INSTITUTE OF PHYSICS PUBLISHING NONLINEARITY

Nonlinearity **19** (2006) 769–794 doi:10.1088/0951-7715/19/4/001

Stochastic models for selected slow variables in large deterministic systems

A Majda[1,2]**, I Timofeyev**[3,4] **and E Vanden-Eijnden**[1]

[1] Courant Institute of Mathematical Sciences, New York University, New York, NY 10012, USA
[2] Center for Atmosphere Ocean Science, New York University, New York, NY 10012, USA
[3] Department of Mathematics, University of Houston, Houston, TX 77204, USA

E-mail: ilya@math.uh.edu

Received 17 May 2005, in final form 16 January 2006
Published 14 February 2006
Online at stacks.iop.org/Non/19/769

Recommended by C Le Bris

Abstract

A new stochastic mode-elimination procedure is introduced for a class of deterministic systems. Under assumptions of ergodicity and mixing, the procedure gives closed-form stochastic models for the slow variables in the limit of infinite separation of timescales. The procedure is applied to the truncated Burgers–Hopf (TBH) system as a test case where the separation of timescale is only approximate. It is shown that the stochastic models reproduce exactly the statistical behaviour of the slow modes in TBH when the fast modes are artificially accelerated to enforce the separation of timescales. It is shown that this operation of acceleration only has a moderate impact on the bulk statistical properties of the slow modes in TBH. As a result, the stochastic models are sound for the original TBH system.

PACS numbers: 02.50.−r, 02.70.Hm, 02.70.Rr, 05.10.Gg

1. Introduction

Multiscale problems have attracted considerable attention in recent years as a result of the significant increase in computational capacity. Yet increasing computational power alone cannot overcome the inherent complexity of multi-scale models due to the existence of many dynamical variables evolving on vastly different scales. The situation is especially frustrating since slowly evolving large-scale structures and their statistical behaviour are often the most interesting, and yet the computational power is wasted on resolving the smallest and fastest variables in the system. This problem has become a common concern in many fields, including atmosphere–ocean sciences, material sciences, molecular dynamics, etc.

[4] Author to whom any correspondence should be addressed.

Statistical theories and stochastic modelling with non-essential degrees of freedom represented stochastically provide computationally feasible alternatives for calculating the statistical evolution of the slow variables, and this topic has received a lot of attention in recent years [1–11]. Recently, a systematic approach to stochastic mode-elimination was developed by the authors in [8,9], hereafter referred to as MTV. Here we generalize MTV by removing an unnecessary assumption in the original procedure. Namely, we show that the preliminary step in the original MTV procedure in which the nonlinear self-interactions of the fast degrees of freedom are represented stochastically can, in fact, be avoided. Here, closed-form stochastic equations for the slow modes are obtained for a class of large deterministic systems, without *a priori* modification of these systems, based only on assumptions of ergodicity and mixing (assumptions 2.1 and 2.2) as well as infinite separation of timescale between fast and slow modes in these systems. Under these assumptions, the existence of closed-form equations for the slow modes is guaranteed by standard adiabatic elimination theorems [12,13], and we shall derive these equations explicitly.

In any practical situation, the timescale separation is not infinite. Yet the MTV procedure has been successfully tested on various prototype models [14–16] and more realistic atmospheric systems [17, 18] with moderate timescale separation. This suggests that the statistical behaviour of the slow modes in these models is rather insensitive to the details of the timescale over which the fast modes evolve. One way to rationalize and test this hypothesis is to introduce a parameter ε into the equations in order to selectively accelerate the motion of the fast modes when $\varepsilon \to 0$ and observe the effect on the behaviour of the slow modes. In this paper, the truncated Burgers–Hopf model (TBH) [19, 20] is utilized as a test case for this approach. TBH has a quadratic invariant (energy), and the dynamics can be accelerated in such a way that this property is preserved. We perform a series of numerical simulations of the full equations with different values of the parameter ε to verify the existence of a dynamics for the slow variables in the limit as $\varepsilon \to 0$. This allows us to compare the statistical behaviour of the slow modes in the original TBH system, in the selectively accelerated TBH system (SA-TBH) and in the stochastic model. Using specific scaling properties of the class of conservative systems that TBH belongs to, we show that the coefficients in the stochastic equations for the slow variables can be estimated from a single numerical simulation of an auxiliary subsystem and then be easily extrapolated to other regimes. This is in the spirit of the computational procedure introduced in [21] (see also [22, 23]) and the seamless MTV procedure used in [17, 18], but considerably lowers the numerical cost.

The rest of the paper is organized as follows. In section 2 we derive closed-form stochastic models in the context of a class of deterministic systems which are energy-conserving, ergodic and mixing. In section 3 the spectral truncation of the inviscid Burgers–Hopf system is introduced and the analytical and statistical properties of this model are discussed briefly. In section 4 the SA-TBH system is introduced in which the arbitrary large separation of timescales between selectively chosen slow and fast modes can be enforced. The properties of SA-TBH are investigated through the direct numerical simulations and compared with those of the original TBH system. In section 5 explicit formulae for the stochastic models for the first Fourier mode (section 5.1) and first and second Fourier modes (section 5.2) are given and compared with simulations of the original TBH and the SA-TBH systems.

2. Stochastic models for deterministic systems

In this section the mode-reduction strategy for a class of deterministic systems is introduced. This strategy is particularly relevant in the context of high-dimensional systems of ODEs arising as projections of conservative partial differential equations. To present the general

treatment, we consider a set of real variables $\{u_k(t)\}_{k\in S}$ with index k varying in some set S. The variables $\{u_k(t)\}_{k\in S}$ can be thought of as coefficients in the some appropriate representation (Fourier, etc), and the set S is the set of indices retained in the Galerkin projection. We also assume that the dependent variables $\{u_k(t)\}_{k\in S}$ can be decomposed in two sets, $\{a_i\}_{i\in S_a}$ and $\{b_j\}_{j\in S_b}$, where $\{a_i\}_{i\in S_a}$ represent the slow essential degrees of freedom and $\{b_j\}_{j\in S_b}$ represent the fast unresolved modes. The indices i and j vary over some index sets $S_a = \{1,\dots,M\}$ and $S_b = \{1,\dots,N\}$. M and N are the numbers of slow and fast variables, respectively.

We consider a general quadratic system of equations for the variables $a = \{a_i\}$ and $b = \{b_i\}$

$$\begin{aligned}
\dot a_i &= \sum_{j,k\in S_a} m^{aaa}_{ijk} a_j a_k + 2\varepsilon^{-1} \sum_{j\in S_a, k\in S_b} m^{aab}_{ijk} a_j b_k + \varepsilon^{-1} \sum_{j,k\in S_b} m^{abb}_{ijk} b_j b_k,\\
\dot b_i &= \varepsilon^{-1} \sum_{j,k\in S_a} m^{baa}_{ijk} a_j a_k + 2\varepsilon^{-1} \sum_{j\in S_a, k\in S_b} m^{bab}_{ijk} a_j b_k + \varepsilon^{-2} \sum_{j,k\in S_b} m^{bbb}_{ijk} b_j b_k,
\end{aligned} \tag{1}$$

where $\varepsilon < 1$ is a parameter measuring the difference in timescales between the slow and fast modes, and we will be interested in the asymptotic behaviour of (1) in the limit as $\varepsilon \to 0$. The right-hand sides in (1) have been explicitly decomposed into the self-interactions of slow modes (a with a), interactions between the slow and fast dynamics (a and b) and fast self-interactions (b with b). The interaction coefficients are denoted as m^{xzy}_{ijk} where each x, y, z stands for a or b. Without the loss of generality we can make the symmetry assumption $m^{xyz}_{ijk} = m^{xzy}_{ikj}$ and we also suppose that the coefficients m^{xyz}_{ijk} satisfy

$$\begin{aligned}
&m^{xyz}_{ijk} + m^{yzx}_{jki} + m^{zxy}_{kij} = 0,\\
&\sum_{\substack{i\in S_x, j\in S_y\\ k\in S_z}} m^{xyz}_{ijk} \frac{\partial}{\partial x_i} y_j z_k = 0.
\end{aligned} \tag{2}$$

The first equation in (2) guarantees that the dynamics in (1) conserves

$$E = \sum_{i\in S_a} a_i^2 + \sum_{i\in S_b} b_i^2 =: |a|^2 + |b|^2, \tag{3}$$

with E fixed by the initial condition for (1). The second equation in (2) ensures that the dynamics in (1) is volume-preserving (Liouville property).

As shown below, to eliminate the fast degrees of freedom we will have to consider an auxiliary subsystem involving only the fast modes to determine the coefficients in the stochastic model for the slow modes. This subsystem is the projection of the original equations in (1) onto the fast modes alone

$$\dot c_i = \sum_{j,k\in S_b} m^{bbb}_{ijk} c_j c_k, \qquad (i \in S_b). \tag{4}$$

By the first equation in (2), the fast subsystem conserves

$$\bar E = \sum_{i\in S_b} c_i^2 =: |c|^2. \tag{5}$$

We will denote by $C_i^{c,\bar E}(t)$ the solution of the fast subsystem (4) for the initial condition $C_i^{c,\bar E}(0) = c_i$ with

$$|c|^2 = \bar E \tag{6}$$

and to proceed, we will make two assumptions about the statistical behaviour of this subsystem.

Assumption 2.1. *(1) is ergodic on the hypersphere defined in (5) with respect to the uniform distribution on this sphere, i.e. for any suitable test function $f : \mathbb{R}^M \to \mathbb{R}$ and almost every initial condition, we have*

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f(C^{c,\bar{E}}(t))\mathrm{d}t = S_M^{-1} \bar{E}^{1-M/2} \int_{|c|^2=\bar{E}} f(c)\mathrm{d}\sigma(c), \tag{7}$$

where S_n is the area of the unit sphere in dimension n and $\mathrm{d}\sigma(c)$ is the surface element (Lebesgue measure) on the sphere $|c|^2 = \bar{E}$.

Notice that, using the co-area formula, the distribution in (7) can be expressed as

$$\begin{aligned} \mathrm{d}\mu_{\bar{E}}(c) &:= S_M^{-1} E^{1-M/2}\, \mathrm{d}\sigma(c) \\ &= S_M^{-1} \bar{E}^{1-M/2} \delta(\bar{E} - |c|^2)\mathrm{d}c, \end{aligned} \tag{8}$$

where $\delta(z)$ is the Dirac delta distribution and $\mathrm{d}c = \prod_{i\in S_b} \mathrm{d}c_i$. This distribution is usually referred to as the microcanonical distribution. In addition we will assume the following:

Assumption 2.2. *The dynamics in (1) is rapidly mixing in the sense that for any suitable test function $g : \mathbb{R}^M \times \mathbb{R}^M \to \mathbb{R}$ and almost every initial condition, we have*

$$\begin{aligned} \lim_{T\to\infty} \frac{1}{T} \int_0^T g(C^{c,\bar{E}}(t), C^{c,\bar{E}}(t+s))\mathrm{d}t &= \int_{|c|^2=\bar{E}} g(c, C^{c,\bar{E}}(s))\mathrm{d}\mu_{\bar{E}}(c) \\ &= G_\infty + G(s) \qquad (s \geqslant 0), \end{aligned} \tag{9}$$

where

$$G_\infty = \int_{\substack{|c|^2=\bar{E} \\ |c'|^2=\bar{E}}} g(c, c')\mathrm{d}\mu_{\bar{E}}(c)\mathrm{d}\mu_{\bar{E}}(c') \tag{10}$$

and $G(s)$ satisfies

$$\int_0^\infty G(s)\mathrm{d}s < \infty. \tag{11}$$

We require that (9) holds for any g such that G_∞ is finite.

Note that the verification of assumptions 2.1 and 2.2 in any specific system is extremely difficult. Here we will empirically verify these assumptions on TBH and SA-TBH via a series of numerical experiments presented in section 4.

