NUMERICAL SCHEMES FOR STOCHASTIC BACKSCATTER IN THE INVERSE CASCADE OF QUASIGEOSTROPHIC TURBULENCE

IAN GROOMS[†‡], YOONSANG LEE[†§], AND ANDREW J. MAJDA[†§]

Abstract. Backscatter is the process of energy transfer from small to large scales in turbulence; it is crucially important in the inverse energy cascades of geophysical turbulence, where the net transfer of energy is from small to large scales. One approach to modeling backscatter in under-resolved simulations is to add a stochastic forcing term. This study, set in the idealized context of the inverse cascade of two-layer quasigeostrophic turbulence, focuses on the importance of spatial and temporal correlation in numerical stochastic backscatter schemes when used with low-order finite-difference spatial discretizations. A minimal stochastic backscatter scheme is developed as a stripped-down version of stochastic superparameterization [Grooms and Majda, J. Comp. Phys. 271, (2014)]. This simplified scheme allows detailed numerical analysis of the spatial and temporal correlation structure of the modeled backscatter. Its essential properties include a local formulation amenable to implementation in finite difference codes and non-periodic domains, and tunable spatial and temporal correlations. Experiments with this scheme in the idealized context of homogeneous two-layer quasigeostrophic turbulence demonstrate the need for stochastic backscatter to be smooth at the coarse grid scale when used with low-order finite difference schemes in an inverse-cascade regime. In contrast, temporal correlation of the backscatter is much less important for achieving realistic energy spectra. It is expected that the spatial and temporal correlation properties of the simplified backscatter schemes examined here will inform the development of stochastic backscatter schemes in more realistic models.

Key words. Backscatter; Quasigeostrophic Turbulence; Stochastic Modeling

1. Introduction. Backscatter is the process of energy transfer from small to large scales in turbulence. Quasigeostrophic (QG) and two-dimensional (2D) turbulence, which are idealized models of large-scale geophysical turbulence, both include regimes with a net transfer of kinetic energy from small to large scales [5, 7]. While the dynamical regime of the atmosphere precludes significant net transfer to large scales, there is evidence of an inverse transfer of kinetic energy in the oceans [36, 37]. Ocean models used in coupled climate simulations are currently unable to resolve the full range of scales that correspond to quasigeostrophic dynamics, and must model the effects of the unresolved scales. When the dynamics transfer net energy upscale across the model grid scale, a backscatter model is crucial to maintain an accurate energy level for the resolved scales.

Representations of backscatter are included in many engineering and atmospheric turbulence models, where the backscatter is much weaker. For example, backscatter is incorporated in large-eddy simulations (LES) of three-dimensional isotropic turbulence either by including a stochastic term in the equations for the resolved dynamics [20, 24, 25, 35], or through nonlinear deterministic models [2, 10, 26]. In larger-scale geophysical models backscatter has been also represented by stochastic forcing terms [3, 9, 17, 18, 33] and by nonlinear deterministic models [15–17, 29–31, 33]. These models have been applied in situations where the net transfer of kinetic energy is either downscale or negligible (see [16] for a possible exception), whereas, as noted above, application to non-eddy-resolving ocean models requires the ability to represent a net

[†]Center for Atmosphere Ocean Science, Courant Institute of Mathematical Sciences, New York University, 251 Mercer St., New York, NY 10012

[‡]Corresponding author, email address: grooms@cims.nyu.edu. I. Grooms is supported by the NSF CMG program grant DMS-1025468, and by ONR MURI grant N00014-12-1-0912.

[§]Y. Lee and A. J. Majda are supported by ONR MURI grant N00014-12-1-0912.

upscale transfer. Backscatter schemes are also often formulated in terms of spherical harmonics [3, 9, 18], which is appropriate for applications to fluids in a spherical shell, like the atmosphere. But application to domains with boundaries, like the ocean, precludes the use of Fourier modes or spherical harmonics in specifying the backscatter model; a local approach is required, similar to those in [17, 20, 24, 25, 33, 35].

This paper studies the interplay between low-order finite-difference numerics and the spatial and temporal correlation properties of stochastic backscatter using a stochastic backscatter model with a local formulation in an idealized setting with net upscale energy transfer. The stochastic backscatter used here is based on a simplified form of stochastic superparameterization [12–14, 23] that has previously been applied to quasigeostrophic turbulence in the inverse cascade regime. This original, simplified scheme was initially developed in [13], section 3.3, and is described in §2; its spatial correlation structure is analyzed and modified in §2.1, and temporal correlation is added in §2.2. A main initial result of the current investigation is that direct application of the original method, where the backscatter forcing is strongest at the coarse grid scale, in a low-order finite-difference code gives very poor results, much worse than in a model with spectral numerics (results in §5). The method was originally developed in a setting where the under-resolved model used a spectral discretization [12, 13]; the huge impact of simply changing the numerics underscores the sensitivity of models to the details of backscatter parameterizations in the inverse cascade regime.

The failure of the original method in the context of a low-order code results primarily from the fact that the low-order model is unable to efficiently transfer energy injected near the coarse grid scale to larger scales. Energy injected near the coarse grid scale, failing to cascade to larger scales, accumulates until the simulation is eventually swamped by small-scale energy. This is remedied by modifying the spatial correlation structure of the backscatter scheme so that it injects zero energy at the coarse grid scale, and smoothly transitions to the original scheme at larger, more well-resolved scales. The spatial correlation structure is modified using a simple local spatial average. The improvements obtained by spatial smoothing suggest that results might be further improved by adding temporal correlation to the stochastic backscatter; this is natural since the real, physical forcing of the large scales by the unresolved small scales has a finite decorrelation time, unlike the white-noise backscatter of the original scheme. However, our experiments in §6 find minimal differences between simulations with temporally-correlated backscatter and simulations with white-noise backscatter. We only evaluate the performance of the schemes based on time-averaged energy spectra, and it is possible that temporal correlation may improve other aspects of the simulations. The kinetic energy spectra in the models with spatially-smoothed backscatter are quite accurate, but their potential energy spectra are not because the backscatter scheme only models the inverse transfer of kinetic energy and not the direct transfer of potential energy from large to small scales. We briefly demonstrate in §7 that the full stochastic superparameterization scheme of [13] with spatial and temporal correlation is able to produce accurate kinetic and potential energy spectra.

The paper is organized as follows. Section 2 presents and analyses the stochastic kinetic energy backscatter scheme, including the incorporation of spatial and temporal correlation. Section 3 presents a high resolution simulation for comparison with low resolution models in following sections. The results of low resolution models without backscatter are presented in §4, to emphasize the need for backscatter. Sections 5 and 6 show the results of backscatter simulations with and without spatial and tem-

poral correlations; §7 presents the results of the full stochastic superparameterization scheme of [13] with the new spatial and temporal correlation properties. The results are reviewed and conclusions are offered in §8.

2. The Stochastic Backscatter Model. Quasigeostrophic (QG) and incompressible two-dimensional (2D) fluid dynamics are governed by conservation of a scalar q, which is vorticity in 2D and potential vorticity in QG:

$$\partial_t q + \nabla_h \cdot (\boldsymbol{u} q) = \dot{q} \tag{2.1}$$

where $\dot{q}$ denotes forcing and dissipation terms and $\nabla_h = (\partial_x, \partial_y)$. The scalar q is related to a streamfunction ψ for the incompressible horizontal velocity $\boldsymbol{u} = (-\partial_y \psi, \partial_x \psi)$ by an elliptic inversion

$$\nabla_h^2 \psi + \mathbb{L}\psi = q \tag{2.2}$$

where $\mathbb{L}$ is a linear operator. In 2D dynamics $\mathbb{L} = 0$ and in QG it is a second-order negative semi-definite differential operator in z. Sections 3-7 will focus on the simplest model of geophysical turbulence, the two-equal-layer QG model [40], where the vertical direction is discretized into two levels $j = 1, 2$ and where the elliptic inversion is

$$\nabla_h^2 \psi_j + \frac{k_d^2}{2}(\psi_i - \psi_j) = q_j, \quad i \neq j. \tag{2.3}$$

The deformation radius k_d^{-1} is related to the nonzero eigenvalue of $\mathbb{L}$. The discussion generalizes to multiple vertical levels in a straightforward manner.

In the two-layer case there is a scalar conservation equation of the form (2.4) for each layer:

$$\partial_t q_j + \nabla \cdot (\boldsymbol{u}_j q_j) = \dot{q}_j. \tag{2.4}$$

The unforced, inviscid dynamics with $\dot{q}_j = 0$ conserve energy in the form

$$\frac{1}{2} \iint \sum_j |\boldsymbol{u}_j|^2 + \frac{k_d^2}{2} (\psi_1 - \psi_2)^2 \, \mathrm{d}x\mathrm{d}y \tag{2.5}$$

where the integral of $k_d^2(\psi_1 - \psi_2)^2/4$ is the potential energy.

An equation for the large-scale dynamics can be derived by applying a low-pass spatial filter, denoted by $\overline{(\cdot)}$, to (2.4)

$$\partial_t \overline{q}_j + \nabla \cdot (\overline{\boldsymbol{u}}_j \, \overline{q}_j) = \nabla \cdot (\overline{\boldsymbol{u}_j q_j} - \overline{\boldsymbol{u}}_j \, \overline{q}_j) + \dot{\overline{q}}_j. \tag{2.6}$$

Examples of low-pass filters include truncation to a set of large-scale Fourier modes, or a local Gaussian convolution, among others [22] (see also §3 and §6).

