

Efficient stochastic superparameterization for geophysical turbulence

Ian Grooms and Andrew J. Majda *

*Department of Mathematics and Climate, Atmosphere, Ocean Science, Courant Institute of Mathematical Sciences, New York, NY 10012

Contributed by Andrew J. Majda

Efficient computation of geophysical turbulence, such as occurs in the atmosphere and ocean, is a formidable challenge for the following reasons: the complex combination of waves, jets, and vortices; significant energetic backscatter from unresolved small scales to resolved large scales; lack of dynamical scale separation between large and small scales; and small-scale instabilities, conditional on the large scales, which do not saturate. Nevertheless, efficient methods are needed to allow large ensemble simulations of sufficient size to provide meaningful quantifications of uncertainty in future predictions and past reanalyses through data assimilation and filtering. Here a class of efficient stochastic superparameterization algorithms is introduced. In contrast with conventional superparameterization, the method here (i) does not require the simulation of nonlinear eddy dynamics on periodic embedded domains, (ii) includes a better representation of unresolved small-scale instabilities, and (iii) allows efficient representation of a much wider range of unresolved scales. The simplest algorithm implemented here radically improves efficiency by representing small-scale eddies at and below the limit of computational resolution by a suitable one-dimensional stochastic model of random-direction plane waves. In contrast to heterogeneous multiscale methods, the methods developed here do not require strong scale separation or conditional equilibration of local statistics. The simplest algorithm introduced here shows excellent performance on a difficult test suite of prototype problems for geophysical turbulence with waves, jets, and vortices with a speedup of several orders of magnitude compared with direct simulation.

waves, jets, vortices | stochastic backscatter | random plane waves

Introduction

One of the foremost challenges of modern applied mathematics is to guide successful methods of accounting for unresolved scales in computational models of multiscale turbulent systems without scale separation. Examples of such systems include atmospheric and oceanic fluid dynamics, stellar- and geodynamos, mantle convection, and confined plasmas, among others. In many of these systems direct resolution of all relevant scales in numerical simulations is impossible given current computers and will remain so for the foreseeable future. The problem is compounded by the need to run large ensembles of simulations to quantify the uncertainty in predictions.

The approach here to modeling the effects of unresolved scales is founded on a multiscale method, called ‘superparameterization’ (SP), developed for capturing the effects of unresolved cloud processes in atmospheric convection [1, 2, 3]. SP deals with unresolved scales by partially resolving them: high-resolution, horizontally periodic computational domains are embedded within the grid cells of a low-resolution global atmospheric model; computational savings are realized by drastically simplifying both the coupling between the large and small scales and the detailed dynamics of the small scales themselves – the embedded domains are reduced to having only one horizontal coordinate. Despite its success in a variety of problems [3, 4, 5], traditional SP is still extremely

expensive, and does not admit straightforward application to other multiscale turbulent systems.

The initial successes of SP, given the drastic simplification of the large-small coupling and of the small-scale dynamics, suggests that further computational savings might be had, without decreasing performance, by making further simplifications of the small-scale dynamics. Xing, Majda, and Grabowski [6] have pursued this line of reasoning by developing sparse space-time SP algorithms using embedded domains that do not fill the spatio-temporal grid of the large scale model. We follow a similar line of reasoning, but in a different direction, pursuing the idea that the small scales might be efficiently modeled stochastically, yet still retaining the multiscale structure of SP. The algorithm presented here is inspired by the mathematical test model for superparameterization of [7], the stochastic Gaussian Closure of [8], and random-direction plane waves in turbulent diffusion [9, 10, 11], and results in a semi-analytical, nonlinear, stochastic closure for the unresolved dynamics based on random sampling of unidirectional, small-scale, unstable plane waves. Unlike conventional SP, our approach does not require computation on embedded domains (although such domains are formally present in the theory), and as a result is extremely computationally efficient. In contrast with other multiscale methods like the heterogeneous multiscale methods (HMM [12]), stochastic SP requires neither spatial nor temporal scale separation, nor conditional equilibration of the small-scale dynamics.

In this article we describe the implementation of stochastic SP in a difficult, paradigm model of geophysical turbulence with an inverse cascade of energy from small to large scales, turbulent dispersive waves, and coherent jets and vortices: two-layer quasigeostrophic (QG) dynamics. The approach is tested in a numerical model whose coarse resolution is such that any parameterization, to be successful, must simultaneously model the stochastic backscatter of kinetic energy from small to large scales in an inverse cascade, and the forward/direct cascade of potential energy from large to small scales. The success in this setting suggests that stochastic SP may have application in fields more diverse than two-layer QG dynamics, for example in atmosphere-ocean modeling, astrophysical turbulence, mantle convection, confined plasmas, etc. – any setting with complex multiscale interactions and turbulent unresolved scales.

Reserved for Publication Footnotes

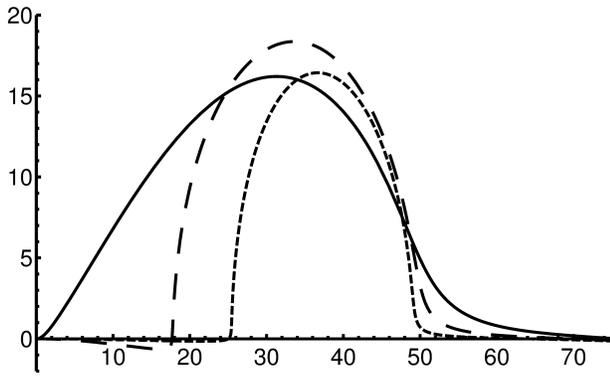


Fig. 1. Growth rates of the most unstable waves as functions of k_x for linear baroclinic instability about the state of rest, $\bar{q}_j = 0$. High latitude (solid), mid latitude (long dash), and low latitude (short dash). The Nyquist wavenumber of the coarse grid in the stochastic SP tests is $|k_x| = 32$, and the deformation wavenumber is $k_d = 50$.

