A new mode reduction strategy for the generalized Kuramoto–Sivashinsky equation

IMA Journal of Applied Mathematics (2015) **80**, 273–301
doi:10.1093/imamat/hxt041
Advance Access publication on 24 October 2013

A new mode reduction strategy for the generalized Kuramoto–Sivashinsky equation

M. Schmuck
Department of Chemical Engineering, Imperial College, London SW7 2AZ, UK and Department of Mathematics, Imperial College, London SW7 2AZ, UK

M. Pradas
Department of Chemical Engineering, Imperial College, London SW7 2AZ, UK

G. A. Pavliotis
Department of Mathematics, Imperial College, London SW7 2AZ, UK

and

S. Kalliadasis*
Department of Chemical Engineering, Imperial College, London SW7 2AZ, UK
*Corresponding author: s.kalliadasis@imperial.ac.uk

[Received on 14 December 2012; revised on 1 August 2013; accepted on 24 September 2013]

Consider the generalized Kuramoto–Sivashinsky (gKS) equation. It is a model prototype for a wide variety of physical systems, from flame-front propagation, and more general front propagation in reaction–diffusion systems, to interface motion of viscous film flows. Our aim is to develop a systematic and rigorous low-dimensional representation of the gKS equation. For this purpose, we approximate it by a renormalization group equation which is qualitatively characterized by rigorous error bounds. This formulation allows for a new stochastic mode reduction guaranteeing optimality in the sense of maximal information entropy. Herewith, noise is systematically added to the reduced gKS equation and gives a rigorous and analytical explanation for its origin. These new results would allow one to reliably perform low-dimensional numerical computations by accounting for the neglected degrees of freedom in a systematic way. Moreover, the presented reduction strategy might also be useful in other applications where classical mode reduction approaches fail or are too complicated to be implemented.

Keywords: generalized Kuramoto–Sivashinsky equation; renormalization group method; stochastic mode reduction.

1. Introduction

We consider abstract evolution equations of the form:

$$\begin{aligned} \frac{\partial u}{\partial t} + B(u,u) + Au = 0, \\ u(x,0) = u_0(x), \end{aligned} \tag{1.1}$$

where A denotes a general linear operator and B represents a non-linear term of Burgers' type, i.e. $B(u) = B(u,u) = uu_x$. Well-known equations in this class include, e.g. the viscous Burgers equation,

the Korteweg–de Vries equation and the Benney–Lin equation. We start by performing a formal renormalization group (RG) approach for the general form in (1.1). We subsequently focus on a rigorous low-dimensional reduction of the generalized Kuramoto–Sivashinsky (gKS) equation:

$$\begin{aligned} \partial_t u + \lambda u u_x + \kappa u_{xx} + \delta u_{xxx} + \nu u_{xxxx} = 0 \quad &\text{in } \mathscr{P}_\alpha \times]0, T[, \\ u(x,0) = g(x) \quad &\text{in } \mathscr{P}_\alpha, \end{aligned} \tag{1.2}$$

where $\mathscr{P}_\alpha :=]-\alpha\pi, \alpha\pi[$ is a periodic domain with $\alpha := L/2\pi$ for an arbitrary period $L > 0$ while the solution $u(x,t) : \mathscr{P}_\alpha \times]0, T[\to \mathbb{R}$ of (1.2) represents, for example, the fluctuations around a fixed mean height of a 1D surface above a substrate point x at time t as, e.g. in a thin film flowing down a vertical wall, see e.g. Kalliadasis *et al.* (2011) and Pradas *et al.* (2011a, 2013). We also take $g(x) \in H^q(\mathscr{P}_\alpha)$ for $q \geqslant 4$, a periodic initial condition, i.e.

$$g(x + L) = g(x).$$

The gKS equation is of the type (1.1) with

$$B(u) := B(u,u) := \lambda u u_x \quad \text{and} \quad A := (\kappa \partial_x^2 + \delta \partial_x^3 + \nu \partial_x^4). \tag{1.3}$$

A rigorous dimensional reduction of the gKS equation is of special interest because it is not a Hamiltonian system and does not have an intrinsic invariant measure. This makes direct application of stochastic mode reduction strategies difficult; see Stinis (2004) for instance.

It is noteworthy that the gKS equation retains the fundamental elements of any non-linear process that involves wave evolution: the simplest possible non-linearity uu_x, instability and energy production u_{xx}, stability and energy dissipation u_{xxxx} and dispersion u_{xxx}. We note that the non-linearity arises effectively from the non-linear correction to the phase speed, a non-linear kinematic effect that captures how larger waves move faster than smaller ones. In the context of thin-film flows (Duprat *et al.*, 2009; Tseluiko *et al.*, 2010a,b; Tseluiko & Kalliadasis, 2012), the terms uu_x, u_{xx}, u_{xxx} and u_{xxxx} are due to the interfacial kinematics associated with mean flow, inertia, viscosity and surface tension, respectively, with the corresponding parameters, λ, κ, δ and ν, all positive and measuring the relative importance of these effects. The 'strength of the non-linearity', λ, in particular, is associated with the scaling for the velocity (and hence time). In addition, $\int_{\mathscr{P}_\alpha} u \, \mathrm{d}x = u_0$, a measure of the volume of the liquid, a conservation property for systems whose spatial average does not drift. A simplified form of (1.2) is obtained by appropriately rescaling u, x and t, which is equivalent to setting $\lambda = \kappa = \nu = 1$ and keeping the same notation for dimensionless quantities (Duprat *et al.*, 2009; Tseluiko *et al.*, 2010a,b; Tseluiko & Kalliadasis, 2012).

As with many non-linear time-dependent problems in science and engineering, equations of the form (1.1) are too complex to be fully resolved and the influence of neglected degrees of freedom is not clear a priori. This problem exists independently of spatial dimensions for (1.1) and hence for the gKS equation also. The reliable resolution of high-dimensional problems is a well-known issue in computational science where one can numerically only deal with a finite number of degrees of freedom.

Hence, there is a strong need for (finite-dimensional/) dimensionally reduced formulations, which in turn would allow for studies of long time behaviour of physical systems. Modelling of the ocean-atmosphere, which mainly generates our weather, is one important example: One has a characteristic timescale of several years for the ocean in contrast to several days governing atmospheric structures such as cyclones. As a consequence, a characteristic feature of many physical systems is the presence of fast and slow degrees of freedom. The relevant information of a system's long-time behaviour is often primarily contained in the slow modes. For Hamiltonian systems, such a mode-reduced mathematical

formulation is generally obtained by the Mori–Zwanzig or optimal prediction techniques as described later on. Here we focus on non-linear equations not showing a Hamiltonian-like structure as exemplified by Equation (1.2) and we provide a systematic (e.g. via our RG method and maximum entropy principle) and rigorous (e.g. via error estimates) framework for the reliable derivation of low-dimensional (/slow-mode) representations of such equations.

First, we recall the general, often ad hoc, approximation of decomposing the problem of interest into fast w and slow v modes. For Equation (1.1) such a purely formal splitting, i.e. $u \approx u^\varepsilon = v^\varepsilon + w^\varepsilon$, reads in standard notation applied in the literature as

$$\begin{aligned} \frac{\partial}{\partial t} v^\varepsilon &= f(v^\varepsilon, w^\varepsilon), \\ \frac{\partial}{\partial t} w^\varepsilon &= \frac{1}{\varepsilon} g(v^\varepsilon, w^\varepsilon), \end{aligned} \tag{1.4}$$

where the small parameter $0 < \varepsilon \ll 1$ mediates the timescale separation. Mode reduction strategies, such as 'adiabatic elimination' (VanKampen, 1985), invariant manifolds (Foias *et al.*, 1988a) and optimal prediction (Chorin *et al.*, 1998), are tools to eliminate the fast modes and derive 'appropriate' equations for the slow modes only. We remark that, especially for systems with spatio-temporal chaos (like the gKS equation Chow & Hwa, 1995; Tseluiko *et al.*, 2010a), such a reduction needs to be carefully performed in order to not lose the relevant dynamical characteristics of the full system. Also the study in Marion & Temam (1989) emphasizes the importance of careful finite-dimensional approximations with computational schemes by exploiting the structure of Galerkin methods. The strategy of defining an invariant manifold is almost classical by now. For example, in Foias *et al.* (1988b) the existence of an inertial manifold for the KS equation (obtained from (1.2) with $\delta = 0$) is shown. An inertial manifold is a finite-dimensional, exponentially attracting, positively invariant, Lipschitz manifold. The principle idea is to determine a map $\Phi : \mathbb{V} \to \mathbb{W}$ such that we can rewrite Equation (1.1) in the low-dimensional form:

$$\partial_t v + PB(v + \Phi(v), v + \Phi(v)) + APv = 0, \tag{1.5}$$

where $P : \mathbb{H} \to \mathbb{V}$ and $Q := (I - P) : \mathbb{H} \to \mathbb{W}$ are projections onto the orthogonal subspaces $\mathbb{V}$ and $\mathbb{W}$, respectively, such that $\mathbb{H} = \mathbb{V} \oplus \mathbb{W}$. A strategy to determine Φ in general Galerkin spaces is, for example, suggested in Foias *et al.* (1988a) for the KS equation. The RG approach performed here can also be understood as a formal and feasible procedure to derive an asymptotic invariant manifold; see (1.8) and (1.9). Further analytical results are the characterization of a global attracting set for the Kuramoto–Sivashinsky equation by the so-called background flow method (Collet *et al.*, 1993b; Goodman, 1994) and via a 'capillary Burgers equation' (Otto, 2009), where the latter also forms the best known bound in this context. In Akrivis *et al.* (2013) and Collet *et al.* (1993a), the analyticity of solutions is studied.

Open questions and answers to the classical separation (1.4): (i) *Is the splitting* (1.4) *and the approximation of u by* u^ε *valid and in which sense?* This question is often not answered in the literature where, from the outset, a separation (1.4) is assumed; see Boghosian *et al.* (1999) and Chow & Hwa (1995) for instance. These studies heuristically motivate a timescale mediation (or separation into slow and fast scales) of the form (1.4). The present work aims to provide a rigorous foundation in Theorem 3.1 by the following estimate:

$$\|u - v^\varepsilon\|^2_{L^2(\mathscr{P}_\alpha)}(T) \leqslant C\varepsilon^2 + \exp(CT)(\varepsilon^{1/4} + \varepsilon). \tag{1.6}$$

If we suppose that u and u^ε satisfy a Gevrey regularity characterized by a parameter $\sigma > 0$, then we can improve (1.6) in the following way:

$$\|u - v^\varepsilon\|^2_{L^2(\mathscr{P}_\alpha)}(T) \leqslant C\varepsilon^2 + \exp(CT)\left(\varepsilon^{1/4}\exp\left(-\frac{\sigma}{\varepsilon^{1/4}}\right) + \varepsilon\right). \tag{1.7}$$

It should be pointed out that these estimates also account for the reduction to the slow degrees of freedom v^ε too and not just for the error between u and u^ε.

(ii) *How can we account for the fast degrees of freedom w^ε in an equation for the slow modes v^ε only?* For this purpose, we apply an abstract RG approach extended to general multiscale problems; see Chen *et al.* (1996), Moise &Temam (2000) and Moise & Ziane (2001). The RG method was first introduced in quantum field theory as a tool to perform scale transformations. The method then became popular with Wilson's work on the Kondo problem (Wilson, 1975). It can formally provide the separation (1.4). This means that we first obtain an approximation for v^ε of the form

$$\partial_t v^\varepsilon + A_v v^\varepsilon + P_N B(v^\varepsilon, v^\varepsilon) = -\varepsilon G_\varepsilon(U(t), v^\varepsilon), \tag{1.8}$$

where $\varepsilon G_\varepsilon(U, v^\varepsilon)$ is a perturbation 'force' originating from the renormalization method and $U = V + W$ is a solution of the RG equations

$$\begin{aligned} \partial_t V + A_v V + P_N B(V, V) &= 0, \\ \partial_t W + Q_N B_1(V, W) &= 0, \end{aligned} \tag{1.9}$$

where $V = P_N U$ and $W = (I - P_N)U =: Q_N U$ are projections onto the normalized slow and fast manifolds, respectively. Since we can analytically solve for W, we end up with an equation for the slow variable v^ε only. The above estimates (1.6) and (1.7) then make the reduction (1.8) rigorous. Moreover, Equation $(1.9)_2$ can be interpreted as the map $\Phi^\varepsilon(v^\varepsilon)$ onto the asymptotic invariant manifold.

It should also be pointed out that at this stage the RG approximation (1.8) alone is not satisfactory since the fast variable W contained in U is of infinite dimension and hence cannot entirely be resolved numerically. We give an answer to this problem after the last question (iii) by the principle of maximum entropy. Moreover, question (ii) is of particular relevance here, since the fast modes prevent the existence of a canonical invariant measure. Such a measure makes classical reduction methods such as Mori–Zwanzig and optimal prediction more feasible and does not require the choice of a less physically founded non-invariant measure; see also question (iii) where a different methodology is proposed.

