Backscatter in energetically-constrained Leith parameterizations

Ian Grooms

Department of Applied Mathematics, University of Colorado, Boulder, CO 80309.

Abstract

This paper combines two strains in the literature on subgrid-scale parameterizations for eddying ocean models to develop six new parameterizations and test them in an idealized quasigeostrophic model. The first strain develops nonlinear, Smagorinsky-like Leith scalings for the viscous coefficient based on reasoning about the turbulent forward enstrophy cascade of geophysical turbulence. The second introduces backscatter whose amplitude is scaled to re-inject a portion of the energy dissipated by other parameterizations. In the new parameterizations developed here the backscatter and viscous coefficients depend on each other and are set to simultaneously absorb the forward enstrophy (or potential enstrophy) cascade and backscatter a portion of the dissipated energy. The addition of backscatter to Leith-scaled nonlinear viscosity improves the simulations at resolutions from 4 km to 24 km by increasing the total kinetic energy and by reducing the fraction of the total energy dissipation rate associated with the net effect of viscosity and backscatter. Versions that use biharmonic viscosity to absorb the enstrophy cascade perform better than versions using a harmonic viscosity, and purely viscous closures perform better than closures that dissipate both kinetic and potential energy. Stochastic and deterministic backscatter schemes are developed, and though similar, the deterministic schemes perform slightly better.

Keywords: Backscatter, Leith viscosity, Stochastic

Email address: ian.grooms@colorado.edu (Ian Grooms)

1. Introduction

Modeling the effects of unresolved length scales in simulations of fluid flow is crucially important for accurate predictions in many contexts. Smagorinsky (1963) developed an early model – a parameterization – based on Kolmogorov's theory of three-dimensional homogeneous isotropic turbulence. In that theory, the fluid is forced at a large scale, and nonlinear interactions transfer energy through an inertial range of scales to a dissipation scale where energy is removed by viscosity. The cascade rate ϵ is the rate of energy transfer, and has dimensions of length squared over time cubed, i.e. L^2/T^3. (Here and throughout, density is assumed constant and the total mass is scaled out of kinetic energy, leaving a 'kinetic energy' with dimensions of velocity squared.) If one assumes that the scale at which viscous dissipation arrests the energy cascade depends only on the cascade rate ϵ and the viscosity ν, dimensional analysis predicts that this dissipation scale, called the Kolmogorov scale, is proportional to $(\nu^3/\epsilon)^{1/4}$. Smagorinsky's parameterization applies when the grid scale lies in the inertial range of scales; the parameterization increases the viscosity ν until the resulting Kolmogorov scale is resolved on the grid.

The behavior of two-dimensional (2D) turbulence is qualitatively different from that of three-dimensional turbulence: Energy tends to be transferred from the forcing scale to larger scales, while enstrophy (the square of vorticity) is transferred towards smaller scales. Leith (1996) developed a parameterization for 2D turbulence that is conceptually similar to Smagorinsky's parameterization, but is based on 2D turbulence theory. The cascade rate η for enstrophy has dimensions of T^{-3}, and the viscous dissipation scale is proportional to $(\nu^3/\eta)^{1/6}$. Leith's parameterization applies when the grid scale lies in the enstrophy inertial range of scales; it increases the viscosity ν until the dissipation scale for enstrophy is resolved on the grid.

The oceans' aspect ratio and density stratification conspire to make the large-scale circulation nearly 2D, so Fox-Kemper and Menemenlis (2008) developed Leith's parameterization for ocean models. The quasigeostrophic (QG) approximation is less restrictive and more appropriate than the 2D approximation, and has its own turbulence theory which replaces the enstrophy cascade with a potential enstrophy cascade (Charney, 1971). Bachman et al. (2017) developed a parameterization in the vein of Leith's parameterization, but based on a cascade of potential enstrophy in QG turbulence; to distinguish this parameterization from the earlier ones, the one based on 2D turbulence is called 2D-Leith while the one based on QG turbulence is called QG-Leith. The performance of these closures in an eddying global ocean model was explored by Pearson et al. (2017).

Viscous parameterizations, both nonlinear viscosities like the Leith parameterization and constant biharmonic hyperviscosities (Semtner and Mintz, 1977; Böning and Budich, 1992), tend to dissipate too much energy. The problem can be compounded by poor numerical resolution of the baroclinic instability processes that are largely responsible for ocean mesoscale eddy generation (Barham et al., 2018; Barham and Grooms, 2019). A similar problem of over-dissipative closures occurs in atmospheric models, and Shutts (2005) developed a parameterization that re-injects energy to offset the energy spuriously dissipated by other parameterizations. Shutt's parameterization is a backscatter parameterization because it transfers energy into the resolved scales. Backscatter parameterizations whose primary aim is to mitigate the spurious dissipation of energy resulting from numerics and other parameterizations were further developed in the atmospheric modeling context by Berner et al. (2008, 2009), among others, and were introduced in ocean modelling by Jansen and Held (2014). The Stochastic Kinetic Energy Backscatter Scheme (SKEBS) of Berner et al. (2009) was recently adapted for ocean models by Storto and Andriopoulos (2021).

Instantaneous backscatter occurs in the forward cascade of 3D turbulence even when the mean energy transfer is from large to small scales, and it also occurs in the forward enstrophy cascade range of 2D and QG turbulence, where the mean energy transfer is very small. Backscatter parameterizations that attempt to represent physical backscatter processes have appeared throughout the turbulence modeling community starting from the work of Bertoglio (1985) and Leith (1990), and in the context of geophysical models from at least the work of Mason and Thomson (1992). By 2005, physical backscatter parameterizations had been developed for atmospheric models by Schumann (1995) and Frederiksen and Davies (1997), among others. Shutt's backscatter parameterization differs from these in that it is not a parameterization of a physical backscatter process per se; rather, it uses backscatter to mitigate spurious dissipation that arises in other parts of the model. Within the ocean modelling context, a popular approach to parameterizing backscatter sets the amplitude of backscatter based on a prognostic model of subgrid-scale kinetic energy (e.g. Jansen et al., 2015; Klöwer et al., 2018; Jansen et al., 2019; Juricke et al., 2019, 2020b). Whether these particular backscatter parameterizations are explicitly designed to represent a physical backscatter process or whether they are designed to make up for the energetic failings of other parameterizations is not always clear, nor does the distinction necessarily have a large practical import, since backscatter parameterizations in practice end up accomplishing, to some extent, both ends. In contrast, Bachman (2019) developed a backscatter parameterization that is clearly physical in nature. It ties the backscatter rate to a dissipation rate, but instead of re-injecting energy to correct spurious dissipation the backscatter represents a physical process whereby potential energy that is transferred to unresolved scales via baroclinic instability cascades back to the resolved scales via the QG inverse cascade.