In the next section we review the relevant properties of two-layer QG turbulence and describe high resolution reference simulations. Subsequent sections develop the theory, implementation, and performance of stochastic superparameterization in a coarse-resolution model; the article ends with a brief concluding discussion.

Two-Layer QG Turbulence

We test stochastic SP in the setting of two-equal-layer, rigid-lid, quasigeostrophic turbulence forced by an imposed, baroclinically unstable, horizontally uniform, vertically sheared zonal (x -direction) flow. The governing equations are

$$\begin{aligned} \partial_t q_j &= -\nabla \cdot (\mathbf{u}_j q_j) + (-1)^j \partial_x q_j - \Pi_j \partial_x \psi_j - \delta_{j2} r \nabla^2 \psi_j - \nu \nabla^8 q_j, \\ q_j &= \nabla^2 \psi_j + \frac{k_d^2}{2} (-1)^j (\psi_1 - \psi_2), \quad \mathbf{u}_j = \nabla^\perp \psi_j, \quad j = 1, 2 \end{aligned} \quad [1]$$

where q_j is the potential vorticity in the upper ($j = 1$) and lower ($j = 2$) layers, k_d is the deformation wavenumber (k_d^{-1} is the deformation radius), δ_{j2} is a Kronecker delta, $\Pi_j = \beta - k_d^2 (-1)^j$ is the mean meridional (y -direction) potential vorticity gradient arising from the mean shear ($k_d^2 (-1)^j$) and from the variation of the Coriolis parameter with latitude (β), the coefficient r specifies the strength of linear bottom friction (Ekman drag) and ν is the hyperviscous Reynolds number. The equations are posed in a 2π -periodic domain. When $\beta < k_d^2$ the state of rest $q_j = 0$ is linearly unstable to Rossby waves of the form $q_j = \hat{q}_j \exp\{i(k_x x + k_y y - ct)\}$. The most unstable modes occur for $k_y = 0$; for $\beta \ll k_d^2$ the unstable range is approximately $|k_x| \in (\sqrt{\beta/2}, k_d)$ with peak instability at $|k_x| \approx 0.6k_d$, though modes with $|k_x| \geq k_d$ are slightly destabilized by bottom friction. Growth rates of linear instability for the three model configurations detailed below are shown as functions of k_x with $k_y = 0$ in Fig. 1. The dynamics can also be described in terms of barotropic and baroclinic modes, the former being given by the vertical average $q_t = (q_1 + q_2)/2 = \nabla^2 \psi_t$ and the latter by the vertical difference $q_c = (q_1 - q_2)/2 = (\nabla^2 - k_d^2) \psi_c$. The basic phenomenology for this paradigm model of geophysical turbulence is discussed in [13]: potential energy ($k_d^2 (\psi_1 - \psi_2)^2 / 2$) generated at large scales cascades downscale towards the deformation radius, where it is converted to barotropic kinetic energy. This barotropic kinetic energy then cascades upscale and is absorbed by bottom friction, with the inverse cascade

arrested at a scale determined by β and drag. The dynamics generate a meridional heat flux (proportional to the domain-integral of $v_t \psi_c$) which acts to erode the imposed potential vorticity gradient. Thorough investigations of the parameter space are provided in [14, 15].

Three model configurations are investigated, corresponding to low ($\beta = k_d^2/2$, $r = 1$), medium ($\beta = k_d^2/4$, $r = 4$), and high ($\beta = 0$, $r = 16$) latitudes; the three reference solutions use a resolution of 512 points in each direction, which equals the highest resolution used in [14, 15], and adaptive, fourth-order, semi-implicit Runge-Kutta time integration [16] which treats the hyperviscous terms implicitly. In every simulation $\nu = 1.5 \times 10^{-16}$ and $k_d = 50$; the nonlinear advection terms are dealiased using the 3/2-rule, which means that they are equivalent to simulations at 768^2 using the 2/3-rule – this allows a slightly longer time step.

Figure 2 shows snapshots of the upper-layer potential vorticity q_1 from the three reference simulations. The low- and mid-latitude dynamics organize into six and four zonal (x -direction) jets, respectively, with vortical eddies, filaments, and waves superimposed, and the high-latitude dynamics organize into a sea of vortices and filaments of various sizes.

Stochastic Superparameterization

Theory. We apply a Reynolds average to the governing equations (1) to arrive at the following ‘mean’ equations

$$\begin{aligned} \partial_t \bar{q}_j &= -\nabla \cdot (\overline{\mathbf{u}_j q_j}) + (-1)^j \partial_x \bar{q}_j - \Pi_j \partial_x \bar{\psi}_j - \delta_{j2} r \nabla^2 \bar{\psi}_j - \nu \nabla^8 \bar{q}_j, \\ \bar{q}_j &= \nabla^2 \bar{\psi}_j + \frac{k_d^2}{2} (-1)^j (\bar{\psi}_1 - \bar{\psi}_2), \quad \bar{\mathbf{u}}_j = \nabla^\perp \bar{\psi}_j, \quad j = 1, 2. \end{aligned} \quad [2]$$

