

S. I. Dolaptchiev · U. Achatz · I. Timofeyev

Stochastic closure for local averages in the finite-difference discretization of the forced Burgers equation

Received: date / Accepted: date

Abstract We present a new approach for the construction of stochastic subgrid scale parameterizations. Starting from a high-resolution finite-difference discretization of some model equations, the new approach is based on splitting the model variables into fast, small-scale and slow, large-scale modes by averaging the model discretization over neighboring grid cells. After that, the fast modes are eliminated applying a stochastic mode reduction procedure. This procedure is a generalization of the mode reduction strategy proposed by Majda, Timofeyev & Vanden-Eijnden, in that it allows for oscillations in the closure assumption. The new parameterization is applied to the forced Burgers equation and is compared with a Smagorinsky type subgrid scale closure.

Keywords subgrid scale modeling, stochastic parameterization, coarse-grained models, large eddy simulation

PACS First · Second · More

1 Introduction

The construction of reduced models, describing only a small number of large-scale, low-frequency patterns in the atmosphere-ocean system, is a topic of ongoing research. Such simplified models can provide a tool for gaining insight into the comprehensive general circulation models (GCMs) with millions of degrees of freedom. Reduced models can also be viewed as a numerically attractive alternative to GCMs for performing paleoclimate or ensemble simulations. Such models can be formulated by projecting the governing equations on some suitable basis functions, e.g., Empirical-Orthogonal-Functions (EOFs) [36, 1, 2, 22] or principal interaction patterns [4, 21]. Alternative approaches in the construction of reduced models include the regression-fitting of a linear operator [5, 39] or the derivation of reduced model equations applying multiple scales asymptotic techniques [23, 12].

S. I. Dolaptchiev
Institut für Atmosphäre und Umwelt
Goethe-Universität Frankfurt
Altenhöferallee 1, 60438 Frankfurt am Main, Germany
E-mail: Dolaptchiev@iau.uni-frankfurt.de

U. Achatz
Institut für Atmosphäre und Umwelt
Goethe-Universität Frankfurt, Frankfurt am Main, Germany
Altenhöferallee 1, 60438 Frankfurt am Main, Germany
E-mail: Achatz@iau.uni-frankfurt.de

I. Timofeyev
Department of Mathematics
University of Houston, Houston, Texas
E-mail: ilya@math.uh.edu

A main issue in all strategies for the construction of reduced models is the parameterization of the unresolved scales or subgrid scales (SGS) (e.g. [3]). Such closure problems appear in a variety of atmosphere-ocean applications and recently stochastic methods are successfully utilized for SGS parameterizations there [6, 31, 11, 32, 18]. Stochastic parameterizations include stochastic differential equations (SDEs) (e.g. [17, 15, 9]), Markov chains (e.g. [26, 8, 19]) and randomization of deterministic closure (e.g. [6, 31]).

Majda, Timofeyev and Vanden-Eijnden proposed a systematic self-consistent mathematical framework known as the MTV stochastic mode reduction strategy for parameterizing unresolved modes in climate models using SDEs [24, 25, 28]. The main assumptions in the general MTV procedure (see [24] for details) are: a scale separation in time between resolved/unresolved modes and ergodicity and mixing in the underlying system. Often the derivation of an explicit form of the SGS model can be considerably simplified by the ad hoc approximation that the term describing fast self-interactions in the equation for the unresolved modes can be modeled by a particular stochastic process, namely a stationary diagonal Ornstein-Uhlenbeck (OU) process. The parameters of the OU process are estimated from data. With these parameters as input the MTV procedure systematically derives all deterministic and stochastic SGS correction terms in the equation for the resolved modes. This semi-analytical approach differs from other studies [1, 2, 5, 11, 20] where an ad hoc approximation is made in the equation for resolved variables. There, a certain type of model for the interactions between resolved/unresolved modes is imposed and its coefficients are determined by some fitting procedure.

The assumption, made in the MTV, of a scale separation in time between the resolved and unresolved modes has to be verified for the particular application in order to test the applicability of the theory. Indeed, there are empirical approaches for SGS parameterization (e.g., [1, 5, 8]), which do not require a scale separation. However, such methods share the problem with the ad hoc assumption on the interactions discussed in the previous paragraph. Further, the MTV stochastic mode reduction strategy has the advantage that it is rigorously valid in the asymptotic limit of infinite scale separation. Moreover, a number of studies [25, 28, 15, 14] demonstrated that the MTV approach can be successfully applied in a variety of models ranging from simple to realistic atmospheric models with more or less pronounced time scale separation.

In the MTV studies we mentioned the resolved modes are defined by some leading global basis functions such as Fourier harmonics or EOFs. These modes also evolve on the largest time scales. However, the explicit formulation of the model equations using global basis functions effectively limits the number of large-scale modes one can consider in atmospheric applications.

The main motivation for the current approach is the necessity to develop efficient numerical methods for long-term integration of coupled atmosphere-ocean models. Ocean models are often discretized using finite-difference or finite-volume schemes due to complex boundaries (e.g. [16]). Recently, finite-difference discretizations are becoming more popular for global models of the atmosphere as well [33, 29, 34], while they have traditionally been in use for regional atmospheric modelling. When considering the coupled system, one has to develop reduced, coarse-grained models for computing the locally averaged wind stress in the atmospheric boundary layer, which acts as a primary forcing mechanism the atmosphere exhibits on the ocean. Therefore, the technique developed in this paper can be utilized to derive the effective coarse-grained model for the local averages of various atmospheric quantities which will considerably accelerate the numerical simulations of the coupled model.

In the current approach the resolved and unresolved modes are defined locally (rather than using some global basis functions) by utilizing a standard averaging method from Large Eddy Simulation (LES). After that, terms involving only SGS modes in the equation for the unresolved modes are parameterized by a stationary and homogeneous in space OU process. The coefficients of this process are estimated from data. Next, all SGS modes are eliminated by applying a stochastic mode reduction procedure. This procedure is a generalization of the MTV mode reduction strategy, because it allows for oscillations in the closure assumption. The new approach provides a framework for constructing SGS parameterizations consistent with the spatial discretization of the model. It also bears the potential for application in scenarios, like LES, where even the number of resolved degrees of freedom is large. The new method is studied by applying it to the stochastically forced Burgers equation.

The outline of this paper is as follows: in section 2 we introduce the discrete version of the forced Burgers model. In section 3 we discuss the stochastic mode reduction procedure and present a reduced model with a stochastic SGS closure. The performance of this model is assessed in section 4, where

it is compared with two other reduced models. A summary and a conclusion are presented in the last section.

2 The forced Burgers equation

2.1 Model formulation

The Burgers equation shows a lot of properties in common with the Navier-Stokes equations, e.g., quadratic nonlinearities, conservation and invariance properties. A stochastically forced Burgers equation provides thus a test model for developing techniques for turbulence parameterization. We consider here the one dimensional Burgers equation with dissipation and random forcing (e.g. [7])

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} \left(\frac{u^2}{2} - \nu \frac{\partial u}{\partial x} \right) = f(x, t) \quad (1)$$

over a periodic domain of size L and with a constant diffusion coefficient ν . The function $f(x, t)$ represents a random forcing within some wavenumber range and will be specified later on. Discretizing eq. (1) in space, one can write

$$\frac{d}{dt} u_i + \frac{F_{i+\frac{1}{2}} - F_{i-\frac{1}{2}}}{\Delta x} = f_i. \quad (2)$$

Where u_i and f_i denote the values of the functions u and f at the center of the discrete cell i ($i \in s = \{0, \dots, N-1\}$) of width $\Delta x = \frac{L}{N}$; $F_{i+\frac{1}{2}}, F_{i-\frac{1}{2}}$ denote the fluxes at the boundaries of cell i . These fluxes have to be approximated using values from neighboring cells, here we apply the discretization of [40] for the quadratic nonlinearities and the standard three point discretization for the diffusion term

$$F_{i+\frac{1}{2}} = \frac{1}{6} (u_{i+1}^2 + u_i u_{i+1} + u_i^2) - \nu \frac{u_{i+1} - u_i}{\Delta x}, \quad (3)$$

$$F_{i-\frac{1}{2}} = \frac{1}{6} (u_i^2 + u_i u_{i-1} + u_{i-1}^2) - \nu \frac{u_i - u_{i-1}}{\Delta x}. \quad (4)$$

The above discretization conserves total momentum and total energy (in the absence of forcing and dissipation). An analysis of its mathematical properties, as well as numerical simulations of different regimes in the inviscid Burgers equation, can be found in [27].

The resolution scale Δx in the discrete model (2) has to be sufficiently small in order to resolve the dissipation scale. In LES the model equations are solved on some much coarser grid, so all effects of SGS have to be parameterized. The discrete model has to be reformulated in terms of resolved and unresolved scales, this is done by averaging the equations. Following the volume-balance procedure of [35], we average over neighboring, fixed in space, ‘‘fine’’ grid cells of the original width Δx . We define a ‘‘coarse’’ grid with a spacing of $n\Delta x$ and cell index set $s_x = \{0, \dots, N_x - 1\}$, where $N_x = \frac{N}{n}$ is the total number of coarse cells (N multiple of n). Let $x_i, i \in s_x$ denote the averaged values of $u_i, i \in s$ on the ‘‘coarse’’ grid, x_i is computed by applying a top-hat filter to u_i . This filtering can be written simply as the arithmetic mean over n neighboring grid cells

$$x_i = \frac{1}{n} \sum_{k=ni}^{n(i+1)-1} u_k. \quad (5)$$

Next, we split u_i into a mean $x_{\hat{i}}$ and a deviation y_i

$$u_i = x_{\hat{i}} + y_i, \quad (6)$$

where the index \hat{i} denotes the coarse cell in which the fine cell i is located. We will refer to $x_{\hat{i}}, y_i$ as resolved, unresolved mode. Applying the averaging operator (5) to the discretized Burgers equation, eq. (2) can be written in terms of $x_{\hat{i}}$ and y_i

$$\frac{d}{dt}x_{\hat{i}} + \frac{F_{n(\hat{i}+1)-\frac{1}{2}} - F_{n\hat{i}-\frac{1}{2}}}{n\Delta x} = F_{\hat{i}}^x, \quad (7)$$

$$\frac{d}{dt}y_i + \frac{F_{i+\frac{1}{2}} - F_{i-\frac{1}{2}}}{\Delta x} - \frac{F_{n(\hat{i}+1)-\frac{1}{2}} - F_{n\hat{i}-\frac{1}{2}}}{n\Delta x} = F_i^y. \quad (8)$$

Here F^x, F^y represent the projections of the forcing on the x, y -modes. We apply a forcing function which projects only on the resolved modes. Such large-scale forcing is motivated by forcing mechanisms in the atmosphere, e.g., baroclinic instability, planetary scale orography and land-ocean thermal contrast. The current approach is applicable to situations where the unresolved modes are forced as well, if the forcing can be modeled as a stochastic process of linear Langevin type. The particular forcing function considered here reads

$$F_{\hat{i}}^x = \sum_{k=k_b}^{k_e} \frac{\alpha_k A}{\sqrt{k\Delta t}} \cos\left(2\pi\left(\frac{k\hat{i}n\Delta x}{L} + \phi_k\right)\right), \quad F_{\hat{i}}^y = 0, \quad (9)$$

with α_k, ϕ_k random numbers chosen from a normal distribution at each time step Δt of integration; k_b and k_e define the wavenumber range of the forcing and A is a constant. In the model presented here, k_b, k_e are always set to some low wavenumbers (either 1, 3 or 4, 7). As we will discuss later, the coarse grid width $n\Delta x$ is sufficiently small to resolve these wavenumbers. Thus, the forcing (9) has the effect, as if the model is forced on the fine grid in the same wavenumber range with a random forcing function f : the projection of f on the unresolved modes is negligible. After testing this in numerical simulations, we consider here the forcing (9). In this way, the model is accelerated, since the forcing has to be determined only over the coarse grid.