Backscatter parameterizations can also be categorized based on the functional form of the parameterization: Backscatter parameterizations are either deterministic or stochastic, or in rare cases, a combination of both. The earliest backscatter parameterizations (e.g. Bertoglio, 1985; Leith, 1990; Mason and Thomson, 1992) were stochastic. Inspiration for a deterministic parameterization of backscatter goes back to at least Kraichnan (1976), whose work inspired Sukoriansky et al. (1996) to develop a parameterization for 2D turbulence in the form of a negative viscosity combined with a biharmonic hyperviscosity. In the context of ocean models, Kitsios et al. (2013) and Jansen and Held (2014) developed deterministic backscatter parameterizations based on combinations of viscosity and hyperviscosity. The backscatter scheme described by Shutts (2005) is also deterministic, but is based on the chaotic dynamics of cellular automata rather than on a negative viscosity.

The goal of this paper is to develop new parameterizations that combine the Smagorinsky-like 2D- and QG-Leith parameterizations with backscatter in a theoretically consistent way by building on the basic theory of the 2D and QG turbulent enstrophy cascade, and to compare a wide range of parameterizations in an idealized model across a range of resolutions. The two primary threads in the theoretical development pursued here, viz. absorbing a downscale cascade of enstrophy or potential enstrophy and setting the backscatter amplitude to re-inject a portion of the dissipated energy, have appeared in the literature, as have the components of the implementation (harmonic and biharmonic operators, stochastic backscatter, spatial filters). The practical effect of combining these two theoretical threads is that the amplitudes of backscatter and dissipation are set simultaneously, rather than setting the dissipation coefficient independently of backscatter. The parameterizations are developed in section 2, and the idealized model results are described in section 3.

2. Energetically-constrained Leith parameterizations

The parameterizations developed here rely on quasigeostrophic (QG) theory, but are intended for implementation in primitive-equation ocean models. To clarify the connection, the QG vorticity and buoyancy-anomaly equations are recorded here

$$\partial_t \omega + \boldsymbol{u} \cdot \nabla \omega - f_0 \partial_z w + \beta v = \mathcal{B}_\omega + \mathcal{D}_\omega, \tag{1}$$

$$\partial_t b + \boldsymbol{u} \cdot \nabla b + w N^2(z) = \mathcal{B}_b + \mathcal{D}_b. \tag{2}$$

The vertical component of relative vorticity is $\omega = \partial_x v - \partial_y u = \nabla^2 \psi$ where u and v are the zonal and meridional components of velocity, respectively, and ψ is the QG streamfunction. The gradient ∇ acts horizontally, and w is the vertical component of velocity. The Coriolis parameter f_0 is twice the angular rotation rate of the planet projected onto the local vertical direction, and β is the meridional gradient of planetary vorticity. The mean buoyancy is $\bar{b}(z)$ and the buoyancy frequency is $N(z) = (\partial_z \bar{b})^{1/2}$; the buoyancy anomaly is related to the streamfunction via $b = f_0 \partial_z \psi$. The terms $\mathcal{B}_{\omega,b}$ and $\mathcal{D}_{\omega,b}$ represent backscatter and dissipation terms, respectively.

The vorticity and buoyancy-anomaly equations can be combined into a single equation for QG potential vorticity

$$\partial_t q + \boldsymbol{u} \cdot \nabla q + \beta v = \mathcal{B}_\omega + \mathcal{D}_\omega + \partial_z \left[\frac{f_0}{N^2} \left(\mathcal{B}_b + \mathcal{D}_b \right) \right] \tag{3}$$

where

$$q = \nabla^2 \psi + \partial_z \left(\frac{f_0^2}{N^2} \partial_z \psi \right) \tag{4}$$

is the QG potential vorticity (QG PV). The energy conserved by the dynamics is a combination of kinetic and available potential energies

$$\mathrm{E} = \mathrm{KE} + \mathrm{APE} = \frac{1}{2} \int_V \left[|\nabla \psi|^2 + \frac{f_0^2}{N^2} \left(\partial_z \psi \right)^2 \right] \mathrm{d}V. \tag{5}$$

The parameterizations appearing in the vorticity equation directly affect the kinetic energy budget, while the parameterizations appearing in the buoyancy-anomaly equation directly affect the available potential energy budget.

The following subsections develop deterministic and stochastic parameterizations that consistently incorporate backscatter into the 2D and QG

Leith parameterizations. In all cases it is crucial to achieve a scale separation between the backscatter and the dissipation: any energy backscattered at the grid scale is immediately dissipated, which defeats the purpose, and nonlinear processes that might otherwise transfer energy from the grid scale to larger scales are badly resolved by discretizations. The deterministic parameterizations developed here achieve this scale separation by using a combination of a dissipative biharmonic term and a backscattering harmonic term following Sukoriansky et al. (1996) and Jansen and Held (2014), while the stochastic parameterizations use either a harmonic or biharmonic dissipation term in combination with a stochastic backscatter that is designed to avoid backscattering at the grid scale.

The derivations in the following subsections follow the traditional approach to deriving the Smagorinsky and Leith parameterizations. Expressions for the viscous coefficients are derived under the assumption that these coefficients are constant and that the turbulence is homogeneous, but the derived expressions are then used to define spatially-varying viscous coefficients for use in simulations of inhomogeneous dynamics. The gap between the assumptions of the derivation and the use of the expressions is mitigated somewhat by smoothing the coefficients before use (cf. (28)).

