

SUBGRID-SCALE CLOSURE FOR THE INVISCID BURGERS-HOPF EQUATION *

S. I. DOLAPTCHEV [†], I. TIMOFEYEV [‡], AND U. ACHATZ [§]

Abstract. A method is presented for constructing effective stochastic models for the time-evolution of spatial averages in finite-difference discretizations of partial differential equations. This method relies on the existence of a time-scale separation in the dynamics of the spatial averages and fine-grid variables. The spatial averages, thus, are treated as the slow variables in the system and a stochastic mode reduction strategy can be applied to derive a closed form effective stochastic model. A conservative discretization of the Burgers-Hopf equation is used as particular example to illustrate the approach. An advantage over heretofore applications of stochastic mode reduction to spectrally discretized models is that the resulting closure is local and thus remains applicable even if the number of slow variables is large.

Key words. stochastic parameterization, coarse-grained modeling, stochastic mode elimination
AMS subject classifications. 60J60, 60H10, 62M99

1. Introduction

Reduced stochastic modeling in spatially extended systems has been an important research topic for many decades with atmospheric sciences being one of the major application areas. In particular, reduced models describing the behavior of leading patterns of atmospheric variability [24, 3, 4, 1, 2, 14] present a viable alternative in reducing the computational complexity of coupled atmosphere-ocean models. In the last few decades reduced modeling in spatially extended systems received an increased attention in the mathematical community as well. In particular, various techniques have been proposed to model the behavior of essential degrees of freedom in different applied settings. For instance, such techniques include derivation of reduced dynamics in multiscale stochastic differential equations (SDEs) [16, 8], optimal prediction [5, 6], reduced Markov chain modeling [19, 7, 12, 22], empirical estimation of linear reduced models [4, 13].

In this paper we apply a stochastic mode reduction strategy, proposed by Majda, Timofeyev & Vanden-Eijnden [16, 17, 21] and known as MTV, to derive an effective stochastic model for spatial averages in a finite-difference or finite-volume discretization of partial differential equations. Such discretizations are widely used in ocean modeling, regional atmosphere modeling and more recently in global atmosphere models. The coupled atmosphere-ocean system presents considerable computational challenges since the two components evolve on vastly different time-scales. The atmospheric component is much faster and requires a much smaller time-step and thus typically consumes more of the computational time. The approach presented in this paper is designed to eliminate this limitation by considering effective stochastic models for atmospheric discretizations averaged over some pre-defined spatial box. The subgrid-scale closure is local, i.e. its effect on some local average only needs information from the immediate neighborhood. Note that in former applications of stochastic mode reduction to spectrally discretized models [11, 15, 10, 18] always the

*

[†] Institut für Atmosphäre und Umwelt, Goethe-Universität Frankfurt, Altenhöferallee 1, 60438 Frankfurt am Main, Germany, (Dolapchev@iau.uni-frankfurt.de)

[‡] Department of Mathematics, University of Houston, Houston, Texas, (ilya@math.uh.edu).

[§] Institut für Atmosphäre und Umwelt, Goethe-Universität Frankfurt, Altenhöferallee 1, 60438 Frankfurt am Main, Germany, (Achatz@iau.uni-frankfurt.de)

information from all resolved modes is required. In contrast to these a SGS closure as proposed here does not become impracticable if the number of resolved modes is large. The new method bears thus the potential for application in turbulence parameterization in the context of large eddy simulations.

To construct the effective model we introduce a coarse grid and define new dynamic variables as averages of the sub-grid scales over the coarse box. The sub-grid scales are equivalent to the fine-grid variables in this context. Next, we apply the stochastic mode reduction to eliminate the subgrid-scale variables from the equation. To mimic the properties of some conservative finite-difference/finite-volume discretizations, we apply our methodology to a prototype chaotic system [20], obtained as a finite-difference discretization of the inviscid Burgers-Hopf equation. We demonstrate that the stochastic mode reduction technique can be successfully utilized to obtain an effective stochastic model for spatial averages over the coarse grid. In particular, we compare and contrast the statistical behavior of the new coarse-grained model with the one from the full model (direct numerical simulation) and from two purely empirical coarse-grained models.

The outline of this paper is as follows: in section 2 we introduce the discrete version of the inviscid Burgers model. In section 3 we discuss the stochastic mode reduction technique and present a new reduced model with stochastic subgrid-scale closure. The performance of this model is assessed in section 4 by comparing it with the full model and with two other coarse-grained models with purely empirical Ornstein-Uhlenbeck type closures. A summary and a conclusion are presented in the last section.

2. The inviscid Burgers-Hopf equation

2.1. The discrete model

We consider here the one dimensional inviscid Burgers-Hopf equation

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} \frac{u^2}{2} = 0 \quad (2.1)$$

over a periodic domain of size L . Discretizing eq. (2.1) in space, one can write

$$\frac{d}{dt} u_i + \frac{F_{i+\frac{1}{2}} - F_{i-\frac{1}{2}}}{\Delta x} = 0. \quad (2.2)$$

Where u_i denotes the values of the function u at the center of the discrete cell i ($i \in s = \{0, \dots, N-1\}$) of width $\Delta x = \frac{L}{N}$ and $F_{i+\frac{1}{2}}, F_{i-\frac{1}{2}}$ denote the fluxes at the boundaries of cell i . These fluxes have to be approximated using values from neighboring cells, we use the discretization proposed in [27]

$$F_{i+\frac{1}{2}} = \frac{1}{6} (u_{i+1}^2 + u_i u_{i+1} + u_i^2), \quad (2.3)$$

$$F_{i-\frac{1}{2}} = \frac{1}{6} (u_i^2 + u_i u_{i-1} + u_{i-1}^2). \quad (2.4)$$

The above discretization conserves total momentum, total energy and the Liouville property. We refer to [20] for an analysis of the mathematical properties of the finite-difference scheme. The numerical approximations presented in [20] are

characterized by strong chaos with ergodic behavior. This motivated us to use the discretized equation (2.2) with (2.3),(2.4) as a test model for the stochastic mode reduction procedure.

A typical problem, encountered in modeling processes involving interactions on a wide range of spatial scales, is that one cannot afford to solve the discrete model on a sufficiently small grid but only on some much coarser grid. The model equations have to be reformulated in terms of resolved and unresolved (subgrid-scale) variables. In large eddy simulation this is done by averaging the equations. Following the volume-balance procedure of [23] we average over “ n ” 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). The configuration of the coarse and fine grid cells is depicted in fig. 2.1. Let $x_i, i \in s_x$ denote the averaged value 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

$$x_i = \frac{1}{n} \sum_{k=ni}^{n(i+1)-1} u_k. \quad (2.5)$$

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

$$u_i = x_{\hat{i}} + y_i, \quad (2.6)$$

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, respectively. Applying the averaging operator (2.5) to the discretized Burgers-Hopf equation, eq. (2.2) can be written in terms of $x_{\hat{i}}$ and y_i

$$\frac{d}{dt} x_{\hat{i}} + \frac{F_{n(\hat{i}+1)-\frac{1}{2}} - F_{n\hat{i}-\frac{1}{2}}}{n\Delta x} = 0, \quad (2.7)$$

$$\frac{d}{dt} y_i + \frac{F_{i+\frac{1}{2}} - F_{i-\frac{1}{2}}}{\Delta x} - \frac{F_{n(\hat{i}+1)-\frac{1}{2}} - F_{n\hat{i}-\frac{1}{2}}}{n\Delta x} = 0. \quad (2.8)$$

The flux terms in (2.7) and (2.8) 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

$$\dot{x}_i = \sum_{j \in s_x} \sum_{k \in s_x} B_{ijk}^{xxx} x_j x_k + \sum_{j \in s_x} \sum_{k \in s} B_{ijk}^{xxy} x_j y_k + \sum_{j \in s} \sum_{k \in s} B_{ijk}^{xyy} y_j y_k, \quad i \in s_x, \quad (2.9)$$

$$\dot{y}_i = \sum_{j \in s_x} \sum_{k \in s_x} B_{ijk}^{yxx} x_j x_k + \sum_{j \in s_x} \sum_{k \in s} B_{ijk}^{yxy} x_j y_k + \sum_{j \in s} \sum_{k \in s} B_{ijk}^{yyy} y_j y_k, \quad i \in s. \quad (2.10)$$

The terms of the form $B^{\alpha\beta\gamma}$ 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. In (2.9), (2.10) the indices of the summation signs go formally over s or s_x , however, one has to keep in mind that due to the local form of the discretization the tensors with the interaction coefficients are sparse and involve only “neighboring” indices. The explicit form of the tensors can be found in appendix A.

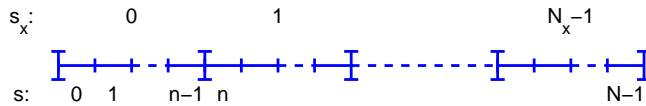


FIG. 2.1. Schematic representation of the fine grid cells (small intervals) and coarse grid cells (large intervals) introduced in section 2.1. The indexing of the fine/coarse cells is denoted by the integers corresponding to s/s_x , respectively.