(iii) *What kind of information do we need to carry over from the (infinite-dimensional) fast degrees of freedom to the (finite-dimensional) slow ones and how?* To this end, we derive a stochastic evolution equation for the resolved (slow) variable by properly including necessary information from the unresolved (fast) variable by a maximum information entropy principle introduced in Jaynes (1957a,b) and Rosenkrantz (1989). This principle does not require statistical data to define all Fourier modes. It turns out that the asymptotic behaviour in time of a weighted variance of the fast modes is sufficient. The necessity of such a strong assumption relies on the fact the gKS equation does not have an infinite-dimensional invariant measure and that we only account for spatial randomness. Via this entropy principle (Theorem 4.1), we then conclude that the Fourier modes of the fast variable W in $(1.9)_2$ are Gaussian distributed with zero mean. Hence, we rigorously obtain a noisy gKS equation by applying the random variable $U = v^\varepsilon + W$ in the deterministic equation for the slow variable (1.8). Herewith, our analysis explains how to rigorously add a random force to the gKS equation. Furthermore, our derivation further

shows that the induced noise accounts for the unresolved degrees of freedom and hence becomes less important for an increasing number of grid points in computations.

The approach proposed here provides an alternative to the Mori–Zwanzig formalism (Zwanzig, 1961, 1973; Mori, 1965) which advantageously makes use of a Hamiltonian (Zwanzig, 1961; Mori, 1965) or extended Hamiltonian structure (Zwanzig, 1973). Mori–Zwanzig techniques and related optimal prediction methods (Chorin *et al.*, 1998) generally rely on a canonical probability distribution (invariant measure) which exists naturally for Hamiltonian systems. In principle, one can also apply these techniques to systems that lack an invariant measure. However, the methodology becomes much more involved in such situations and it is not clear how to choose the required non-invariant measure unlike with systems with a canonical invariant measure. The canonical probability density for a Hamiltonian $H(u)$ is $\rho(u) := Z^{-1}\exp(-\beta H(u))$, where β is the inverse temperature and Z is a normalization constant referred to as the partition function. The Mori–Zwanzig formalism then is based on a projection operator P that projects functions in L^2 onto a subspace that only depends on the resolved degrees of freedom. With respect to the canonical density ρ such a projection operator P can be defined by the conditional expectation

$$[Pf](v) := \mathbb{E}[f|v] = \frac{\int f(v,w)\rho(v,w)\,\mathrm{d}w}{\int \rho(v,w)\,\mathrm{d}w}, \tag{1.10}$$

where $f \in L^2$, and v is the resolved and w the unresolved variable. The projection P and Dyson's formula for evolution operators then provide an equation for the resolved modes v only. Moreover, (1.10) is the conditional expectation of f given v and hence is the best least square approximation of f by a function of v. Therefore, the projection P guarantees optimality which is the key idea in the optimal prediction method. However, neither a Hamiltonian structure nor an invariant measure exists for the gKS equation. Therefore, it is not obvious how to derive standard optimality statements relying on a conditional probability argument (Stinis, 2004). In contrast to such a conditional probability approach, we achieve optimality in the sense of maximum information entropy. However, we remark that one can also define other projections than (1.10).

The purpose of the present article is threefold:

1. To reliably perform a (stochastic) mode reduction for the full gKS equation in contrast to Stinis (2004) where a truncated problem is studied. The principal idea is based on an abstract RG approach, as emphasized earlier. We derive error estimates (Theorem 3.1) for this reduction and hence provide rigorous support for the heuristic motivation of a noisy, low-dimensional approximation deducted in Chow & Hwa (1995) by the standard RG method in physics (Wilson, 1975).
2. To rigorously support Stinis' assumption of Gaussian distributed Fourier modes (Stinis, 2004). To this end, we derive a probability distribution (Theorem 4.1) for the fast modes by the principle of maximum information entropy.
3. The findings in (1) and (2) form the bases for a new stochastic mode reduction strategy. We are able to reduce the fast variable by an equation for the slow variable only. The information of the fast modes enters as a random variable W via a force term into the slow mode equations. We are not aware of any previous work that utilizes the RG method in the context of stochastic mode reduction.

We introduce basic notation and well-known results in Section 1.2. A formal derivation of an RG equation for the gKS equation follows in Section 2. In Section 3, we obtain error estimates to rigorously

verify the approximation derived in Section 2. In Section 4, we reduce the fast modes by a mode reduction strategy based on the maximum information entropy principle. Finally, in Section 6 we close with conclusions and perspectives.

1.1 *The gKS equation*

The KS equation is a paradigmatic model for the study of low-dimensional spatio-temporal chaos or weak/dissipative turbulence as defined by Manneville (1990). This type of turbulence is often characterized by formation of clearly identifiable, localized coherent structures in what appears to be a randomly disturbed system, as is e.g. the case with Rayleigh–Bénard convection (Shraiman, 1986). The KS equation was first proposed as a model for pattern formation in reaction–diffusion systems by (Kuramoto & Tsuzuki, 1976). Its derivation is based on a generalized, time-dependent Ginzburg–Landau equation. Sivashinsky (1979) derived the KS equation as an asymptotic approximation of a diffusional–thermal flame model. The equation also describes small-amplitude waves on the surface of a thin film flowing down a planar inclined wall (e.g. Homsy, 1974; Kalliadasis *et al.*, 2011).

With the addition of the dispersive term, u_{xxx}, the KS equation becomes the gKS equation. Like the KS equation, it has been reported for a wide variety of systems, from plasma waves with dispersion due to finite ion banana width (Cohen *et al.*, 1976) to a thin film flowing down a planar wall for near-critical conditions (e.g. Kalliadasis *et al.*, 2011; Saprykin *et al.*, 2005). The studies in Duprat *et al.* (2009), Tseluiko *et al.* (2010a,b) and Tseluiko & Kalliadasis (2012) have developed a coherent-structure theory for the interaction of the solitary-pulse solutions of the gKS equation. In Duprat *et al.* (2009) and Tseluiko *et al.* (2010a), the theory was shown to be in agreement with experiments using a thin film coating a vertical fiber, another hydrodynamic system where the gKS equation can be applicable.

The well-posedness of (1.1) is established, for example, in Tadmor (1986) in the class of generalized Burgers equations which consist of a quadratic non-linearity and arbitrary linear parabolic part. The article (Larkin, 2004) verifies solvability of the gKS equation in bounded domains and studies its limit towards the Korteweg–de Vries equation. In the context of long-time and large-space considerations, there are recent analytical attempts to verify an 'equipartition principle' in the power spectrum of periodic solutions by deriving bounds on their space average of $|u|$ and certain derivatives of it; see Giacomelli & Otto (2005) and Otto (2009). Such a spectral characterization is reminiscent of white noise.

An interesting work that applies the optimal prediction to the KS equation is that of Stinis (2004). Since this approach requires a non-invariant measure, the author constructs a Gibbs measure for the required initial distribution through inference from empirical data (obtained by a computational approach). This allows one then to define the conditional expectation providing optimality by an orthogonal projection of the unresolved modes to the resolved ones. However, this approach already assumes a Gaussian distribution from the outset. For this strategy, one also needs to work with the truncated KS equation. Sufficient numerical data are then required in advance for a reliable construction of an initial distribution.

1.2 *Notation*

Functions $u \in H^s(\mathscr{P}_\alpha)$ for $s \geqslant 1$ can be represented by their Fourier series:

$$u(x) = \sum_{k\in\mathbb{Z}} u_k \exp\left(\mathrm{i}\frac{k}{\alpha}x\right), \quad \bar{u}_k = u_{-k}, \tag{1.11}$$

where H^s denotes here the usual periodic Sobolev space with finite norm,

$$\|f\|_{H^s}^2 := \sum_{k\in\mathbb{Z}} (1+|k|^2)^s |\hat{f}(k)|^2. \tag{1.12}$$

Furthermore, the square root of the latter quantity is a norm on $H^s(\mathscr{P}_\alpha)$ equivalent to the usual one. We denote, for $s \geqslant 0$,

$$\dot{H}^s(\mathscr{P}_\alpha) := \left\{ u \in H^s(\mathscr{P}_\alpha) \,\middle|\, \int_{\mathscr{P}_\alpha} u \,\mathrm{d}x = 0 \right\}. \tag{1.13}$$

The subspace of $\dot{H}^s(\mathscr{P}_\alpha)$ spanned by the set

$$\{\mathrm{e}^{\mathrm{i}(k/\alpha)x} \mid k \in \mathbb{Z}, -N \leqslant k \leqslant N\} \tag{1.14}$$

is denoted by H_N^s. For a given integer N we define the projections $v := P_N u$ and $w := Q_N u := (I - P_N)u$, respectively, by

$$\begin{aligned} v = P_N u &= \sum_{|k|\leqslant N} u_k \exp\left(\mathrm{i}\frac{k}{\alpha}x\right), \\ w = Q_N u &= \sum_{|k|>N} u_k \exp\left(\mathrm{i}\frac{k}{\alpha}x\right). \end{aligned} \tag{1.15}$$

Let us mention that the gKS equation preserves mass as already noted in Section 1, i.e.

$$\int_{\mathscr{P}_\alpha} u \,\mathrm{d}x = u_0, \tag{1.16}$$

where u_0 is the zeroth Fourier mode. We remark that P_N is an orthogonal projection with respect to H_N^s, which means

$$\int_{\mathscr{P}_\alpha} (P_N u - u)\phi \,\mathrm{d}x = 0 \quad \text{for all } \phi \in H_N^s. \tag{1.17}$$

The projection P_N enjoys the following well-known property (Jackson, 1930; Maday & Quarteroni, 1988), i.e. for $k \geqslant s$, $k \geqslant 0$ it holds

$$\|u - P_N u\|_{H^s} \leqslant C N^{s-k} \|u\|_{H^k} \quad \text{for all } u \in H_N^k(\mathscr{P}_\alpha). \tag{1.18}$$

Next, we introduce Gevrey spaces. For $\sigma \geqslant 0$ and $s \geqslant 0$, we say that a function f is in the Gevrey space $G_{\sigma,s}$ if and only if

$$\|f\|_{G_{\sigma,s}}^2 := \sum_{k\in\mathbb{Z}} (1+|k|^2)^s \exp(2\sigma\sqrt{1+|k|^2}) |f_k|^2 < \infty, \tag{1.19}$$

where f_k denote the Fourier coefficients of f. Note that if $\sigma = 0$, then $H^s = G_{0,s}$. Moreover, it can be readily proved (see Kalisch & Raynaud (2007)) that, for $u \in G_{\sigma,s}$, the following inequality holds:

$$\|u - P_N u\|_{H^s} \leqslant N^{s-k} \exp(-\sigma N) \|u\|_{G_{\sigma,k}}. \tag{1.20}$$

2. Formal derivation of a reduced gKS equation

As noted in Section 1, we adapt RG approaches (Moise &Temam, 2000; Moise & Ziane, 2001) to the gKS equation.

2.1 *Projections into fast and slow equations*

We apply the projections P_N and Q_N defined in (1.15) to equation (1.1) and obtain the following coupled system for v and w:

$$\begin{aligned} \partial_t v + P_N B(v+w) + A_v v = 0, \quad &\text{where } A_v = P_N A = A P_N, \\ \partial_t w + Q_N B(v+w) + A_w w = 0, \quad &\text{where } A_w = Q_N A = A Q_N. \end{aligned} \tag{2.1}$$

We define $\varepsilon = 1/N^4$, where N is large enough (see error estimates, i.e. Theorem 3.1) and set

$$\begin{aligned} &\tilde{A}_v = A_v \quad \text{on } P_N \dot{H}^s = H^s_N, \\ &\tilde{A}_w = \varepsilon A_w = \frac{A_w}{N^4} \quad \text{on } Q_N \dot{H}^s = \dot{H}^s \setminus H^s_N. \end{aligned} \tag{2.2}$$

The eigenvectors of $\tilde{A}_v$ are the functions $\exp(\mathrm{i}(k/\alpha)x)$, $k \in \mathbb{Z}$, $|k| \leqslant N$ with eigenvalues

$$\rho^v_k := -\nu \left|\frac{k}{\alpha}\right|^2 - \mathrm{i}\delta \left(\frac{k}{\alpha}\right)^3 + \kappa \left|\frac{k}{\alpha}\right|^4.$$

Correspondingly, the eigenvectors of $\tilde{A}_w$ are the functions $\exp(\mathrm{i}(k/\alpha)x)$, $k \in \mathbb{Z}$, $|k| > N$ with eigenvalues,

$$\rho^w_k := \frac{1}{N^4}\left(-\nu \left|\frac{k}{\alpha}\right|^2 - \mathrm{i}\delta \left(\frac{k}{\alpha}\right)^3 + \kappa \left|\frac{k}{\alpha}\right|^4\right). \tag{2.3}$$

Remark 2.1 The RG method is formally applied here as if these operators $\tilde{A}_v$ and $\tilde{A}_w$ were independent of ε. This technical step of scaling the linear operator and its subsequent treatment is part of the abstract RG approach introduced in Moise &Temam (2000) and Moise & Ziane (2001) in the context of fluid dynamics.