2.1. Biharmonic 2D Leith + E

The biharmonic 2D Leith parameterization sets the vorticity dissipation term to

$$\mathcal{D}_\omega = -\nabla^2 \left(\nu_4 \nabla^2 \omega\right). \tag{6}$$

(In a primitive-equation model one might use $\nabla^2(\nu_4\nabla^2\boldsymbol{u})$ or $\nabla\cdot(\sqrt{\nu_4}\nabla(\nabla\cdot(\sqrt{\nu_4}\nabla\boldsymbol{u})))$.) Re-injection of spuriously-dissipated energy is accomplished using a negative-viscosity harmonic operator

$$\mathcal{B}_\omega = \overline{\nabla^2(\nu_2\overline{\omega})} \tag{7}$$

where $\nu_2 \leq 0$. In a primitive-equation model one might use $\nabla\cdot\overline{(\nu_2\nabla\overline{\boldsymbol{u}})}$. There is no backscatter or dissipation in the buoyancy-anomaly equation: $\mathcal{B}_b = \mathcal{D}_b = 0$. The overline in (7) represents a self-adjoint low-pass spatial filter whose purpose is to enhance the scale separation between the dissipative action of the biharmonic term and the backscattering action of the harmonic term. The number of applications of the filter could be adjusted if needed to enhance scale separation.

To motivate the spatial filter, consider an equispaced grid with grid scale Δx. Ignoring discretization errors and spatial variability of the viscous coefficients, the combined influence of the (unfiltered) harmonic and biharmonic terms on a discrete Fourier mode with wavenumber $\boldsymbol{k}$ is a linear growth (or decay) with rate

$$-\nu_2 k^2 - \nu_4 k^4 \tag{8}$$

where $k = |\boldsymbol{k}|$. Wavenumbers with $0 < k < (-\nu_2/\nu_4)^{1/2}$ are forced while wavenumbers with $k > (-\nu_2/\nu_4)^{1/2}$ are damped. In order to successfully absorb an enstrophy cascade, one wants a reasonably-wide range of scales to be damped; e.g., one might want to ensure that all modes within a factor of two of the grid scale are damped. This sets a limit on the allowable values of the viscous coefficients. The smallest unambiguously representable Fourier mode on the grid has wavenumber $\pi/\Delta x$ in each direction, so the requirement that the crossover between forcing and damping occurs at a wavenumber less than half the grid wavenumber leads to the following constraint on the viscous coefficients

$$\left(-\frac{\nu_2}{\nu_4}\right)^{1/2} < \frac{\pi}{2\Delta x} \Rightarrow -\nu_2 < \nu_4 \left(\frac{\pi}{2\Delta x}\right)^2. \tag{9}$$

This effectively limits the allowable backscatter rates.

The filter allows this constraint to be relaxed. A simple moving-average filter with a three-point stencil in each direction with weights [1/4, 1/2, 1/4] multiplies a discrete Fourier mode with wavenumber $\boldsymbol{k} = (k_x, k_y)$ by a factor of $\cos(k_x\Delta x/2)\cos(k_y\Delta x/2)$, i.e. it leaves the largest scales unchanged and zeroes out the smallest scales. The standard second-order discretization of the Laplacian multiplies a discrete Fourier mode with wavenumber $\boldsymbol{k} = (k_x, k_y)$ by a factor of $-(4/\Delta x^2)(\sin^2(k_x\Delta x/2)+\sin^2(k_y\Delta x/2))$. For this specific filter the combined influence of the discrete harmonic and biharmonic terms on a discrete Fourier mode with wavenumber $\boldsymbol{k}$ is a linear growth (or decay) with rate

$$-\nu_2\left(\cos\left(\frac{k_x\Delta x}{2}\right)\cos\left(\frac{k_y\Delta x}{2}\right)\right)^2\left[\frac{4}{\Delta x^2}\left(\sin^2\left(\frac{k_x\Delta x}{2}\right)+\sin^2\left(\frac{k_y\Delta x}{2}\right)\right)\right] - \nu_4\left[\frac{4}{\Delta x^2}\left(\sin^2\left(\frac{k_x\Delta x}{2}\right)+\sin^2\left(\frac{k_y\Delta x}{2}\right)\right)\right]^2. \tag{10}$$

To ensure that wavenumbers between $\pi/\Delta x$ and $\pi/(2\Delta x)$ are damped, the viscous coefficients must satisfy the following constraint

$$-\nu_2 < \frac{4\nu_4}{\Delta x^2}. \tag{11}$$

The largest allowable backscatter coefficient is twice as large with the filtering as without. If, in a primitive equation model, the range of wavenumbers that are damped needs to be increased, or if the number of applications of the filter changes, the upper bound on $-\nu_2$ can be modified accordingly.

Bachman (2019) similarly used a spatial filter to enhance the scale separation between backscatter and dissipation, though both dissipation and backscatter were achieved using harmonic operators. Both Berloff (2018) and Juricke et al. (2020a) used spatial filters to define their backscatter schemes instead of a combination of harmonic and biharmonic terms. In a model with non-uniform grid spacing the weights would be updated to account for the unequal grid cell sizes.