The flux terms in (7) and (8) contain both resolved and unresolved modes. We split these terms depending on the modes involved and obtain the following system of ODEs for the discretized model equations

$$\begin{aligned} \dot{x}_i = F_i^x &+ \sum_{j \in s_x} L_{ij}^{xx} x_j + \sum_{j \in s_x} \sum_{k \in s_x} B_{ijk}^{xxx} x_j x_k + \sum_{j \in s_x} \sum_{k \in s} B_{ijk}^{xxy} x_j y_k \\ &+ \sum_{j \in s} \sum_{k \in s} B_{ijk}^{xyy} y_j y_k + \sum_{j \in s} L_{ij}^{xy} y_j, \quad i \in s_x, \end{aligned} \quad (10)$$

$$\begin{aligned} \dot{y}_i &= \sum_{j \in s_x} L_{ij}^{yx} x_j + \sum_{j \in s_x} \sum_{k \in s_x} B_{ijk}^{yxx} x_j x_k + \sum_{j \in s_x} \sum_{k \in s} B_{ijk}^{yxy} x_j y_k \\ &+ \sum_{j \in s} \sum_{k \in s} B_{ijk}^{yyy} y_j y_k + \sum_{j \in s} L_{ij}^{yy} y_j, \quad i \in s. \end{aligned} \quad (11)$$

Here we have introduced the MTV notation for the interaction terms; the explicit form of the different terms in the equations above is given in appendix A. The terms denoted with $B^{\alpha\beta\gamma}$ arise from the quadratic nonlinearities; the first superscript α indicates the modes on which they project, the second two, β and γ , the two modes involved. The terms of the form $L^{\alpha\beta}$ result from the dissipation term; again the first superscript denotes the modes on which it projects and the second indicates the mode involved. In (10), (11) the indices of the summation signs go over s or s_x , but one has to keep in mind that due to the local form of the discretization only 'neighboring' indices are involved, implying that the tensors with the interaction coefficients are sparse.

2.2 Model setup

Equations (10), (11) represent another form of the discretized model (2), we will refer to this model as the full model and to its integration as direct numerical simulation (DNS). We integrate the full model in time using a 3rd order Runge-Kutta method with the following choice of parameters: $\Delta t = 10^{-2}$, $\nu = 0.02$, $L = 100$, $N = 512$, $A = \sqrt{2}10^{-2}$, $k_b = 1$, $k_e = 3$. We will refer to this setup as our reference setup. Model output is written every 40 time steps and we generate a time series with 1.5×10^6 entries. Most

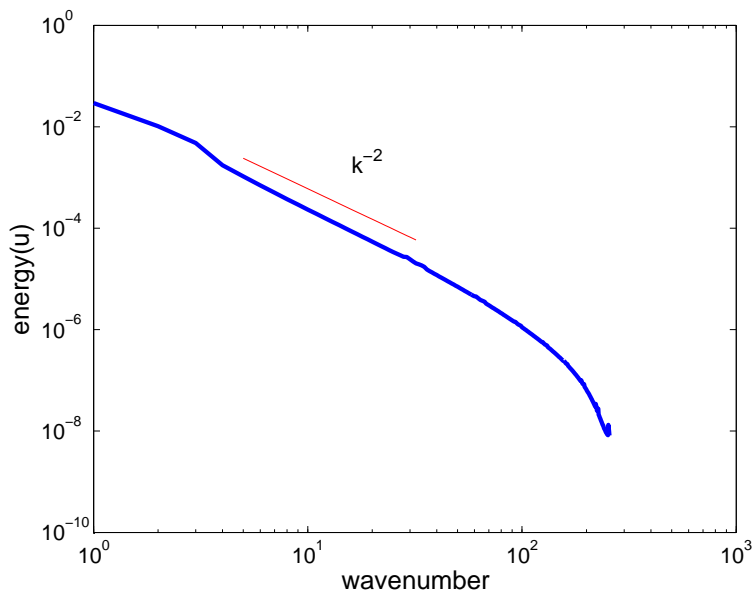


Fig. 1 Power spectrum of u from DNS with the reference setup, a line with a k^{-2} slope is also drawn.

of the time the solution is characterized by up to three shocks which propagate in the domain and merge (e.g. see the figures in [10])

The power spectrum of u from DNS is shown in figure 1. Overall three regions are visible in the spectrum: an energy containing range for wave numbers $k \leq 3$, an inertial range with a k^{-2} slope and a dissipation range. Our goal is to derive from (10), (11) a reduced model for the x -modes only and parameterize the effects from the y -modes. The separation of u in x - and y -modes depends on the choice of the averaging interval n in (5). The averaging interval defines a characteristic length scale for the x variables. This scale should be on the one hand far from the dissipation scale and on the other hand small enough to resolve the large-scale forcing in the model. We set the parameter n to 16; in the setup described above it corresponds to resolving the x variables up to wave number $k_{max} = N/(2n) = 16$ (32 coarse grid cells). In this way we can resolve the energy containing range and only a part of the inertial range; the effects from the remaining scales have to be modeled.

The MTV stochastic mode reduction procedure requires a scale separation in time between the x - and the y -modes. The spatially averaged time autocorrelation functions (ACFs) of these modes are presented in fig. 2. Because of the spatial homogeneity of the model, all statistical moments for x we present in the paper are computed by averaging over coarse grid cells (the statistics of the y -modes show some inhomogeneity inside each coarse cell, but this is not crucial here). The different decay time scales (computed by evaluating the area under the curves in fig. 2) indicate that the x -modes are about five times slower than the y -modes. In general, the autocorrelation time scales of x and y will depend on n and on the time scales set by the forcing and dissipation.

In order to exclude bifurcation behavior in the Burgers model like the one reported for the Lorenz-96 model [30], we performed a series of experiments by varying separately the control parameters of the system: the magnitude of the forcing and the diffusion coefficient. Runs were performed with $A = 0.1, 0.25, 0.3, 0.4$ and $\nu = 0.01, 0.015, 0.025, 0.03, 0.04$. The various simulations did not show any qualitative difference from the statistical behavior of the solution in the reference setup with $A = 0.2, \nu = 0.02$.

3 Stochastic mode reduction

3.1 Closure assumptions

In this section we discuss the assumptions made in order to carry out the stochastic mode reduction procedure. First, we observe that the variables y inside each coarse cell are subject to a constraint: they

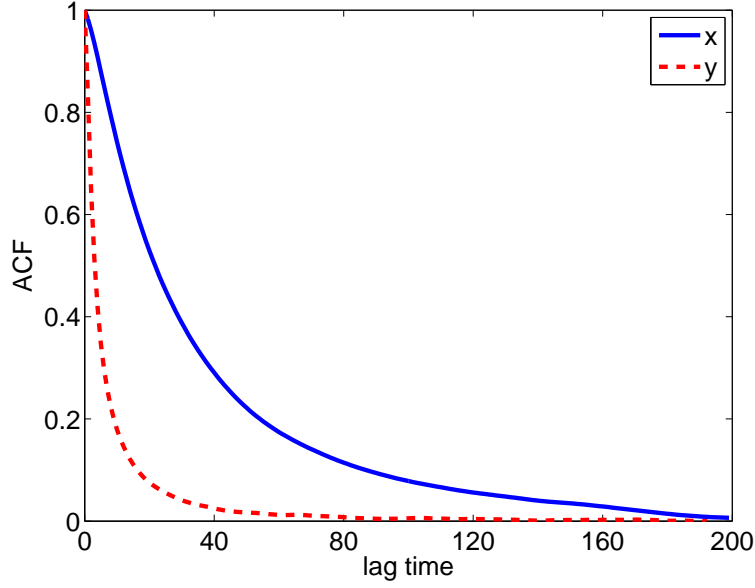


Fig. 2 Spatially averaged time autocorrelation function (ACF) for the x - and y -modes from DNS with the reference setup.

have to sum to zero and thus only $n - 1$ y -modes are independent inside a coarse cell. As a consequence instead of N y -modes in the whole domain we have only $N_z = N - N_x$ independent modes, which we denote with \hat{y} . These modes are found by transforming the y -variables inside each coarse cell into Fourier space and skipping the zero mode. The transformation can be written in a formal way as

$$\hat{y}_k = \hat{T}_{kj} y_j, \quad (12)$$

$$y_j = \hat{R}_{jk} \hat{y}_k, \quad (13)$$

with $\hat{T} \in \mathbb{R}^{N_z, N}$, $\hat{R} \in \mathbb{R}^{N, N_z}$, see appendix B for details. Here and in the following paragraphs we make use of Einstein's summation convention. Next, we assume that in Fourier space the term $L^{yy}y + B^{yyy}yy$ from (11) can be modeled as an Ornstein-Uhlenbeck (OU) process in \hat{y} with drift and diffusion matrices Γ and Σ

$$\hat{T}_{ij} \left(L_{jk}^{yy} y_k + B_{jkl}^{yyy} y_k y_l \right) = \Gamma_{ij} \hat{y}_j + \Sigma_{ij} \dot{W}_j, \quad (14)$$

where W_j denotes a Wiener process. Further, we assume, that the OU process couples only \hat{y} -modes corresponding to the same coarse cell. Due to the spatial homogeneity of the model, we will have the same OU parameters in each coarse cell. Thus, Γ, Σ can be represented as block-diagonal matrices, built from N_x blocks of matrices $g, s \in \mathbb{R}^{(n-1), (n-1)}$

$$\Gamma = \begin{pmatrix} g & \dots & 0 \\ \vdots & g & \vdots \\ 0 & & \ddots \end{pmatrix}, \quad \Sigma = \begin{pmatrix} s & \dots & 0 \\ \vdots & s & \vdots \\ 0 & & \ddots \end{pmatrix}. \quad (15)$$

Using standard techniques from time series analysis, the matrices Γ and Σ can be estimated from data. This is done in the following way: we compute from the reference simulations a time series of y and of the term $L^{yy}y + B^{yyy}yy$. These time series are transformed into Fourier space according to (12). From the new time series we estimate the parameters of the OU process applying a standard maximum-likelihood-approach [38].

In general Γ has complex eigenvalues: the negative real parts assure a stable OU process and the imaginary parts describe an oscillation. For all estimates of Γ reported here, the eigenvalue analysis revealed a stable drift matrix. If Γ can be diagonalized, then real matrices $\Lambda, U, U^{-1} \in \mathbb{R}^{N_z, N_z}$ can be found, such that $\Gamma = U\Lambda U^{-1}$, where Λ has a real block-diagonal form with the real parts of the eigenvalues of Γ on the main diagonal and the imaginary parts on the upper/lower diagonal, see appendix C for details. We introduce a new variable z_i ($i \in s_z = \{0, \dots, N_z - 1\}$), defined by the transformation matrices T and R in the following way

$$z_i = T_{ik}y_k = U_{ij}^{-1}\hat{T}_{jk}y_k, \quad (16)$$

$$y_i = R_{ik}z_k = \hat{R}_{ij}U_{jk}z_k. \quad (17)$$

Applying T to (11), making use of (14) and expressing the y -variables in terms of z , we obtain from (10), (11) the following system of SDEs

$$dx_i = \left(F_i^x + L_{ij}^{xx}x_j + B_{ijk}^{xxx}x_jx_k + B_{ijk}^{xxy}x_jR_{kl}z_l + B_{ijk}^{xyy}R_{jl}z_lR_{km}z_m + L_{ij}^{xy}R_{jl}z_l \right) dt, \quad (18)$$

$$dz_i = \left(T_{ij}L_{jk}^{yx}x_k + T_{ij}B_{jkl}^{yxx}x_kx_l + T_{ij}B_{jkl}^{yxy}x_kR_{lm}z_m + \Lambda_{ij}z_j \right) dt + \hat{\Sigma}_{ij}dW_j. \quad (19)$$

Here we have introduced $\hat{\Sigma} = U^{-1}\Sigma$. The classical MTV procedure for mode elimination [24] assumes a diagonal diffusion coefficient in the noise term from (19). This assumption is also made in the current approach and $\hat{\Sigma}$ is approximated by an effective diagonal matrix with diagonal elements $\sigma_i = (\sum_j \hat{\Sigma}_{ij}^2)^{1/2}$. This approximation was verified in numerical simulations of the model (18),(19) in the setup from section 2.2 with full and approximated diffusion coefficient which showed no significant differences in the statistical behavior. Finally, equations (18),(19) can be formally written in a more compact form as

$$dx_i = \left(F_i^x + L_{ij}^{xx}x_j + B_{ijk}^{xxx}x_jx_k + \frac{1}{\varepsilon}B_{ijk}^{xxz}x_jz_k + \frac{1}{\varepsilon}B_{ijk}^{xzz}z_jz_k + \frac{1}{\varepsilon}L_{ij}^{xz}z_j \right) dt, \quad (20)$$

$$dz_i = \left(\frac{1}{\varepsilon}L_{ik}^{zx}x_k + \frac{1}{\varepsilon}B_{ikl}^{zxx}x_kx_l + \frac{1}{\varepsilon}B_{ikl}^{zxx}x_kz_l + \frac{1}{\varepsilon^2}\Lambda_{ij}z_j \right) dt + \frac{1}{\varepsilon}\sigma_i dW_i. \quad (21)$$

Here we have introduced a small parameter $\varepsilon \ll 1$, which is a measure of the scale separation in time between the x and z (or y) modes. Introducing ε is a technical step in order to carry out the stochastic mode reduction strategy and ε is set to one in all final results. The different ε factors in front of the terms in (20), (21) indicate the different time scales of these terms. It was shown that the stochastic mode reduction strategy can be applied for a wide range of models, where ε ranges from quite small up to order one [25,28,15,14]. Equations (20),(21) represent the full model with Ornstein-Uhlenbeck closure and we will refer to its integrations as OU-DNS. The DNS and OU-DNS are compared in the next section.