The subgridscale flux of potential vorticity can be further decomposed as follows

$$\overline{\boldsymbol{u}_j q_j} - \overline{\boldsymbol{u}}_j \, \overline{q}_j = \left(\overline{\overline{\boldsymbol{u}}_j \, \overline{q}_j} - \overline{\boldsymbol{u}}_j \, \overline{q}_j\right) + \overline{\overline{\boldsymbol{u}}_j q_j'} + \overline{\boldsymbol{u}_j' \overline{q}_j} + \overline{\boldsymbol{u}_j' q_j'} \tag{2.7}$$

where the prime denotes the small-scale part $(\cdot)' = (\cdot) - \overline{(\cdot)}$. These terms can all be modeled separately [2, 8, 10, 21]; we instead model their combined effect while formally assuming the filter is a Reynolds average, i.e. that it has the same properties

as a statistical average. This approach follows the development of stochastic superparameterization [12, 13, 23], and has been used e.g. by [35] in the development of a stochastic backscatter scheme in incompressible three-dimensional turbulence. Under this assumption the divergence of the subgridscale potential vorticity flux becomes

$$\nabla \cdot (\overline{\boldsymbol{u}_j' q_j'}) = \nabla \cdot \left(\overline{\boldsymbol{u}_j' \nabla_h^2 \psi_j'} \right) + \frac{k_d^2}{2} \nabla \cdot (\overline{\boldsymbol{u}_j' (\psi_i' - \psi_j')}). \tag{2.8}$$

The horizontal Laplacian of ψ_j' is the subgridscale vorticity $\nabla_h^2 \psi_j' = \zeta_j' = \partial_x v_j' - \partial_y u_j'$. The subgridscale potential vorticity flux $\overline{\boldsymbol{u}_j' q_j'}$ includes the vorticity flux $\overline{\boldsymbol{u}_j' \zeta_j'}$, whose divergence can be further expanded as

$$\nabla_h \cdot (\overline{\boldsymbol{u}_j' \zeta_j'}) = \left(\partial_x^2 - \partial_y^2 \right) \overline{u_j' v_j'} + \partial_{xy} \left(\overline{(v_j')^2} - \overline{(u_j')^2} \right). \tag{2.9}$$

This is the curl of the divergence of the Reynolds stress tensor $\overline{\boldsymbol{u}_j' \otimes \boldsymbol{u}_j'}$. It has been shown by [36] that the subgridscale vorticity flux divergence is responsible for the inverse kinetic energy cascade in two-layer QG turbulence. The stochastic kinetic energy backscatter scheme developed here models the subgridscale vorticity flux divergence, ignoring the remaining part of the subgridscale potential vorticity flux.

The effect of subgridscale Reynolds stresses can be modeled in general as the sum of a deterministic drain on the resolved scales plus a stochastic forcing [18, 19, 22]. Careful analysis shows that the deterministic drain can be modeled to a good approximation as a hyperviscous drain of momentum, which in this context implies that the drain component of the vorticity flux divergence takes the form $\nu_{2n}(-\nabla^2)^n \overline{\zeta}_j$ where $n \geq 0$ depends on the grid scale and $\nu_{2n} > 0$ is the turbulent hyperviscosity coefficient [18]. Eddy-permitting ocean models typically use a biharmonic viscosity ($n = 2$), which we adopt here, leaving the coefficient ν_4 as a tunable parameter.

The stochastic backscatter scheme developed here is based on the 'uncorrelated' closure from [13, section 3.3], which is a simplified version of the full stochastic SP scheme of [13]. The subgridscale velocity in this closure is based on modeling the subgridscale dynamics by randomly-oriented plane waves with wavenumbers in the direction θ. In this closure the subgridscale velocity is assumed to be completely independent of the large-scale flow. Note that the assumption of independence is relaxed in the full stochastic superparameterization developed in [13]; the spatial and temporal correlation properties developed in this section for the backscatter scheme are incorporated into stochastic superparameterization in §7. Under the assumption of independence from the large-scale flow, the multiscale description of the subgridscale velocity field in [13] results in subgrid stresses of the form

$$\overline{u_j' v_j'} = \frac{E_0}{2} \sin(2\theta), \quad \overline{(v_j')^2} - \overline{(u_j')^2} = E_0 \cos(2\theta) \tag{2.10}$$

where E_0 is a constant subgridscale kinetic energy and θ is a random field, white in space and time and with a uniform distribution $\theta \in [0, 2\pi)$. These stresses would also be generated by subgridscale velocities with amplitude $\sqrt{E_0}$ in the direction orthogonal to θ, i.e.

$$u_j' = \sqrt{E_0} \sin(\theta), \quad v_j' = \sqrt{E_0} \cos(\theta). \tag{2.11}$$

The direction θ is depth-independent, although a more general vertical correlation structure is clearly possible. In this scheme the Reynolds stresses are thus identical in

each layer; in the full stochastic superparameterization scheme the Reynolds stresses are not completely set by the angle θ, and can be different in each layer even when θ is the same in each layer. The original, 'uncorrelated' backscatter scheme of [13], combined with a biharmonic drain, thus takes the form

$$\nabla \cdot (\overline{\boldsymbol{u}_j' \zeta_j'}) = \frac{E_0}{2} \left(\partial_x^2 - \partial_y^2\right) \sin(2\theta) + E_0 \partial_{xy} \cos(2\theta) + \nu_4 \nabla_h^4 \overline{\zeta}_j \tag{2.12}$$

The backscatter amplitude E_0 and hyperviscosity coefficient ν_4 are treated here as tunable parameters, although if a prediction were available for the subgridscale kinetic energy, e.g. through the subgridscale energy model of [13, section 3.4], it could be used to provide a spatio-temporally variable prediction of E_0. The subgridscale terms in this scheme are completely independent of the large-scale variables; this is not entirely realistic (Nadiga found that the subgridscale potential vorticity flux is correlated with a nonlinear function of the large-scale gradients [27,28]), but it allows the spatial and temporal correlation structure to be analyzed more readily.

2.1. Spatial Correlation. We now demonstrate that the model (2.10) leads to a random forcing of the large-scale dynamics with a k^5 spectrum. The large-scale potential vorticity equation (2.6) has the form

$$\partial_t \overline{q}_j = -\nabla_h \cdot (\overline{\boldsymbol{u}}_j \overline{q}_j) + \frac{E_0}{2} \left(\partial_y^2 - \partial_x^2\right) \sin(2\theta) - E_0 \partial_{xy} \cos(2\theta) - \nu_4 \nabla^4 \overline{\zeta}_j + \dot{\overline{q}}. \tag{2.13}$$

Recall that the random field θ has uniform distribution on the periodic interval $[0, 2\pi)$ and that it is spatially white, i.e. the values at different horizontal locations are independent. These properties of θ imply that the terms $\sin(2\theta)$ and $\cos(2\theta)$ are delta-correlated (white) in space, have zero mean, and are uncorrelated with each other since

$$\mathbb{E}\left[\sin(2\theta)\cos(2\theta)\right] = \frac{1}{2\pi} \int \sin(2\theta)\cos(2\theta) \mathrm{d}\theta = 0 \tag{2.14}$$

where $\mathbb{E}$ denotes the expecation.

Recalling that the Fourier transform of a delta-correlated (white) random field is uniform, the Fourier transform of the backscatter forcing terms is proportional to

$$\frac{1}{2}(k_x^2 - k_y^2) A_k + k_x k_y B_k \tag{2.15}$$

where A_k and B_k are uncorrelated random variables with zero mean and equal variance. The expected value of the square amplitude of the Fourier transform is thus proportional to k^4, and the azimuthally-averaged (1D) spectrum is proportional to k^5. This implies that the stochastic backscatter scheme forces most strongly at the coarse grid scale.

For comparison, a k^4 spectrum is expected for momentum backscatter in homogeneous isotropic three-dimensional turbulence [19, 22, 35]. Schumann [35] points out that although this can be justified by appeal to the Eddy-Damped Quasi-Normal Markovian (EDQNM) model of turbulence, it is also expected based solely on the assumption that the components of the Reynolds stress are independent, and spatially uncorrelated. These are also the only assumptions that go into the proof of the k^5 spectrum for the current model, which suggests that the k^5 behavior is a robust quality of the backscatter spectrum in quasigeostrophic turbulence.

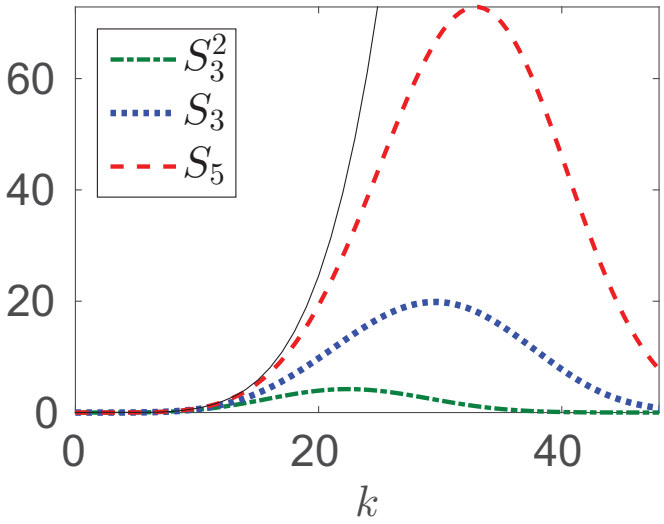

FIG. 2.1. *Angle-integrated (1D) backscatter forcing spectra that result from the three local-average smoothers: the three-point average $\mathcal{S}_3$, the five-point average $\mathcal{S}_5$, and the three-point average applied twice $\mathcal{S}_3^2$. The plots assume a 96×96 coarse grid, with Nyquist wavenumber 48. The thin black line shows an unsmoothed k^5 forcing spectrum.*

The backscatter scheme developed above was shown in [13] to give good results when used with a low resolution model with spectral numerics, but in §5 we demonstrate that it performs poorly in a model with a second order finite-difference discretization. The poor performance is a unique aspect of modeling in a regime where the energy backscattered from unresolved scales to scales near the coarse grid scale is then transferred to even larger scales in an inverse cascade; the nonlinear dynamics of the inverse cascade are not as well represented in the second-order finite difference model as in the spectral model, especially near the coarse grid scale.