Having defined and motivated the filtering term, consider the enstrophy budget for the dynamics. Enstrophy is injected by the wind forcing, cascades towards small scales, and is dissipated by viscosity at small scales. In a statistically steady state, the rate of injection, the cascade rate, and the dissipation rate are all equal. The global-average enstrophy dissipation rate associated with the parameterization is

$$\frac{1}{V}\int_V \nu_4(\nabla^2\omega)^2 \mathrm{d}V. \tag{12}$$

When a backscatter scheme is present we assume it to act at small scales (though at larger scales than dissipation), so that the cascade rate equals the combined backscatter and dissipation rates. The global-average enstrophy dissipation and backscatter rates combine to form the net enstrophy dissipation rate, which is

$$\frac{1}{V}\int_V \nu_2|\nabla\overline{\omega}|^2 + \nu_4(\nabla^2\omega)^2 \mathrm{d}V. \tag{13}$$

As usual in the derivation of Smagorinsky and Leith parameterizations, we assume that the local net enstrophy dissipation rate is instantaneously balanced by the local cascade rate η so that the final expressions for the dissipation and backscatter coefficients can vary in space and time. Henceforth therefore, assume a local equilibrium such that the local enstrophy cascade rate η is equal to the local net enstrophy dissipation rate

$$\eta = \nu_2 |\nabla \overline{\omega}|^2 + \nu_4 (\nabla^2 \omega)^2. \tag{14}$$

Formulas for the coefficients ν_2 and ν_4 will be obtained by imposing two constraints: one on the energy cascade and one on the enstrophy cascade. The enstrophy cascade condition is that the enstrophy cascade should terminate within the range of scales represented on the grid. In the non-backscattering case one assumes that the dissipation scale for enstrophy depends only on the biharmonic coefficient ν_4 and the enstrophy cascade rate η, and dimensional analysis then implies that the dissipation scale must be proportional to

$$\ell_\eta = \left(\frac{\nu_4^3}{\eta}\right)^{1/12}. \tag{15}$$

In the backscattering case there is a second dimensional coefficient ν_2 that could, in principle, play a role in setting the dissipation scale for enstrophy. The application of the spatial filter effectively removes the backscatter term from the smallest resolved scales, which allows us to use the standard assumption that the dissipation scale for enstrophy depends only on the biharmonic coefficient ν_4 and the enstrophy cascade rate η. To ensure that the enstrophy cascade is absorbed on the grid, the grid wavenumber $\pi/\Delta x$ is set proportional to one over the dissipation scale for enstrophy with proportionality constant Υ

$$\frac{\pi}{\Delta x} = \Upsilon \left(\frac{\nu_4^3}{\eta}\right)^{-1/12}. \tag{16}$$

Inserting the appropriate expression for η and simplifying leads to the following constraint

$$\nu_2 |\nabla \overline{\omega}|^2 + \nu_4 (\nabla^2 \omega)^2 = \nu_4^3 \left(\frac{\pi}{\Upsilon \Delta x}\right)^{12}. \tag{17}$$

Notice that when $\nu_2 = 0$ we recover exactly the biharmonic 2D-Leith scaling (Fox-Kemper and Menemenlis, 2008)

$$\nu_4 = \left(\frac{\Upsilon \Delta x}{\pi}\right)^6 |\nabla^2 \omega|. \tag{18}$$

A second constraint is obtained by requiring the backscatter term to re-inject a fraction c_K of the energy that is dissipated by the biharmonic term:

$$\nu_2 \overline{\omega}^2 = -\nu_4 c_K |\nabla \omega|^2. \tag{19}$$

(The expression for the local kinetic energy backscatter rate assumes that the filter commutes with spatial derivatives.) This leads to a local net energy dissipation rate of

$$\epsilon = \nu_2 \bar{\omega}^2 + \nu_4 |\nabla \omega|^2 = \nu_4 (1 - c_K) |\nabla \omega|^2. \tag{20}$$

Setting $c_K = 1$ re-injects all of the energy that is dissipated by the biharmonic term so that the combined effect has no net energy dissipation $\epsilon = 0$, while setting $c_K = 0$ leads to no backscatter.

The solution to the system is

$$\begin{aligned} \nu_4 &= \left(\frac{\Upsilon \Delta x}{\pi} \right)^6 \left[(\nabla^2 \omega)^2 - c_K \frac{|\nabla \omega|^2 |\nabla \overline{\omega}|^2}{\overline{\omega}^2} \right]^{1/2}, \quad (21) \\ \nu_2 &= -\nu_4 c_K \frac{|\nabla \omega|^2}{\overline{\omega}^2}. \quad (22) \end{aligned}$$

There are several practical considerations that prevent direct use of the foregoing expressions. For example, the biharmonic term can be imaginary, which indicates that the constraints on the energy and enstrophy cascades cannot both be satisfied by any real combination of harmonic and biharmonic coefficients. Similarly, the coefficients produced by the above formulas may violate the constraint (11) that ensures scale separation between backscatter and dissipation on the grid, or the formulas may generate a biharmonic coefficient large enough to cause numerical instability.

To obtain a practical method, we first compare the energetic constraint (22) to the scale-separation constraint (11), and choose the constraint that produces a smaller value of ν_2 for a given ν_4. This produces a backscatter coefficient

$$\nu_2 = -m \nu_4 \tag{23}$$

where

$$m = \begin{cases} c_K \frac{|\nabla \omega|^2}{\overline{\omega}^2} & \text{where } c_K \Delta x^2 |\nabla \omega|^2 < 4 \overline{\omega}^2 \\ \frac{4}{\Delta x^2} & \text{where } c_K \Delta x^2 |\nabla \omega|^2 \geq 4 \overline{\omega}^2 \end{cases} \tag{24}$$

Choosing the smaller of the two constraints means that the negative viscosity backscatters as much as possible without violating the scale separation between backscatter and dissipation. We then substitute this expression into the enstrophy constraint (17) and solve for the biharmonic coefficient, which yields the expression

$$\left(\frac{\Upsilon\Delta x}{\pi}\right)^6 \left[(\nabla^2\omega)^2 - m|\nabla\overline{\omega}|^2\right]_+^{1/2}. \tag{25}$$

The argument of the square root is truncated to non-negative values, and this truncation is reflected in the notation $[\cdot]_+$. In practice one wants to avoid having the biharmonic coefficient become too small, so we assume some minimum value ν_4^{min} has been specified. There is also an upper limit ν_4^{max} on the allowable values of the biharmonic coefficient associated with numerical stability. The final formulas are therefore

$$\begin{aligned} \nu_4 &= \max\left\{\nu_4^{\text{min}}, \min\left\{\nu_4^{\text{max}}, \left(\frac{\Upsilon\Delta x}{\pi}\right)^6 \left[(\nabla^2\omega)^2 - m|\nabla\overline{\omega}|^2\right]_+^{1/2}\right\}\right\} && (26)\\ \nu_2 &= -m\nu_4. && (27) \end{aligned}$$

Direct application of these formulas can yield viscous coefficients with extreme spatial variability. To avoid this, the results of the above expressions could be smoothed using the spatial filter before use. This would have the effect of reducing dissipation in places where it was needed though; to smooth the fields while maintaining large values, the values produced by the above formulas were modified as follows before use:

$$\nu_4 \mapsto \left[\overline{\nu_4^2}\right]^{1/2} \tag{28}$$

and similarly for ν_2, though maintaining its negative sign.