3.2 Numerical tests of the asymptotic limit

The reduced stochastic model, introduced in the next section, represents the asymptotic limit $\varepsilon \rightarrow 0$ of the OU-DNS. Such limit corresponds to infinitely large time scale separation between the x and z (or y) modes. This motivated us to study the limit $\varepsilon \rightarrow 0$ by systematically decreasing the value of ε in (20),(21) and running the resulting model.

Figures 3(a),(b),(c) and table 1 show that for $\varepsilon = 1$ the OU closure reasonably approximates the power spectrum, ACF and the kurtosis of x from the full model. As ε gets smaller we observe, that the overestimation of the power at higher wavenumbers in the energy spectrum of the OU-DNS is slightly reduced, but there is tendency of small overestimation at small wavenumbers. In the ACF (kurtosis) there is a small tendency to overestimate (underestimate) the moments as $\varepsilon \rightarrow 0$. However, we can conclude that the asymptotic limit of the OU-DNS is close to the case $\varepsilon = 1$ and we can apply the stochastic mode reduction strategy. The results from the mode elimination are summarized in the next section.

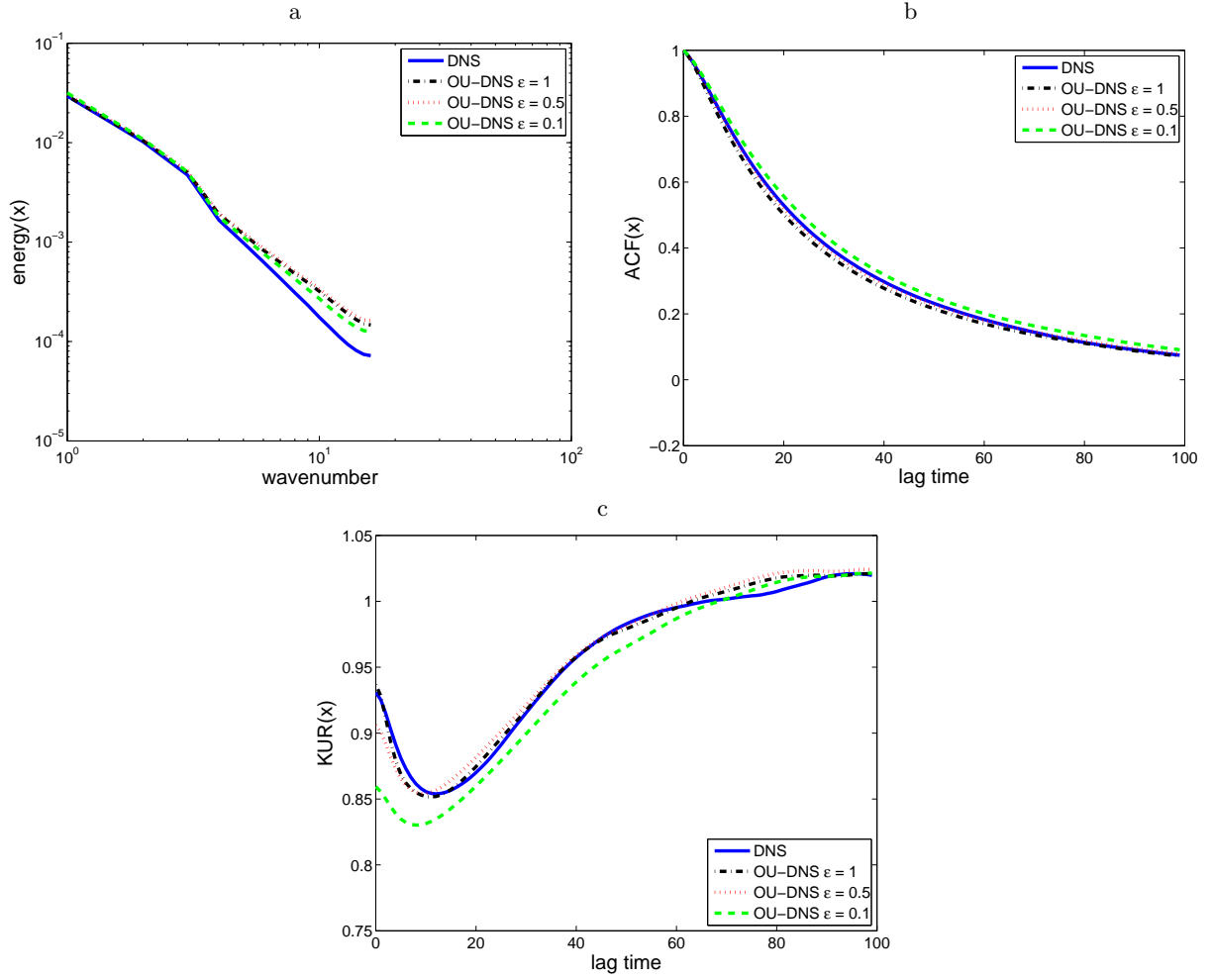


Fig. 3 Energy spectra, autocorrelation function (ACF) and kurtosis (KUR) of x for: DNS and OU-DNS with different values of ε . Reference setup. In the energy spectra all curves nearly overlap for small wavenumbers. The kurtosis for lag s is given by $K(s) = \langle x(t+s)^2 x^2(t) \rangle \{ \langle x^2(t) \rangle^2 + 2 \langle x(t+s)x(t) \rangle^2 \}^{-1}$, where $\langle \cdot \rangle$ denotes a time average.

Table 1 Variance (VAR) and integrated absolute value of the ACF (Int. ACF) of x from DNS and from OU-DNS with different values of ε . The relative error of the OU-DNS wrt. the DNS value is given in brackets.

| | DNS | $\varepsilon = 1.$ | $\varepsilon = 0.5$ | $\varepsilon = 0.1$ |
|-----------|--------|--------------------|---------------------|---------------------|
| VAR | 0.0247 | 0.0260 (0.05) | 0.0269 (0.09) | 0.0268 (0.09) |
| Int. ACF. | 141.2 | 133.8 (0.05) | 139.8 (0.01) | 151.4 (0.07) |

3.3 The reduced stochastic model with subgrid scale closure

The stochastic mode reduction is carried out by performing an asymptotic expansion in the Kolmogorov backward equation corresponding to the SDEs (20),(21). For the detailed discussion of the procedure we refer to [24]; the generalization for an oscillating OU process is summarized in appendix C. After eliminating the SGS modes, we obtain the following effective stochastic differential equation for x

$$dx_i(t) = \sum_{j,k=i-1}^{i+1} B_{ijk}^{xxx} x_j x_k dt + \sum_{j=i-1}^{i+1} L_{ij}^{xx} x_j dt + F_i^x dt$$

$$\begin{aligned}
& + \sum_{j=i-2}^{i+2} M_{ij} x_j dt + \sum_{j=i-2}^{i+2} Q_{ijk} x_j x_k dt + \sum_{j,k,l=i-2}^{i+2} C_{ijkl} x_j x_k x_l dt \\
& + \sigma^{(1)} dW_i^{(1)} + \sigma_i^{(2)}(x) dW_i^{(2)},
\end{aligned} \tag{22}$$

where $W^{(1)}, W^{(2)}$ denote vectors with N_x entries of independent Wiener processes. We will refer to (22) as the reduced stochastic model (RSM). The first three terms on the rhs. of (22) result from the full model when all SGS modes are neglected. The other terms in (22) arise from the SGS model, they represent: linear, quadratic and cubic corrections, additive noise and correlated additive and multiplicative noise. In order to stress the local form of the SGS model, we explicitly give here the intervals for the summation signs. The explicit form of the SGS parameterization is given in appendix D. The RSM conserves total momentum, as the full model does.

4 Results for the reduced stochastic model

In the following section we study the performance of the RSM by comparing its statistics with those from the full model (DNS) and from two other reduced models. These two models are: the bare truncation model and a model with Smagorinsky SGS parameterization. We will refer to these models as BRT and SMG, respectively. The BRT is the RSM model with no SGS correction terms, the SMG incorporates a purely diffusive type parameterization and is a benchmark model in LES. Details of the SMG can be found in the appendix E. Unless otherwise stated, the simulations presented here have the reference setup. Although the RSM allows a larger time step than the DNS (because of the coarser resolution), the time step was not changed.

It is important to note that the BRT differs from the SMG with switched off SGS model (SMG-no-SGS). In particular, the last model would be unstable for the present resolution. The difference between the two models is in the dissipation term. For the BRT and SMG-no-SGS this term is proportional to $\nu/(n\Delta^2 x)$ (see eq. (31)) and $\nu/(n\Delta x)^2$, respectively. Thus the BRT can be considered as a SMG-no-SGS running with an increased value of the diffusion coefficient $\nu_{eff} = n\nu$. This effective diffusion is sufficiently large for the model to be stable.

4.1 The reference setup

The spectra of the different models for the reference setup are shown in fig. 4(a). From all three reduced models the spectrum of the RSM is the closest one to the DNS spectrum. Comparing the RSM with the BRT, it can be stated that the inclusion of the SGS correction terms improves significantly the representation of the energy cascade in the inertial range. These corrections describe a considerable backscatter of energy, because the BRT is far more dissipative. At lower wavenumbers the RSM slightly overestimates the energy compared with SMG, but in the inertial range it is the only model which can produce a continuous power-law slope not far from the one of the DNS.

From fig. 4(b) it is visible that all three models can capture the decay in the ACF from the DNS. The SMG is here the model which produces the best match, the RSM slightly overestimates the time scale (see also table 2). We also computed some higher order statistical moments for x . The lagged third order moments in the reduced models are vanishing, which is consistent with the DNS. The normalized lagged fourth order moments (kurtosis) of the reduced models are displayed in fig. 4(c). The kurtosis is a measure for deviations from Gaussianity, since for a Gaussian distribution it is equal to one. Fig. 4(c) shows small departures from Gaussianity in the DNS at all lags. All reduced models can to some extent reproduce the structure of the kurtosis. The largest differences are visible at lag zero, where the RSM underestimates and the SMG, BRT overestimate the kurtosis.