2.1. Stochastic models for small ε

Propositions 2.3 and 2.4 describe the limiting behaviour of $a(t)$ in (1) in the limit $\varepsilon \to 0$. The stochastic model which captures the dynamics of $a(t)$ in this limit is given in (21).

For any test function $\varphi : \mathbb{R}^M \to \mathbb{R}$, let

$$u^E(a, b, t) = \varphi(a(t)), \tag{12}$$

where $a(t)$ denotes the solution of the first equation in (1) for the initial condition $(a(0), b(0)) = (a, b)$, with $|a|^2 + |b|^2 = E$. We have the following proposition.

Proposition 2.3. *Under assumptions 2.1 and 2.2, for any test function φ and any $T \in [0, \infty)$, we have*

$$\lim_{\varepsilon\to 0} \sup_{0\leqslant t\leqslant T} |u^E(a, b, t) - \bar{u}(a, t)| = 0, \tag{13}$$

where $\bar{u}(a,t)$ *satisfies*

$$\bar{u}_t = L\bar{u}, \qquad \bar{u}(a,0) = \varphi(a), \tag{14}$$

Here $L = L_1 + L_2$ *with*

$$\begin{aligned} L_1 &= \sum_{i,j,k\in S_a} m_{ijk}^{aaa} a_j a_k \frac{\partial}{\partial a_i} \\ L_2 &= \int_0^\infty \mathrm{d}t \int \mathrm{d}\mu_{E-|a|^2}(c)(K_1+K_2)K_3, \end{aligned} \tag{15}$$

where

$$\begin{aligned} K_1 &= 2\sum_{\substack{i,j\in S_a\\ k\in S_b}} m_{ijk}^{aab} a_j c_k \frac{\partial}{\partial a_i} + \sum_{\substack{i\in S_a\\ j,k\in S_b}} m_{ijk}^{abb} c_j c_k \frac{\partial}{\partial a_i} \\ K_2 &= 2\sum_{\substack{j\in S_a\\ i,k\in S_b}} m_{ijk}^{bab} a_j c_k \frac{\partial}{\partial c_i} + \sum_{\substack{i\in S_b\\ j,k\in S_a}} m_{ijk}^{baa} a_j a_k \frac{\partial}{\partial c_i} \\ K_3 &= 2\sum_{\substack{i,j\in S_a\\ k\in S_b}} m_{ijk}^{aab} a_j C_k^{c,E-|a|^2}(t) \frac{\partial}{\partial a_i} + \sum_{\substack{i\in S_a\\ j,k\in S_b}} m_{ijk}^{abb} C_j^{c,E-|a|^2}(t) C_k^{c,E-|a|^2}(t) \frac{\partial}{\partial a_i}. \end{aligned} \tag{16}$$

The proposition is formally established in section 2.2 by singular perturbation analysis of the backward equation associated with (1) following what was done in [9, 23, 24]. A rigorous proof of the proposition can be made by generalizing the proof procedure in [13] (see also [12, 25, 26]). Assumptions 2.1 and 2.2 guarantee that the integrals in (15) exists and are finite, i.e. L_2 is a well-defined (elliptic) operator (see also proposition 2.4).

In (16), the operators K_1, K_2 and K_3 are averaged over the sphere of energy $\bar{E} = E - |a|^2$ whose radius changes as the slow variables a evolve. Using the scaling properties of the solutions of (4), the resulting dependence in a of K_1, K_2 and K_3 can be put in a more explicit form that is convenient for the calculations. We state this result as follows:

Proposition 2.4. *An equivalent expression for* L *given in proposition 2.3 is*

$$L = L_1 + \sum_{i\in S_a} B_i(a)\frac{\partial}{\partial a_i} + \sum_{i,j\in S_a} \frac{\partial}{\partial a_i} D_{ij}(a) \frac{\partial}{\partial a_j}. \tag{17}$$

Here

$$B_i(a) = -(1-2N^{-1})\mathcal{E}^{-1}(a) \sum_{j\in S_a} D_{ij}(a) a_j, \tag{18}$$

where $\mathcal{E}(a) := N^{-1}(E-|a|^2)$ *and*

$$D_{ij}(a) = \mathcal{E}^{1/2}(a) \int_0^\infty \mathrm{d}t \int \mathrm{d}\mu_N(c) P_i(c) P_j(C^{c,N}(t)), \tag{19}$$

with

$$P_i(c) = 2\sum_{\substack{j\in S_a\\ k\in S_b}} m_{ijk}^{aab} a_j c_k + \mathcal{E}^{1/2}(a) \sum_{j,k\in S_b} m_{ijk}^{abb} c_j c_k. \tag{20}$$

Note that the Itô stochastic differential equation (SDE) corresponding to L is

$$\mathrm{d}a_i = \sum_{j,k\in S_a} m_{ijk}^{aaa} a_j a_k \,\mathrm{d}t + B_i(a)\mathrm{d}t + \sum_{j\in S_a} \frac{\partial}{\partial a_j} D_{ij}(a)\mathrm{d}t + \sqrt{2}\sum_{j\in S_a} \sigma_{ij}(a)\mathrm{d}W_j, \tag{21}$$

where W_j is a M-dimensional Wiener process and $\sigma_{ij}(a)$ satisfies

$$\sum_{k\in S_a} \sigma_{ik}(a)\sigma_{jk}(a) = D_{ij}(a). \tag{22}$$

The stochastic models in 21 will be utilized as effective models for the slow dynamics in the applications below.

The proof proposition 2.4 is given in section 2.3. It uses a rescaling on the sphere of radius $\sqrt{N}$ of the fast subsystem. After this rescaling system (4) must be solved with an initial condition consistent with $|C^{c,N}(0)|^2 = |c|^2 = N$, which is therefore independent of a. In other words, the diffusion tensor $D_{ij}(a)$ can be estimated for all a from a single calculation with (4); this can be made more evident by writing (19) as

$$\begin{aligned} D_{ii'}(a) = {} & 4\mathcal{E}^{1/2}(a) \sum_{\substack{j,j'\in S_a\\ k,k'\in S_b}} m_{ijk}^{aab} m_{i'j'k'}^{aab} a_j a_{j'} Q_{kk'}^{(1)} + 2\mathcal{E}(a) \sum_{\substack{j\in S_a\\ k,j',k'\in S_b}} m_{ijk}^{aab} m_{i'j'k'}^{abb} a_j Q_{kj'k'}^{(2)} \\ & + 2\mathcal{E}(a) \sum_{\substack{j'\in S_a\\ k',j,k\in S_b}} m_{ijk}^{abb} m_{i'j'k'}^{aab} a_{j'} Q_{k'jk}^{(2)} + \mathcal{E}^{3/2}(a) \sum_{j,k,j',k'\in S_b} m_{ijk}^{abb} m_{i'j'k'}^{abb} Q_{jkj'k'}^{(3)}, \end{aligned} \tag{23}$$

where

$$\begin{aligned} Q_{kk'}^{(1)} &= \int_0^\infty \mathrm{d}t \int \mathrm{d}\mu_N(c) c_k C_{k'}^{c,N}(t) \\ Q_{kj'k'}^{(2)} &= \int_0^\infty \mathrm{d}t \int \mathrm{d}\mu_N(c) c_k C_{j'}^{c,N}(t) C_{k'}^{c,N}(t) \\ Q_{jkj'k'}^{(3)} &= \int_0^\infty \mathrm{d}t \int \mathrm{d}\mu_N(c) c_j c_k C_{j'}^{c,N}(t) C_{k'}^{c,N}(t). \end{aligned} \tag{24}$$

The integrals with respect to the equilibrium distribution $\mu_N(c)$ correspond to microcanonical averages of the fast subsystem on the energy shell $\bar{E} = N$. $Q_{kk'}^{(1)}$ is the area under the graph of the two-point autocorrelation function between fast variables k and k'. Similarly, $Q_{kj'k'}^{(2)}$, involves a lagged mixed third moment, and $Q_{jkj'k'}^{(3)}$ involves a lagged fourth moment. These terms are evaluated numerically from a single realization of the fast subsystem.

A straightforward consequence of proposition 2.4 is that the equilibrium distribution of a is

$$\mathrm{d}\mu(a) = Z^{-1}(E - |a|^2)_+^{N/2-1}\,\mathrm{d}a, \tag{25}$$

where $Z = S_{N+M} S_N^{-1} E^{(N+M)/2-1}$. It can be checked by direct calculation that (25) is indeed annihilated by the adjoint of L. Therefore, (25) is precisely the reduced distribution obtained by integrating the microcanonical distribution of the original system (1) over the fast variables b, which indicates that the effective equation in a leads to the same equilibrium distribution as the original equation in (1). As a result, the restriction that $T < \infty$ in proposition 2.3 seems unnecessary.

Note that in the limit where the number of fast modes tends to infinity, $N \to \infty$, provided that we scale the total energy as $E = N/\beta$ (β plays the role of an inverse temperature), the invariant distribution in (25) tends to a Gaussian

$$\mathrm{d}\mu(a) \to \bar{Z}^{-1}\mathrm{e}^{-\beta|a|^2}\,\mathrm{d}a, \tag{26}$$

where $\bar{Z}$ is a normalization factor.

2.2. Formal proof of proposition 2.3

The function $u^E(a,b,t)$ defined in (12) satisfies the backward equation

$$\frac{\partial u^E}{\partial t} = \left(\frac{1}{\varepsilon^2}L_3 + \frac{1}{\varepsilon}L_2 + L_1\right)u^E, \tag{27}$$

where

$$\begin{aligned}
L_1 &= \sum_{i,j,k\in S_a} m_{ijk}^{aaa} a_j a_k \frac{\partial}{\partial a_i} \\
L_2 &= \sum_{i\in S_a}\left(2\sum_{j\in S_a,k\in S_b} m_{ijk}^{aab} a_j b_k + \sum_{j,k\in S_b} m_{ijk}^{abb} b_j b_k\right)\frac{\partial}{\partial a_i} \\
&\quad + \sum_{i\in S_b}\left(\sum_{j,k\in S_a} m_{ijk}^{baa} a_j a_k + 2\sum_{j\in S_a,k\in S_b} m_{ijk}^{bab} a_j b_k\right)\frac{\partial}{\partial b_i} \\
L_3 &= \sum_{i,j,k\in S_b} m_{ijk}^{bbb} b_j b_k \frac{\partial}{\partial b_i}.
\end{aligned} \tag{28}$$

Look for a power series representation of the function u^E

$$u^E = u_0 + \varepsilon u_1 + \varepsilon^2 u_2 + \cdots$$

substitute this series into (27) and collect succesive powers of ε:

$$\begin{aligned}
L_3 u_0 &= 0 \\
L_3 u_1 &= -L_2 u_0 \\
L_3 u_2 &= \frac{\partial u_0}{\partial t} - L_2 u_1 - L_3 u_0 \\
L_3 u_3 &= \ldots.
\end{aligned} \tag{29}$$

The operator L_3 is the backward operator for the fast subsystem defined in (4). The first equation in (29) belongs to the null space of this operator. By assumption 2.1, this is equivalent to requiring that

$$u_0(a,t) = (\mathbf{P}u_0)(a,t), \tag{30}$$

where

$$(\mathbf{P}u_0)(a,t) := \int_{|c|^2=E-|a|^2} u_0(a,c,t)\mathrm{d}\mu_{E-|a|^2}(c) \tag{31}$$

denotes the expectation with respect to the microcanonical distribution (8). The second equation in (29) requires a solvability condition, namely, that the right-hand side belongs to the range of the operator L_3 or, equivalently, that it be orthogonal to the null-space of its adjoint, i.e. we must have

$$\mathbf{P}L_2 u_0 = \mathbf{P}L_2\mathbf{P}u_0 = 0. \tag{32}$$

Taking into account the particular form of L_2, this condition translates into the following two conditions

$$\begin{aligned}
&\int_{|c|^2=E-|a|^2} c_j \,\mathrm{d}\mu_{E-|a|^2}(c) = 0, \\
&\int_{|c|^2=E-|a|^2} c_j c_k \,\mathrm{d}\mu_{E-|a|^2}(c) = 0, \qquad \text{if } j\neq k,
\end{aligned} \tag{33}$$

which are automatically satisfied by parity. Therefore we can can solve for u_1 the second equation in (29). Formally

$$u_1 = -L_3^{-1} L_2 u_0. \tag{34}$$

To give an explicit expression for L_3^{-1} applied to any function $f(c)$ such that $\mathbf{P}f = 0$ (such as $-L_2 u_0$ by (32)) consider the following equation for $g = g(c,t)$:

$$\frac{\partial g}{\partial t} = L_3 g - f = \sum_{i,j,k \in S_b} m_{ijk}^{bbb} c_j c_k \frac{\partial}{\partial c_i} g - f, \qquad g(c,0) = 0. \tag{35}$$