Note that for the same vorticity field the biharmonic viscous coefficient produced by this new parameterization is smaller than the biharmonic 2D-Leith coefficient. It is counterintuitive to *reduce* the biharmonic coefficient when backscatter is introduced. Intuitively, one expects from (14) that the local enstrophy cascade rate η will remain constant, so the introduction of enstrophy backscatter $\nu_2|\nabla\bar{\omega}|^2$ should also lead to a correspondingly greater gross enstrophy dissipation $\nu_4(\nabla^2\omega)^2$ so that the sum of backscatter and gross dissipation, i.e. the net dissipation, can remain constant to match the cascade rate. This is in fact what happens, as noted at the end of section 3.4: The gross enstrophy dissipation rate increases when backscatter is introduced. Nevertheless, the coefficient ν_4 remains approximately the same with and without backscatter, which indicates that changes in the vorticity field rather than changes in ν_4 are responsible for maintaining the balance of the net enstrophy dissipation rate. This also explains how the coefficient ν_4 can remain the same: The value produced by (21) is only smaller than the classical biharmonic Leith coefficient (18) if the resolved vorticity field in both expressions is the same. In simulations with and without backscatter the resolved vorticity field is different, and as reported at the end of section 3 these differences conspire to keep the values of ν_4 produced by the classical and backscattering Leith schemes nearly identical.

2.2. Biharmonic QG Leith + E

The energetically-constrained biharmonic QG Leith parameterization adds backscatter to the QG-Leith parameterization of Bachman et al. (2017). It uses the same vorticity dissipation and backscatter terms as the energetically-constrained biharmonic 2D Leith parameterization, given in equations (6) and (7). The parameterization adds diffusion of buoyancy in the form

$$\mathcal{D}_b = -\nu_4 \nabla^4 b \tag{29}$$

which leads to a biharmonic horizontal dissipation of QG PV in equation (3).

In the QG-Leith perspective the cascade to be absorbed by dissipation at small scales is one of potential enstrophy ($q^2/2$) rather than enstrophy ($\omega^2/2$). The local potential enstrophy cascade rate η_q is assumed to balance the local net potential enstrophy dissipation rate

$$\eta_q = \nu_2 \nabla \bar{q} \cdot \nabla \bar{\omega} + \nu_4 (\nabla^2 q)^2. \tag{30}$$

This expression is derived under the traditional assumption that the viscous coefficients are constant, and also under the assumption that the spatial filter commutes with the derivatives. The condition that the dissipation scale of the potential enstrophy cascade be proportional to the grid scale leads to the constraint

$$\nu_2 \nabla \bar{q} \cdot \nabla \bar{\omega} + \nu_4 |\nabla q|^2 = \nu_4^3 \left(\frac{\Upsilon \Delta x}{\pi} \right)^{12}. \tag{31}$$

The form of the KE backscatter rate remains the same as for the 2D version, and one could require the backscatter rate to be equal to a fraction of the KE dissipation rate associated with the biharmonic vorticity diffusion term. The biharmonic buoyancy-anomaly diffusion in QG-Leith leads to a dissipation of APE that is not present in the 2D version, so a more general backscatter constraint would be to make the backscatter rate proportional to some combination of the KE and APE dissipation rates. This produces the constraint

$$\nu_2\bar{\omega}^2 = -\nu_4\left[c_K|\nabla\omega|^2 + c_P\frac{(\nabla^2 b)^2}{N^2}\right] \tag{32}$$

where the second term in the square brackets is proportional to the local APE dissipation rate associated with the biharmonic term. The local net KE dissipation rate associated with the parameterization has the same form as the 2D version (20). There is no APE backscatter associated with this parameterization, so the local net APE dissipation rate is simply $\nu_4(\nabla^2 b)^2/N^2$, while the local dissipation rate of total energy is

$$\epsilon_{\text{KE+APE}} = \nu_4\left[(1-c_K)|\nabla^2\omega|^2 + (1-c_P)\frac{(\nabla^2 b)^2}{N^2}\right]. \tag{33}$$

The GM+E parameterization (Bachman, 2019) and the approach of Jansen et al. (2019) both similarly recycle dissipated APE into KE backscatter.

The two constraints (31) and (32) form a system of two equations for two unknowns. As in the 2D case, the exact solution is not practical, and applying the same practical constraints as in the 2D case produces the following recipe for ν_2 and ν_4

$$\nu_4 = \max\left\{\nu_4^{\min}, \min\left\{\nu_4^{\max}, \left(\frac{\Upsilon\Delta x}{\pi}\right)^6\left[(\nabla^2 q)^2 - m\nabla\bar{q}\cdot\nabla\overline{\omega}\right]_+^{1/2}\right\}\right\} \tag{34}$$

$$\nu_2 = -m\nu_4 \tag{35}$$

where

$$m = \begin{cases} \frac{c_K|\nabla\omega|^2 + c_P(\nabla^2 b)^2/N^2}{\overline{\omega}^2} & \text{where } \Delta x^2(c_K|\nabla\omega|^2 + c_P(\nabla^2 b)^2/N^2) < 4\overline{\omega}^2 \\ \frac{4}{\Delta x^2} & \text{where } \Delta x^2(c_K|\nabla\omega|^2 + c_P(\nabla^2 b)^2/N^2) \geq 4\overline{\omega}^2 \end{cases} \tag{36}$$

The results of the above expressions are smoothed before use via (28) as in the 2D case. Unlike in the 2D case, the biharmonic coefficient can actually be larger than the classical QG-Leith scaling obtained by setting $m = 0$ because $\nabla\bar{q} \cdot \nabla\bar{\omega}$ can in principle be negative.