4.2 Resolution tests

In order to assess the new SGS parameterization, we performed simulations by changing some of the parameters in the reference setup. The RSM was derived in a systematic way for a particular choice of the OU parameters and one important question arising is, how universal these parameters are. The

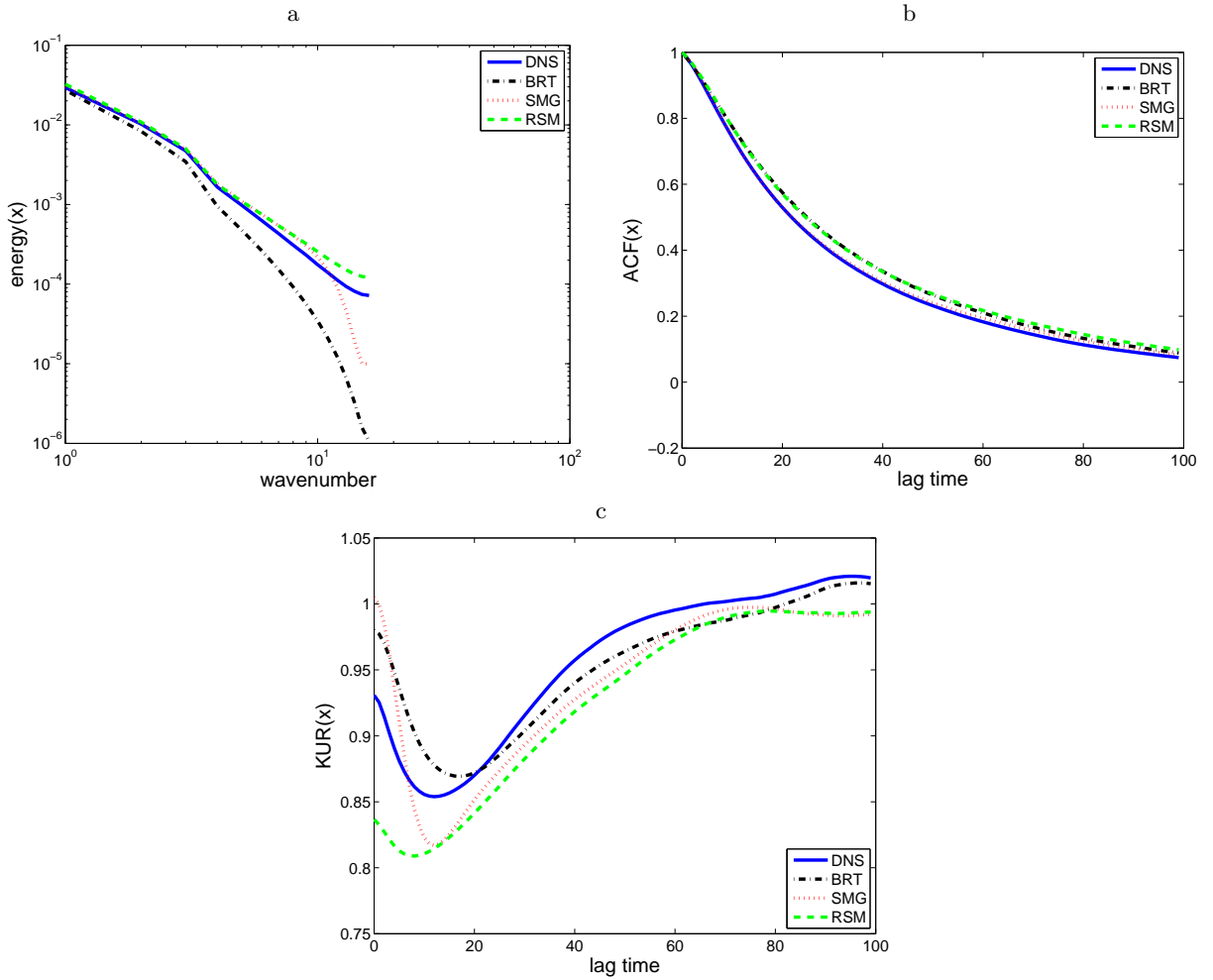


Fig. 4 Energy spectra, autocorrelation function (ACF) and kurtosis (KUR) of x for: DNS, bare truncation model (BRT), Smagorinsky SGS model (SMG) and reduced stochastic model (RSM). Reference setup.

OU closure models a term involving only SGS modes. From a turbulence viewpoint the behavior of the SGS motion remains qualitatively similar for different large-scale regimes, i.e., it extracts energy from the large-scale flow, so one might expect that the OU parameters will also be valid for different regimes in the Burgers model.

We performed simulations with the reduced models for different numbers of coarse cells, but running the RSM with the SGS parameterization for the reference setup. For a particular resolution on the fine grid, increasing (lowering) the number of coarse cells in the DNS corresponds to lowering (increasing) the number of SGS modes, which have to be parameterized in the reduced model. The number of coarse cells N_x is controlled by the averaging interval n and we have performed simulations with $n = 8, 32$ corresponding to $N_x = 64, 16$ (where $n = 16$ and $N_x = 32$ are the values from the reference setup).

The results from these runs are depicted in fig. 5 and fig. 6. The RSM model performs reasonably well in reproducing the spectrum of the full model, although the SGS model was not recomputed for the new setups. The RSM captures considerably better the inertial range than the other two reduced models do. All three models can mimic the decay of the ACF in the DNS, but for $n = 32$, the BRT and RSM overestimate the decay time scale. The large-scale structure of the kurtosis is captured by the models, however for small lags they are showing deviations qualitatively similar to those from the reference setup simulation.

In table 2 we have summarized the results for the variance and ACF from the different simulations. As a measure for the quality of describing the ACF, we considered the integrated absolute value of the

Table 2 Variance (VAR) and integrated absolute value of the ACF (Int. ACF) for various models and different averaging intervals n . The relative error wrt. the DNS value is given in brackets.

| VAR | DNS | BTR | SMG | RSM |
|----------|--------|---------------|---------------|---------------|
| n=8 | 0.0254 | 0.0231 (0.09) | 0.0255 (0.01) | 0.0266 (0.05) |
| n=16 | 0.0247 | 0.0206 (0.17) | 0.0256 (0.04) | 0.0272 (0.10) |
| n=32 | 0.0230 | 0.0166 (0.28) | 0.0247 (0.07) | 0.0268 (0.17) |
| Int. ACF | DNS | BTR | SMG | RSM |
| n=8 | 138.2 | 146.5 (0.06) | 135.4 (0.02) | 140.7 (0.02) |
| n=16 | 141.2 | 155.0 (0.10) | 145.7 (0.03) | 159.7 (0.13) |
| n=32 | 147.5 | 162.3 (0.10) | 151.6 (0.03) | 177.9 (0.21) |

Table 3 Relative error in the 4th and 6th statistical moments for various reduced models when compared with DNS. Shown are the results for different averaging intervals n .

| error 4th moment | BTR | SMG | RSM |
|------------------|------|------|------|
| n = 8 | 0.14 | 0.06 | 0.04 |
| n = 16 | 0.27 | 0.16 | 0.09 |
| n = 32 | 0.44 | 0.28 | 0.20 |
| error 6th moment | BTR | SMG | RSM |
| n = 8 | 0.16 | 0.22 | 0.01 |
| n = 16 | 0.31 | 0.41 | 0.02 |
| n = 32 | 0.52 | 0.61 | 0.14 |

ACF. From the table it is seen, that the RSM and SMG models overestimate the variance, however the SMG is more accurate. Both models are considerably better than the BTR, which is too dissipative. From the values of the integrated ACF, we can confirm the observations from the figures: all models can capture the ACF and there is a tendency for RSM and BRT to overestimate the ACF for increasing n .

In table 3 the relative errors for 4th and 6th order statistical moments at lag zero are presented; the odd moments nearly vanish and are not shown. Overall it can be stated that the RSM captures higher moments much more accurately than the SMG and BRT.

By inspecting the ACF of the x - and y -modes from the DNS, we found that the scale separation between the resolved and the SGS modes decreases for increasing n . This can explain the diminishing accuracy of the RSM for larger n (see tables 2,3), since the RSM is rigorously valid in the asymptotic limit for infinitely large scale separation. We refitted the OU closure parameters for the simulations with $n = 8, 32$ and recomputed the SGS parameterization terms in the RSM. The new correction terms improve the performance of the RSM by reducing the relative error by approximately 20% of the original value (not shown).

4.3 Forcing in the intermediate range

We addressed the question, how sensitive the results presented above are with respect to the scales forced in the model. For that purpose we excited in the random forcing function (9) modes within an intermediate wavenumber range $k = 4 - 7$ (instead of $k = 1 - 3$ from the reference setup). Some of the results are depicted in fig. 7. In the new regime the BRT completely fails to capture the statistics of the full model, whereas the RSM and SMG have still some skill. The RSM systematically overestimates the power at all scales, however in the inertial range it is closer to the DNS than the SMG (the latter model showing deviations of more than one order magnitude). The performance of the RSM and SMG reproducing the lagged 2th and 4th order moments is qualitatively similar to the one from the resolution tests discussed above. We have to stress, that for the regime presented here, the SGS parameterization in the RSM was not recomputed again.

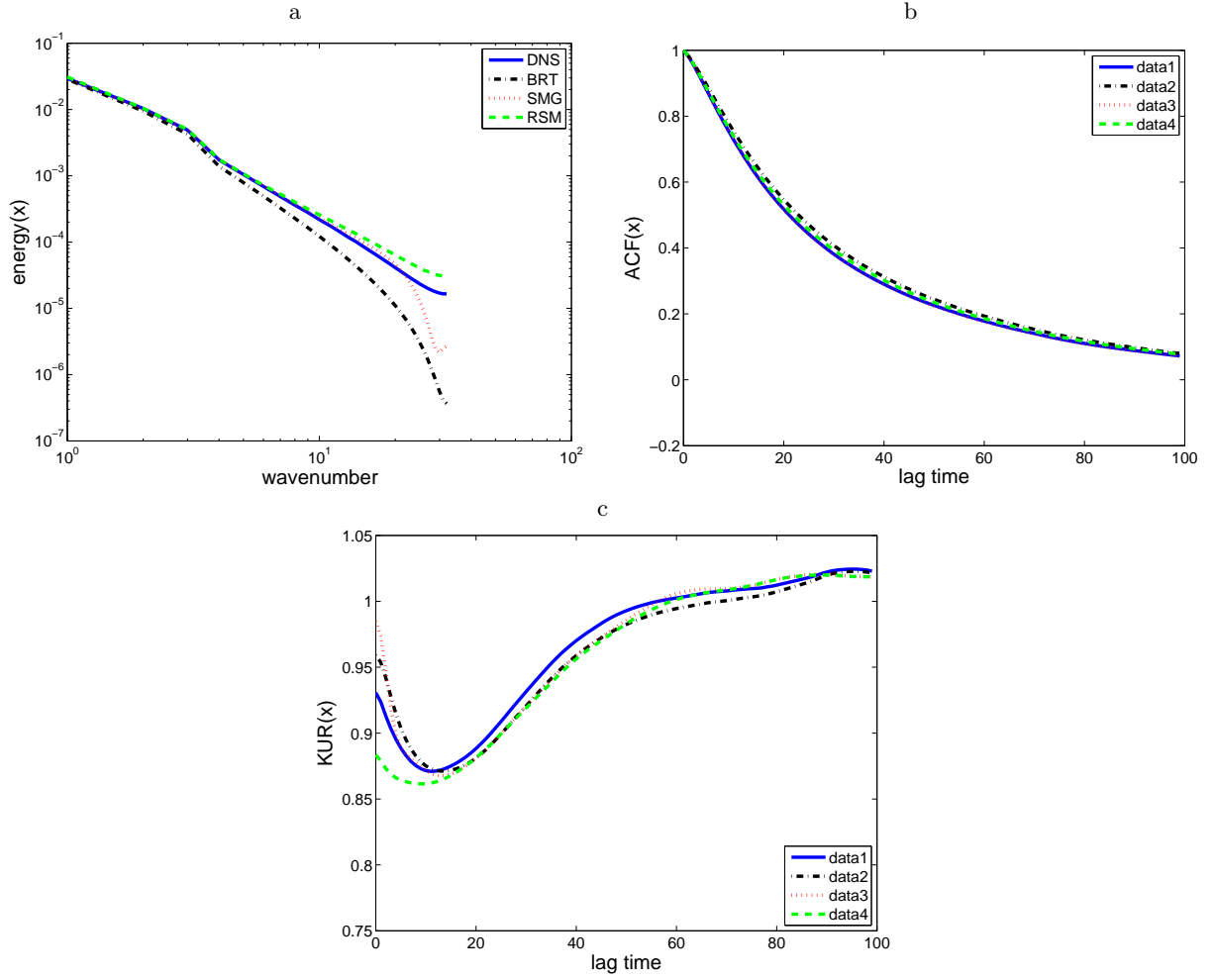


Fig. 5 As in fig. 4 but for an averaging interval $n = 8$.

