0} = x_0 \end{cases} \quad (3.2)$$

and  $\tilde{X}_t$  solve

$$\begin{cases} \frac{d}{dt}\tilde{X}_t = B(\tilde{X}_t, \tilde{X}_t) \\ \tilde{X}_t|_{t=0} = \Pi_{\ker A}x_0. \end{cases} \quad (3.3)$$

By taking successive time derivatives of (3.3), we see that

$$\frac{d^j}{dt^j}\tilde{X}_t|_{t=0}$$

is a homogeneous  $j + 1$  degree polynomial in  $\Pi_{\ker A}x_0$ . Therefore, (3.1) and the condition (1.8) imply that  $\exists C_J \geq 1$  and  $j \leq J$  such that

$$\left| \Pi_{\ker A^\perp} \frac{d^j}{dt^j} \tilde{X}_t|_{t=0} \right| \geq \frac{2}{C_J} K^{j+1}.$$

It follows that for  $\delta$  sufficiently small there holds

$$\left| \Pi_{\ker A^\perp} \frac{d^j}{dt^j} X_t|_{t=0} \right| \geq \frac{1}{C_J} K^{j+1}. \quad (3.4)$$

**Step 1** (growth for the deterministic dynamics): The first step quantifies how condition (1.8) implies growth of  $\Pi_{\ker A^\perp} X_t$  for short times.

**Lemma 3.2.** *Let  $\delta \in (0, 1)$  be small enough so that (3.4) holds and fix  $\tau = K^{-1}$ . There is a constant  $c_* > 0$  depending on  $C_J$  and  $J$  such that*

$$\frac{1}{\tau} \int_0^\tau |\Pi_{\ker A^\perp} X_t| dt \geq c_* K. \quad (3.5)$$

*Proof.* We first claim that there exists  $\gamma_0 \in (0, 1)$  so that if for some  $1 \leq \ell \leq J$  and  $\gamma \in (0, \gamma_0)$  there holds

$$\left| \Pi_{\ker A^\perp} \frac{d^\ell}{dt^\ell} X_t|_{t=0} \right| \geq \frac{\gamma}{C_J} K^{\ell+1}, \quad (3.6)$$

then there is  $t_0 \in \{0, 3\gamma^2 K^{-1}\}$  such that

$$\left| \Pi_{\ker A^\perp} \frac{d^{\ell-1}}{dt^{\ell-1}} X_t|_{t=t_0} \right| \geq \frac{\gamma^3}{C_J} K^\ell. \quad (3.7)$$

The claim is trivial if the desired bound holds for  $t_0 = 0$ . So, suppose otherwise and expand to first order to obtain

$$\left| \Pi_{\ker A^\perp} \frac{d^{\ell-1}}{dt^{\ell-1}} X_t \right| \geq \frac{\gamma}{C_J} K^{\ell+1} t - \frac{\gamma^3}{C_J} K^\ell - C t^2 K^{\ell+2}$$

for some constant  $C \geq 1$  that does not depend on  $\ell$  (it depends on  $J$  and the size of  $B$  on the unit sphere). Thus, for  $t_0 = 3\gamma^2 K^{-1}$  we have

$$\left| \Pi_{\ker A^\perp} \frac{d^{\ell-1}}{dt^{\ell-1}} X_t|_{t=t_0} \right| \geq \frac{2\gamma^3}{C_J} K^\ell - 9C\gamma^4 K^\ell.$$

The bound (3.7) then follows for  $\gamma_0 < (9CC_J)^{-1}$ .

Using (3.4) and iterating the claim we just proved, it is straightforward to show that if

$$\gamma < \min((9CC_J)^{-1}, 1/\sqrt{2}),$$

then there exists  $t_0 \in [0, 6\gamma^2 K^{-1}]$  such that

$$|\Pi_{\ker A^\perp} X_{t_0}| \geq \frac{\gamma^{3^J}}{C_J} K.$$

Taylor expanding to first order at  $t = t_0$  then gives

$$|\Pi_{\ker A^\perp} X_{t_0+t}| \geq \frac{\gamma^{3^J}}{C_J} K - tCK^2$$

and hence for  $\tau_1 = (1/2)\gamma^{3^J} (KCC_J)^{-1}$  we have

$$\frac{1}{\tau_1} \int_0^{\tau_1} |\Pi_{\ker A^\perp} X_{t_0+t}| dt \geq \frac{\gamma^{3^J}}{2C_J} K.$$

Supposing that  $\gamma$  is small enough so that  $\tau_1 + t_0 \leq K^{-1}$ , it follows that for  $\tau = K^{-1}$  there holds

$$\frac{1}{\tau} \int_0^\tau |\Pi_{\ker A^\perp} X_t| dt \geq \frac{\gamma^{3^J}}{2C_J} K \frac{\tau_1}{\tau} \geq \frac{\gamma^{2 \cdot 3^J}}{4CC_J^2} K,$$

which completes the proof.  $\square$

**Step 2** (approximating with  $X_t$ ): In order to make use of Lemma 3.2 we need to show that  $X_t$  is a sufficiently good approximation of  $x_t$  for  $t \lesssim K^{-1}$ . To this end, we have the following lemma.

**Lemma 3.3.** *Let  $x_0$  and  $K$  be as defined at the beginning of proof and set  $\tau = K^{-1}$ . With  $X_t$  given by (3.2) and  $x_t$  given by (1.1), there are  $K_* \geq 1$  and  $C > 0$  (both independent of  $x_0$ ) so that for  $K \geq K_*$  there holds*

$$\mathbf{P} \left( \sup_{0 \leq t \leq \tau} |X_t - x_t| \leq C \right) \geq \frac{1}{2}.$$

*Proof.* The error  $X_t - x_t$  solves

$$d(X_t - x_t) = L_{X_t}(X_t - x_t)dt + AX_t dt - A(X_t - x_t)dt - B(X_t - x_t, X_t - x_t)dt - \sigma dW_t,$$

where the operator  $L_x$  is as defined in (1.9). In what follows, denote by  $S_X(t, s)$  the two-time linear propagator of the time-inhomogeneous ODE associated to  $L_{X_t}$ , i.e.

$$\begin{aligned} \frac{d}{dt} S_X(t, s)v &= L_{X_t} S_X(t, s)v \\ S_X(s, s)v &= v. \end{aligned}$$

Since  $|X_t| \leq 2K$  for all  $t$  there is a constant  $C_1$  that does not depend on  $K$  such that for  $t > s$ ,

$$\|S_X(t, s)\| \leq e^{C_1 K(t-s)}. \quad (3.8)$$

Therefore, there is  $C_2 > 0$  independent of  $K$  so that for  $T \leq \tau = K^{-1}$  we have

$$\begin{aligned} \sup_{0 \leq t \leq T} |X_t - x_t| &\lesssim \int_0^t e^{CK(t-s)} (|X_s - x_s|^2 + |X_s| + |X_s - x_s|) ds + \sup_{0 \leq t \leq \tau} \left| \int_0^t S_X(t, s) \sigma dW_s \right| \\ &\leq C_2 \left( K^{-1} \sup_{0 \leq t \leq T} |X_t - x_t|^2 + 1 + \sup_{0 \leq t \leq T} \left| \int_0^t S_X(t, s) \sigma dW_s \right| \right). \end{aligned} \quad (3.9)$$



By the Doob martingale inequality, the Itô isometry, and (3.8), there is  $C_3 > 0$  depending on  $\sigma$  but independent of  $K$  and so that for all  $R \geq 1$  there holds

$$\mathbf{P} \left( \sup_{0 \leq t \leq \tau} \left| \int_0^t S_X(t, s) \sigma dW_s \right| \geq R/\sqrt{K} \right) \lesssim \frac{K}{R^2} \int_0^\tau e^{2K(\tau-s)} \|\sigma\|^2 ds \leq \frac{C_3}{R^2}.$$

Therefore, there is  $R_* \geq 1$  independent of  $K$  so that  $\mathbf{P}(\Omega_0) \geq 1/2$  if we define  $\Omega_0$  as

$$\Omega_0 = \left\{ \omega \in \Omega : \sup_{0 \leq t \leq \tau} \left| \int_0^t S_X(t, s) \sigma dW_s \right| \leq R_*/\sqrt{K} \right\}.$$

Fix  $\omega \in \Omega_0$  and let  $T_\omega$  be the maximal time such that  $\sup_{0 \leq t \leq T_\omega} |X_t(\omega) - x_t(\omega)| \leq 2C_2$ , where  $C_2$  is as in (3.9). Since  $x_t$  and  $X_t$  take values continuously in time,  $T_\omega > 0$ . Moreover, by (3.9) and the definition of  $\Omega_0$ , for  $T \leq \min(T_\omega, \tau)$  we have

$$\sup_{0 \leq t \leq T} |X_t - x_t| \leq \frac{4C_2^3}{K} + C_2 + \frac{C_2 R_*}{\sqrt{K}}. \quad (3.10)$$

Thus,  $T_\omega \geq \tau$  as soon as  $K \geq \max(4R_*^2, 8C_2^2)$ . This completes the proof.  $\square$

With Lemma 3.2 and Lemma 3.3 in hand, the proof of Proposition 3.1 follows quickly. Indeed, let  $x_0$ ,  $\delta > 0$ , and  $K \geq 1$  be as defined at the beginning of the proof and  $K_*$  be as in Lemma 3.3. We need to show that there is  $c > 0$  so that for all  $K \geq K_*$  there holds

$$\frac{1}{\tau} \int_0^\tau \mathbf{E} |\Pi_{\ker A^\perp} x_t|^2 dt \geq cK^2,$$

where as before  $\tau = K^{-1}$ . First, by Chebyshev's inequality we have for all  $K \geq 1$  and  $\delta' > 0$ ,

$$\delta' K^2 \mathbf{P} \left( \frac{1}{\tau} \int_0^\tau |\Pi_{\ker A^\perp} x_t|^2 dt \geq \delta' K^2 \right) \leq \frac{1}{\tau} \int_0^\tau \mathbf{E} |\Pi_{\ker A^\perp} x_t|^2 dt.$$

Let  $\Omega_1$  be the set such that

$$\sup_{0 \leq t \leq \tau} |X_t - x_t| \leq C,$$

where  $C$  is as in Lemma 3.3, which implies  $\mathbf{P}(\Omega_1) > 1/2$  (note that while  $\Omega_1$  can depend on  $K$  and  $x_0$ , the associated estimates do not). By Lemma 3.2, for  $\omega \in \Omega_1$  we have

$$\int_0^\tau |\Pi_{\ker A^\perp} x_t|^2 dt \geq \frac{c_*^2}{2} K^2 \tau - C\tau.$$

Therefore, assuming also  $K_* \geq 2\sqrt{C}/c_*$ , for  $\delta' < c_*^2/4$  and  $K \geq K_*$  we have

$$\frac{1}{2} \delta' K^2 \leq \frac{1}{\tau} \int_0^\tau \mathbf{E} |\Pi_{\ker A^\perp} x_t|^2 dt,$$

which implies Assumption 1.  $\square$

## 4 Unstable equilibria in the kernel

In this section, we consider the case where  $\mathcal{U} = \ker(A) \cap \mathbb{S}^{n-1}$  consists entirely of unstable equilibria of the conservative dynamics (i.e.,  $B(x, x) = 0$  for every  $x \in \ker A$ ) and prove Theorems 1.2 and 1.3. As in Section 3, the proofs are based on a two-step procedure that consists of first deducing growth of the damped modes for a suitable approximate solution and second justifying the approximation on a long enough time-scale to verify Assumption 1. In the present setting, for an initial condition  $x \in B_K$  (where  $B_K$  is as in the statement of Assumption 1), the approximation of the damped modes that we consider is obtained simply by linearizing  $B$  around the equilibrium  $\Pi_{\ker A} x$ .

Recall that for  $x \in \mathbb{R}^n$  we define  $L_x : \mathbb{R}^n \rightarrow \mathbb{R}^n$  by  $L_x v = B(x, v) + B(v, x)$  and that we denote  $L_x^\perp = \Pi_{\ker A^\perp} L_x \Pi_{\ker A^\perp}$ . In the setting of Theorems 1.2 and 1.3, for any  $x \in \mathcal{U}$ ,  $L_x^\perp$  either has an eigenvalue  $\lambda$  with  $\operatorname{Re}(\lambda) > 0$  or an unstable Jordan block corresponding to  $\operatorname{Re}(\lambda) = 0$ . In studying the properties of linearized solutions we must consider separately these two scenarios. In Section 4.1 we prove the necessary growth and approximation estimates in the spectrally unstable case, and in Section 4.2 we treat the Jordan block unstable case. In Section 4.3 we use the results of Sections 4.1 and 4.2 to complete the proofs of Theorems 1.2 and 1.3.

### 4.1 Spectrally unstable estimates

We begin by considering the case where for  $z \in \mathcal{U}$ ,  $L_z^\perp$  is spectrally unstable. In this setting, the result concerning growth of the damped modes for the linear approximation is given as follows. Recall from Section 1 that we denote the Jordan normal form of  $L_z^\perp$  by

$$L_z^\perp = P_z J_z^\perp P_z^{-1}.$$

**Lemma 4.1.** *Let  $z \in \ker A$  and  $r \in (0, 1)$ . Assume that the eigenvalue  $\lambda = \lambda_R + i\lambda_I$  of  $L_{z/|z|}^\perp$  with largest real part is such that  $\lambda_R > 0$ . Suppose further that there exists a generalized eigenvector  $v = v_R + iv_I$  corresponding to eigenvalue  $\lambda$  such that  $\operatorname{Ran}(\sigma) \cap \{v_R, v_I\} \setminus \{0\} \neq \emptyset$  and, defining  $V = \operatorname{span}\{P_{z/|z|}^{-1} v\}$ , there holds*

$$\Pi_V J_{z/|z|}^\perp x = \lambda \Pi_V x \quad \forall x \in \mathbb{C}^n. \quad (4.1)$$

Let  $Y_t : [0, \infty) \rightarrow \ker A^\perp$  solve

$$\begin{cases} dY_t = L_z^\perp Y_t dt + \Pi_{\ker A^\perp} \sigma dW_t \\ Y_t|_{t=0} = Y_0 \in \ker A^\perp. \end{cases} \quad (4.2)$$

For any  $\epsilon \in (0, 1)$ , there is  $K_*(\epsilon) \geq 1$  and constants  $c_*, \beta > 0$  that do not depend on  $\epsilon$ ,  $r$ , or  $|z|$  so that for  $|z| \geq K_*$  and

$$\tau = \frac{(1/2 + r + \epsilon) \log(|z|)}{\lambda_R |z|}$$

there holds

$$\mathbf{P} \left( \frac{1}{\tau} \int_0^\tau |Y_t(\omega)| dt \geq c_* |z|^r \right) \geq \beta. \quad (4.3)$$

**Remark 9.** The condition (4.1) just says that  $v$  is the first generalized eigenvector in a Jordan chain corresponding to eigenvalue  $\lambda$ .

**Remark 10.** It follows directly from the proof below that if  $\sigma$  is invertible, then for any  $C > 0$  the constants  $c_*$  and  $\beta$  can be chosen uniformly for  $\lambda_R \geq C^{-1}$  and  $\|P_{z/|z|}\| + \|P_{z/|z|}^{-1}\| \leq C$ . The main observation here is that when  $\sigma$  is invertible the constant  $c_1$  in the proof below depends only on  $\|\sigma^{-1}\|$  and  $\|P_{z/|z|}\|$ .

*Proof.* We will consider the case where  $\lambda_I \neq 0$ , as the situation where  $\lambda_I = 0$  follows from the same argument. Without loss of generality, we may suppose that  $v_R \in \text{Ran}(\sigma)$ . Let  $\tilde{Y}_t = P_{z/|z|}^{-1} Y_t \in \mathbb{C}^n$ . For simplicity of notation we write  $J^\perp$  for  $J_{z/|z|}^\perp$  and  $P$  for  $P_{z/|z|}$ . Then,  $\tilde{Y}_t$  solves

$$\begin{cases} d\tilde{Y}_t = |z|J^\perp \tilde{Y}_t dt + P^{-1} \Pi_{\ker A^\perp} \sigma dW_s \\ \tilde{Y}_0 = P^{-1} Y_0 \in \mathbb{C}^n. \end{cases} \quad (4.4)$$

Since

$$|\tilde{Y}_t| = |P^{-1} Y_t| \leq \|P^{-1}\|_{\mathbb{C}^n \rightarrow \mathbb{C}^n} |Y_t|, \quad (4.5)$$

it suffices to prove (4.3) with  $Y_t(\omega)$  replaced by  $\tilde{Y}_t(\omega)$ . Define the subspace  $V = \text{span}\{P^{-1}v\}$  of  $\mathbb{C}^n$  and let  $t_* = (\lambda_R|z|)^{-1}$ . The plan is to first show that for every  $R > 0$  there exists  $\beta_1 > 0$  so that

$$\mathbf{P} \left( |\Pi_V \tilde{Y}_{t_*}| \geq \frac{R}{\sqrt{\lambda_R|z|}} \right) \geq \beta_1. \quad (4.6)$$

We will then prove that there is  $R \geq 1$  and  $c_* > 0$  so that

$$\mathbf{P} \left( \frac{1}{\tau} \int_0^{\tau-t_*} |\tilde{Y}_t| dt \geq c_* |z|^r \mid |\Pi_V \tilde{Y}_0| \geq \frac{R}{\sqrt{\lambda_R|z|}} \right) \geq \frac{1}{2}. \quad (4.7)$$

Together, (4.6) and (4.7) yield the bound (4.3) for  $\tilde{Y}_t$ .

We now prove (4.6). The formula for  $\tilde{Y}_t$  reads

$$\tilde{Y}_t = e^{|z|J^\perp t} \tilde{Y}_0 + \int_0^t e^{|z|J^\perp(t-s)} P^{-1} \Pi_{\ker A^\perp} \sigma dW_s. \quad (4.8)$$

By the Itô isometry, the variance of  $\Pi_V \tilde{Y}_{t_*}$  is given by

$$\text{Var}(\Pi_V \tilde{Y}_{t_*}) = \int_0^{t_*} \left\| \Pi_V e^{|z|J^\perp(t_*-s)} P^{-1} \Pi_{\ker A^\perp} \sigma \right\|_F^2 ds, \quad (4.9)$$

where  $\|\cdot\|_F$  denotes the Frobenius norm on  $\mathbb{C}^{n \times n}$ . Observe now that for any  $t \geq 0$  there holds

$$\Pi_V e^{|z|J^\perp t} P^{-1} v_R = \frac{1}{2} e^{\lambda|z|t} P^{-1} v,$$

which gives

$$\left| \Pi_V e^{|z|J^\perp t} P^{-1} v_R \right|^2 \geq \frac{e^{2|z|\lambda_R t}}{4}. \quad (4.10)$$

Since  $v_R \in \text{Ran}(\sigma)$ , by (4.10) we have

$$\left\| \Pi_V e^{|z|J^\perp(t_*-s)} P^{-1} \Pi_{\ker A^\perp} \sigma \right\|_{\mathbb{C}^n \rightarrow \mathbb{C}^n}^2 \geq c_1 e^{2|z|\lambda_R(t_*-s)} \quad (4.11)$$

for some  $c_1 > 0$  depending only on  $\sigma$  and  $v_R$ . Thus, from (4.9) and the equivalence of norms in finite dimensions there holds

$$\text{Var}(\Pi_V \tilde{Y}_{t_*}) \gtrsim \int_0^{t_*} \left\| \Pi_V e^{|z|J^\perp(t_*-s)} P^{-1} \Pi_{\ker A^\perp} \sigma \right\|_{\mathbb{C}^n \rightarrow \mathbb{C}^n}^2 ds \gtrsim c_1 \int_0^{t_*} e^{2|z|\lambda_R(t_*-s)} ds \gtrsim \frac{c_1}{\lambda_R|z|}. \quad (4.12)$$

The claim (4.6) then follows from (4.12) and the fact that the real and imaginary parts of  $\Pi_V \tilde{Y}_{t_*}$  are both Gaussian.

We now turn to (4.7). First, note that

$$\left| \Pi_V e^{|z|J^\perp t} \tilde{Y}_0 \right|^2 = \left| e^{|z|J^\perp t} \Pi_V \tilde{Y}_0 \right|^2 = e^{2|z|\lambda_R t} \left| \Pi_V \tilde{Y}_0 \right|^2.$$

Therefore,  $|\Pi_V \tilde{Y}_0| \geq R/\sqrt{\lambda_R|z|}$  implies that

$$\frac{1}{\tau} \int_0^{\tau-t_*} |\Pi_V e^{|z|J^\perp t} \tilde{Y}_0| dt \geq \frac{R}{\tau \sqrt{\lambda_R|z|}} \int_0^{\tau-t_*} e^{|z|\lambda_R t} dt \geq \frac{R|z|^{r+\epsilon}}{6e \log(|z|) \sqrt{\lambda_R}}, \quad (4.13)$$

where in the second inequality we have assumed that  $K_* \geq 2e$ . Taking  $K_*(\epsilon)$  even larger to ensure that  $|z|^\epsilon \geq \log(|z|)$ , it follows from (4.13) and

$$\frac{1}{\tau} \int_0^{\tau-t_*} |\tilde{Y}_t| dt \geq \frac{1}{\tau} \int_0^{\tau-t_*} |\Pi_V e^{|z|J^\perp t} \tilde{Y}_0| dt - \frac{1}{\tau} \int_0^\tau \left| \int_0^t \Pi_V e^{|z|J^\perp(t-s)} P^{-1} \Pi_{\ker A^\perp} \sigma dW_s \right| dt,$$

that to complete the proof of (4.7) it suffices to show that

$$\mathbf{P} \left( \frac{1}{\tau} \int_0^\tau \left| \int_0^t \Pi_V e^{|z|J^\perp(t-s)} P^{-1} \Pi_{\ker A^\perp} \sigma dW_s \right| dt \leq \frac{R|z|^{r+\epsilon}}{12e \log(|z|) \sqrt{\lambda_R}} \right) \geq \frac{1}{2} \quad (4.14)$$

for some  $R \geq 1$ . By the Itô isometry, we have

$$\begin{aligned} \frac{1}{\tau} \mathbf{E} \int_0^\tau \left| \int_0^t \Pi_V e^{|z|J^\perp(t-s)} P^{-1} \Pi_{\ker A^\perp} \sigma dW_s \right|^2 dt &\leq \frac{1}{\tau} \int_0^\tau \left( \int_0^t \|\Pi_V e^{|z|J^\perp(t-s)} P^{-1} \Pi_{\ker A^\perp} \sigma\|_F^2 ds \right) dt \\ &\lesssim \frac{1}{\tau} \int_0^\tau \left( \int_0^t e^{2|z|\lambda_R(t-s)} \|P^{-1}\|^2 \|\sigma\|^2 ds \right) dt \\ &\lesssim \frac{\|P^{-1}\| \|\sigma\| |z|^{r+\epsilon}}{\sqrt{\lambda_R} \log(|z|)}. \end{aligned}$$

Then, (4.14) follows by taking  $R$  sufficiently large and using Chebyshev's inequality, completing the proof.  $\square$

We now use Lemma 4.1 to prove the time-averaged growth estimate (2.13) required by Assumption 1 when the initial condition  $x \in \mathbb{R}^n$  is such that  $\Pi_{\ker A} x$  is a spectrally unstable equilibrium point for  $B$ . In what follows, for  $x \in \mathbb{R}^n$  we write  $z = \Pi_{\ker A} x$  and  $y = x - z = \Pi_{\ker A^\perp} x$ .