2.2. Statistics in the full model

Equations (2.9), (2.10) represent another form of the discretized model (2.2),(2.3),(2.4), 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: $\Delta t=0.02$, $L=100$, $N=256$. The initial condition represents a random distribution of u , characterized by zero total momentum and a total energy of 1.716.

The time autocorrelation functions of the x - and y -modes are presented in fig. 2.2, the averaging interval n from (2.5) is set to 16. Because of the spatial homogeneity of the model, all statistical moments in the paper represent a spatial average. From the figure it is visible that the x -modes evolve on a slower time scale than the y -modes. By computing the area under the curves we estimated a decay time scale of 66 and 12 for the x - and y -modes, respectively. This implies a moderate scale separation in time of about 5.5. The scale separation is proportional to the averaging interval n : for larger values of n the x/y -modes become slower/faster.

In fig. 3.1(b) the autocorrelation function of the x -modes from DNS for larger lags is displayed. We observe oscillations in the tail with periods much larger than the estimated decay time scale. It was found that the magnitude of these oscillations depends on n as well. Long-period oscillations are visible in the tails of all DNS autocorrelation functions reported in the paper. Although characterized by small amplitudes, these oscillations appear to be an inherent property of the low-frequency dynamics of the resolved modes.

Fig. 2.3 depicts the spatial correlation pattern in the DNS. We see that the full variable u_i is uncorrelated in space. The subgrid-scale variable y_i is also to a good approximation uncorrelated, except of small negative correlation values for grid points corresponding to the same coarse cells. This is due to the constraint that the y_i 's sum to zero inside the coarse cell. From the pattern of x_i we can conclude that the correlations between different coarse cells are negligible.

3. Stochastic mode reduction

The stochastic mode reduction [16, 17, 21, 18] can be thought of as a three step procedure. In the first step, one has to impose certain assumptions on the statistical behavior of the full model (2.9), (2.10). In particular, there should exist a time-scale separation (at least moderate) between the resolved and unresolved variables. Essentially this means that we can treat the resolved and unresolved variables as the slow and fast degrees of freedom, respectively. For the discrete model we consider, this assumption is verified in sec. 2.2 and fig. 2.2. In addition, the stochastic mode reduction also requires that the term $B_{ijk}^{yyy} y_j y_k$ is the most dominant term in the equation for fast variables. The second step in the stochastic mode reduction procedure is a stochastic modeling step for the term $B_{ijk}^{yyy} y_j y_k$ (see (3.1)). This step allows explicit analytical calculations of the coefficients in the reduced model. Finally, the third step

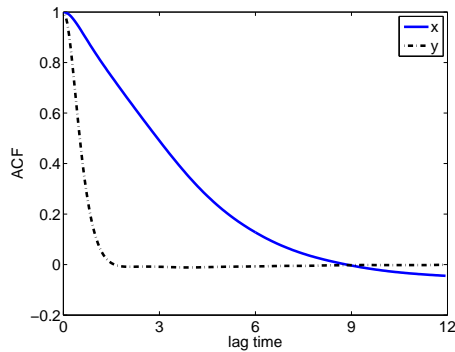


FIG. 2.2. Time autocorrelation functions for the resolved (x) and unresolved (y) modes from DNS. The averaging interval n is set to 16; the time unit on the abscissa is 20 nondimensional model time units.

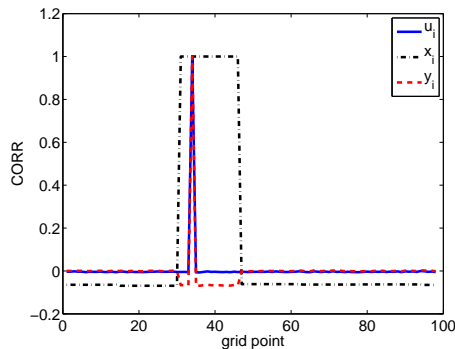


FIG. 2.3. Spatial correlation from the DNS of u_i , resolved mode x_i and unresolved mode y_i ; all for an arbitrary chosen grid point $i=36$. The averaging interval n is set to 16; grid points with $i=32, \dots, 48$ are located inside the same coarse cell.

consists of further accelerating the fast variables, which allows to compute the reduced equations in the limit of infinite time-scale separation. Typically, the full model under consideration has only a moderate separation of time scales between the fast and slow variables. Nevertheless, the mode reduction procedure has been successfully tested on many models with moderate time-scale separation which suggests that often the slow variables are not particularly sensitive to the time-scale of the fast modes. The third step can be formalized by introducing explicitly a scale separation parameter ε into the equations. This yields a modified (also called selectively accelerated [18]) model of the same dimensionality as the original full model. One then can test the existence of the limit of the modified model as $\varepsilon \rightarrow 0$.

If the self-interactions of the slow variables in the equation for slow variables (the term $B_{ijk}^{xxx} x_j x_k$) are absent in the full model, then the modified model can also be obtained by introducing a coarse-grained time $t \rightarrow \varepsilon t$. On the other hand, if these terms are non-zero then the time rescaling alone will not yield the modified equations in the proper form and an additional assumption on the behavior of these terms must be imposed. In particular, one must assume that the self-interactions are rescaled

as $B_{ijk}^{xxx} x_j x_k \rightarrow \varepsilon B_{ijk}^{xxx} x_j x_k$ under the time rescaling $t \rightarrow \varepsilon t$ (assumption A3 on p.896 in [16]). This assumption ensures that the self-interactions of the slow variables are not affected by the stochastic mode reduction. One can motivate this by the fact that the resulting reduced model is compared directly with the original full model. Moreover, the behavior of the modified model for different values of ε is also compared with the behavior of the full model. Thus, a problem can potentially arise because the full model and the modified model are compared on different time-scales, but the self-interactions of the slow variables should not be affected by this time-scale difference.

Another way of looking at the introduction of different powers of ε in front of different terms in modified model is as a modeling step in order to carry out the stochastic mode reduction strategy. Different powers of ε in front of various terms are necessary to take into account the different time scales and magnitudes of these terms. For instance, since we assume that the fast variables are primarily driven by the self-interactions of fast modes, these should be the leading terms in the equation for y -variables. Next, since the stationary averages $\langle B_{ijk}^{xxy} x_j y_k \rangle$ and $\langle B_{ijk}^{xyy} y_j y_k \rangle$ are zero, ε^{-1} is required in front of these terms in order to obtain non-trivial stochastic corrections from the stochastic mode-reduction procedure. Thus, ε^{-1} is also required in front of $B_{ijk}^{yxx} x_j x_k$ and $B_{ijk}^{yyx} y_j x_k$ terms to preserve the conservation of energy. Finally, since the full model is directly compared with the stochastically parametrized model and the reduced model (derived in the limit $\varepsilon \rightarrow 0$) the nonlinear self-interactions of the slow variables $B_{ijk}^{xxx} x_j x_k$ should not be affected by the stochastic mode-reduction.

Therefore, introducing ε can be considered as a technical step to emphasize the influence of various terms in the equations when the fast variables are accelerated even further. On the other hand, this step should also be consistent with the structure of the equations (e.g. the conservation of energy). Finally, the original full-model is compared directly with the modified model (the stochastically parametrized model in (3.2), (3.3)) and the reduced model and this is often referred to as ‘‘setting ε to one’’.

3.1. Closure assumptions

The stochastic mode reduction procedure assumes that the term $B_{ijk}^{yyy} y_j y_k$ in (2.10) can be modeled as an Ornstein-Uhlenbeck (OU) process in y

$$\sum_{j \in s} \sum_{k \in s} B_{ijk}^{yyy} y_j y_k = -\gamma y_i + \sigma \dot{W}_i. \quad (3.1)$$

Here γ and σ denote the drift and diffusion coefficient of the OU process and W_i is a Wiener process. Because of the spatial homogeneity of the model, γ and σ does not show any spatial variations (are not dependent on the index i).

By introducing (3.1) we approximate the fast self-interactions between subgrid-scale modes only in the equation for the y -modes. This is in contrast with traditional subgrid-scale parameterizations, where the terms involving subgrid-scale modes in the evolution equation for x are expressed in terms of resolved modes by assuming some particular closure. In our approach, the explicit form of this closure is derived by the MTV stochastic mode reduction procedure and is rigorously valid in the limit of infinitely large time scale separation between the resolved and subgrid-scale modes. We want to stress that the OU process from (3.1) does not couple neighboring subgrid-scale modes, since on longer time scales these modes are uncorrelated, see fig. 2.3. It is thus in line with the MTV stochastic mode reduction strategy which presupposes a diagonal OU process. As shown in [9] interactions between different subgrid-scale

modes can be incorporated in the ansatz (3.1) by choosing, instead of a scalar drift coefficient, a drift matrix.

Assuming (3.1) and properly introducing ε in (2.9), (2.10), we obtain the following modified model

$$dx_i = \left(B_{ijk}^{xxx} x_j x_k + \frac{1}{\varepsilon} B_{ijk}^{xxy} x_j y_k + \frac{1}{\varepsilon} B_{ijk}^{xyy} y_j y_k \right) dt, \quad (3.2)$$

$$dy_i = \left(\frac{1}{\varepsilon} B_{ijk}^{yxx} x_j x_k + \frac{1}{\varepsilon} B_{ijk}^{yxy} x_j y_k - \frac{1}{\varepsilon^2} \gamma y_i \right) dt + \frac{1}{\varepsilon} \sigma dW_i. \quad (3.3)$$

In the equations above and in the following paragraphs we make use of Einstein's summation convention and $\varepsilon \ll 1$ is a small parameter which is a measure of the scale separation in time between the x and y modes. We will refer to the simulation with the modified model from (3.2),(3.3) as OU-DNS.