4.4 Budget analysis

One advantage of the SGS parameterization derived by stochastic mode reduction is that the different SGS corrections can be attributed to different physical processes [24,25,28]. Performing a budget analysis for the different SGS correction terms, the importance of different physical interactions can be assessed in a systematic way [15,14]. In order to evaluate the contributions of various SGS corrections, we regroup these terms as

$$\begin{aligned}
 dx_i(t) = & \sum_{j,k=i-1}^{i+1} B_{ijk}^{xxx} x_j x_k dt + \sum_{j=i-1}^{i+1} L_{ij}^{xx} x_j dt + F_i^x dt + \lambda_a \left(\sum_{j=i-2}^{i+2} M_{ij}^a x_j dt + \sigma^{(1)} dW_i^{(1)} \right) \\
 & + \lambda_m \left(\sum_{j=i-2}^{i+2} M_{ij}^m x_j dt + \sum_{j,k,l=i-2}^{i+2} C_{ijkl} x_j x_k x_l dt \right) + \lambda_l \sum_{j=i-2}^{i+2} M_{ij}^l x_j dt \\
 & + \lambda_l \lambda_m \sum_{j=i-2}^{i+2} Q_{ijk} x_j x_k dt + \sigma_i^{(2)}(x, \lambda_m, \lambda_l) dW_i^{(2)},
 \end{aligned} \tag{23}$$

where $\lambda_a = \lambda_m = \lambda_l = 1$ corresponds to the complete reduced stochastic model. Interactions between the terms B^{xyy} and B^{yxy} in the full model are called additive triad interactions. They give rise to linear corrections and additive noise in the SGS parameterization. We will refer to these SGS

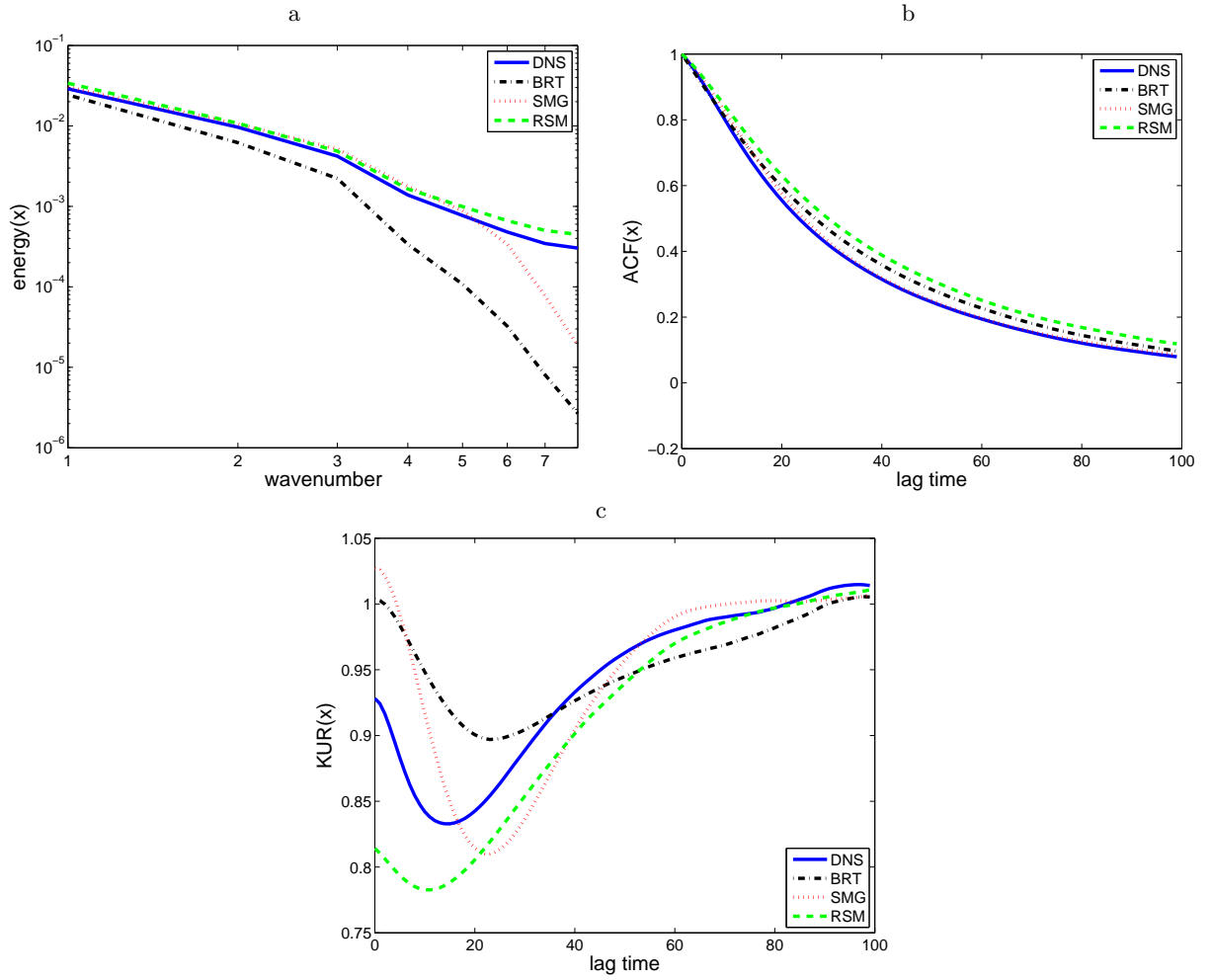


Fig. 6 As in fig. 4 but for an averaging interval $n = 32$.

terms as additive corrections and we place λ_a in front of them to denote their origin. Multiplicative SGS correction terms, indicated by λ_m , include linear and cubic terms, as well as multiplicative noise. These terms are produced by the multiplicative triad interactions between B^{xxy} and B^{yxx} . Linear correction terms and additive noise in SGS terms, which arise from interactions between L^{yx} and L^{xy} , are called augmented linearity and are denoted with λ_l . Interactions between L^{yx} and B^{xxy} , as well as between L^{xy} and B^{yxx} , create quadratic correction terms in the reduced model. We indicate these terms by the factors λ_l and λ_m in front of them, because they can arise only when both augmented linearity and multiplicative corrections are included. The explicit form of the SGS correction terms described above is given in appendix D.

It was already shown that for the reference setup the BRT model ($\lambda_a = \lambda_m = \lambda_l = 0$) can capture the ACF and to some extent also the variance of the DNS. This motivated us to perform a number of runs by systematically adding one type or a combination of two types of SGS correction terms to the BRT. All parameters in these runs were chosen as in the reference setup.

Looking at the ACFs, presented in fig. 8(a), one can say that the additive and multiplicative corrections alone can describe the decay time scale. However, the inclusion of the augmented linearity in a model with additive corrections spoils the performance, see fig. 8(b). Only models which include either multiplicative and additive corrections or multiplicative corrections and augmented linearity can still capture the ACF. Figure 9 demonstrates that the multiplicative corrections and augmented linearity are the most dominant terms in SGS parameterization. The model containing these corrections is able perfectly to match the spectrum of the full RSM, the error in the variance is only about 1%.

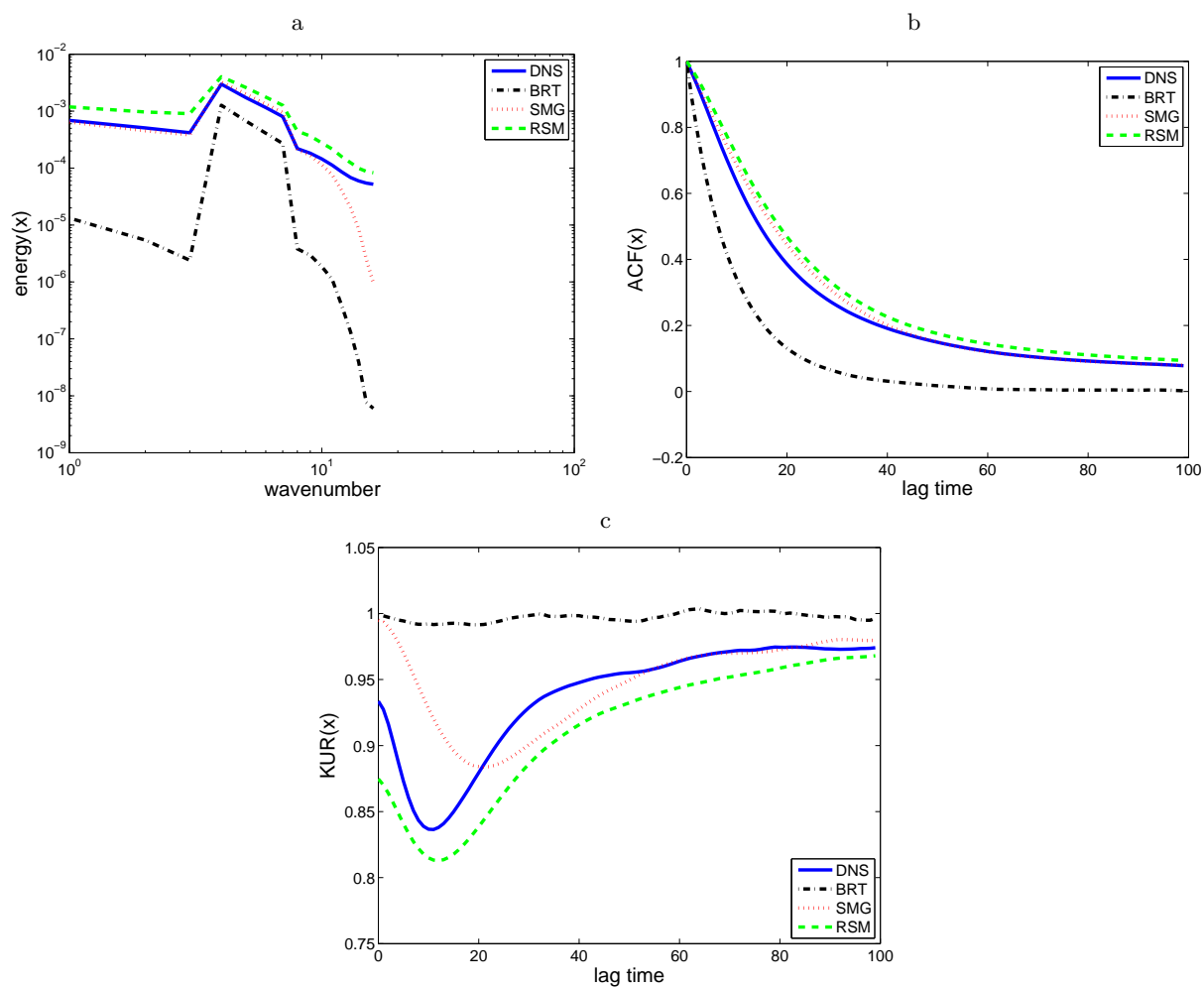


Fig. 7 As in fig. 4 but with a forcing at modes $k = 4 - 7$.

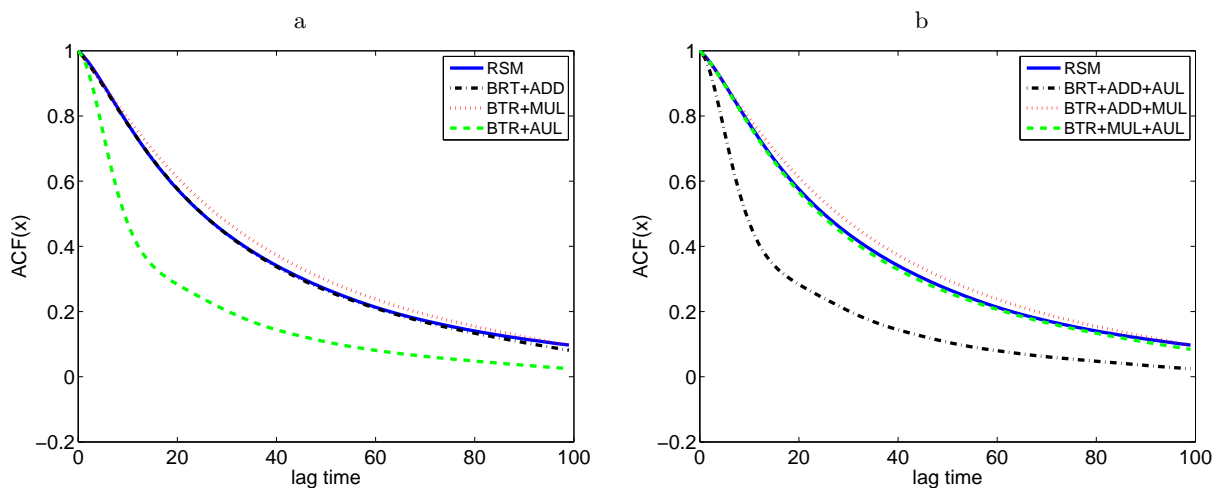


Fig. 8 Autocorrelation function of x computed with the reduced stochastic model including different correction terms: bare truncation (BRT), additive triads (ADD), multiplicative triads (MUL), augmented linearity (AUL) and full reduced stochastic model (RSM).