**Lemma 4.2.** *Suppose that  $B(x, x) = 0$  for every  $x \in \ker A$  and let  $x_0 \in \mathbb{R}^n$  be such that  $L_{z_0/|z_0|}^\perp$  has maximally unstable eigenvalue  $\lambda = \lambda_R + i\lambda_I$  with  $\lambda_R > 0$ . Suppose further that there exists a generalized eigenvector  $v = v_R + iv_I$  satisfying the conditions of Lemma 4.1. Fix  $r \in (0, 1/4)$  and for  $K \geq 1$  set*

$$\eta(K) = 10 \left( \frac{(1/2 + r) \log(K)}{\lambda_R K} \right).$$

There exist  $K_* \geq 1$ ,  $c_* > 0$ , and a universal constant  $\delta_* \in (0, 1/4]$  so that if

$$|y_0| \leq \delta |z_0|^r \text{ and } (1 - \delta)K \leq |x_0| \leq (1 + \delta)K$$

for  $\delta \in (0, \delta_*]$  and  $K \geq K_*$ , then there holds

$$\frac{1}{\eta(K)} \int_0^{\eta(K)} \mathbf{E} |y_t|^2 dt \geq c_* K^{2r}. \quad (4.15)$$

Moreover, if  $\sigma$  is invertible,  $r_0 \in (0, 1/4)$  is fixed, and  $C_0 \geq 1$  is such that  $\lambda_R \geq C_0^{-1}$  and

$$\|P_{z_0/|z_0|}^{-1}\| + \|P_{z_0/|z_0|}\| \leq C_0, \quad (4.16)$$

then the constants  $c_*$  and  $K_*$  can be chosen to depend only on  $C_0$  and  $r_0$  for  $r \leq r_0$ .

*Proof.* We first assume only that  $\{v_R, v_I\} \setminus \{0\} \cap \text{Ran}(\sigma) \neq \emptyset$ . For  $\epsilon \in (0, 1)$  to be chosen, let

$$\tau = \frac{(1/2 + r + \epsilon) \log(|z_0|)}{\lambda_R |z_0|} \quad (4.17)$$

and suppose that

$$\int_0^\tau \mathbf{E}|y_t|^2 dt \leq \tau \delta_1 K^{2r} \quad (4.18)$$

for some  $\delta_1 \in (0, 1)$ . We will obtain a contradiction for  $\delta_1$  sufficiently small.

The first step is to use the contradiction hypothesis (4.18) to obtain bounds on  $|z_t - z_0|$ . Since  $B(z_t, z_t) = 0$  by assumption, we have

$$dz_t = (\Pi_{\ker A}(B(y_t, z_t) + B(z_t, y_t) + B(y_t, y_t) - Ay_t)dt + \Pi_{\ker A}\sigma dW_t. \quad (4.19)$$

Using (4.18), the Cauchy-Schwarz inequality,  $\mathbf{E}|x_t| \lesssim K$  for  $t \lesssim 1$  (this follows from (A.5)), and Doob's martingale inequality we obtain

$$\mathbf{E} \sup_{0 \leq t \leq \tau} |z_t - z_0| \lesssim \sqrt{\delta_1} K^{1+r} \tau + \sqrt{\tau} \lesssim \max(\sqrt{\delta_1}, K_*^{-1/2}) K^{1+r} \tau. \quad (4.20)$$

Define

$$\Omega_0 = \left\{ \omega \in \Omega : \int_0^\tau |y_t|^2 dt \leq \sqrt{\delta_1} \tau K^{2r}, \sup_{0 \leq t \leq \tau} |z_t - z_0| \leq K^{1+r} \tau \right\}$$

and let  $\beta, c_* > 0$  be as in Lemma 4.1 applied with  $z = z_0$  and the chosen  $r \in (0, 1/4)$ . Recall here that  $\beta$  and  $c_*$  do not depend on  $r$  or  $\epsilon$ . By (4.18) and (4.20), for  $\delta_1$  sufficiently small and  $K_*$  sufficiently large depending only on  $\beta$  there holds

$$\mathbf{P}(\Omega_0) > 1 - \beta/2. \quad (4.21)$$

Let  $Y_t$  solve

$$\begin{cases} dY_t = L_{z_0}^\perp Y_t dt + \Pi_{\ker A^\perp} \sigma dW_t \\ Y_t|_{t=0} = \Pi_{\ker A^\perp} x_0. \end{cases}$$

We will show that the exact solution  $y_t$  is well approximated by the linearized dynamics  $Y_t$  on the set  $\Omega_0$ . The difference  $Y_t - y_t$  solves

$$\begin{aligned} \frac{d}{dt}(Y_t - y_t) &= \Pi_{\ker A^\perp}(B(Y_t, z_0) + B(z_0, Y_t) - B(y_t, z_t) - B(z_t, y_t) - B(y_t, y_t) + Ay_t) \\ &= L_{z_0}^\perp(Y_t - y_t) + \Pi_{\ker A^\perp}(B(y_t, z_0 - z_t) + B(z_0 - z_t, y_t) - B(y_t, y_t) + Ay_t). \end{aligned}$$

Therefore,

$$Y_t - y_t = \int_0^t e^{L_{z_0}^\perp(t-s)} \Pi_{\ker A^\perp}(B(y_s, z_0 - z_s) + B(z_0 - z_s, y_s) - B(y_s, y_s) + Ay_s) ds. \quad (4.22)$$

Now, from the Jordan canonical form, for  $t \leq \tau$  and

$$C_1 = \|P_{z_0/|z_0|}\| \|P_{z_0/|z_0|}^{-1}\|$$

there holds

$$\|e^{L_{z_0}^\perp t}\| \lesssim C_1(1 + (|z_0|t)^n)e^{\lambda_R|z_0|t} \lesssim C_1(1 + \lambda_R^{-n})|\log(K)|^n e^{\lambda_R|z_0|t}.$$

Thus, by applying Young's convolution inequality in (4.22), for  $\omega_0 \in \Omega_0$  we have the estimate

$$\begin{aligned} \int_0^\tau |Y_t(\omega_0) - y_t(\omega_0)| dt &\lesssim C_1(1 + \lambda_R^{-n})|\log(K)|^n \left( \int_0^\tau e^{\lambda_R|z_0|t} dt \right) \\ &\quad \times \left( \int_0^\tau (|y_t(\omega_0)|^2 + |y_t(\omega_0)| + |y_t(\omega_0)||z_0 - z_t(\omega_0)|) dt \right) \\ &\lesssim C_1(1 + \lambda_R^{-n-1})|\log(K)|^n K^{r-1/2+\epsilon} \left( \delta_1^{1/4} \tau K^{2r} + \delta_1^{1/4} \tau K^r + \delta_1^{1/4} \tau^2 K^{1+2r} \right) \\ &\lesssim C_1(1 + \lambda_R^{-n-2})|\log(K)|^{n+1} K^\epsilon K^{2r-1/2} (\delta_1^{1/4} K^r \tau), \end{aligned} \quad (4.23)$$

where in the last inequality above we have assumed that  $K$  is large enough so that  $|z_0| \geq K/2$  (and consequently  $K\tau \lesssim \lambda_R^{-1} \log(K)$ ). Assuming  $r \leq r_0 < 1/4$ , we may take  $\epsilon = 1/4 - r_0 > 0$  to obtain

$$\int_0^\tau |Y_t(\omega_0) - y_t(\omega_0)| dt \leq C_1 C_2 \delta_1^{1/4} \tau K^r \quad (4.24)$$

for some constant  $C_2 > 0$  satisfying

$$C_2 \lesssim (1 + \lambda_R^{-n-2}) \sup_{K \geq 1} \{K^{-\epsilon} \log(K)^{n+1}\}. \quad (4.25)$$

With (4.24) established we are now ready to use Lemma 4.1 to complete the proof. Applying Lemma 4.1 and using again  $|z_0| \geq K/2$ , we obtain that for  $K$  sufficiently large depending only on  $\epsilon$  there holds

$$\mathbf{P} \left( \int_0^\tau |Y_t(\omega)| dt \geq \frac{c_*}{2} \tau K^r \right) \geq \mathbf{P} \left( \int_0^\tau |Y_t(\omega)| dt \geq c_* \tau |z_0|^r \right) \geq \beta, \quad (4.26)$$

where the constants  $c_*$  and  $\beta$  are as defined after (4.20). From  $\mathbf{P}(\Omega_0) \geq 1 - \beta/2$  we thus have

$$\mathbf{P} \left( \Omega_0 \cap \left\{ \int_0^\tau |Y_t(\omega)| dt \geq \frac{c_*}{2} \tau K^r \right\} \right) \geq \frac{\beta}{2}. \quad (4.27)$$

By (4.27), the reverse triangle inequality, Cauchy-Schwarz, and (4.24) we deduce that for  $C_1 C_2 \delta_1^{1/4} \leq c_*/4$  there holds

$$\mathbf{E} \int_0^\tau |y_t|^2 dt \geq \frac{\beta c_*^2}{32} \tau K^{2r}. \quad (4.28)$$

Taking  $\delta_1$  even smaller to ensure  $\delta_1 < \frac{\beta c_*^2}{32}$  gives the desired contradiction with (4.18).

In the calculations above,  $\delta_1$  is chosen small depending on  $C_1, C_2, \beta$ , and  $c_*$ , while  $K$  is chosen sufficiently large depending only on  $\beta$  and  $\epsilon = 1/4 - r_0$ . We conclude that there is a constant  $c'_*(C_1, C_2, \beta, c_*) > 0$  and  $K_*(\beta, r_0)$  so that for  $K \geq K_*$  there holds

$$\mathbf{E} \int_0^\tau |y_t|^2 dt \geq c'_* \tau K^{2r}. \quad (4.29)$$

To obtain (4.15) from (4.29), observe that  $|y_0| \leq \delta |z_0|^r$  and  $(1 - \delta)K \leq |x_0| \leq (1 + \delta)K$  imply that  $(1 - 2\delta)K \leq |z_0| \leq (1 + \delta)K$  for  $\delta$  small enough. Therefore, taking  $\delta_*$  sufficiently small and  $K_*$  perhaps larger yields

$$\frac{1}{40} \eta(K) \leq \tau \leq \eta(K),$$

which when combined with (4.29) gives

$$\mathbf{E} \int_0^{\eta(K)} |y_t|^2 dt \geq \frac{c'_*}{40} \eta(K) K^{2r} := c''_* \eta(K) K^{2r}. \quad (4.30)$$

It remains only to argue that if  $\sigma$  is invertible and  $C_0 \geq 1$  is such that  $\lambda_R \geq C_0^{-1}$  and (4.16) holds, then  $K_*$  and the constant  $c''_*$  in (4.30) can be taken to depend only on  $C_0$  and  $r_0$ . Since in the proof of (4.29) we took  $K_* = K_*(\beta, r_0)$  and  $c'_* = c'_*(C_1, C_2, \beta, c_*)$ , it suffices to show that  $\beta$ ,  $C_1$ ,  $C_2$ , and  $c_*$  can be taken to depend only on  $C_0$  and  $r_0$ . By Lemma 4.1 and Remark 10, both  $\beta$  and  $c_*$  depend only on  $C_0$  when  $\sigma$  is invertible. Regarding  $C_1$  and  $C_2$ , following the proof above we see that (4.16) and  $\lambda_R \geq C_0^{-1}$  imply  $C_1 \leq C_0^2$  and

$$C_2 \lesssim (1 + C_0^{n+2}) \sup_{K \geq 1} \{K^{1/4-r_0} \log(K)^{n+1}\}.$$

This completes the proof.  $\square$

## 4.2 Jordan block unstable estimates

In this section, we consider the case where for each  $z \in \ker A$  the eigenvalues of  $L_{z/|z|}^\perp$  all have non-positive real part, but there exists an unstable Jordan block of size greater than or equal to two corresponding to an eigenvalue  $\lambda$  with  $\operatorname{Re}(\lambda) = 0$ . In other words, there exists  $1 \leq J \leq n - 2$  such that

$$t^J \lesssim \|e^{J_{z/|z|}^\perp t}\|_{\mathbb{R}^n \rightarrow \mathbb{R}^n} \lesssim (1 + t^J) \quad (4.31)$$

for all  $t \geq 0$ . Note that when (4.31) holds there necessarily exists a generalized eigenvector  $v = v_R + iv_I$  of  $L_{z/|z|}^\perp$  corresponding to eigenvalue  $\lambda$  such that, defining  $V = \operatorname{span}\{P_{z/|z|}^{-1}v\}$ , there holds both

$$\Pi_V J_{z/|z|}^\perp x = \lambda J_{z/|z|}^\perp \Pi_V x \quad \forall x \in \mathbb{C}^n \quad (4.32)$$

and

$$\left| e^{J_{z/|z|}^\perp t} P_{z/|z|}^{-1} v \right| \gtrsim t. \quad (4.33)$$

In this setting, the analogue of Lemma 4.1 is stated as follows.

**Lemma 4.3.** *Let  $z \in \ker A$  and assume that  $L_{z/|z|}^\perp$  has an unstable Jordan block in the sense that (4.31) holds. Suppose that there exists a generalized eigenvector  $v = v_R + iv_I$  satisfying (4.32) and (4.33) above as well as  $\{v_R, v_I\} \setminus \{0\} \cap \operatorname{Ran}(\sigma) \neq \emptyset$ . Let  $\tilde{v}$  be the generalized eigenvector such that  $(L_{z/|z|}^\perp - \lambda)v = \tilde{v}$  and define  $\tilde{V} = \operatorname{span}\{P_{z/|z|}^{-1}\tilde{v}\}$ . For  $p \in (0, 2/3)$ , set  $r = 1 - 3p/2 > 0$  and*

$$\tau(|z|) = |z|^{-p}.$$

*There are constants  $c_*, \beta > 0$  that do not depend on  $|z|$  so that the solution to (4.2) satisfies*

$$\mathbf{P} \left( \frac{1}{\tau(|z|)} \int_0^{\tau(|z|)} |\Pi_{\tilde{V}} P_{z/|z|}^{-1} Y_t(\omega)| dt \geq c_* |z|^r \right) \geq \beta. \quad (4.34)$$

**Remark 11.** The assumptions above imply that  $v$  is the first vector in a Jordan chain of length greater than or equal to two corresponding to eigenvalue  $\lambda$ . Thus,  $\tilde{v}$  is a generalized eigenvector in the same chain and for any  $x \in \mathbb{C}^n$  there holds

$$\Pi_{\tilde{V}} e^{J_{z/|z|}^\perp t} x = e^{\lambda t} (\Pi_{\tilde{V}} x + t \Pi_V x). \quad (4.35)$$

**Remark 12.** Similar to Lemma 4.1, if  $\sigma$  is invertible and  $\|P_{z/|z}\| + \|P_{z/|z}^{-1}\| \leq C$ , then  $c_*$  and  $\beta$  can be chosen depending only on  $C$  and  $\sigma$ . This follows directly from the proof below after noting that when  $\sigma$  is invertible, the constant  $c_1$  in (4.39) satisfies  $c_1 \gtrsim (\|\sigma^{-1}\| \|P_{z/|z}\|)^{-2}$ .

*Proof.* As in the proof of Lemma 4.1, we assume that  $v_R \in \text{Ran}(\sigma)$  and  $\lambda \neq 0$  (so that  $\lambda$  is pure imaginary); the case  $\lambda = 0$  is a straightforward variation. Denote  $P_{z/|z}$  and  $J_{z/|z}^\perp$  by  $P$  and  $J^\perp$ , respectively. Let  $\tilde{Y}_t = P^{-1}Y_t \in \mathbb{C}^n$ , which is given by the formula

$$\tilde{Y}_t = e^{|z|J^\perp t} \tilde{Y}_0 + \int_0^t e^{|z|J^\perp(t-s)} P^{-1} \Pi_{\ker A^\perp} \sigma dW_s. \quad (4.36)$$

We will follow the same general strategy as in the proof of Lemma 4.1.

We first show that for every  $R > 0$  there exists  $\beta_1 > 0$  so that

$$\mathbf{P} \left( |\Pi_V \tilde{Y}_{\tau/2}| \geq R\sqrt{\tau} \right) \geq \beta_1. \quad (4.37)$$

Since  $v_R \in \text{Ran}(\sigma)$  and

$$|\Pi_V e^{|z|J^\perp(t-s)} P^{-1} v_R| \gtrsim 1, \quad (4.38)$$

we have

$$\text{Var}(\Pi_V \tilde{Y}_{\tau/2}) = \int_0^{\tau/2} \left\| \Pi_V e^{|z|J^\perp(t-s)} P^{-1} \Pi_{\ker A^\perp} \sigma \right\|_F^2 ds \geq c_1 \tau \quad (4.39)$$

for a constant  $c_1$  depending  $v_R$  and  $\sigma$ . The bound (4.37) now follows from the fact that  $\Pi_V \tilde{Y}_{\tau/2}$  is Gaussian.

Next, as in Lemma 4.1, to complete the proof it suffices to show that there is  $c_* > 0$  and  $R$  sufficiently large so that

$$\mathbf{P} \left( \frac{1}{\tau} \int_0^{\tau/2} |\Pi_V \tilde{Y}_t| dt \geq c_* |z|^r \left| |\Pi_V \tilde{Y}_0| \geq R\sqrt{\tau} \right. \right) \geq \frac{1}{2}. \quad (4.40)$$

First, note the elementary fact that for any  $a, b \in \mathbb{R}$  and  $T > 0$  there holds

$$\int_0^T |a + bt| dt \gtrsim |b| T^2 \quad (4.41)$$

with the implicit constant independent of  $a, b$  or  $T$ . One can see this easily by dividing the integral into  $t \leq \min(-\frac{a}{b}, T)$  and  $t \geq \min(-\frac{a}{b}, T)$ . By (4.41), (4.35), and  $|e^{\lambda t}| = 1$  we have

$$\frac{1}{\tau} \int_0^{\tau/2} |\Pi_V e^{|z|J^\perp t} \tilde{Y}_0| dt = \frac{1}{\tau} \int_0^{\tau/2} |\Pi_V \tilde{Y}_0 + t|z|| \Pi_V \tilde{Y}_0| dt \gtrsim R |z| \tau^{3/2}. \quad (4.42)$$

Moreover, by the Itô isometry,

$$\begin{aligned} \frac{1}{\tau} \mathbf{E} \int_0^{\tau/2} \left| \int_0^t \Pi_V e^{|z|J^\perp(t-s)} P^{-1} \Pi_{\ker A^\perp} \sigma dW_s \right|^2 dt &\lesssim \frac{\|P^{-1}\| \|\sigma\|}{\tau} \int_0^{\tau/2} \left( \int_0^t (1 + |z|s)^2 ds \right)^{1/2} dt \\ &\lesssim \|P^{-1}\| \|\sigma\| |z| \tau^{3/2}. \end{aligned} \quad (4.43)$$

Using the reverse triangle inequality and Chebyshev's inequality as in the proof of (4.7), the estimates (4.42) and (4.43) together yield, for  $R \gg \|P^{-1}\| \|\sigma\|$ ,

$$\mathbf{P} \left( \frac{1}{\tau} \int_0^{\tau/2} |\Pi_V \tilde{Y}_t| dt \gtrsim R |z| \tau^{3/2} \left| |\Pi_V \tilde{Y}_0| \geq R\sqrt{\tau} \right. \right) \geq \frac{1}{2}.$$

Since  $r(p)$  is such that  $|z| \tau^{3/2} = |z|^r$  we obtain (4.40), completing the proof.  $\square$



We now turn to the analogue of Lemma 4.2 in the Jordan block unstable case. The idea is the same as in the spectrally unstable case. However, due to the slower timescale of the instability (i.e.,  $p < 1$  in Lemma 4.4) we need to make use of the cancellation

$$\Pi_{\ker A}(B(y, z) + B(z, y)) = 0 \quad \forall z \in \ker A, y \in \ker A^\perp \quad (4.44)$$

assumed in Theorem 1.3. We have not assumed (4.44) in Theorem 1.2 since, as we will show in Lemma 4.6, the cancellation condition is automatically satisfied in the case that  $\dim(\ker A) = 1$  due to  $B(x, x) \cdot x = 0$ .

**Lemma 4.4.** *Suppose that  $B(x, x) = 0$  for every  $x \in \ker A$  and that the cancellation condition (4.44) is satisfied. Let  $x_0 \in \mathbb{R}^n$  be such that  $L_{z_0/|z_0|}^\perp$  is Jordan block unstable in the sense that (4.31) holds and suppose that there exists a generalized eigenvector  $v$  satisfying the conditions in Lemma 4.3. Fix any  $r \in (0, 1/7)$  and for  $K \geq 1$  set*

$$\eta(K) = 4K^{\frac{2r-2}{3}}. \quad (4.45)$$

There exists  $c_* > 0$  and a universal constant  $\delta_* \in (0, 1)$  so that if

$$|y_0| \leq \delta|z_0|^r \text{ and } K/2 \leq |x_0| \leq 2K$$

for  $\delta \in (0, \delta_*)$  and  $K \geq 1$ , then

$$\frac{1}{\eta(K)} \int_0^{\eta(K)} \mathbf{E}|y_t|^2 dt \geq c_* K^{2r}.$$

Moreover, if  $\sigma$  is invertible and

$$\|P_{z_0/|z_0|}\| + \|P_{z_0/|z_0|}^{-1}\| \leq C_0 \quad (4.46)$$

for some  $C_0 \geq 1$ , then  $c_*$  can be chosen depending only on  $\sigma$  and  $C_0$ .