The stationary solution of this equation can be expressed formally as $g = L_3^{-1} f$. On the other hand, solving (35) by the method of characteristics, we obtain

$$g(c,t) = \int_0^t f(C^{c,\bar{E}}(s))\mathrm{d}s, \tag{36}$$

where $C^{c,\bar{E}}(s)$ denotes the solution of (4) for the initial condition $C^{c,\bar{E}}(0) = c$ with $|c|^2 = \bar{E}$. Since $\mathbf{P}f = 0$, assumption 2.2 guarantees the convergence of this integral as $t \to \infty$ and therefore

$$(L_3^{-1} f)(c) = \lim_{t\to\infty} g(c,t) = \int_0^\infty f(C^{c,\bar{E}}(s))\mathrm{d}t. \tag{37}$$

Finally the equation for u_0 is obtained from the solvability condition for the last equation in (29). Using (34), this equation is

$$\frac{\partial u_0}{\partial t} = \mathbf{P}L_2 u_1 + L_3 u_0 = -\mathbf{P}L_2 L_3^{-1} L_2 \mathbf{P} u_0 + L_1 u_0, \tag{38}$$

where we have taken into account that $\mathbf{P}u_0 = u_0$, $\mathbf{P}L_1 u_0 = L_1 u_0$ (since L_3 does not depend on b) and $\mathbf{P}L_3 = 0$. The explicit form of the operator at the right-hand side of (38) is the operator L defined in (15), which terminates our formal proof.

Note that this proof is formal because we have not shown that the function $u_1 + \varepsilon u_2 + \cdots$ entering in $u^E - u_0 = \varepsilon u_1 + \varepsilon^2 u_2 \cdots$ remains bounded. Showing this can, in principle, be done by adapting the arguments in [13].

2.3. Proof of proposition 2.4

The proof makes use of rather tedious but otherwise completely straightforward algebraic manipulations. Therefore, for the sake of brevity, we only outline the calculations to be done and leave the details to the (courageous) reader. The proof consists of two steps.

First, the self-adjoint part of the operator L_2 can be obtained by moving the differentiation in a_i from K_1 to the right-most place. This leads to the diffusive term in (17), and some remainder, which leads to the drift term in (17). The diffusive part of the operator one obtains this way can be written as

$$D_{ij}(a) = \int_0^\infty \mathrm{d}t \int \mathrm{d}\mu_{E-|a|^2}(c)\, \bar{P}_i(c) \bar{P}_j(C^{c,N\mathcal{E}(a)}(t)), \tag{39}$$

with

$$\bar{P}_i(c) = 2 \sum_{\substack{j\in S_a \\ k \in S_b}} m_{ijk}^{aab} a_j c_k + \sum_{j,k\in S_b} m_{ijk}^{abb} c_j c_k. \tag{40}$$

The remainder leading to the drift term in (17) contains terms involving derivatives with respect to a_i of the distribution $\mu_{E-|a|^2}$. These terms can be calculated using the explicit expression

$$\mathrm{d}\mu_{E-|a|^2}(c) = S_N^{-1} (E - |a|^2)^{1-N/2} \delta(E - |a|^2 - |c|^2)\mathrm{d}c$$

together with the relation

$$\frac{\partial}{\partial a_i} S_N^{-1}(E-|a|^2)^{1-N/2}\delta(E-|a|^2-|c|^2) = -2a_i S_N^{-1}(E-|a|^2)^{1-N/2}\delta'(E-|a|^2-|c|^2) + (N-2)S_N^{-1}a_i(E-|a|^2)^{-N/2}\delta(E-|a|^2-|c|^2),$$

where $\delta'(\cdot)$ is the distributional derivative of $\delta(\cdot)$. These terms in the remainder can be combined with terms arising from K_2 after integration by part in c (no other term arise due to the second property of m_{ijk}^{xyz} in (2)), which can be computed using

$$\frac{\partial}{\partial c_i} S_N^{-1}(E-|a|^2)^{1-N/2}\delta(E-|a|^2-|c|^2) = -2c_i S_N^{-1}(E-|a|^2)^{1-N/2}\delta'(E-|a|^2-|c|^2).$$

The terms proportional to $\delta'(\cdot)$ in the remainder cancel exactly using the first property of m_{ijk}^{xyz} in (2), and one is left with an additional drift term which can be expressed in terms of $D_{ij}(a)$ precisely as in (18).

Second, the expression in (39) for $D_{ij}(a)$ can be simplified using the following scaling property of the solutions of (4):

$$C^{\mathcal{E}^{1/2}(a)c,N\mathcal{E}(a)}(t) = \mathcal{E}^{1/2}(a)C^{c,N}(\mathcal{E}^{1/2}(a)t). \tag{41}$$

Using this property in (39) allows us to change integration variables in c and t and obtain (19).

3. Truncated Burgers–Hopf system

The spectral truncation of the inviscid Burgers–Hopf model (TBH), introduced recently in [19, 20], exhibits many of the desirable properties found in more complex systems but has the virtue of allowing a relatively complete analysis of statistical properties and extensive numerical studies. The model is constructed by projecting the inviscid Burgers–Hopf equation on a finite number of Fourier modes in periodic geometry

$$(u_\Lambda)_t + \tfrac{1}{2}P_\Lambda(u_\Lambda^2)_x = 0. \tag{42}$$

Here P_Λ is the projection operator

$$P_\Lambda f(x) \equiv f_\Lambda(x) = \sum_{|k|\leqslant\Lambda} \hat{f}_k \mathrm{e}^{\mathrm{i}kx}, \tag{43}$$

where

$$\hat{f}_k = \int_0^{2\pi} f(x)\mathrm{e}^{-\mathrm{i}kx}\,\mathrm{d}x. \tag{44}$$

Here and elsewhere in the paper the Burgers–Hopf equation is considered on a 2π-periodic domain. With the expansion

$$u_\Lambda(x,t) = \sum_{|k|\leqslant\Lambda} \hat{u}_k(t)\mathrm{e}^{\mathrm{i}kx}, \qquad \hat{u}_{-k}(t) = \hat{u}_k^*(t) \tag{45}$$

the equations in (42) can be recast as a finite-dimensional system of equations for the Fourier amplitudes, $\hat{u}_k$ with $|k| \leqslant \Lambda$

$$\dot{\hat{u}}_k = -\frac{\mathrm{i}k}{2} \sum_{\substack{k+p+q=0 \\ |p|,|q|\leqslant\Lambda}} \hat{u}_p^*\hat{u}_q^*. \tag{46}$$

The statistical properties of the equations in (42) or (46) were analysed in detail in a series of papers [19, 20, 27]. In particular, it is easy to show that this system has three conserved quantities, momentum

$$M = \frac{1}{2\pi}\int_0^{2\pi} u_\Lambda \,\mathrm{d}x = \hat{u}_0, \tag{47}$$

energy

$$E = \frac{1}{4\pi}\int_0^{2\pi} u_\Lambda^2 \,\mathrm{d}x = \frac{1}{2}|\hat{u}_0|^2 + \sum_{k=1}^{\Lambda} |\hat{u}_k|^2 \tag{48}$$

and Hamiltonian

$$H = \frac{1}{12\pi}\int_0^{2\pi} P_\Lambda(u_\Lambda^3)\mathrm{d}x = \frac{1}{6}\sum_{\substack{k+p+q=0\\ |p|,|q|\leqslant\Lambda}} \hat{u}_k\hat{u}_p\hat{u}_q. \tag{49}$$

The vector field in (46) is volume-preserving and in [19, 20] an equilibrium statistical theory for the TBH based on the first two conserved quantities defined by momentum and energy, M and E alone, was developed. In [27] the role of the Hamiltonian (49) was examined. It was demonstrated that for a fixed energy level the distribution of H is sharply peaked near $H = 0$, provided that the truncation size, Λ, is large enough. Moreover, the marginal on each Fourier mode $k \neq 0$ of the microcanonical distribution on constant $M = 0$, E and H approaches the Gaussian probability distribution

$$\mathrm{d}\mu(\hat{u}_k) = C_\beta \mathrm{e}^{-\beta|\hat{u}_k|^2}\,\mathrm{d}\hat{u}_k \tag{50}$$

in the limit $\Lambda \to \infty$, provided that E grows linearly with Λ, $E = \Lambda/\beta$ for some β playing the role of an inverse temperature. This implies equipartition of energy in this limit

$$\forall k \neq 0: \qquad \mathrm{var}\{\mathrm{Re}\,\hat{u}_k\} = \mathrm{var}\{\mathrm{Im}\,\hat{u}_k\} = \frac{1}{2\beta}. \tag{51}$$

These predictions were verified for a wide variety of regimes and random and deterministic initial data. In addition, it was demonstrated that for low wave numbers an empirical scaling law for correlation times, defined as the area under the normalized auto-correlation functions for mode $\hat{u}_k$, holds

$$\mathrm{corr\_time}\{\hat{u}_k\} \sim |k|^{-1}. \tag{52}$$

The correlation functions the first five modes in the TBH system are depicted in figure 1. The correlation times are given in table 1.

4. Selectively accelerated TBH systems

The results in section 2 hold provided that the timescale separation between fast and slow variables is infinite. In practical applications the timescale separation between these two groups of variables is moderate, at best. Therefore, for any such system it is in principle necessary to verify the applicability of asymptotic expansions *a priori*. A systematic way to address this issue is to artificially accelerate the dynamics of the fast variables and observe the effect this induces on the statistical behaviour of the slow variables. This procedure can be implemented as follows on TBH: (i) fix a wave number $\Lambda_1 < \Lambda$ such that any mode with

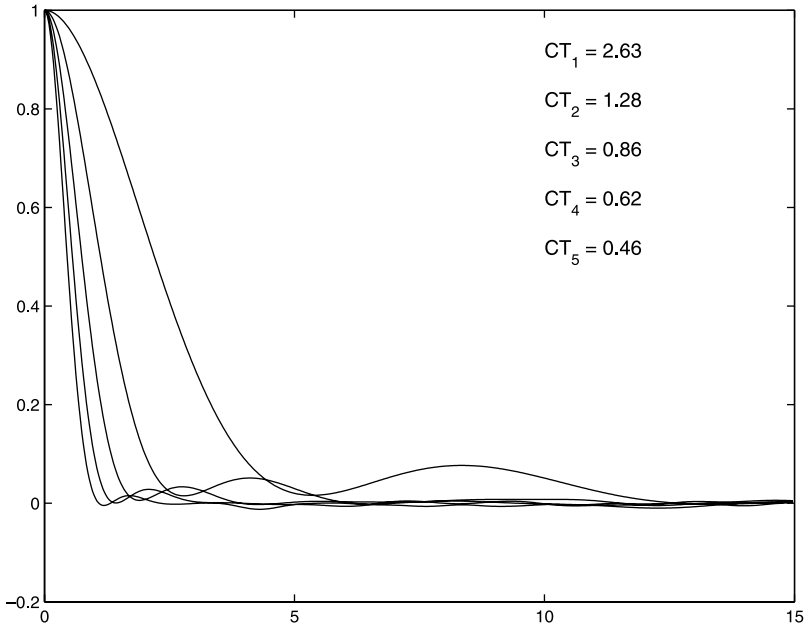

Figure 1. Normalized correlation function of $\mathrm{Re}\,\hat{u}_k$, $k = 1, \ldots, 5$.