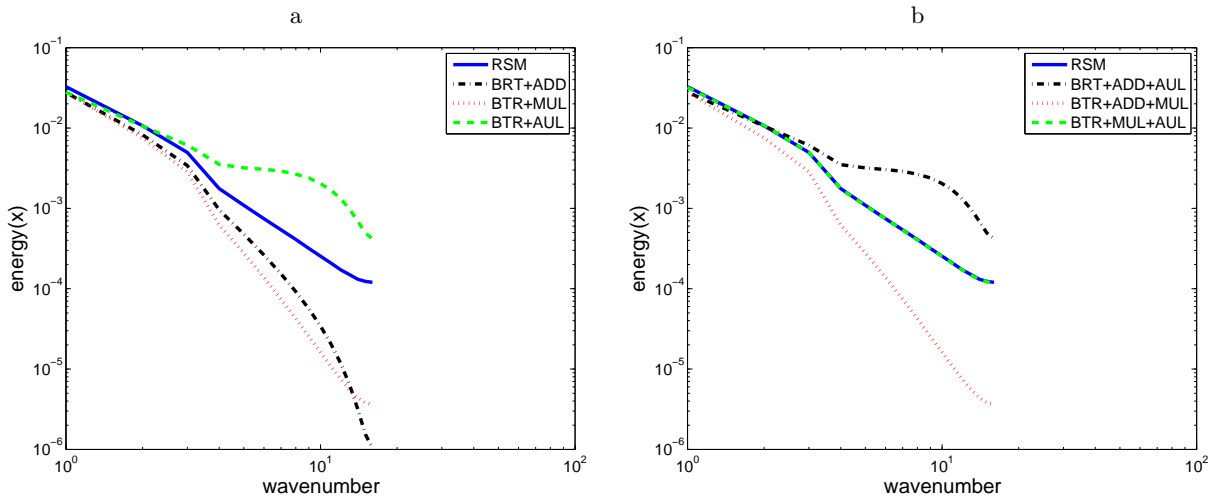


Fig. 9 Spectra of x computed with the reduced stochastic model including different correction terms: bare truncation (BRT), additive triads (ADD), multiplicative triads (MUL), augmented linearity (AUL) and full reduced stochastic model (RSM). The curves for the BRT+MUL+AUL model and RSM overlap

5 Summary and conclusions

In the present paper we presented an approach for the construction of a stochastic SGS parameterization inferred from the finite-difference discretization of the model equations. The SGS model includes linear and nonlinear deterministic corrections, as well as additive and multiplicative noise terms. The particular form of the SGS correction terms is derived in a systematic way from the discretized model by applying a stochastic mode reduction. This procedure is a generalization of the MTV mode reduction strategy [24,25,28], in that it allows for oscillations in the closure assumptions. The SGS parameterization is rigorously valid in the limit of infinite scale separation between resolved and unresolved modes. The SGS model does not require any regression fitting of the resolved modes; it is based on the assumption, that the terms involving only SGS modes in the equations for the SGS modes can be modeled as a low-dimensional Ornstein-Uhlenbeck (OU) process. The parameters of this process were estimated from a time series of the SGS modes.

The new method was applied for the construction of a reduced stochastic model (RSM) with SGS closure for the forced Burgers equation. The performance of the RSM was assessed by comparing its statistics with those from two other reduced models: a reduced model without a SGS parameterization (BRT) and a model with Smagorinsky SGS closure (SMG) (the BRT differs from the SMG with switched off SGS closure). Inspecting the energy spectra of the models, we found that the RSM is the only model, that can produce a continuous power-law slope in the inertial range, close to the DNS spectrum. All reduced models are able to capture the structure of the autocorrelation function (ACF) and lagged kurtosis. The SMG is performing best in simulating the variance and the decay time scale of the ACF, however, the difference to the RSM is moderate. For the standard setup the relative errors in the variance/ACF are: 4/3% for SMG and 10/13% for RSM, respectively. The RSM shows the lowest errors in the simulation of the fourth and sixth moments: 9/2% for RSM and 16/41% for SMG, respectively.

The parameters of the low-dimensional OU process entering in the SGS parameterization for the RSM were estimated from data for a particular model setup. In practical applications when different model regimes are considered, it is desirable to have a universal SGS model and to avoid the necessity of recomputing new SGS correction terms. This motivated us to perform simulations with the RSM by varying separately the number of resolved modes and the forcing range, but using the SGS model from the standard setup. The RSM was still able to capture main statistical properties of the DNS. Its performance and the performance of the other two reduced models were qualitatively similar to the results from the standard setup.

We performed a budget analysis for the RSM. This analysis suggests that the dominant SGS terms are the augmented linearity and multiplicative corrections. It was shown further, that a RSM without

the additive SGS corrections is almost perfectly able to reproduce the behavior of the full RSM. A study with the inviscid Burgers equation, which will be presented somewhere else [13], shows that there the additive SGS corrections are the most important. These corrections result from nonlinear interactions between SGS modes. Interactions of this type are the most numerous, which explains the dominance of the additive corrections in the inviscid model. However, in the dissipative case, considered here, the magnitude of the SGS modes is considerably smaller than the one of the resolved modes. Thus terms linear in the SGS modes become important; such terms describe linear interactions with SGS modes or interactions between resolved and SGS modes. They give rise to multiplicative corrections and augmented linearity in the RSM. These terms were also found to be the dominant corrections in a EOF-based reduced quasigeostrophic model [14].

One possible further extension of the current approach is to incorporate additional information on the SGS self-interactions in the closure assumption: e.g. by increasing the dimensionality of the OU process. In the present model the OU process couples only SGS modes inside a particular coarse cells but it can be easily modified to take into account interactions between SGS modes in neighboring coarse cells. Some preliminary work in this direction gave promising results. In the future the authors plan to apply the present approach to a simplified atmospheric model. The application to turbulence parameterization might be another interesting line of research to follow.

Acknowledgements We are thankful to Andrew Majda and Christian Franzke for useful discussions. We thank Illia Horenko and the second anonymous reviewer for their comments and suggestions which helped to improve the manuscript. U. A. thanks Deutsche Forschungsgemeinschaft for partial support through the MetStröm Priority Research Program (SPP 1276) and through Grant Ac71/4-1; I. T. thanks the NSF for partial support through grants DMS-0713793 and DMS-1109582.

A Tensor notation

The quadratic interaction terms from (10), (11) are given by

$$\sum_{j \in s_x} \sum_{k \in s_x} B_{ijk}^{xxx} x_j x_k = -\frac{1}{6n\Delta x} (x_{i+1}^2 + x_i x_{i+1} - x_i x_{i-1} - x_{i-1}^2), \quad (24)$$

$$\sum_{j \in s} \sum_{k \in s} B_{ijk}^{yyy} y_j y_k = -\frac{1}{6n\Delta x} (y_{ir}^2 + y_{ir} y_{ir+1} + y_{ir+1}^2 - y_{il}^2 - y_{il} y_{il-1} - y_{il-1}^2), \quad (25)$$

$$\sum_{j \in s_x} \sum_{k \in s} B_{ijk}^{xxy} x_j y_k = -\frac{1}{6n\Delta x} (2x_i y_{ir} + x_i y_{ir+1} + x_{i+1} y_{ir} + 2x_{i+1} y_{ir+1} - 2x_i y_{il} - x_i y_{il-1} - x_{i-1} y_{il} - 2x_{i-1} y_{il-1}), \quad (26)$$

$$\sum_{j \in s_x} \sum_{k \in s} B_{ijk}^{yxy} x_j y_k = -\frac{1}{6\Delta x} \left(2x_{[\frac{i+1}{n}]} y_{i+1} + x_{[\frac{i+1}{n}]} y_i + x_{[\frac{i}{n}]} y_{i+1} - 2x_{[\frac{i-1}{n}]} y_{i-1} - x_{[\frac{i-1}{n}]} y_i - x_{[\frac{i}{n}]} y_{i-1} \right) \quad (27)$$

$$\begin{aligned} & + \frac{1}{6n\Delta x} (2x_{[\frac{i}{n}]} y_{ir} + x_{[\frac{i}{n}]} y_{ir+1} + x_{[\frac{i}{n}]+1} y_{ir} + 2x_{[\frac{i}{n}]+1} y_{ir+1} \\ & - 2x_{[\frac{i}{n}-1]} y_{il-1} - x_{[\frac{i}{n}]} y_{il-1} - x_{[\frac{i}{n}-1]} y_{il} - 2x_{[\frac{i}{n}]} y_{il}), \\ \sum_{j \in s_x} \sum_{k \in s_x} B_{ijk}^{yxx} x_j x_k & = -\frac{1}{6\Delta x} \left(x_{[\frac{i+1}{n}]}^2 + x_{[\frac{i}{n}]} x_{[\frac{i+1}{n}]} - x_{[\frac{i-1}{n}]}^2 - x_{[\frac{i}{n}]} x_{[\frac{i-1}{n}]} \right) \\ & + \frac{1}{6n\Delta x} \left(x_{[\frac{i}{n}]+1}^2 + x_{[\frac{i}{n}]} x_{[\frac{i}{n}]+1} - x_{[\frac{i}{n}-1]}^2 - x_{[\frac{i}{n}]} x_{[\frac{i}{n}-1]} \right). \end{aligned} \quad (28)$$

In the equations above the index $[i/n]$ denotes the coarse cell corresponding to the fine cell i and the index $ir(il)$ marks the fine cell at the right (left) boundary of a coarse cell

$$ir = \begin{cases} ([\frac{i}{n}] + 1)n - 1 & \text{if } i \in s \\ (i + 1)n - 1 & \text{if } i \in s_x \end{cases}, \quad (29)$$

$$il = \begin{cases} ([\frac{i}{n}])n - 1 & \text{if } i \in s \\ in - 1 & \text{if } i \in s_x \end{cases}. \quad (30)$$

The linear interaction terms from (10), (11) are given by

$$\sum_{j \in s_x} L_{ij}^{xx} x_j = \frac{\nu}{n\Delta x^2} (x_{i+1} - 2x_i + x_{i-1}), \quad (31)$$

$$\sum_{j \in s} L_{ij}^{xy} y_j = \frac{\nu}{n\Delta x^2} (y_{ir+1} - y_{ir} - y_{il} + y_{il-1}), \quad (32)$$

$$\sum_{j \in s_x} L_{ij}^{yx} x_j = \frac{\nu}{\Delta x^2} (x_{[\frac{i+1}{n}] + 1} + x_{[\frac{i-1}{n}] - 1} - 2x_{[\frac{i}{n}]} - \frac{\nu}{n\Delta x^2} (x_{[\frac{i}{n}] + 1} - 2x_{[\frac{i}{n}]} + x_{[\frac{i}{n}] - 1})), \quad (33)$$

$$\sum_{j \in s} L_{ij}^{yx} y_j = \frac{\nu}{\Delta x^2} (y_{i+1} - 2y_i + y_{i-1}) - \frac{\nu}{n\Delta x^2} (y_{ir+1} - y_{ir} + y_{il} - y_{il-1}). \quad (34)$$

B Fourier Transformation

We introduce the following notation for the y -modes

$$y_i \rightarrow \hat{y}_j^i, \quad (35)$$

where to each fine cell index i ($i \in s$) we can assign a corresponding coarse cell \hat{i} and an index $j \in \{0, \dots, n-1\}$, denoting the position of the fine cell inside cell \hat{i} . The Fourier transformation of the modes inside a coarse grid cell can be written as

$$\hat{y}_k^{\hat{i}} = \frac{1}{n} \sum_{j=0}^{n-1} e^{-i\frac{2\pi}{n}jk} \hat{y}_j^{\hat{i}}, \quad (36)$$

$$\hat{y}_j^{\hat{i}} = \sum_{k=1}^{n-1} e^{i\frac{2\pi}{n}jk} \hat{y}_k^{\hat{i}}. \quad (37)$$

Here $\iota = \sqrt{-1}$ and k in (37) takes values only between 1 and $n-1$, because we skip the mode $\hat{y}_0^{\hat{i}}$, which is always zero due to the definition of the y -variables. Thus by transforming $\hat{y}_j^{\hat{i}}$, $j = 0, \dots, n-1$ according to (37), we end with a complex sequence of $n-1$ elements: $\hat{y}_k^{\hat{i}}$, $k = 1, \dots, n-1$. Since $\hat{y}_j^{\hat{i}}$ is real, only $n-1$ real numbers, which we denote with $\hat{y}_j^{\hat{i}}$, are required to represent the complex sequence $\hat{y}_k^{\hat{i}}$ (instead of $2(n-1)$). From (36),(37) we can determine real matrices \bar{T}, \bar{R} ($\bar{T} \in \mathbb{R}^{n-1, n}$, $\bar{R} \in \mathbb{R}^{n, n-1}$) satisfying

$$\hat{y}_k^{\hat{i}} = \bar{T}_{kj} \hat{y}_j^{\hat{i}}, \quad (38)$$

$$\hat{y}_j^{\hat{i}} = \bar{R}_{jk} \hat{y}_k^{\hat{i}}. \quad (39)$$

Finally, the matrix $\hat{T}(\hat{R})$ from (12), (13) is constructed by replicating N_x times $\bar{T}(\bar{R})$ on the diagonal of a $N_z \times N(N \times N_z)$ zero matrix.