*Proof.* We will consider the case where  $\sigma$  is invertible and (4.46) holds. The proof when one only assumes that  $\{v_R, v_I\} \setminus \{0\} \cap \text{Ran}(\sigma) \neq \emptyset$  follows from exactly the same argument. Let  $K \geq 1$  be such that  $K/2 \leq |x_0| \leq 2K$  and suppose for the sake of contradiction that

$$\mathbf{E} \int_0^\eta |y_t|^2 dt \leq \delta_1 \eta K^{2r} \quad (4.47)$$

for  $\delta_1 \in (0, 1)$  and  $\eta = \eta(K)$  given by (4.45). As in the proof of Lemma 4.2, we will obtain a contradiction for  $\delta_1$  sufficiently small.

By the cancellation condition (4.44), the equation for  $z_t$  is given by

$$dz_t = \Pi_{\ker A}(B(y_t, y_t) - Ay_t)dt + \Pi_{\ker A}\sigma dW_t. \quad (4.48)$$

It follows that

$$\mathbf{E} \sup_{0 \leq t \leq \eta} |z_t - z_0| \lesssim \eta \sqrt{\delta_1} K^{2r} + \sqrt{\eta} \lesssim K^{\frac{r-1}{3}}, \quad (4.49)$$

where we have used that the choices of  $\eta$  and  $r$  are such that

$$\eta K^{2r} \lesssim \sqrt{\eta} \lesssim K^{\frac{r-1}{3}}.$$

Let  $c_*, \beta > 0$  be as in Lemma 4.3 applied with  $z = z_0$  and  $p = (2 - 2r)/3$ . Recall from Remark 12 that  $c_*$  and  $\beta > 0$  depend only on  $C_0$  and  $\sigma$ . By (4.47) and (4.49), for  $\delta_1$  sufficiently small and  $R$  sufficiently large, both depending on  $\beta$ , we have

$$\mathbf{P} \left( \Omega_0 = \left\{ \omega \in \Omega : \int_0^\eta |y_t|^2 dt \leq \sqrt{\delta_1} \eta K^{2r}, \sup_{0 \leq t \leq \eta} |z_t - z_0| \leq RK^{\frac{r-1}{3}} \right\} \right) \geq 1 - \frac{\beta}{2}. \quad (4.50)$$

Now, as in the proof of Lemma 4.2, let  $Y_t$  solve

$$\begin{cases} dY_t = L_{z_0}^\perp Y_t dt + \Pi_{\ker A^\perp} \sigma dW_t \\ Y_t|_{t=0} = \Pi_{\ker A^\perp} x_0. \end{cases}$$

Let  $\tilde{v}$  and  $\tilde{V}$  be as in Lemma 4.3. By the Jordan canonical form, we have

$$\|\Pi_{\tilde{V}} P_{z_0/|z_0|}^{-1} e^{L_{z_0}^\perp t}\| \leq C_0(1 + |z_0|t) \leq C_0(1 + 2Kt).$$

Using this in (4.22) we obtain, for  $\omega_0 \in \Omega_0$ ,

$$\begin{aligned} \int_0^\eta |\Pi_{\tilde{V}} P_{z_0/|z_0|}^{-1} (Y_t(\omega_0) - y_t(\omega_0))| dt &\lesssim C_0 \left( \int_0^\eta (1 + Kt) dt \right) \left( \delta_1^{1/4} \eta K^{2r} + R \delta_1^{1/4} \eta K^{\frac{r-1}{3}} K^r \right) \\ &\lesssim C_0 \delta_1^{1/4} \eta K^r (K\eta) \left( \eta K^r + R \eta K^{\frac{r-1}{3}} \right) \\ &\lesssim C_0 R \delta_1^{1/4} \eta K^r (K\eta) \eta K^r, \end{aligned}$$

where in the last line we noted that trivially  $K^{\frac{r-1}{3}} \leq K^r$ . Observe now that the restriction  $r < 1/7$  and the formula for  $\eta$  imply that

$$(K\eta) \eta K^r \leq 1,$$

and thus we have

$$\int_0^\eta |\Pi_{\tilde{V}} P_{z_0/|z_0|}^{-1} (Y_t(\omega_0) - y_t(\omega_0))| dt \lesssim C_0 R \delta_1^{1/4} \eta K^r. \quad (4.51)$$

We now use (4.51) and Lemma 4.3 to complete the proof. Suppose that  $\delta$  is small enough so that  $K/4 \leq |z_0| \leq 2K$ . Then, the choice of  $\eta(K)$  ensures that

$$|z_0|^{\frac{2r-2}{3}} \leq \eta(K) \leq 8|z_0|^{\frac{2r-2}{3}}.$$

Thus, from Lemma 4.3 we have

$$\mathbf{P} \left( \frac{1}{\eta(K)} \int_0^{\eta(K)} |\Pi_{\tilde{V}} P_{z_0/|z_0|}^{-1} Y_t| dt \geq \frac{c_*}{32} K^r \right) \geq \beta. \quad (4.52)$$

It follows from (4.51), (4.52), and  $\mathbf{P}(\Omega_0) \geq 1 - \beta/2$  that for  $\delta_1$  sufficiently small depending only on  $c_*$ ,  $\beta$ , and  $C_0$  there holds

$$\mathbf{E} \int_0^\eta |\Pi_{\tilde{V}} P_{z_0/|z_0|}^{-1} y_t|^2 dt \geq \left( \frac{c_*}{64} \right)^2 \frac{\beta}{2} \eta K^{2r}.$$

Therefore,

$$\mathbf{E} \int_0^\eta |y_t|^2 dt \geq \frac{1}{C_0^2} \left( \frac{c_*}{64} \right)^2 \frac{\beta}{2} \eta K^{2r}.$$

We obtain a contradiction by taking  $\delta_1$  perhaps even smaller to guarantee

$$\delta_1 \leq \frac{1}{C_0^2} \left( \frac{c_*}{64} \right)^2 \frac{\beta}{4}.$$

Since  $c_*$  and  $\beta$  depend only on  $C_0$  and  $\sigma$  we obtain

$$\mathbf{E} \int_0^\eta |y_t|^2 dt \geq c \eta K^{2r}$$

for a constant  $c$  depending only on  $C_0$  and  $\sigma$ , which completes the proof.  $\square$

**Remark 13.** If the cancellation condition (4.44) is assumed in the spectrally unstable case, so that (4.19) can be replaced with (4.48), one can show that any  $r \in (0, 1)$  is permissible in Lemma 4.2.

### 4.3 Concluding the proofs of Theorems 1.2 and 1.3

In this section we use the results from Sections 4.1 and 4.2 to prove Theorems 1.2 and 1.3.

#### 4.3.1 $\dim(\ker A) = 1$

In this section we prove Theorem 1.2. It is a special case of the result below, stated for more general assumptions on  $\sigma$ . Recall that we denote  $\mathcal{U} = \ker A \cap \mathbb{S}^{n-1}$  and for  $x \in \mathcal{U}$  write

$$L_x^\perp = P_x J_x^\perp P_x^{-1}$$

for the Jordan normal form of  $L_x^\perp$ .

**Theorem 4.5.** *Suppose that  $\mathcal{U} = \{x_0, -x_0\}$  for some unit vector  $x_0$  and that for each  $x \in \mathcal{U}$  there holds*

$$B(x, x) = 0, \quad \lim_{t \rightarrow \infty} \left\| e^{tL_x^\perp} \right\| = \infty. \quad (4.53)$$

Moreover, let  $\sigma$  satisfy the following conditions (which hold trivially when  $\text{rank}(\sigma) = n$ ).

- If  $x \in \mathcal{U}$  is such that  $L_x^\perp$  has an eigenvalue with positive real part, then there is a generalized eigenvector  $v = v_R + iv_I$  associated with the eigenvalue  $\lambda$  of  $L_x^\perp$  with maximal real part such that  $\{v_R, v_I\} \setminus \{0\} \cap \text{Ran}(\sigma) \neq \emptyset$  and

$$\Pi_{\text{span}\{P_x^{-1}v\}} J_x^\perp y = \lambda J_x \Pi_{\text{span}\{P_x^{-1}v\}} y \quad \forall y \in \mathbb{C}^n. \quad (4.54)$$

- If  $x \in \mathcal{U}$  is such that  $t^J \lesssim \|e^{L_x^\perp t}\| \lesssim 1 + t^J$  for some  $1 \leq J \leq n - 2$ , then there is a generalized eigenvector  $v = v_R + iv_I$  associated with an eigenvalue  $\lambda$  of  $L_x^\perp$  with  $\text{Re}(\lambda) = 0$  that satisfies  $\{v_R, v_I\} \setminus \{0\} \cap \text{Ran}(\sigma) \neq \emptyset$ , (4.54), and  $|e^{J_x^\perp t} P_x^{-1}v| \gtrsim t$ .

Then, there exists at least one stationary measure  $\mu$  and  $\langle x \rangle^p \in L^1(d\mu)$  for all  $p < 1/3$ .

We begin by showing that the cancellation condition (4.44) is automatically satisfied in one dimension due to  $B(x, x) \cdot x = 0$ .

**Lemma 4.6.** *Let  $\Pi : \mathbb{R}^n \rightarrow \mathbb{R}^n$  be a projection onto a one-dimensional subspace of  $\mathbb{R}^n$ . Suppose  $B : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  is a bilinear function satisfying  $B(x, x) \cdot x = 0$  and*

$$B(\Pi x, \Pi x) = 0 \quad \forall x \in \mathbb{R}^n.$$

Then,

$$\Pi B(\Pi x, \Pi^\perp x) + \Pi B(\Pi^\perp x, \Pi x) = 0$$

for every  $x \in \mathbb{R}^n$ .

*Proof.* The property  $B(x, x) \cdot x = 0$  remains true after any orthogonal coordinate transform, and so without loss of generality we may assume that  $\Pi$  is the projection onto the subspace  $\{(x_1, 0, \dots, 0) : x_1 \in \mathbb{R}\}$ . In this setting, we need to show that

$$B_1(\Pi x, \Pi^\perp x) + B_1(\Pi^\perp x, \Pi x) = 0$$

for any  $x \in \mathbb{R}^n$ . The condition  $B(\Pi x, \Pi x) = 0$  implies that  $\partial_{x_1}^2 B(x, x) = 0$ . Hence, differentiating  $B(x, x) \cdot x = 0$  twice with respect to  $x_1$  gives

$$\partial_{x_1} [B_1(x, x)] = 0.$$

Substituting  $x = \Pi x + \Pi^\perp x$  we find

$$\partial_{x_1} (B_1(\Pi x, \Pi^\perp x) + B_1(\Pi^\perp x, \Pi x)) = 0.$$

Noting that  $B_1(\Pi x, \Pi^\perp x) + B_1(\Pi^\perp x, \Pi x) = 0$  when  $x_1 = 0$  completes the proof.  $\square$

We are now ready to prove Theorem 4.5:

*Proof of Theorem 4.5.* Let  $\mathcal{U} = \{x_0, -x_0\}$  for  $x_0 \in \ker A \cap \mathbb{S}^{n-1}$ . Fix  $r < 1/7$  and for  $\delta \in (0, 1)$  to be chosen let

$$B_K^\delta = \{x \in \mathbb{R}^n : |\Pi_{\ker A^\perp} x|^2 \leq \delta |\Pi_{\ker A} x|^{2r} \text{ and } (1 - \delta)K^2 \leq |x|^2 \leq (1 + \delta)K^2\}.$$

For any  $x \in \mathbb{R}^n \setminus \{0\}$  there is  $c > 0$  and  $j \in \{1, 2\}$  such that  $\Pi_{\ker A} x = c(-1)^j x_0$ . Therefore, defining for  $j \in \{1, 2\}$  the sets

$$B_{K,j}^\delta = \{x \in B_K^\delta : \Pi_{\ker A} x = c(-1)^j x_0 \text{ for some } c > 0\},$$

we have  $B_K^\delta = B_{K,1}^\delta \cup B_{K,2}^\delta$ . The assumptions in (4.53) imply that  $x_0$  and  $-x_0$  are both equilibria of  $B$  with  $L_{x_0}^\perp$  and  $L_{-x_0}^\perp$  spectrally or Jordan block unstable. Observe now that  $r < 1/7$  is always permitted in Lemmas 4.2 and 4.4. Moreover, the associated  $\eta$  always satisfies

$$\eta(K) \lesssim K^{-4/7}.$$

Therefore, by Lemmas 4.2 and 4.4 (note that we may apply Lemma 4.4 in the present one-dimensional setting due to Lemma 4.6) there are constants  $\delta > 0$ ,  $c_* > 0$ , and  $K_* \geq 1$  along with functions  $\eta_j(K)$  satisfying  $\lim_{K \rightarrow \infty} \sup_{j=1,2} \eta_j(K) = 0$  such that for  $K \geq K_*$  there holds

$$x_0 \in B_{K,j} \implies \frac{1}{\eta_j(K)} \int_0^{\eta_j(K)} |\Pi_{\ker A^\perp} x_t|^2 dt \geq c_* K^{2r}.$$

Thus, Assumption 1 is satisfied for any  $r < 1/7$ . Theorem 4.5 then follows from Lemma 2.1.  $\square$

**Remark 14.** Let  $\mu$  be the stationary measure constructed in Theorem 4.5. If  $L_{x_0}^\perp$  and  $L_{-x_0}^\perp$  are both spectrally unstable, then by Remark 13 and Lemma 4.6 it holds that  $\langle x \rangle^p \in L^1(d\mu)$  for every  $p > 0$ .

### 4.3.2 $\dim(\ker A) > 1$

*Proof of Theorem 1.3.* We will give the details only for the spectrally unstable case, i.e., the case where there exists  $C_0 \geq 1$  so that for every  $z \in \mathcal{U}$  there is a maximally unstable eigenvalue  $\lambda(z)$  of  $L_z^\perp$  satisfying  $\operatorname{Re}(\lambda(z)) > 0$ . Fix  $r < 1/4$  and for  $\delta \in (0, 1/4)$ ,  $K \geq 1$ , and  $z \in \mathcal{U}$ , define the sets

$$B_K^\delta = \{x \in \mathbb{R}^n : |\Pi_{\ker A^\perp} x|^2 \leq \delta |\Pi_{\ker A} x|^{2r}, (1 - \delta)K^2 \leq |x|^2 \leq (1 + \delta)K^2\}$$

and

$$B_{K,z}^\delta = B_K^\delta \cap \{x \in \mathbb{R}^n : \Pi_{\ker A} x / |\Pi_{\ker A} x| = z \in \mathcal{U}\}.$$

Since  $\mathcal{U}$  is compact and the eigenvalues of a matrix vary continuously with respect its entries, we have

$$0 < \lambda_- := \min_{z \in \mathcal{U}} \operatorname{Re}(\lambda(z)) \leq \max_{z \in \mathcal{U}} \operatorname{Re}(\lambda(z)) := \lambda_+ < \infty.$$

Therefore, by Lemma 4.2 and (1.10) there exist  $\delta_* \in (0, 1/4)$ ,  $K_* \geq 1$ , and  $c_* > 0$  so that for every  $z \in \mathcal{U}$  and  $K \geq K_*$ , defining

$$\eta(K, z) = 10 \left( \frac{(1/2 + r) \log(K)}{\operatorname{Re}(\lambda(z)) K} \right),$$

there holds

$$x_0 \in B_{K,z}^{\delta_*} \implies \frac{1}{\eta(K, z)} \mathbf{E} \int_0^{\eta(K, z)} |\Pi_{\ker A^\perp} x_t|^2 dt \geq c_* K^{2r}.$$

We used here the statement at the end of Lemma 4.2 about the dependence of the constants when  $\sigma$  is invertible and (4.46) holds. Since

$$10 \left( \frac{(1/2 + r) \log(K)}{\lambda_+ K} \right) \leq \eta(K, z) \leq 10 \left( \frac{(1/2 + r) \log(K)}{\lambda_- K} \right)$$

for every  $z \in \mathcal{U}$  it follows that for

$$\eta_*(K) := 10 \left( \frac{(1/2 + r) \log(K)}{\lambda_- K} \right)$$

and  $K \geq K_*$  there holds

$$x_0 \in B_K^{\delta_*} \implies \frac{1}{\eta_*(K)} \mathbf{E} \int_0^{\eta_*(K)} |\Pi_{\ker A^\perp} x_t|^2 dt \geq \frac{\lambda_-}{\lambda_+} c_* K^{2r}.$$

Thus, Assumption 1 is satisfied for any  $r < 1/4$ , which due to Lemma 2.2 completes the proof.  $\square$

## 5 Sabra and Galerkin Navier-Stokes

### 5.1 Statement and proof of general result

In this section we state and prove the general theorem that will be used to obtain Theorem 1.4 announced earlier.

**Theorem 5.1.** *Let  $\text{rank}(\sigma) = n$  and suppose that  $\ker A = V_1 \oplus V_2$  for orthogonal subspaces of  $\mathbb{R}^n$  satisfying the following properties.*

- For any  $x \in V_1 \cup V_2$ ,  $B(x, x) = 0$ , i.e.,  $V_1$  and  $V_2$  consist of deterministic equilibria.
- There is  $C > 0$  so that

$$\max_{j=1,2} \sup_{x \in V_j \cap \mathbb{S}^{n-1}} (\|P_{x,j}\| + \|P_{x,j}^{-1}\|) \leq C, \quad (5.1)$$

where  $P_{x,j} J_j^\perp P_{x,j}^{-1}$  denotes the Jordan canonical form of  $\Pi_{V_j^\perp} L_x \Pi_{V_j^\perp}$ , with  $L_x$  as defined in (1.9).

- There is  $\lambda_{\min} > 0$  such that for any  $j \in \{1, 2\}$  and  $x \in V_j \cap \mathbb{S}^{n-1}$  there is an eigenvalue  $\lambda$  of  $\Pi_{V_j^\perp} L_x \Pi_{V_j^\perp}$  with  $\text{Re}(\lambda) \geq \lambda_{\min}$ .
- There exists  $c > 0$  so that for any  $v_1 \in V_1$  and  $v_2 \in V_2$  there holds

$$\Pi_{\ker A} B(v_1, v_2) + \Pi_{\ker A} B(v_2, v_1) = 0 \quad (5.2)$$

and

$$|\Pi_{\ker A^\perp} B(v_1, v_2) + \Pi_{\ker A^\perp} B(v_2, v_1)| \geq c |v_1| |v_2|. \quad (5.3)$$

Then, there exists at least one stationary measure  $\mu$  and  $\langle x \rangle^p \in L^1(d\mu)$  for every  $p < 2/3$ .

The proof of Theorem 5.1 will proceed roughly as follows. As before we will verify Assumption 1. For initial conditions  $x_0$  near  $\ker A$  with  $\min(|\Pi_{V_1} x_0|, |\Pi_{V_2} x_0|)$  sufficiently large, we use (5.3) and arguments similar to those in Section 3 to obtain growth of the damped modes. If instead  $x_0$  is concentrated in one of the  $V_j$ , we proceed similarly to Section 4 and use the spectral instability to deduce growth into  $V_j^\perp$ . This either causes the damped modes to grow directly or the solution to enter a region where  $\min(|\Pi_{V_1} x_t|, |\Pi_{V_2} x_t|)$  is large enough to subsequently apply (5.3) as in the first case. The cancellation (5.2) is used throughout to justify certain approximations.

We begin with a lemma that describes growth of the damped modes for initial conditions with  $|\Pi_{V_1} x_0|$  and  $|\Pi_{V_2} x_0|$  both sufficiently large.