Table 1. Correlations times (area under the graphs) of the first five correlation functions in TBH. The correlations times multiplied by $|k|$ is approximately constant, consistent with (52).

	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$
corr_time	2.63	1.28	0.86	0.62	0.46
corr_time $\times\, k$	2.63	2.56	2.58	2.48	2.29

$|k| \leqslant \Lambda_1$ is considered as slow, and any mode with $\Lambda_1 < |k| \leqslant \Lambda$ is considered as fast; (ii) modify (46) as

$$\begin{aligned}
\dot{\hat{u}}_k &= -\frac{\mathrm{i}k}{2} \sum_{S_{\mathrm{ss}}(k)} \hat{u}_p^* \hat{u}_q^* - \frac{\mathrm{i}k}{\varepsilon} \sum_{S_{\mathrm{sf}}(k)} \hat{u}_p^* \hat{u}_q^* - \frac{\mathrm{i}k}{2\varepsilon} \sum_{S_{\mathrm{ff}}(k)} \hat{u}_p^* \hat{u}_q^*, \ |k| \leqslant \Lambda_1 \\
\dot{\hat{u}}_k &= -\frac{\mathrm{i}k}{2\varepsilon} \sum_{S_{\mathrm{ss}}(k)} \hat{u}_p^* \hat{u}_q^* - \frac{\mathrm{i}k}{\varepsilon} \sum_{S_{\mathrm{sf}}(k)} \hat{u}_p^* \hat{u}_q^* - \frac{\mathrm{i}k}{2\varepsilon^2} \sum_{S_{\mathrm{ff}}(k)} \hat{u}_p^* \hat{u}_q^*, \ \Lambda_1 < |k| \leqslant \Lambda,
\end{aligned} \tag{53}$$

where

$$\begin{aligned}
S_{\mathrm{ss}}(k) &= \{p, q : k + p + q = 0, |p| \leqslant \Lambda_1, |q| \leqslant \Lambda_1\}, \\
S_{\mathrm{sf}}(k) &= \{p, q : k + p + q = 0, \Lambda_1 < |p| \leqslant \Lambda, |q| \leqslant \Lambda_1\}, \\
S_{\mathrm{ff}}(k) &= \{p, q : k + p + q = 0, \Lambda_1 < |p| \leqslant \Lambda, \Lambda_1 < |q| \leqslant \Lambda\}.
\end{aligned} \tag{54}$$

The original TBH system in (45) is recovered by setting $\varepsilon = 1$.

The SA-TBH systems in (53) preserve the energy in (48) for all ε and all choices of $S_{\mathrm{ss}}(k)$, $S_{\mathrm{sf}}(k)$ and $S_{\mathrm{ff}}(k)$. Furthermore, (48) is the only surviving conserved quantity and there is no cubic integral of motion such as (49) for the modified system when $\varepsilon < 1$. If the dynamics is ergodic and mixing on the surface of constant energy, then equilibrium statistical mechanics

predicts that the equilibrium distribution is the microcanonical distribution on the surface of constant energy. Then (53) (written in terms of $\mathrm{Re}\,\hat{u}_k$ and $\mathrm{Im}\,\hat{u}_k$) is a special case of (1), and its behaviour as $\varepsilon \to 0$ can be analysed by means of propositions 2.3 and 2.4.

Since the timescale separation is only approximate in the original TBH, the grouping into fast and slow modes in SA-TBH (53) is somewhat arbitrary. In section 4.1, we consider the first Fourier coefficient as the only slow mode of the TBH system; in section 4.2, we consider the first two Fourier coefficient as slow modes and compare the results with the original TBH system.

4.1. One slow mode, $\hat{u}_1$

Taking the first Fourier coefficient as the only slow mode amounts to choosing $\Lambda_1 = 1$, in which case the SA-TBH system in (53) reduces to

$$\begin{aligned} \dot{\hat{u}}_1 &= -\frac{\mathrm{i}}{2\varepsilon} \sum_{\substack{p+q+1=0 \\ 2\leqslant|p|,|q|\leqslant\Lambda}} \hat{u}_p^* \hat{u}_q^*, \\ \dot{\hat{u}}_k &= -\frac{\mathrm{i}k}{2\varepsilon}(\hat{u}_{k+1}\hat{u}_1^* + \hat{u}_{k-1}\hat{u}_1) - \frac{\mathrm{i}k}{2\varepsilon^2} \sum_{\substack{k+p+q=0 \\ 2\leqslant|p|,|q|\leqslant\Lambda}} \hat{u}_p^* \hat{u}_q^*, \qquad (k \geqslant 2), \end{aligned} \tag{55}$$

where we used the reality condition in (45) to simplify the right-hand side of the second equation. Note that the slowest of the fast modes, $\hat{u}_2$, is only twice as fast as the designated slow mode, $\hat{u}_1$, when $\varepsilon = 1$

$$\frac{\mathrm{corr\_time}\{\hat{u}_1\}}{\mathrm{corr\_time}\{\hat{u}_2\}} \approx 2.$$

Of course, the situation changes when $\varepsilon \to 0$, and we now investigate the effect of this operation on the statistical behaviour of $\hat{u}_1$.

We perform direct numerical simulations of the equations in (55) with three values of ε

$$\varepsilon = 0.5, \qquad \varepsilon = 0.25, \qquad \varepsilon = 0.1,$$

and compare them with the original system with $\varepsilon = 1$. The other parameters were chosen to be

$$\Lambda = 20, \qquad E = 0.4 \qquad (\beta = 50). \tag{56}$$

(Note that the energy can be chosen arbitrarily since the dynamics on one energy shell can be re-mapped onto another energy shell by appropriate rescaling of time (see (41)).) The time-step has to be adjusted for smaller values of ε to conserve energy up to appropriate precision. In simulations with $\varepsilon = 0.1$, the energy is conserved up to 10^{-4} absolute error, 10^{-3} relative error. All statistics are computed as time-averages from a single microcanonical realization of length $T \approx 10^5$.

The behaviour of correlation functions for various values of ε is presented in figure 2. There is a significant difference between the simulations with $\varepsilon = 1$ and $\varepsilon = 0.5$; a small difference between $\varepsilon = 0.5$ and $\varepsilon = 0.25$ and almost no difference between $\varepsilon = 0.25$ and $\varepsilon = 0.1$. This demonstrates the convergence of the correlation functions in the limit as $\varepsilon \to 0$. However, the shape of the correlation functions in this limit also shows that the artificial acceleration in TBH does have an effect on the dynamics. The correlation functions at $\varepsilon = 0.5, \ldots, 0.1$ are close to exponential, while the correlation function in the original TBH systems has a complicated shape with the 'bump' in the middle and smoother behaviour at zero. In addition, the correlation function changes very fast with ε when ε is close to 1, as can be seen in figure 3. Nevertheless, the mean decay rate is reproduced correctly by the simulations

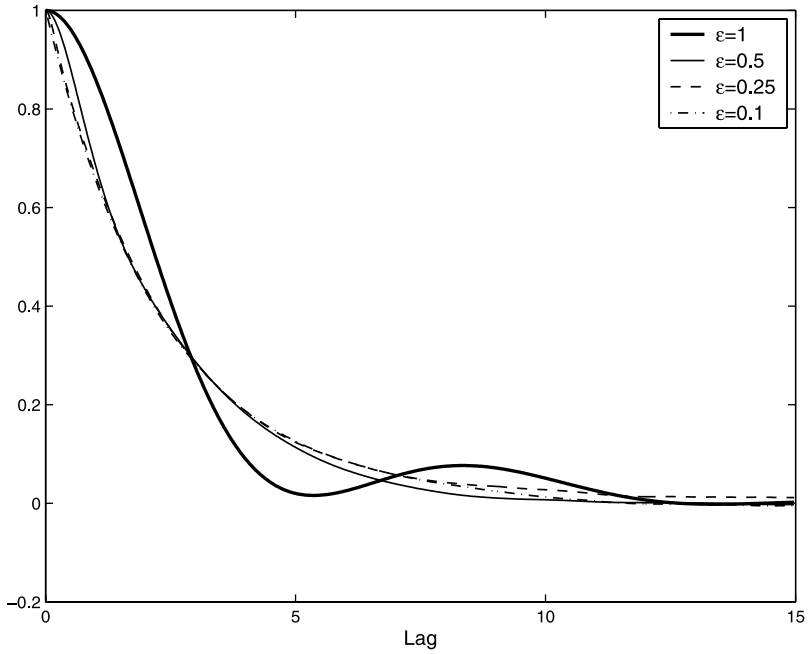

Figure 2. Normalized correlation function of $\mathrm{Re}\,\hat{u}_1$ from the simulations of the SA-TBH system in (55) (one slow mode). Solid bold line: $\varepsilon = 1$; solid line: $\varepsilon = 0.5$; dashed line: $\varepsilon = 0.25$; dash–dotted line: $\varepsilon = 0.1$. The graphs for $\varepsilon = 0.25$ and $\varepsilon = 0.1$ overlap almost completely.

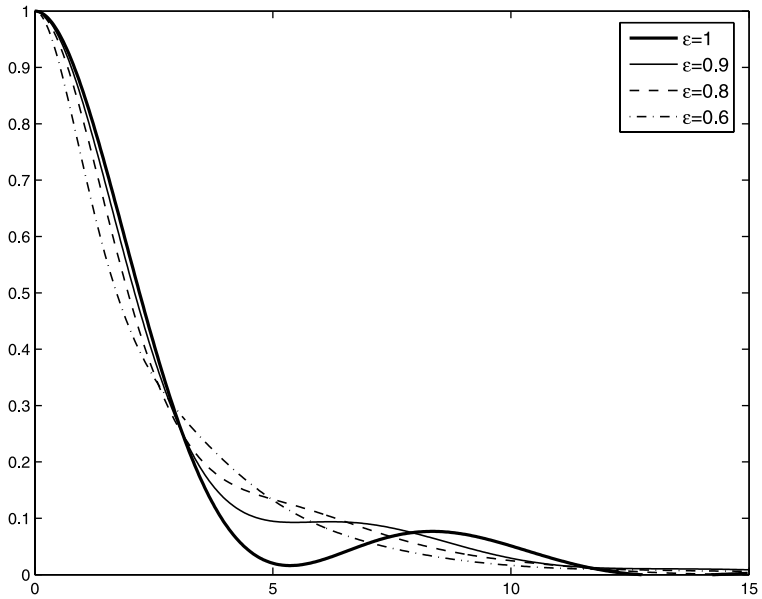

Figure 3. Normalized correlation function of $\mathrm{Re}\,\hat{u}_1$ from the simulations of the SA-TBH system in (55) (one slow mode). Solid bold line: $\varepsilon = 1$; solid line: $\varepsilon = 0.9$; dashed line: $\varepsilon = 0.8$; dash–dotted line: $\varepsilon = 0.6$. The shape of the correlation function seems to evolve continuously with ε and by $\varepsilon = 0.6$ and the 'bump' structure of the original model with $\varepsilon = 1$ is completely lost.