C Stochastic mode reduction for an oscillating Ornstein-Uhlenbeck process

The drift matrix Γ in the Ornstein-Uhlenbeck process (14) has eigenvalues $\lambda_j = -\hat{\gamma}_j + i\hat{\omega}_j$ ($\hat{\gamma}_j, \hat{\omega}_j \in \mathbb{R}$), with $\hat{\gamma}_j > 0$ for a stable OU process. The complex eigenvalues ($\hat{\omega}_j \neq 0$) form complex-conjugate pairs λ_j, λ_k ($k \neq j$) such that $\lambda_j^* = \lambda_k$. If Γ can be diagonalized, we can find a real Jordan canonical form $\Gamma = U\Lambda U^{-1}$: here $\Lambda, U, U^{-1} \in \mathbb{R}^{N_z, N_z}$ and Λ has a block diagonal structure with either blocks of 2×2 matrices

$$\begin{pmatrix} -\hat{\gamma}_j & \hat{\omega}_j \\ -\hat{\omega}_j & -\hat{\gamma}_j \end{pmatrix} \quad (40)$$

or with the scalar $\hat{\gamma}_j$ on the main diagonal and zero everywhere else. The classical MTV strategy [24] for mode elimination can be applied to the OU-DNS model (20), (21) if Λ in (21) is a real diagonal matrix. This corresponds to the case when the drift matrix Γ in the Ornstein-Uhlenbeck closure has real eigenvalues only. However, our estimate for Γ has complex eigenvalues as well (we did not succeed in finding a matrix Γ with real eigenvalues only such that the corresponding OU-DNS model performs well). This requires an extension of the mode reduction strategy for matrices Λ with the block structure as in (40).

Next, we introduce a vector $\boldsymbol{\gamma}$ containing the absolute values of the real parts of the eigenvalues of Γ , the i th component ($i \in s_z$) is defined by

$$\gamma_i = |A_{ii}|. \quad (41)$$

In addition, we introduce a vector $\boldsymbol{\omega}$ with the imaginary parts of the eigenvalues as components

$$\omega_i = \begin{cases} A_{ii+1}, & \text{if } A_{ii+1} \neq 0 \\ A_{ii-1}, & \text{if } A_{ii-1} \neq 0 \\ 0, & \text{else} \end{cases}. \quad (42)$$

Later on we will make use of a vector $\boldsymbol{\sigma}$ which has as i th component σ_i from (21).

The Kolmogorov backward equation corresponding to the stochastic differential equations (20),(21) can be written as

$$\frac{\partial \rho}{\partial t} = \frac{1}{\varepsilon^2} \mathcal{L}_1 \rho + \frac{1}{\varepsilon} \mathcal{L}_2 \rho + \mathcal{L}_3 \rho, \quad (43)$$

where we have adopted the notation from [24] (see eq.(4.23) there) but with the modified operator \mathcal{L}_1

$$\mathcal{L}_1 = - \left(\sum_{i,j} A_{ij} z_j \frac{\partial}{\partial z_i} + \frac{\sigma_i^2}{2} \frac{\partial^2}{\partial z_i^2} \right). \quad (44)$$

The order ε^{-2} equation from (43) defines an auxiliary stochastic process with an invariant probability distribution π given by

$$\pi(\mathbf{z}) = C_0 \exp\left\{-\frac{1}{2} \mathbf{z}^T D^{-1} \mathbf{z}\right\}. \quad (45)$$

Here \mathbf{z} is a vector with z_i as i th component, C_0 is a normalization constant and the matrix D is given by

$$D_{ij} = \begin{cases} \sigma_i^2 \frac{2\gamma_i^2 + \omega_i^2}{4\gamma_i(\gamma_i^2 + \omega_i^2)} + \sigma_{i-1}^2 \frac{\omega_i^2}{4\gamma_i(\gamma_i^2 + \omega_i^2)}, & j = i, \omega_i > 0 \\ \sigma_i^2 \frac{2\gamma_i^2 + \omega_i^2}{4\gamma_i(\gamma_i^2 + \omega_i^2)} + \sigma_{i+1}^2 \frac{\omega_i^2}{4\gamma_i(\gamma_i^2 + \omega_i^2)}, & j = i, \omega_i < 0 \\ \frac{\sigma_i^2}{2\gamma_i}, & j = i, \omega_i = 0 \\ (\sigma_i^2 - \sigma_{i-1}^2) \frac{\omega_i^2}{4(\gamma_i^2 + \omega_i^2)}, & j = i + 1, \omega_i > 0 \\ (\sigma_{i+1}^2 - \sigma_i^2) \frac{\omega_i^2}{4(\gamma_i^2 + \omega_i^2)}, & j = i - 1, \omega_i < 0 \\ 0, & \text{else} \end{cases}. \quad (46)$$

However, in order to satisfy the solvability condition ([24], (4.29)), we require for the components in $\boldsymbol{\sigma}$

$$\sigma_i = \sigma_{i-1} = \frac{1}{2}(\sigma_i + \sigma_{i-1}), \quad \text{if } \omega_i > 0. \quad (47)$$

Thus (46) reduces to

$$D_{ij} = \begin{cases} \frac{\sigma_i^2}{2\gamma_i}, & j = i \\ 0, & \text{else} \end{cases}. \quad (48)$$

The stochastic mode reduction procedure requires the computation of the inverse operator \mathcal{L}_1^{-1} . In particular, it could be shown that

$$\mathcal{L}_1^{-1} z_i = \int_0^\infty d\tau z_i(\tau) = L_{ij} z_j, \quad (49)$$

$$\mathcal{L}_1^{-1} z_i z_j = \int_0^\infty d\tau z_i(\tau) z_j(\tau) = P_{ijkl} z_k z_l, \quad (50)$$

where

$$L_{ij} = \begin{cases} \frac{\gamma_i}{\gamma_i^2 + \omega_i^2}, & j = i \\ \frac{\omega_i}{\gamma_i^2 + \omega_i^2}, & j = i + \text{sgn}(\omega_i), \omega_i \neq 0 \\ 0, & \text{else} \end{cases} \quad (51)$$

$$P_{ijkl} = \begin{cases} \kappa_{ij}(\gamma_i + \gamma_j) \left((\gamma_i + \gamma_j)^2 + \omega_i^2 + \omega_j^2 \right), & k = i, l = j \\ \kappa_{ij}\omega_i \left((\gamma_i + \gamma_j)^2 + \omega_i^2 - \omega_j^2 \right), & k = i + \text{sgn}(\omega_i), l = j, \omega_i \neq 0 \\ \kappa_{ij}\omega_j \left((\gamma_i + \gamma_j)^2 + \omega_j^2 - \omega_i^2 \right), & k = i, l = j + \text{sgn}(\omega_j), \omega_j \neq 0 \\ 2\kappa_{ij}\omega_i\omega_j(\gamma_i + \gamma_j), & k = i + \text{sgn}(\omega_i), l = j + \text{sgn}(\omega_j), \omega_{i,j} \neq 0 \\ 0, & \text{else} \end{cases} \quad (52)$$

$$\kappa_{ij} = \left\{ \left((\gamma_i + \gamma_j)^2 + (\omega_i - \omega_j)^2 \right) \left((\gamma_i + \gamma_j)^2 + (\omega_i + \omega_j)^2 \right) \right\}^{-1}, \quad (53)$$

and we have used the sign function

$$\text{sgn}(x) = \begin{cases} 1 & x \geq 0 \\ -1 & x < 0 \end{cases}. \quad (54)$$

With the expressions from (48),(49) and (50) it is straightforward to carry out the mode reduction procedure. The explicit form of the SGS correction terms is given in the next section.

D Coefficients in the reduced stochastic model

As discussed in section 4.4, the different SGS correction terms in the RSM can be regrouped depending on their origin. The resulting deterministic terms are given by

$$\sum_{j=i-2}^{i+2} M_{ij}^a x_j = \sum_j \sum_{m,n,p,q} 2B_{imn}^{xzz} B_{pjq}^{zxx} P_{mnpq} D_{qq} x_j, \quad (55)$$

$$\sum_{j=i-2}^{i+2} M_{ij}^m x_j = \sum_j \sum_{k,l,n} B_{iln}^{xzx} B_{ljk}^{xzx} x_j L_{nk} D_{kk}, \quad (56)$$

$$\sum_{j,k,l=i-2}^{i+2} C_{ijkl} x_j x_k x_l = \sum_{j,k,l} \sum_{p,q} B_{ilp}^{xzx} B_{qjk}^{zxx} L_{pq} x_j x_k x_l, \quad (57)$$

$$\sum_{j=i-2}^{i+2} M_{ij}^l x_j = \sum_j \sum_{l,m} L_{im}^{xz} L_{lj}^{zx} x_j L_{ml}, \quad (58)$$

$$\sum_{j=i-2}^{i+2} Q_{ijk} x_j x_k = \sum_{j,k} \sum_{l,n} \left(B_{ikn}^{xzx} L_{lj}^{zx} L_{nl} x_j x_k + L_{in}^{xz} B_{ljk}^{zxx} L_{nl} x_j x_k \right). \quad (59)$$

For the construction of the stochastic correction terms we neglected the correlations of noise terms in different cells, i.e., we assume a diagonal diffusion matrix. The resulting noise terms in the RSM are given by

$$\sigma^{(1)} = 2 \sqrt{\sum_{j,k,m,n} B_{ijk}^{xzz} B_{imn}^{xzz} P_{mnjk} D_{jj} D_{kk}}, \quad (60)$$

$$\sigma_i^{(2)}(x, \lambda_m, \lambda_l) = \sqrt{2 \sum_{k=(i-1)n}^{(i+2)n-1} \left(\lambda_l \sum_p L_{ip}^{xz} \hat{L}_{pk} + \lambda_m \sum_{j=i-1}^{i+1} \sum_p B_{ijp}^{xzx} x_j \hat{L}_{pk} \right)^2}. \quad (61)$$

Here the matrix \hat{L} is defined by the decomposition

$$L^s = \hat{L} \hat{L}^T, \quad (62)$$

where L^s is the symmetric part of the matrix $\bar{L} = LD : L^s = \frac{1}{2}(\bar{L} + \bar{L}^T)$.