In the theory of large-eddy simulation (LES) of homogeneous isotropic three-dimensional turbulence it is common to identify the solution of a low resolution model with a low-pass filtered version of the true dynamics, as in our derivation of the large-scale dynamics (2.6). If the low resolution model uses a spectral numerical method then it is natural to use a Fourier truncation as the filter, projecting the true dynamics onto the spectrum of Fourier modes represented by the low resolution model. In contrast, it is more appropriate to identify low-order finite difference models with a locally smoothed version of the true dynamics, e.g. convolution with a Gaussian. Leslie & Quarini [22] showed that the sharp k^4 backscatter spectrum is associated with a Fourier truncation filter, whereas the backscatter spectrum associated with a smooth Gaussian filter is attenuated near the coarse grid scale, implying that the spectrum of the backscatter forcing in a low-order finite difference code should be attenuated near the coarse grid scale. This can be heuristically associated with the fact that low order finite-difference methods represent the dynamics of the smallest resolved scales less accurately than spectral methods.

To adapt the backscatter scheme for use in the finite-difference code we apply spatial smoothing to the Reynolds stresses. This spatial smoothing should be defined

by a local average in order to be easily implementable in a finite-difference code. Also, since the Nyquist wavenumber is represented on the coarse grid but its dynamics are not resolved we require the smoothing operator to damp the spectrum of the Reynolds stresses to zero at the Nyquist wavenumber. There is only one symmetric three-point moving average that completely removes the Nyquist wavenumber, and its coefficients are $1/4$, $1/2$, $1/4$. We denote this smoothing operator by $\mathcal{S}_3$. Application of this average in the x and y directions modifies the Fourier coefficients of the backscatter forcing terms to

$$\cos^2\left(\frac{k_x\Delta x}{2}\right)\cos^2\left(\frac{k_y\Delta y}{2}\right)\left[\frac{1}{2}(k_x^2-k_y^2)A_k+k_xk_yB_k\right] \tag{2.16}$$

where Δx and Δy are the grid spacing and A_k and B_k are as in equation (2.15). The 2D forcing spectrum is proportional to $\cos^4(k_x\Delta x/2)\cos^4(k_y\Delta y/2)k^4$, which is no longer isotropic, but reduces the forcing to zero for modes at the Nyquist wavenumbers $k_x=\pi/\Delta x$ and $k_y=\pi/\Delta y$. This particular choice of smoothing is ad hoc, but serves to reduce the forcing smoothly to zero at the coarse grid scale without affecting the forcing at large scales, and is easily implemented. In §5 we compare to two alternate smoothers: double application of the $1/4$, $1/2$, $1/4$ average, and application of a five-point moving average with coefficients $-1/16$, $1/4$, $5/8$, $1/4$, and $-1/16$. The five-point average is denoted $\mathcal{S}_5$. Whereas the three-point smoother damps the coefficient of a Fourier mode by a factor of $\cos^2(k_x\Delta x/2)\cos^2(k_y\Delta y/2)$, the five-point smoother damps the coefficient of a Fourier mode by

$$\frac{1}{4}\cos^2(k_x\Delta x/2)(3-\cos(k_x\Delta x))\cos^2(k_y\Delta y/2)(3-\cos(k_y\Delta y)). \tag{2.17}$$

This smoothing operator sets the Nyquist wavenumbers to zero, but damps the remaining modes less than the three-point average.

The angle-averaged (1D) forcing spectra that result from these three smoothing operators are plotted in figure 2.1 (i.e. the angle-averages of (2.16), (2.17), and a version of (2.16) with both $\cos^2$ replaced by $\cos^4$), assuming a 96×96 coarse grid. There are significant differences between these forcing spectra, both in the amplitude of the peak forcing and in the wavenumber of the peak forcing. Nevertheless, in §5 we find that there is little qualitative difference in the results for these three different smoothers. The main difference is that the optimal forcing amplitude E_0 is different for the three filters; it is higher for the double smoother, and lower for the five-point smoother.

It is worth noting that the finite difference approximation to the derivatives of the Reynolds stress terms in equation (2.12) does not reproduce exactly a k^5 forcing spectrum due to truncation errors. For example, the standard centered second-order difference approximation of ∂_x^2 multiplies a single Fourier mode by $-2\sin^2(k_x\Delta x/2)/\Delta x^2$ rather than $-k_x^2$. However, as shown in §5 the smoothing effect of the second-order finite-difference approximation is not sufficient to give good results in the finite difference code without additional smoothing of the Reynolds stresses.

Leith [20], Mason & Thomson [25], and Schumann [35] developed stochastic backscatter schemes for low-order codes. Only Mason & Thomson included spatial smoothing of the Reynolds stresses (via the same three-point average used here), but Schumann noted that it was not necessary since good results could be obtained without it. We expect that the primary difference is that in their applications the net energy transfer was downscale. Energy backscattered near the grid scale of the

coarse model simply dissipates away under the linear action of coarse-model viscosity, and this process is accurately represented even by low-order numerics, whereas in the inverse cascade it must be nonlinearly transferred to larger scales, and low-order numerics do not represent these nonlinear dynamics accurately near the grid scale.

2.2. Temporal Correlation. Leith [20], Mason & Thomson [25], and two of the current authors [12, 13] have developed backscatter schemes where the subgrid-scale stresses are white (delta-correlated) in time; this is implemented simply by updating the random variables at each time step [12, 13, 20], or every other time step [25]. Of course, real subgrid stresses have a nonzero decorrelation time because they are generated by deterministic subgridscale dynamics, and Schumann [35], following Bertoglio [4], incorporated a nonzero decorrelation time in his model of stochastic backscatter.

To give the subgrid stresses a realistic nonzero decorrelation time we model the random field θ at each coarse grid point by a Weiner process on the circle

$$\mathrm{d}\theta = \sigma \mathrm{d}W. \tag{2.18}$$

We prove in the appendix that this leads to an exponential lagged autocorrelation with a decorrelation time of $(2\sigma^2)^{-1}$ for the modeled Reynolds stress terms $\overline{u'v'}$ and $\overline{(v')^2} - \overline{(u')^2}$. The proof makes use of the Fokker-Planck equation for (2.18) on a circle, which has the uniform distribution as a stationary solution. The evolution of the Dirac delta under the action of the Fokker-Planck equation is given in terms of its Fourier series, which together with some trigonometric identities allows us to prove that the lagged autocorrelation function of $\sin(2\theta)$ or $\cos(2\theta)$ is exponential with decay scale $(2\sigma^2)^{-1}$. A natural physical timescale that can be associated with this is the eddy turnover time ℓ/U, where U is the RMS velocity and ℓ is a characteristic eddy length scale. In §6 the decorrelation time of the Reynolds stresses is measured directly, and agrees well with a turnover time based on the deformation radius $\ell = k_d^{-1}$.

Although the spatially-discretized fluid equations with stochastic backscatter are SDEs, the time integration methods used with these backscatter schemes are ad hoc, because the SDEs are generally not in the standard form for which numerical methods are well studied, i.e. $\mathrm{d}X = f(X,t)\mathrm{d}t + g(X,t)\mathrm{d}W$ where W is a Weiner process. Leith [20] notes that in order for his method to be robust to changing time steps the amplitude of the forcing should vary as $(\Delta t)^{-1/2}$ where Δt is the time step size; since the forcing is multiplied by Δt in the time integration scheme the overall method is analogous to the Euler method for Itō SDEs, which multiplies the stochastic forcing term by $(\Delta t)^{1/2}$ instead of Δt. Although the backscatter scheme developed here does not result in a standard-form SDE, we find similar behavior in §6: when the backscatter is white in time the scheme is robust to changing time steps when the amplitude E_0 is scaled by the square root of the time step.

We implement the backscatter forcing as follows. When the forcing is white in time the integration scheme multiplies the forcing terms by $(\Delta t)^{1/2}$ instead of Δt and the forcing terms are held constant through all stages of the Runge–Kutta time step. When the forcing has an exponential decorrelation as described above, the integration scheme multiplies the forcing terms by Δt as usual, and the angle θ evolves under equation (2.18) during the internal stages of the Runge–Kutta time step. We find in §6 that both of these schemes are robust to changes in the time step.