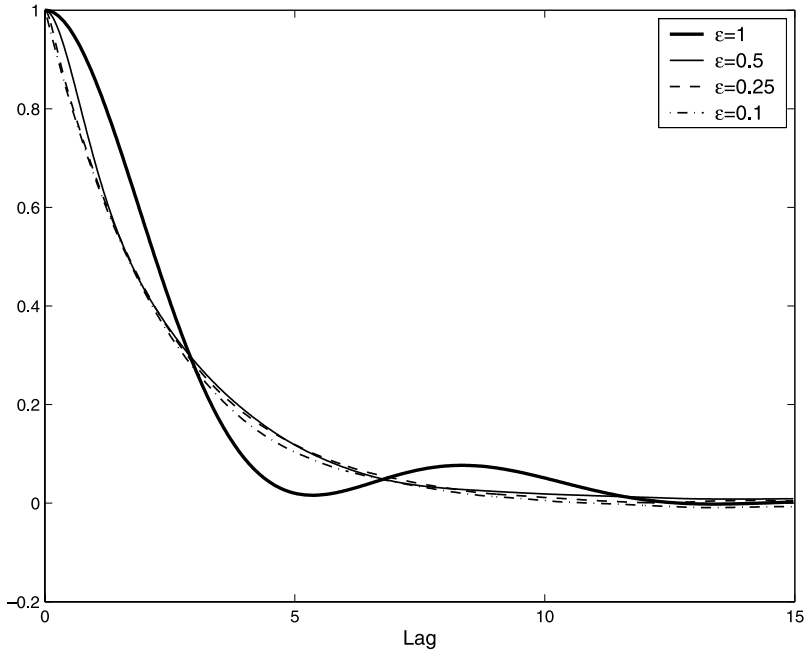

Figure 4. Normalized correlation function of $\mathrm{Re}\,\hat{u}_1$ from the simulations of the SA-TBH system in (57) (two slow modes). Solid bold line: $\varepsilon = 1$; solid line: $\varepsilon = 0.5$; dashed line: $\varepsilon = 0.25$; dash–dotted line: $\varepsilon = 0.1$.

of the accelerated model. The correlation times (area under the graph of the corresponding normalized correlation function) of $\mathrm{Re}\,\hat{u}_1$ in the simulation with $\varepsilon = 1$ and $\varepsilon = 0.1$ are

$$\mathrm{corr\_time}\{\hat{u}_1\}^{\varepsilon=1} \approx 2.63, \qquad \mathrm{corr\_time}\{\hat{u}_1\}^{\varepsilon=0.1} \approx 2.4.$$

4.2. Two slow modes, $\hat{u}_1$ and $\hat{u}_2$

Here the first two Fourier coefficients are selected as slow modes, which amounts to taking $\Lambda_1 = 2$. In this case the SA-TBH system in (53) reduces to

$$\begin{aligned}
\dot{\hat{u}}_1 &= -\mathrm{i}\hat{u}_2\hat{u}_1^* - \frac{\mathrm{i}}{2\varepsilon} \sum_{\substack{p+q+1=0 \\ 3\leqslant|p|,|q|\leqslant\Lambda}} \hat{u}_p^*\hat{u}_q^*, \\
\dot{\hat{u}}_2 &= -\mathrm{i}\hat{u}_1^2 - \frac{\mathrm{i}}{\varepsilon} \sum_{\substack{p+q+2=0 \\ 3\leqslant|p|,|q|\leqslant\Lambda}} \hat{u}_p^*\hat{u}_q^*, \qquad (57) \\
\dot{\hat{u}}_k &= -\frac{\mathrm{i}k}{2\varepsilon}(\hat{u}_{k+1}\hat{u}_1^* + \hat{u}_{k-1}\hat{u}_1) - \frac{\mathrm{i}k}{2\varepsilon}(\hat{u}_{k+2}\hat{u}_2^* + \hat{u}_{k-2}\hat{u}_2) - \frac{\mathrm{i}k}{2\varepsilon^2} \sum_{\substack{p+q+k=0 \\ 3\leqslant|p|,|q|\leqslant\Lambda}} \hat{u}_p^*\hat{u}_q^* \qquad (k \geqslant 3).
\end{aligned}$$

We perform simulations of this system with three values of ε

$$\varepsilon = 0.5, \qquad \varepsilon = 0.25, \qquad \varepsilon = 0.1$$

and the same values of $\beta = 50$, $\Lambda = 20$ and $E = 0.4$ as in the previous example (see (56)). The correlation functions of $\mathrm{Re}\,\hat{u}_1$ and $\mathrm{Re}\,\hat{u}_2$ are shown in figures 4 and 5. The trend is very similar

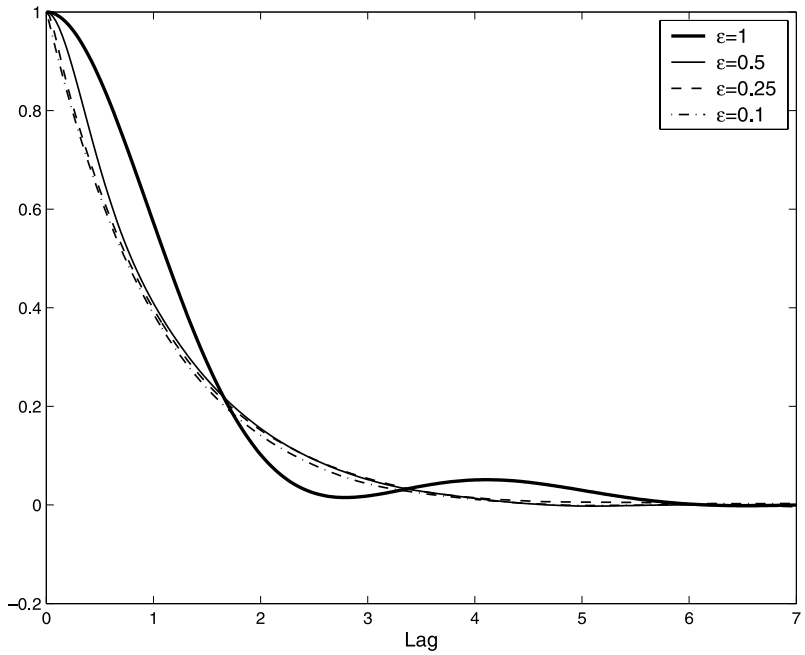

Figure 5. Normalized correlation function of $\mathrm{Re}\,\hat{u}_2$ from the simulations of the SA-TBH system in (57) (two slow modes). Solid bold line: $\varepsilon = 1$; solid line: $\varepsilon = 0.5$; dashed line: $\varepsilon = 0.25$; dash–dotted line: $\varepsilon = 0.1$. Note that the abscissa scale is different from that of figure 4.

to what was observed in section 4.2. The numerical results corroborate that the dynamics of the slow modes in (57) has a limit as $\varepsilon \to 0$, even though the acceleration has an effect on the shape of correlation functions. However, the bulk properties of these correlations functions, such as the correlation times, are rather insensitive to ε:

$$\text{corr\_time}\{\hat{u}_1\}^{\varepsilon=1} \approx 2.63, \qquad \text{corr\_time}\{\hat{u}_1\}^{\varepsilon=0.1} \approx 2.3,$$
$$\text{corr\_time}\{\hat{u}_2\}^{\varepsilon=1} \approx 1.28, \qquad \text{corr\_time}\{\hat{u}_2\}^{\varepsilon=0.1} \approx 1.04.$$

5. Stochastic models for TBH

In this section we use propositions 2.3 and 2.4 to derive effective SDEs for $\hat{u}_1$ and $\{\hat{u}_1, \hat{u}_2\}$ by taking the limit as $\varepsilon \to 0$ on (55) and (57), respectively. In principle, the solutions of these SDEs should behave similarly to the solutions of (55) and (57) at small ε which we described in sections 4.1 and 4.2. In practice, however, additional discrepancies can be introduced because the coefficients in the SDEs are obtained numerically with finite precision only. We discuss this issue below.

5.1. Stochastic model for $\hat{u}_1$

In order to write the SDEs in a more compact form we denote the slow variables as

$$a = (a_1, a_2) \equiv (\mathrm{Re}\,\hat{u}_1, \mathrm{Im}\,\hat{u}_1). \tag{58}$$

Then the SDE in (21) obtained from (55) in the limit as $\varepsilon \to 0$ can be written explicitly as:

$$\mathrm{d}a_k = B(a)a_k\,\mathrm{d}t + H(a)a_k\,\mathrm{d}t + \sqrt{2}\sigma(a)\mathrm{d}W_k(t), \qquad (k = 1, 2), \tag{59}$$

where

$$B(a) = -(1 - N^{-1})\mathcal{E}^{-1}(a)(\mathcal{E}^{1/2}(a)I_2|a|^2 + \mathcal{E}^{3/2}(a)I_f),$$
$$H(a) = -N^{-1}\mathcal{E}^{-1/2}(a)|a|^2 I_2 + 2\mathcal{E}^{1/2}(a)I_2 - 3N^{-1}\mathcal{E}^{1/2}(a)I_f, \tag{60}$$
$$\sigma^2(a) = \mathcal{E}^{1/2}(a)|a|^2 I_2 + \mathcal{E}^{3/2}(a)I_f,$$

where $B(a)$ and $H(a)$ are the drift and Itô terms in (21), respectively. The first and second terms in $B(a)$ arise, respectively, from the terms involving $Q^{(1)}_{kk'}$ and $Q^{(3)}_{jkj'k'}$ in (24). $\mathcal{E}(a)$ denotes the energy per mode of the fast subsystem, i.e.

$$\mathcal{E}(a) = N^{-1}(E - |a|^2), \tag{61}$$

where $N = 2\Lambda - 2$ is the number of fast degrees of freedom and E is the total energy of the full TBH model.

The last term involving $Q^{(3)}_{jkj'k'}$ in (23) can be recast as the cross-correlation between right-hand sides of the slow variables projected onto the fast dynamics alone. Therefore, in (60), we have also defined

$$I_2 = I[\mathrm{Re}\,\hat{u}_2, \mathrm{Re}\,\hat{u}_2] = I[\mathrm{Im}\,\hat{u}_2, \mathrm{Im}\,\hat{u}_2],$$
$$I_f = I[f^r, f^r] = I[f^i, f^i]. \tag{62}$$

where $I[\cdot,\cdot]$ is a short-hand notation for the area under the graph of a correlation function

$$I[g,h] = \int_0^\infty \langle g(t)h(t+\tau)\rangle_t \,\mathrm{d}\tau, \tag{63}$$

where $\langle\cdot\rangle_t$ denotes the temporal average and

$$f^r(t) = \mathrm{Re}\left(-\frac{\mathrm{i}}{2}\sum_{\substack{p+q+1=0\\ 2\leqslant|p|,|q|\leqslant\Lambda}} \hat{u}_p^*\hat{u}_q^*\right),$$
$$f^i(t) = \mathrm{Im}\left(-\frac{\mathrm{i}}{2}\sum_{\substack{p+q+1=0\\ 2\leqslant|p|,|q|\leqslant\Lambda}} \hat{u}_p^*\hat{u}_q^*\right) \tag{64}$$

denote the real and the imaginary parts of the right-hand side of the equation for $\hat{u}_1$ in (55). The various correlation functions in the expressions above must be computed on the fast subsystem (4), which in the present situation corresponds to a TBH system with wave numbers $2 \leqslant |k| \leqslant \Lambda$

$$\dot{\hat{u}}_k = -\frac{\mathrm{i}k}{2}\sum_{\substack{k+p+q=0\\ 2\leqslant|p|,|q|\leqslant\Lambda}} \hat{u}_p^*\hat{u}_q^*, \qquad (2 \leqslant |k| \leqslant \Lambda). \tag{65}$$

The derivation of (59) is somewhat tedious but straightforward. In addition to the general assumptions stated in section 2 the derivation utilizes specific properties of the TBH system to simplify the expressions for the stochastic model further. First, it uses the fact that the correlations of the real and imaginary parts of the same mode are identical by symmetry. Second, it utilizes the property (verified to a very good precision for large Λ in [19, 20]) that the joint distributions of any two modes is Gaussian with a diagonal correlation matrix. Therefore, all third moments can be neglected, which implies, in particular, that the terms involving $Q^{(2)}_{ijk}$ in (23) are zero. Finally, it uses the fact that the cross-correlation between f^r and f^i and $\mathrm{Re}\,\hat{u}_2$ and $\mathrm{Im}\,\hat{u}_2$ are negligible

$$\langle f^r(0)\, f^i(t)\rangle_t = \langle \mathrm{Re}\,\hat{u}_2(0)\,\mathrm{Im}\,\hat{u}_2(t)\rangle_t \approx 0, \tag{66}$$

which is also verified to a very good precision in the simulations (see figure 6).