The coupling to small scales appears in the mean potential vorticity flux divergence $\nabla \cdot (\overline{\mathbf{u}_j q_j}) = \nabla \cdot (\bar{\mathbf{u}}_j \bar{q}_j) + \nabla \cdot (\overline{\mathbf{u}'_j q'_j})$; the eddy component is given by

$$\begin{aligned} \nabla \cdot (\overline{\mathbf{u}'_j q'_j}) &= \frac{k_d^2 (-1)^j}{2} \nabla \cdot (\overline{\mathbf{u}'_j (\psi'_1 - \psi'_2)}) \\ &+ (\partial_x^2 - \partial_y^2) \overline{u'_j v'_j} + \partial_{xy} \left(\overline{(v'_j)^2} - \overline{(u'_j)^2} \right). \end{aligned} \quad [3]$$

The ‘eddy’ equations are derived simply by subtracting the mean equations (2) from the full equations (1)

$$\begin{aligned} \partial_t q'_j &= -\nabla \cdot (\mathbf{u}'_j q'_j)' - (\bar{\mathbf{u}}_j - (-1)^j \partial_x) \partial_x q'_j \\ &- \mathbf{u}'_j \cdot \nabla \bar{Q}_j - \delta_{j2} r \nabla^2 \psi'_j - \nu \nabla^8 q'_j, \\ q'_j &= \nabla^2 \psi'_j + \frac{k_d^2}{2} (-1)^j (\psi'_1 - \psi'_2), \quad \mathbf{u}'_j = \nabla^\perp \psi'_j, \quad j = 1, 2 \end{aligned} \quad [4]$$

where $\bar{Q}_j = \Pi_j y + \bar{q}_j$.

We impose scale separation by taking the eddy equation (4) to apply on domains embedded at each point of the large-scale domain. This is done by introducing new coordinates \tilde{x} , \tilde{y} , and τ for the embedded domains, requiring the mean variables to have no dependence on the new coordinates, and interpreting the average $\overline{(\cdot)}$ as an average over the new coordinates. The mean variables in the eddy equations are constant and have constant derivatives, as at a point [8, 7]; the approximation is therefore called the ‘point approximation.’ Unlike the SP framework of [1, 2, 17], the point approximation includes horizontal gradients of large-scale quantities in the small-scale equations, allowing correct representation of a wider range of small-scale instabilities (baroclinic instability, for example, is precluded in traditional SP).

The above equations provide a potential foundation for a deterministic SP implementation, where the eddy equations

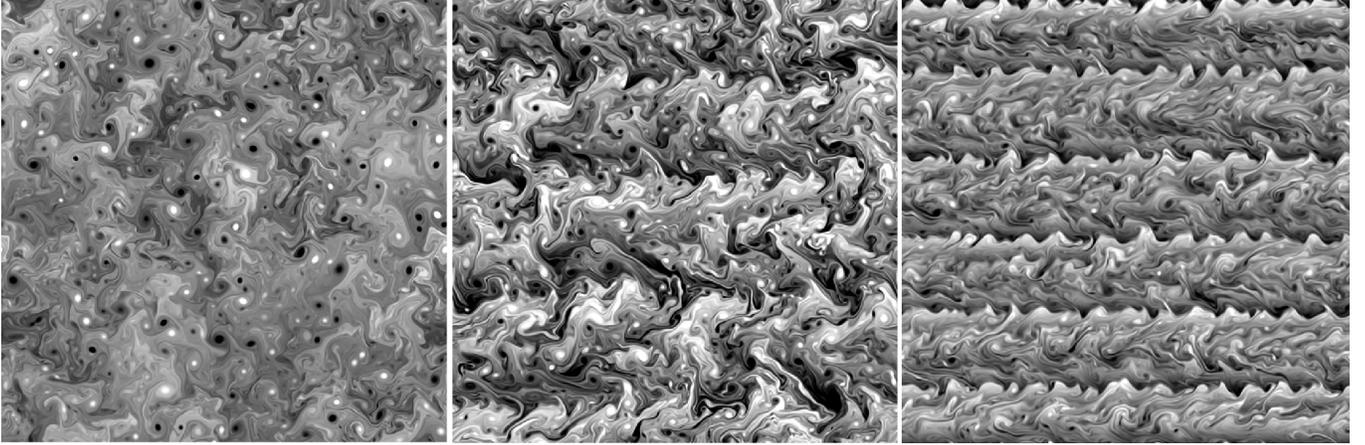


Fig. 2. Snapshots of upper layer potential vorticity q_1 from the reference simulations at high (left), medium (center), and low latitudes (right).

are solved on horizontally periodic embedded domains. However, reduction to one fixed horizontal dimension is not possible here, since the advective nonlinearity responsible for the turbulence reduces to zero in one horizontal coordinate. Furthermore, we consider a coarse resolution grid of 64×64 points with a Nyquist wavenumber of 32. In this case embedded domains that completely fill the large-scale computational grid have a minimum wavenumber of 64, and will not resolve any of the linear instability because $64 > kd$; the sparse-space methods of [6] are thus not applicable to this problem. Rather, the embedded domains in a deterministic SP implementation would have to cover more area than the entire large-scale domain in order to minimally interact with the large scales, resulting in a complete loss of computational efficiency.

To overcome these deficiencies of deterministic SP in this problem, we replace the nonlinear, deterministic eddy equations (4) by the following quasi-linear, stochastic model

$$\partial_\tau q'_j = [F - \Gamma q'_j] - (\bar{\mathbf{u}}_j - (-1)^j) \partial_{\tilde{x}} q'_j - \mathbf{u}'_j \cdot \nabla \bar{Q}_j - \delta_{j2} \tau \tilde{\nabla}^2 \psi'_j - \nu \tilde{\nabla}^8 q'_j \quad [5]$$

where F is additive stochastic forcing and Γ is a positive-definite pseudo-differential operator. This approximation is fundamental to our method, and assumes that the eddies are turbulent; our method should not be expected to work in situations with weakly nonlinear or non-turbulent eddies.