We can now rewrite (2.1) as

$$\begin{aligned} \partial_t v + \tilde{A}_v v + P_N B(v+w) = 0, \\ \partial_t w + \frac{1}{\varepsilon}\tilde{A}_w w + Q_N B(v+w) = 0. \end{aligned} \tag{2.4}$$

For convenience, we additionally define

$$u = \begin{pmatrix} v \\ w \end{pmatrix}, \quad L = \begin{pmatrix} 0 \\ \tilde{A}_w \end{pmatrix}, \quad \mathscr{A} = \begin{pmatrix} \tilde{A}_v \\ 0 \end{pmatrix}, \quad F(u) = \begin{pmatrix} -P_N B(v+w) \\ -Q_N B(v+w) \end{pmatrix}, \tag{2.5}$$

and hence rewrite (2.4) in the following compact way:

$$\partial_t u + \frac{1}{\varepsilon} Lu + \mathscr{A} u = F(u). \tag{2.6}$$

For the subsequent RG analysis, we introduce the fast time scale $s = t/\varepsilon$, and we define $\tilde{u}(s) = u(\varepsilon s)$. We set $\tilde{v}(s) = P\tilde{u}(s)$, $\tilde{w}(s) = Q\tilde{u}(s)$. In these new variables, (2.4) becomes

$$\begin{aligned} \partial_s \tilde{v} + \varepsilon \tilde{A}_v \tilde{v} + \varepsilon P_N B(\tilde{v} + \tilde{w}) = 0, \\ \partial_s \tilde{w} + \tilde{A}_w \tilde{w} + \varepsilon Q_N B(\tilde{v} + \tilde{w}) = 0 \end{aligned} \tag{2.7}$$

or (2.6),

$$\partial_s \tilde{u} + L\tilde{u} + \varepsilon \mathscr{A} \tilde{u} = \varepsilon F(\tilde{u}). \tag{2.8}$$

2.2 *Perturbation expansion: the RG equation*

We now formally apply the RG method and additionally omit the dependence of L and $\mathscr{A}$ on N as in Moise &Temam (2000) and Moise & Ziane (2001). For simplicity, we also assume that either (i) L, ν, κ are not proportional to π or that (ii) $N^2 \geqslant \frac{8}{7}(\alpha^2 \nu/\kappa)$, where N denotes the largest Fourier mode in the Galerkin approximation.

We make the ansatz of a naive perturbation expansion,

$$\tilde{u}^\varepsilon = \tilde{u}^0 + \varepsilon \tilde{u}^1 + \varepsilon^2 \tilde{u}^2 + \cdots \tag{2.9}$$

for $\tilde{u}$ in (2.8). After substituting (2.9) into (2.8), we formally obtain the following sequence of problems:

$$\begin{aligned} \partial_s \tilde{u}^0 + L\tilde{u}^0 = 0, \\ \partial_s \tilde{u}^1 + L\tilde{u}^1 = F(\tilde{u}^0) - \mathscr{A}\tilde{u}^0 \end{aligned} \tag{2.10}$$

and so on.

Formally, the solution of $(2.10)_1$ for the initial condition $\tilde{u}^0(0) = u_0$ is

$$\tilde{u}^0(s) = \exp(-Ls)u_0. \tag{2.11}$$

Equation (2.11) can be equivalently written as

$$\begin{aligned} \tilde{v}^0(s) = v_0, \\ \tilde{w}^0(s) = \exp(-\tilde{A}_w s)w_0. \end{aligned} \tag{2.12}$$

We solve equation $(2.10)_2$ with the variation of constants formula

$$\tilde{u}^1(s) = \exp(-Ls) \int_0^s \exp(L\sigma)[F(\exp(-L\sigma)u_0) - \mathscr{A} \exp(-L\sigma)u_0] \, \mathrm{d}\sigma, \tag{2.13}$$

where $\tilde{u}^1(0) = 0$, since we are interested in approximations up to $O(\varepsilon)$ such that $\tilde{u}^1(0)$ is irrelevant and can be taken to be zero; see Moise & Ziane (2001). We note that $\mathscr{A} \exp(-L\sigma) = \mathscr{A}$ and we decompose

the rest of the integrand in (2.13) as

$$\exp(L\sigma)F(\exp(-L\sigma)u_0) - \mathscr{A}v_0 =: F_{\mathrm{R}}(u_0) + \tilde{F}_{\mathrm{NR}}(\sigma, u_0), \tag{2.14}$$

where $F_{\mathrm{R}}(u_0)$ represents the part independent of σ on the left-hand side of (2.14) and $\tilde{F}_{\mathrm{NR}}$ the rest. Using standard RG terminology, we refer to F_{R} as the 'resonant' and $\tilde{F}_{\mathrm{NR}}$ as the 'non-resonant' term.

Using (2.11), (2.13) and (2.14) in (2.9) provides the following Duhamel's form of the formal perturbation expansion for $\tilde{u} = \tilde{u}^{\varepsilon}$,

$$\tilde{u}^{\varepsilon}(s) = \exp(-Ls)\left(u_0 + \varepsilon s F_{\mathrm{R}}(u_0) + \varepsilon \int_0^s \tilde{F}_{\mathrm{NR}}(\sigma, u_0)\,\mathrm{d}\sigma\right) + \mathscr{O}(\varepsilon^2). \tag{2.15}$$

The key idea is now to remove the secular term $\varepsilon s F_{\mathrm{R}}(u_0)$, which grows in time. To this end, we define the 'renormalized function' $\tilde{U} = \tilde{U}(s)$ as the solution of

$$\begin{aligned} \partial_s \tilde{U} &= \varepsilon F_{\mathrm{R}}(\tilde{U}), \\ \tilde{U}(0) &= u_0. \end{aligned} \tag{2.16}$$

The equation for the slow variable $U(t) = \tilde{U}(t/\varepsilon)$ correspondingly satisfies

$$\begin{aligned} \partial_t U &= F_{\mathrm{R}}(U), \\ U(0) &= u_0. \end{aligned} \tag{2.17}$$

Let us derive the explicit form of the RG equation for our problem. With the expressions for L and F, and the identity $u_0 = v_0 + w_0$, we obtain

$$\exp(L\sigma)F(\exp(-L\sigma)u_0) = \exp(L\sigma)\begin{pmatrix} -P_N B(v_0 + \exp(-L\sigma)w_0) \\ -Q_N B(v_0 + \exp(-L\sigma)w_0) \end{pmatrix}. \tag{2.18}$$

Next we identify the resonant terms, i.e. $F_{\mathrm{R}}(u_0)$. With the Fourier series expansion

$$\phi(x) = \sum_{k\in\mathbb{Z}} \exp\left(\mathrm{i}\frac{k}{\alpha}x\right)\phi_k, \tag{2.19}$$

we have,

$$\begin{aligned} B(\phi,\psi) &= \mathrm{i}\lambda \sum_{k\in\mathbb{Z}} \exp\left(\mathrm{i}\frac{k}{\alpha}x\right)\phi_k \sum_{l\in\mathbb{Z}} \exp\left(\mathrm{i}\frac{l}{\alpha}x\right)\frac{l}{\alpha}\psi_l \\ &= \mathrm{i}\lambda \sum_{j\in\mathbb{Z}} \exp\left(\mathrm{i}\frac{j}{\alpha}x\right) \sum_{k+l=j} \left(\phi_k \frac{l}{\alpha}\right)\psi_l. \end{aligned} \tag{2.20}$$

As a consequence, we end up with the expressions,

$$Q_N B(v_0, \exp(-L\sigma)w_0) = \mathrm{i}\lambda \sum_{|j|>N} \exp\left(\mathrm{i}\frac{j}{\alpha}x\right) \sum_{\substack{k+l=j \\ |k|\leqslant N<|l|}} \left(v_{0k}\frac{l}{\alpha}\right) \exp(-\sigma\rho_l^w)w_{0l},$$

$$\begin{aligned} Q_N B(\exp(-L\sigma)w_0, \exp(-L\sigma)w_0) = \mathrm{i}\lambda &\sum_{|j|>N} \exp\left(\mathrm{i}\frac{j}{\alpha}x\right) \\ &\times \sum_{\substack{k+l=j \\ |k|,|l|>N}} \left(\exp(-\sigma\rho_k^w)w_{0k}\frac{l}{\alpha}\right) \exp(-\sigma\rho_l^w)w_{0l}. \end{aligned} \tag{2.21}$$

The resonant terms in the first sum are the terms for which $\rho_l^w = \rho_j^w$ holds. We note that one also needs to look at the skew-symmetric bilinear form $Q_N B(\exp(-L\sigma)w_0, v_0)$, which leads to the same resonance condition; this means $(-|l/\alpha|^2 - \mathrm{i}\delta(l/\alpha)^3 + \kappa|l/\alpha|^4) = \left(-|j/\alpha|^2 - \mathrm{i}\delta(j/\alpha)^3 + \kappa|j/\alpha|^4\right)$. Since $\nu, \delta, \kappa > 0$, the following set characterizes the resonant indices:

$$R_1(j) := \{(k,l) \mid k = 0,\ j = l,\ |l| > N\}. \tag{2.22}$$

The condition $\rho_k^w + \rho_l^w = \rho_j^w$ characterizes the resonant terms in the second sum of (2.21), i.e. $(k/\alpha)^n + (l/\alpha)^n = (j/\alpha)^n$ for $n = 2, 3, 4$ needs to hold at the same time. Assuming that this condition holds for $n = 2$, we immediately obtain an additional requirement $|j/\alpha|^4 = |k/\alpha|^4 + |l/\alpha|^4 + 2|k/\alpha|^2|l/\alpha|^2$, which holds with respect to the set of resonant indices defined by

$$R_2(j) := \{(k,l) \mid k = 0, l = j, |k|, |l| > N\} \cup \{(k,l) \mid l = 0,\ k = j, |k|, |l| > N\} = \emptyset, \tag{2.23}$$

since $|k|, |l| > N$. For a rigorous and detailed proof we refer the reader to the Appendix. We immediately recognize that (2.23) also justifies our assumption on the case $n = 2$ above.

These considerations determine the resonant part of F by

$$F_{\mathrm{R}}(u_0) = \begin{bmatrix} -P_N B(v_0) - \tilde{A}_v v_0 \\ -Q_N B_1(v_0, w_0) \end{bmatrix}, \tag{2.24}$$

where B_1 is given by its Fourier series expansions for the corresponding index set $R_1(j)$, i.e.

$$Q_N B_1(v_0, w_0) = 2\mathrm{i}\lambda \sum_{|j|>N} \mathrm{e}^{\mathrm{i}(j/\alpha)x} \left(v_{00}\frac{j}{\alpha}\right) w_{0j}. \tag{2.25}$$

Equation (2.14) and the above consideration give the non-resonant term by

$$\tilde{F}_{\mathrm{NR}}(\sigma, u_0) = \begin{bmatrix} -P_N B(v_0 + \mathrm{e}^{-\tilde{A}_w\sigma}w_0) + P_N B(v_0) + \tilde{A}_v v_0 \\ -Q_N \tilde{B}_1(v_0, w_0) - Q_N \tilde{B}_2(w_0) \end{bmatrix}, \tag{2.26}$$

where $Q_N\tilde{B}_1$ and $Q_N\tilde{B}_2$, respectively, are defined by their Fourier series expansions

$$\begin{aligned} Q_N\tilde{B}_1(v_0,w_0) &= \mathrm{i}\lambda \sum_{|j|>N} \mathrm{e}^{\mathrm{i}(j/\alpha)x} \sum_{\substack{k+l=j \\ |l|\neq|j| \\ |k|\leqslant N<|l|}} \left(\left(v_{0k}\frac{j}{\alpha}\right) w_{0l} + \left(w_{0l}\frac{j}{\alpha}\right) v_{0k}\right) \mathrm{e}^{(\rho_j^w-\rho_l^w)\sigma}, \\ Q_N\tilde{B}_2(w_0) &= \mathrm{i}\lambda \sum_{|j|>N} \mathrm{e}^{\mathrm{i}(j/\alpha)x} \sum_{\substack{k+l=j \\ |k/\alpha|^n+|l/\alpha|^n\neq|j/\alpha|^n \text{ for } n=2,3,4 \\ |k|,|l|>N}} \left(w_{0k}\frac{j}{\alpha}\right) w_{0l}\, \mathrm{e}^{(\rho_j^w-\rho_k^w-\rho_l^w)\sigma}. \end{aligned} \tag{2.27}$$

With (2.24) the RG equation for our problem is in the fast time scale,

$$\begin{aligned} \partial_s\tilde{V} + \varepsilon\tilde{A}_v\tilde{V} + \varepsilon P_N B(\tilde{V}) &= 0, \\ \partial_s\tilde{W} + \varepsilon Q_N B_1(\tilde{V},\tilde{W}) &= 0, \end{aligned} \tag{2.28}$$

or after rescaling by $t=\varepsilon s$, and defining $V=P_NU$, $W=Q_NU$,

$$\begin{aligned} \partial_t V + A_v V + P_N B(V) &= 0, \\ \partial_t W + Q_N B_1(V,W) &= 0. \end{aligned} \tag{2.29}$$

REMARK 2.2

(1) The above considerations for the resonant and non-resonant terms can easily be extended to situations where we replace the linear spatial differential operator A with pseudodifferential operators $P(\partial/\partial x)$ with symbol $p(\xi)$ of the form

$$\mathrm{Re}\, p(\mathrm{i}\xi) \geqslant c|\xi|^{\nu}, \quad |\xi|\to\infty, \tag{2.30}$$

where $\nu > \frac{3}{2}$. The requirement (2.30) on the $P(\partial/\partial x)$ guarantees the well-posedness (of such generalized Burgers equations) (Tadmor, 1986). One only needs to adapt the sets for the resonant indices; see (2.22) and (2.23).

(2) Note that the V-equation in the RG equation (2.29) is simply the Galerkin approximation of the gKS equation (1.1).