3. An Eddy-Resolving Baseline Simulation. Development of the backscatter model is provided here in the context of the two-layer quasigeostrophic (QG)

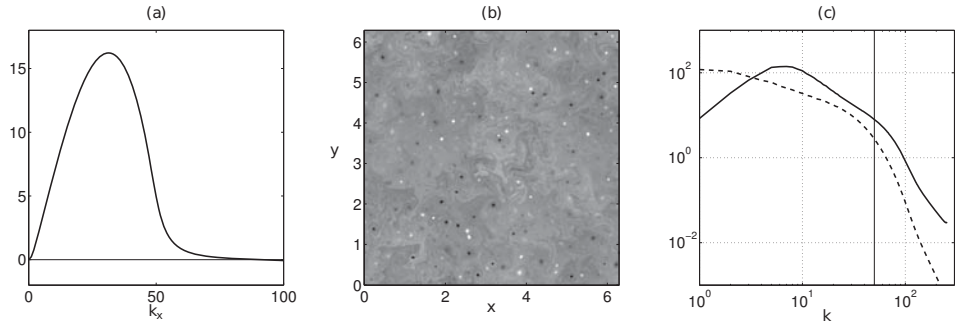

FIG. 3.1. *Properties of the high-resolution reference simulation: (a) Growth rate of linear baroclinic instability, (b) Snapshot of upper-layer potential vorticity q_1, (c) Time-averaged kinetic (solid) and potential (dashed) energy spectra. The thin vertical line in (c) indicates the deformation wavenumber $k_d = 50$.*

model

$$\begin{aligned}
\partial_t q_1 &= -\nabla\cdot(\boldsymbol{u}_1 q_1) - \partial_x q_1 - k_d^2 v_1 - \nu_8\nabla^8 q_1,\\
\partial_t q_2 &= -\nabla\cdot(\boldsymbol{u}_2 q_2) + \partial_x q_2 + k_d^2 v_2 - \nu_8\nabla^8 q_2 - r\nabla^2\psi_2 \qquad (3.1)\\
q_1 &= \nabla^2\psi_1 + \frac{k_d^2}{2}(\psi_2-\psi_1), \quad q_2 = \nabla^2\psi_2 - \frac{k_d^2}{2}(\psi_2-\psi_1),
\end{aligned}$$

where q_j is the potential vorticity in the upper ($j = 1$) and lower ($j = 2$) layers, $\nabla^2\psi_j = \zeta_j$ is the relative vorticity, the velocity-streamfunction relation is $u_j = -\partial_y\psi_j$, $v_j = \partial_x\psi_j$, k_d is the deformation wavenumber (k_d^{-1} is the deformation radius), the coefficient r specifies the strength of linear bottom friction (Ekman drag) and ν_8 is the hyperviscous Reynolds number. A large-scale zonal (east-west, x-direction) vertical shear is applied with equal and opposite unit velocities in the upper and lower layers; this imposed velocity is associated with the terms $(-1)^j(\partial_x q_j + k_d^2 v_j)$. The imposed vertical shear leads to a baroclinic instability that drives the system dynamics [40]. The doubly-periodic, two-layer QG model driven by imposed shear is a classical idealized model of synoptic-scale turbulence in the atmosphere and oceans though there are, of course, many realistic features of ocean dynamics that are lacking from this model (e.g. topography, lateral boundaries, diabatic effects, etc.).

A high resolution reference simulation of equations (3.1) in a 2π-periodic domain with a 512×512 computational grid was run to provide a baseline for comparison with the low resolution simulations of the next sections. The simulations use a Fourier spectral method with semi-implicit fourth-order Runge–Kutta time integration; details of the numerical method are exactly the same as [13]. As in [12, 13], the friction parameter is $r = 16$ and the deformation wavenumber is $k_d = 50$, but we here use a smaller viscous coefficient $\nu = 10^{-17}$, which is the same as used by [39]. The different viscous coefficient has minimal impact on the results compared to [13], affecting only the smallest scales. The simulation starts from a low-energy initial condition, and statistics are computed after the total kinetic energy is saturated, which occurs at approximately $t = 4$.

Figure 3.1(a) shows the growth rate of the linear instability that drives the turbulence. Specifically, it shows the growth rate of Fourier modes of the linearized system (3.1); the most unstable modes at fixed k_x have $k_y = 0$ so only these modes are shown. The most unstable mode has $k_x \approx 31.3$, and the instability drops to negligible rates

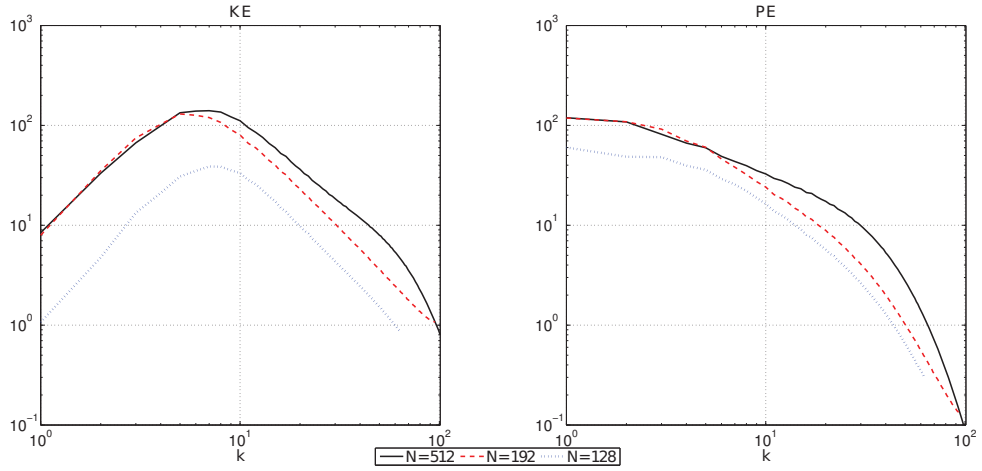

FIG. 4.1. *Time-averaged kinetic (left) and potential (right) energy spectra from simulations with the spectral code at* $N = 512$, 192, *and* 128.

for $k > k_d = 50$. Figure 3.1(b) is a snapshot of the upper-layer potential vorticity q_1, which is dominated by small-scale coherent vortices. Figure 3.1(c) shows the time-averaged kinetic (solid) and potential (dashed) energy spectra; the thin vertical line indicates the deformation wavenumber $k_d = 50$. The imposed shear is equivalent to the effect of an imposed streamfunction in the upper and lower layers of the form $(-1)^j y$, which is associated with a meridional potential vorticity gradient of the form $(-1)^j k_d^2$. The dynamics generate a net potential vorticity flux $v_j q_j$ that acts counter to this imposed gradient. The net potential vorticity flux is equal and opposite in each layer, and is proportional to the area integral of $(v_1 + v_2)(\psi_1 - \psi_2)/4$, which we refer to as the 'heat flux'. The reference simulation generates a net, time-averaged heat flux of 226.

As described in §2, the backscatter amplitude E_0 can be related to the subgridscale kinetic energy via equation (2.11); a similar relation holds for the full stochastic superparameterization scheme of [13]. Although we treat E_0 as a tunable parameter in the following sections, it is of interest to diagnose the subgridscale kinetic energy for comparison with the tuned values of E_0. To diagnose the subgrid kinetic energy we define a low-pass filter $\overline{(\cdot)}$ by multiplying the Fourier coefficients of the high resolution reference simulation by $1 - (k/48)^{12}$ for $k \leq 48$, and by 0 for $k > 48$, where $k = (k_x^2 + k_y^2)^{1/2}$. This has the effect of smoothly transitioning from large scales with $k < 48$ to small scales with $k \geq 48$, where 48 is the Nyquist wavenumber of the coarse model grid. A corresponding high-pass filter is defined by $(\cdot)' = (\cdot) - \overline{(\cdot)}$. For comparison with E_0 we diagnose

$$(u_j')^2 + (v_j')^2 \tag{3.2}$$

from the high resolution reference simulation. The distribution is skewed to large values with a median of 57 and a mean of 268. We tested low-pass filters with exponents 8 and 10 instead of 12, and found that the diagnosed subgridscale energy is not very sensitive to the exponent in the filter,

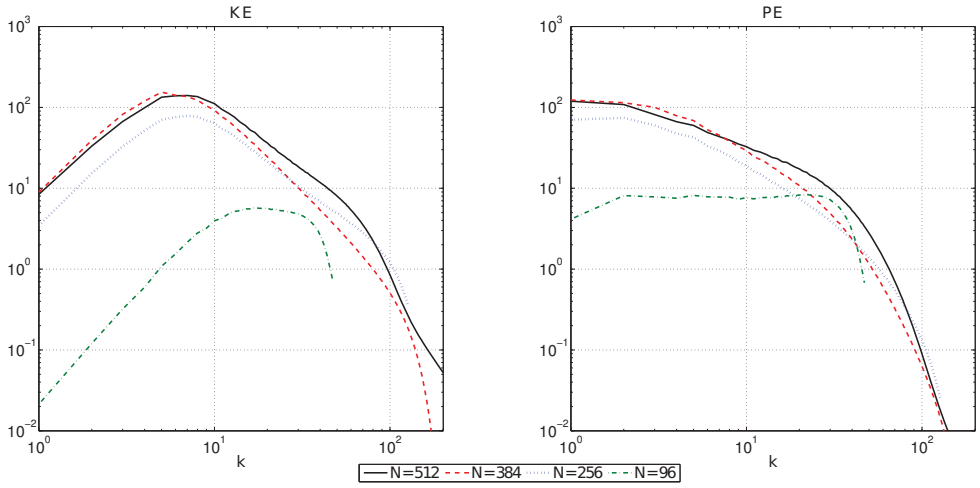

FIG. 4.2. *Time-averaged kinetic (left) and potential (right) energy spectra from simulations with the spectral code at $N = 512$, and from the FD code at $N = 384$, 256, and 96.*

4. Low-Order and Spectral Model Results at Eddy-Permitting Resolutions without Backscatter. In this section we describe a low-order code for the two-layer QG equations that is loosely analogous to large-scale ocean models, at least insofar as it uses a second-order finite difference discretization of the nonlinear terms; we call this code the 'FD' (finite-difference) code to distinguish it from the code that uses fully spectral numerics. We then compare simulations of the spectral and FD codes at various low resolutions and without backscatter schemes, emphasizing the differences between the codes and the need for backscatter.