Finally, we want to comment on the estimation of the OU parameters γ and σ . These parameters are determined using a time series of y_i from the DNS and assuming that the OU closure dominates in (3.3). In this case γ can be estimated (see e.g. [17, 21]) as the reciprocal of the area under the graph of the modulus of the autocorrelation function for y_i . Then, the diffusion coefficient σ is found from the variance of y_i (denoted with $\langle y_i y_i \rangle$) by using that in equilibrium

$$\langle y_i y_i \rangle = \frac{\sigma^2}{2\gamma}. \quad (3.4)$$

The quality of this estimate is discussed in the next section.

3.2. Numerical tests of the closure assumption

We have assessed the approximation made in (3.1) by running the OU-DNS with $\varepsilon=1$ and comparing the results with the DNS. The autocorrelation function for the x - and y -modes from the two simulations are depicted in fig. 3.1. The structure of the autocorrelation function of the subgrid-scale modes from the DNS is approximated in the OU-DNS by an exponential with similar decay rate, see fig. 3.1 (a). The OU-DNS is not able to reproduce the long-period oscillations in the autocorrelation function of x from the DNS, see fig. 3.1 (b). However, it can roughly capture the large-scale envelop of it. Table 3.1 shows that the OU closure is able to reproduce the variance of the subgrid-scale modes and captures reasonably well the variance of the resolved modes.

The reduced stochastic model introduced in section 3.3 represents the asymptotic limit $\varepsilon \rightarrow 0$ of the OU-DNS (3.2),(3.3). Such limit corresponds to infinitely large time scale separation between the x - and y -modes. In order to study the asymptotic limit, we performed numerical simulations with (3.2),(3.3) by systematically decreasing the value of ε and thus accelerating the y -modes. The corresponding autocorrelation functions of x are displayed in fig. 3.2. It is visible that as ε decreases the autocorrelation function remains close to the $\varepsilon=1$ run (note the narrow time window chosen in fig. 3.2, in order to make differences visible). Further, the deviations between $\varepsilon=0.1$ and $\varepsilon=0.5$ are smaller than those between $\varepsilon=0.5$ and $\varepsilon=1$. Similar convergence is also observed in the variance of x . The numerical results suggest that after the mode elimination procedure the resulting reduced model for x will behave similarly to the OU-DNS with $\varepsilon=0.1$ which is close to the modified model with $\varepsilon=1$.

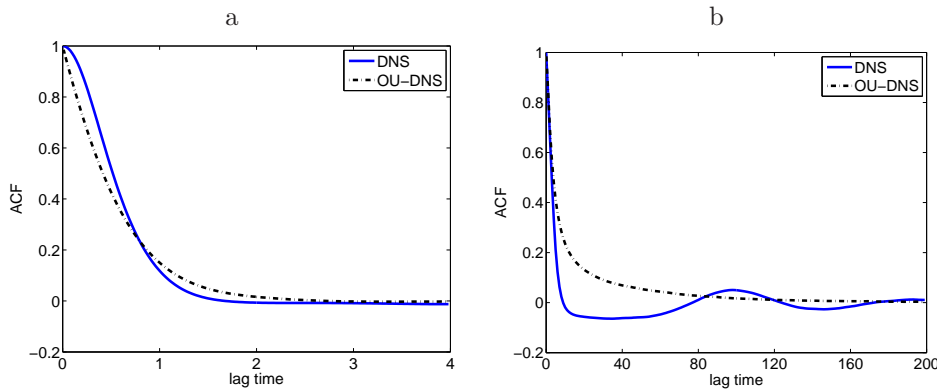


FIG. 3.1. Time autocorrelation function of unresolved modes (a) and resolved modes (b) from DNS and OU-DNS with $\varepsilon = 1$. Averaging interval and time unit on the abscissa as for fig. 2.2.

TABLE 3.1. Variance from DNS and OU-DNS.

Variance	y	x
DNS	1.2616×10^{-2}	7.8692×10^{-4}
OU-DNS $\varepsilon = 1$.	1.2619×10^{-2}	7.2649×10^{-4}
rel. error	< 0.01	0.08

3.3. The reduced stochastic model with subgrid scale closure

The stochastic mode reduction procedure allows us to eliminate the subgrid-scale modes in the OU-DNS (3.2),(3.3) and derive effective equations for the resolved modes only. The procedure is carried out by performing an asymptotic expansion in the Kolmogorov backward equation corresponding to the SDEs (3.2),(3.3), see [16]. The resulting coarse-grained model reads

$$\begin{aligned}
 dx_i(t) = & \sum_{j,k=i-1}^{i+1} B_{ijk}^{xxx} x_j x_k dt + \sum_{j=i-1}^{i+1} M_{ij} x_j dt + \sum_{j,k,l=i-2}^{i+2} C_{ijkl} x_j x_k x_l dt \\
 & + \sum_{j=i-\frac{1}{2}}^{i+\frac{1}{2}} P_{ij} dW_j^{(1)} + \sum_{j=i-\frac{1}{2}}^{i+\frac{1}{2}} D_{ij}(x) dW_j^{(2)},
 \end{aligned} \quad (3.5)$$

where $W_j^{(1)}, W_j^{(2)}$ denote independent Wiener processes. We will refer to (3.5) as the reduced stochastic model. The first term on the rhs. of (3.5) describes the bare truncation model, which is derived from the full model when all subgrid-scale modes are neglected. The other terms in (3.5) arise from the subgrid-scale model; they represent linear and cubic deterministic corrections as well as additive and multiplicative noise (last two terms). Note the local form of the subgrid-scale model as visible from the summation intervals. This is in contrast to the heretofore applications of MTV to spectrally discretized models. There the closure for any resolved mode involves all other resolved modes. It thus becomes impracticable as the number of resolved

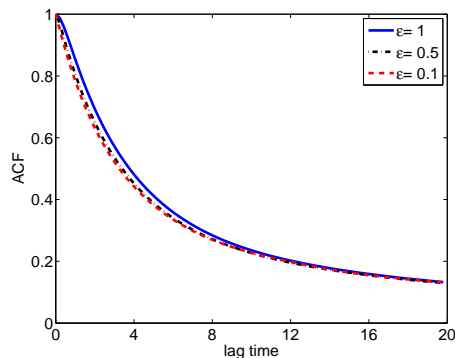


FIG. 3.2. Time autocorrelation function of the resolved modes from OU-DNS for different values of ε . Averaging interval and time unit on the abscissa as for fig. 2.2.

TABLE 4.1. Integrated normalized cross-correlation function $C_{i,i+k}(\tau)/C_{i,i}(0)$ (Int. CCF) from DNS for different lags k ; here $C_{i,i+k}(\tau) = \langle x_i(t)x_{i+k}(t+\tau) \rangle$. For $N_x = 16$ only cross-correlations up to lag 8 are required due to the spatial homogeneity.

Lag k	0	1	2	3	4
Int. CCF	1.3149	0.2000	0.0260	-0.0750	-0.1325
Lag k	5	6	7	8	
Int. CCF	-0.1721	-0.1883	-0.2231	-0.2221	

modes gets too large. The explicit form of the subgrid-scale parameterization is given in appendix B. The reduced stochastic model conserves total momentum as the full model does. The energy ($\sum_{i \in s_x} x_i^2$) is not conserved exactly, but fluctuates around a constant mean value. This, however, is also the case in the full model, where energy is transferred between the resolved and unresolved scales.

4. Results for the reduced stochastic model

In the following section we study the performance of the reduced stochastic model by comparing it with the full model, the linear-closure stochastic model and the stochastic model of multivariate OU type.

4.1. The multivariate OU model and the linear-closure stochastic model

Since the derived reduced stochastic model represents a local parameterization, it is natural to ask the question if a global closure, coupling all resolved modes, can outperform it. As a zero order global closure we consider a multivariate OU process (MVOU)

$$dx_i = \Gamma_{ij}x_j dt + \Sigma_{ij}dW_j, \quad (4.1)$$

where $\Gamma, \Sigma \in \mathbf{R}^{N_x \times N_x}$ denote the drift and diffusion matrix. Empirical models of global form, similar to (4.1), are often used in atmospheric modeling, e.g., [26, 4]. It can be shown that the lagged cross-correlation of the process in (4.1) satisfies

$$C(\tau) = e^{\Gamma\tau} C(0), \quad (4.2)$$

where $C(\tau) = \langle \vec{x}(t)\vec{x}(t+\tau)^T \rangle$ denotes the cross-correlation matrix at lag τ . By integrating (4.2) in time the following useful relation for Γ can be obtained

$$\Gamma^{-1} = - \int_0^{\infty} C(\tau) C^{-1}(0) d\tau. \quad (4.3)$$

In order to find the model parameters in (4.1), first the matrices $C(0)$ and $C(\tau)$ (with different lags τ) are estimated from a time series of the resolved variables. Then, the drift matrix Γ is determined from (4.3) such that to match the integrated lagged cross-correlation of the full model. Finally, the diffusion matrix Σ is found from the fluctuation-dissipation-relation to match the covariance of the full model

$$\Gamma C(0) + C(0)\Gamma^T = -\Sigma\Sigma^T. \quad (4.4)$$

For known Γ and Σ all statistical moments of the model from (4.1) can be computed analytically. The models of the type in (4.1) typically perform very well if the correlation functions of the slow variables are close to exponential. In this case, the estimation procedure described above provides damping and forcing parameters consistent with the large-scale statistical features of the slow variables and the multivariate OU model is likely to reproduce the statistics of the slow variables. On the other hand, if the correlation functions of the slow variables are non-exponential, then the model (4.1) is more likely to fail.