The special structure of the renormalization equation $(2.29)_2$ for the unresolved (fast) variable allows one to give an explicit expression for its solution. After rewriting $(2.29)_2$ as

$$\partial_t W_j(t) + 2\mathrm{i}\lambda\frac{j}{\alpha}V_0(t)W_j(t) = 0, \tag{2.31}$$

where $V_0(t)=\text{const.}$ due to conservation of mass (1.16), we immediately obtain the solution

$$W_j(t) = c_W^j\, \mathrm{e}^{\mathrm{i}2\lambda(j/\alpha)V_0t}, \quad c_W^j := W_j(0). \tag{2.32}$$

With (2.32) the solution of $(2.29)_2$ becomes

$$W(x,t) = \sum_{|j|>N} c_W^j\, \mathrm{e}^{\mathrm{i}(j/\alpha)(x+2\lambda V_0t)}. \tag{2.33}$$

Equation (2.33) shows that there is no restriction on the definition of the mass V_0. In the context of stochastic mode reduction the situation is different; see Section 4.

2.3 *Construction of approximate/renormalized solutions*

In order to define renormalized solutions, we have to first determine the non-resonant term $\tilde{F}_{\mathrm{NR}}(\sigma, u_0)$ given by (2.26). In fact, we are interested in

$$F_{\mathrm{NR}}(s, U) = \int_0^s \tilde{F}_{\mathrm{NR}}(\sigma, U)\, \mathrm{d}\sigma. \tag{2.34}$$

Let

$$\begin{aligned}
PF_{\mathrm{NR}}(s, U) &= 2\mathrm{i}\lambda \sum_{|j|\leqslant N} \mathrm{e}^{\mathrm{i}(j/\alpha)x} \sum_{\substack{k+l=j \\ |k|\leqslant N<|l|}} \frac{\mathrm{e}^{-\rho_l^w s}}{\rho_l^w} V_k \frac{j}{\alpha} W_l + \mathrm{i}\lambda \sum_{|j|\leqslant N} \mathrm{e}^{\mathrm{i}(j/\alpha)x} \sum_{\substack{k+l=j \\ |k|,|l|>N}} \frac{\mathrm{e}^{-(\rho_k^w+\rho_l^w)s}}{\rho_k^w+\rho_l^w} W_k \frac{j}{\alpha} W_l, \\
QF_{\mathrm{NR}}(s, U) &= -2\mathrm{i}\lambda \sum_{|j|>N} \mathrm{e}^{\mathrm{i}(j/\alpha)x} \sum_{\substack{k+l=j \\ |k|\leqslant N<|l| \\ |l|\neq|j|}} \frac{\mathrm{e}^{(\rho_j^w-\rho_l^w)s}-1}{\rho_j^w-\rho_l^w} V_k \frac{j}{\alpha} W_l \\
&\quad - \mathrm{i}\lambda \sum_{|j|>N} \mathrm{e}^{\mathrm{i}(j/\alpha)x} \sum_{\substack{k+l=j \\ |k/\alpha|^n+|l/\alpha|^n\neq|j/\alpha|^n \text{ for } n=2,3,4 \\ |k|,|l|>N}} \frac{\mathrm{e}^{(\rho_j^w-\rho_k^w-\rho_l^w)s}-1}{\rho_j^w-\rho_k^w-\rho_l^w} W_k \frac{j}{\alpha} W_l.
\end{aligned} \tag{2.35}$$

Now, we are able to define the approximate solution suggested by the RG theory. We obtain

$$u^\varepsilon(t) = \mathrm{e}^{-L(t/\varepsilon)}(U(t) + \varepsilon F_{\mathrm{NR}}(t/\varepsilon, U(t))), \tag{2.36}$$

or with respect to fast $\bar{w}^\varepsilon$ and slow variables $\bar{v}^\varepsilon$,

$$\begin{aligned}
v^\varepsilon &= P_N u^\varepsilon = V(t) + \varepsilon PF_{\mathrm{NR}}(t/\varepsilon, U(t)), \\
w^\varepsilon &= Q_N u^\varepsilon = \mathrm{e}^{-Q_N A t}(W(t) + \varepsilon QF_{\mathrm{NR}}(t/\varepsilon, U(t))).
\end{aligned} \tag{2.37}$$

We note that the initial data are defined by

$$\begin{aligned}
v^\varepsilon(0) &= V(0) + \varepsilon PF_{\mathrm{NR}}(0, U(0)) = v_0 + \varepsilon PF_{\mathrm{NR}}(0, u_0), \\
w^\varepsilon(0) &= W(0) + \varepsilon QF_{\mathrm{NR}}(0, U(0)) = w_0.
\end{aligned} \tag{2.38}$$

3. The renormalized gKS equation and approximation error

After inserting (2.36) into (1.1), we obtain the following perturbed gKS equation:

$$\partial_t u^\varepsilon + Au^\varepsilon + B(u^\varepsilon, u^\varepsilon) = -\varepsilon R_\varepsilon(U(t)), \tag{3.1}$$

where A and B are defined by (1.3) and R_ε is given by

$$\begin{aligned} R_\varepsilon = {} & B(\mathrm{e}^{-L(t/\varepsilon)}U(t), \mathrm{e}^{-L(t/\varepsilon)}F_{\mathrm{NR}}(t/\varepsilon, U(t))) + B(\mathrm{e}^{-L(t/\varepsilon)}F_{\mathrm{NR}}(t/\varepsilon, U(t)), \mathrm{e}^{-L(t/\varepsilon)}U(t)) \\ & + B(\mathrm{e}^{-L(t/\varepsilon)}F_{\mathrm{NR}}(t/\varepsilon, U(t)), \mathrm{e}^{-L(t/\varepsilon)}F_{\mathrm{NR}}(t/\varepsilon, U(t))) - APF_{\mathrm{NR}}(t/\varepsilon, U(t)) \\ & - \mathrm{e}^{-L(t/\varepsilon)}\delta_U F_{\mathrm{NR}}(t/\varepsilon, U(t))\partial_t U. \end{aligned} \tag{3.2}$$

Next, we study estimates on the approximate solutions u^ε of Equation (3.1). In a first step, we need to investigate the non-resonant part F_{NR} of the approximate solutions.

LEMMA 3.1 Let $p \geqslant 2$ and let the initial condition satisfy $g \in H^q(\mathscr{P}_\alpha)$ with $q \geqslant 4$. Assume that the solution of the RG equation (2.29) satisfies $U(t) \in H^p(\mathscr{P}_\alpha)$ for all $t > 0$. For N large enough there exist two uniform constants c_1 and C_2, where C_2 depends on the initial conditions, and c_1 depends only on $\mathscr{P}_\alpha$, but both independent of N, such that the following estimates are true for all $t > 0$:

$$\begin{aligned} \|P_N F_{\mathrm{NR}}(t/\varepsilon, U(t))\|_{H^p} &\leqslant C_2 \mathrm{e}^{-c_1 N^4 t}, \\ \|\mathrm{e}^{-Q_N A t} Q_N F_{\mathrm{NR}}(t/\varepsilon, U(t))\|_{H^p} &\leqslant C_2 \mathrm{e}^{-c_1 N^4 t}. \end{aligned} \tag{3.3}$$

REMARK 3.1 *(Initial conditions)* We note that the regularity assumed above in Lemma 3.1 and in the results below is slightly higher since $\mathrm{g} \in H^p(\mathscr{P}_\alpha)$ would be enough. This regularity assumption enters via an argument based on Gronwall's inequality.

Proof. We see that c_1, and C_2 represent generic constants independent of N (or ε). We first derive estimate $(3.3)_1$. With the expression $(2.35)_1$ we immediately obtain

$$\|P_N F_{\mathrm{NR}}(t/\varepsilon, U(t))\|_{H^p} \leqslant C_2 \mathrm{e}^{-c_1 N^4 t}(\|V \cdot \nabla W\|_{H^p} + \|W \cdot \nabla V\|_{H^p} + \|W \cdot \nabla W\|_{H^p}), \tag{3.4}$$

where we used the fact that we have the following bound:

$$\frac{\mathrm{e}^{-\rho_l^w t/\varepsilon}}{\rho_l^w} = \frac{\mathrm{e}^{-(-\nu|l/\alpha|^2 - \mathrm{i}\delta(l/\alpha)^3 + \kappa|l/\alpha|^4)t}}{1/N^4(-\nu|l/\alpha|^2 - \mathrm{i}\delta(l/\alpha)^3 + \kappa|l/\alpha|^4)} \leqslant C_2 \mathrm{e}^{-c_1 N^4 t}. \tag{3.5}$$

The last inequality follows due to $N < |l|$. The second estimate $(3.3)_2$ can be obtained in the same way by using the inequalities $1 - \mathrm{e}^{-x} \leqslant x$ for all $x \geqslant 0$ and $x\,\mathrm{e}^{-x} \leqslant 1/e$ for all $x \geqslant 0$. We refer the interested reader to Moise &Temam (2000) for a deeper consideration. □

The bounds of Lemma 3.1 allow us to control R_ε in the spirit of Moise &Temam (2000).

LEMMA 3.2 For $N > 0$ and for initial conditions $g \in H^q(\mathscr{P}_\alpha)$ for $q \geqslant 4$, there exist two constants c_1 and C_2 independent of N, such that the following estimate holds true for all $t \geqslant 0$:

$$\|R_\varepsilon(t)\|_{L^2} \leqslant C_2\, \mathrm{e}^{-c_1 N^4 t}. \tag{3.6}$$

Proof. The proof follows in the same way as the proof of Lemma 3.1. We only need to take into account the expression of R_ε and apply Lemma 3.1. □

Subsequently, we write $\|\cdot\|$ and $(\cdot, \cdot)$ for the $L^2(\mathscr{P}_\alpha)$-norm and the $L^2(\mathscr{P}_\alpha)$-scalar product, respectively.

LEMMA 3.3 For $0 < T^* < \infty$ and initial conditions g as in Lemma 3.2, there exists an $0 < \varepsilon^* < \infty$ such that, for $0 \leqslant \varepsilon := 1/N^4 \leqslant \varepsilon^*$, solutions to Equation (3.1) satisfy $u^\varepsilon \in L^\infty(0, T^*; L^2(\mathscr{P}_\alpha)) \cap L^2(0, T^*; H^2(\mathscr{P}_\alpha))$.

Proof. We give here the elements of the proof for the case $\kappa > \nu$ and refer the reader to Larkin (2004) and Tadmor (1986) where stronger regularity (e.g. $u^\varepsilon \in L^\infty(0; T; H^2(P_\alpha))$) and existence results can be found. We formally test equation (3.1) with u^ε and using the periodicity of $\mathscr{P}_\alpha$, i.e. $(\lambda/6)(\partial_x (u^\varepsilon)^3, 1) = 0$, such that

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u^\varepsilon\|^2 + (\kappa - \nu)\|\partial_x^2 u^\varepsilon\|^2 \leqslant \frac{\varepsilon^2}{2}\|u^\varepsilon\|^2 + \frac{1}{2}\|R_\varepsilon(U)\|^2, \tag{3.7}$$

where we used the inequality $\|\nabla u^\varepsilon\|^2 \leqslant \|\Delta u^\varepsilon\|^2$, which holds in the periodic case (see Tadmor, 1986, p. 3). After defining

$$\begin{aligned} \beta &:= 2\varepsilon^2, \\ \gamma &:= \tfrac{1}{2}\|R_\varepsilon(U)\|^2, \end{aligned} \tag{3.8}$$

we multiply (3.7) by $\exp(-\int_0^t \beta \,\mathrm{d}s)$ such that

$$\frac{1}{2}\exp(-\beta t)\frac{\mathrm{d}}{\mathrm{d}t}\|u^\varepsilon\|^2 \leqslant \exp(-\beta t)\frac{\beta}{2}\|u^\varepsilon\|^2 + \exp(-\beta t)\gamma(t). \tag{3.9}$$

Since

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\exp(-\beta t)\frac{1}{2}\|u^\varepsilon\|^2\right) = -\beta\exp(-\beta t)\frac{1}{2}\|u^\varepsilon\|^2 + \exp(-\beta t)\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u^\varepsilon\|, \tag{3.10}$$

we can rewrite (3.9) as

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\exp(-\beta t)\frac{1}{2}\|u^\varepsilon\|^2\right) \leqslant \exp(-\beta t)\gamma(t), \tag{3.11}$$

and subsequent integration together with Lemma 3.2 gives

$$\|u^\varepsilon(T)\|^2 \leqslant C\exp(\beta T)\int_0^T \exp((\beta - C/\varepsilon)t)\,\mathrm{d}t \leqslant \frac{C}{\beta + C/\varepsilon}\exp(\beta T). \tag{3.12}$$

For aribtrary $0 < T^* < \infty$ we can choose $0 \leqslant \varepsilon \leqslant \varepsilon^* := 1/(\exp(\beta T^*) - \beta/CC_\infty)$ where the constant C_∞ is chosen such that

$$\frac{C}{\beta + C/\varepsilon}\exp(\beta T^*) \leqslant C_\infty < C\exp(\beta T^*)/\beta. \tag{3.13}$$

□

The reduced equation for the resolved (slow) modes v^ε alone follows immediately after using $(2.37)_1$ and $(2.29)_1$, i.e. $V(t) = v^\varepsilon(t) - \varepsilon P F_{\mathrm{NR}}(t/\varepsilon, U(t))$,

$$\partial_t v^\varepsilon + A_v v^\varepsilon + P_N B(v^\varepsilon, v^\varepsilon) = -\varepsilon G_\varepsilon(U(t), v^\varepsilon), \tag{3.14}$$

where the induced force term G_ε is defined by

$$\begin{aligned} G_\varepsilon(U(t), v^\varepsilon) := {} & P_N B(\mathrm{e}^{-L(t/\varepsilon)} v^\varepsilon, \mathrm{e}^{-L(t/\varepsilon)} P F_{\mathrm{NR}}(t/\varepsilon, U(t))) + P_N B(\mathrm{e}^{-L(t/\varepsilon)} P F_{\mathrm{NR}}(t/\varepsilon, U(t)), \mathrm{e}^{-L(t/\varepsilon)} v^\varepsilon) \\ & + P_N B(\mathrm{e}^{-L(t/\varepsilon)} P F_{\mathrm{NR}}(t/\varepsilon, U(t)), \mathrm{e}^{-L(t/\varepsilon)} P F_{\mathrm{NR}}(t/\varepsilon, U(t))) \\ & - A_v P F_{\mathrm{NR}}(t/\varepsilon, U(t)) - \mathrm{e}^{-L(t/\varepsilon)} \delta_U P F_{\mathrm{NR}}(t/\varepsilon, U(t)) \partial_t U, \end{aligned} \tag{3.15}$$

and $U(t) = V(t) + W(t) = v^\varepsilon(t) + W(t)$ is the solution of the RG equation (2.29).