2.3. Biharmonic 2D Leith + Stochastic Backscatter

The foregoing methods backscattered kinetic energy using a deterministic negative harmonic viscosity. Many other backscatter schemes in the literature use stochastic formulations of backscatter, and it is of interest to compare the approaches. This subsection develops a parameterization with a combination of biharmonic diffusion of vorticity and stochastic backscatter of QG PV. The biharmonic viscosity takes the usual form

$$\mathcal{D}_\omega = -\nabla^2 \nu_4 \nabla^2 \omega \tag{37}$$

and there is no diffusion of buoyancy anomalies $\mathcal{D}_b = 0$.

One could in principle follow the approach of the previous subsections by defining stochastic forcings of vorticity and buoyancy, $\mathcal{B}_\omega$ and $\mathcal{B}_b$. The stochastic forcing of vorticity would correspond to a stochastic kinetic energy backscatter, while the stochastic forcing of buoyancy would correspond to a stochastic backscatter of available potential energy. In practice though it is impossible to separate stochastic forcing of KE and APE in a QG system, unless the backscatter is barotropic (i.e. depth-independent), which is overly limiting.

To understand this it is helpful to think in terms of the practical implementation of a stochastic forcing of QG PV. At the discrete level, a stochastic increment is constructed and added to the QG PV q at every time step, where the amplitude of the increment is proportional to the square root of the time step size $\sqrt{\Delta t}$. When the noise is temporally uncorrelated, the mean rates of energy and enstrophy injection associated with the forcing are the mean energy and enstrophy of the increment, divided by the time step size. Notice that scaling the increment by $\sqrt{\Delta t}$ leads to energy and enstrophy injection rates are independent of Δt. A peculiarity of the QG system is that a QG PV increment Δq generally has both kinetic and available potential energy; the only way to construct a QG PV increment that has only kinetic and no available potential energy is for the QG PV increment to be barotropic. Even if one formally constructs the stochastic increment to q as a combination of a stochastic increment to vorticity plus zero stochastic increment to buoyancy, the resulting QG PV increment ends up backscattering both KE and APE.

It is worth briefly noting that in a primitive equation model it is possible to increment the momentum and buoyancy separately, and thereby construct a stochastic forcing that increments KE and APE separately. But if the flow is approximately geostrophic, then geostrophic adjustment processes will rapidly convert a purely KE increment into a combination of KE and APE. Deterministic backscatter via negative viscosity presumably behaves similarly: although the backscatter is purely kinetic, geostrophic adjustment processes rapidly convert some of this kinetic energy to potential energy. At scales larger than the deformation radius, one expects that a large fraction of the backscattered KE will end up in APE after adjustment processes. The stochastic scheme of Grooms (2016) backscatters only APE; the foregoing reasoning suggests that some fraction of this backscattered APE will convert to KE on the resolved scales.

The stochastic increment of QG PV is constructed to backscatter total energy (rather than kinetic energy) at the desired rate. In order to connect to a potential implementation in a primitive-equation model, an increment is constructed for the QG streamfunction ψ, and the implied QG PV increment is derived from this ψ increment.

Let S be a Gaussian random field that is depth-independent and uncorrelated in time and in the horizontal directions. The horizontal Fourier transform of this field has a uniform 2D spectrum. If S were used directly as an increment to ψ, the associated 1D kinetic energy forcing spectrum would be proportional to k^3, which is strongest at the smallest scales, and violates the principle of separation of scales between backscatter and dissipation. (The 1D spectrum results from averaging the 2D spectrum, which depends on both k_x and k_y, around circles so that the resulting 1D spectrum is a function only of $k = |\boldsymbol{k}|$.) To create such a scale separation, the field S is spatially filtered twice $\overline{\overline{S}}$. Using the spatial filter from section 2.1 (and ignoring discretization errors in forming KE from ψ on the grid), the kinetic energy forcing spectrum associated with a ψ increment of $\overline{\overline{S}}$ would be proportional to $k^3(\cos(k_x\Delta x/2)\cos(k_y\Delta x/2))^4$. The peak of the 1D KE forcing spectrum occurs at a wavenumber close to $\pi/(2\Delta x)$, i.e. the forcing peaks at a scale approximately twice as large as the grid scale. More applications of the filter could be used to push the peak of the forcing spectrum to larger scales, but this would come at extra computational expense as well as communication costs in a parallel code. The use of spatial filters to achieve scale separation between the stochastic backscatter and dissipation goes back at least to Schumann (1995), and its importance was underscored by Grooms et al. (2015).

The amplitude of the stochastic forcing should be set so that it injects energy back into the system at a desired rate. To achieve this goal it is convenient to normalize the basic stochastic increment so that it has unit energy. Define

$$S_0 = \left(\frac{1}{2}\mathbb{E}\left[|\nabla\overline{\overline{S}}|^2\right]\right)^{1/2} \tag{38}$$

where $\mathbb{E}$ denotes the mean across realizations of the random field. The normalizing constant S_0 can be calculated analytically or numerically, and depends only on the shape of the grid, the form of the spatial filter, and the discretization of the Laplacian; the computation is done offline. We define the increment to QG PV to be

$$\Delta q = \sqrt{\Delta t}\nabla\cdot\left(A(x,y,z,t)\nabla\left(\frac{\overline{\overline{S}}}{S_0}\right)\right). \tag{39}$$