While in (4.1) the non-linear self-interactions of the slow variables are not included in the effective model similarly to [26, 4], there are also studies which emphasize the importance of these terms in reduced models [1, 2, 25, 7]. This motivated us to consider the linear-closure stochastic model (LCSM), which is based on an ad-hoc approximation of all subgrid-scale interactions in (2.9) by an OU process with drift α and diffusion β

$$dx_i = B_{ijk}^{xxx} x_j x_k dt + \alpha x_i dt + \beta dW_i. \quad (4.5)$$

The inclusion of the non-linear self-interactions in (4.5) might improve the representation of non-Gaussian features of the full model, where the multivariate OU model from (4.1) will definitely fail. We have experimented with OU closures coupling neighboring grid cells, however it turned out that the particular choice in (4.5) (no coupling) gives nearly optimal results. The diagonal choice for the damping and diffusion in (4.5) is supported by the weak statistical coupling between different slow variables. In particular, cross-correlation times (integral of the lagged cross-correlation functions $\langle x_i(t)x_j(t+\tau) \rangle$) are presented in Table 4.1. The behavior of the cross-correlations indicates that the coupling is weak for spatial lags greater than zero. Further, the estimate for Γ in (4.1) reveals a diagonally dominant damping matrix. Therefore, omitting the non-diagonal terms does not have a significant affect on the model in (4.5).

By comparing the statistical behavior of the slow variables from a simulation with and without the $B_{ijk}^{xxx} x_j x_k$ in the full model (2.9), (2.10), we found that these interaction terms are rather weak and do not significantly affect the behavior of the slow variables. In particular, the marginal equilibrium distribution and the autocorrelation function are not affected by these terms (results not shown). This can be viewed as a numerical justification for the missing ε^{-1} factor in front of the non-linear slow self-interaction terms in eq. (3.2) of the modified model (see also the discussion in the beginning of section 3).

The closure parameters α and β from (4.5) are estimated with the same procedure as the coefficients in (3.1) (see section 3.1), but now from the variance and the integrated modulus of the autocorrelation function for x_i . Since the subgrid-scale terms dominate in (2.9), the estimation procedure guarantees that the linear-closure stochastic model will capture the variance and the large-scale decay behavior in the autocorrelation function from the DNS.

Overall, based on the statistical tests for the slow variables in the full model reported above, we expect the behavior of the empirical coarse-grained models (4.1) and (4.5) to be qualitatively similar, especially if the variance is considered. Differences in the simulated autocorrelation function may emerge due to the different estimation procedures for the drift coefficients.

4.2. Statistics of the reduced models

The reduced models are integrated with the parameter setup from section 2.2. In order to study how the number of subgrid-scale modes affects the statistics of the resolved modes, we perform simulations for different averaging intervals: $n = 8, 16, 32$. Note, that the parameters from (3.1) γ and σ , entering in the reduced stochastic model, have been determined only once (for $n = 16$) and then been used in all three cases. In contrast to this, the linear-closure stochastic model and the multivariate OU model require new estimation of the closure parameters for each different case.

The autocorrelation functions of all three reduced models are shown in fig. 4.1. The linear-closure stochastic model and the reduced stochastic model are able to capture only the overall decay behavior in the autocorrelation function of the DNS, the multivariate OU model completely underestimates the decay time scale. None of the models is able to reproduce the oscillations in the autocorrelation function. We take a closer look on the differences between the reduced stochastic model and the other coarse-grained models. Figure 4.2 displays the autocorrelation functions of the models for small lags. It is visible that the reduced stochastic model can capture the slope of the DNS curve at lags close to zero for all values of n , whereas the linear-closure stochastic model and multivariate OU model fail. In fig. 4.3 a semilog plot of the absolute value of the autocorrelation functions is presented, in order to make discrepancies in the tails visible. The linear-closure stochastic model and the reduced stochastic model perform poorly for the case $n = 32$. However, the reduced stochastic model captures roughly the low-frequency envelop of the DNS for $n = 16$ and perfectly for $n = 8$. In the latter two cases the linear-closure stochastic model is damping too strongly at large lags. Again it is visible that for all different values of n the decorrelation time scale is completely underestimated by the multivariate OU model.

Figure. 4.4 shows the simulated kurtosis with the models. For a given time series $x(t)$ the kurtosis at lag s is defined as

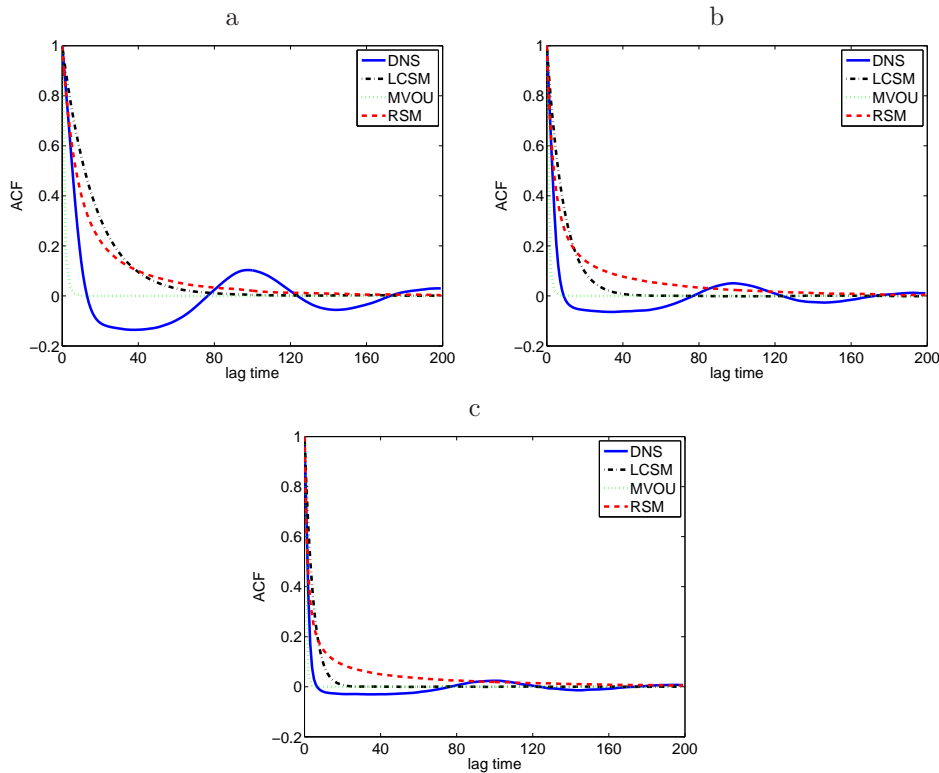


FIG. 4.1. Time autocorrelation function of the resolved modes from DNS, linear-closure stochastic model (LCSM), multivariate OU model (MVOU) and reduced stochastic model (RSM) for $n=32$ (a), $n=16$ (b) and $n=8$ (c).

$$K(s) = \frac{\langle x(t+s)^2 x^2(t) \rangle}{\langle x^2(t) \rangle^2 + 2 \langle x(t+s)x(t) \rangle^2}, \quad (4.6)$$

where $\langle \cdot \rangle$ denotes a time average. The kurtosis measures deviations from Gaussianity, since for a Gauss process it is equal to one (see the MVOU lines in fig. 4.4). Looking at the DNS profiles in fig. 4.4, we observe the same structure for all n : a minimum of kurtosis at smaller lags, whereas at larger lags the kurtosis takes a constant value close to one. It is visible from the figures that for decreasing n the magnitude of the minimum in the DNS increases indicating weak non-Gaussianity. The reduced stochastic model is able to capture the position of this minimum and to some extent its magnitude. The linear-closure stochastic model completely underestimates the minimum and shifts it to much larger lags.

Deviations from Gaussian statistics in the linear-closure stochastic model simulations are entirely due to the bare truncation part in (4.5). Thus, one can conclude that the additional subgrid-scale correction terms in the reduced stochastic model, such as cubic terms and multiplicative noise, significantly improve the representation of the kurtosis.