**Lemma 5.2.** Fix  $r \in (0, 1]$  and  $\delta_0 \in (0, 1)$ . There are  $c_*(\delta_0) > 0$  and  $K_*(\delta_0) \geq 1$  so that for any  $x_0 \in \mathbb{R}^n$  satisfying

$$K/2 \leq |x_0| \leq 2K, \quad |\Pi_{\ker A^\perp} x_0|^2 \leq \delta K^{2r}, \quad \min(|\Pi_{V_1} x_0|, |\Pi_{V_2} x_0|) \geq \delta_0^{1/8} K^r$$

for  $0 \leq \delta \leq \epsilon \delta_0^{3/4}$  and  $K \geq K_*$ , where  $\epsilon$  is a sufficiently small constant independent of  $\delta_0$ , there holds

$$\frac{1}{K^{-1}} \mathbf{E} \int_0^{K^{-1}} |\Pi_{\ker A^\perp} x_t|^2 dt \geq c_* K^{2r}.$$

*Proof.* Let  $X_t$  solve

$$\begin{cases} \frac{d}{dt} X_t = B(X_t, X_t) \\ X_0 = x_0 \end{cases} \quad (5.4)$$

and define  $\eta(K) = K^{-1}$ . We claim that there is  $c_*(\delta_0) > 0$  so that for all  $K \geq 1$ ,  $\delta \ll \delta_0^{3/4}$ , and  $x_0$  as in the statement of the lemma there holds

$$\frac{1}{\eta(K)} \int_0^{\eta(K)} |\Pi_{\ker A^\perp} X_t|^2 dt \geq c_* K^{2r}. \quad (5.5)$$

From here the lemma follows by taking  $K_*$  large enough so that  $c_* K_*^{2r} \gg 1$  and applying Lemma 3.3. We now prove (5.5). For  $\gamma \in (0, 1)$  to be chosen sufficiently small and  $\tilde{\eta}(K) = \gamma \delta_0^{1/4} K^{-1}$ , suppose for the sake of contradiction that

$$\frac{1}{\tilde{\eta}} \int_0^{\tilde{\eta}} |\Pi_{\ker A^\perp} X_t|^2 dt \leq \delta_1 K^{2r} \quad (5.6)$$

for  $\delta_1 \in (0, 1)$ . By performing a Taylor expansion and using (5.3),  $|\Pi_{\ker A^\perp} X_0| \leq \sqrt{\delta} K^r$ , and  $|X_t| \leq 2K$  we obtain, for  $t \leq \tilde{\eta}$ ,

$$\begin{aligned} |\Pi_{\ker A^\perp} X_t| &= \left| \Pi_{\ker A^\perp} X_0 + t \Pi_{\ker A^\perp} B(X_0, X_0) + \Pi_{\ker A^\perp} \int_0^t (t-s) \frac{d}{ds} B(X_s, X_s) ds \right| \\ &\geq ct |\Pi_{V_1} X_0| |\Pi_{V_2} X_0| - C \sqrt{\delta} K^r - CK \tilde{\eta} \int_0^t |B(X_s, X_s)| ds, \end{aligned} \quad (5.7)$$

where in the second inequality we used that

$$\frac{d}{ds} B(X_s, X_s) = B(B(X_s, X_s), X_s) + B(X_s, B(X_s, X_s)).$$

The goal is now to bound the integral in (5.7). First, by writing

$$X_s = \Pi_{\ker A} X_s + \Pi_{\ker A^\perp} X_s = \Pi_{V_1} X_s + \Pi_{V_2} X_s + \Pi_{\ker A^\perp} X_s$$

and using the triangle inequality we deduce

$$|B(X_s, X_s)| \lesssim K |\Pi_{\ker A} X_s - \Pi_{\ker A} X_0| + K |\Pi_{\ker A^\perp} X_s| + |\Pi_{V_1} X_0| |\Pi_{V_2} X_0|. \quad (5.8)$$

Now, by (5.6) and (5.2), for all  $t \leq \tilde{\eta}$  there holds

$$\begin{aligned} |\Pi_{\ker A} X_t - \Pi_{\ker A} X_0| &\leq \int_0^t |B(\Pi_{\ker A} X_s, \Pi_{\ker A^\perp} X_s)| ds + \int_0^t |B(\Pi_{\ker A^\perp} X_s, X_s)| ds \\ &\lesssim K \int_0^t |\Pi_{\ker A^\perp} X_s| ds \lesssim \gamma \delta_0^{1/4} K^r. \end{aligned} \quad (5.9)$$

Putting this bound into (5.8) and using (5.6) again gives, for  $t \leq \tilde{\eta}$ ,

$$\int_0^t |B(X_s, X_s)| ds \lesssim \gamma \delta_0^{1/4} K^r + \tilde{\eta} |\Pi_{V_1} X_0| |\Pi_{V_2} X_0|. \quad (5.10)$$

Inserting (5.10) in (5.7) and integrating the resulting bound over  $[0, \tilde{\eta}]$  yields, for  $\gamma$  sufficiently small (and new constants  $c$  and  $C$  which may change from line to line),

$$\begin{aligned} \frac{1}{\tilde{\eta}} \int_0^{\tilde{\eta}} |\Pi_{\ker A^\perp} X_t| dt &\geq c\tilde{\eta} |\Pi_{V_1} X_0| |\Pi_{V_2} X_0| - C\sqrt{\delta} K^r - C\gamma^2 \sqrt{\delta_0} K^r - C\gamma\tilde{\eta} |\Pi_{V_1} X_0| |\Pi_{V_2} X_0| \\ &\geq \tilde{\eta} |\Pi_{V_1} X_0| |\Pi_{V_2} X_0| (c - C\gamma) - C\sqrt{\delta} K^r - C\gamma^2 \sqrt{\delta_0} K^r \\ &\geq c\gamma \delta_0^{3/8} K^r - C\sqrt{\delta} K^r. \end{aligned}$$

In the last inequality we have used the fact that  $|\Pi_{V_j} X_0| \gtrsim K$  for some  $j \in \{1, 2\}$  when  $\delta \ll 1$ . With  $\gamma$  now fixed we may take  $\delta \ll \gamma^2 \delta_0^{3/4}$  to obtain, for some new constant  $c \in (0, 1)$ ,

$$\frac{1}{\tilde{\eta}} \int_0^{\tilde{\eta}} |\Pi_{\ker A^\perp} X_t|^2 dt \geq c\gamma^2 \delta_0^{3/4} K^{2r}.$$

We obtain a contradiction with (5.6) for  $\delta_1 \leq (c/2)\gamma^2 \delta_0^{3/4}$ , and so we conclude

$$\frac{1}{\eta} \int_0^\eta |\Pi_{\ker A^\perp} X_t|^2 dt \geq \frac{\tilde{\eta}}{\eta} \frac{c}{2} \gamma^2 \delta_0^{3/4} K^{2r} = \frac{c}{2} \gamma^3 \delta_0 K^{2r},$$

which implies (5.5) and completes the proof.  $\square$

*Proof of Theorem 5.1.* Fix any  $r \in (0, 1/4)$  and for  $\delta \in (0, 1/100)$  and  $K \geq 2$  define

$$B_K^\delta = \{x \in \mathbb{R}^n : |\Pi_{\ker A^\perp} x|^2 \leq \delta |\Pi_{\ker A} x|^{2r}, (1 - \delta)K^2 \leq |x|^2 \leq (1 + \delta)K^2\}.$$

Let

$$\ell(j) = \begin{cases} 1 & j = 2 \\ 2 & j = 1. \end{cases}$$

We split the set  $B_K^\delta$  as

$$B_K^\delta = B_{K,1}^\delta \cup B_{K,2}^\delta \cup B_{K,3}^\delta,$$

where

$$B_{K,j}^\delta = \{x \in B_K^\delta : |\Pi_{V_{\ell(j)}} x| \leq \delta^{1/8} K^r\} \quad \text{if } j \in \{1, 2\}$$

and

$$B_{K,3}^\delta = \{x \in B_K^\delta : \min(|\Pi_{V_1} x|, |\Pi_{V_2} x|) > \delta^{1/8} K^r\}.$$

By Lemma 2.2, to complete the proof it suffices to show that there are  $K_* \geq 1$ ,  $\delta > 0$ ,  $c_* > 0$ , and times  $\{\eta_j\}_{j=1}^3$  with  $\lim_{K \rightarrow \infty} \sup_j \eta_j(K) = 0$  so that for  $K \geq K_*$  there holds

$$x_0 \in B_{K,j}^\delta \implies \frac{1}{\eta_j(K)} \int_0^{\eta_j(K)} |\Pi_{\ker A^\perp} x_t|^2 dt \geq c_* K^{2r}. \quad (5.11)$$

Due to Lemma 5.2, for all  $\delta$  sufficiently small there is  $c_*(\delta) > 0$  and  $K_*(\delta) \geq 1$  so that (5.11) is satisfied for  $j = 3$  by taking  $\eta_3(K) = K^{-1}$ . Thus we must only consider the case where  $j \in \{1, 2\}$ .

Let  $x_0 \in B_{K,j}^\delta$  for  $j \in \{1, 2\}$  and fix any  $\bar{r}$  with  $r < \bar{r} < 1/4$ . Suppose that the maximally unstable eigenvalue of  $\Pi_{V_j^\perp} L_{\Pi_{V_j} x_0} / |\Pi_{V_j} x_0| \Pi_{V_j^\perp}$  has real part  $\lambda \geq \lambda_{\min} > 0$  and define, for  $\epsilon \in (0, 1)$  to be chosen,

$$\tau_1 = \frac{(1/2 + \bar{r} + \epsilon) \log(|\Pi_{V_j} x_0|)}{\lambda |\Pi_{V_j} x_0|}$$

and  $\tau = \tau_1 + K^{-1}$ . Let the approximate solution  $Y_t : [0, \infty) \rightarrow V_j^\perp$  solve

$$\begin{cases} dY_t = (\Pi_{V_j^\perp} L_{\Pi_{V_j} x_0} \Pi_{V_j^\perp}) Y_t dt + \Pi_{V_j^\perp} \sigma dW_t \\ Y_0 = \Pi_{V_j^\perp} x_0. \end{cases} \quad (5.12)$$

By Lemma 4.1 (with  $\Pi_{V_j} x_0$  and  $V_j$  playing the roles of  $z$  and  $\ker A$ , respectively), there are  $K_j(\epsilon) \geq 1$  and  $c_j, \beta_j > 0$  that do not depend on  $\epsilon, \bar{r}$ , or  $x_0$  so that for  $|\Pi_{V_j} x_0| \geq K_j$  there holds

$$\mathbf{P} \left( \frac{1}{\tau_1} \int_0^{\tau_1} |Y_t(\omega)| dt \geq c_j |\Pi_{V_j} x_0|^{\bar{r}} \right) \geq \beta_j. \quad (5.13)$$

Towards a contradiction, suppose that

$$\mathbf{E} \int_0^\tau |\Pi_{\ker A^\perp} x_t|^2 dt \leq \delta \tau K^{2r} \quad (5.14)$$

for  $K \geq K_* \geq \delta^{-1}$ . Note that since  $\tau/\tau_1 \lesssim 1$  this implies

$$\mathbf{E} \int_0^{\tau_1} |\Pi_{\ker A^\perp} x_t|^2 dt \lesssim \delta \tau_1 K^{2r}. \quad (5.15)$$

The condition (5.2) and the fact that  $V_1 \cup V_2$  consists of deterministic equilibria imply that the equation for  $z_t = \Pi_{\ker A} x_t$  is exactly (4.19). Thus, using (5.15), the proof of (4.20) applies and gives

$$\mathbf{E} \sup_{0 \leq t \leq \tau_1} |z_t - z_0| \lesssim \sqrt{\delta} K^{1+r} \tau_1 + \sqrt{\tau_1} \lesssim \sqrt{\delta} \log(K) K^r, \quad (5.16)$$

where in the second inequality we used the assumption that  $K \geq \delta^{-1}$ . Define

$$\Omega_0 = \left\{ \omega \in \Omega : \int_0^{\tau_1} |\Pi_{V_j^\perp} x_t|^2 dt \leq \delta^{1/8} \tau_1 (\log(K))^2 K^{2r}, \sup_{0 \leq t \leq \tau_1} |z_t - z_0| \leq \delta^{1/4} K^r \log(K) \right\}.$$

Since  $|\Pi_{V_{\ell(j)}} x_0| \leq \delta^{1/8} K^r$  by the definition of  $B_{K,j}^\delta$ , it follows from (5.15) and (5.16) that for  $\delta$  sufficiently small depending on  $\beta_j$  there holds

$$\mathbf{P}(\Omega_0) \geq 1 - \frac{\beta_j}{2}. \quad (5.17)$$

Obtaining estimates on  $Y_t(\omega) - \Pi_{V_j^\perp} x_t(\omega)$  for  $\omega \in \Omega_0$  as in proof of Lemma 4.2 (we make the choice  $\epsilon = 1/4 - \bar{r}$ ) and then using (5.13), we deduce that there is  $c_0 \in (0, 1)$  depending only on  $c_j$  so that for  $\delta$  sufficiently small and  $K_*$  sufficiently large there holds

$$\mathbf{P} \left( \frac{1}{\tau_1} \int_0^{\tau_1} |\Pi_{V_j^\perp} x_t(\omega)| dt \geq c_0 K^{\bar{r}} \right) \geq \frac{\beta_j}{2}. \quad (5.18)$$

It follows that

$$\mathbf{P} \left( \frac{1}{\tau_1} \int_0^{\tau_1} |\Pi_{\ker A^\perp} x_t(\omega)| dt \geq \frac{c_0}{2} K^{\bar{r}} \right) \geq \frac{\beta_j}{4} \quad \text{or} \quad \mathbf{P} \left( \frac{1}{\tau_1} \int_0^{\tau_1} |\Pi_{V_{\ell(j)}} x_t(\omega)| dt \geq \frac{c_0}{2} K^{\bar{r}} \right) \geq \frac{\beta_j}{4}. \quad (5.19)$$



In the first case, we immediately obtain a contradiction to (5.15) for  $\delta$  sufficiently small depending on  $c_0$  and  $\beta_j$ . In the second case, define the stopping time

$$\bar{\tau}(\omega) = \inf \left\{ t \geq 0 : \min(|\Pi_{V_1} x_t|, |\Pi_{V_2} x_t|) \geq \frac{c_0}{2} K^{\bar{r}}, K/2 \leq |x_t| \leq 2K, |\Pi_{\ker A^\perp} x_t|^2 \leq \delta^{1/8} K^{2\bar{r}} \right\}.$$

Now, using  $|\Pi_{V_{\ell(j)}} x_0| \leq \delta^{1/8} K^r$  and  $|\Pi_{\ker A^\perp} x_0| \leq \sqrt{\delta} K^r$  we can show

$$\begin{aligned} \sup_{0 \leq t \leq \tau_1} |\Pi_{\ker A^\perp} x_t| &\lesssim \sqrt{\delta} K^r + K^{1+r} \delta^{1/8} \tau_1 + \tau_1 K \sup_{0 \leq t \leq \tau_1} |z_t - z_0| + \sup_{0 \leq t \leq \tau_1} |W_t| \\ &+ \tau_1 \left( \sup_{0 \leq t \leq \tau_1} |z_t - z_0| \right)^2 + \int_0^{\tau_1} |\Pi_{\ker A^\perp} x_t|^2 dt \\ &+ \sqrt{\tau_1} \left( \sup_{0 \leq t \leq \tau_1} |z_t - z_0| + K \right) \left( \int_0^{\tau_1} |\Pi_{\ker A^\perp} x_t|^2 dt \right)^{1/2}. \end{aligned}$$

It follows then from (5.15), (5.16), and  $K \geq \delta^{-1}$  that for  $\delta$  sufficiently small there holds

$$\mathbf{P} \left( \sup_{0 \leq t \leq \tau_1} |\Pi_{\ker A^\perp} x_t| \leq \delta^{1/16} (\log(K))^3 K^r \right) \geq 1 - \frac{\beta_j}{32}. \quad (5.20)$$

By Lemma A.1, (5.20), (5.16), and assuming the second case in (5.19), for  $K_*$  sufficiently large depending on  $\beta_j$ ,  $r$ , and  $\bar{r}$  we have

$$\mathbf{P}(\bar{\tau} \leq \tau_1) \geq \frac{\beta_j}{8}.$$

Thus, by Lemma 5.2 and the strong Markov property, there is  $c'_0$  depending on  $c_0$  so that for all  $\delta$  sufficiently small and  $K$  sufficiently large (both depending only on  $c_0$ ) there holds

$$\begin{aligned} \frac{1}{\tau} \mathbf{E} \int_0^\tau |\Pi_{\ker A^\perp} x_t|^2 dt &\geq \frac{1}{\tau} \int_\Omega \int_0^{\tau - \tau_1 \wedge \bar{\tau}} D(x_{\tau_1 \wedge \bar{\tau} + t}) dt d\mathbf{P} \\ &\geq \frac{1}{\tau} \frac{\beta_j}{8} \inf_{\bar{\tau}(\omega) \leq \tau_1} \int_0^{K^{-1}} \mathcal{P}_t D(x_{\bar{\tau}}(\omega)) dt \\ &\geq \frac{1}{\tau K} \frac{\beta_j}{8} c'_0 K^{2\bar{r}} \gtrsim \frac{\beta_j c'_0 K^{2\bar{r}}}{\log(K)}. \end{aligned}$$

Taking  $K$  large enough so that  $K^{2\bar{r}} \log(K)^{-1} \geq K^{2r}$  yields a contradiction with (5.14) for  $\delta$  small enough. Overall, we have shown that for all  $\delta$  sufficiently small there are  $K_* \geq 1$  and  $c_* > 0$  so that, for  $j \in \{1, 2\}$  and  $K \geq K_*$ ,

$$x_0 \in B_{K,j}^\delta \implies \frac{1}{\tau(x_0)} \mathbf{E} \int_0^{\tau(x_0)} |\Pi_{\ker A^\perp} x_t|^2 dt \geq c_* K^{2r}.$$

The desired bound (5.11) then follows for  $j \in \{1, 2\}$  by setting

$$\eta_1(K) = \eta_2(K) = \frac{C \log(K)}{\lambda_{\min} K}$$

for some  $C$  sufficiently large. □

## 5.2 Applications to the Sabra shell model and Galerkin Navier-Stokes

In this section we first apply Theorem 5.1 to prove Theorem 1.4 on the 2d Galerkin Navier-Stokes equations and then we give an application of Theorem 5.1 to the Sabra shell model.

*Proof of Theorem 1.4.* Note that the nonlinear structure implies that  $L_x = \Pi_{V_j^\perp} L_x \Pi_{V_j^\perp}$  for any  $x \in V_j$  and  $j \in \{1, 2\}$ . By translation invariance, the linearization around  $A \cos \ell x_1 + B \sin \ell x_1$  is unitarily conjugate to the linearization around  $\sqrt{A^2 + B^2} \cos \ell x_1$ , and so the uniformity of eigenvalues and  $\|P\|$  of the Jordan canonical form follows immediately on each  $V_j$  once it is verified for  $\cos \ell x_1$  (by discrete rotation invariance, the analysis in the  $x_2$  direction is also the same as the  $x_1$  direction). That for any  $\ell \geq 2$ , the linearization of  $\cos \ell x_1$  in  $\mathbb{T}^2$  has an eigenvalue with positive real part in the conservative system is a well-known variation of the classical results of Meshalkin and Sinai [29]. It remains only to verify the conditions (5.2) and (5.3). For  $v_1 = \alpha \cos \ell x_1 + \beta \sin \ell x_1$  and  $v_2 = \sin kx_2$  (this is sufficient by translation invariance), we may compute

$$\begin{aligned} B(v_1, v_2) + B(v_2, v_1) &= \frac{k}{\ell} (-\alpha \sin \ell x_1 + \beta \cos \ell x_1) \cos kx_2 - \frac{\ell}{k} \cos kx_2 (-\alpha \sin \ell x_1 + \beta \cos \ell x_1) \\ &= \left( \frac{k}{\ell} - \frac{\ell}{k} \right) (-\alpha \sin \ell x_1 + \beta \cos \ell x_1) \cos kx_2. \end{aligned}$$

For our choices of  $k, \ell$  conditions (5.2) and (5.3) follow immediately and hence Theorem 5.1 applies.  $\square$

The Sabra shell model was first introduced in [25]. Here we consider the model truncated to finite dimensions. Denoting the dependent variable  $(u_1, \dots, u_J) \in \mathbb{C}^J$ , the equation reads

$$\begin{aligned} du_m &= i2^m \left( \overline{u_{m+1}} u_{m+2} - \frac{\delta}{2} \overline{u_{m-1}} u_{m+1} - \frac{\delta-1}{4} u_{m-2} u_{m-1} \right) \\ &\quad - \delta 2^{2m} u_m + q_m dW_t^{(m;R)} + ip_m dW_t^{(m;I)}, \end{aligned} \tag{5.21}$$

where  $q_m, p_m$  are real parameters and  $\delta \in (0, 2) \setminus \{1\}$ . The boundary conditions are  $u_{-1} = u_0 = u_{J+1} = u_{J+2} = 0$ . When  $\delta \in (0, 1)$  the system has just one positive invariant and is considered a model for 3d turbulence. If instead  $\delta \in (1, 2)$  then there are two positive invariants and the equations are meant to capture properties of 2d turbulence. For additional discussion of Sabra and other shell models, see [12]. Rewriting the system in real variables  $u_m = a_m + ib_m$  and introducing parameters  $c_m \in \{0, 1\}$  that determine whether or not there is damping on shell  $m$ , we obtain the system

$$\begin{aligned} da_m &= 2^m (a_{m+2} b_{m+1} - a_{m+1} b_{m+2}) + \delta 2^{m-1} (a_{m-1} b_{m+1} - a_{m+1} b_{m-1}) \\ &\quad + (\delta - 1) 2^{m-2} (a_{m-2} b_{m-1} + a_{m-1} b_{m-2}) - \delta 2^{2m} c_m a_m + q_m dW_t^{(m;R)}, \\ db_m &= 2^m (a_{m+1} b_{m+2} + b_{m+1} b_{m+2}) - \delta 2^{m-1} (a_{m-1} a_{m+1} + b_{m+1} b_{m-1}) \\ &\quad - (\delta - 1) 2^{m-2} (a_{m-2} a_{m-1} - b_{m-1} b_{m-2}) - \delta 2^{2m} c_m b_m + p_m dW_t^{(m;I)}. \end{aligned} \tag{5.22}$$

**Theorem 5.3.** *Assume that  $\delta \in (1/4, 1)$ ,  $c_1 = c_2 = 0$ ,  $c_m > 0$  for  $3 \leq m \leq J$ , and  $q_m, p_m \neq 0$  for all  $m$ . Then, system (5.22) admits a unique invariant measure  $\mu$  and*

$$\int_{\mathbb{R}^{2J}} (|a| + |b|)^p \mu(da, db) < \infty$$

for every  $p < 2/3$ .