LEMMA 3.4 For $0 < T^* < \infty$ and g as in Lemma 3.2, there exists an $0 \leqslant \varepsilon^* < \infty$ such that, for $0 \leqslant \varepsilon = 1/N^4 \leqslant \varepsilon^*$, solutions to (3.14) satisfy $v^\varepsilon := P_N u^\varepsilon \in L^\infty(0, T^*; L^2(\mathscr{P}_\alpha)) \cap L^2(0, T^*; H^2(\mathscr{P}_\alpha))$.

Proof. The proof is similar to the proof of Lemma 3.3. □

The same arguments as those for Lemma 3.2 lead to the following lemma.

LEMMA 3.5 For $N > 0$ and $v_0 = P_N g \in H^q(\mathscr{P}_\alpha)$ for $q \geqslant 4$, there exist two constants c_1 and C_2 independent of N, such that the following estimate holds true for all $t \geqslant 0$:

$$\|G_\varepsilon(t)\|_{L^2} \leqslant C_2 \, \mathrm{e}^{-c_1 N^4 t}. \tag{3.16}$$

The following theorem gives qualitative information about the RG approach by quantifying the error between (3.14) and (1.1).

THEOREM 3.1 Let $g \in H^4(\mathscr{P}_\alpha)$, $\varepsilon = 1/N^4$, and suppose that $u, u^\varepsilon \in L^\infty(0, T; H^2(\mathscr{P}_\alpha))$. Then the difference between the reduced solution v^ε and the exact solution of the gKS equation (1.1) satisfies the following error estimate:

$$\|u(T) - v^\varepsilon(T)\|^2_{L^2(\mathscr{P}_\alpha)} \leqslant C\varepsilon^2 + \exp(CT)(\varepsilon^{1/4} + \varepsilon). \tag{3.17}$$

If we suppose that $u, u^\varepsilon \in L^\infty(0, T; G_{\sigma,2}(\mathscr{P}_\alpha))$, then we can improve (3.17) in the following way:

$$\|u(T) - v^\varepsilon(T)\|^2_{L^2(\mathscr{P}_\alpha)} \leqslant C\varepsilon^2 + \exp(CT)\left(\varepsilon^{1/4} \exp\left(-\frac{\sigma}{\varepsilon^{1/4}}\right) + \varepsilon\right) \tag{3.18}$$

REMARK 3.2

1. The exponential growth in time is not surprising; see, for example, estimate (3.12) in the proof of Lemma 3.3. This estimate motivates the definition of a new variable $h(x, t; \eta) := \exp(-\eta t) u(x, t)$. Tadmor (1986) verifies global existence for such a decayed variable h and a conservative form of the KS equation.
2. The assumption $u, u^\varepsilon \in L^\infty(0, T; H^2(\mathscr{P}_\alpha))$ is a direct consequence of a priori estimates, which are derived by analogous steps as in the proof of Lemma 3.3, and of imposing initial conditions $u_0, u_0^\varepsilon \in H^2(\mathscr{P}_\alpha)$.

Proof. The error $\|u - v^\varepsilon\|_{L^2(\mathscr{P}_\alpha)}$ can be bounded using the triangle inequality by

$$\|u - v^\varepsilon\|_{L^2(\mathscr{P}_\alpha)} \leqslant \|u - u^\varepsilon\|_{L^2(\mathscr{P}_\alpha)} + \|u^\varepsilon - v^\varepsilon\|_{L^2(\mathscr{P}_\alpha)}, \tag{3.19}$$

where the first term on the right-hand side in (3.19) represents the approximation error from the RG method (*RG error*) and the second term accounts for the truncation error (*Tr error*). For notational brevity, we introduce the error variables

$$E^\varepsilon_{\mathrm{RG}} := u - u^\varepsilon \quad \text{and} \quad E^\varepsilon_{\mathrm{Tr}} := u^\varepsilon - v^\varepsilon. \tag{3.20}$$

Step 1: (*RG error*) The equation for the error variable $e^\varepsilon_{\mathrm{RG}}$ reads

$$\partial_t E^\varepsilon_{\mathrm{RG}} + [\kappa\partial_x^2 + \delta\partial_x^3 + \nu\partial_x^4]E^\varepsilon_{\mathrm{RG}} + E^\varepsilon_{\mathrm{RG}}\partial_x u + u^\varepsilon\partial_x E^\varepsilon_{\mathrm{RG}} = \varepsilon R_\varepsilon(U). \tag{3.21}$$

First, we test (3.21) with $-\partial_x^2 E^\varepsilon_{\mathrm{RG}}$, i.e.,

$$\begin{aligned}&\partial_t(\partial_x E^\varepsilon_{\mathrm{RG}}, \partial_x E^\varepsilon_{\mathrm{RG}}) - (\kappa\partial_x^2 E^\varepsilon_{\mathrm{RG}}, \partial_x^2 E^\varepsilon_{\mathrm{RG}}) - (\delta\partial_x^3 E^\varepsilon_{\mathrm{RG}}, \partial_x^2 E^\varepsilon_{\mathrm{RG}}) - (\nu\partial_x^4 E^\varepsilon_{\mathrm{RG}}, \partial_x^2 E^\varepsilon_{\mathrm{RG}})\\ &\quad - (E^\varepsilon_{\mathrm{RG}}\partial_x u, \partial_x^2 E^\varepsilon_{\mathrm{RG}}) - (u^\varepsilon\partial_x E^\varepsilon_{\mathrm{RG}}, \partial_x^2 E^\varepsilon_{\mathrm{RG}}) = -(\varepsilon R_\varepsilon(U), \partial_x^2 E^\varepsilon_{\mathrm{RG}}).\end{aligned} \tag{3.22}$$

Then, we use the test function $E^\varepsilon_{\mathrm{RG}}$,

$$\begin{aligned}&\partial_t(E^\varepsilon_{\mathrm{RG}}, E^\varepsilon_{\mathrm{RG}}) + (\kappa\partial_x^2 E^\varepsilon_{\mathrm{RG}}, E^\varepsilon_{\mathrm{RG}}) + (\delta\partial_x^3 E^\varepsilon_{\mathrm{RG}}, E^\varepsilon_{\mathrm{RG}}) + (\nu\partial_x^4 E^\varepsilon_{\mathrm{RG}}, E^\varepsilon_{\mathrm{RG}})\\ &\quad + (E^\varepsilon_{\mathrm{RG}}\partial_x u, E^\varepsilon_{\mathrm{RG}}) + (u^\varepsilon\partial_x E^\varepsilon_{\mathrm{RG}}, E^\varepsilon_{\mathrm{RG}}) = +(\varepsilon R_\varepsilon(U), E^\varepsilon_{\mathrm{RG}}).\end{aligned} \tag{3.23}$$

Next, we add up (3.22) and (3.23) and apply the Sobolev embedding theorem and standard inequalities to end up with

$$\begin{aligned}&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}[\|E^\varepsilon_{\mathrm{RG}}\|^2 + \|\partial_x E^\varepsilon_{\mathrm{RG}}\|^2] + (\nu - 3\alpha)[\|\partial_x^2 E^\varepsilon_{\mathrm{RG}}\|^2 + \|\partial_x^3 E^\varepsilon_{\mathrm{RG}}\|^2]\\ &\quad \leqslant C_{\mathrm{RG}}(\kappa, \alpha, \varepsilon, \|u\|_{H^1}, \|\partial_x u\|_{H^1}, \|u^\varepsilon\|_{H^1}, \|\partial_x u^\varepsilon\|_{H^1})\|E^\varepsilon_{\mathrm{RG}}\|^2_{H^1}.\end{aligned} \tag{3.24}$$

After defining

$$\tilde{C}_{\mathrm{RG}} = 2C_{\mathrm{RG}}, \tag{3.25}$$

we can multiply (3.24) by $\exp(-\int_0^t \tilde{C}_{\mathrm{RG}}\,\mathrm{d}s)$ such that

$$\frac{1}{2}\exp\left(-\int_0^t \tilde{C}_{\mathrm{RG}}\,\mathrm{d}s\right)\frac{\mathrm{d}}{\mathrm{d}t}\|E^\varepsilon_{\mathrm{RG}}\|^2 \leqslant \exp\left(-\int_0^t \tilde{C}_{\mathrm{RG}}\,\mathrm{d}s\right)\frac{\tilde{C}_{\mathrm{RG}}}{2}\|E^\varepsilon_{\mathrm{RG}}\|^2. \tag{3.26}$$

Applying a corresponding identity based on the product rule as (3.10) in the proof of Lemma 3.3, we can simplify (3.26) to

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\exp\left(-\int_0^t \tilde{C}_{\mathrm{RG}}\,\mathrm{d}s\right)\frac{1}{2}\|E^\varepsilon_{\mathrm{RG}}\|^2\right) \leqslant 0, \tag{3.27}$$

which further reduces by assumptions of Theorem 3.1 and after integration to

$$\frac{1}{2}\|E^\varepsilon_{\mathrm{RG}}(T)\|^2 \leqslant \frac{1}{2}\exp\left(-\int_0^T \tilde{C}_{\mathrm{RG}}\,\mathrm{d}s\right)\|E^\varepsilon_{\mathrm{RG}}(0)\|^2 \leqslant C\|E^\varepsilon_{\mathrm{RG}}(0)\|^2. \tag{3.28}$$

In order to get a bound controlled by ε on the right-hand side of (3.28), we have to take into account the definition of the initial data (2.38), i.e.

$$\|E^{\varepsilon}_{\mathrm{RG}}(0)\|^2 = \varepsilon^2 \|PF_{\mathrm{NR}}(0,g)\|^2 \leqslant C\varepsilon^2, \tag{3.29}$$

since $g \in H^2(\mathscr{P}_\alpha)$ and hence $PF_{\mathrm{NR}}(0,g) \in H^1(\mathscr{P}_\alpha)$ and its norm is bounded independently of $\varepsilon = 1/N^4$ in $H^1(\mathscr{P}_\alpha)$. Hence, we can conclude that

$$\|E^{\varepsilon}_{\mathrm{RG}}(t)\|^2_{L^2(\mathscr{P}_\alpha)} \leqslant C\varepsilon^2, \tag{3.30}$$

which holds uniformly in time.