The probability density functions of the models are depicted in fig. 4.5. Overall,

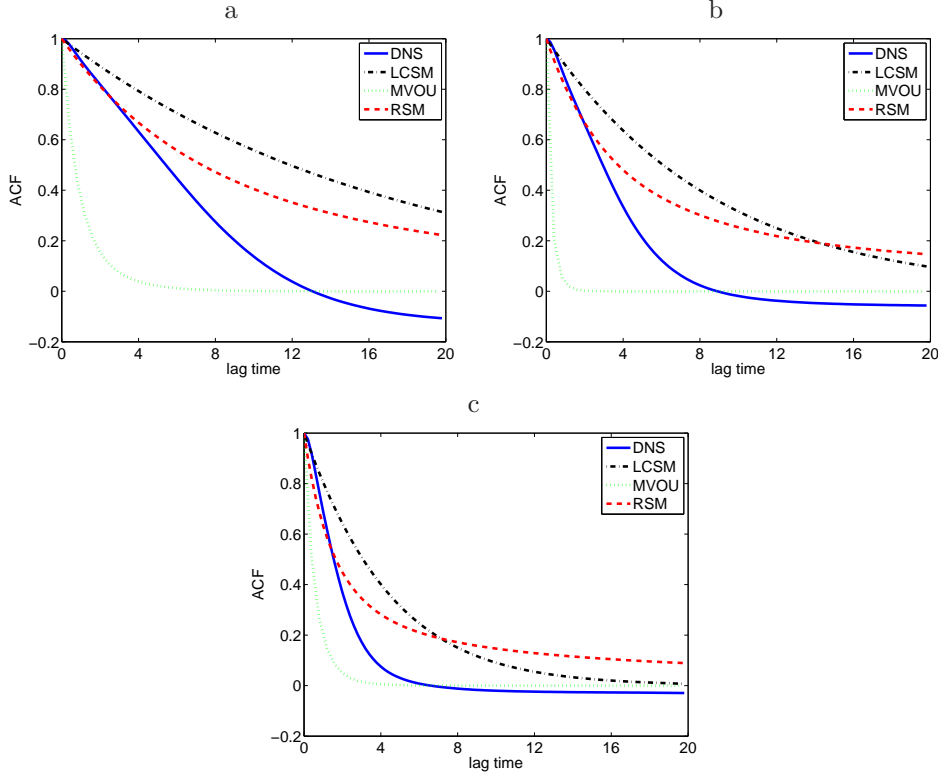


FIG. 4.2. As fig. 4.1 but for small lags.

it can be stated that the multivariate OU model performs best, indicating that the probability density distribution of the full model is close to Gaussian. The linear-closure stochastic model and the reduced stochastic model show similar precision. For an averaging interval $n = 32$ the reduced stochastic model is doing better than the linear-closure stochastic model, but its performance worsens for smaller values of n , due to the decreased scale separation between the x - and y -modes.

4.3. Budget analysis

One advantage of the subgrid-scale parameterization derived by stochastic mode reduction is that the different subgrid-scale corrections can be attributed to different physical processes [16, 17, 21, 11]. Performing a budget analysis for the different subgrid-scale correction terms the importance of different physical interactions can be assessed in a systematic way [11, 10]. In order to evaluate the contributions of various subgrid-scale corrections, we regroup the different terms in (3.5) as

$$\begin{aligned}
 dx_i(t) = & \lambda_b \sum_{j,k=i-1}^{i+1} B_{ijk}^{xxx} x_j x_k dt + \lambda_a \left(\sum_{j=i-1}^{i+1} M_{ij}^a x_j dt + \sum_{j=i-\frac{1}{2}}^{i+\frac{1}{2}} P_{ij} dW_j^{(1)} \right) \\
 & + \lambda_m \left(\sum_{j=i-1}^{i+1} M_{ij}^m x_j dt + \sum_{j,k,l=i-2}^{i+2} C_{ijkl} x_j x_k x_l dt + \sum_{j=i-\frac{1}{2}}^{i+\frac{1}{2}} D_{ij}(x) dW_j^{(2)} \right), \quad (4.7)
 \end{aligned}$$

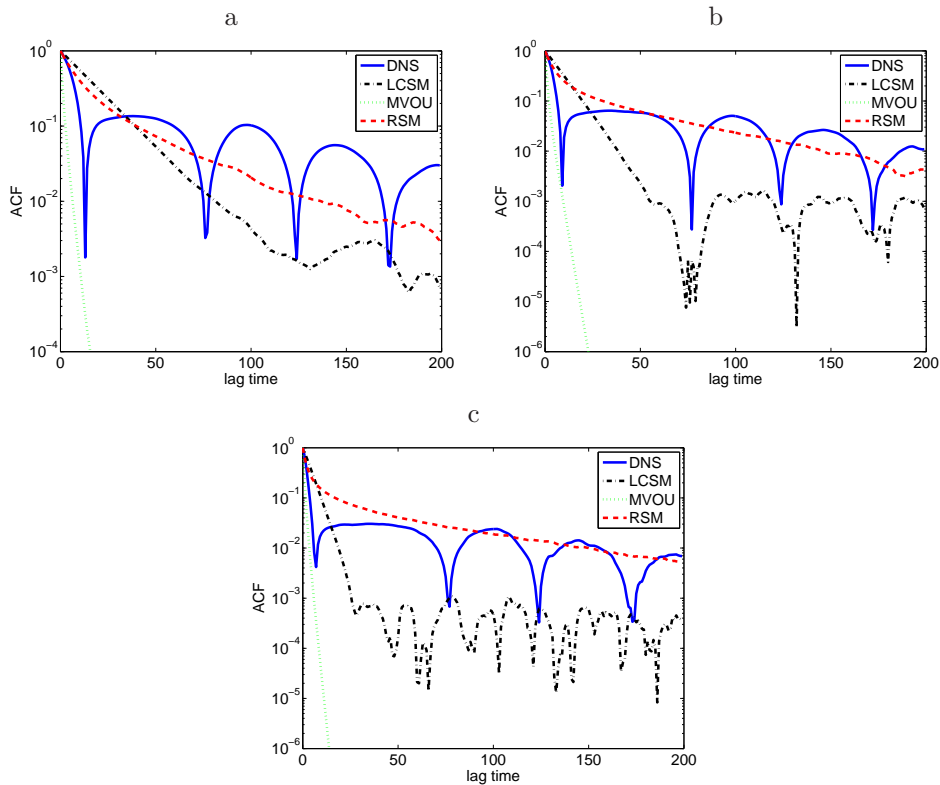


FIG. 4.3. Semilog plot of the absolute value of the time autocorrelation function of x from DNS, linear-closure stochastic model (LCSM), multivariate OU model (MVOU) and reduced stochastic model (RSM) for $n=32$ (a), $n=16$ (b) and $n=8$ (c).

with $\lambda_b = \lambda_a = \lambda_m = 1$ for the complete reduced stochastic model. Interactions involving 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 subgrid-scale parameterization. We will refer to these subgrid-scale terms as additive corrections and we place λ_a in front of them to denote their origin. Multiplicative subgrid-scale correction terms include linear and cubic terms, as well as multiplicative noise. These terms are multiplied by λ_m and they are due to multiplicative triad interactions involving B^{xxy} and B^{yxx} . Finally, we mark the bare truncation term with an λ_b -factor in front of it. The explicit form of the subgrid-scale corrections discussed above can be found in appendix B.

We perform simulations with a single type interactions switched on, i.e., just one among λ_b, λ_a or λ_m is $= 1$, and all others 0. The results are summarized in fig. 4.6. First, we see that the bare truncation model (BRT), a model without any subgrid-scale corrections, fails completely to reproduce the autocorrelation function and kurtosis of the DNS. Further, a reduced model with only additive triad corrections captures the autocorrelation function and probability density function of the reduced stochastic model, indicating that the additive triads are the dominant interactions. However, by definition such model does not show any deviations from Gaussianity (since this is a purely Gaussian process), see fig 4.6 (b). The statistics from simulations with

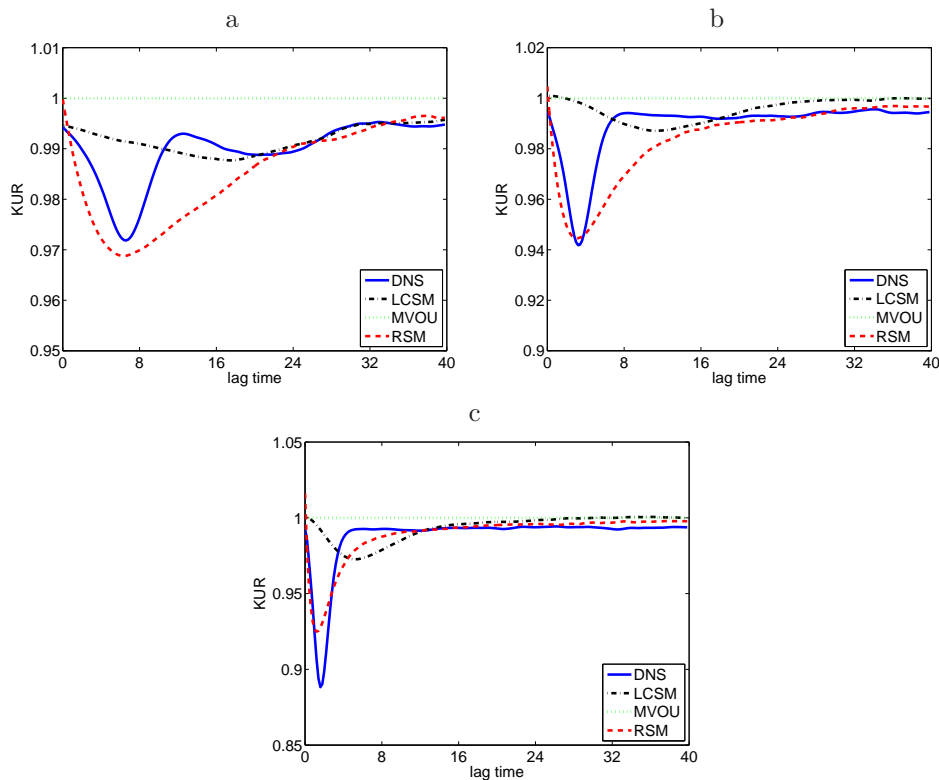


FIG. 4.4. Kurtosis of the resolved modes from DNS, linear-closure stochastic model (LCSM), multivariate OU model (MVOU) and reduced stochastic model (RSM) for $n=32$ (a), $n=16$ (b) and $n=8$ (c). Note the different scaling for the ordinate in the figures.