*Proof.* We denote the solution  $(a, b) = (a_1, \dots, a_J, b_1, \dots, b_J) \in \mathbb{R}^J \times \mathbb{R}^J$  and the natural canonical basis vectors by  $\{\hat{a}_m\}_{m=1}^J, \{\hat{b}_m\}_{m=1}^J$ . Observe that (5.22) takes the form of (1.1) with

$$\ker A = \{(a_1, a_2, 0, \dots, 0, b_1, b_2, 0, \dots, 0) : a_1, b_1, a_2, b_2 \in \mathbb{R}\},$$

and the drift  $B$  given by

$$\begin{aligned} B\left((a, b), (\tilde{a}, \tilde{b})\right) &= \sum_{m=1}^J [2^m(a_{m+2}\tilde{b}_{m+1} - a_{m+1}\tilde{b}_{m+2}) + \delta 2^{m-1}(a_{m-1}\tilde{b}_{m+1} - a_{m+1}\tilde{b}_{m-1}) \\ &\quad + (\delta - 1)2^{m-2}(a_{m-2}\tilde{b}_{m-1} + a_{m-1}\tilde{b}_{m-2})] \hat{a}_m \\ &\quad + \sum_{m=1}^J [2^m(a_{m+1}\tilde{b}_{m+2} + b_{m+1}\tilde{b}_{m+2}) - \delta 2^{m-1}(a_{m-1}\tilde{a}_{m+1} + b_{m+1}\tilde{b}_{m-1}) \\ &\quad - (\delta - 1)2^{m-2}(a_{m-2}\tilde{a}_{m-1} - b_{m-1}\tilde{b}_{m-2})] \hat{b}_m. \end{aligned} \quad (5.23)$$

We will verify the conditions of Theorem 5.1. For  $j = 1, 2$  let

$$V_j = \text{span}\{\hat{a}_j, \hat{b}_j\}.$$

It is immediate from the lack of self-interactions in (5.23) that each  $V_j$  consists entirely of deterministic equilibria, and so the first condition in Theorem 5.1 is satisfied. Similarly, since the  $m$ 'th coordinates of  $(a, b)$  and  $(\tilde{a}, \tilde{b})$  do not show up in  $B \cdot \hat{a}_m$  or  $B \cdot \hat{b}_m$  it is easy to see that (5.2) is satisfied. To verify the lower bound (5.3) we compute, for  $v_1 = a_1 \hat{a}_1 + b_1 \hat{b}_1 \in V_1$  and  $v_2 = a_2 \hat{a}_2 + b_2 \hat{b}_2 \in V_2$ ,

$$\begin{aligned} |\Pi_{\ker A^\perp}(B(v_1, v_2) + B(v_2, v_1))| &= |\Pi_{\text{span}\{\hat{a}_3, \hat{b}_3\}}(B(v_1, v_2) + B(v_2, v_1))| \\ &= |2(\delta - 1)| |(a_1 b_2 + a_2 b_1) \hat{a}_3 - (a_1 a_2 - b_1 b_2) \hat{b}_3| \\ &= |2(\delta - 1)| \sqrt{(a_1 b_2 + a_2 b_1)^2 + (a_1 a_2 - b_1 b_2)^2} \\ &= |2(\delta - 1)| \sqrt{(a_1^2 + b_1^2)(a_2^2 + b_2^2)} \\ &= |2(\delta - 1)| |v_1| |v_2|. \end{aligned}$$

It remains only to check the second and third conditions. This requires computing the linearized operators  $\Pi_{V_j^\perp} L_x \Pi_{V_j^\perp}$  for  $x \in V_j$  with  $|x| = 1$ . For  $j = 1$ , let

$$x = (\bar{a}_1, 0, \dots, 0, \bar{b}_1, 0, \dots, 0)$$

for  $\bar{a}_1, \bar{b}_1 \in \mathbb{R}$  satisfying  $\sqrt{\bar{a}_1^2 + \bar{b}_1^2} = 1$ . For general  $(a, b) \in \mathbb{R}^J \times \mathbb{R}^J$  we compute

$$\Pi_{V_1^\perp} L_x \Pi_{V_1^\perp}(a, b) = \begin{pmatrix} 0 & 0 & -2\delta \bar{b}_1 & 2\delta \bar{a}_1 \\ 0 & 0 & -2\delta \bar{a}_1 & -2\delta \bar{b}_1 \\ 2(\delta - 1)\bar{b}_1 & 2(\delta - 1)\bar{a}_1 & 0 & 0 \\ -2(\delta - 1)\bar{a}_1 & 2(\delta - 1)\bar{b}_1 & 0 & 0 \end{pmatrix} \begin{pmatrix} a_2 \\ b_2 \\ a_3 \\ b_3 \end{pmatrix},$$

with the components not shown being zero. The eigenvalues and associated eigenvectors of the matrix above are given by

$$\lambda_{+,1} = 2\sqrt{\delta(1-\delta)}, \quad E_{+,1} = \left\{ \begin{pmatrix} -\bar{a}_1 \sqrt{\frac{\delta}{\delta-1}} \\ \bar{b}_1 \sqrt{\frac{\delta}{\delta-1}} \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} \bar{b}_1 \sqrt{\frac{\delta}{\delta-1}} \\ -\bar{a}_1 \sqrt{\frac{\delta}{\delta-1}} \\ 1 \\ 0 \end{pmatrix} \right\}, \quad (5.24)$$

$$\lambda_{-,1} = -2\sqrt{\delta(1-\delta)}, \quad E_{-,1} = \left\{ \begin{pmatrix} \bar{a}_1 \sqrt{\frac{\delta}{\delta-1}} \\ -\bar{b}_1 \sqrt{\frac{\delta}{\delta-1}} \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -\bar{b}_1 \sqrt{\frac{\delta}{\delta-1}} \\ -\bar{a}_1 \sqrt{\frac{\delta}{\delta-1}} \\ 1 \\ 0 \end{pmatrix} \right\}. \quad (5.25)$$

For  $x = \bar{a}_2 \hat{a}_2 + \bar{b}_2 \hat{b}_2$  with  $|x| = 1$  we similarly have

$$\Pi_{V_2^\perp} L_x \Pi_{V_2^\perp}(a, b) = \begin{pmatrix} 0 & 0 & 2\bar{b}_2 & -2\bar{a}_2 & 0 & 0 \\ 0 & 0 & 2\bar{a}_2 & 2\bar{b}_2 & 0 & 0 \\ 2(\delta-1)\bar{b}_2 & 2(\delta-1)\bar{a}_2 & 0 & 0 & -4\delta\bar{b}_2 & 4\delta\bar{a}_2 \\ -2(\delta-1)\bar{a}_2 & 2(\delta-1)\bar{b}_2 & 0 & 0 & -4\delta\bar{a}_2 & -4\delta\bar{b}_2 \\ 0 & 0 & 4(\delta-1)\bar{b}_2 & 4(\delta-1)\bar{a}_2 & 0 & 0 \\ 0 & 0 & -4(\delta-1)\bar{a}_2 & 4(\delta-1)\bar{b}_2 & 0 & 0 \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \\ a_3 \\ b_3 \\ a_4 \\ b_4 \end{pmatrix}.$$

Defining  $c_\delta = 5\delta - 4\delta^2 - 1 > 0$ , the eigenvalues and associated eigenvectors are given by

$$\lambda_0 = 0, \quad E_0 = \left\{ \begin{pmatrix} \frac{-4\bar{a}_2\bar{b}_2\delta}{\delta-1} \\ \frac{2\bar{b}_2^2\delta - 2\bar{a}_2^2\delta}{\delta-1} \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} \frac{2\bar{b}_2^2\delta - 2\bar{a}_2^2\delta}{\delta-1} \\ \frac{4\bar{a}_2\bar{b}_2\delta}{\delta-1} \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\}, \quad (5.26)$$

$$\lambda_{+,2} = 2\sqrt{c_\delta}, \quad E_{+,2} = \left\{ \begin{pmatrix} \frac{-\bar{a}_2\bar{b}_2}{\delta-1} \\ \frac{\bar{b}_2^2 - \bar{a}_2^2}{2(\delta-1)} \\ \frac{\bar{a}_2\bar{b}_2}{\delta-1} \\ \frac{-\bar{a}_2\sqrt{c_\delta}}{2(\delta-1)} \\ \frac{\bar{b}_2\sqrt{c_\delta}}{2(\delta-1)} \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} \frac{\bar{b}_2^2 - \bar{a}_2^2}{2(\delta-1)} \\ \frac{\bar{a}_2\bar{b}_2}{\delta-1} \\ \frac{\bar{b}_2\sqrt{c_\delta}}{2(\delta-1)} \\ \frac{\bar{a}_2\sqrt{c_\delta}}{2(\delta-1)} \\ 1 \\ 0 \end{pmatrix} \right\}, \quad (5.27)$$

and

$$\lambda_{-,2} = 2\sqrt{c_\delta}, \quad E_{-,2} = \left\{ \begin{pmatrix} \frac{-\bar{a}_2\bar{b}_2}{\delta-1} \\ \frac{\bar{b}_2^2 - \bar{a}_2^2}{2(\delta-1)} \\ \frac{\bar{a}_2\bar{b}_2}{\delta-1} \\ \frac{\bar{a}_2\sqrt{c_\delta}}{2(\delta-1)} \\ \frac{-\bar{b}_2\sqrt{c_\delta}}{2(\delta-1)} \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} \frac{\bar{b}_2^2 - \bar{a}_2^2}{2(\delta-1)} \\ \frac{\bar{a}_2\bar{b}_2}{\delta-1} \\ \frac{-\bar{b}_2\sqrt{c_\delta}}{2(\delta-1)} \\ \frac{-\bar{a}_2\sqrt{c_\delta}}{2(\delta-1)} \\ 1 \\ 0 \end{pmatrix} \right\}. \quad (5.28)$$

Since, for each  $j$ ,  $\lambda_{+,j}$  is positive and independent of  $x \in V_j \cap \mathbb{S}^{n-1}$ , we see that the third condition of Theorem 5.1 is satisfied. Lastly, (5.1) follows from the formula for the eigenvectors given in (5.24)-(5.28). This completes the proof of Theorem 5.3.  $\square$

## 6 Lorenz-96 with a two-dimensional kernel

In this section, we consider the stochastic Lorenz-96 system for  $x_t = (x_{t,1}, \dots, x_{t,n}) \in \mathbb{R}^n$  (with  $n \geq 6$ ) defined by

$$dx_{t,j} = -a_j x_{t,j} dt + B_j(x_t, x_t) dt + \sigma_j dW_t^{(j)}, \quad (6.1)$$

where  $x_{t,k} = x_{t,k+n}$  (periodic conditions),  $a_j \geq 0$ , and

$$B_j(x, x) = (x_{j+1} - x_{j-2})x_{j-1}. \quad (6.2)$$

Consistent with our earlier notation, we write  $A = \text{diag}(a_1, \dots, a_n)$ . A consequence of Theorem 4.5 is that (6.1) admits an invariant measure when  $a_1 = 0$ ,  $\sigma_{n-1} \neq 0$ , and  $a_j > 0$  for all  $2 \leq j \leq n$ . Indeed, in this case  $B(x, x) = 0$  for every  $x \in \ker A$  and moreover it is straightforward to check that for  $z = (z_0, 0, \dots, 0) \in \ker A$  the linearized operator  $L_z^\perp$  is Jordan block unstable with

$$(|z|t)^3 \lesssim \|e^{L_z^\perp t}\| \lesssim 1 + (|z|t)^3 \quad \text{and} \quad |e^{L_z^\perp t} e_{n-1}| \gtrsim |z|t,$$

where  $e_{n-1}$ , which denotes the usual canonical basis vector, is a generalized eigenvector that is the last element of a Jordan chain. Our goal in this section is to show that an invariant measure can in fact also be constructed using our methods in the more degenerate case where  $a_1 = a_2 = 0$ . The main result is stated precisely as follows.

**Theorem 6.1.** *Let  $6 \leq n < \infty$ . The stochastic Lorenz-96 system with  $a_1 = a_2 = 0$  and  $a_j > 0$  for  $3 \leq j \leq n$  admits an invariant measure  $\mu_*$  provided that  $\sigma_n, \sigma_{n-1} \neq 0$ . Moreover, we have the moment bound*

$$\int_{\mathbb{R}^n} |x|^p \mu_*(dx) < \infty$$

for every  $0 < p < 1/3$ .

As in the earlier sections, we will prove Theorem 6.1 by verifying Assumption 1 using suitable approximation arguments for solutions in the vicinity of  $\ker A$ . To this end, for  $K \gg 1$  and  $\delta \in (0, 1)$  we split the set

$$B_{K,\delta} = \{x \in \mathbb{R}^n : |\Pi_{\ker A^\perp} x| \leq \delta |\Pi_{\ker A} x|^{1/7} \text{ and } K/2 \leq |x| \leq 2K\}$$

as

$$B_{K,\delta} = B_{K,\delta}^1 \cup B_{K,\delta}^2 \cup B_{K,\delta}^3,$$

where, for some small parameter  $\delta_1 \in (0, 1)$ ,

$$B_{K,\delta}^1 = \{x \in B_{K,\delta} : |x_1| \geq K/\sqrt{32}\},$$

$$B_{K,\delta}^2 = \{x \in B_{K,\delta} : \delta_1 K^{1/7} \leq |x_1| < K/\sqrt{32}\},$$

and

$$B_{K,\delta}^3 = \{x \in B_{K,\delta} : |x_1| < \delta_1 K^{1/7}\}.$$

Note that  $|x_2| \geq K/\sqrt{32}$  for  $x \in B_{K,\delta}^2 \cup B_{K,\delta}^3$ .

In the region  $B_{K,\delta}^1$ , we can use a treatment similar to that used for Jordan block unstable equilibria in Theorem 1.2. Specifically, we show that a large  $x_1$  induces a significant growth in  $x_n$  through the interaction  $\dot{x}_n = x_1 x_{n-1} + \dots$ . Here we rely on the fact that  $x_{n-1}$  is being driven by a Brownian motion (since  $\sigma_{n-1} \neq 0$ ), which ensures it is non-trivial with high probability. In the region  $B_{K,\delta}^2$ , we can use a treatment similar to that used in Theorem 1.1, by noting that  $\dot{x}_3 = -x_1 x_2 + \dots$  and hence if both  $x_1$  and  $x_2$  are sufficiently large, then  $x_3$  will rapidly grow. The region  $B_{K,\delta}^3$  is the region that is most different from previous cases. Here, the Jordan block instability of the equilibrium  $e_2$  excites  $x_1$ , which is still in  $\ker A$ . Heuristically, we show that solutions which start in  $B_{K,\delta}^3$  are basically ejected into  $B_{K,\delta}^2$ , where they are subsequently ejected into  $(\ker A)^\perp$ .

By Lemma 2.2, Theorem 6.1 is a direct consequence of the following time-averaged coercivity estimates.

**Proposition 6.2.** Let  $\tau_1(K) = \tau_3(K) = K^{-4/7}$  and  $\tau_2(K) = K^{-1}$ . There exist  $K_* \geq 1$ ,  $c_* > 0$ , and  $\delta, \delta_1 \in (0, 1)$  so that if  $K \geq K_*$  and  $x_0 \in B_{K,\delta}^j$  then

$$\frac{1}{\tau_j(K)} \int_0^{\tau_j(K)} |\Pi_{\ker A^\perp} x_t|^2 dt \geq c_* K^{2/7}.$$

Consequently, Assumption 1 is satisfied with  $r = 1/7$ .

As discussed above the region  $B_{K,\delta}^3$  is the most involved. The main difficulty here is to deduce growth of  $\Pi_{\ker A^\perp} X_t$  for a suitable approximate solution when  $X_0 \in B_{K,\delta}^3$ , as proved in the next lemma.

**Lemma 6.3.** Let  $X_t$  solve

$$\begin{cases} dX_{t,n} = X_{t,1}X_{t,n-1}dt + \sigma_n dW_t^{(n)} \\ dX_{t,1} = X_{t,2}X_{t,n}dt \\ dX_{t,3} = -X_{t,1}X_{t,2}dt \\ dX_{t,j} = 0 \end{cases} \quad j \notin \{n, 1, 3\}. \quad (6.3)$$

with initial condition  $X_0 \in B_{K,\delta}^3$  and some  $\sigma_n \neq 0$ . For  $\delta$  and  $\delta_1$  chosen sufficiently small, there are constants  $\beta, c_* > 0$  (independent of  $X_0$ ) so that for all  $K$  sufficiently large and  $\tau = K^{-4/7}$  there holds

$$\mathbf{P} \left( \frac{1}{\tau} \int_0^\tau |X_{t,3}| dt \geq c_* K^{4/7} \right) \geq \beta. \quad (6.4)$$

*Proof.* Without loss of generality we set  $\sigma_n dW_t^{(n)} = dW_t$  for a standard Brownian motion  $W_t$ . We also write  $r = 1/7$ , so that  $\tau = K^{-4r}$  and moreover from the definition of  $B_{K,\delta}^3$  we have

$$\sum_{j=3}^n |X_{0,j}|^2 \leq \delta^2 (|X_{0,1}|^2 + |X_{0,2}|^2)^r \leq 4\delta^2 K^{2r}, \quad (6.5)$$

$$|X_{0,1}| \leq \delta_1 K^r. \quad (6.6)$$

For  $R \gg 1$  to be chosen we split into the cases  $|X_{0,n}| \geq R\sqrt{\tau}$  and  $|X_{0,n}| < R\sqrt{\tau}$ . In the former, we approximate  $X_{t,n} \approx X_{0,n}$  and in the latter we approximate  $X_{t,n} \approx X_{0,n} + W_t$ .

**Case 1** ( $|X_{0,n}| \geq R\sqrt{\tau}$ ): Write  $X_{t,n} = X_{0,n} + E_t$ , where  $E_t$  is an error to be controlled. Substituting this into the system we have

$$\begin{cases} dE_t = X_{t,1}X_{0,n-1}dt + dW_t \\ dX_{t,1} = X_{0,2}(X_{0,n} + E_t)dt \\ dX_{t,3} = -X_{t,1}X_{0,2}dt. \end{cases} \quad (6.7)$$

Thus,

$$X_{t,1} = X_{0,1} + tX_{0,2}X_{0,n} + X_{0,2} \int_0^t E_s ds \quad (6.8)$$

and

$$E_t = tX_{0,n-1}X_{0,1} + \frac{t^2}{2}X_{0,n-1}X_{0,n}X_{0,2} + X_{0,n-1}X_{0,2} \int_0^t (t-s)E_s ds + W_t.$$

Recalling  $|X_0| \leq 2K$  and using the bounds (6.5) and (6.6) we have

$$|E_t| \leq 2\delta\delta_1 t K^{2r} + 2\delta t^2 K^{r+1} |X_{0,n}| + 4\delta t K^{r+1} \int_0^t |E_s| ds + |W_t|, \quad (6.9)$$

and therefore

$$\mathbf{E} \sup_{0 \leq t' \leq t} |E_{t'}| \leq 2t\delta\delta_1 K^{2r} + 2\delta t^2 K^{1+r} |X_{0,n}| + 4\delta t K^{1+r} \int_0^t \mathbf{E} \sup_{0 \leq t' \leq s} |E_{t'}| ds + \sqrt{t} \quad (6.10)$$

By Grönwall's Lemma,  $|X_{0,n}| \geq RK^{-2/7}$ , and the definitions of  $r$  and  $\tau$ , it follows that

$$\begin{aligned} \mathbf{E} \sup_{0 \leq t \leq \tau} |E_t| &\leq (2\tau\delta\delta_1 K^{2r} + 2\delta\tau^2 K^{1+r} |X_{0,n}| + \sqrt{\tau}) \exp(4\delta\tau^2 K^{1+r}) \\ &\leq C(K^{-2/7} + \delta |X_{0,n}|) \\ &\leq 2C \max(\delta, R^{-1}) |X_{0,n}|, \end{aligned}$$

where  $C$  is a constant that does not depend on  $\delta$ ,  $\delta_1$ , or  $K$ . Putting in  $R = \delta^{-1/2}$  we conclude

$$\mathbf{E} \sup_{0 \leq t \leq \tau} |E_t| \leq 2C\sqrt{\delta} |X_{0,n}|. \quad (6.11)$$

The goal is now to use (6.11) to show that  $X_{t,3}$  must grow. By (6.11) and Chebyshev's inequality, for  $\delta$  sufficiently small we have

$$\mathbf{P} \left( \sup_{0 \leq t \leq \tau} |E_t| \leq \delta^{1/4} |X_{0,n}| \right) \geq \frac{1}{2}. \quad (6.12)$$

Suppose that  $\omega \in \Omega$  is such that

$$\sup_{0 \leq t \leq \tau} |E_t(\omega)| \leq \delta^{1/4} |X_{0,n}|. \quad (6.13)$$

We will show that in this case, for  $\delta$  sufficiently small, there is  $c_* \in (0, 1]$  (independent of  $\omega$ ,  $X_0$ ,  $K$ ) such that

$$\int_0^\tau |X_{t,3}(\omega)| dt \geq c_* \tau K^{4/7}, \quad (6.14)$$

which is sufficient to imply (6.4) with  $\beta = 1/2$ . From (6.8) and (6.13) we have (suppressing now the dependence on  $\omega$  from the notation)

$$|X_{t,1} - tX_{0,2}X_{0,n}| \leq |X_{0,1}| + \delta^{1/4} |X_{0,2}| |X_{0,n}| t \quad \forall t \in [0, \tau]. \quad (6.15)$$

Applying this bound in the formula for  $X_{t,3}$  and using  $|X_{0,3}| \leq K^r$  we get, for  $t \in [0, \tau]$ ,

$$\begin{aligned} |X_{t,3}| &\geq \left| \int_0^t s X_{0,2}^2 X_{0,n} ds \right| - |X_{0,2}| \int_0^t (|X_{0,1}| + s\delta^{1/4} |X_{0,2}| |X_{0,n}|) ds - K^r \\ &\geq \frac{t^2}{2} |X_{0,2}|^2 |X_{0,n}| - t |X_{0,2}| |X_{0,1}| - \delta^{1/4} t^2 |X_{0,2}|^2 |X_{0,n}| - K^r \\ &\geq \frac{t^2}{4} |X_{0,2}|^2 |X_{0,n}| - t |X_{0,2}| |X_{0,1}| - K^r, \end{aligned}$$

where in the last inequality we have assumed that  $\delta$  is sufficiently small. We thus have

$$|X_{t,3}| \geq \frac{\tau^2}{16} |X_{0,2}|^2 |X_{0,n}| - \tau |X_{0,2}| |X_{0,1}| - K^r \quad \forall t \in [\tau/2, \tau]. \quad (6.16)$$

Using (6.16),  $|X_{0,1}| \leq K^r$ ,  $\delta^{-1/2} K^{-2r} \leq |X_{0,n}| \leq 2\delta K^r$ , and  $K/\sqrt{32} \leq |X_{0,2}| \leq 2K$  it follows that for  $t \in [\tau/2, \tau] = [K^{-4r}/2, K^{-4r}]$  and  $\delta$  sufficiently small there holds

$$|X_{t,3}| \geq \frac{\delta^{-1/2} K^{2-10r}}{512} - 2K^{1-3r} - K^r \geq \frac{1}{512} \delta^{-1/2} K^{4/7} - 3K^{4/7},$$

where we have noted that the choice  $r = 1/7$  implies  $2 - 10r = 1 - 3r = 4/7$ . Hence, for  $\delta$  sufficiently small we have  $|X_{t,3}| \geq K^{4/7}$  for  $t \in [\tau/2, \tau]$ , and so

$$\int_0^\tau |X_{t,3}| dt \geq \frac{1}{2} \tau K^{4/7},$$

which proves (6.14).