Step 2: (*Tr error*) We derive an estimate for the error variable $E^{\varepsilon}_{\mathrm{Tr}} := u^\varepsilon - v^\varepsilon = u^\varepsilon - P_N u^\varepsilon$. From (3.1) and (3.14), the error $E^{\varepsilon}_{\mathrm{Tr}}$ satisfies the equation

$$\partial_t E^{\varepsilon}_{\mathrm{Tr}} + [\kappa\partial_x^2 + \delta\partial_x^3 + \nu\partial_x^4]E^{\varepsilon}_{\mathrm{Tr}} + P_N[u^\varepsilon u^\varepsilon_x] - v^\varepsilon v^\varepsilon_x = \varepsilon(P_N R_\varepsilon(U) - G_\varepsilon(U,v^\varepsilon), v^\varepsilon), \tag{3.31}$$

which can be rewritten, for all $\phi \in H^2_N$, as

$$\begin{aligned}&\partial_t(E^{\varepsilon}_{\mathrm{Tr}},\phi) + ([\kappa\partial_x^2 + \delta\partial_x^3 + \nu\partial_x^4]E^{\varepsilon}_{\mathrm{Tr}},\phi)\\ &\quad + (\{P_N[u^\varepsilon u^\varepsilon_x] - v^\varepsilon v^\varepsilon_x\}, E^{\varepsilon}_{\mathrm{Tr}}) = \varepsilon(\{P_N R_\varepsilon(U) - G_\varepsilon(U,v^\varepsilon)\},\phi).\end{aligned} \tag{3.32}$$

Choosing $\phi = E^{\varepsilon}_{\mathrm{Tr}}$ allows one to estimate (3.32) in the following way:

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|E^{\varepsilon}_{\mathrm{Tr}}\|^2 + \nu\|\partial_x^2 E^{\varepsilon}_{\mathrm{Tr}}\|^2 \leqslant C(\alpha,\kappa)\|E^{\varepsilon}_{\mathrm{Tr}}\|^2 + \alpha\|\partial_x E^{\varepsilon}_{\mathrm{Tr}}\|^2 + (\mathrm{I}) + (\mathrm{II}), \tag{3.33}$$

where we define

$$\begin{aligned}(\mathrm{I}) &:= (P_N[u^\varepsilon u^\varepsilon_x] - vv_x, E^{\varepsilon}_{\mathrm{Tr}}),\\ (\mathrm{II}) &:= \varepsilon(\|P_N R_\varepsilon(U)\| + \|G_\varepsilon(U,v^\varepsilon)\|)\|E^{\varepsilon}_{\mathrm{Tr}}\|.\end{aligned} \tag{3.34}$$

Let us first control term (I), which means

$$\begin{aligned}|(\mathrm{I})| &\leqslant |(-P_N u^\varepsilon (P_N u^\varepsilon)_x + u^\varepsilon u^\varepsilon_x, E^{\varepsilon}_{\mathrm{Tr}})| + |(P_N u^\varepsilon (P_N u^\varepsilon)_x - v^\varepsilon v^\varepsilon_x, E^{\varepsilon}_{\mathrm{Tr}})|\\ &\leqslant \tfrac{1}{2}|(\partial_x((u^\varepsilon - P_N u^\varepsilon)(u^\varepsilon + P_N u^\varepsilon)), E^{\varepsilon}_{\mathrm{Tr}})| + 1/4|(\partial_x(P_N u^\varepsilon + v^\varepsilon), E^{\varepsilon}_{\mathrm{Tr}})|\\ &\leqslant C(\|u^\varepsilon\|_{H^1} + \|P_N u^\varepsilon\|_{H^1})\|u^\varepsilon - P_N u^\varepsilon\|_{H^1}\|E^{\varepsilon}_{\mathrm{Tr}}\| + C\|E^{\varepsilon}_{\mathrm{Tr}}\|^2,\end{aligned} \tag{3.35}$$

where we used the embedding $H^1_N(\mathscr{P}_\alpha)$ into $L^\infty(\mathscr{P}_\alpha)$, i.e. $\|\partial_x(P_N u^\varepsilon + v^\varepsilon)\|_{L^\infty} \leqslant C$. The second term (II) immediately becomes

$$|(\mathrm{II})| \leqslant \varepsilon \exp(-ct/\varepsilon) C\|E^{\varepsilon}_{\mathrm{Tr}}\|. \tag{3.36}$$

Define

$$\begin{aligned}\gamma(t) &:= C(\|u^\varepsilon\|, \|v^\varepsilon\|, \|u^\varepsilon\|_{H^1}, \|P_N u^\varepsilon\|_{H^1}, \|u^\varepsilon\|_{H^2})\varepsilon^{1/4} + C(\|u^\varepsilon\|, \|v^\varepsilon\|)\varepsilon\exp(-Ct/\varepsilon),\\ \beta &:= 2(C(\alpha,\kappa) + C)\end{aligned} \tag{3.37}$$

and mulitply (3.33) with $\exp(-\int_0^t \beta\,\mathrm{d}s)$ such that

$$\frac{1}{2}\exp(-\beta t)\frac{\mathrm{d}}{\mathrm{d}t}\|E^{\varepsilon}_{\mathrm{Tr}}\|^2 \leqslant \frac{\beta}{2}\exp(-\beta t)\|E^{\varepsilon}_{\mathrm{Tr}}\|^2 + \gamma(t)\exp(-\beta t). \tag{3.38}$$

Using again a corresponding identity to (3.10) in Lemma 3.3, we can rewrite (3.38) as

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\exp(-\beta t)\frac{1}{2}\|E^{\varepsilon}_{\mathrm{Tr}}\|^2\right) \leqslant \gamma(t)\exp(-\beta t), \tag{3.39}$$

which becomes after integration with respect to time

$$\begin{aligned}\frac{1}{2}\exp(-\beta T)\|E^{\varepsilon}_{\mathrm{Tr}}\|^2(T) &\leqslant C\int_0^T \{(\varepsilon^{1/4} + \varepsilon\exp(-Ct/\varepsilon))\exp(-\beta t)\}\,\mathrm{d}t \\ &\leqslant \frac{C\varepsilon^{1/4}}{\beta}(1-\exp(\beta T)) + \frac{C\varepsilon}{\beta + C/\varepsilon}(1-\exp(-(\beta + C/\varepsilon)T)).\end{aligned} \tag{3.40}$$

In the remaining part, we need to improve (3.40) with the help of Gevrey spaces. To this end, we remark that the factor $\varepsilon^{1/4}$ in (3.37) relies on the interpolation estimate (1.18). If we assume that solutions u^{ε} are in $G_{\sigma,s}$, then we improve (1.18) by (1.20). As a consequence, we are able to rewrite inequality (3.40) by

$$\begin{aligned}\frac{1}{2}\|E^{\varepsilon}_{\mathrm{Tr}}\|^2(T) &\leqslant C\int_0^T \left(\varepsilon^{1/4}\exp\left(\beta(T-t) - \frac{\sigma}{\varepsilon^{1/4}}\right) + \varepsilon\exp(-Ct/\varepsilon + \beta(T-t))\right)\mathrm{d}t \\ &\leqslant \frac{C\varepsilon^{1/4}}{\beta}\exp(\beta T) + \frac{C\varepsilon}{\beta + c/\varepsilon}\exp(\beta T).\end{aligned} \tag{3.41}$$

□

4. Stochastic mode reduction

In this section, the renormalized equations (2.29), (2.31) from Sections 2.1 to 3 allow for a rigorous stochastic mode reduction similar in spirit to the Mori–Zwanzig one (Zwanzig, 1961; Mori, 1965) but for systems not satisfying an extended or generalized Hamiltonian structure (Zwanzig, 1973) and without a canonical invariant measure.

As in the Mori–Zwanzig formalism, we assign a stochastic process to the unresolved modes. This is done by applying the Jaynes maximum entropy principle; see Jaynes (1957a,b) and Rosenkrantz (1989). This seems a reasonable approach for our problem since we do not have a canonically induced probability density. Hence maximizing the information entropy for the probability density of Fourier modes is equivalent to maximizing the multiplicity of Fourier modes. Multiplicity means the number of different ways a certain state in a system can be achieved. States in a system with the highest multiplicity can be realized by nature in the largest number of ways. Hence, the probability density functions with maximum entropy are optimal statistical descriptions.

It should also be noted that a system at equilibrium will most probably be found in the state of highest multiplicity since fluctuations from that state will be usually too small to measure. The probability distribution may also be obtained from experiments as statistical data. In Stinis (2004), the probability distribution is constructed by a conditional expectation obtained from previously computed samples which are used to fit an a priori assumed Gaussian distribution.

Finally, we emphasize that the maximum entropy principle can also be applied to problems where one lacks deterministic data as a consequence of not enough experimental data to fix all degrees of freedom. A common approach to model such uncertainty is to use white noise. The maximum entropy method turns out to be an attractive alternative because it allows one to systematically add noise to the gKS equation over the Equation (3.14), which is obtained by the evolutionary RG method.

However, since we apply the entropy maximization principle (Jaynes, 1957b) on an approximate equation, we already neglect information from the beginning and hence have to account for this by an asymptotic in time characterization of the fast modes for example; see Assumption (A). This assumption might be improved or adapted appropriately in other applications. Subsequently, $(\Omega, \mathcal{F}, \mathbb{P})$ denotes the usual probability space with sample space Ω, σ-algebra $\mathcal{F}$ and probability measure $\mathbb{P}$.

4.1 *Problem induced probability density by maximizing information entropy*

With the considerations at the beginning of Section 4, we assign a probability distribution to the unresolved degrees of freedom W based on the following:

Assumptions:

(A) For a probability measure $\mathbb{P}_j$ with density f_j and

$$\mathscr{C}_N(\tilde{w}_j^{\varepsilon,0}) := \tfrac{1}{2}(\tilde{w}_j^{\varepsilon,0})^2, \tag{4.1}$$

where $\tilde{w}_j^{\varepsilon,0}(t) = w_j^{\varepsilon,0}(t,\omega)$ denotes a realization for $\omega \in \Omega$ of the jth Fourier mode of the leading order term $w^{\varepsilon,0}$ of w^ε in $(2.37)_2$, i.e. $w^{\varepsilon,0}(x,t) = \mathrm{e}^{-Q_N t/\varepsilon} W(x,t)$, we assume that it holds asymptotically in time that

$$\mathbb{E}_j\left[\frac{\partial}{\partial t}\mathscr{C}_N(w_j^{\varepsilon,0})\right] = \int_{-\infty}^{\infty} f_j(\tilde{w}_j^{\varepsilon,0})\frac{\partial}{\partial t}\mathscr{C}_N(\tilde{w}_j^{\varepsilon,0})\,\mathrm{d}\tilde{w}_j^{\varepsilon,0} = \delta_j(t) := -\rho_j^w \mathrm{e}^{-2\rho_j^w t} W_j^2(0), \tag{4.2}$$

i.e. there is a $t_0 \geqslant 0$ such that (4.2) holds for $t > t_0$. We call $\delta_j(t)$ a dissipation rate and $\mathbb{E}_j$ denotes the expectation with respect to the probability $\mathbb{P}_j$.

(B) Under **(A)** the probability $\mathbb{P}_j[W_j \leqslant \tilde{w}_j^{\varepsilon,0}] = F(\tilde{w}_j^{\varepsilon,0})$ with density f_j, i.e.

$$F(\tilde{w}_j^{\varepsilon,0}) := \int_{-\infty}^{\tilde{w}_j^{\varepsilon,0}} f_j(r)\,\mathrm{d}r, \tag{4.3}$$

has maximum *information entropy* $\mathbb{S}_I(f_j)$,

$$\mathbb{S}_I(f_j) = -\int_{-\infty}^{\infty} f_j(r)\log\left(\frac{f_j(r)}{\nu(r)}\right)\,\mathrm{d}r, \tag{4.4}$$

where $f_j(r)$ denotes the probability density of the jth Fourier mode of the unresolved variable W and ν is an according *invariant measure* which is defined on background information intrinsically given by the physical origin of W.

REMARK 4.1

(1) The idea of deriving probability distributions for multiscale evolution problems by maximizing the information entropy seems to go back to Lorenz (1996). The energy argument in Lorenz (1996), which assumes that the fast modes reached already the stationary state, does not provide here enough information to fix the Lagrange multiplier λ_1 associated with this energy constraint. We impose Assumption **(A)** instead. Note that we take slightly more information into account by using $w^{\varepsilon,0}$ instead of W, which does not decay as fast as $w^{\varepsilon,0}$.

(2) A mechanical system governed by the Hamiltonian $H(q,p)$ canonically induces an invariant measure by the density distribution function $f(q,p) := (1/Z(\beta))\,\mathrm{e}^{-\beta H(q,p)}$.

(3) Equation (4.2) accounts for the fact that we do not have an invariant measure to the fast modes. For simplicity, we also neglect a possible randomness in time. This is a further reason for the assumption in (4.2).

In information theory, an entropy related to (4.4) was originally introduced by Shannon (1948) to measure the maximum information content in a message. The Assumptions **(A)** and **(B)** above account for the lack of a free energy and a Hamiltonian for which the thermodynamic equilibrium (invariant measure) can be achieved via the gradient flow with respect to the Wasserstein distance (Jordan *et al.*, 1998). In fact, it should be noted that minimizing the free energy with respect to constant internal energy is equivalent to maximizing the entropy.

THEOREM 4.1 Under Assumptions **(A)** and **(B)**, it follows that the unresolved modes W_k for $|k| > N$ obtained by Equation (2.31) are normally distributed with zero mean, i.e. $\mu_k = 0$, and variance $\sigma_k^2 = 1/2\lambda_k\rho_k$, where $\lambda_k := 1/2\delta_j(t)$ is a Langrange multiplier.

REMARK 4.2 Instead of **(A)**, one can make the following assumption **(A*)**: For large enough times $t > 0$, it holds that

$$\mathbb{E}_j\left[\frac{\partial}{\partial t}\mathscr{C}_N(w_j^{\varepsilon,0})\right] = \sigma^2. \tag{4.5}$$

This immediately leads to the result that the fast modes satisfy $W_k \sim \mathscr{N}(0,\sigma^2)$ where the variance can be defined by the power spectral density as in the case of complete uncertainty; see also Section 5.

To keep the considerations simple, we only account for a spatial random process and keep the time deterministic in Theorem 4.1 (and Assumption **(A)**).

Proof. To maximize (4.4) under Assumptions **(A)** and **(B)**, we apply the following constraints:

$$\textbf{(CI)} \quad \begin{cases} \displaystyle\int_{-\infty}^{\infty} f_k(r)\,\mathrm{d}r = 1, \\ \mathbb{E}_k[\partial_t\mathscr{C}_N(w_k^{\varepsilon,0})] := \displaystyle\int f_k(\tilde{w}_k^{\varepsilon,0})\partial_t\mathscr{C}_N(\tilde{w}_k^{\varepsilon,0})\,\mathrm{d}\tilde{w}_k^{\varepsilon,0} = \delta_k(t), \end{cases} \tag{4.6}$$

where $(4.6)_2$ is a consequence of assumption (i).