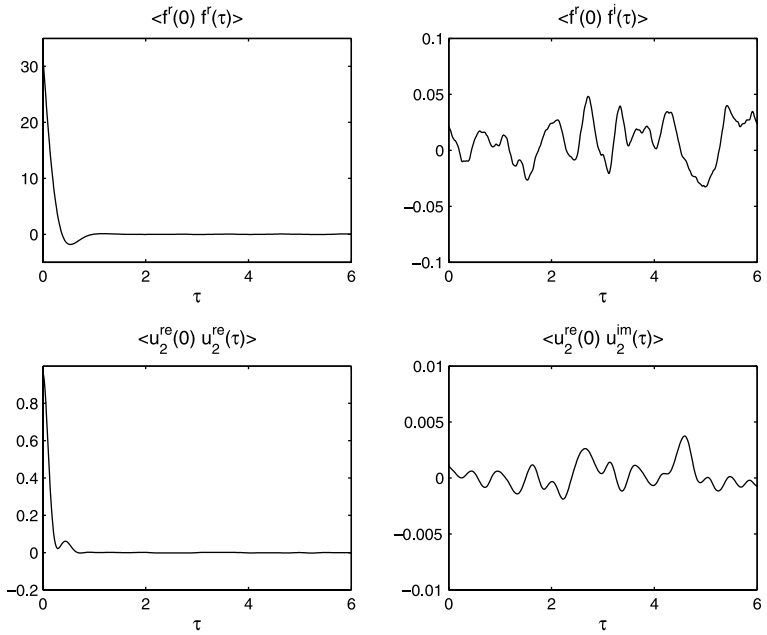

Figure 6. Autocorrelation functions from the simulations of the fast subsystem in (65) with $\Lambda = 20$ and $\beta = 1/2$. Left panel—correlation functions of f^r and $\mathrm{Re}\,\hat{u}_2$; right panel—cross-correlations between the real and imaginary components $\langle f^r(0) f^i(\tau)\rangle$ and $\langle \hat{u}_2^r(0)\hat{u}_2^i(\tau)\rangle$.

Table 2. Area (integral with respect to t) under the graph of the corresponding correlation function computed from the simulations of the auxiliary fast subsystem for $\Lambda = 20$; fluctuations are within 1.5%.

	Run #1	Run #2	Run #3	Run #4
I_2	0.141	0.139	0.141	0.14
I_f	4.35	4.27	4.3	4.28

5.1.1. Numerical evaluation of the coefficients. To evaluate the various coefficient in (59), we integrate the fast subsystem in (65) numerically with $\Lambda = 20$ and energy $E = 38$ ($\beta = 1/2$), so that $\mathrm{var}\{\mathrm{Re}\,\hat{u}_k\} = \mathrm{var}\{\mathrm{Im}\,\hat{u}_k\} = 1$ and compute all necessary two-time statistics. The simulations were run for the total time of $T \approx 120\,000$ with a sufficiently small time-step to conserve energy up to 10^{-4} relative error. In order to test the robustness of numerical estimates we performed several runs with different random initial conditions and verified that the time-averages from individual microcanonical realizations coincide. Note that the tail behaviour of the correlation functions makes significant contributions to the correlation functions in (62), and therefore must be computed accurately over relatively long time-lags.

The numerical estimates for the integrals of the two-time statistics in the simulations with of the fast subsystem in (65) are presented in table 2. The comparison of selected cross-correlations with auto-correlation functions of $\mathrm{Re}\,\hat{u}_2$ and f^r for the regime $\Lambda = 20$ is presented in figure 6. The difference between auto-correlation and cross-correlation functions is two orders of magnitude, which indicates that the correlation matrix is approximately diagonally consistent with (66).

Table 3. Estimates of the area (integral with respect to t) under the graph of the corresponding correlation function computed from the simulations of the auxiliary fast subsystem for $\Lambda = 20$ and $\Lambda = 40$.

	$\Lambda = 20$	$\Lambda = 40$
I_2	0.14	0.092
I_f	4.3	6.1

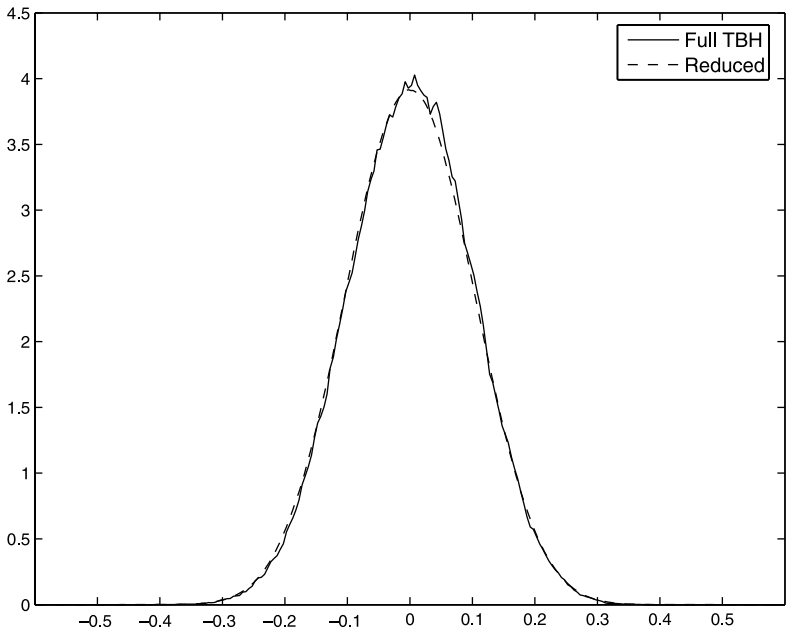

Figure 7. Marginal probability density function of $\mathrm{Re}\,\hat{u}_1$ in the simulations of the original TBH system (46) with $\Lambda = 20$ (——) and the corresponding SDE in (59) (- - - -).

We also verified that the coefficients in (59) can be estimated for a larger value of Λ, using $\Lambda = 40$ instead of $\Lambda = 20$. The estimates for I_2 and I_f obtained from simulations of (65) on the energy surface with $\beta = 1/2$ are presented in table 3.

5.1.2. Statistical behaviour of the stochastic model. The comparison between the direct numerical simulations of the original TBH system in (46) and the SDEs in (59) is depicted in figures 7–9. The one-time statistics is Gaussian in both simulations with perfect agreement between the simulations of the full TBH system and the stochastic model. To illustrate this, a marginal distribution of $\mathrm{Re}\,\hat{u}_1$ is presented in figure 7.

Unlike the one-time statistics, the correlation functions of $\mathrm{Re}\,\hat{u}_1$ and $\mathrm{Im}\,\hat{u}_1$ differ considerably between the original TBH system and the stochastic model. The detailed structure of the correlations is no longer represented in the stochastic model. Instead, correlation functions of $\mathrm{Re}\,\hat{u}_1$ and $\mathrm{Im}\,\hat{u}_1$ are exponentials with the averaged rate of decay reflecting the decorrelation times of the full model. But, as expected, the correlation functions of the stochastic model agree with the simulations of the SA-TBH system in (55) within a few per cent. The correlation functions of $\mathrm{Re}\,\hat{u}_1$ for truncation sizes $\Lambda = 20$ and $\Lambda = 40$ in the simulation of the original TBH system, the SA-TBH system with $\varepsilon = 0.1$ and the corresponding stochastic

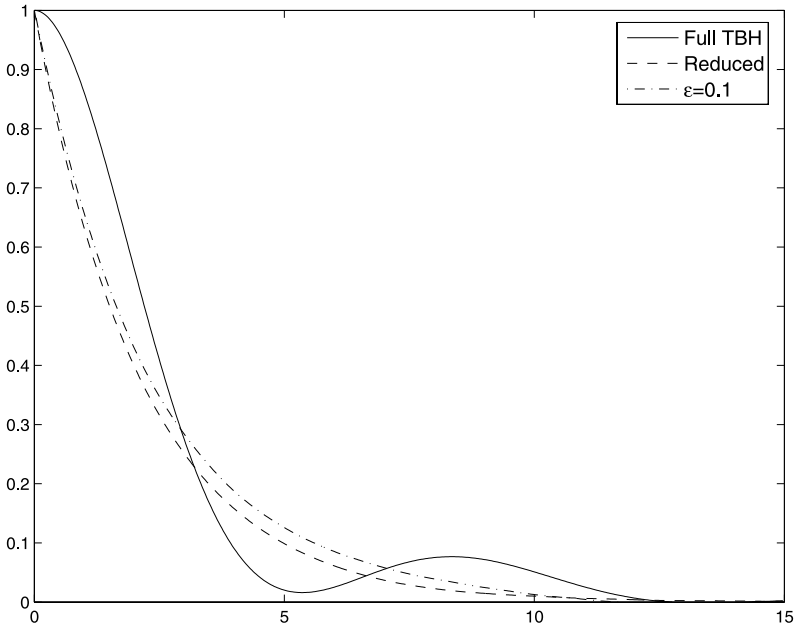

Figure 8. Normalized correlation function of Re $\hat{u}_1$ in the simulation of the original TBH system in (46) with $\Lambda = 20$ (——), the SA-TBH system in (55) with $\varepsilon = 0.1$ (— · —) and the SDE in (59) (- - - -).

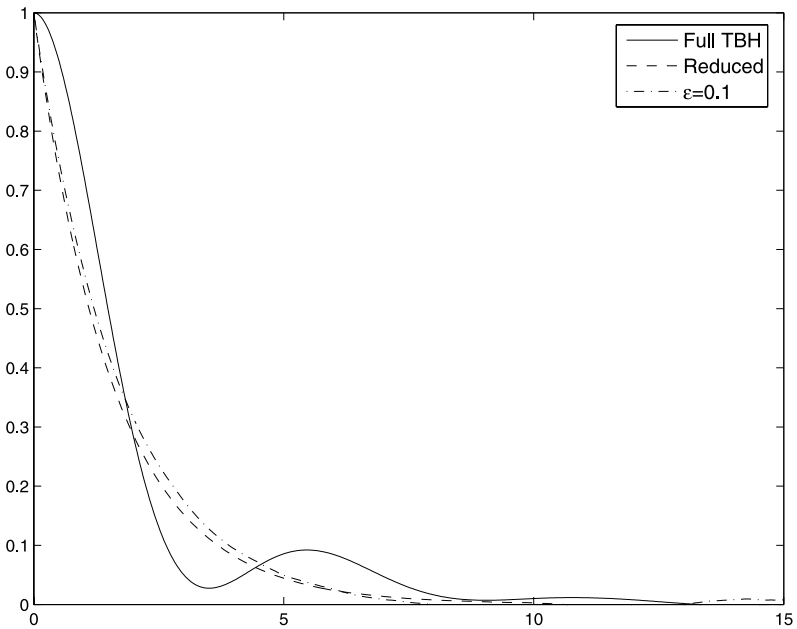

Figure 9. Normalized correlation function of Re $\hat{u}_1$ in the simulation of the original TBH system in (46) with $\Lambda = 40$ (——), the SA-TBH system in (55) with $\varepsilon = 0.1$ (— · —) and the SDE in (59) (- - - -).

Table 4. Estimates for decorrelation times (1/area of the normalized correlation function) of $\mathrm{Re}\,\hat{u}_1$ in simulations with $\Lambda = 20$ and $\Lambda = 40$.

	$\Lambda = 20$	$\Lambda = 40$
Original TBH	2.63	1.81
SDE in (59)	2.17	1.61
SA-TBH ($\varepsilon = 0.1$)	2.38	1.84

Table 5. Estimates of the area (integral with respect to t) under the graph of the corresponding correlation function computed from the simulations of the auxiliary fast subsystem for $\Lambda = 20$ and $\Lambda = 40$.