two types of interactions are displayed in fig. 4.7. We observe that a model including additive corrections and multiplicative ones, i.e. $\lambda_a = \lambda_m = 1$ and $\lambda_b = 0$, reproduces not only the autocorrelation function and the probability density function of the full reduced stochastic model but also the kurtosis. Clearly the closure is effectively controlling the statistics of the reduced stochastic model.

5. Summary and conclusions

We present a framework for the construction of stochastic subgrid-scale parameterizations consistent with the finite-difference discretization of the model equations. The subgrid-scale model is based on the assumption of a time scale separation between resolved and subgrid-scale modes. It assumes further that the fast self-interactions between subgrid-scale modes in the equations for the subgrid-scale modes can be modeled as an one dimensional Ornstein-Uhlenbeck (OU) process. The particular form of the subgrid-scale correction terms is derived in a systematic way from the discretized model by applying the Majda, Timofeyev & Vanden-Eijnden stochastic mode reduction strategy [16, 17, 21]. The subgrid-scale parameterization is rigorously valid in the limit of infinite scale separation between resolved and unresolved modes. It includes linear and cubic deterministic corrections, as well as additive and multiplicative noise terms. The subgrid-scale corrections are determined without any regression fitting of the resolved modes, they require only the variance and integrated autocorrelation

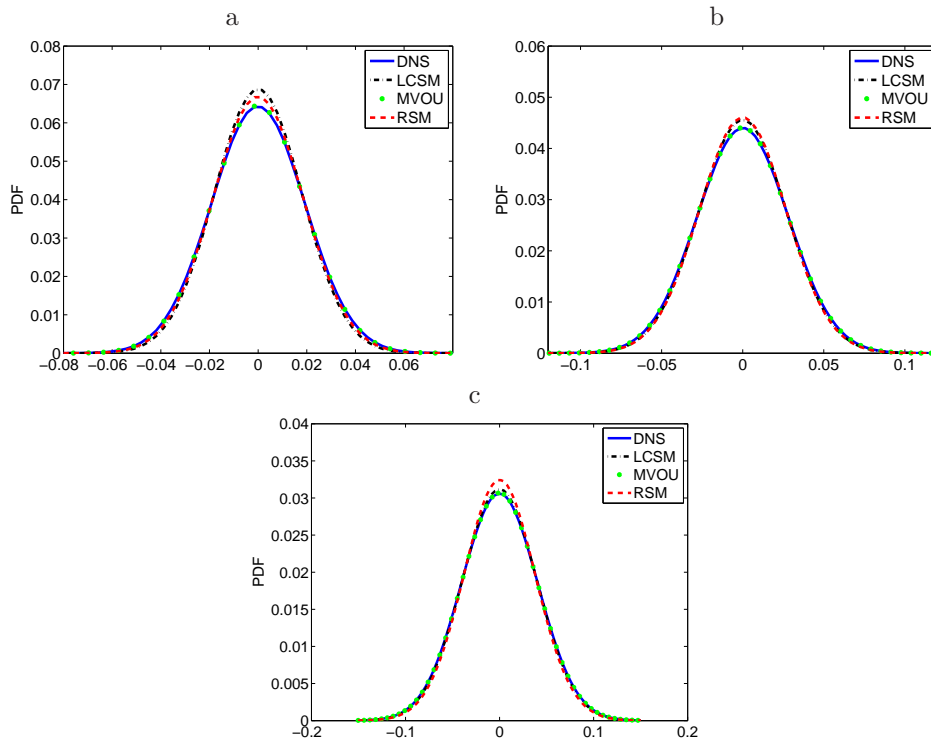


FIG. 4.5. Probability density function of the resolved modes from DNS, linear-closure stochastic model (LCSM), multivariate OU model (MVOU) and reduced stochastic model (RSM) for $n=32$ (a), $n=16$ (b) and $n=8$ (c). Note the different scaling for the axis in the figures.

function of the subgrid-scale modes as input.

As a test model for the new framework we use the inviscid Burgers-Hopf equation. We consider an energy and momentum conserving finite-difference discretization of the model equation [20] and construct for it a stochastic subgrid-scale parameterization. The resulting reduced stochastic model has been compared with the full model (DNS) and with two purely empirical coarse-grained models. The first such model is the multivariate OU model, constructed in such a way as to match the covariance and integrated lagged cross-correlations of the full model. The second empirical model is the linear-closure stochastic model which incorporates one dimensional OU closure in addition to the non-linear self-interactions of the resolved variables. In order to assess the quality of the different subgrid-scale parameterizations, we performed simulations by varying the number of subgrid-scale modes which have to be modeled and comparing the resulting statistics with those from the full model.

All coarse-grained models capture the probability density function from the DNS, with the multivariate OU model performing best. However, the latter model completely underestimates the autocorrelation decay time scale. The linear-closure stochastic model and the reduced stochastic model, on the other hand, can capture the overall decay behavior in the autocorrelation function. Beyond this, the reduced stochastic model exhibits several advantages. (1) It captures the autocorrelation function at very small lags. (2) If the oscillations in the autocorrelation function of the

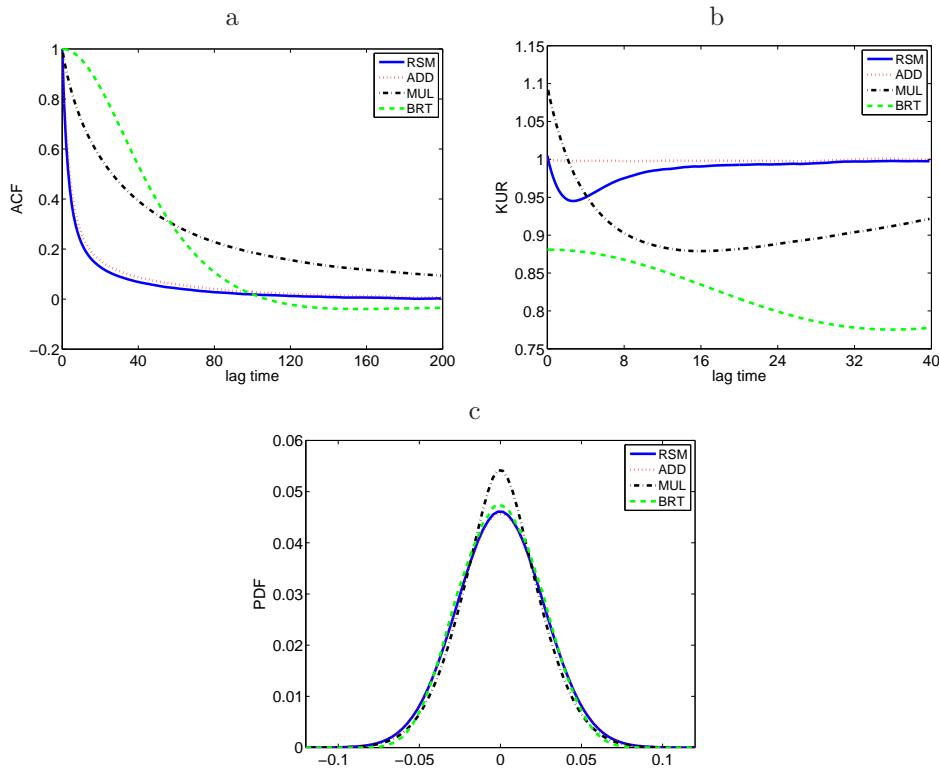


FIG. 4.6. Autocorrelation function (a), kurtosis (b) and probability density function (c) computed with different correction terms in the reduced stochastic model: bare truncation (BRT), additive triads (ADD), multiplicative triads (MUL) and full reduced stochastic model (RSM). All models with $n = 16$. See text for an explanation of the different correction terms.