If it were possible to separately increment the velocity and buoyancy fields, then this QG PV increment would result from a random, non-divergent, isotropic Gaussian velocity field with amplitude A and zero buoyancy anomaly; indeed, one might implement the parameterization in a primitive-equation model by forming a velocity increment $\sqrt{\Delta t}\,A(x,y,z,t)\nabla(\overline{\overline{S}}/S_0) = (\Delta v, -\Delta u)$. When A is depth-independent the mean local energy injection rate associated with this QG PV increment is

$$\epsilon_{\text{back}} = A^2. \tag{40}$$

Note that the dimensions of A are $L/T^{3/2}$; the unusual fractional scaling is associated with the factor of $\sqrt{\Delta t}$ in the definition of the QG PV increment. Precisely separating this into a KE backscatter rate and an APE backscatter rate requires solving for the associated increment to ψ and then computing the KE and APE backscatter rates from the ψ increment; this is done in the code for diagnostic purposes, but is not required to implement the parameterization. When A is constant the local enstrophy backscatter rate associated with the QG PV increment is

$$\eta_{\text{back}} = A^2 Z \tag{41}$$

where Z can be computed analytically or numerically, and depends only on the shape of the grid, the form of the spatial filter, and the discretization of the Laplacian; the computation of Z is done offline. The dimensions of Z are L^{-2}.

We assume that the local net enstrophy dissipation rate associated with the parameterization is balanced by the local enstrophy cascade rate η

$$\eta = -A^2 Z + \nu_4(\nabla^2\omega)^2. \tag{42}$$

Requiring this to be absorbed by the biharmonic term on the grid leads to the constraint

$$-A^2 Z + \nu_4(\nabla^2\omega)^2 = \nu_4^3\left(\frac{\Upsilon\Delta x}{\pi}\right)^{12}. \tag{43}$$

Requiring the local energy backscatter rate to be proportional to the local rate of KE dissipation associated with the biharmonic term leads to the constraint

$$A^2 = \nu_4 c_K|\nabla\omega|^2. \tag{44}$$

As with the previous, deterministic parameterizations, we want to provide upper and lower bounds on ν_4, so the final form of the parameterization is

$$\nu_4 = \max\left\{\nu_4^{\min}, \min\left\{\nu_4^{\max}, \left(\frac{\Upsilon\Delta x}{\pi}\right)^6\left[(\nabla^2\omega)^2 - c_K Z|\nabla\omega|^2\right]_+^{1/2}\right\}\right\} \tag{45}$$

$$A = [\nu_4 c_K]^{1/2}|\nabla\omega|. \tag{46}$$

The coefficients A and ν_4 are smoothed via (28) before use.

2.4. Biharmonic QG Leith + Stochastic Backscatter

The stochastic method from the foregoing section can be updated from '2D' to 'QG' by adding biharmonic diffusion of buoyancy anomalies. The expressions for the KE, APE, and enstrophy dissipation rates remain the same as in section 2.2, and the expressions for the enstrophy and net energy backscatter rates remain as in the previous section. The updated formulas for the backscatter amplitude and biharmonic coefficient are

$$\nu_4 = \max\left\{\nu_4^{\min}, \min\left\{\nu_4^{\max}, \left(\frac{\Upsilon\Delta x}{\pi}\right)^6\left[(\nabla^2 q)^2 - Z\left(c_K|\nabla\omega|^2 + c_P\frac{(\nabla^2 b)^2}{N^2}\right)\right]_+^{1/2}\right\}\right\} \tag{47}$$

$$A = \left[\nu_4\left(c_K|\nabla\omega|^2 + c_P\frac{(\nabla^2 b)^2}{N^2}\right)\right]^{1/2}. \tag{48}$$

2.5. Harmonic 2D and QG Leith + Stochastic Backscatter

The deterministic parameterizations of sections 2.1 and 2.2 achieved scale separation between the backscatter and dissipation by using a combination of harmonic and biharmonic terms; the scale separation was enhanced through the use of a spatial filter. The backscatter scheme in the GM+E parameterization (Bachman, 2019) uses a combination of harmonic terms and a spatial filter to achieve scale separation; the backscatter is also confined to the barotropic part of the flow. The backscatter scheme of Juricke et al. (2020a) uses a combination of a harmonic operator and a spatial filter to achieve dissipation and backscatter on a non-rectangular grid.

A similar deterministic parameterization could be developed here using only a harmonic term and a spatial filter. Following Juricke et al. (2020a) we could develop a scheme of the form

$$\mathcal{B}_\omega + \mathcal{D}_\omega = \nu_2\nabla^2\omega - \alpha\nu_2\nabla^2\overline{\omega}. \tag{49}$$

If we ignore discretization errors and spatial variability of the viscous coefficients, the influence of the combined parameterization on a discrete Fourier mode with wavenumber $\boldsymbol{k}$ is a linear growth (or decay) with rate

$$-\nu_2\left(1-\alpha\left(1-\frac{\Delta x^2 k^2}{\pi^2}\right)\right)k^2 \tag{50}$$

where, for simplicity, it was assumed that the spatial filter has Fourier symbol $1-(\Delta x k/\pi)^2$. This is exactly equivalent to a combination of a harmonic backscatter with coefficient $\nu_2(1-\alpha)$ and a biharmonic dissipation with coefficient $\nu_2\alpha(\Delta x/\pi)^2$. (The scheme only backscatters for $\alpha > 1$.) The resulting scheme would be very similar to what was developed in sections 2.1 and 2.2, so we do not develop such a scheme here. Note that this similarity enables the scalings of ν_2 and ν_4 developed in the preceding sections to be converted into scalings for ν_2 and α in the backscatter scheme of Juricke et al. (2020a).