The stochastically-approximated eddy equation (5) has constant coefficients in \tilde{x} and \tilde{y} so the evolution of Fourier modes is decoupled. To overcome the difficulties imposed by using periodic embedded domains (e.g. that the discrete Fourier spectrum may miss unstable eddy modes, as discussed above and in [18, 8]) we represent the eddy variables as homogeneous random functions in formally infinite domains, and make use of the stochastic Fourier transform

$$q'_j = \iint \hat{q}_{j,\mathbf{k}} e^{i\mathbf{k} \cdot \tilde{\mathbf{x}}} dW_{j,\mathbf{k}}. \quad [6]$$

The average $\overline{(\cdot)}$, which includes a spatial average, becomes equivalent to an ensemble average, i.e. $\overline{(\cdot)}$ is a deterministic quantity. Our method therefore produces a deterministic model of the eddy terms in the mean equations; we show below

how to make a stochastic approximation of this deterministic closure whose mean value reduces to the deterministic closure.

The eddy equations for a single Fourier mode are

$$d\hat{q}_{j,\mathbf{k}} = - \left[\gamma_{\mathbf{k}} + \nu k^8 + i(\bar{\mathbf{u}}_j \cdot \mathbf{k} - (-1)^j k_x) \right] \hat{q}_{j,\mathbf{k}} d\tau - \left[(i\mathbf{k} \times \nabla \bar{Q}_j) - \delta_{j2} \tau k^2 \right] \hat{\psi}_{j,\mathbf{k}} d\tau + \sigma_{j,\mathbf{k}} dW_{j,\mathbf{k}} \quad [7]$$

where $k = |\mathbf{k}|$, and $W_{j,\mathbf{k}}$ are independent, complex Weiner processes. We write this as a linear system of Itô stochastic differential equations for $\hat{\psi}_{j,\mathbf{k}}$, and use Itô's lemma to derive a real linear system of four ordinary differential equations for the covariance

$$\begin{aligned} \frac{d}{d\tau} \mathbf{c}_{\mathbf{k}} &= \mathbf{M}_{\mathbf{k}} \mathbf{c}_{\mathbf{k}} + \Sigma_{\mathbf{k}}, \\ \mathbf{c}_{\mathbf{k}} &= \mathbb{E} \left[(|\hat{\psi}_{1,\mathbf{k}}|^2, \mathcal{R}\{\hat{\psi}_{1,\mathbf{k}} \hat{\psi}_{2,\mathbf{k}}^*\}, \mathcal{I}\{\hat{\psi}_{1,\mathbf{k}} \hat{\psi}_{2,\mathbf{k}}^*\}, |\hat{\psi}_{2,\mathbf{k}}|^2) \right] \\ \Sigma_{\mathbf{k}} &= \mathbb{E} \left[(|\sigma_{1,\mathbf{k}}|^2, \mathcal{R}\{\sigma_{1,\mathbf{k}} \sigma_{2,\mathbf{k}}^*\}, \mathcal{I}\{\sigma_{1,\mathbf{k}} \sigma_{2,\mathbf{k}}^*\}, |\sigma_{2,\mathbf{k}}|^2) \right] \quad [8] \end{aligned}$$

where \mathbb{E} denotes the expectation, \mathcal{R} and \mathcal{I} denote the real and imaginary parts of a complex number, and $*$ the complex conjugate. The linear propagator $\mathbf{M}_{\mathbf{k}}$ incorporates the local values and gradients of the large-scale variables, so the eddy statistics will respond to local large-scale conditions; the form of $\mathbf{M}_{\mathbf{k}}$ is given in the supplementary material. We note that the size of $\mathbf{c}_{\mathbf{k}}$ is the square of the number of dependent variables in the system. Systems with more dependent variables (e.g. systems with more vertical layers) will thus have larger $\mathbf{c}_{\mathbf{k}}$, but the form of equation (8) will remain the same.

The utility of this equation stems from the fact that the eddy terms in the mean equation are derivable as integrals over the Fourier covariance, via Plancherel's theorem; for example,

$$\begin{aligned} \overline{u'_1 \psi'_2} &= i \frac{\epsilon k_d^2}{2} \int_0^{\epsilon^{-1}} \iint_{-\infty}^{\infty} \mathbb{E} \left[k_y \hat{\psi}_1^* \hat{\psi}_2 \right] d\mathbf{k} d\tau \\ &= \frac{\epsilon k_d^2}{2} \int_0^{\epsilon^{-1}} \iint_{-\infty}^{\infty} k_y \mathbb{E} \left[\mathcal{I}\{\hat{\psi}_{1,\mathbf{k}} \hat{\psi}_{2,\mathbf{k}}^*\} \right] d\mathbf{k} d\tau. \quad [9] \end{aligned}$$

Thus, to compute the eddy terms in the mean equations one must specify an initial condition for $\mathbf{c}_{\mathbf{k}}$, an integration length for the time average ϵ^{-1} , the autocorrelation of the stochastic forcing $\Sigma_{\mathbf{k}}$, and the additional damping $\gamma_{\mathbf{k}}$.