Most large-scale ocean models use low-order numerical methods based on finite differences, finite elements, or finite volumes. Also, when the spatial grid is sufficient to allow some eddy variability ('eddy-permitting'), they typically use biharmonic diffusion of momentum. The FD code solves the large-scale QG equations (2.13) in the following form

$$\partial_t \overline{q}_1 = -\nabla \cdot (\overline{\boldsymbol{u}}_1 \, \overline{q}_1) - \partial_x \overline{q}_1 - k_d^2 \overline{v}_1 - \nu_4 \nabla^4 \overline{\zeta}_1, \tag{4.1}$$

$$\partial_t \overline{q}_2 = -\nabla \cdot (\overline{\boldsymbol{u}}_2 \, \overline{q}_2) + \partial_x \overline{q}_2 + k_d^2 \overline{v}_2 - \nu_4 \nabla^4 \overline{\zeta}_2 - r \nabla^2 \overline{\psi}_2 \tag{4.2}$$

using the energy- and enstrophy-conserving second-order finite-difference discretization of the Jacobian developed by Arakawa [1]. All other spatial derivatives are approximated using spectral Fourier methods, and in this section there is no backscatter, i.e. E_0=0. Time integration is achieved using the third-order Runge–Kutta scheme of [6], paired with a PI.3.4 adaptive time step algorithm [38].

We run simulations using both the spectral and FD codes on $N \times N$ grids with at $N = 96$, 128, 192, 256, and 384; the viscous coefficients ν_8 and ν_4 are tuned at each resolution to give results with reasonable energy spectra. The goal is to find the minimal resolution necessary to achieve good results in each code. Figure 4.1 shows the time-averaged kinetic and potential energy spectra from simulations with the spectral code at $N = 512$, 192, and 128. The simulation with $N = 192$ used $\nu_8 = 10^{-14}$, and the simulation with $N = 128$ used $\nu_8 = 5 \times 10^{-13}$. Although the simulation with $N = 192$ is not perfect, the most energetic modes are approximately

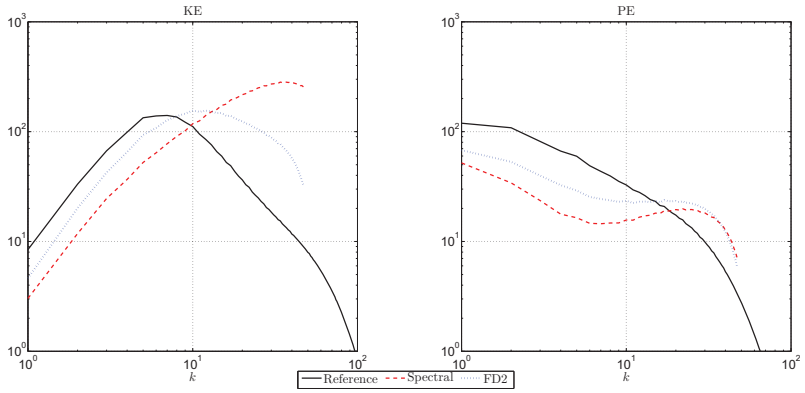

FIG. 5.1. *Time-averaged kinetic (left) and potential (right) energy spectra from simulations with the FD code at $N = 96$, where the backscatter terms are differentiated using spectral (dashed) and second-order finite difference (dotted) approximations. The solid lines are from the reference simulation, and the Reynolds stresses are not spatially smoothed.*

correct and the mean heat flux is 238, quite close to the true value of 226. At $N = 128$ the total energy drops significantly, and the flux drops to 109, approximately half the correct value; the poor results at $N = 128$ cannot be ameliorated by tuning the viscosity.

Figure 4.2 shows the time-averaged kinetic and potential energy spectra from simulations with the spectral code at $N = 512$ and from the FD code at $N = 384$, 256, and 96. The simulation with N =384 used $\nu_4 = 5 \times 10^{-7}$, the simulation with N =256 used $\nu_8 = 2 \times 10^{-7}$, and the simulation with $N = 96$ used $\nu_4 = 1.5 \times 10^{-5}$. There is similar behavior to the spectral model results, but at higher resolution: the results at $N = 384$ are acceptable (the heat flux is 245), but the results at $N = 256$ have too little energy and too little heat flux (172), which can't be improved by tuning the viscosity. The results at $N = 96$ are terrible, but they are included for comparison with results in the following sections that use $N = 96$ and have a backscatter parameterization.

The significant differences between the performance of the under-resolved spectral and FD codes are striking. The spectral code performs much better at low resolution than the FD code, which underscores the importance of developing parameterizations in the more difficult and more realistic context of the FD code. The spectral and FD codes have different viscous operators and Runge–Kutta integration schemes, but we have verified that the discretization of the advection terms is primarily responsible for the differences in performance by running simulations using the same biharmonic viscosity and Runge–Kutta scheme as the FD code but using the spectral Jacobian (not shown).

5. Stochastic Backscatter: Spatial Smoothing. In this section we examine the effect of spatial smoothing on the backscatter model. The FD code is used to

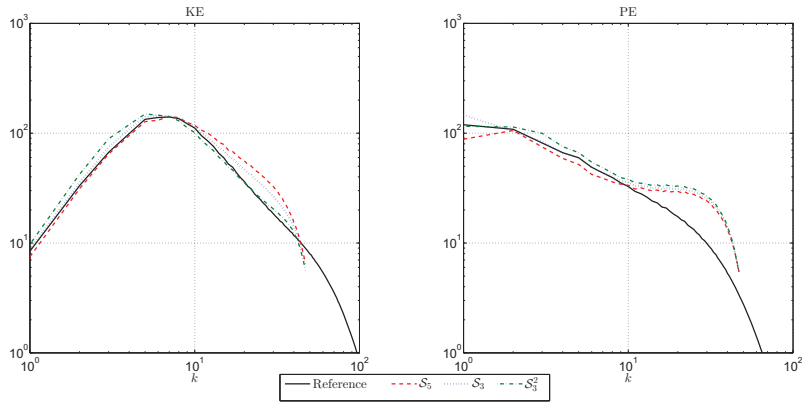

FIG. 5.2. *Time-averaged kinetic (left) and potential (right) energy spectra from simulations with the FD code at $N = 96$, where the backscatter terms are smoothed by three different local averages, denoted $\mathcal{S}_5$ (least smoothing), $\mathcal{S}_3$ (more smoothing), and $\mathcal{S}_3^2$ (most smoothing).*

solve the equations

$$\partial_t \overline{q}_1 = -\nabla \cdot (\overline{\boldsymbol{u}}_1 \overline{q}_1) + \frac{E_0}{2}\left(\partial_y^2 - \partial_x^2\right)\left[\mathcal{S}\sin(2\theta)\right] - E_0\partial_{xy}\left[\mathcal{S}\cos(2\theta)\right] - \partial_x \overline{q}_1 - k_d^2 \overline{v}_1 - \nu_4 \nabla^4 \overline{\zeta}_1, \quad (5.1)$$

$$\partial_t \overline{q}_2 = -\nabla \cdot (\overline{\boldsymbol{u}}_2 \overline{q}_2) + \frac{E_0}{2}\left(\partial_y^2 - \partial_x^2\right)\left[\mathcal{S}\sin(2\theta)\right] - E_0\partial_{xy}\left[\mathcal{S}\cos(2\theta)\right] + \partial_x \overline{q}_2 + k_d^2 \overline{v}_2 - r\nabla^2 \overline{\psi}_2 - \nu_4 \nabla^4 \overline{\zeta}_2 \quad (5.2)$$

where $\mathcal{S}$ denotes a smoothing operator. We denote no smoothing by $\mathcal{S}_0$, the three-point moving average with coefficients $1/4$, $1/2$, $1/4$ in each direction by $\mathcal{S}_3$, and the five-point moving average with coefficients $-1/16$, $1/4$, $5/8$, $1/4$, $-1/16$ in each direction by $\mathcal{S}_5$. Double application of the three-point moving average is denoted $\mathcal{S}_3^2$.

In all simulations in this section θ is a random field that is white in time. Following [12,13] the random field θ is updated at the beginning of each Runge–Kutta time step, and held constant through all stages. The FD code uses a time step $\Delta t = 5 \times 10^{-5}$; for comparison, the spectral simulation with $N = 512$ adaptively chooses a time step on the order of 2.5×10^{-5}, so $\Delta t = 5 \times 10^{-5}$ should be sufficiently small in the coarse code at $N = 96$. Furthermore, as shown in the next section the results are robust to varying the step size by a factor of 2 in either direction.

In all simulations both E_0 and ν_4 are tuned to optimize the kinetic and energy spectra in comparison with the spectral reference simulation. Overall the results are less sensitive to viscosity than to the backscatter amplitude, so for the sake of brevity we choose to present results only for $\nu_4 = 2 \times 10^{-5}$. This value gives reasonable results with all types of smoothing operator.

Figure 5.1 shows the time-averaged kinetic and potential energy spectra from simulations with no smoothing; in one simulation the derivatives of the backscatter terms are evaluated using a spectral numerical method, and in the other they are

evaluated using centered second-order finite differences. Both simulations use the same backscatter amplitude $E_0 = 76$. Neither simulation gives good results (and the results can't be qualitatively improved by using different E_0 or ν_4), but the simulation with spectral differencing gives worse results. This simply illustrates the beneficial smoothing effect of finite differences, but shows that this smoothing by itself is insufficient. All remaining simulations in this section, and in §6 and §7, use second order finite differences to approximate the derivatives acting on the Reynolds stress terms.

Figure 5.2 shows the time-averaged kinetic and potential energy spectra from simulations where three different spatial smoothing operators are applied to the Reynolds stress terms: the five-point local average $\mathcal{S}_5$, the three-point local average $\mathcal{S}_3$, and the three-point average applied twice $\mathcal{S}_3^2$. The $\mathcal{S}_5$ simulation uses $E_0 = 69$, the $\mathcal{S}_3$ simulation uses $E_0 = 92$, and the $\mathcal{S}_3^2$ simulation uses $E_0 = 123$; these values are between the median value of 57 and the mean value of 268 diagnosed from the reference simulation. The primary difference in the results is in the kinetic energy spectra, at scales between the peak and the grid scale. As might be expected, the simulation with the least smoothing ($\mathcal{S}_5$) has more kinetic energy at these scales than the simulation with the most smoothing ($\mathcal{S}_3^2$).