**Case 2** ( $|X_{0,n}| \leq R\sqrt{\tau} = \delta^{-1/2}\sqrt{\tau}$ ): Let  $X_{t,n} = X_{0,n} + W_t + E_t$ , where again  $E_t$  is an error to be bounded. Computations similar to those of Case 1 give

$$\mathbf{E} \sup_{0 \leq t \leq \tau} \frac{|E_t|}{\sqrt{t}} \leq CR\delta \leq C\sqrt{\delta}, \quad (6.17)$$

where  $C$  is a constant that does not depend on  $\delta$  or  $K$ . Now we justify the growth of  $X_{t,3}$ , which is also similar to above. We have

$$X_{t,1} = X_{0,1} + X_{0,2}X_{0,n}t + X_{0,2} \int_0^t W_s ds + X_{0,2} \int_0^t E_s ds. \quad (6.18)$$

Without loss of generality we may assume that  $X_{0,2}X_{0,n} \geq 0$ . By the scaling and support theorems for Brownian motion, there exists  $\alpha > 0$  that does not depend on  $K$  such that

$$\mathbf{P} \left( X_{0,2} \int_0^t W_s ds \geq |X_{0,2}| \tau^{3/2} \quad \forall t \in [\tau/4, \tau] \right) \geq \alpha. \quad (6.19)$$

By (6.17) and (6.19), if  $\delta$  is small enough we have

$$\mathbf{P} \left( X_{0,2} \int_0^t W_s ds \geq |X_{0,2}| \tau^{3/2} \quad \forall t \in [\tau/4, \tau] \quad \text{and} \quad \sup_{0 \leq t \leq \tau} |E(t)|/\sqrt{t} \leq \delta^{1/4} \right) \geq \frac{\alpha}{2}. \quad (6.20)$$

Let  $\omega \in \Omega$  be such that the two bounds in (6.20) hold true. We will prove that for such an  $\omega$  one has

$$\int_0^\tau |X_{t,3}(\omega)| dt \geq c_* \tau K^{4/7}$$

for  $c_*$  sufficiently small. First, there is nothing to show if

$$\int_0^\tau |X_{t,3}(\omega)| dt \geq \frac{\tau}{8000} K^{4/7},$$

so suppose otherwise. In this case, there exists  $t_0 \in [\tau/4, \tau/2]$  is such that  $|X_{t_0,3}| \leq K^{4/7}/2000$ . Then, for  $t \in [t_0, \tau]$  there holds

$$\begin{aligned} |X_{t,3}| &\geq \left| \int_{t_0}^t X_{0,2}^2 \left( X_{0,n}s + \int_0^s W_{s'} ds' \right) ds \right| \\ &\quad - \int_{t_0}^t |X_{0,2}| \left( |X_{0,1}| + |X_{0,2}| \int_0^s |E_{s'}| ds' \right) ds - \frac{K^{4/7}}{2000} \\ &\geq |X_{0,2}|^2 (t - t_0) \tau^{3/2} - t |X_{0,2}| |X_{0,1}| - \delta^{1/4} |X_{0,2}|^2 t^{5/2} - \frac{K^{4/7}}{2000}, \end{aligned}$$

where in obtaining the final inequality we have noted that  $X_{0,n}$  and  $\int_0^s W_{s'} ds'$  have the same sign for  $s \geq t_0$  since  $X_{0,n}X_{0,2} \geq 0$ . Taking  $\delta$  sufficiently small and  $t \in [3\tau/4, \tau]$  to absorb the third term by the first we obtain

$$|X_{t,3}| \geq \frac{K^2 \tau^{5/2}}{256} - 2\delta_1 \tau K^{1+r} - \frac{K^{4/7}}{2000} \geq \frac{K^{4/7}}{512} - 2\delta_1 K^{4/7}, \quad (6.21)$$



where we have recalled also that  $K^2/32 \leq X_{0,2}^2 \leq 4K^2$ ,  $|X_{0,1}| \leq \delta K^r$ , and  $2 - 10r = 1 - 3r = 4/7$ . For  $\delta_1$  sufficiently small we conclude that  $|X_{t,3}| \geq K^{4/7}/1024$  for  $t \in [3\tau/4, \tau]$ . Thus,

$$\mathbf{P} \left( \int_0^\tau |X_{t,3}| dt \geq \frac{1}{4096} \tau K^{4/7} \right) \geq \frac{\alpha}{2},$$

which completes the proof.  $\square$

We are now ready to prove Proposition 6.2.

*Proof of Proposition 6.2.* Let  $x_0 \in B_{K,\delta}$  for  $K$  to be taken sufficiently large and  $\delta, \delta_1$  chosen appropriately. As before, set  $r = 1/7$  for the sake of simplifying the presentation of the estimates. Let  $\tau_j(K)$  be as given in the statement of the proposition. There are three cases to consider.

**Case 1** ( $x_0 \in B_{K,\delta}^1$ ): Consider the approximate solution  $X_t$  defined by

$$\begin{cases} dX_{t,n} = X_{t,1}X_{t,n-1}dt + \sigma_n dW_t^{(n)} \\ dX_{t,n-1} = \sigma_{n-1} dW_t^{(n-1)} \\ dX_{t,j} = 0 \end{cases} \quad j \notin \{n, n-1\} \quad (6.22)$$

and initial condition  $X_0 = x_0$ . We have

$$X_{t,n} = X_{0,n} + tX_{0,1}X_{0,n-1} + \sigma_{n-1}X_{0,1} \int_0^t W_s^{(n-1)} ds + \sigma_n W_t^{(n)}.$$

Similar to as in Case 2 from the proof of Lemma 6.3, using the support theorem for Brownian motion we can show that

$$\mathbf{P} \left( |X_{t,n}| \geq K\tau_1^{3/2} \quad \forall t \in [\tau_1/2, \tau_1] \right) \geq \alpha \quad (6.23)$$

for some  $\alpha > 0$  that does not depend on  $K$  or  $\delta$ . Consequently, since  $K\tau_1^{3/2} = K^r$ ,

$$\mathbf{P} \left( \frac{1}{\tau_1} \int_0^{\tau_1} |X_{t,n}| dt \geq \frac{1}{2} K^r \right) \geq \alpha. \quad (6.24)$$

Suppose now for the sake of contradiction that

$$\mathbf{E} \int_0^{\tau_1} |\Pi_{\ker A^\perp} x_t|^2 dt = \mathbf{E} \sum_{j=3}^n \int_0^{\tau_1} |x_{t,j}|^2 dt \leq \delta \tau_1 K^{2r}. \quad (6.25)$$

The error  $X_{t,n} - x_{t,n}$  solves

$$d(X_{t,n} - x_{t,n}) = x_{0,1}(X_{t,n-1} - x_{t,n-1})dt + x_{t,n-1}(x_{0,1} - x_{t,1})dt + a_n x_{t,n} dt + x_{t,n-2} x_{t,n-1} dt \quad (6.26)$$

with zero initial condition, and so

$$\begin{aligned} |X_{t,n} - x_{t,n}| &\leq 2K \int_0^t |X_{s,n-1} - x_{s,n-1}| ds + \int_0^t |x_{s,n-1}| |x_{0,1} - x_{s,1}| ds \\ &\quad + |a_n| \int_0^t |x_{s,n}| ds + \int_0^t |x_{s,n-2}| |x_{s,n-1}| ds. \end{aligned} \quad (6.27)$$

We now obtain bounds on  $|X_{t,n-1} - x_{t,n-1}|$  and  $|x_{0,1} - x_{t,1}|$ . By (6.25) and  $\mathbf{E}|x_t| \lesssim K$  for  $t \leq 1$  we have

$$\mathbf{E} \sup_{0 \leq t \leq \tau_1} |x_{0,1} - x_{t,1}| \leq C \left( \sqrt{\delta} \tau_1 K^{1+r} + \sqrt{\tau_1} \right) \leq C(\sqrt{\delta} + K_*^{r-1}) K^{1-3r}. \quad (6.28)$$

Moreover, a straightforward application of (6.25) yields

$$\mathbf{E} \sup_{0 \leq t \leq \tau_1} |X_{t,n-1} - x_{t,n-1}| \leq C\sqrt{\delta}K^{-2r}. \quad (6.29)$$

Let

$$\begin{aligned} \Omega_1 &= \left\{ \omega \in \Omega : \sum_{j=3}^n \int_0^{\tau_1} |x_{s,j}|^2 ds \leq \sqrt{\delta}K^{-2r} \right\}, \\ \Omega_2 &= \left\{ \omega \in \Omega : \sup_{0 \leq t \leq \tau_1} |x_{0,1} - x_{t,1}| \leq K^{1-3r} \right\}, \\ \Omega_3 &= \left\{ \omega \in \Omega : \sup_{0 \leq t \leq \tau_1} |X_{t,n-1} - x_{t,n-1}| \leq \delta^{1/4}K^{-2r} \right\}, \\ \tilde{\Omega} &= \Omega_1 \cap \Omega_2 \cap \Omega_3. \end{aligned}$$

By (6.25), (6.28), and (6.29) for  $\delta$  sufficiently small and  $K_*$  sufficiently large we have  $\mathbf{P}(\tilde{\Omega}) \geq 1 - \alpha/2$ . Now, by (6.27) and  $\tau_1 = K^{-4r}$ , for  $\omega \in \tilde{\Omega}$  we have

$$\sup_{0 \leq t \leq \tau_1} |X_{t,n} - x_{t,n}| \leq C\delta^{1/4}K^{1-6r} = C\delta^{1/4}K^r,$$

and hence

$$\mathbf{P} \left( \int_0^{\tau_1} |X_{t,n} - x_{t,n}| dt \leq \tau_1 C\delta^{1/4}K^r \right) \geq 1 - \frac{\alpha}{2}. \quad (6.30)$$

Combining (6.24) and (6.30) we see that for  $\delta$  sufficiently small there holds

$$\mathbf{P} \left( \int_0^{\tau_1} |x_{t,n}| dt \geq \frac{\tau_1}{4} K^r \right) \geq \frac{\alpha}{2}.$$

This is enough to yield a contradiction for  $\delta$  sufficiently small.

**Case 2** ( $x_0 \in B_{K,\delta}^2$ ): Consider the approximate solution  $X_t$  defined simply by

$$\begin{cases} dX_{t,3} = -X_{t,1}X_{t,2}dt \\ dX_{t,j} = 0 & j \neq 3 \end{cases} \quad (6.31)$$

with initial condition  $X_0 = x_0$ . We have then

$$X_{t,3} = X_{0,3} - tX_{0,1}X_{0,2}$$

so that the bounds on  $x_0 \in B_{K,\delta}^2$  imply

$$\frac{1}{\tau_2} \int_0^{\tau_2} |X_{t,3}| dt \geq \frac{\tau_2}{64} \delta_1 K^{1+r} - \delta K^r = \left( \frac{\delta_1}{64} - \delta \right) K^r.$$

Taking  $\delta \ll \delta_1$  yields

$$\frac{1}{\tau_2} \int_0^{\tau_2} |X_{t,3}| dt \geq \frac{\delta_1}{128} K^r. \quad (6.32)$$

The error satisfies

$$d(X_{t,3} - x_{t,3}) = (x_{t,1} - x_{0,1})x_{t,2}dt + x_{0,1}(x_{t,2} - x_{0,2})dt + a_3x_{t,3}dt - x_{t,4}x_{t,2}dt - \sigma_3 dW_t^{(3)}. \quad (6.33)$$

Supposing for contradiction that

$$\mathbf{E} \int_0^{\tau_2} |\Pi_{\ker A^\perp} x_t|^2 dt = \mathbf{E} \sum_{j=3}^n \int_0^{\tau_2} |x_{t,j}|^2 dt \leq \delta \tau_2 K^{2r} \quad (6.34)$$

we easily derive

$$\mathbf{E} \sup_{0 \leq t \leq \tau_2} |x_{t,1} - x_{0,1}| + \mathbf{E} \sup_{0 \leq t \leq \tau_2} |x_{t,2} - x_{0,2}| \leq C(\sqrt{\delta} \tau_2 K^{1+r} + \sqrt{\tau_2}) \leq C \max(\sqrt{\delta}, K_*^{-1/2}) K^r.$$

Therefore, choosing  $K_* = \delta^{-1}$  and defining

$$\begin{aligned} \Omega_1 &= \{\omega \in \Omega : \sup_{0 \leq t \leq \tau_2} |x_{t,1} - x_{0,1}| + \sup_{0 \leq t \leq \tau_2} |x_{t,2} - x_{0,2}| \leq \delta^{1/4} K^r\}, \\ \Omega_2 &= \{\omega \in \Omega : \sup_{0 \leq t \leq \tau_2} |\sigma_3| |W_t^{(3)}| \leq \delta^{1/4} K^r\}, \\ \Omega_3 &= \{\omega \in \Omega : \sum_{j=3}^n \int_0^{\tau_2} |x_{t,j}|^2 dt \leq \sqrt{\delta} \tau_2 K^{2r}\}, \end{aligned}$$

we have  $\mathbf{P}(\Omega_1 \cap \Omega_2 \cap \Omega_3) \geq 1/2$  for  $\delta$  taken sufficiently small. Let  $\omega \in \Omega_1 \cap \Omega_2 \cap \Omega_3$ . Returning to (6.33) we obtain

$$\sup_{0 \leq t \leq \tau_2} |X_{t,3}(\omega) - x_{t,3}(\omega)| \leq C \delta^{1/4} K^r. \quad (6.35)$$

Hence,

$$\mathbf{P} \left( \int_0^{\tau_2} |X_{t,3} - x_{t,3}| dt \leq C \tau_2 \delta^{1/4} K^r \right) \geq \frac{1}{2}. \quad (6.36)$$

By choosing  $\delta \ll \delta_1^4$ , (6.36) and (6.32) combined are enough to yield a contradiction.

**Case 3** ( $x_0 \in B_{K,\delta}^3$ ): Now we turn to the final case. Let  $X_t$  be as given in Lemma 6.3 and define  $\bar{x}_{t,j} = X_{t,j} - x_{t,j}$ . Observe that

$$\begin{cases} d\bar{x}_{t,n} = x_{0,n-1} \bar{x}_{t,1} dt + x_{t,1}(x_{0,n-1} - x_{t,n-1}) dt + a_n x_{t,n} + x_{t,n-2} x_{t,n-1} dt \\ d\bar{x}_{t,1} = x_{0,2} \bar{x}_{t,n} dt + x_{t,n}(x_{0,2} - x_{t,2}) dt + x_{t,n-1} x_{t,n} dt - \sigma_1 dW_t^{(1)}. \end{cases} \quad (6.37)$$

Let

$$\begin{aligned} F(t) &= x_{t,1}(x_{0,n-1} - x_{t,n-1}) + a_n x_{t,n} + x_{t,n-2} x_{t,n-1}, \\ G(t) &= x_{t,n}(x_{0,2} - x_{t,2}) + x_{t,n-1} x_{t,n}, \end{aligned}$$

and  $S(t)$  be the group generated by the corresponding (constant) linearization matrix:

$$S(t) := \exp \left( t \begin{pmatrix} 0 & x_{0,n-1} \\ x_{0,2} & 0 \end{pmatrix} \right).$$

Then, we have

$$\begin{pmatrix} \bar{x}_{t,n} \\ \bar{x}_{t,1} \end{pmatrix} = \int_0^t S(t-s) \begin{pmatrix} F(s) \\ G(s) \end{pmatrix} ds - \int_0^t S(t-s) \begin{pmatrix} 0 \\ \sigma_1 dW_s^{(1)} \end{pmatrix}. \quad (6.38)$$

Note that since  $\tau_3 K^{(1+r)/2} \lesssim 1$ , for any  $s \leq t \leq \tau_3$  there holds

$$\|S(t-s)\| \leq \exp \left( (t-s) \sqrt{|x_{0,n-1}| |x_{0,2}|} \right) \leq \exp \left( (t-s) C K^{(1+r)/2} \right) \lesssim 1. \quad (6.39)$$

Thus,

$$\mathbf{E} \sup_{0 \leq t \leq \tau_3} \left| \int_0^t S(t-s) \begin{pmatrix} 0 \\ \sigma_1 dW_s^{(1)} \end{pmatrix} \right|^2 \lesssim \tau_3 = K^{-4r}. \quad (6.40)$$

Suppose now for the sake of contradiction that

$$\mathbf{E} \int_0^{\tau_3} \sum_{j=3}^n |x_{t,j}|^2 dt \leq \delta \tau_3 K^{2r}. \quad (6.41)$$

For  $R \geq 1$ , let  $\Omega_0 \subseteq \Omega$  be the set where the following bounds hold:

$$\begin{aligned} \sup_{0 \leq t \leq \tau_3} \left| \int_0^t S(t-s) \begin{pmatrix} 0 \\ \sigma_1 dW_s^{(1)} \end{pmatrix} \right| &\leq R\sqrt{\tau_3}, \\ \int_0^{\tau_3} \sum_{j=3}^n |x_{t,j}|^2 dt &\leq \sqrt{\delta} \tau_3 K^{2r}, \\ \sup_{0 \leq t \leq \tau_3} |W_t| &\leq R\sqrt{\tau_3}, \\ \sup_{0 \leq t \leq \tau_3} |x_{t,1}| &\leq (\delta_1^{1/4} + \delta^{1/4}) K^{4r} + 1. \end{aligned}$$

By (6.40), (6.41), and

$$\mathbf{E} \sup_{0 \leq t \leq \tau_3} |x_{t,1}| \lesssim \delta_1 K^r + \sqrt{\delta} \tau_3 K^{1+r} + \sqrt{\tau_3} \lesssim (\delta_1 + \sqrt{\delta}) K^{4r} + K^{-2r},$$

for  $R, K$  sufficiently large and  $\delta, \delta_1$  sufficiently small we have  $\mathbf{P}(\Omega_0) \geq 1 - \beta/2$ , where  $\beta$  is as given in Lemma 6.3. Note that for  $\omega \in \Omega_0$  we have the additional estimates

$$\sup_{0 \leq t \leq \tau_3} |x_{t,n-1} - x_{0,n-1}| \leq CRK^{-2r}, \quad (6.42)$$

$$\sup_{0 \leq t \leq \tau_3} |x_{t,2} - x_{0,2}| \leq CR(\delta^{1/4} K^r + K^{-2r}). \quad (6.43)$$

Moreover, by (6.42), (6.43), (6.38), and (6.39) for  $\omega \in \Omega_0$  there holds

$$\sup_{0 \leq t \leq \tau_3} |\bar{x}_{t,n}| + \sup_{0 \leq t \leq \tau_3} |\bar{x}_{t,1}| \leq CR. \quad (6.44)$$

Observe now that

$$d(X_{t,3} - x_{t,3}) = -\bar{x}_{t,1} x_{t,2} dt + x_{t,1} (x_{t,2} - x_{0,2}) dt + \bar{x}_{t,1} (x_{t,2} - x_{0,2}) dt - x_{t,4} x_{t,2} dt + a_3 x_{t,3} - \sigma_3 dW_t^{(3)}, \quad (6.45)$$

which together with the estimates above gives, for  $\omega \in \Omega_0$ ,

$$\begin{aligned} \sup_{0 \leq t \leq \tau_3} |X_{t,3} - x_{t,3}| &\leq \tau_3 \sup_{0 \leq t \leq \tau_3} (|\bar{x}_{t,1} x_{t,2}| + |x_{t,1}| |x_{t,2} - x_{0,2}| + |\bar{x}_{t,1}| |x_{t,2} - x_{0,2}|) \\ &\quad + \int_0^{\tau_3} (|x_{t,4} x_{t,2}| + |a_3| |x_{t,3}|) dt + \sup_{0 \leq t \leq \tau_3} |\sigma_3 W_t^{(3)}| \\ &\leq CK^{4r} (\delta^{1/4} + R^2 K_*^{-r}). \end{aligned}$$

This error estimate (with  $\delta$  taken sufficiently small and  $K_*$  taken sufficiently large), together with Lemma 6.3 on the growth of the approximate solution, allows us to obtain a contradiction as in our earlier arguments. This completes the proof of Proposition 6.2 (and hence also of Theorem 6.1).  $\square$

## 7 Stochastic triad model with non-trivial, invariant conservative dynamics in the kernel

In this section we prove Theorem 1.6. It is sufficient to prove the result after rotating coordinates so that  $\ker A = \{x_2 = 0\}$ . In these new coordinates, the nonlinearity becomes

$$B(x, y) = \begin{pmatrix} x_1 y_3 \\ -x_2 y_3 \\ (x_2 - x_1)(y_2 + y_1) \end{pmatrix}. \quad (7.1)$$

Henceforth in this section,  $x_t$  denotes the solution to (1.1) with  $n = 3$ ,  $B$  given by (7.1), the non-negative definite matrix  $A$  such that  $\ker A = \{x_2 = 0\}$ , and  $\sigma \in \mathbb{R}^{3 \times 3}$  satisfying  $\text{rank}(\sigma) = 3$ .

The dynamical system  $\dot{x} = B(x, x)$ , with  $B$  given by (7.1), has equilibria at  $(0, 0, \pm a)$  for any  $a > 0$  and the stable/unstable manifold of each fixed point is joined to the other via a heteroclinic connection. The unstable manifold of  $(0, 0, a)$  is tangent to  $\ker A$  and the associated heteroclinic connections with the stable manifold of  $(0, 0, -a)$  lie entirely in  $\ker A$ . The present example thus distinguishes itself from the previous ones in that there exist nontrivial conservative dynamics in  $\ker A$ .

As in the earlier examples, our plan to prove Theorem 1.6 is to show that the Markov semigroup generated by (1.1) satisfies Assumption 1. Again as before, we will deduce the growth required by (2.13) by establishing it instead for a suitable approximate solution. The idea is to study the linearization of  $\Pi_{\ker A} B(\cdot, \cdot)$  around  $Z_t = (X_{t,1}, 0, X_{t,3}) \in \ker A$  solving

$$\begin{cases} \frac{d}{dt} X_{t,1} = X_{t,1} X_{t,3} \\ \frac{d}{dt} X_{t,3} = -X_{t,1}^2 \\ (X_{0,1}, 0, X_{0,3}) = \Pi_{\ker A} x_0. \end{cases} \quad (7.2)$$

Since  $(0, 0, -a)$  attracts all points on the circle  $\{(x_1, 0, x_3) \in \mathbb{R}^3 : x_1^2 + x_3^2 = a^2\}$  except  $(0, 0, a)$  and has an unstable manifold perpendicular to  $\ker A$ , one expects this linearization to grow exponentially fast provided the noise has a nonzero projection onto both  $(1, 0, 0)$  and  $(0, 1, 0)$ . Besides arguments analogous to those in previous sections used to study the linearization around (7.2), we construct a local Lyapunov function to estimate exit times of the process from the vicinity of the unstable fixed points  $(0, 0, a)$ .