Hence, maximizing the entropy $\mathbb{S}_I$ subject to the constraints (4.6) leads to

$$\int_{-\infty}^{\infty} \delta f_k(\tilde{w}_k^{\varepsilon,0})\left\{\log\left(\frac{f_k(\tilde{w}_k^{\varepsilon,0})}{\nu(\tilde{w}_k^{\varepsilon,0})}\right) + \nu(\tilde{w}_k^{\varepsilon,0}) + \lambda_0 + \lambda_k\partial_t\mathscr{C}_N(\tilde{w}_k^{\varepsilon,0})\right\}\mathrm{d}\tilde{w}_k^{\varepsilon,0} = 0, \tag{4.7}$$

where λ_0 and λ_k are Lagrange multipliers associated with the constraints (4.6). In order to give $(4.6)_2$ a precise meaning, we write down the explicit form of the equation belonging to each Fourier coefficient of the fast mode variable $w = Q_N u$ solving (2.29).

We briefly show what the constraint $(4.6)_2$ means:

$$\begin{aligned}\mathbb{E}_k[\partial_t \mathscr{C}_N(w_k^{\varepsilon,0})] &= -\int_{-\infty}^{\infty} f_k(\tilde{w}_k^{\varepsilon,0})(\tilde{w}_k^{\varepsilon,0})^2 \left(2\mathrm{i}\lambda \frac{k}{\alpha} V_0 - \frac{1}{\varepsilon}\rho_k^w\right) \mathrm{d}\tilde{w}_k^{\varepsilon,0} \\ &= \delta_k(t). \end{aligned} \tag{4.8}$$

We recall that $V_0 = \text{const}$. Due to conservation of mass. Since (4.7) should hold for arbitrary variations δf_k, we obtain the following expression for the probability density function:

$$f_k(\tilde{w}_k^{\varepsilon,0}) = \frac{1}{Z_k} \nu(\tilde{w}_k^{\varepsilon,0})\, \mathrm{e}^{-\lambda_k \partial_t \mathscr{C}_N(\tilde{w}_k^{\varepsilon,0})}, \tag{4.9}$$

where $Z_k := \mathrm{e}^{\nu(\tilde{w}_k^{\varepsilon,0})+\lambda_0}$ is called the 'partition function', which is determined by the normalization constraint $(4.6)_2$, i.e.

$$Z_k := Z_k(\lambda_k) = \int_{-\infty}^{\infty} \nu_k(\tilde{w}_k^{\varepsilon,0}) \mathrm{e}^{-\lambda_k \partial_t \mathscr{C}_N(\tilde{w}_k^{\varepsilon,0})}\, \mathrm{d}\tilde{w}_k^{\varepsilon,0}. \tag{4.10}$$

Since the constraint $(4.6)_2$ is quadratic in its nature, we represent it by

$$\lambda_k \partial_t \mathscr{C}_N(\tilde{w}_k^{\varepsilon,0}) = -\lambda_k (\tilde{w}_k^{\varepsilon,0})^2 \tilde{\rho}_k^w = -\frac{1}{2\sigma_k^2}((\tilde{w}_k^{\varepsilon,0} - \mu_k)^2 - \mu_k^2), \tag{4.11}$$

where $\tilde{\rho}_k^w := 2\mathrm{i}\lambda(k/\alpha)V_0 - \rho_k^w/\varepsilon$ and

$$\sigma_k^2 = \frac{1}{2\lambda_k \tilde{\rho}_k^w}, \quad \text{and} \quad \mu_k = 0. \tag{4.12}$$

With identities (4.11) and (4.12) the probability density function (4.9) can be written as

$$f_k(\tilde{w}_k^{\varepsilon,0}) = \frac{1}{Z_k} c_{W_k}^{-1} \sigma_k \sqrt{2\pi}\, \mathscr{N}(\mu_k, \sigma_k, \tilde{w}_k^{\varepsilon,0}) \tag{4.13}$$

for $|k| > N$, where $\mathscr{N}$ is the normal distribution given by

$$\mathscr{N}(\mu_k, \sigma_k, \tilde{w}_k^{\varepsilon,0}) = \frac{1}{\sigma_k\sqrt{2\pi}} \mathrm{e}^{-((\tilde{w}_k^{\varepsilon,0} - \mu_k)^2)/2\sigma_k^2}, \tag{4.14}$$

which is characterized by the following moments:

$$\int_{-\infty}^{\infty} \mathscr{N}(\mu,\sigma,w)\, \mathrm{d}w = 1, \quad \int_{-\infty}^{\infty} \mathscr{N}(\mu,\sigma,w) w\, \mathrm{d}w = \mu \quad \text{and} \quad \int_{-\infty}^{\infty} \mathscr{N}(\mu,\sigma,w) w^2\, \mathrm{d}w = \sigma^2 + \mu^2. \tag{4.15}$$

The first property in (4.15) together with the normalization condition $(4.6)_1$ allow us to define the partition function Z by

$$Z_k = c_{W_k}^{-1}\sigma_k\sqrt{2\pi} \tag{4.16}$$

for $|k| > N$.

REMARK 4.3 The measure $\nu(\mathbf{w}) := \prod_{|k|>N}(1/c_{W_k})$ is the probability density function if we only have a priori information. Usually, it is a non-trivial task and basic considerations of symmetries are required to find this measure ν.

The probability density function f_k admits then the simple form as a product of Gaussian distributions, i.e.

$$f_k(\tilde{w}_k^{\varepsilon,0}) = \mathcal{N}(\mu_k, \sigma_k, \tilde{w}_k^{\varepsilon,0}) \tag{4.17}$$

for $|k| > N$. With the second and third property in (4.15) and the constraint $(4.6)_2$, i.e. (4.8), we obtain, for all $|k| > N$,

$$\delta_k(t) = \mathbb{E}_k[\partial_t \mathscr{C}_N(\omega_k)] = -(\tilde{\rho}_k^w\{\sigma_k^2 + \mu_k^2\}). \tag{4.18}$$

We conclude with (4.18) that the Lagrange parameter λ_k is

$$\lambda_k(t) = \mathrm{e}^{2\rho_k^w t}/(2\rho_k^w W_0^2(0)). \tag{4.19}$$

The important information contained in formula (4.19) and $(4.12)_1$ is that we do not have to assert to each Fourier mode k its standard deviation $k^2\sigma_W^2$. We only need to determine once the Lagrange parameter λ_1 via (4.19). From (4.18) and the property $\bar{\mu_k} = \mu_{-k}$, we obtain that the mean satisfies $\mu_k = 0$ for all $|k| > N$. □

Hence, the approach of maximizing the generalized information entropy allows one to systematically determine the probability distribution function $f_k(w_k)$ of the Fourier modes for the unresolved degrees of freedom W. The stochastic partial differential equation for the resolved degrees of freedom is then obtained by computing the probability distribution of W as the inverse Fourier transform of the sum of normally distributed unresolved Fourier modes and by assuming that the probability distribution for W derived in the long-time regime also holds for the unresolved modes of the initial conditions.

We emphasize that the RG approach suggests multiplicative noise as a compensation for the unresolved modes unlike the commonly obtained additive noise by Mori–Zwanzig's mode reduction (Zwanzig, 1973). Moreover, an estimate (3.6), which can be correspondingly derived by additionally accounting for the Galerkin error, shows that the influence of the stochastic force decreases for decreasing $\varepsilon := 1/N^4$.

5. Direct approach: Replacement of G_ε by white noise

The result of Lemma 3.5 also enables for a direct approach to model the unresolved degrees of freedom as completely unknown. Such a kind of complete uncertainty is generally described by white noise

$W(x)$ with zero mean and a variance equal to the power spectral density. It is very common and widely accepted to model uncertainty by white noise. Hence, we replace $\varepsilon G_\varepsilon(U(t), v^\varepsilon)$ in Equation (3.14) by

$$\mathcal{N}_\varepsilon(x,t) := \varepsilon \exp(-Ct/\varepsilon) W(x), \tag{5.1}$$

where $W(x) \in L^2(\mathcal{P}_\alpha)$ is the Gaussian random variable as motivated above, i.e. with zero mean μ and suitable variance σ. It is immediately clear that $\mathcal{N}_\varepsilon$ is a compatible replacement of G_ε since (5.1) satisfies a bound corresponding to the one in Lemma 3.5. One can follow Stinis' approach (Stinis, 2004), for example, in order to determine μ and σ by a maximum likelihood method.

6. Discussion and conclusions

We have formally developed a new stochastic mode reduction strategy with a rigorous basis by obtaining appropriate error estimates. The analysis can be summarized in three key steps as follows:

(1) *RG method*: The RG technique (Moise &Temam, 2000; Moise & Ziane, 2001) turns out to be a formal and feasible method to decompose the gKS equation into slow v^ε and fast variables w^ε, respectively. The equation for the slow modes v^ε represents a Galerkin approximation of the gKS equation plus an additional perturbed force term $\varepsilon G_\varepsilon$ which also depends on the infinite-dimensional renormalized fast modes W. An important property of the RG technique is that it can be easily extended to higher space dimensions; see Moise &Temam (2000) with respect to the RG method and Biswas & Swanson (2007) for an existence theory of the KS equation in higher space dimensions. We also remark that the dispersion term, i.e. u_{xxx}, does not affect the mode reduction analysis.

(2) *Error bounds*: We rigorously characterize the formal RG method (1) by qualitative error estimates (Theorem 3.1). These estimates further allow for an additional direct mode reduction strategy which is much simpler and straightforward but not as systematic. The basic idea is to replace the perturbed force term $\varepsilon G_\varepsilon$ directly by white noise. A physical motivation for such a simplified reduction is the fact that white noise is a well-accepted random model for complete uncertainty.

(3) *Maximum entropy principle*: Owing to the lack of a Hamiltonian structure and an invariant measure, we apply Jaynes' maximum entropy principle (Jaynes, 1957a,b) to define the renormalized fast modes W as a random variable. This random variable then, together with the renormalized approximation of the slow variable v^ε, provides a systematic explanation for the appearance of a noisy low-dimensional gKS equation. In contrast to optimal prediction we obtain optimality in the sense of maximum entropy here.

There are three main features of the new low-dimensional gKS equations:

(i) *Reliable and efficient numerics*: The low-dimensional formulation developed here should allow for reliable (since information from the unresolved degrees of freedom is included) and efficient (since low-dimensional) numerical approximations. In fact, we systematically account for the unresolved degrees of freedom by the steps (1) and (2). This is especially of importance since the choice of slow and fast variables depends on the physical problem and is often not clear. For instance, by considering the gKS in large domains, it is possible to introduce a further scale which accounts for the unstable modes. Hence, we can study three different scales such as 'unstable modes', 'slow stable modes' and 'fast stable modes'. The main question is then how to account for the unstable and the fast stable modes in an equation for the resolved slow modes only.

Moreover, the error estimates from step (2) provide a qualitative measure on how to choose the dimension of the slow variable. This is also the main advantage of mode reduction considerations over pure convergence analyses of Galerkin approximations (e.g. numerical schemes) where one completely

neglects the unresolved degrees of freedom. Hence, straightforward discretization strategies might lose model relevant information in the neglected degrees of freedom. This is a major motivation to include rigorous mode reduction strategies as an important part of the development of computational schemes. We further remark that this is a major reason why the addition of noise to deterministic partial differential equations shows good results and is currently a topic of increasing interest. It is also important to emphasize that one of the key points for the presented methodology to be computationally efficient is precisely because we add the noise *a posteriori* after solving the reduced model, something which is computationally simpler than solving the full system at every time step.

(ii) *No Hamiltonian structure; no invariant measure*: Many classical mode reduction strategies rely either on a Hamiltonian structure or an invariant measure. Based on the three steps (1)–(3), the new asymptotic reduction strategy circumvents such dependencies. For example, when classical optimal prediction methods (Chorin *et al.*, 1998) fail because of such deficiencies, the stochastic renormalization provides optimality in the sense of maximum information entropy and hence proves to be a promising alternative.

(iii) *The role of noise*: We gain a rigorous understanding of the origin of noise and the way it appears in the gKS equation. This is especially of interest due to numerical evidence provided together with a heuristic motivation in Chow & Hwa (1995) for instance.

Clearly, there are open questions and future perspectives. For example, motivated by the comparative study initiated by Stinis (2006), it would be of interest to numerically analyse and compare available mode reduction strategies such as adiabatic elimination (VanKampen, 1985), invariant manifolds (Foias *et al.*, 1988a) and optimal prediction (Chorin *et al.*, 1998) with the new RG approach developed here. Since the statistically based optimal prediction (Stinis, 2004) is performed for a truncated KS equation, it provides a convenient set-up for comparison with the new and more generally applicable method suggested here.

Another question is how can we apply the RG method to the derivation of a low-dimensional approximation for a gKS equation investigated under three scales, i.e. ‘slow unstable modes’, ‘slow stable modes’ and ‘fast stable modes’ or to explore the possibility of obtaining low-dimensional approximations of equations where noise is present from the outset, e.g. Pradas *et al*. (2011b, 2012).

The RG method is based on a natural splitting into linear and non-linear terms by the variation of constants formula. Recent studies, e.g. by Holden *et al.* (2011a,b), make use of such a splitting via a suitable numerical scheme for equations with Burgers’ non-linearity. Hence, the reliability and efficiency of the renormalized low-dimensional gKS equation motivate the application of such numerical splitting strategies to the new reduced equations derived here. Finally, we emphasize that efficient low-dimensional approximations are of great interest for numerical scrutiny of long time asymptotes. We shall examine these and related issues in future studies.