For the explicit computation of the SGS correction terms we utilized the MATLAB symbolic toolbox. We obtain the following expressions (rounded to three decimal places)

$$\sum_{j=i-2}^{i+2} M_{ij}^a x_j = \frac{a^2}{n} (.203 \cdot 10^{-5} x_{i-2} + .854 \cdot 10^{-4} x_{i-1} - .183 \cdot 10^{-3} x_i + .939 \cdot 10^{-4} x_{i+1} + .180 \cdot 10^{-5} x_{i+2}) \quad (63)$$

$$\sum_{j=i-2}^{i+2} M_{ij}^m x_j = \frac{a^2}{n^2} (-.108 \cdot 10^{-3} x_{i-2} - .149 \cdot 10^{-3} x_{i-1} + .596 \cdot 10^{-3} x_i - .230 \cdot 10^{-3} x_{i+1} - .108 \cdot 10^{-3} x_{i+2}) \quad (64)$$

$$\sum_{j,k,l=i-2}^{i+2} C_{ijkl} x_j x_k x_l = \frac{a^2}{n} (-.404 x_{i-2}^2 x_{i-1} - .404 x_{i-2} x_{i-1}^2 + 5.23 x_{i-1}^3 + 3.46 x_i x_{i-1}^2 - 2.66 x_i^2 x_{i-1} + .407 x_{i+1} x_i x_{i-1} + .205 x_{i+1}^2 x_{i-1} - .202 x_{i-2}^2 x_i - .202 x_{i-1} x_{i-2} x_i - 10.5 x_i^3 - 2.64 x_{i+1} x_i^2 + 3.47 x_{i+1}^2 x_i - .205 x_{i+2} x_{i+1} x_i - .205 x_{i+2}^2 x_i + .202 x_{i-1}^2 x_{i+1} + 5.23 x_{i+1}^3 - .409 x_{i+2} x_{i+1}^2 - .409 x_{i+2}^2 x_{i+1}) \quad (65)$$

$$\sum_{j=i-2}^{i+2} M_{ij}^l x_j = \frac{b^2}{n} (-.202 x_{i-2} - 1.36 x_{i-1} + 3.13 x_i - 1.36 x_{i+1} - .205 x_{i+2}) \quad (66)$$

$$\sum_{j=i-2}^{i+2} Q_{ijk} x_j x_k = \frac{ab}{n} (.404 x_{i-2} x_{i-1} + .480 x_{i-1}^2 - 2.58 x_i x_{i-1} + .253 \cdot 10^{-2} x_{i+1} x_{i-1} + .202 x_{i-2} x_i - .697 \cdot 10^{-2} x_i^2 + 2.58 x_{i+1} x_i - .205 x_{i+2} x_i - .473 x_{i+1}^2 - .409 x_{i+2} x_{i+1} + .202 x_{i-2}^2 + .202 x_{i-2} x_{i-1} - 1.49 x_{i-1}^2 + 1.56 x_i x_{i-1} - .313 \cdot 10^{-2} x_i^2 - 1.56 x_{i+1} x_i + 1.50 x_{i+1}^2 - .205 x_{i+2} x_{i+1} - .205 x_{i+2}^2) \quad (67)$$

$$\sigma^{(1)} dW_i^{(1)} = \frac{a}{n} 1.208 \cdot 10^{-4} dW_i^{(1)} \quad (68)$$

In the equations above $a = -1/(6\Delta x)$ and $b = \nu/\Delta x^2$.

E Smagorinsky SGS model

After applying a spatial filter of width Δ_f , the Burgers equation can be written as

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} \left(\frac{u^2}{2} - \nu \frac{\partial u}{\partial x} \right) = f(x, t) - \frac{1}{2} \frac{\partial \tau}{\partial x}, \quad (69)$$

where τ denotes the SGS stress resulting from the filtering procedure. The Smagorinsky SGS model [37] relates the SGS stress to the deformation of the filtered variable u in the following way

$$\tau = -2C_s \Delta_f^2 \left| \frac{\partial u}{\partial x} \right| \frac{\partial u}{\partial x}. \quad (70)$$

Here C_s denotes the Smagorinsky constant; a typical value from the literature is $C_s = 0.2$. In the numerical simulations we choose for the filter width $\Delta_f = \Delta x$.

References

1. Achatz, U., Branstator, G.: A two-layer model with empirical linear corrections and reduced order for studies of internal climate variability. *J. Atmos. Sci.* **56**, 3140–3160 (1999)
2. Achatz, U., Opsteegh, J.D.: Primitive-Equation-Based Low-Order Models with Seasonal Cycle. Part I: Model Construction. *J. Atmos. Sci.* **60**, 465–477 (2003)
3. Achatz, U., Schmitz, G.: On the Closure Problem in the Reduction of Complex Atmospheric Models by PIPs and EOFs: A Comparison for the Case of a Two-Layer Model with Zonally Symmetric Forcing. *Journal of Atmospheric Sciences* **54**, 2452–2474 (1997). DOI 10.1175/1520-0469(1997)054<2452:OTCPIT>2.0.CO;2
4. Achatz, U., Schmitz, G., Greisiger, K.M.: Principal Interaction Patterns in Baroclinic Wave Life Cycles. *Journal of Atmospheric Sciences* **52**, 3201–3213 (1995). DOI 10.1175/1520-0469(1995)052<3201:PIPIBW>2.0.CO;2
5. Branstator, G., Haupt, S.E.: An Empirical Model of Barotropic Atmospheric Dynamics and Its Response to Tropical Forcing. *Journal of Climate* **11**, 2645–2667 (1998). DOI 10.1175/1520-0442(1998)011<2645:AEMOBA>2.0.CO;2
6. Buizza, R., Miller, M., Palmer, T.N.: Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System. *Quarterly Journal of the Royal Meteorological Society* **125**, 2887–2908 (1999). DOI 10.1002/qj.49712556006
7. Chekhlov, A., Yakhot, V.: Kolmogorov turbulence in a random-force-driven Burgers equation. *Physical Review E* **51**, 2739–+ (1995). DOI 10.1103/PhysRevE.51.R2739
8. Crommelin, D., Vanden-Eijnden, E.: Subgrid-Scale Parameterization with Conditional Markov Chains. *Journal of Atmospheric Sciences* **65**, 2661–2675 (2008). DOI 10.1175/2008JAS2566.1
9. Culina, J., Kravtsov, S., Monahan, A.H.: Stochastic Parameterization Schemes for Use in Realistic Climate Models. *Journal of Atmospheric Sciences* **68**, 284–299 (2011). DOI 10.1175/2010JAS3509.1
10. Das, A., Moser, R.D.: Optimal large-eddy simulation of forced Burgers equation. *Physics of Fluids* **14**, 4344–4351 (2002). DOI 10.1063/1.1516212
11. DelSole, T.: Stochastic models of quasigeostrophic turbulence. *Surveys in Geophysics* (2004)
12. Dolaptchiev, S.I., Klein, R.: Planetary geostrophic equations for the atmosphere with evolution of the barotropic flow. *Dynamics of Atmospheres and Oceans* **46**, 46–61 (2009)
13. Dolaptchiev, S.I., Timofeyev, I., Achatz, U.: Subgrid-scale closure for the inviscid Burgers-Hopf equation. submitted to *Commun. Math. Sci.*
14. Franzke, C., Majda, A.J.: Low-Order Stochastic Mode Reduction for a Prototype Atmospheric GCM. *J. Atmos. Sci.* **63**, 457–479 (2006)
15. Franzke, C., Majda, A.J., Vanden-Eijnden, E.: Low-Order Stochastic Mode Reduction for a Realistic Barotropic Model Climate. *J. Atmos. Sci.* **62**, 1722–1745 (2005)
16. Haidvogel, Beckmann: *Numerical Ocean Circulation Modeling*. Imperial College Press, London (1999)
17. Hasselmann, K.: Stochastic climate models, Part I, Theory. *Tellus* **28**, 473–+ (1976)
18. Horenko, I.: On the Identification of Nonstationary Factor Models and Their Application to Atmospheric Data Analysis. *Journal of Atmospheric Sciences* **67**, 1559–1574 (2010). DOI 10.1175/2010JAS3271.1
19. Horenko, I., Dolaptchiev, S.I., Eliseev, A.V., Mokhov, I.I., Klein, R.: Metastable Decomposition of High-Dimensional Meteorological Data with Gaps. *Journal of Atmospheric Sciences* **65**, 3479–+ (2008). DOI 10.1175/2008JAS2754.1
20. Horenko, I., Klein, R., Dolaptchiev, S.I., Schtte, C.: Automated generation of reduced stochastic weather models I: simultaneous dimension and model reduction for time series analysis. *SIAM Journal on Multiscale Modelling and Simulation* **6**(4), 1125–1145 (2008)
21. Kwasniok, F.: The reduction of complex dynamical systems using principal interaction patterns. *PHYSICA D* **92**, 28–60 (1996)
22. Kwasniok, F.: Reduced Atmospheric Models Using Dynamically Motivated Basis Functions. *Journal of Atmospheric Sciences* **64**, 3452–+ (2007). DOI 10.1175/JAS4022.1
23. Majda, A., Klein, R.: Systematic multi-scale models for the tropics. *J. Atmos. Sci.* **60**, 393–408 (2003)
24. Majda, A., Timofeyev, I., Vanden-Eijnden, E.: A mathematical framework for stochastic climate models. *Commun. Pure Appl. Math.* **54**, 891–974 (2001)
25. Majda, A., Timofeyev, I., Vanden-Eijnden, E.: A priori tests of a stochastic mode reduction strategy. *Physica D* **170**, 206–252 (2002)
26. Majda, A.J., Khouider, B.: Stochastic and mesoscopic models for tropical convection. *Proceedings of the National Academy of Science* **99**, 1123–1128 (2002). DOI 10.1073/pnas.032663199
27. Majda, A.J., Timofeyev, I.: Statistical Mechanics for Truncations of the Burgers-Hopf Equation: A Model of Intrinsic Stochastic Behavior with Scaling. *Milan jour. math.* **70**, 39–96 (2002)
28. Majda, A.J., Timofeyev, I., Vanden-Eijnden, E.: Systematic Strategies for Stochastic Mode Reduction in Climate. *J. Atmos. Sci.* **60**, 1705–1722 (2003)
29. Majewski, D., Liermann, D., Prohl, P., Ritter, B., Buchhold, M., Hanisch, T., Paul, G., Wergen, W., Baumgardner, J.: The Operational Global Icosahedral Hexagonal Gridpoint Model GME: Description and High-Resolution Tests. *Monthly Weather Review* **130**, 319 (2002). DOI 10.1175/1520-0493(2002)130<0319:TOGIHG>2.0.CO;2
30. Orrell, D., Smith, L.: Visualising bifurcations in high dimensional systems: the spectral bifurcation diagram. *International journal of bifurcation and chaos* **13**, 3015–3027 (2003)
31. Palmer, T.N.: A nonlinear dynamical perspective on model error: A proposal for non-local stochastic-dynamic parameterization in weather and climate prediction models. *Quarterly Journal of the Royal Meteorological Society* **127**, 279–304 (2001). DOI 10.1002/qj.49712757202

32. Plant, R.S., Craig, G.C.: A Stochastic Parameterization for Deep Convection Based on Equilibrium Statistics. *Journal of Atmospheric Sciences* **65**, 87–+ (2008). DOI 10.1175/2007JAS2263.1
33. Rípodas, P., Gassmann, A., Förstner, J., Majewski, D., Giorgetta, M., Korn, P., Kornbluh, L., Wan, H., Zängl, G., Bonaventura, L., Heinze, T.: Icosahedral Shallow Water Model (ICOSWM): results of shallow water test cases and sensitivity to model parameters. *Geoscientific Model Development* **2**, 231–251 (2009)
34. Satoh, M., Matsuno, T., Tomita, H., Miura, H., Nasuno, T., Iga, S.: Nonhydrostatic icosahedral atmospheric model (NICAM) for global cloud resolving simulations. *Journal of Computational Physics* **227**, 3486–3514 (2008). DOI 10.1016/j.jcp.2007.02.006
35. Schumann, U.: Subgrid scale model for finite difference simulations of turbulent flows in plane channels and annuli. *Journal of Computational Physics* **18**, 376–404 (1975)
36. Selten, F.M.: An Efficient Description of the Dynamics of Barotropic Flow. *Journal of Atmospheric Sciences* **52**, 915–936 (1995). DOI 10.1175/1520-0469(1995)052<0915:AEDOTD>2.0.CO;2
37. Smagorinsky, J.: General Circulation Experiments with the Primitive Equations. *Monthly Weather Review* **91**, 99–165 (1963)
38. Sorensen, D., Gianola, D.: *Likelihood, Bayesian and MCMC methods in quantitative genetics*. Springer, New York (2010)
39. Winkler, C.R., Newman, M., Sardeshmukh, P.D.: A Linear Model of Wintertime Low-Frequency Variability. Part I: Formulation and Forecast Skill. *Journal of Climate* **14**, 4474–4494 (2001). DOI 10.1175/1520-0442(2001)014<4474:ALMOWL>2.0.CO;2
40. Zabusky, N.J., Kruskal, M.D.: Interaction of "Solitons" in a Collisionless Plasma and the Recurrence of Initial States. *Physical Review Letters* **15**, 240–243 (1965). DOI 10.1103/PhysRevLett.15.240