A 'zero-order' approach is adopted whereby the eddies are required to relax towards a specified equilibrium in the

absence of a mean flow. The equilibrium is set to have an isotropic 1D total energy spectrum (kinetic plus potential) proportional to $k^{-5/3}$ for $k < k_d$ and to k^{-3} for $k > k_d$, which is consistent with the theory of QG turbulence (see, e.g. [13]). In addition, the potential energy spectrum is chosen to be half the kinetic energy spectrum for $k < k_d$ and zero for $k > k_d$ (consistent with the reference simulations), the kinetic energy is equal in each layer, and the equilibrium does not bias the eddy terms. These conditions are sufficient to specify the equilibrium as

$$\mathbf{c}_{k,eq} \equiv \begin{cases} (0, 0, 0) & \text{for } k \leq k_0 \\ A \frac{1}{3k^{14/3}} \left(1, \frac{k_d^2 - k^2}{k_d^2}, 0, 1 \right) & \text{for } k_0 < k \leq k_d \\ A \frac{k_d^{4/3} e^{-\alpha^2(k-k_d)^2}}{3k^6} (1, 0, 0, 1) & \text{for } k > k_d \end{cases} \quad [10]$$

To preclude the unrealistic inclusion of overly large scales on the formally infinite embedded domains, we introduce the large-scale cutoff wavenumber k_0 , beyond which scale the eddies have no variation; a natural but not mandatory choice is to set k_0 equal to the Nyquist wavenumber of the coarse grid. The exponential decay at large k is added to approximate the effect of high-wavenumber damping. The constant of proportionality A is related to the total eddy energy and is considered to be constant on the large-scale domain, although it could be modeled by a large-scale prognostic equation in inhomogeneous settings [19]. The forcing Σ_k is specified by

$$\Sigma_k = -\mathbf{M}_{k,0} \mathbf{c}_{k,eq} \quad [11]$$

where $\mathbf{M}_{k,0}$ denotes \mathbf{M}_k including only bottom friction, hyperviscosity, γ_k and β . To complete the specification of the stochastic model of the eddies we choose $1/\gamma_k$ proportional to the eddy turnover time, given by $1/\sqrt{k^3 E(k)}$ where $E(k)$ is the 1D energy spectrum; thus, $\gamma_k = \gamma_0 e^{-\alpha^2(k-k_d)^2}$ for $k > k_d$ and $\gamma_k = \gamma_0 (k/k_d)^{2/3}$ for $k \leq k_d$. The stochastic approximation in the eddy equation (5) thus specifies F in terms of the Fourier transform of $\sigma_{j,k} dW_{j,k}$. Although γ_k , and hence Γ are completely specified, $\sigma_{j,k}$ and hence F are not completely specified; only the autocorrelation Σ_k appears in the theory.

Although we have developed the theory of stochastic SP for the specific case of two-layer QG dynamics, our methods generalize to a large class of turbulent systems including those with quadratic and cubic nonlinearities as appear in hydro- and magnetohydrodynamics.

Implementation. In the low-resolution experiments we choose $\gamma_0 = 50$, which is strong enough to damp the linear instability of the imposed background shear. This choice is motivated by the finding that the eddy turnover time is faster than the instability timescale in the multiscale analysis of [19], but the results are only weakly sensitive to the choice of γ_0 . The large scale dynamics generate local conditions whose small-scale instability is more than sufficient to overcome the damping in the eddy equations: the small scales do not equilibrate.

The exponential decay scale for the equilibrium spectrum is set to $1/\sqrt{\alpha} = 128$, but the results are not sensitive to this choice. The low-wavenumber cutoff k_0 is set equal to the Nyquist wavenumber of the coarse grid, i.e. $k_0 = 32$. Although this choice implies that there is no formal scale separation between the large- and small-scale dynamics, there is a practical scale separation that results from the fact that the large-scale dynamics near the grid scale are not correctly represented due to truncation errors. One might therefore choose to set k_0 even smaller than the Nyquist wavenumber of the coarse grid, but we do not pursue that approach here. This practical scale separation is exemplified by the need to use a larger coefficient of hyperviscosity ($\nu = 2 \times 10^{-10}$) on the coarse grid to prevent the buildup of grid-scale noise, especially since the

Nyquist wavenumber of the coarse grid is approximately equal to the scale of peak linear instability (see figure 1).

We specify the initial conditions for the eddy covariance to equal the equilibrium forcing. One might alternatively initialize the eddies to zero, or attempt to track the state of the covariance from one large-scale time step to the next, although the latter choice would significantly increase the cost. The time average of the covariance evolution is given by

$$\epsilon \int_0^{\epsilon^{-1}} \mathbf{c}_k(\tau) d\tau = \left[\phi_1(\mathbf{M}_k/\epsilon) - \frac{1}{\epsilon} \phi_2(\mathbf{M}_k/\epsilon) \mathbf{M}_{k,0} \right] \mathbf{c}_{k,eq} \quad [12]$$

$$\phi_1(\mathbf{A}) = \mathbf{A}^{-1} \left[e^{\mathbf{A}} - \mathbf{I} \right], \quad [13]$$

$$\phi_2(\mathbf{A}) = \mathbf{A}^{-2} \left[e^{\mathbf{A}} - \mathbf{I} - \mathbf{A} \right] = \mathbf{A}^{-1} \left[\phi_1(\mathbf{A}) - \mathbf{I} \right]. \quad [14]$$

This result assumes \mathbf{M}_k to be nonsingular, which is true except on a set of measure zero in \mathbf{k} , which does not affect the value of the integrals over \mathbf{k} that define the eddy terms (e.g. equation 9). Furthermore, the choice of γ_k can always be altered on a set of measure zero to render \mathbf{M}_k nonsingular for all \mathbf{k} , although this is not necessary. The length of the time average is just over twice the length of the time-step used in the coarse equations, i.e. $\epsilon^{-1} = 5 \times 10^{-4}$; this allows extra time for the eddies to respond to the local mean, since we are re-initializing the eddies at each time step. Although we have not performed a full sensitivity analysis, we note that performance degrades with averaging times much shorter than the coarse-grid time step.