DNS are small, the reduced stochastic model captures also the low-frequency envelop. The linear-closure stochastic model, on the other hand, decays always too fast at large lags. (3) The inspection of the lagged kurtosis of the models showed that the reduced stochastic model is able to capture better than the linear-closure stochastic model some weakly non-Gaussian features in the DNS. (4) The statistics of the resolved modes show large changes in contrast to those of the subgrid-scale modes, when the total number of resolved modes N_x is varied in the full model. Thus, the subgrid-scale parameterization of the purely empirical coarse-grained models has to be recomputed for each case considered here. However, such recomputation is not required for the reduced stochastic model, since for this parameterization the dependency on N_x (and on the averaging interval n) is known. Of course, for the Burgers-Hopf equation, considered here, one can find a simple scaling relation between the parameters in the purely empirical closures and N_x , in order to avoid recomputing these parameters. However, this might not be the case, if more complex models are considered. (5) Another practical advantage of the reduced stochastic model is that the empirical coefficients, entering the closure, can be inferred from a shorter time series. In the reduced stochastic model these coefficients depend only on the fast subgrid-scale modes, in the purely empirical reduced models, on the other hand, on the slow resolved modes, characterized by large decorrelation time scales.

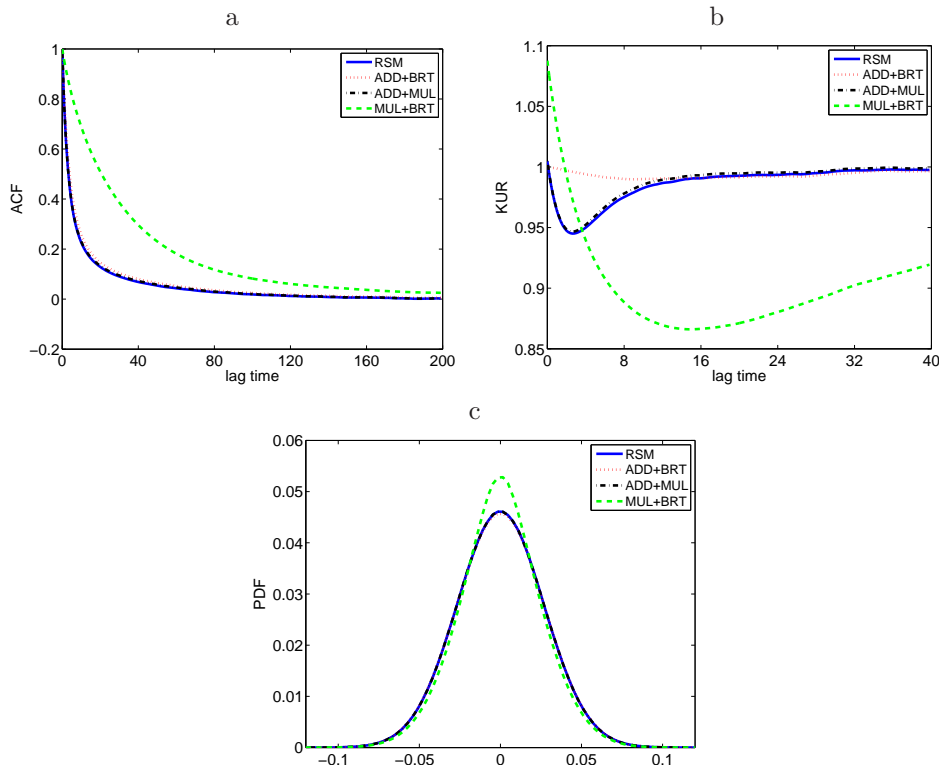


FIG. 4.7. Notation as in fig. 4.6.

An interesting result was that the statistics of the reduced stochastic model is effectively controlled by the closure. By performing a budget analysis it was found that the linear deterministic term and additive noise are the dominant subgrid-scale corrections in the model. These terms result from nonlinear interactions between subgrid-scale modes. Next in importance comes the multiplicative noise, only followed by the deterministic terms from the barely truncated model.

One possibility to improve the present approach is by incorporating interactions between different subgrid-scale modes in the OU process from the closure assumption. Work on the Burgers equation with forcing and dissipation suggests that the subgrid-scale self-interactions can be modeled by an OU process coupling small number of neighboring grid cells [9].

Acknowledgement. We would like to thank Andrew Majda and Christian Franzke for useful discussions. We thank the two reviewers 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.

Appendix A. Interaction coefficients in the discretized Burgers-Hopf equation.

The quadratic interaction terms from (2.9), (2.10) are given by

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

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

$$\sum_{j \in s_x} \sum_{k \in s} B_{ijk}^{xxy} x_j y_k = -\frac{1}{6n\Delta x} (2x_i y_{ir} + x_i y_{ir+1} + x_{i+1} y_{ir} + 2x_{i+1} y_{ir+1} - 2x_i y_{il} - x_i y_{il-1} - x_{i-1} y_{il} - 2x_{i-1} y_{il-1}), \quad (\text{A.3})$$

$$\sum_{j \in s_x} \sum_{k \in s} B_{ijk}^{yxy} x_j y_k = -\frac{1}{6\Delta x} \left(2x_{[\frac{i+1}{n}]} y_{i+1} + x_{[\frac{i+1}{n}]} y_i + x_{[\frac{i}{n}]} y_{i+1} - 2x_{[\frac{i-1}{n}]} y_{i-1} - x_{[\frac{i-1}{n}]} y_i - x_{[\frac{i}{n}]} y_{i-1} \right) \quad (\text{A.4})$$

$$\begin{aligned} & + \frac{1}{6n\Delta x} (2x_{[\frac{i}{n}]} y_{ir} + x_{[\frac{i}{n}]} y_{ir+1} + x_{[\frac{i}{n}]+1} y_{ir} + 2x_{[\frac{i}{n}]+1} y_{ir+1} - 2x_{[\frac{i}{n}]-1} y_{il-1} - x_{[\frac{i}{n}]} y_{il-1} - x_{[\frac{i}{n}]-1} y_{il} - 2x_{[\frac{i}{n}]} y_{il}), \\ & \sum_{j \in s_x} \sum_{k \in s_x} B_{ijk}^{yxx} x_j x_k = -\frac{1}{6\Delta x} \left(x_{[\frac{i+1}{n}]}^2 + x_{[\frac{i}{n}]} x_{[\frac{i+1}{n}]} - x_{[\frac{i-1}{n}]}^2 - x_{[\frac{i}{n}]} x_{[\frac{i-1}{n}]} \right) \quad (\text{A.5}) \\ & + \frac{1}{6n\Delta x} \left(x_{[\frac{i}{n}]+1}^2 + x_{[\frac{i}{n}]} x_{[\frac{i}{n}]+1} - x_{[\frac{i}{n}]-1}^2 - x_{[\frac{i}{n}]} x_{[\frac{i}{n}]-1} \right). \end{aligned}$$

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

$$ir = \begin{cases} ([\frac{i}{n}] + 1)n - 1 & \text{if } i \in s \\ (i + 1)n - 1 & \text{if } i \in s_x \end{cases}, \quad (\text{A.6})$$

$$il = \begin{cases} ([\frac{i}{n}])n - 1 & \text{if } i \in s \\ in - 1 & \text{if } i \in s_x \end{cases}. \quad (\text{A.7})$$

Appendix B. Subgrid-scale coefficients in the reduced stochastic model.

As discussed in section 4.3, the different subgrid-scale correction terms can be regrouped depending on their origin. The resulting deterministic terms are given by

$$\sum_{j=i-2}^{i+2} M_{ij}^a x_j = \frac{\sigma^2}{2\gamma^2} \sum_j \sum_{m,n} B_{imn}^{xyy} B_{mjn}^{yxy} x_j, \quad (\text{B.1})$$

$$\sum_{j=i-2}^{i+2} M_{ij}^m x_j = \frac{\sigma^2}{2\gamma^2} \sum_j \sum_{k,l} B_{ilk}^{xxy} B_{ljk}^{xxy} x_j, \quad (\text{B.2})$$

$$\sum_{j,k,l=i-2}^{i+2} C_{ijkl} x_j x_k x_l = \frac{1}{\gamma} \sum_{j,k,l} \sum_p B_{ilp}^{xxy} B_{pjk}^{yxx} x_j x_k x_l. \quad (\text{B.3})$$