All the coarse simulations have relatively accurate kinetic energy spectra, but the potential energy spectra all have a pronounced bump at small scales. This pileup of potential energy at small scales may be related to the dual energy cascade of quasigeostrophic turbulence [34, 36]: potential energy is transferred to scales near the deformation radius, where it converted to kinetic energy before being cascaded back to large scales. In the low resolution model simulations the coarse grid Nyquist wavenumber is $k = 48$, and the dynamics of potential energy conversion near the deformation wavenumber $k_d = 50$ are poorly resolved. Thus, potential energy cascades to small scales, but piles up instead of being converted to kinetic energy. We show in §7 that the full superparameterization closure, which includes terms responsible for absorption of the downscale cascade of potential energy, removes this bump in the potential energy spectrum.

The simulations all have similar heat flux, which is too large. In comparison with the reference simulation's heat flux of 226, $\mathcal{S}_5$ generates 292, $\mathcal{S}_3$ generates 310, and $\mathcal{S}_3^2$ generates 312. The fact that the heat flux is too large for all of these simulations may be related to the bump in the potential energy spectrum, and like that bump it is improved by the full superparameterization algorithm in §7.

The main results of this section are that it is crucial to smooth the k^5 backscatter spectrum in a low-order code, and that good results can be obtained with several easily-implemented local averages. The degree of smoothing applied to the Reynolds stresses is inversely proportional to the value of E_0 required to achieve good results: more smoothing requires larger E_0.

6. Stochastic Backscatter: Temporal Correlation. In this section we incorporate temporal correlation into the backscatter by making the random angle θ a Weiner process on the circle, as described in §2.2. We set the decorrelation time by comparison with the high resolution reference simulation by measuring the decorrelation time of the subgridscale Reynolds stress terms. The measured decorrelation time is very close to the eddy turnover time $(Uk_d)^{-1}$ where U is the RMS velocity, which suggests that the decorrelation time could be set using the eddy turnover time without having to measure it directly. This is useful because the turnover time can be estimated from the resolved large-scale dynamics, without the need to diagnose it from a high resolution reference simulation.

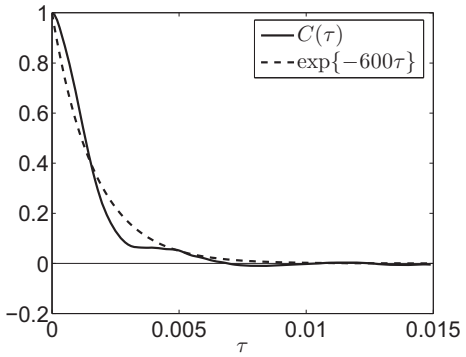

FIG. 6.1. *Empirical lagged autocorrelation for the subgridscale Reynolds stresses (solid) compared to an exponential approximation exp$\{-600\tau\}$ (dashed).*

Specifically, we define a low-pass filter $\overline{(\cdot)}$ (as in §3) by multiplying the Fourier coefficients of the high resolution reference simulation by $1-(k/48)^{12}$ for $k \leq 48$, and by 0 for $k\,48$, where $k = (k_x^2 + k_y^2)^{1/2}$. This has the effect of smoothly transitioning from large scales with $k < 48$ to small scales with $k \geq 48$, where 48 is the Nyquist wavenumber of the coarse model grid. The low-pass filtered fields correspond loosely to the solution of a finite-difference model on a 96×96 grid, and they obey equation (2.6). The backscatter model is intended to approximate the subgridscale vorticity flux divergence, which has the form

$$\nabla \cdot \left(\overline{\boldsymbol{u}_j \zeta_j} - \overline{\boldsymbol{u}}_j \overline{\zeta}_j\right) = \left(\partial_x^2 - \partial_y^2\right)\left(\overline{u_j v_j} - \overline{u}_j \overline{v}_j\right) + \partial_{xy}\left[\left(\overline{(v_j)^2} - \overline{(u_j)^2}\right) - \left(\overline{v}_j^2 - \overline{u}_j^2\right)\right]. \quad (6.1)$$

Time series of the subgridscale Reynolds stress terms $\overline{u_j v_j} - \overline{u}_j \overline{v}_j$ and $(\overline{(v_j)^2} - \overline{(u_j)^2}) - (\overline{v}_j^2 - \overline{u}_j^2)$ were computed from the reference simulation along a line of 16 equispaced points at constant x in each layer. The lagged autocorrelation function was then computed at for each Reynolds stress term at each location. Recall that the lagged autocorrelation function for a zero-mean stationary stochastic process $A(t)$ is defined as

$$C(\tau) = \frac{\mathbb{E}\left[A(t)A(t+\tau)\right]}{\mathrm{Var}\left[A(t)\right]} \quad (6.2)$$

where Var$[\cdot]$ denotes the variance. The expectation $\mathbb{E}$ and the variance are computed using a time average. An approximate decorrelation time is computed from the lagged autocorrelation function as

$$\tau_d = \int_0^\infty C(\tau)\mathrm{d}\tau, \quad (6.3)$$

which is exact for for exponential lagged autocorrelation functions $C(\tau) = \exp\{-\tau/\tau_d\}$.

Figure 6.1 shows the average of all the lagged autocorrelation functions for each Reynolds stress term and at each location compared to an approximate exponential lagged autocorrelation function with $\tau_d = 1/600$, which fits the data reasonably well. This decorrelation time implies a value of $\sigma^2 = 300$ for the temporally correlated

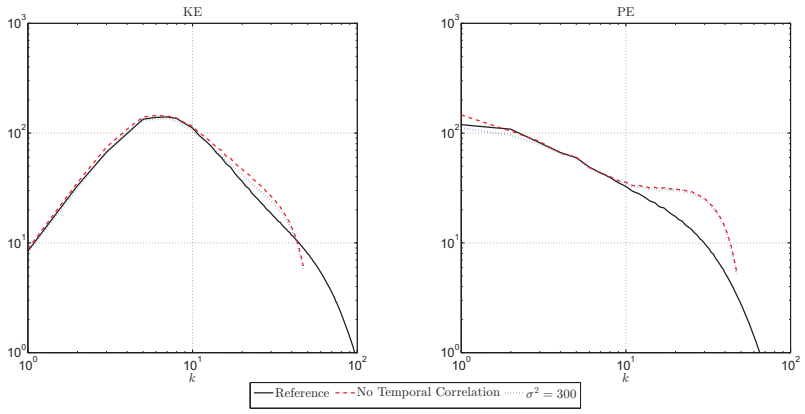

FIG. 6.2. *Time-averaged kinetic (left) and potential (right) energy spectra from simulations with the FD code at $N = 96$, where the backscatter terms are white in time (dashed) or have an exponential decorrelation time (dotted).*

model of θ (equation (2.18), and this value is used in simulations. For comparison, the eddy turnover time $(Uk_d)^{-1}$ is approximately 1/650.

To test the effect of the addition of temporal correlation we run a simulation with the three-point average $\mathcal{S}_3$, time step $\Delta t = 5 \times 10^{-5}$, and viscosity $\nu_4 = 2 \times 10^{-5}$, and with empirical decorrelation time corresponding to $\sigma^2 = 300$. Figure 6.2 compares this simulation to the analogous result with no temporal correlation. The kinetic and potential energy spectra are essentially indistinguishable, but there is a small difference in heat flux: the simulation with temporal correlation has a heat flux of 293, as compared to a heat flux of 310 without temporal correlation (the reference simulation has a heat flux of 226). The main difference is in the optimal backscatter amplitude: the simulation without temporal correlation uses $E_0 = 92$, whereas the simulation with temporal correlation uses $E_0 = 1625$. Part of the discrepancy is related to the fact that E_0 is scaled by the square root of the time step in the white noise scheme, but not in the scheme with temporal correlation.

The robustness of the schemes with and without temporal correlation has been tested by running simulations using time steps half as long and twice as long as those given above. In both cases the results are so similar that the energy spectra are indistinguishable in plots (not shown). Similarly, the heat flux varies by less than 1% as the time step is changed.

Overall the addition of temporal correlation seems to have minimal effect on the simulation results. The optimal backscatter amplitude $E_0 = 1625$ for the scheme with temporal correlation is much too large in comparison with the values diagnosed from the reference simulation (median 57, mean 268). On the other hand the Reynolds stresses diagnosed from the reference simulation clearly have a nonzero decorrelation time. The addition of temporal correlation is clearly not as crucial to success as the addition of spatial correlation, although it remains unclear whether it may have some hidden benefit to the large-scale dynamics.

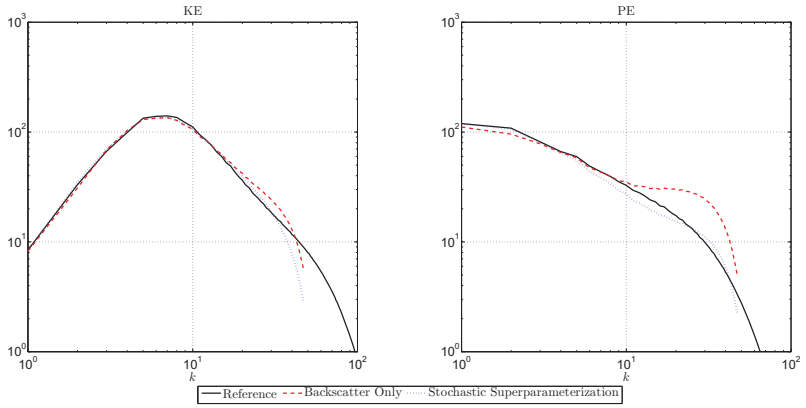

FIG. 7.1. *Time-averaged kinetic (left) and potential (right) energy spectra from simulations with the FD code at $N = 96$, comparing the simplified backscatter scheme (dashed) to the full stochastic superparameterization (dotted).*

7. Stochastic Superparameterization. In this section we incorporate spatial and temporal smoothing into the full stochastic superparameterization scheme of [12, 13]. The primary goal is to demonstrate that the full stochastic superparameterization improves the potential energy spectrum and net heat flux as compared to the simplified backscatter scheme of the previous sections.