Calculation of the eddy terms by the Fourier integrals over \mathbf{k} , as in equation (9), results in a deterministic, nonlinear closure for the eddies in terms of the mean variables which, if implemented directly, is still expensive due to the quadrature required. (Although such an approach is much less expensive than traditional SP.) Additionally, it is often advantageous to include an element of stochasticity in models of unresolved eddies, particularly when modeling stochastic backscatter from unresolved scales into resolved ones. We reduce the cost and randomize the algorithm by computing the Fourier integrals that define the eddy terms using a random integration method based on sampling unidirectional plane waves with random directions. Specifically, we re-write the 2D integrals in polar coordinates and integrate in k along one azimuthal direction which is randomly chosen at each coarse grid point and time step. We approximate the polar eddy Fourier integrals by a midpoint-rule quadrature in k using nodes with integer values $k = 32, \dots, 256$ (from the Nyquist wavenumber of the coarse grid to the Nyquist wavenumber of the DNS). The deterministic closure can be reproduced by integrating the polar Fourier integral along a large number of azimuthal directions chosen randomly from a uniform distribution; we present results only for integrals in one randomly chosen azimuthal direction at each coarse-grid point. The use of one-dimensional integrals is similar to the conventional SP practice of using reduced dimensional embedded domains, although such a practice would not work in this setting, as noted earlier. The use of one-dimensional integrals is also motivated by the success of random plane waves in models of turbulent diffusion [9, 10, 11].

We further reduce the cost by noting that the Fourier integrals depend on the mean variables only through three scalar parameters: $\bar{\mathbf{u}}_c \cdot \hat{\mathbf{k}}$, $\nabla \bar{\omega}_c \times \hat{\mathbf{k}}$, and $\nabla \bar{\omega}_t \times \hat{\mathbf{k}}$ where $\hat{\mathbf{k}}$ is a unit vector in the direction of the azimuthal Fourier line integral and $\bar{\omega}_j = \nabla^2 \bar{\psi}_j$ is the local large-scale vorticity. The mean barotropic velocity $\bar{\mathbf{u}}_t$ has no impact on the eddy fluxes since it amounts to uniform translation in the eddy domains. Rather

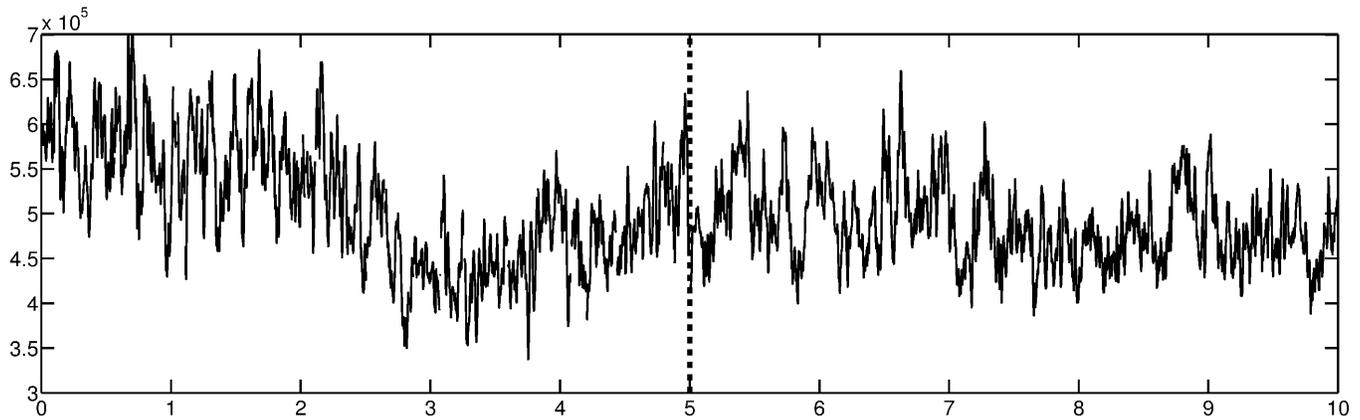


Fig. 3. Time series of the heat flux generated by high-resolution DNS (left) and stochastic SP (right) solutions at high latitude ($\beta = 0$, $r = 16$).

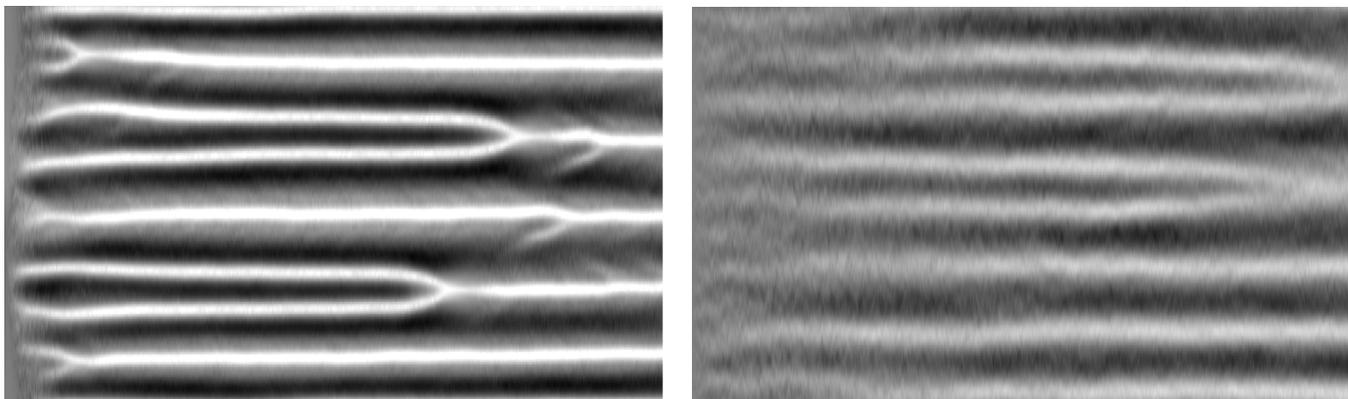


Fig. 4. Time-series of zonally-averaged zonal barotropic velocity at low latitude ($\beta = k_d^2/2$, $r = 1$) from the high-resolution DNS (left) and stochastic SP (right)