### 7.1 Local Lyapunov function

**Lemma 7.1.** *For  $K \geq 1$ , let*

$$\mathcal{B}_K = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : K/2 \leq x_3 \leq 2K, \quad |x_1| \leq K, \quad |x_2| \leq K^{1/4}\}.$$

*There exists  $\gamma \in (0, 1)$  such that for all  $K$  sufficiently large there is a smooth, strictly positive function  $V_K : \mathbb{R}^3 \rightarrow \mathbb{R}$  such that for  $x \in \mathcal{B}_K$ ,*

$$\mathcal{L}V_K \leq -\gamma K V_K \quad (7.3)$$

*and*

$$\gamma K^{-1} \leq V_K \leq \gamma^{-1} \sqrt{K}. \quad (7.4)$$

*Specifically, for some  $R \geq 1$  sufficiently large,*

$$V_K = \frac{1}{|x_1|} \chi_T(x_1) + \sqrt{K} \left(1 - \frac{K}{32} \frac{|x_1|^2}{R^2}\right) \chi_D(x_1),$$

*where for an arbitrary smooth cutoff  $\varphi : [0, \infty) \rightarrow [0, 1]$  with  $\varphi(y) = 1$  for  $y \leq 1/2$ ,  $\varphi(y) = 0$  for  $y \geq 1$ , and  $\varphi'(y) \leq 0$ , we define*

$$\chi_D(x_1) = \varphi\left(\frac{\sqrt{K}|x_1|}{4R}\right), \quad \chi_T(x_1) = 1 - \varphi\left(\frac{\sqrt{K}|x_1|}{R}\right).$$

**Remark 15.** The cutoff  $\chi_T$  refers to ‘transport’ as it is in the region  $|x_1| \gtrsim K^{-1/2}$  wherein the conservative dynamics (i.e. the first order terms in the generator) will be the most significant. The cutoff  $\chi_D$  refers to ‘diffusive’, as it is in the region  $|x_1| \ll K^{-1/2}$  in which the noise (i.e. the second order terms in the generator) will be dominant.

*Proof.* Since  $\ker A = \{x_2 = 0\}$ , there exist  $a_1, a_2, a_3 \in \mathbb{R}$  such that

$$Ax \cdot \nabla = x_2 \sum_{j=1}^3 a_j \partial_{x_j}.$$

Defining  $\Lambda = \sigma \sigma^T$ , we can thus write the generator as

$$\mathcal{L} = \frac{1}{2} \sum_{i,j=1}^3 \Lambda_{ij} \partial_{x_i} \partial_{x_j} + x_1 x_3 \partial_{x_1} - x_2 x_3 \partial_{x_2} + (x_2^2 - x_1^2) \partial_{x_3} - x_2 \sum_{j=1}^3 a_j \partial_{x_j}. \quad (7.5)$$

Note that  $\Lambda_{11} > 0$  since  $\sigma$  is assumed full rank. Let  $\varphi : [0, \infty) \rightarrow [0, 1]$  be a smooth cutoff with  $\varphi(y) = 1$  for  $y \leq 1/2$ ,  $\varphi(y) = 0$  for  $y \geq 1$  and  $\varphi'(y) \leq 0$ . For  $R \geq 1$  to be chosen sufficiently large independently of  $K$ , define

$$\chi_T(x_1) = 1 - \varphi\left(\frac{\sqrt{K}|x_1|}{R}\right), \quad V_{K,T}(x_1) = \chi_T(x_1)|x_1|^{-1}.$$

For  $x \in \mathcal{B}_K$  we compute

$$\begin{aligned} \mathcal{L}V_{K,T} &= -x_3 V_{K,T} + \chi_T \frac{\Lambda_{11}}{|x_1|^3} + \chi_T \frac{a_1 x_1 x_2}{|x_1|^3} \\ &\quad \left( -\Lambda_{11} \frac{x_1}{|x_1|^3} + \frac{x_1 x_3}{|x_1|} - \frac{x_2 a_1}{|x_1|} \right) \partial_{x_1} \chi_T + \frac{\Lambda_{11}}{2|x_1|} \partial_{x_1}^2 \chi_T \\ &\leq -\frac{K}{2} V_{K,T} + \left( \frac{\Lambda_{11}}{|x_1|^2} + \frac{|a_1||x_2|}{|x_1|} \right) V_{K,T} \\ &\quad + C \frac{\sqrt{K}}{R} \left( \frac{1}{|x_1|^2} + |x_3| + \frac{|x_2|}{|x_1|} \right) \mathbf{1}_{RK^{-1/2}/2 \leq |x_1| \leq RK^{-1/2}} \\ &\quad + C \frac{K}{R^2} \frac{1}{|x_1|} \mathbf{1}_{RK^{-1/2}/2 \leq |x_1| \leq RK^{-1/2}}, \end{aligned}$$

where  $C$  is a constant that depends only on  $A$ ,  $\sigma$ , and the choice of cutoff  $\varphi$ . We will continue to denote by  $C$  such a constant, though it may change line-to-line. From the support properties of  $\chi_T$  and  $x \in \mathcal{B}_K$  we then obtain that for  $R$  large depending only on  $\Lambda_{11}$  and  $|a_1|$  there holds

$$\mathcal{L}V_{K,T} \leq -\frac{K}{4} V_{K,T} + C \frac{K^{3/2}}{R} \mathbf{1}_{RK^{-1/2}/2 \leq |x_1| \leq RK^{-1/2}}. \quad (7.6)$$

Now define

$$\chi_D(x_1) = \varphi\left(\frac{\sqrt{K}|x_1|}{4R}\right), \quad V_{K,D}(x_1) = \chi_D(x_1) \sqrt{K} \left(1 - \frac{K}{32} \frac{|x_1|^2}{R^2}\right)$$

and note that

$$\frac{\sqrt{K}}{2} \chi_D \leq V_{K,D} \leq \sqrt{K} \chi_D.$$

For  $x \in \mathcal{B}_K$  we now compute

$$\begin{aligned}
\mathcal{L}V_{K,D} &= \frac{K^{3/2}\chi_D}{16R^2} \left( -\frac{\Lambda_{11}}{2} - |x_1|^2 x_3 + a_1 x_1 x_2 \right) \\
&\quad + \partial_{x_1} \chi_D \left( -\Lambda_{11} \frac{K^{3/2}}{16R^2} x_1 + \sqrt{K} \left( 1 - \frac{K|x_1|^2}{32R^2} \right) (x_1 x_3 - a_1 x_2) \right) \\
&\quad + \partial_{x_1}^2 \chi_D \frac{\Lambda_{11}}{2} \sqrt{K} \left( 1 - \frac{K|x_1|^2}{32R^2} \right) \\
&\leq -\frac{K\Lambda_{11}}{32R^2} V_{K,D} - \frac{K^{5/2}\chi_D}{32R^2} |x_1|^2 + C \frac{K^{5/4}}{R} \chi_D \\
&\quad + C \frac{\sqrt{K}}{R} \left( \frac{K^{3/2}}{R^2} |x_1| + \sqrt{K} |x_2| \right) \mathbf{1}_{2RK^{-1/2} \leq |x_1| \leq 4RK^{-1/2}} \\
&\quad + C \frac{K^{3/2}}{R^2} \mathbf{1}_{2RK^{-1/2} \leq |x_1| \leq 4RK^{-1/2}},
\end{aligned}$$

where in the inequality we noted that  $x_1 x_3 \partial_{x_1} \chi_D \leq 0$  for  $x \in \mathcal{B}_K$ . Taking  $K$  large enough so that

$$C \frac{K^{5/4}}{R} \leq \frac{\Lambda_{11} K^{3/2}}{128R^2}$$

it follows that

$$\mathcal{L}V_{K,D} \leq -\frac{K\Lambda_{11}}{64R^2} V_{K,D} - \frac{K^{5/2}\chi_D}{32R^2} |x_1|^2 + C \left( \frac{K^{3/2}}{R^2} + \frac{K^{5/4}}{R} \right) \mathbf{1}_{2RK^{-1/2} \leq |x_1| \leq 4RK^{-1/2}}. \quad (7.7)$$

The plan is now to add (7.6) and (7.7). Upon doing this, for  $K$  and  $R$  sufficiently large the second term in (7.7) absorbs the second term in (7.6) and the first term in (7.6) absorbs the third term in (7.7). In particular, defining

$$V_K = V_{K,D} + V_{K,T}$$

we have

$$\mathcal{L}V_K \leq -K \min \left( \frac{1}{8}, \frac{\Lambda_{11}}{64R^2} \right) V_K. \quad (7.8)$$

This completes the proof.  $\square$

The next lemma uses Lemma 7.1 to obtain estimates on the exit times from neighborhoods of the north pole equilibria  $x_1 = x_2 = 0, x_3 > 0$ .

**Lemma 7.2.** *Let  $x_0 \in \mathbb{R}^3$  satisfy*

$$|x_0| = K, \quad x_{0,3} > 0, \quad |x_{0,1}| < \delta K, \quad \text{and} \quad |x_{0,2}| < \delta K^{1/4}$$

for  $K \geq 1$  and  $\delta \in (0, 1)$ . Define the stopping time

$$\tau(\omega) = \inf \{ t \geq 0 : |x_{t,1}(\omega)| \geq \delta K, |x_{t,2}(\omega)| \leq K^{1/4} \}.$$

There exists  $C_0 \geq 1$  so that for all  $K$  sufficiently large and  $\delta$  sufficiently small there holds

$$\mathbf{P} \left( \tau \leq \frac{C_0 \log K}{K} \right) \geq \frac{1}{2}. \quad (7.9)$$

*Proof.* First note that taking at least  $\delta^2 \leq 7/32$  gives  $3K/4 \leq x_{0,3} \leq K$ , so we may assume that  $x_0 \in \mathcal{B}_K$  as defined in Lemma 7.1. Define  $\mathcal{B}_{K,\delta} \subseteq \mathcal{B}_K$  by

$$\mathcal{B}_{K,\delta} = \{x \in \mathbb{R}^3 : K/2 < x_3 < 2K, \quad |x_1| < \delta K, \quad |x_2| < K^{1/4}\}$$

and let

$$\tilde{\tau}(\omega) = \inf\{t \geq 0 : x_t(\omega) \in \mathcal{B}_{K,\delta}^c\}.$$

Let  $V_K$  be as given in Lemma 7.1. By Dynkin's formula, for any  $t \geq 0$  there holds

$$\mathbf{E}e^{\gamma K t \wedge \tilde{\tau}} V_K(x_{t \wedge \tilde{\tau}}) \leq V_K(x_0) + \mathbf{E} \int_0^{t \wedge \tilde{\tau}} e^{\gamma K s} (\mathcal{L} + \gamma K) V_K(x_s) ds. \quad (7.10)$$

Since  $x_s(\omega) \in \mathcal{B}_K$  for  $s \leq \tilde{\tau}(\omega)$ , it follows from Lemma 7.1 that

$$\mathbf{E}e^{\gamma K t \wedge \tilde{\tau}} \leq \gamma^{-2} K^{3/2}. \quad (7.11)$$

From (7.11) and Chebyshev's inequality we obtain, for any  $C_0 > 0$ ,

$$\mathbf{P}\left(\tilde{\tau} \geq \frac{C_0 \log K}{K}\right) \leq \gamma^{-2} K^{3/2 - \gamma C_0}. \quad (7.12)$$

Hence, for  $C_0 \geq 2/\gamma$  and  $K$  sufficiently large there holds

$$\mathbf{P}\left(\tilde{\tau} \leq \frac{C_0 \log K}{K}\right) \geq \frac{3}{4}. \quad (7.13)$$

Now set  $t_* = C_0 \log(K)/K$ . By (7.13) and the definitions of  $\tilde{\tau}$  and  $\tau$  we have

$$\begin{aligned} \mathbf{P}(\tau \leq t_*) &\geq \frac{3}{4} - \mathbf{P}\left(\{|x_{\tilde{\tau},2}| \geq K^{1/4}\} \cap \{\tilde{\tau} \leq t_*\}\right) \\ &\quad - \mathbf{P}\left(\{x_{\tilde{\tau},3} \geq 2K \text{ or } x_{\tilde{\tau},3} \leq K/2\} \cap \{\tilde{\tau} \leq t_*\}\right), \end{aligned}$$

and so to complete the proof it suffices to show that

$$\mathbf{P}\left(\{|x_{\tilde{\tau},2}| \geq K^{1/4}\} \cap \{\tilde{\tau} \leq t_*\}\right) + \mathbf{P}\left(\{x_{\tilde{\tau},3} \geq 2K \text{ or } x_{\tilde{\tau},3} \leq K/2\} \cap \{\tilde{\tau} \leq t_*\}\right) \leq \frac{1}{4}. \quad (7.14)$$

To bound the first term we begin by using Dynkin's formula to obtain

$$\mathbf{E}|x_{t_* \wedge \tilde{\tau},2}|^2 = |x_{0,2}|^2 + \mathbf{E} \int_0^{t_* \wedge \tilde{\tau}} (\Lambda_{22} - 2x_{s,3}|x_{s,2}|^2 - 2a_2|x_{s,2}|^2) ds. \quad (7.15)$$

Since  $a_2 \geq 0$  and  $x_{s,3} > 0$  for  $s \leq \tilde{\tau}$  it follows that

$$\mathbf{E}|x_{t_* \wedge \tilde{\tau},2}|^2 \leq |x_{0,2}|^2 + \Lambda_{22} t_* \leq \delta^2 \sqrt{K} + \Lambda_{22} t_*.$$

Thus,

$$\mathbf{P}\left(\{|x_{\tilde{\tau},2}| \geq K^{1/4}\} \cap \{\tilde{\tau} \leq t_*\}\right) \leq \frac{1}{\sqrt{K}} \mathbf{E}|x_{t_* \wedge \tilde{\tau},2}|^2 \leq \delta^2 + CK^{-1/2}, \quad (7.16)$$

which implies

$$\mathbf{P}\left(\{|x_{\tilde{\tau},2}| \geq K^{1/4}\} \cap \{\tilde{\tau} \leq t_*\}\right) \leq \frac{1}{8} \quad (7.17)$$



for  $\delta$  sufficiently small and  $K$  sufficiently large. To bound the second term in (7.14), observe that for  $K \geq 8$  and  $\delta^2 \leq 1/8$  we have

$$\mathbf{P}(\{x_{\tilde{\tau},3} \geq 2K \text{ or } x_{\tilde{\tau},3} \leq K/2\} \cap \{\tilde{\tau} \leq t_*\}) \leq \mathbf{P}\left(\sup_{0 \leq t \leq t_*} ||x_t|^2 - K^2| \geq K^2/2\right).$$

Now, for  $K$  sufficiently large, Lemma A.1 implies

$$\mathbf{P}\left(\sup_{0 \leq t \leq t_*} ||x_t|^2 - K^2| \geq K^2/2\right) \leq \frac{1}{8},$$

which completes the proof.  $\square$

## 7.2 Growth of approximate solution

The next lemma gives the growth of an approximate solution for initial conditions that are in the vicinity of  $\ker A$  but not too close to the north pole equilibria.

**Lemma 7.3.** Fix  $r, \epsilon, \delta \in (0, 1/4)$  and let  $X_t$  solve

$$\begin{cases} dX_{t,1} = X_{t,1}X_{t,3}dt \\ dX_{t,2} = -X_{t,2}X_{t,3}dt + \sum_{j=1}^3 \sigma_{2j}dW_t^{(j)} \\ dX_{t,3} = -X_{t,1}^2dt \end{cases} \quad (7.18)$$

with an initial condition  $X_0 \in \mathbb{R}^3$  that satisfies

$$|X_{0,2}| \leq (|X_{0,1}|^2 + |X_{0,3}|^2)^{r/2}$$

and at least one the bounds

$$X_{0,3} \leq 0 \quad \text{or} \quad |X_{0,1}| \geq \delta|X_0|. \quad (7.19)$$

There exist  $K_*(\epsilon) \geq 1$ ,  $C_0(\epsilon, \delta) \geq 1$ ,  $c_*(\epsilon, \delta) > 0$ , and  $\beta > 0$  so that for  $|X_0| \geq K_*$  and

$$\tau = \frac{C_0(\epsilon, \delta)}{|X_0|} + \frac{(1/2 + r + \epsilon) \log(|X_0|)}{(1 - \epsilon)|X_0|}$$

there holds

$$\mathbf{P}\left(\frac{1}{\tau} \int_0^\tau |X_{t,2}|dt \geq c_*|X_0|^r\right) \geq \beta. \quad (7.20)$$

**Remark 16.** Observe that the dynamics of  $Z_t := (X_{t,1}, 0, X_{t,3}) \in \ker A$  is decoupled from  $X_{t,2}$  and satisfies

$$|Z_t|^2 = |Z_0|^2 \quad (7.21)$$

for all  $t \geq 0$ .

*Proof.* By (7.21) and  $|X_{0,2}| \leq |Z_0|^r$ , for  $K_*(\epsilon)$  sufficiently large we have  $|Z_t| \geq (1 - \epsilon/2)|X_0|$  for all  $t \geq 0$ . It follows then by (7.19) and rescaling  $Z_t$  back to the unit circle that there exists  $C_0(\epsilon, \delta)$  such that for  $\tau_1 = C_0/|X_0|$  there holds

$$X_{t,3} \leq -(1 - \epsilon)|X_0| \quad (7.22)$$

for all  $t \geq \tau_1$ . Now, we may assume without loss of generality that  $\sum_{j=1}^3 \sigma_{2j}dW_t^{(j)} = dW_t$  for a standard Brownian motion  $W_t$ . The formula for  $X_{t,2}$  then reads

$$X_{t,2} = e^{-\int_0^t X_{s,3}ds} X_{0,2} + \int_0^t e^{-\int_s^t X_{s',3}ds'} dW_{s'}. \quad (7.23)$$

Since  $X_{t,3}$  is deterministic, we have that  $\int_0^\tau X_{t,2} dt$  is a Gaussian random variable with variance

$$\begin{aligned} \mathbf{E} \left| \int_0^\tau X_{t,2} dt - \int_0^\tau e^{-\int_0^t X_{s,3} ds} X_{0,2} dt \right|^2 &= \mathbf{E} \left| \int_0^\tau \int_0^t e^{-\int_s^t X_{s',3} ds'} dW_s dt \right|^2 \\ &= \mathbf{E} \int_0^\tau \left| \int_s^\tau e^{-\int_s^t X_{s',3} ds'} dt \right|^2 ds. \end{aligned}$$

Using  $X_{s,3} \leq |X_0|$  for any  $s$  and  $X_{s,3} \leq -(1-\epsilon)|X_0|$  for  $s \geq \tau_1$  we obtain, for  $K_*$  sufficiently large, the lower bound

$$\mathbf{E} \int_0^\tau \left| \int_s^\tau e^{-\int_s^t X_{s',3} ds'} dt \right|^2 ds \gtrsim_{\epsilon, \delta} \frac{\tau}{|X_0|^2} e^{2(1-\epsilon)|X_0|(\tau-\tau_1)} = \tau |X_0|^{2r+2\epsilon-1}. \quad (7.24)$$

It follows that there are  $c_*(\epsilon, \delta), \beta > 0$  sufficiently small so that

$$\mathbf{P} \left( \frac{1}{\tau} \int_0^\tau |X_{t,2}| dt \geq c_* \frac{|X_0|^{r+\epsilon}}{\sqrt{\log |X_0|}} \right) \geq \beta. \quad (7.25)$$

This completes the proof for  $K_*(\epsilon)$  sufficiently large.  $\square$

### 7.3 Justifying the approximation

Theorem 1.6 follows immediately from Lemma 2.2 and the following proposition.

**Proposition 7.4.** Fix  $r \in (0, 1/4)$  and for  $\delta \in (0, 1/4)$  define

$$B_K^\delta = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : |x_2|^2 \leq \delta(|x_1|^2 + |x_3|^2)^r, (1-\delta)K^2 \leq |x|^2 \leq (1+\delta)K^2\}.$$

For  $\delta$  sufficiently small there exist  $\eta(K) \approx \log(K)/K$ ,  $c_* > 0$ , and  $K_* \geq 1$  so that for any  $K \geq K_*$  large there holds

$$x_0 \in B_K^\delta \implies \frac{1}{\eta(K)} \mathbf{E} \int_0^{\eta(K)} |x_{t,2}|^2 dt \geq c_* K^{2r}. \quad (7.26)$$

*Proof.* Let

$$S_{\delta_1} = \{x \in \mathbb{R}^3 : x_3 > 0, |x_1| < \delta_1 |x|\}$$

and

$$B_K^{\delta, \delta_1} = S_{\delta_1}^c \cap \{x \in \mathbb{R}^3 : |x_2|^2 \leq (|x_1|^2 + |x_3|^2)^r, (1-2\delta)K^2 \leq |x|^2 \leq (1+2\delta)K^2\}.$$

We first show that it is sufficient to prove that if  $\delta$  sufficiently small then for every  $\delta_1 \in (0, 1/4)$  there is  $\eta_1(K) \approx_{\delta_1} \log(K)/K$  and  $c > 0$  so that for  $K$  sufficiently large there holds

$$x_0 \in B_K^{\delta, \delta_1} \implies \frac{1}{\eta_1(K)} \int_0^{\eta_1(K)} \mathcal{P}_t D(x_0) dt \geq c K^{2r}, \quad (7.27)$$

where  $D : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  is defined by  $D(x) = |x_2|^2$ . Fix  $x_0 \in B_K^\delta$  and let

$$\tilde{\tau}(\omega) = \inf \left\{ t \geq 0 : x_t(\omega) \in B_K^{\delta, \delta_1} \right\}.$$

By Lemma 7.2 and Lemma A.1, there are  $\delta_1, C_1 > 0$  so that for all  $\delta$  small enough, if  $K$  is taken sufficiently large depending on  $\delta$  there holds

$$\mathbf{P} \left( \tilde{\tau}_1 \leq \frac{C_1 \log(K)}{K} \right) \geq \frac{1}{4}. \quad (7.28)$$

Let now  $\eta_1$  and  $c$  be as in (7.27) applied with  $\delta_1$  chosen so that (7.28) holds. Define

$$\eta(K) = \eta_1(K) + \frac{C_1 \log(K)}{K}.$$

By (7.28), the strong Markov property and (7.27), for  $K$  sufficiently large we have

$$\begin{aligned} \mathbf{E} \int_0^\eta |x_{t,2}|^2 dt &\geq \int_\Omega \int_0^{\eta - \bar{\tau} \wedge \eta} D(x_{\bar{\tau} \wedge \eta + t}(\omega)) dt d\mathbf{P} \\ &\geq \frac{1}{4} \inf_{\bar{\tau}(\omega) \leq C_1 \log(K)/K} \int_0^{\eta_1} \mathcal{P}_t D(x_{\bar{\tau}}(\omega)) dt \\ &\geq \frac{c}{4} \eta_1(K) K^{2r}. \end{aligned}$$

The bound (7.26) then follows since  $\eta_1(K) \approx \eta(K)$ .