	$\Lambda = 20$	$\Lambda = 40$
I_2	0.099	0.063
I_3	0.076	0.048
I_{f_1}	4.18	6
I_{f_2}	9.1	13.35

Table 6. Estimates for correlation times (1/area of the normalized correlation function) of $\mathrm{Re}\,\hat{u}_1$ and $\mathrm{Re}\,\hat{u}_2$ in simulations with $\Lambda = 20$ and $\Lambda = 40$.

	$\Lambda = 20$		$\Lambda = 40$	
	$\hat{u}_1$	$\hat{u}_2$	$\hat{u}_1$	$\hat{u}_2$
Original TBH	2.63	1.27	1.82	0.88
SDE	2.14	0.95	1.59	0.71
Accelerated TBH ($\varepsilon = 0.1$)	2.24	1.03	1.69	0.82

models is depicted in figures 8 and 9, respectively. In addition, the decorrelation times are presented in table 4.

Finally, we observed in the simulations that the original TBH system, the SA-TBH system and the stochastic model are nearly Gaussian (i.e. non-Gaussian corrections are too small to be measured accurately). This means that the two-time statistical behaviour of the system is completely specified by the correlation function presented above.

5.2. *Stochastic model for $\hat{u}_1$ and $\hat{u}_2$*

Denoting the slow modes as

$$a = (a_1, a_2, a_3, a_4) = (\mathrm{Re}\,\hat{u}_1, \mathrm{Im}\,\hat{u}_1, \mathrm{Re}\,\hat{u}_2, \mathrm{Im}\,\hat{u}_2), \tag{67}$$

the stochastic model for $\hat{u}_1$, $\hat{u}_2$ can be written compactly as follows:

$$\mathrm{d}a_k = L_k(a)\mathrm{d}t + B_k(a)\mathrm{d}t + H_k(a)\mathrm{d}t + (\sigma(a)\mathrm{d}W)_k, \qquad (k = 1, \ldots, 4). \tag{68}$$

The explicit forms of the coefficients $L(a)$, etc are given in the appendix, see (A.1)–(A.6). Despite the complex appearance of these coefficients, the SDEs in (68) are explicit and can be

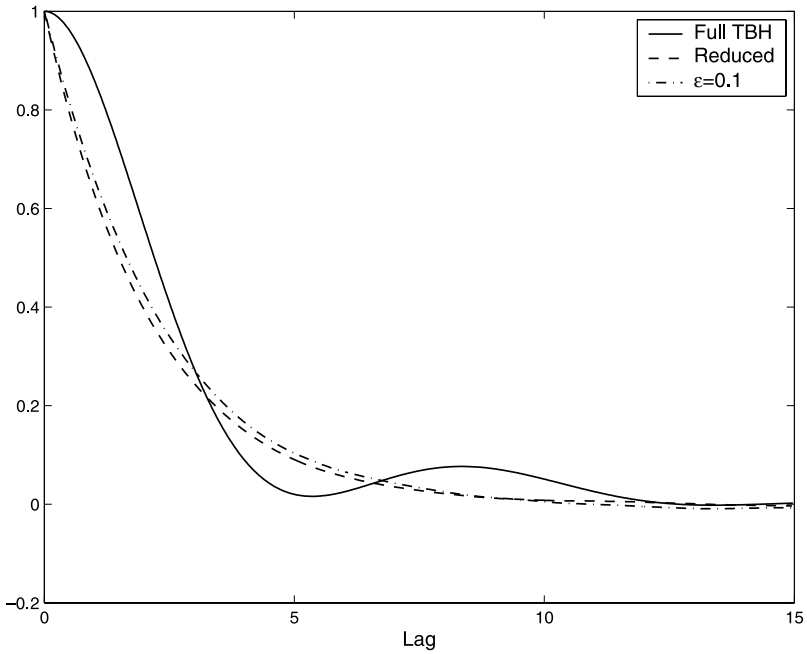

Figure 10. Normalized correlation function of Re $\hat{u}_1$ in the simulation of the original TBH system in (46) with $\Lambda = 20$ (——), the SA-TBH system in (57) with $\varepsilon = 0.1$ (— · —) and the SDE in (68) (- - - -).

easily simulated numerically. These coefficients involve the following correlation functions

$$\begin{aligned} I_2 &= I[\operatorname{Re}\hat{u}_2, \operatorname{Re}\hat{u}_2] = I[\operatorname{Im}\hat{u}_2, \operatorname{Im}\hat{u}_2], \\ I_3 &= I[\operatorname{Re}\hat{u}_3, \operatorname{Re}\hat{u}_3] = I[\operatorname{Im}\hat{u}_3, \operatorname{Im}\hat{u}_3], \\ I_{f_1} &= I[f_1^r, f_1^r] = I[f_1^i, f_1^i], \\ I_{f_2} &= I[f_2^r, f_2^r] = I[f_2^i, f_2^i], \end{aligned} \tag{69}$$

where the notation $I[f, h]$ was introduced in (63) and $f_{1,2}^{r,i}$ are the corresponding parts of the right-hand sides for modes $\hat{u}_{1,2}^{r,i}$ defined as follows:

$$\begin{aligned} f_1^r(t) &= \operatorname{Re}\left(-\frac{\mathrm{i}}{2}\sum_{\substack{p+q+1=0 \\ 2\leqslant|p|,|q|\leqslant\Lambda}} \hat{u}_p^*\hat{u}_q^*\right), \\ f_1^i(t) &= \operatorname{Im}\left(-\frac{i}{2}\sum_{\substack{p+q+1=0 \\ 2\leqslant|p|,|q|\leqslant\Lambda}} \hat{u}_p^*\hat{u}_q^*\right), \end{aligned} \tag{70}$$

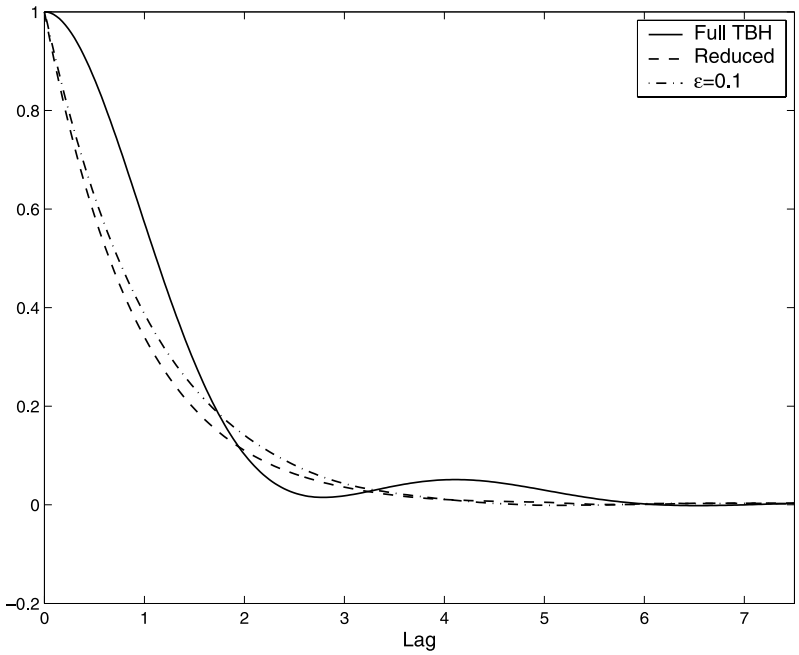

Figure 11. Normalized correlation function of Re $\hat{u}_2$ in the simulation of the original TBH system in (46) with $\Lambda = 20$ (——), the SA-TBH system in (57) with $\varepsilon = 0.1$ (— · —) and the SDE in (68) (- - - -).

$$f_2^r(t) = \mathrm{Re}\left(-\frac{i}{2}\sum_{\substack{p+q+2=0\\ 2\leqslant|p|,|q|\leqslant\Lambda}} \hat{u}_p^*\hat{u}_q^*\right),$$
$$f_2^i(t) = \mathrm{Im}\left(-\frac{i}{2}\sum_{\substack{p+q+2=0\\ 2\leqslant|p|,|q|\leqslant\Lambda}} \hat{u}_p^*\hat{u}_q^*\right). \tag{71}$$

5.2.1. Numerical evaluation of correlations. The correlation in (69) must be computed from the fast subsystem, which is identical to (65) except that the wave numbers all run over $3 \leqslant |k| \leqslant \Lambda$. Similarly to the previous section, resolving correlation functions for high wavenumbers is a somewhat challenging computational task. We performed several runs with different initial conditions and estimated the decorrelation times in (69) up to 1.5% relative error. Numerical estimates for parameters in (69) are presented in table 5.

5.2.2. Statistical behaviour of the stochastic model. The correlation functions obtained from the simulations with $\Lambda = 20$ and $\Lambda = 40$ of the original TBH system in (46), the SA-TBH system in (57) with $\varepsilon = 0.1$ and the SDE in (68) are shown in figures 10, 11, 12, 13, respectively. In addition, the correlation times for $\hat{u}_1$ and $\hat{u}_2$ are presented in table 6. The correlation functions in the simulations of the SA-TBH system and the stochastic model are in very good agreement, but they are exponential and do not reproduce the complicated 'bump' structure of

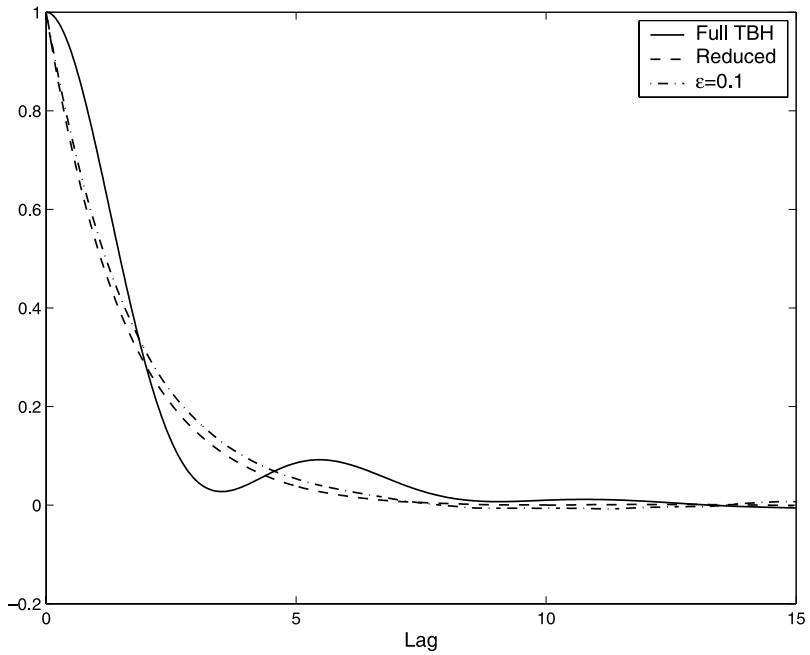

Figure 12. Normalized correlation function of Re $\hat{u}_1$ in the simulation of the original TBH system in (46) with $\Lambda = 40$ (——), the SA-TBH system in (57) with $\varepsilon = 0.1$ (— · —) and the SDE in (68) (- - - -).

the original TBH system. Nevertheless, the overall decay rate of the correlation functions is captured correctly by both the SA-TBH system and the stochastic model. The discrepancies between the modified system are approximately 6.8% for mode Re $\hat{u}_1$ and 12.3% for mode Re $\hat{u}_2$. Finally, consistent with the results presented in section 5.1, the two-time statistics is nearly Gaussian in all the three models.

6. Conclusions

A modified stochastic mode-reduction strategy for conservative systems was presented. One of the main advantages of the current approach is that no *ad hoc* modifications of the underlying equations are necessary. Under assumptions of mixing and ergodicity, the procedure gives closed-form SDEs for the slow dynamics which are exact in the limit of infinite timescale separation between fast and slow modes. Only bulk statistical quantities of the fast dynamics enter the stochastic equations as coefficients and these can be computed for all energy levels from a single microcanonical realization on an auxiliary subsystem.