Acknowledgement

We thank Dirk Blömker (U Augsburg) for drawing our attention to the reference Blömker *et al.* (2002).

Funding

We acknowledge financial support from EPSRC Grant No. EP/H034587, EU-FP7 ITN Multiflow and ERC Advanced Grant No. 247031.

References

Akrivis, G., Papageorgiou, D. T. & Smyrlis, Y.-S. (2013) On the analyticity of certain dissipative–dispersive systems. *Bull. London Math. Soc.*, **45**, 52–60.

Biswas, A. & Swanson, D. (2007) Existence and generalized Gevrey regularity of solutions to the Kuramoto–Sivashinsky equation in. *J. Differential Equations*, **240**, 145 –163.

Blömker, D., Gugg, C. & Maier-Paape, S. (2002) Stochastic Navier–Stokes equation and renormalization group theory. *Physica D*, **137**, 137–152.

Boghosian, B. M., Chow, C. C. & Hwa, T. (1999) Hydrodynamics of the Kuramoto–Sivashinsky equation in two dimensions. *Phys. Rev. Lett.*, **83**, 5262–5265.

Chen, L.-Y., Goldenfeld, N. & Oono, Y. (1996) Renormalization group and singular perturbations: multiple scales, boundary layers, and reductive perturbation theory. *Phys. Rev. E*, **54**, 376–394.

Chorin, A. J., Kast, A. P. & Kupferman, R. 1998 Optimal prediction of under-resolved dynamics. *Proc. Natl Acad. Sci. USA*, **95**, 4094–4098.

Chow, C. & Hwa, T. (1995) Defect-mediated stability: an effective hydrodynamic theory of spatiotemporal chaos. *Physica D*, **84**, 494–512.

Cohen, B. I., Krommes, J. A., Tang, W. M. & Rosenbluth, M. N. (1976) Nonlinear saturation of the dissipative trapped-ion mode by mode coupling. *Nucl. Fusion*, **16**, 971–992.

Collet, P., Eckmann, J.-P., Epstein, H. & Stubbe, J. (1993a) Analyticity for the Kuramoto–Sivashinsky equation. *Physica D*, **67**, 321–326.

Collet, P., Eckmann, J.-P., Epstein, H. & Stubbe, J. (1993b) A global attracting set for the Kuramoto–Sivashinsky equation. *Commun. Math. Phys.*, **152**, 203–214.

Duprat, C., Giorgiutti-Dauphiné, F., Tseluiko, D., Saprykin, S. & Kalliadasis, S. (2009) Liquid film coating a fiber as a model system for the formation of bound states in active dispersive–dissipative nonlinear media. *Phys. Rev. Lett.*, **103**, 234501.

Foias, C., Jolly, M. S., Kevrekidis, I. G. & Titi, E. S. (1988a) On the computation of inertial manifolds. *Phys. Lett. A*, **131**, 433–436.

Foias, C., Nicolaenko, B., Sell, G. R. & Temam, R. (1988b) Inertial manifolds for the Kuramoto–Sivashinsky equation and an estimate of their lowest dimension. *J. Math. Pure. Appl.*, **67**, 197–226.

Giacomelli, L. & Otto, F. (2005) New bounds for the Kuramoto–Sivashinsky equation. *Commun. Pure. Appl. Math.*, **58**, 297–318.

Goodman, J. (1994) Stability of the Kuramoto–Sivashinsky and related systems. *Commun. Pure Appl. Math.*, **47**, 293–306.

Holden, H., Karlsen, K. H., Risebro, N. H. & Tao, T. (2011a) Operator splitting For the KdV equation. *Math. Comput.*, **80**, 821–846.

Holden, H., Lubich, Ch. & Risebro, N. H. (2011b) Operator splitting for partial differential equations with Burgers nonlinearity. arXiv:1102.4218, 1–12.

Homsy, G. M. (1974) Model equations for wavy viscous film flow. *Lect. Appl. Math.*, **15**, 191–194.

Jackson, D. (1930) *The Theory of Approximations*. Providence, RI: AMS Colloquium Publications.

Jaynes, E. T. (1957a) Information theory and statistical mechanics. *Phys. Rev.*, **106**, 620–630.

Jaynes, E. T. (1957b) Information theory and statistical mechanics II. *Phys. Rev.*, **108**, 171–190.

Jordan, R., Kinderlehrer, D. & Otto, F. (1998) The variational formulation of the Fokker–Planck equation. *SIAM J. Math. Anal.*, **29**, 1.

Kalisch, H. & Raynaud, X. (2007) On the rate of convergence of a collocation projection of the KdV equation. *ESAIM Math. Model. Numer. Anal.*, **41**, 95–110.

Kalliadasis, S., Ruyer-Quil, C., Scheid, B. & Velarde, M. G. (2012) *Falling Liquid Films*. Springer Series on Applied Mathematical Sciences, vol. 176. London: Springer.

Kuramoto, Y. & Tsuzuki, T. (1976) Persistent propagation of concentration waves in dissipative media far from thermal equilibrium. *Progr. Theoret. Phys.*, **55**, 356–369.

Larkin, N. (2004) Korteweg–de Vries and Kuramoto–Sivashinsky equations in bounded domains. *J. Math. Anal. Appl.*, **297**, 169–185.

Lorenz, E. N. (1996) Predictability–a problem partly solved. *ECMWF Seminar on Predictability, 1995*. United Kingdom: Reading, pp. 1–18.

Maday, Y. & Quarteroni, A. (1988) Error analysis for spectral approximation of the Korteweg–de Vries equation. *ESAIM Math. Model. Numer. Anal.*, **22**, 499–529.

Manneville, P. (1990) *Dissipative Structures and Weak Turbulence*. New York: Academic Press.

Marion, M. & Temam, R. (1989) Nonlinear Galerkin methods. *SIAM J. Numer. Anal.*, **26**, 1139–1157.

Moise, I. & Temam, R. (2000) Renormalization group method: application to Navier–Stokes equation. *Discrete Contin. Dyn. Syst.*, **6**, 191–210.

Moise, I. & Ziane, M. (2001) Renormalization group method. Applications to partial differential equations. *J. Dynam. Differential Equations*, **13**, 275–321.

Mori, H. (1965) Transport, collective motion, and Brownian motion. *Progr. Theoret. Phys.*, **33**, 423–455.

Otto, F. (2009) Optimal bounds on the Kuramoto–Sivashinsky equation. *J. Funct. Anal.*, **257**, 2188–2245.

Pradas, M., Kalliadasis, S., Nguyen P.-K. & Bontozoglou, V. (2013) Bound-state formation in interfacial turbulence: direct numerical simulations and theory. *J. Fluid Mech.*, 716:R2.

Pradas, M., Pavliotis, G. A., Kalliadasis, S., Papageorgiou, D. T. & Tseluiko, D. (2012) Additive noise effects in active nonlinear spatially extended systems. *European J. Appl. Math.*, **23**, 563–591.

Pradas, M., Tseluiko, D. & Kalliadasis, S. (2011a) Rigorous coherent-structure theory for falling liquid films: Viscous dispersion effects on bound-state formation and self-organization. *Phys. Fluids*, **23**, 044104.

Pradas, M., Tseluiko, D., Kalliadasis, S., Pagageorgiou, D. T. & Pavliotis, G. A. (2011b) Noise induced state transitions, intermittency, and universality in the noisy Kuramoto–Sivashinksy equation. *Phys. Rev. Lett.*, **106**, 060602.

Rosenkrantz, R. D. (ed.) (1989) *Jaynes, E.T.: Papers on Probability, Statistics and Statistical Physics*. Dordrecht: Kluwer Academic Publisher.

Saprykin, S., Demekhin, E. A. & Kalliadasis, S. (2005) Self-organization of two-dimensional waves in an active dispersive–dissipative nonlinear medium. *Phys. Rev. Lett.*, **94**, 224101.

Shannon, C. E. (1948) The mathematical theory of communication. *Bell Syst. Tech. J.*, **27**, 379–423, 623–656. Reprint.

Shraiman, B. I. (1986) Order, disorder, and phase turbulence. *Phys. Rev. Lett.*, **57**, 325–328.

Sivashinsky, G. (1979) On self-turbulization of a laminar flame. *Acta Astronaut.*, **6**, 569–591.

Stinis, P. (2004) Stochastic optimal prediction for the Kuramoto–Sivashinsky equation. *Multiscale Model. Simul.*, **2**, 613–638.

Stinis, P. (2006) A comparative study of two stochastic mode reduction methods. *Physica D*, **213**, 197–213.

Tadmor, E. (1986) The well-posedness of the Kuramoto–Sivashinsky equation. *SIAM J. Math. Anal.*, **17**, 884.

Tseluiko, D. & Kalliadasis, S. (2012) Weak interaction of solitary pulses in active dispersive–dissipative nonlinear media. *IMA J. Appl. Math.*, 1–26. doi:10.1093/imamat/hxs064.

Tseluiko, D., Saprykin, S., Duprat, C., Giorgiutti-Dauphiné, F. & Kalliadasis, S. (2010a) Pulse dynamics in low-Reynolds-number interfacial hydrodynamics: Experiments and theory. *Physica D*, **239**, 2000–2010.

Tseluiko, D., Saprykin, S. & Kalliadasis, S. (2010b) Interaction of solitary pulses in active dispersive-dissipative media. *Proc. Est. Acad. Sci.*, **59**, 139–144.

van Kampen, N. G. (1985) Elimination of fast variables. *Phys. Rep.*, **124**, 69–160.

Wilson, K. G. 1975 The renormalizaton group: critical phenomena and the Kondo problem. *Rev. Mod. Phys.*, **47**, 773–840.

Zwanzig, R. (1961) Memory effects in irreversible thermodynamics. *Phys. Rev.*, **124**, 983–992.

Zwanzig, R. (1973) Nonlinear generalized Langevin equations. *J. Stat. Phys.*, **9**, 215–220.

Appendix

We prove the following lemma:

LEMMA A.1 $R_2(j)$ is an empty set.

Proof. Let us first recall the definition of the second resonance, that is,

$$\rho_k^w + \rho_l^w = \rho_j^w, \tag{A.1}$$

where the indices satisfying (A.1) belong to $R_2(j)$. We further remind the reader of the convention according to (2.3)

$$\rho_k^w = \frac{1}{N^4}\left(-\nu\left(\frac{k}{\alpha}\right)^2 - \mathrm{i}\delta\left(\frac{k}{\alpha}\right)^3 + \kappa\left(\frac{k}{\alpha}\right)^4\right). \tag{A.2}$$

Taking the imaginary parts of (A.1), one obtains

$$\left(\frac{k}{\alpha}\right)^n + \left(\frac{l}{\alpha}\right)^n = \left(\frac{j}{\alpha}\right)^n, \tag{A.3}$$

for $n = 3$. Taking the real parts, one only gets

$$-\nu\left(\left(\frac{k}{\alpha}\right)^2 + \left(\frac{l}{\alpha}\right)^2\right) + \kappa\left(\left(\frac{k}{\alpha}\right)^4 + \left(\frac{l}{\alpha}\right)^4\right) = -\nu\left(\frac{j}{\alpha}\right)^2 + \kappa\left(\frac{j}{\alpha}\right)^4. \tag{A.4}$$

In what follows, we transform (A.4) into an expression which convinces us that there are no indices k and l that satisfy (A.4). To this end, we make use of the fact that $j = k + l$ which reads after taking the square on each side as

$$j^2 = (k+l)^2 = k^2 + l^2 + 2kl. \tag{A.5}$$

Multiplying now (A.4) by α^2 gives

$$-k^2 - l^2 + \frac{\kappa}{\alpha^2\nu}(k^4 + l^4) = -j^2 + \frac{\kappa}{\alpha^2\nu}j^4, \tag{A.6}$$

and after applying (A.5) on the right-hand side, we obtain,

$$-k^2 - l^2 + \frac{\kappa}{\alpha^2\nu}(k^4 + l^4) = -k^2 - l^2 - 2kl + \frac{\kappa}{\alpha^2\nu}(k^2 + l^2 + 2kl)^2, \tag{A.7}$$

and hence becomes

$$r - 3kl = 2(k^2 + l^2), \tag{A.8}$$

where we set $r := \alpha^2\nu/\kappa$, which is positive.

Equation (A.8) cannot be satisfied by any integers k and l if (i) $N^2 \geqslant \frac{8}{7}r$ or if (ii) L, ν and κ are not proportional to π. □

REMARK A.1 We note that without assuming either that (i) L, ν, κ are not proportional to π or that (ii) $N^2 \geqslant \frac{8}{7}\alpha^2\nu/\kappa$, we obtain two explicit solutions for k and l via (A.8) over a depressed cubic equation, i.e.

$$k = (r(-64 \mp \sqrt{64^2 + 392^2 r/27}))^{1/3} - \frac{392r}{3(r(-64 \mp \sqrt{64^2 + 392^2 r/27}))^{1/3}},$$

where the same expression also defines l. Herewith, it remains to check whether, for a given $r \in \mathbb{R}$, the solutions k and l are integers and whether they satisfy $|k + l| > N$, $|k| > N$ and $|l| > N$.