Stochastic superparameterization makes use of a formal multiscale ansatz to model the dynamics of the unresolved scales in turbulence; the subgrid dynamics are approximated by local linear stochastic models that respond to the local large-scale velocity and vorticity gradient [12, 13, 23]. It also makes use of a field of random angles, and the subgridscale Reynolds stress terms have the form

$$\begin{aligned} \overline{u_j' v_j'} &= \frac{A}{2} F_{u,j}(\overline{u}_1, \overline{u}_2, \nabla_h \overline{q}_1, \nabla_h \overline{q}_2, \theta; \epsilon) \sin(2\theta), \\ \overline{(v_j')^2} - \overline{(u_j')^2} &= A F_{u,j}(\overline{u}_1, \overline{u}_2, \nabla_h \overline{q}_1, \nabla_h \overline{q}_2, \theta; \epsilon) \cos(2\theta). \end{aligned} \tag{7.1}$$

In the simplified backscatter model the functions $F_{u,j}$ are identical constants (compare to equation (2.10)). Note that the Reynolds stresses in stochastic superparameterization are not identical in the upper and lower layers, like they are in the simplified model. The simplified backscatter model ignores part of the potential vorticity flux divergence $k_d^2 \overline{\boldsymbol{u}_j'(\psi_i' - \psi_j')}$ where $i \neq j$. In stochastic superparameterization these components of the potential vorticity flux take the form

$$\begin{aligned} \overline{u_1'(\psi_1' - \psi_2')} &= A F_b(\overline{u}_1, \overline{u}_2, \nabla_h \overline{q}_1, \nabla_h \overline{q}_2, \theta; \epsilon) \sin(\theta), \\ \overline{v_1'(\psi_1' - \psi_2')} &= A F_b(\overline{u}_1, \overline{u}_2, \nabla_h \overline{q}_1, \nabla_h \overline{q}_2, \theta; \epsilon) \cos(\theta), \\ \overline{u_2'(\psi_2' - \psi_1')} &= -\overline{u_1'(\psi_1' - \psi_2')}, \quad \overline{v_2'(\psi_2' - \psi_1')} = -\overline{v_1'(\psi_1' - \psi_2')}. \end{aligned} \tag{7.2}$$

In stochastic superparameterization the functions $F_{u,j}$ and F_b are computed from solutions of the local stochastic models of the subgridscale dynamics; the stochastic subgrid dynamics evolve for a time length ϵ^{-1} while the large scales are held fixed, and ϵ is considered a tunable parameter. The stochastic subgridscale dynamics are

assumed to evolve on length scales corresponding to wavenumbers with magnitude $k = \sqrt{k_x^2 + k_y^2} \geq 48$, which is the Nyquist wavenumber of the coarse grid. The coefficient A is a tunable parameter analogous to E_0.

We apply the stochastic superparameterization algorithm exactly as described in [13] with the following modifications: (i) the Reynolds stress terms and the potential vorticity flux are smoothed using the three-point average $\mathcal{S}_3$, and their derivatives are approximated with centered second-order finite differences, and (ii) the random angles θ evolve according to equation (2.18) with the empirical value $\sigma^2 = 300$. Figure 7.1 compares the results of the full stochastic superparameterization to the simplified backscatter scheme; both simulations use the empirical decorrelation time corresponding to $\sigma^2 = 300$, and the three-point average $\mathcal{S}_3$. The stochastic superparameterization scheme uses $A = 17500$, and $\epsilon = 800$. The kinetic energy spectra of the simplified scheme and the superparameterization are very similar, differing only near the coarse grid scale. The potential energy spectrum of the superparameterization scheme does not have the bump at small scales, and is a good fit to the potential energy spectrum of the reference simulation. The superparameterization has a net heat flux of 280, which is smaller than the heat flux of 293 from the simplified scheme, but still bigger than the value of 226 from the reference simulation. Overall, the full stochastic superparameterization scheme with spatial and temporal correlations improves over the simplified backscatter scheme. Simulations using stochastic superparameterization without temporal correlation are very similar, provided the amplitude A is retuned (not shown).

8. Conclusions. Kinetic energy backscatter is the process of kinetic energy transfer from small to large scales in turbulence. We have developed stochastic schemes to model backscatter in the inverse kinetic energy cascade of quasigeostrophic turbulence, where backscatter drives a net transfer from small to large scales. The main proposed application of these schemes is in eddy-permitting ocean modeling, where low-resolution models do not resolve the energy transfer from subgrid scales. In addition to improving the accuracy of low-resolution models, stochastic backscatter is expected to improve ensemble filtering and prediction schemes [3, 11].

In §2.1 we showed that a simplified stochastic backscatter scheme from [13], based on stochastic superparameterization, corresponds to an isotropic random forcing with spectrum proportional to k^5 where k is the wavenumber magnitude, i.e. it forces very strongly near the coarse grid scale. This backscatter spectrum is analogous to the backscatter spectrum for smaller-scale three-dimensional isotropic turbulence [19,22,35]. In §5 we showed that the scheme of [13] performs poorly in a low resolution model using second-order finite differences to approximate the nonlinear advection terms, whereas it was shown in [13] that the scheme performed well in a low resolution model with spectral numerics. We argue that the poor performance is due to the fact that the energy backscattered near the coarse grid scale is supposed to be transferred to larger scales by the nonlinear dynamics, but the low-order model is not able to represent this process accurately so energy accumulates near the grid scale, ultimately ruining the simulation.

The use of a low order numerical method requires the numerical solution to be relatively smooth, even in a low resolution model, and we showed in §5 that very good results can be obtained by smoothing the forcing with a local average. This has the effect of damping the backscatter forcing spectrum to zero at the coarse grid scale, so that the backscatter injects energy at a scale where the nonlinear dynamics are better resolved by the low-order numerics. Smoothing methods for non-equispaced grids

developed in [32, 41] could be used in more realistic applications of our backscatter schemes.

The spatially-smoothed backscatter scheme of §5 corresponds to a forcing that is delta-correlated (white) in time. Since the backscatter is supposed to model the nonlinear effects of unresolved deterministic dynamics, it is more realistic for the backscatter model to have a nonzero decorrelation time. We showed in §2.2 and in the Appendix how to add temporal correlation to the backscatter scheme in such a way that the backscatter has an exponential lagged autocorrelation function. The appropriate decorrelation timescale was diagnosed from a high resolution reference simulation in §6, and agreed with the eddy turnover time $(Uk_d)^{-1}$ where U is the RMS eddy velocity and k_d^{-1} is the deformation radius. Incorporation of the nonzero decorrelation time had a minimal impact on the results, although it remains possible that there is some hidden benefit. For example, the short-term prediction and filtering performance of the model might be improved through the use of nonzero temporal decorrelation time.

The stochastic backscatter scheme with spatial and temporal correlation gives good results for the time-averaged kinetic energy spectrum, but the results invariably include a bump in the potential energy spectrum at small scales. This bump is related to the fact that the stochastic kinetic energy backscatter scheme does not include a mechanism for absorbing the downscale cascade of potential energy. In §7 we showed that incorporating spatial smoothing and temporal correlation into the full stochastic superparameterization algorithm of [12, 13] gives excellent results overall, matching both the kinetic and potential energy spectra quite accurately.

Appendix. In this section we prove the following theorem

THEOREM A.1. *The Reynolds stress terms modeled as* $\sin(2\theta)$ *and* $\cos(2\theta)$ *using*

$$d\theta = \sigma\, dW \tag{A.1}$$

have stationary lagged autocorrelation functions of the form $exp\{-2\sigma^2\tau\}$.

Proof. The Fokker-Planck equation for (A.1) is simply

$$\partial_t p = \frac{\sigma^2}{2}\partial_\theta^2 p. \tag{A.2}$$

On a periodic domain of size 2π the invariant measure is uniform $\overline{p} = (2\pi)^{-1}$. The stationary lagged autocorrelation function for the sine term is defined as

$$C(\tau) = 2\mathbb{E}\left[\sin(2\theta_t)\sin(2\theta_{t+\tau})\right] = 2\mathbb{E}\left[\sin(2\theta_t)\sin(2(\theta_t+\Delta_\tau))\right] \tag{A.3}$$

where τ is the lag, θ_t denotes θ at time t, and $\Delta_\tau = \theta_{t+\tau} - \theta_t$. Dependence on t drops out in the stationary state, i.e. the long time limit.