In any realistic system, the separation of timescale is only approximate. In this case, the stochastic model captures the behaviour of the slow modes in a system where the fast modes have been artificially accelerated. This viewpoint allows us, at least in principle, to test the validity and relevance of the stochastic model by assessing the impact of the artificial acceleration on the original dynamics. This approach was tested here on the TBH system. It was shown that the statistical properties of the slow modes in the SA-TBH system are, in the bulk if not in the detail, similar to the properties of these modes in the original TBH system. As a result, the stochastic models with only one or two modes retained out of 102 perform

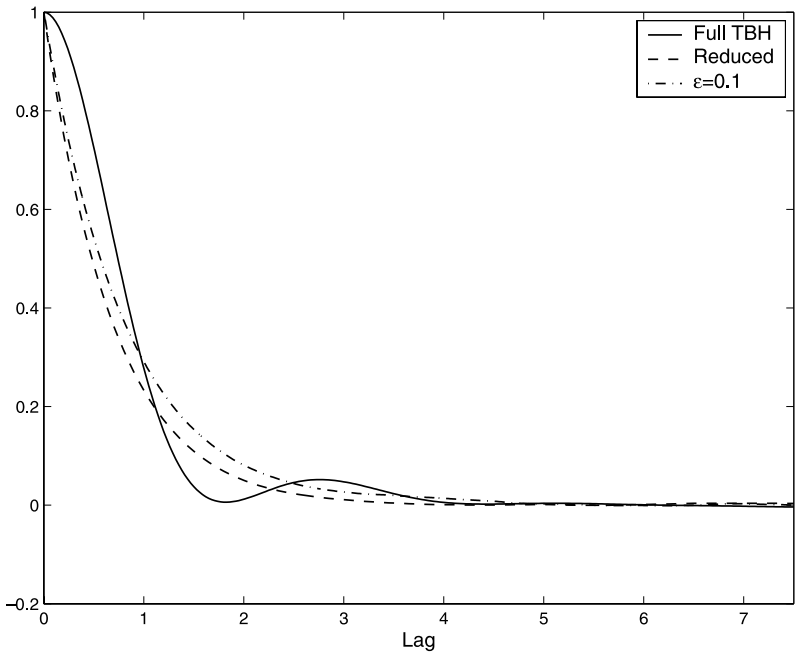

Figure 13. Normalized correlation function of Re $\hat{u}_2$ in the simulation of the original TBH system in (46) with $\Lambda = 40$ (——), the SA-TBH system in (57) with $\varepsilon = 0.1$ (— · —) and the SDE in (68) (- - - -).

surprisingly well. The transportability of these conclusions to other systems is difficult to test, but they offer hope that the stochastic mode-elimination approach is applicable to problems without substantial timescale separation, as is the case in most applications of interest.

Acknowledgments

AM acknowledges the support of the NSF for Grant DMS-9972865, the ONR for Grant N00014-96-1-0043 and the NSF-CMG for Grant DMS02-22133. IT acknowledges the support of the NSF for Grants DMS-0405944 and ATM-0417867 and the DOE for Grant DE-FG02-04ER25645. EV-E acknowledges the support of the NSF for Grants DMS01-01439, DMS02-09959 and DMS02-39625 and of the ONR for Grant N-00014-04-1-0565.

Appendix A. Coefficients in (68)

Here we give the explicit forms of the coefficients in (68). $L(a)$ represents the nonlinear self-interactions between the slow modes, $\hat{u}_1$ and $\hat{u}_2$, i.e.

$$L(a) = \begin{pmatrix} a_1a_4 - a_2a_3 \\ -a_1a_3 - a_2a_4 \\ 2a_1a_2 \\ a_2^2 - a_1^2 \end{pmatrix} \equiv \begin{pmatrix} \hat{u}_1^r\hat{u}_2^i - \hat{u}_1^i\hat{u}_2^r \\ -\hat{u}_1^r\hat{u}_2^r - \hat{u}_1^i\hat{u}_2^i \\ 2\hat{u}_1^r\hat{u}_1^i \\ (\hat{u}_1^i)^2 - (\hat{u}_1^r)^2 \end{pmatrix} \tag{A.1}$$

$B(a)$ is the drift term in (21)

$$B_i(a) = -\frac{(1-2/N)}{\mathcal{E}(a)} \sum_{j=1}^{4} D_{ij}(a) a_j.$$

Here it assumes the following explicit form:

$$B(a) = -\frac{(1-2/N)}{\mathcal{E}(a)} \begin{pmatrix} C_1(a)a_1 + C_3(a)a_3 + C_4(a)a_4 \\ C_1(a)a_2 - C_4(a)a_3 + C_3(a)a_4 \\ C_3(a)a_1 - C_4(a)a_2 + C_2(a)a_3 \\ C_4(a)a_1 + C_3(a)a_2 + C_2(a)a_4 \end{pmatrix}, \tag{A.2}$$

where

$$\begin{aligned} C_1(a) &= \mathcal{E}^{1/2}(a)(a_3^2 + a_4^2)I_3 + \mathcal{E}^{3/2}(a)I_{f_1}, \\ C_2(a) &= 4\mathcal{E}^{1/2}(a)((a_1^2 + a_2^2)I_3 + \mathcal{E}^{3/2}(a)I_{f_2} + (a_3^2 + a_4^2)I_4), \\ C_3(a) &= 2\mathcal{E}^{1/2}(a)(a_1a_3 + a_2a_4)I_3, \\ C_4(a) &= 2\mathcal{E}^{1/2}(a)(a_1a_4 - a_2a_3)I_3. \end{aligned}$$

$D(a)$ is the diffusion matrix, which in the present case can be written as follows:

$$D(a) = \begin{pmatrix} C_1(a) & 0 & C_3(a) & C_4(a) \\ 0 & C_1(a) & -C_4(a) & C_3(a) \\ C_3(a) & -C_4(a) & C_2(a) & 0 \\ C_4(a) & C_3(a) & 0 & C_2(a) \end{pmatrix}. \tag{A.3}$$

Due to the special symmetries of the diffusion matrix it is easy to find the Cholesky decomposition of matrix in (A.3)

$$\sigma(a) = \begin{pmatrix} \sigma_1(a) & 0 & 0 & 0 \\ 0 & \sigma_1(a) & 0 & 0 \\ \sigma_2(a) & \sigma_3(a) & \sigma_4(a) & 0 \\ -\sigma_3(a) & \sigma_2(a) & 0 & \sigma_4(a) \end{pmatrix} \tag{A.4}$$

with

$$\begin{aligned} \sigma_1(a) &= \sqrt{C_1(a)}, & \sigma_2(a) &= \frac{C_3(a)}{\sqrt{C_1(a)}}, \\ \sigma_3(a) &= -\frac{C_4(a)}{\sqrt{C_1(a)}}, & \sigma_4(a) &= \sqrt{C_2(a) - \sigma_2^2(a) - \sigma_3^2(a)}. \end{aligned} \tag{A.5}$$

Finally, $H(a)$ is the Itô term given by

$$H_k(a) = \sum_{j=1}^{4} \frac{\partial}{\partial a_j} D_{kj}(a). \tag{A.6}$$

References

[1] Achatz U and Branstator G 1999 A two-layer model with empirical linear corrections and reduced order for studies of internal climate variability *J. Atmos. Sci.* **56** 3140–60

[2] Branstator G and Haupt S E 1995 An empirical model of barotropic atmospheric dynamics and its response to tropical forcing *J. Climate* **11** 2645–67

[3] Chorin A J, Kast A P and Kupferman R 1998 Optimal prediction of underresolved dynamics *Proc. Natl Acad. Sci. USA* **95** 4094–8

[4] Chorin A J, Kast A P and Kupferman R 1999 Unresolved computation and optimal prediction *Commun. Pure Appl. Math.* **52** 1231–54
[5] DelSole T and Farrel B F 1996 Quasi-linear equilibration of a thermally maintained, stochastically excited jet in a quasigeostrophic model *J. Atmos. Sci.* **53** 1781–97
[6] Egger J 2001 Master equations for climate parameter sets *Climate Dyn.* **18** 169–77
[7] Egger J and Hoinka K-P 2002 Covariance analysis of the global atmospheric axial angular momentum budget *Mon. Weather Rev.* **130** 1063–70
[8] Majda A J, Timofeyev I and Vanden-Eijnden E 1999 Models for stochastic climate prediction *Proc. Natl Acad. Sci. USA* **96** 14687–91
[9] Majda A J, Timofeyev I and Vanden-Eijnden E 2001 A mathematics framework for stochastic climate models *Commun. Pure Appl. Math.* **54** 891–974
[10] Sardeshmukh P D and Whitaker J A 1998 A linear theory of extratropical synoptic eddy statistics *J. Atmos. Sci.* **55** 237–58
[11] Schutte C, Walter J, Hartmann C and Huisinga W 2004 An averaging principle for fast degrees of freedom exhibiting long-term correlations *Multiscale Model. Simul.* **2** 501–26
[12] Kurtz T G 1973 A limit theorem for perturbed operators semigroups with applications to random evolution *J. Funct. Anal.* **12** 55–67
[13] Papanicolaou G 1976 Some probabilistic problems and methods in singular perturbations *Rocky Mountain J. Math.* **6** 653–73
[14] Majda A J and Timofeyev I 2004 Low dimensional chaotic dynamics versus intrinsic stochastic noise: a paradigm model *Physica* D **199** 339–68
[15] Majda A J, Timofeyev I and Vanden-Eijnden E 2002 A priori tests of a stochastic mode reduction strategy *Physica* D **170** 206–52
[16] Majda A J, Timofeyev I and Vanden-Eijnden E 2003 Systematic strategies for stochastic mode reduction in climate *J. Atmos. Sci.* **60** 1705–22
[17] Franzke C and Majda A J 2005 Low-order stochastic mode reduction for a prototype atmospheric GCM *J. Atmos. Sci.* submitted
[18] Franzke C, Majda A J and Vanden-Eijnden E 2005 Low-order stochastic mode reduction for a realistic barotropic model climate *J. Atmos. Sci.* **62** 1722–45
[19] Majda A J and Timofeyev I 2000 Remarkable statistical behavior for truncated Burgers–Hopf dynamics *Proc. Natl Acad. Sci. USA* **97** 12413–17
[20] Majda A J and Timofeyev I 2002 Statistical mechanics for truncations of the Burgers–Hopf equation: a model for intrinsic stochastic behavior with scaling *Milan J. Math.* **70** 39–96
[21] Vanden-Eijnden E 2003 Numerical techniques for multiscale dynamical systems with stochastic effects *Commun. Math. Sci.* **1** 385–91
[22] Fatkullin I and Vanden-Eijnden E 2004 A computational strategy for multiscale systems with application to Lorenz 96 model *J. Comput. Phys.* **200** 605–38
[23] Liu D, E W and Vanden-Eijnden E 2005 Analysis of multiscale methods for stochastic differential equations *Commun. Pure Appl. Math.* **58** 1544–85
[24] Papanicolaou G 1977 Introduction to the asymptotic analysis of stochastic equations *Modern modeling of continuum phenomena (Ninth Summer Sem. Appl. Math., Rensselaer Polytechnic Institute, Troy, NY, 1975*) (*Lectures in Applied Mathematics* vol 16) (Providence, RI: American Mathematical Society)
[25] Khasminsky R Z 1966 On stochastic processes defined by differential equations with a small parameter *Theory Probab. Appl.* **11** 211–28
[26] Khasminsky R Z 1966 A limit theorem for the solutions of differential equations with random right-hand sides *Theory Probab. Appl.* **11** 390–406
[27] Abramov R, Kovacic G and Majda A J 2003 Hamiltonian structure and statistically relevant conserved quantities for the truncated Burgers–Hopf equation *Commun. Pure Appl. Math.* **56** 1–46