than repeatedly calculate similar values for the eddy terms at every time step and grid point of the coarse simulation, we pre-compute the eddy terms as functions of the three scalar parameters, using 101 equispaced nodes for each scalar parameter. The range of values of the three scalar parameters over which solutions are pre-computed is chosen to encompass the variation seen in the low-resolution simulations. The eddy terms in the mean equation are evaluated using linear interpolation based on these pre-computed values.

The large-scale equations (2) are solved using the same methods as the high-resolution reference simulations, but on a grid of 64×64 points and with a fixed time step of 2×10^{-4} . The eddy terms are evaluated using new random directions at the beginning of each time step, and are held constant for the duration of the step. The inflated hyperviscous Reynolds number $\nu = 2 \times 10^{-10}$ and the time step 2×10^{-4} are kept the same in all three test cases, leaving A as the only tunable parameter. After minimal tuning, the results of the stochastic superparameterization algorithms are presented for the low-latitude case using $A = 1.5 \times 10^3$, for the mid-latitude case using $A = 6 \times 10^3$, and for the high-latitude case using $A = 2 \times 10^4$.

Results. The most striking feature of the high-latitude test case ($\beta = 0$, $r = 16$) is the appearance of strong vortex cores (Fig. 2, left panel) which are unresolved on the coarse grid (Fig. S1, left panel). However, the net poleward heat flux (proportional to the domain integral of $v_t \psi_c$) is generated primarily by dynamics at larger scales, as discussed by [14], and these scales are resolved on the coarse grid. Figure 3

demonstrates that the time series of the heat flux generated by the coarse-resolution stochastic SP solution in the high-latitude case (right half) has a nearly identical character to that generated by the high-resolution reference solution (left half), despite the complete lack of small-scale vortex cores on the coarse grid. The time-averaged 1D energy spectra of the high-resolution and stochastic SP solutions also show good agreement (Fig. S2).

The reference solution in the mid-latitude test case ($\beta = k_d^2/4$, $r = 4$) includes strong vortex cores and intermittently broken barotropic zonal jets (Fig. 2, center panel); these jets constitute a barrier to transport and limit the poleward (meridional) heat flux, resulting in a net flux an order of magnitude smaller than the high-latitude simulation (Fig. S3). Figure 5 compares the time-averaged 1D energy spectra for the stochastic SP solution and the reference solution at mid-latitudes, demonstrating good agreement between kinetic and potential energies, and a peak at $k = 4$ corresponding to the four barotropic zonal jets that develop. Although the total energy content is similar, the peak of the kinetic energy spectrum in the stochastic SP solution is weaker than the reference solution; the jets in the stochastic SP solution are also more intermittent (Fig. S4). The result is that the heat flux generated by the stochastic SP solution is about 50% too large (Fig. S3). Generating a heat flux correct to within 50% on a grid with 1/64 as many points constitutes a resounding success; it is likely, however, that the result could be improved by allowing variation of the A across the large-scale domain.

The low-latitude case ($\beta = k_d^2/2$, $r = 1$) is particularly difficult because (i) the band of linear instability responsi-

ble for the turbulence is almost completely unrepresented on the coarse grid (see Fig. 1; indeed, the hyperviscous diffusion is sufficient to completely stabilize the linear instability on the coarse grid) and (ii) the coarse grid allows only approximately ten points per pair of east-west jets, making a single jet barely resolved. Figure 4 compares the time evolution of the zonally-averaged zonal barotropic velocity $\int u_t dx$ from the stochastic SP and reference solutions in the low latitude test case. The reference simulation initially develops eight jets, but later transitions to six, while the stochastic SP simulation initially develops seven jets and later transitions to five. As with the mid-latitude case, the jets comprise a strong barrier to heat transport, and the net transport is two hundred times smaller than the high-latitude case (Fig. S5). Nevertheless, the stochastic SP solution is able to correctly predict the heat transport to within about 30%. In addition, the time-averaged 1D spectra are closely matched (Fig. S6).

Discussion and Conclusions

The main result of this article is the development of an efficient method for the stochastic parameterization of unresolved scales in a multiscale, turbulent dynamical system. The method is based on the ideas of superparameterization [1, 2], stochastic eddy modelling [8], and on the test model for superparameterization of [7]. In contrast with conventional SP, (i) our method does not require the solution of prognostic eddy equations on embedded domains, although such domains are formally present in the theory, (ii) includes a better representation of certain small-scale instabilities including baroclinic instability, which is the primary driver of atmospheric and oceanic variability, and (iii) allows efficient representation of a much wider range of eddy scales through the use of formally infinite embedded domains. Compared with HMM [12], our method does not require scale separation in space or time, and does not require the equilibration of local eddy statistics. The stochastic nature of the parameterization suggests that it will improve the ability of coarse-resolution models to assimilate observational data [18, 20, 21], and the computational efficiency enables larger ensemble sizes for improved uncertainty quantification. The success of our approach in simultaneously parameterizing the quasigeostrophic inverse kinetic

energy cascade from small scales and forward potential energy cascade to small scales, and the flexibility of the stochastic superparameterization framework suggest that the method may prove useful in a wide range of multiscale, turbulent dynamical systems. The implementation here through random plane waves with re-initialization of the eddies at each large-scale time step is the simplest of several strategies that retain computational efficiency. Other options for implementing stochastic superparameterization, as suggested here, should be pursued in the future.