For the explicit computation of the subgrid-scale correction terms we utilized the MATLAB symbolic toolbox, we obtain the following expressions

$$\sum_{j=i-1}^{i+1} M_{ij}^a x_j = \frac{\sigma^2 a^2}{4\gamma^2 n} (x_{i-1} - 2x_i + x_{i+1}) \left(5 + \frac{1}{n}\right), \quad (\text{B.4})$$

$$\sum_{j=i-1}^{i+1} M_{ij}^m x_j = -\frac{\sigma^2 a^2}{2\gamma^2 n^2} (x_{i-1} - 2x_i + x_{i+1}), \quad (\text{B.5})$$

$$\begin{aligned} \sum_{j,k,l=i-2}^{i+2} C_{ijkl} x_j x_k x_l &= \frac{1}{\gamma} \frac{a^2}{n} (5x_{i-1}^3 + 3x_{i-1}^2 x_i + 3x_{i+1}^2 x_i + 5x_{i+1}^3 - 3x_{i-1} x_i^2 - 10x_i^3 - 3x_i^2 x_{i+1}) \\ &\quad - \frac{1}{\gamma} \frac{a^2}{n^2} (2x_{i-2}^2 x_{i-1} + x_{i-2}^2 x_i + 2x_{i-2} x_{i-1}^2 + x_{i-2} x_{i-1} x_i + x_{i-1}^3 + x_{i+1}^3 \\ &\quad + x_{i+2} x_{i+1} x_i + 2x_{i+1}^2 x_{i+2} + x_{i+2}^2 x_i + 2x_{i+2}^2 x_{i+1} - x_{i-1}^2 x_{i+1} - x_{i-1}^2 x_i \\ &\quad - 3x_{i-1} x_i^2 - 2x_{i-1} x_i x_{i+1} - 2x_i^3 - 3x_i^2 x_{i+1} - x_i x_{i+1}^2 - x_{i+1}^2 x_{i-1}), \end{aligned} \quad (\text{B.6})$$

with $a = -1/(6\Delta x)$. The subgrid-scale stochastic correction terms are defined by

$$\sum_{j=i-\frac{1}{2}}^{i+\frac{1}{2}} P_{ij} dW_j^{(1)} = \frac{\sigma^2}{\sqrt{2}\gamma} \sum_{j,k} B_{ijk}^{xyy} d\tilde{W}_{jk}^{(1)}, \quad (\text{B.7})$$

$$\sum_{j=i-\frac{1}{2}}^{i+\frac{1}{2}} D_{ij}(x) dW_j^{(2)} = \frac{\sigma}{\gamma} \sum_{j,k} B_{ijk}^{xxy} x_j d\tilde{W}_k^{(2)}. \quad (\text{B.8})$$

The right hand sides of the last two equations represent the general form of the noise terms after the mode reduction procedure, see [16]. These terms involve independent Wiener processes $\tilde{W}^{(1)}$ and $\tilde{W}^{(2)}$ with the dimension of $N \times N$ and N , respectively (N number of subgrid-scale modes). This high dimensionality makes (B.7),(B.8) inappropriate for an implementation. However, all terms on the right hand side of (B.7),(B.8) are associated with a flux either at the left or at the right boundary of the coarse cell i . Terms, corresponding to the same flux, are combined and represented using an effective Wiener process ($W^{(1)}, W^{(2)}$) with the low-dimension N_x . Thus, the explicit form of the noise terms in the subgrid-scale model reads

$$\sum_{j=i-\frac{1}{2}}^{i+\frac{1}{2}} P_{ij} dW_j^{(1)} = \frac{\sqrt{5}\sigma^2 a}{2\sqrt{\gamma}\gamma n} (dW_{i+\frac{1}{2}}^{(1)} - dW_{i-\frac{1}{2}}^{(1)}), \quad (\text{B.9})$$

$$\begin{aligned} \sum_{j=i-\frac{1}{2}}^{i+\frac{1}{2}} D_{ij}(x) dW_j^{(2)} &= \frac{a\sigma}{n\gamma} \left(\sqrt{5x_i^2 + 8x_i x_{i+1} + 5x_{i+1}^2} dW_{i+\frac{1}{2}}^{(2)} - \right. \\ &\quad \left. \sqrt{5x_i^2 + 8x_i x_{i-1} + 5x_{i-1}^2} dW_{i-\frac{1}{2}}^{(2)} \right). \end{aligned} \quad (\text{B.10})$$

REFERENCES

- [1] U. ACHATZ AND G. BRANSTATOR, *A two-layer model with empirical linear corrections and reduced order for studies of internal climate variability*, J. Atmos. Sci., 56 (1999), pp. 3140–3160.
- [2] U. ACHATZ AND J. D. OPSTEEGH, *Primitive-Equation-Based Low-Order Models with Seasonal Cycle. Part I: Model Construction.*, J. Atmos. Sci., 60 (2003), pp. 465–477.
- [3] U. ACHATZ, G. SCHMITZ, AND K.-M. GREISIGER, *Principal Interaction Patterns in Baroclinic Wave Life Cycles.*, Journal of Atmospheric Sciences, 52 (1995), pp. 3201–3213.
- [4] G. BRANSTATOR AND S. E. HAUPT, *An Empirical Model of Barotropic Atmospheric Dynamics and Its Response to Tropical Forcing.*, Journal of Climate, 11 (1998), pp. 2645–2667.
- [5] A. J. CHORIN, A. P. KAST, AND R. KUPFERMAN, *Optimal prediction of underresolved dynamics*, Proceedings of the National Academy of Science, 95 (1998), pp. 4094–4098.
- [6] ———, *Unresolved computation and optimal prediction*, Commun. Pure Appl. Math., 52 (1999), pp. 1231–1254.
- [7] D. CROMMELIN AND E. VANDEN-ELJNDEN, *Subgrid-Scale Parameterization with Conditional Markov Chains*, Journal of Atmospheric Sciences, 65 (2008), pp. 2661–2675.
- [8] J. CULINA, S. KRAVTSOV, AND A. H. MONAHAN, *Stochastic Parameterization Schemes for Use in Realistic Climate Models*, Journal of Atmospheric Sciences, 68 (2011), pp. 284–299.
- [9] S. I. DOLAPTCHIEV, U. ACHATZ, AND I. TIMOFEYEV, *Stochastic closure for local averages in the finite-difference discretization of the forced Burgers equation*, Theor. Comp. Fluid Dyn., (2012).
- [10] C. FRANZKE AND A. J. MAJDA, *Low-Order Stochastic Mode Reduction for a Prototype Atmospheric GCM.*, J. Atmos. Sci., 63 (2006), pp. 457–479.
- [11] C. FRANZKE, A. J. MAJDA, AND E. VANDEN-ELJNDEN, *Low-Order Stochastic Mode Reduction for a Realistic Barotropic Model Climate.*, J. Atmos. Sci., 62 (2005), pp. 1722–1745.
- [12] I. HORENKO, S. I. DOLAPTCHIEV, A. V. ELISEEV, I. I. MOKHOV, AND R. KLEIN, *Metastable Decomposition of High-Dimensional Meteorological Data with Gaps*, Journal of Atmospheric Sciences, 65 (2008), pp. 3479–+.
- [13] S. KRAVTSOV, D. KONDRASHOV, AND M. GHIL, *Multilevel regression modeling of nonlinear processes: Derivation and applications to climatic variability*, Journal of Climate, 18(21) (2005), pp. 4404–4424.
- [14] F. KWASNIOK, *Reduced Atmospheric Models Using Dynamically Motivated Basis Functions*, Journal of Atmospheric Sciences, 64 (2007), pp. 3452–+.
- [15] A. MAJDA AND I. TIMOFEYEV, *Low-dimensional chaotic dynamics versus intrinsic stochastic noise: a paradigm model*, Physica D Nonlinear Phenomena, 199 (2004), pp. 339–368.
- [16] A. MAJDA, I. TIMOFEYEV, AND E. VANDEN-ELJNDEN, *A mathematical framework for stochastic climate models*, Commun. Pure Appl. Math., 54 (2001), pp. 891–974.
- [17] ———, *A priori tests of a stochastic mode reduction strategy*, Physica D, 170 (2002), pp. 206–252.
- [18] ———, *Stochastic models for selected slow variables in large deterministic systems*, Nonlinearity, 19 (2006), pp. 769–794.
- [19] A. J. MAJDA AND B. KHOUIDER, *Stochastic and mesoscopic models for tropical convection*, Proceedings of the National Academy of Science, 99 (2002), pp. 1123–1128.
- [20] A. J. MAJDA AND I. TIMOFEYEV, *Statistical Mechanics for Truncations of the Burgers-Hopf Equation: A Model of Intrinsic Stochastic Behavior with Scaling*, Milan jour. math., 70 (2002), pp. 39–96.
- [21] A. J. MAJDA, I. TIMOFEYEV, AND E. VANDEN-ELJNDEN, *Systematic Strategies for Stochastic Mode Reduction in Climate.*, J. Atmos. Sci., 60 (2003), pp. 1705–1722.
- [22] K. NIMSAILA AND I. TIMOFEYEV, *Markov Chain Stochastic Parametrizations of Essential Variables*, SIAM Mult. Mod. Simul., 8(5) (2010), pp. 2079–2096.
- [23] U. SCHUMANN, *Subgrid scale model for finite difference simulations of turbulent flows in plane channels and annuli*, Journal of Computational Physics, 18 (1975), pp. 376–404.
- [24] F. M. SELTEN, *An Efficient Description of the Dynamics of Barotropic Flow.*, Journal of Atmospheric Sciences, 52 (1995), pp. 915–936.
- [25] D. S. WILKS, *Effects of stochastic parametrizations in the Lorenz '96 system*, Quarterly Journal of the Royal Meteorological Society, 131 (2005), pp. 389–407.
- [26] C. R. WINKLER, M. NEWMAN, AND P. D. SARDESHMUKH, *A Linear Model of Wintertime Low-Frequency Variability. Part I: Formulation and Forecast Skill.*, Journal of Climate, 14 (2001), pp. 4474–4494.
- [27] N. J. ZABUSKY AND M. D. KRUSKAL, *Interaction of "Solitons" in a Collisionless Plasma and the Recurrence of Initial States*, Physical Review Letters, 15 (1965), pp. 240–243.