We now prove (7.27). Let  $x_0 \in B_K^{\delta, \delta_1}$  for some  $\delta_1 \in (0, 1/4)$ . For  $\epsilon \in (0, 1/4)$  to be chosen later, define

$$\tau = \frac{C_0(\epsilon, \delta_1)}{|x_0|} + \frac{(1/2 + r + \epsilon) \log(|x_0|)}{(1 - \epsilon)|x_0|},$$

where  $C_0$  is as defined in Lemma 7.3. Suppose now for contradiction that

$$\frac{1}{\tau} \mathbf{E} \int_0^\tau |x_{t,2}|^2 dt \leq \delta_2 K^{2r} \quad (7.29)$$

for some  $\delta_2 \in (0, 1)$  to be chosen sufficiently small. Let  $X_t$  solve (7.18) with initial condition  $x_0$ . By Lemma 7.3 there exists  $c_*(\epsilon, \delta_1), \beta > 0$  so that for  $K$  sufficiently large (depending on  $\epsilon$ ) there holds

$$\mathbf{P} \left( \frac{1}{\tau} \int_0^\tau |X_{t,2}| dt \geq c_* |x_0|^r \right) \geq \beta \quad (7.30)$$

and  $X_{t,3} \leq -(1 - \epsilon)|x_0|$  for all  $t \geq \tau_1 := C_0/|x_0|$ . As in our earlier proofs, the plan is now to proceed by deriving suitable estimates on the error  $|X_{t,2} - x_{t,2}|$ .

We begin by estimating  $|\Pi_{\ker A}(X_t - x_t)|$ . This is slightly more involved than in earlier arguments since  $\Pi_{\ker A} X_t$  is not constant. We denote  $Z_t = (Z_{t,1}, 0, Z_{t,3}) = (X_{t,1}, 0, X_{t,3}) \in \ker A$  and  $Y_t = X_{t,2} \in \ker A^\perp$  and define  $z_t$  and  $y_t$  similarly. Moreover, for  $z \in \ker A$  we define the linear operator  $\tilde{L}_z : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  by

$$\tilde{L}_z(x) = \Pi_{\ker A}(B(z, \Pi_{\ker A} x) + B(\Pi_{\ker A} x, z)) = \Pi_{\ker A} L_z \Pi_{\ker A}.$$

The error  $\bar{z}_t = Z_t - z_t$  then solves

$$d\bar{z}_t = \tilde{L}_{Z_t}(\bar{z}_t) dt - B(\bar{z}_t, \bar{z}_t) dt - B(y_t, y_t) dt + \Pi_{\ker A} \sigma dW_t.$$

For  $f : [0, \infty) \rightarrow \ker A$  and  $h \in \mathbb{R}^3$  we write  $S_f(t, s)h$  for the solution to the problem

$$\begin{cases} \frac{d}{dt} S_f h = \tilde{L}_{f(t)} S_f h, & t > s \\ S_f(s, s)h = h. \end{cases}$$

With this notation,  $\bar{z}_t$  satisfies

$$\bar{z}_t = - \int_0^t S_Z(t, s) [B(\bar{z}_s, \bar{z}_s) + B(y_s, y_s)] ds + \int_0^t S_Z(t, s) \Pi_{\ker A} \sigma dW_s.$$

By Hölder's inequality and Fubini's theorem, for any  $T \leq \tau$  we have

$$\begin{aligned}
\|\bar{z}_t\|_{L^2(0,T)} &\lesssim \left( \int_0^T \int_0^t \|S_Z(t,s)\|^2 \left( \|\bar{z}_s\|^2 + \|y_s\|^2 \right) ds dt \right)^{1/2} \left( \|\bar{z}_t\|_{L^2(0,T)} + \|y_t\|_{L^2(0,T)} \right) \\
&\quad + \left( \int_0^\tau \left| \int_0^t S_Z(t,s) \Pi_{\ker A} \sigma dW_s \right|^2 dt \right)^{1/2} \\
&\lesssim \left( \sup_{0 \leq s \leq \tau} \int_0^\tau \mathbf{1}_{s \leq t} \|S_Z(t,s)\|^2 dt \right)^{1/2} \left( \|\bar{z}_t\|_{L^2(0,T)}^2 + \|y_t\|_{L^2(0,T)}^2 \right) \\
&\quad + \left( \int_0^\tau \left| \int_0^t S_Z(t,s) \Pi_{\ker A} \sigma dW_s \right|^2 dt \right)^{1/2}.
\end{aligned} \tag{7.31}$$

To proceed we need estimates for  $\|S_Z(t,s)\|$ . A straightforward computation shows that the top eigenvalue of the symmetric part of  $\tilde{L}_z$  is bounded above by  $(z_3 + |z|)/2$ . Thus, using that  $Z_{t,3} \leq -(1-\epsilon)|x_0|$  for  $t \geq \tau_1$ , we have

$$\|S_Z(t,s)\| \leq \exp \left( \int_s^t \frac{Z_{t',3} + |Z_{t'}|}{2} dt' \right) \leq e^{C_0} e^{(t-\tau_1)\epsilon|x_0|/2}. \tag{7.32}$$

Consequently,

$$\left( \sup_{0 \leq s \leq \tau} \int_0^\tau \mathbf{1}_{s \leq t} \|S_Z(t,s)\|^2 dt \right)^{1/2} \leq e^{C_0} \left( \int_0^\tau e^{(t-\tau_1)\epsilon|x_0|} dt \right)^{1/2} \leq \frac{e^{C_0}}{\sqrt{\epsilon}} |x_0|^{\epsilon-1/2} \tag{7.33}$$

and

$$\mathbf{E} \int_0^\tau \left| \int_0^t S_Z(t,s) \Pi_{\ker A} \sigma dW_s \right|^2 dt \lesssim \int_0^\tau \int_0^t \|S_Z(t,s)\|^2 ds dt \lesssim \tau \frac{e^{2C_0}}{\epsilon} |x_0|^{2\epsilon-1}. \tag{7.34}$$

For  $R \geq 1$ , define

$$\Omega_1 = \left\{ \omega \in \Omega : \left( \int_0^\tau \left| \int_0^t S_Z(t,s) \Pi_{\ker A} \sigma dW_s \right|^2 dt \right)^{1/2} \leq R\sqrt{\tau}|x_0|^{\epsilon-1/2} \right\}$$

and

$$\Omega_2 = \{ \omega \in \Omega : \int_0^\tau |y_t|^2 dt \leq \sqrt{\delta_2} \tau K^{2r} \}.$$

By (7.34) and (7.29), for  $R(\epsilon, \delta_1, \beta)$  sufficiently large and  $\delta_2(\beta)$  sufficiently small there holds

$$\mathbf{P}(\Omega_1 \cap \Omega_2) \geq 1 - \frac{\beta}{2}. \tag{7.35}$$

By (7.33), (7.31), and  $r < 1/4$ , for  $\omega \in \Omega_1 \cap \Omega_2$  there holds

$$\begin{aligned}
\|\bar{z}_t(\omega)\|_{L^2(0,T)} &\lesssim \frac{e^{C_0}}{\sqrt{\epsilon}} K^{\epsilon-1/2} \left( \|\bar{z}_t(\omega)\|_{L^2(0,T)}^2 + \sqrt{\delta_2} \tau K^{2r} \right) + R\sqrt{\tau} K^{\epsilon-1/2} \\
&\lesssim \frac{e^{C_0}}{\sqrt{\epsilon}} K^{\epsilon-1/2} \|\bar{z}_t(\omega)\|_{L^2(0,T)}^2 + R \frac{e^{C_0}}{\sqrt{\epsilon}} \sqrt{\tau} K^{\epsilon-1/2}
\end{aligned}$$

for any  $T \leq \tau$ . From a standard continuity argument, for  $K$  sufficiently large we have

$$\|\bar{z}_t(\omega)\|_{L^2(0,\tau)} \lesssim R \frac{e^{C_0}}{\sqrt{\epsilon}} \sqrt{\tau} K^{\epsilon-1/2}. \tag{7.36}$$

Now we use (7.36) to bound  $\bar{y}_t := Y_t - y_t$  for  $\omega \in \Omega_1 \cap \Omega_2$ . We have

$$\frac{d}{dt} \bar{y}_t = -Z_{t,3} \bar{y}_t - \bar{z}_{t,3} y_t, \quad (7.37)$$

so that

$$\bar{y}_t = - \int_0^t \exp\left(-\int_s^t Z_{t',3} dt'\right) \bar{z}_{s,3} y_s ds. \quad (7.38)$$

Using the rough bound

$$\exp\left(-\int_s^t Z_{t',3} dt'\right) \leq e^{(t-s)|x_0|}$$

and Young's convolution inequality we obtain

$$\|\bar{y}_t\|_{L^1(0,\tau)} \lesssim \frac{e^{|x_0|\tau}}{|x_0|} \int_0^\tau |\bar{z}_{t,3}| |y_t| dt. \quad (7.39)$$

Thus, utilizing (7.36), for  $\omega \in \Omega_1 \cap \Omega_2$  and  $\epsilon$  small enough we have

$$\begin{aligned} \|\bar{y}_t(\omega)\|_{L^1(0,\tau)} &\lesssim e^{C_0} K^{\frac{1/2+r+\epsilon}{1-\epsilon}-1} R \frac{e^{C_0}}{\sqrt{\epsilon}} \sqrt{\tau} K^{\epsilon-1/2} \delta_2^{1/4} \tau K^r \\ &\lesssim \delta_2^{1/4} R \frac{e^{2C_0}}{\sqrt{\epsilon}} \tau K^r. \end{aligned} \quad (7.40)$$

As in our earlier proofs, (7.40), (7.30), and (7.35) together are enough to obtain a contradiction for  $\delta_2$  sufficiently small. The result is that there is  $c > 0$  so that for  $K$  sufficiently large there holds

$$\frac{1}{\tau} \mathbf{E} \int_0^\tau |x_{t,2}|^2 dt \geq cK^{2r}. \quad (7.41)$$

For  $K$  large and  $\delta$  small one has  $\tau \approx_{\delta_1} \log(K)/K$ , and so from (7.41) the proof is completed by setting  $\eta_1(K) = C_{\delta_1} \log(K)/K$  for some large constant  $C_{\delta_1}$ .  $\square$

## A A basic energy estimate

The following lemma quantifies how  $B(x, x) \cdot x = 0$  and the additive nature of the noise imply that the energy level of a trajectory can only change a small amount in a short time.

**Lemma A.1.** *Fix  $\epsilon \in (0, 1)$ . There exist  $K_*(\epsilon) \geq 1$ ,  $\tau_*(\epsilon) \leq 1$ , and  $C > 0$  (which does not depend on  $\epsilon$ ) such that for  $0 \leq \tau \leq \tau_*$  and any  $x_0 \in \mathbb{R}^n$  with  $|x_0| = K \geq K_*$  there holds*

$$\mathbf{P} \left( \sup_{0 \leq t \leq \tau} \left| |x_t|^2 - K^2 \right| \geq \epsilon K^2 \right) \leq \frac{C\tau}{\epsilon^2 K^2}. \quad (\text{A.1})$$

*Proof.* It suffices to show

$$\mathbf{P} \left( \sup_{0 \leq t \leq \tau} |x_t|^2 \geq (1 + \epsilon/2)|x_0|^2 + R^2 \right) \leq C \frac{\tau |x_0|^2}{R^4} \quad (\text{A.2})$$

and

$$\mathbf{P} \left( \inf_{0 \leq t \leq \tau} |x_t|^2 \leq (1 - \epsilon/2)|x_0|^2 - R^2 \right) \leq C \frac{\tau |x_0|^2}{R^4} \quad (\text{A.3})$$

for some constant  $C$  that does not depend on  $\epsilon$ . Indeed, the desired result follows immediately by taking  $R = \sqrt{\epsilon/2}|x_0|$  in (A.2) and (A.3).

We begin with the proof of (A.2). By Itô's formula and  $B(x, x) \cdot x = 0$ , we have

$$|x_t|^2 - |x_0|^2 = 2 \int_0^t x_s \cdot \sigma dW_s - 2 \int_0^t Ax_s \cdot x_s ds + t \sum_{i,j=1}^n |\sigma_{ij}|^2. \quad (\text{A.4})$$

Thus,

$$\mathbf{E}|x_t|^2 \leq |x_0|^2 + C_\sigma t, \quad (\text{A.5})$$

where we have set  $C_\sigma = \sum_{i,j=1}^n |\sigma_{ij}|^2$ . Using the martingale inequality followed by Itô isometry and (A.5) in (A.4) gives

$$\begin{aligned} \mathbf{P} \left( \sup_{0 \leq t \leq \tau} |x_t|^2 - |x_0|^2 - C_\sigma t \geq R^2 \right) &\leq \mathbf{P} \left( 2 \sup_{0 \leq t \leq \tau} \left| \int_0^t x_s \cdot \sigma dW_s \right| \geq R^2 \right) \\ &\leq \frac{4}{R^4} \mathbf{E} \left| \int_0^\tau x_s \cdot \sigma dW_s \right|^2 \leq \frac{C}{R^4} \int_0^\tau \mathbf{E}|x_s|^2 ds \leq C \frac{\tau |x_0|^2}{R^4}, \end{aligned} \quad (\text{A.6})$$

where in the last inequality we assume that  $K_*$  is sufficiently large. The bound (A.2) then follows provided  $K_* \geq \sqrt{2C_\sigma/\epsilon}$ .

Now we turn to the proof of (A.3). Let  $\tilde{x}_t = e^{\lambda_A t} x_t$ , where  $\lambda_A > 0$  is the largest eigenvalue of  $A$ . Then,

$$d\tilde{x}_t = \lambda_A \tilde{x}_t dt + e^{\lambda_A t} B(x_t, x_t) dt - A \tilde{x}_t dt + e^{\lambda_A t} \sigma dW_t. \quad (\text{A.7})$$

Since  $\tilde{x}_t \cdot e^{\lambda_A t} B(x_t, x_t) = e^{2\lambda_A t} B(x_t, x_t) \cdot x_t = 0$ , another application of Itô's lemma gives

$$d|\tilde{x}_t|^2 = 2\lambda_A |\tilde{x}_t|^2 dt - 2A \tilde{x}_t \cdot \tilde{x}_t dt + 2e^{\lambda_A t} \tilde{x}_t \cdot \sigma dW_t + C_\sigma e^{2\lambda_A t} dt. \quad (\text{A.8})$$

Using  $\lambda_A |\tilde{x}_t|^2 \geq A \tilde{x}_t \cdot \tilde{x}_t$ , (A.8) implies

$$|\tilde{x}_t|^2 \geq |x_0|^2 - 2 \left| \int_0^t e^{\lambda_A s} \tilde{x}_s \cdot \sigma dW_s \right|, \quad (\text{A.9})$$

and hence for  $0 \leq t \leq \tau$  there holds

$$|x_t|^2 \geq e^{-2\lambda_A \tau} |x_0|^2 - 2 \left| \int_0^t e^{2\lambda_A s} x_s \cdot \sigma dW_s \right|. \quad (\text{A.10})$$

Proceeding as in the proof of (A.6) we obtain

$$\mathbf{P} \left( \inf_{0 \leq t \leq \tau} |x_t|^2 \leq e^{-2\lambda_A \tau} |x_0|^2 - R^2 \right) \leq \frac{C}{R^4} \int_0^\tau \mathbf{E}|x_s|^2 ds \leq C \frac{\tau |x_0|^2}{R^4}. \quad (\text{A.11})$$

The desired bound follows provided that  $\tau_*$  is small enough so that  $e^{-2\lambda_A \tau_*} \geq 1 - \epsilon/2$ .  $\square$

No data was used or collected in this work.

## References

- [1] A. Athreya, T. Kolba, and J. Mattingly, *Propagating Lyapunov functions to prove noise-induced stabilization*, Electronic Journal of Probability **17** (2012), 1–38.
- [2] D. Bakry, P. Cattiaux, and A. Guillin, *Rate of convergence for ergodic continuous Markov processes: Lyapunov versus Poincaré*, Journal of Functional Analysis **254** (2008), no. 3, 727–759.
- [3] J. Bedrossian, A. Blumenthal, and S. Punshon-Smith, *The Batchelor spectrum of passive scalar turbulence in stochastic fluid mechanics at fixed Reynolds number*, To appear in Comm. Pure Appl. Math. (2019).
- [4] J. Bedrossian and K. Liss, *Quantitative spectral gaps for hypoelliptic stochastic differential equations with small noise*, Probability and Mathematical Physics **2** (2021), no. 3, 477–532.
- [5] J. Birrell, D. Herzog, and J. Wehr, *Transition from ergodic to explosive behavior in a family of stochastic differential equations*, available on math, arXiv preprint ArXiv.1105.2378.
- [6] T. Bohr, M. H. Jensen, G. Paladin, and A. Vulpiani, *Dynamical systems approach to turbulence*, Cambridge University Press, 2005.
- [7] E. Camrud, Ph.D. Thesis, Iowa State University, in preparation.
- [8] J. A. Cañizo, C. Cao, J. Evans, and H. Yoldaş, *Hypo-coercivity of linear kinetic equations via Harris's theorem*, Kinetic & Related Models **13** (2020), no. 1, 97.
- [9] M. Coti Zelati and M. Hairer, *A noise-induced transition in the Lorenz system*, Communications in Mathematical Physics **383** (2021), no. 3, 2243–2274.
- [10] G. Da Prato and J. Zabczyk, *Ergodicity for infinite-dimensional systems*, London Mathematical Society Lecture Note Series, vol. 229, Cambridge University Press, Cambridge, 1996.
- [11] G. Da Prato and J. Zabczyk, *Ergodicity for infinite dimensional systems*, London Mathematical Society Lecture Note Series, Cambridge University Press, 1996.
- [12] P. D. Ditlevsen, *Turbulence and shell models*, Cambridge University Press, 2010.
- [13] R. Douc, G. Fort, and A. Guillin, *Subgeometric rates of convergence of  $f$ -ergodic strong Markov processes*, Stochastic processes and their applications **119** (2009), no. 3, 897–923.
- [14] J. Földes, N. E. Glatt-Holtz, and D. P. Herzog, *Sensitivity of steady states in a degenerately damped stochastic Lorenz system*, Stochastics and Dynamics **21** (2021), no. 08, 2150055.
- [15] S. Friedlander, N. Glatt-Holtz, and V. Vicol, *Inviscid limits for a stochastically forced shell model of turbulent flow*, Ann. Inst. Henri Poincaré Probab. Stat. **52** (2016), no. 3, 1217–1247.
- [16] U. Frisch, *Turbulence: the legacy of A.N. Kolmogorov*, Cambridge University Press, 1995.
- [17] E. B. Gledzer, *System of hydrodynamic type admitting two quadratic integrals of motion*, Sov. phys. dokl., 1973, pp. 216–217.
- [18] F. Hérau and F. Nier, *Isotropic hypoellipticity and trend to equilibrium for the Fokker-Planck equation with a high-degree potential*, Archive for Rational Mechanics and Analysis **171** (2004), no. 2, 151–218.
- [19] D. Herzog and J. Mattingly, *Noise-induced stabilization of planar flows I*, Electronic Journal of Probability **20** (2015), 1–43.
- [20] D. P. Herzog and J. C. Mattingly, *Ergodicity and Lyapunov functions for Langevin dynamics with singular potentials*, Communications on Pure and Applied Mathematics **72** (2019), no. 10, 2231–2255.
- [21] A. Karimi and M. R. Paul, *Extensive chaos in the Lorenz-96 model*, Chaos: An interdisciplinary journal of nonlinear science **20** (2010), no. 4, 043105.
- [22] E. N. Lorenz, *The nature and theory of the general circulation of the atmosphere*, Vol. 218, World Meteorological Organization Geneva, 1967.
- [23] E. N. Lorenz, *Predictability: A problem partly solved*, Proc. seminar on predictability, 1996.
- [24] E. N. Lorenz and K. A. Emanuel, *Optimal sites for supplementary weather observations: Simulation with a small model*, Journal of the Atmospheric Sciences **55** (1998), no. 3, 399–414.
- [25] V. S. L'vov, E. Podivilov, A. Pomyalov, I. Procaccia, and D. Vandembroucq, *Improved shell model of turbulence*, Physical Review E **58** (1998), no. 2, 1811.
- [26] A. J. Majda, *Introduction to turbulent dynamical systems in complex systems*, Springer, 2016.
- [27] J. C. Mattingly, T. Suidan, and E. Vanden-Eijnden, *Simple systems with anomalous dissipation and energy cascade*, Communications in Mathematical Physics **276** (2007), no. 1, 189–220.
- [28] M. Maurelli, *Non-explosion by Stratonovich noise for ODEs*, Electronic Communications in Probability **25** (2020), 1–10.

- [29] L. Meshalkin and I. G Sinai, *Investigation of the stability of a stationary solution of a system of equations for the plane movement of an incompressible viscous liquid*, Journal of Applied Mathematics and Mechanics **25** (1961), no. 6, 1700–1705.
- [30] M. Scheutzow, *Stabilization and destabilization by noise in the plane*, Stochastic Analysis and Applications **11** (1993), no. 1, 97–113.
- [31] C. Villani, *Hypocoercivity*, Hypocoercivity, American Mathematical Society, 2009.
- [32] B. Williamson, *On sdes with partial damping inspired by the navier-stokes equations*, Ph.D. Thesis, Duke University, May 2019. <https://hdl.handle.net/10161/187732019>.
- [33] M. Yamada and K. Ohkitani, *Lyapunov spectrum of a chaotic model of three-dimensional turbulence*, Journal of the Physical Society of Japan **56** (1987), no. 12, 4210–4213.