ACKNOWLEDGMENTS. The work of I. G. is supported by the NSF **Collaboration in Mathematical Geosciences** program, grant DMS-1025468. The work of A. M. is partially supported by Office of Naval Research grant N00014-11-1-0306.

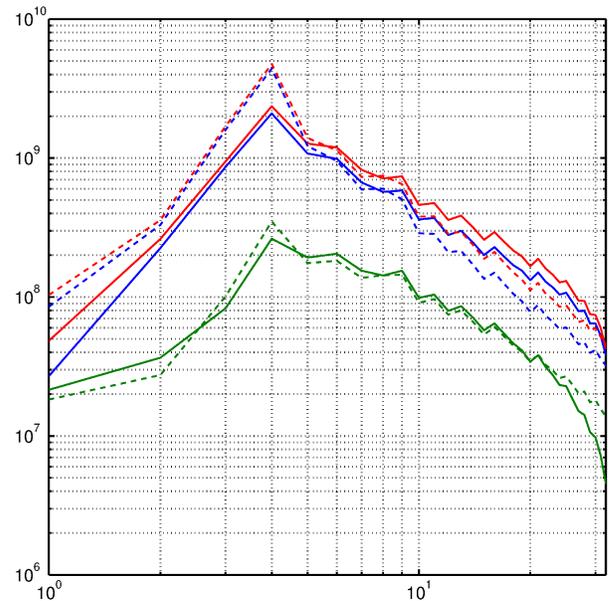


Fig. 5. Time-averaged 1D energy spectra for the high-resolution DNS (dashed) and stochastic SP (solid) solutions at midlatitude ($\beta = k_d^2/4$, $r = 4$). Total energy in red, kinetic in blue, and potential in green.

- Grabowski WW, Smolarkiewicz PK (1999) CRCP: A cloud resolving convection parameterization for modeling the tropical convective atmosphere. *Physica D* 133:171-178.
- Khairoutdinov M, Randall DA (2001) A cloud resolving model as a cloud parameterization in the NCAR community climate system model: Preliminary results. *Geophys. Res. Lett.* 28:3617-3620.
- Khairoutdinov M, Randall DA, DeMott C (2005) Simulations of the atmospheric general circulation using a cloud-resolving model as a superparameterization of physical processes. *J. Atmos. Sci.* 62:2136-2154.
- Khairoutdinov M, DeMott C, Randall DA (2008) Evaluation of the simulated interannual and subseasonal variability in an AMIP-style simulation using the CSU multiscale modeling framework. *J. Clim.* 21:413-431.
- Campin J-M, Hill C, Jones H, Marshall J (2011) Super-parameterization in ocean modeling: Application to deep convection. *Ocean Mod.* 36:90-101.
- Xing Y, Majda AJ, Grabowski WW (2009) New efficient sparse space-time algorithms for superparameterization on mesoscales. *Mon. Wea. Rev.* 137:4307-4324.
- Majda AJ, Grote MJ (2009) Mathematical test models for superparameterization in anisotropic turbulence. *Proc. Natl. Acad. Sci. U.S.A.* 106:5470-5474.
- Majda A (2012) Challenges in climate science and contemporary applied mathematics. *Commun. Pur. Appl. Math.*, 65:920-948.
- Elliott FW Jr, Majda AJ (1995) A new algorithm with plane waves and wavelets for random velocity fields with many spatial scales. *J. Comp. Phys.* 117:146-162.
- Majda AJ (1994) Random shearing direction models for isotropic turbulent diffusion. *J. Stat. Phys.* 75:1153-1165.
- Majda AJ, Kramer PR (1999) Simplified models for turbulent diffusion: Theory, numerical modelling, and physical phenomena. *Physics Reports* 314:237-574.
- E W, Engquist B (2003) The heterogeneous multiscale methods *Comm. Math. Sci.* 1:87-132.
- Salmon R (1998) *Lectures on Geophysical Fluid Dynamics* Oxford University Press, New York.
- Thompson AF, Young WR (2006) Scaling baroclinic eddy fluxes: Vortices and energy balance. *Journal of Physical Oceanography* 36:720-738.
- Thompson AF, Young WR (2007) Two-layer baroclinic eddy heat fluxes: Zonal flows and energy balance. *J. Atmos. Sci.* 64:3214-3231.
- Kennedy CA, Carpenter MH (2003) Additive Runge-Kutta schemes for convection-diffusion-reaction equations. *Applied Numerical Mathematics* 44:139-181.
- Grabowski WW (2004) An improved framework for superparameterization. *J. Atmos. Sci.* 61:1940-1952.
- Harlim J, Majda AJ (2013) Test models for filtering with superparameterization. *Multiscale Modelling & Simulation* in press.
- Grooms I, Majda AJ, Smith KS (2012) Multiscale models for synoptic-mesoscale interactions in the ocean. *Dynamics of Atmospheres and Oceans* 58:95-107.
- Majda AJ, Harlim J (2012) *Filtering Complex Turbulent Systems* Cambridge University Press.
- Majda AJ, Harlim J, Gershgorin B (2010) Mathematical strategies for filtering turbulent dynamical systems. *Discrete Cont. Dyn. S.* 27:441-486.