The proof relies on an explicit expression for the probability density of Δ_τ, which is a Dirac delta distribution at zero lag $\tau = 0$. This can be obtained from the Fourier expansion of the Dirac delta distribution and its explicit evolution under the Fokker-Planck equation (A.1). Specifically, the probability distribution for Δ_τ is

$$p(\Delta_\tau) = \frac{1}{2\pi} + \frac{1}{\pi}\sum_{n=1}^{\infty} e^{-n^2\sigma^2\tau/2}\cos(n\Delta_\tau). \tag{A.4}$$

The proof that the lagged autocorrelation function (A.3) is exponential proceeds as follows. First, simple trigonometric identities imply

$$C(\tau) = 2\mathbb{E}\left[\sin(2\theta_t)\left(\sin(2\theta_t)\cos(2\Delta_\tau) + \cos(2\theta_t)\sin(2\Delta_\tau)\right)\right]. \tag{A.5}$$

Second, the non-anticipating property of solutions to (A.1) implies that θ_t is independent of Δ_τ for $\tau > 0$; the above expression therefore simplifies to

$$C(\tau) = 2\left(\mathbb{E}\left[\sin^2(2\theta_t)\right]\mathbb{E}\left[\cos(2\Delta_\tau)\right] + \mathbb{E}\left[\sin(2\theta_t)\cos(2\theta_t)\right]\mathbb{E}\left[\sin(2\Delta_\tau)\right]\right). \tag{A.6}$$

Since we are interested in the stationary lagged autocorrelation function the distribution of θ_t is considered uniform, which implies that the second term above drops out and the expected value of $\sin^2(2\theta_t)$ is $1/2$, leaving

$$C(\tau) = \mathbb{E}\left[\cos(2\Delta_\tau)\right]. \tag{A.7}$$

This expectation can be written as an integral against the known probability distribution of Δ_τ, which gives

$$C(\tau) = \int_{-\pi}^{\pi}\left(\frac{1}{2\pi} + \frac{1}{\pi}\sum_{n=1}^{\infty} e^{-n^2\sigma^2\tau/2}\cos(n\Delta_\tau)\right)\cos(2\Delta_\tau)\mathrm{d}\Delta_\tau = e^{-2\sigma^2\tau}. \tag{A.8}$$

This proves that the stationary lagged autocorrelation function for the sine term is exponential with decorrelation time scale $(2\sigma^2)^{-1}$; the proof for the cosine term is the same. □

REFERENCES

[1] A. Arakawa, *Computational design for long-term numerical integration of the equations of fluid motion: Two-dimensional incompressible flow. Part I*, J. Comp. Phys., 1 (1966), pp. 119–143.

[2] J. Bardina, J. H. Ferziger, and W. C. Reynolds, *Improved subgrid scale models for large eddy simulation*, AIAA Pap., (1980).

[3] J. Berner, G. J. Shutts, M. Leutbecher, and T. N. Palmer, *A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in the ecmwf ensemble prediction system.*, J. Atmos. Sci., 66 (2009).

[4] J.-P. Bertoglio, *A stochastic subgrid model for sheared turbulence*, in Macroscopic modelling of turbulent flows, Springer, 1985, pp. 100–119.

[5] G. Boffetta and R. E. Ecke, *Two-dimensional turbulence*, Annual Review of Fluid Mechanics, 44 (2012), pp. 427–451.

[6] P. Bogacki and L. F. Shampine, *A 3 (2) pair of runge-kutta formulas*, Applied Mathematics Letters, 2 (1989), pp. 321–325.

[7] J. G. Charney, *Geostrophic turbulence*, J. Atmos. Sci., 28 (1971), pp. 1087–1095.

[8] B. Fox-Kemper and D. Menemenlis, *Can large eddy simulation techniques improve mesoscale rich ocean models?*, in Ocean modeling in an eddying regime, vol. 177 of Geophysical Monograph Series, American Geophysical Union, 2008, pp. 319–337.

[9] J. S. Frederiksen and A. G. Davies, *Eddy viscosity and stochastic backscatter parameterizations on the sphere for atmospheric circulation models*, J. Atmos. Sci., 54 (1997), pp. 2475–2492.

[10] M. Germano, U. Piomelli, P. Moin, and W. H. Cabot, *A dynamic subgrid-scale eddy viscosity model*, Phys. Fluids, 3 (1991), pp. 1760–1765.

[11] I Grooms, Y Lee, and AJ Majda, *Ensemble filtering and low resolution model error: Covariance inflation, stochastic parameterization, and model numerics*, Mon. Wea. Rev., (2015). Submitted.

[12] I. Grooms and A. J. Majda, *Efficient stochastic superparameterization for geophysical turbulence*, Proc. Natl. Acad. Sci. USA, 110 (2013).

[13] ———, *Stochastic superparameterization in quasigeostrophic turbulence*, J. Comp. Phys., 271 (2014).

[14] I. Grooms, A. J. Majda, and K. S. Smith, *Stochastic superparameterization in a quasigeostrophic model of the Antarctic Circumpolar Current*, Ocean Model., 85 (2015), pp. 1–15.

[15] Matthew W Hecht, Darryl D Holm, Mark R Petersen, and Beth A Wingate, *Implementation of the LANS-α turbulence model in a primitive equation ocean model*, J. Comp. Phys., 227 (2008), pp. 5691–5716.

[16] Darryl D Holm and Balasubramanya T Nadiga, *Modeling mesoscale turbulence in the barotropic double-gyre circulation*, J. Phys. Ocean., 33 (2003), pp. 2355–2365.
[17] M. F. Jansen and I. M. Held, *Parameterizing subgrid-scale eddy effects using energetically consistent backscatter*, Ocean Model., (2014).
[18] V. Kitsios, J. S. Frederiksen, and M. J. Zidikheri, *Scaling laws for parameterisations of subgrid eddy–eddy interactions in simulations of oceanic circulations*, Ocean Modelling, 68 (2013), pp. 88–105.
[19] R. H. Kraichnan, *Eddy viscosity in two and three dimensions*, J. Atmos. Sci., 33 (1976), pp. 1521–1536.
[20] C. E. Leith, *Stochastic backscatter in a subgrid-scale model: Plane shear mixing layer*, Physics of Fluids A: Fluid Dynamics (1989-1993), 2 (1990), pp. 297–299.
[21] A. Leonard, *Energy cascade in large-eddy simulations of turbulent fluid flows*, in Turbulent Diffusion in Environmental Pollution Proceedings of a Symposium held at Charlottesville, F.N. Frenkiel and R.E. Munn, eds., vol. 18, Part A of Advances in Geophysics, Elsevier, 1975, pp. 237 – 248.
[22] D. C. Leslie and G. L. Quarini, *The application of turbulence theory to the formulation of subgrid modelling procedures*, J. Fluid Mech., 91 (1979), pp. 65–91.
[23] A. J. Majda and I. Grooms, *New Perspectives on Superparameterization for Geophysical Turbulence*, J. Comp. Phys., 271 (2014).
[24] L. Marstorp, G. Brethouwer, and A. V. Johansson, *A stochastic subgrid model with application to turbulent flow and scalar mixing*, Phys. Fluids, 19 (2007), p. 035107.
[25] P. J Mason and D. J. Thomson, *Stochastic backscatter in large-eddy simulations of boundary layers*, J. Fluid Mech., 242 (1992), pp. 51–78.
[26] C. Meneveau and J. Katz, *Scale-invariance and turbulence models for large-eddy simulation*, Annual Review of Fluid Mechanics, 32 (2000), pp. 1–32.
[27] BT Nadiga, *Orientation of eddy fluxes in geostrophic turbulence*, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 366 (2008), pp. 2489–2508.
[28] ———, *Stochastic vs. deterministic backscatter of potential enstrophy in geostrophic turbulence*, in Stochastic Physics and Climate Modeling, TN Palmer, ed., Cambridge University Press, Cambridge, England, 2010.
[29] B. T. Nadiga and F. Bouchet, *The equivalence of the Lagrangian-averaged Navier-Stokes-α model and the rational large eddy simulation model in two dimensions*, Physics of Fluids, 23 (2011), p. 095105.
[30] Balasubramanya T Nadiga and Len G Margolin, *Dispersive-dissipative eddy parameterization in a barotropic model*, J. Phys. Ocean., 31 (2001), pp. 2525–2531.
[31] Balasubramanya T Nadiga and Steve Shkoller, *Enhancement of the inverse-cascade of energy in the two-dimensional Lagrangian-averaged Navier–Stokes equations*, Physics of Fluids, 13 (2001), pp. 1528–1531.
[32] O'Neill, JJ and Cai, X.-M. and Kinnersley, R, *A generalised stochastic backscatter model: large-eddy simulation of the neutral surface layer*, Q J Roy Met Soc, (2015). In press.
[33] PGL Porta Mana and L. Zanna, *Toward a stochastic parameterization of ocean mesoscale eddies*, Ocean Model., 79 (2014), pp. 1–20.
[34] R. Salmon, *Lectures on Geophysical Fluid Dynamics*, Oxford University Press, New York, 1998.
[35] U. Schumann, *Stochastic backscatter of turbulence energy and scalar variance by random subgrid-scale fluxes*, Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences, 451 (1995), pp. 293–318.
[36] R. B. Scott and B. K. Arbic, *Spectral energy fluxes in geostrophic turbulence: Implications for ocean energetics*, J. Phys. Ocean., 37 (2007), pp. 673–688.
[37] R. B. Scott and F. Wang, *Direct evidence of an oceanic inverse kinetic energy cascade from satellite altimetry*, J. Phys. Ocean., 35 (2005), pp. 1650–1666.
[38] G. Soderlind, *Automatic control and adaptive time-stepping*, Numer. Algorithms, 31 (2002), pp. 281–310.
[39] A. F. Thompson and W. R. Young, *Scaling baroclinic eddy fluxes: Vortices and energy balance*, J. Phys. Ocean., 36 (2006), pp. 720–738.
[40] G. K. Vallis, *Atmospheric and Oceanic Fluid Dynamics*, Cambridge University Press, 2006.
[41] S Weinbrecth and PJ Mason, *Stochastic Backscatter for Cloud-Resolving Models. Part I: Implementation and Testing in a Dry Convective Boundary Layer*, J. Atmos. Sci., 65 (2008), pp. 123–139.