In contrast to the use of filters and a harmonic viscosity, the combination of a harmonic dissipation with a stochastic backscatter is not expected to be effectively the same as a combination of a biharmonic dissipation and a stochastic backscatter. We therefore develop two such parameterizations here, one 2D and one QG. In the 2D case the vorticity dissipation term is

$$\mathcal{D}_\omega = \nabla^2(\nu_2\omega) \tag{51}$$

and there is no diffusion of buoyancy anomalies. The stochastic backscatter takes the same form as developed in section 2.3. The assumption that the local enstrophy cascade rate matches the local net enstrophy dissipation rate thus takes the form

$$\eta = -A^2 Z + \nu_2 |\nabla\omega|^2. \tag{52}$$

Following the same development as the previous sections, the parameterization becomes

$$\nu_2 = \max\left\{\nu_2^{\min}, \min\left\{\nu_2^{\max}, \left(\frac{\Upsilon\Delta x}{\pi}\right)^3 \left[|\nabla\omega|^2 - c_K Z\omega^2\right]_+^{1/2}\right\}\right\}. \tag{53}$$

$$A = \sqrt{\nu_2 c_K}|\omega|. \tag{54}$$

The QG version proceeds as above but adds a harmonic diffusion of buoyancy anomalies

$$\mathcal{D}_b = \nabla^2(\nu_2 b). \tag{55}$$

Requiring the total energy backscatter rate to be proportional to a weighted sum of the KE and APE dissipation rates leads to the following parameterization

$$\nu_2 = \max\left\{\nu_2^{\min}, \min\left\{\nu_2^{\max}, \left(\frac{\Upsilon\Delta x}{\pi}\right)^3 \left[|\nabla q|^2 - Z\left(c_K\omega^2 + c_P\frac{|\nabla b|^2}{N^2}\right)\right]_+^{1/2}\right\}\right\} \tag{56}$$

$$A = \left[\nu_2\left(c_K\omega^2 + c_P\frac{|\nabla b|^2}{N^2}\right)\right]^{1/2}. \tag{57}$$

2.6. Summary

This section has developed six new parameterizations which are summarized here. The '2D' parameterizations dissipate only KE via diffusion of horizontal momentum; the 'QG' parameterizations diffuse both horizontal momentum and buoyancy anomalies.

The two deterministic parameterizations combine harmonic KE backscatter with biharmonic dissipation. These are called biharmonic 2D Leith + E and QG Leith + E, and are abbreviated 2DL4+E and QGL4+E, respectively, where the '+E' is inspired by GM+E (Bachman, 2019). The 2DL4+E parameters are defined by (26), (27), and (24). The QGL4+E parameters are defined by (34), (35), and (36).

Four of the new parameterizations are stochastic. Two of the stochastic parameterizations use a stochastic backscatter paired with a biharmonic

Table 1: Naming Conventions for Parameterizations

Abbrev.	Backscatter	Diffusion	Details
2DL2	None	$\nabla^2(\nu_2\omega)$	§2.5, Eq. (53), (54), $c_K = 0$
2DL2+ES	Stochastic	$\nabla^2(\nu_2\omega)$	§2.5, Eq. (53), (54), $c_K > 0$
QGL2	None	$\nabla^2(\nu_2 q)$	§2.5, Eq. (56), (57), $c_K = c_P = 0$
QGL2+ES	Stochastic	$\nabla^2(\nu_2 q)$	§2.5, Eq. (56), (57), $c_K, c_P > 0$
2DL4	None	$-\nabla^2(\nu_4\nabla^2\omega)$	§2.1, Eq. (26), (27), $c_K = 0$
2DL4+E	$\overline{\nabla^2(\nu_2\bar{\omega})}$	$-\nabla^2(\nu_4\nabla^2\omega)$	§2.1, Eq. (26), (27), $c_K > 0$
2DL4+ES	Stochastic	$-\nabla^2(\nu_4\nabla^2\omega)$	§2.3, Eq. (45), (46), $c_K > 0$
QGL4	None	$-\nabla^2(\nu_4\nabla^2 q)$	§2.2, Eq. (34), (35), $c_K = c_P = 0$
QGL4+E	$\overline{\nabla^2(\nu_2\bar{\omega})}$	$-\nabla^2(\nu_4\nabla^2 q)$	§2.2, Eq. (34), (35), $c_K, c_P > 0$
QGL4+ES	Stochastic	$-\nabla^2(\nu_4\nabla^2 q)$	§2.4, Eq. (47), (48), $c_K, c_P > 0$
2DL4+e	$\overline{\nabla^2(\nu_2\bar{\omega})}$	$-\nabla^2(\nu_4\nabla^2\omega)$	§2.1, Eq. (26) with $c_K = 0$, (27) with $c_K > 0$

dissipation; the 2D version is abbreviated 2DL4+ES, where the trailing S denotes 'stochastic,' and the QG version is abbreviated QGL4+ES. The 2DL4+ES parameters are defined by (45) and (46), while the QGL4+ES parameters are defined by (47) and (48).

Two of the stochastic parameterizations use a stochastic backscatter paired with a harmonic dissipation; the 2D version is abbreviated 2DL2+ES and the QG version is abbreviated QGL2+ES. The 2DL2+ES parameters are defined by (53) and (54), while the QGL2+ES parameters are defined by (56) and (57).

For comparison we also consider the harmonic and biharmonic forms of the 2D-Leith and QG-Leith parameterizations. The harmonic forms of 2D-Leith and QG-Leith are obtained by setting $c_K = c_P = 0$ in (53) and (56), respectively. The biharmonic forms of 2D-Leith and QG-Leith are obtained by setting $m = 0$ in (26) and (34), respectively.

Finally, we consider parameterizations that use Leith scaling for the biharmonic dissipation coefficient ν_4 without including the effect of backscatter on the enstrophy cascade rate. The 2D version is obtained by setting $m = 0$ in (26) and then using (27) to set the backscatter coefficient with $m \neq 0$ set by (24). This scheme is called 2DL4+e to distinguish it from the 2DL4+E scheme developed above.

The naming conventions and details of the parameterizations are summarized in Table 1.

For all of the parameterizations, the dissipation and backscatter coefficients are smoothed using (28) before use. The backscattering schemes all use $c_K = c_P = 1$ except in section 3.5, which explores the sensitivity of the 2DL4+E scheme to c_K. The coefficient Υ in the biharmonic 2D schemes was set to 1.5, and in the biharmonic QG schemes to 1.3. The coefficient Υ in the 2D harmonic schemes was set to 1.3, and in the harmonic QG schemes to 1.1. Section 3.6 explores the sensitivity of the 2DL4+E scheme to Υ. Values of the minimum and maximum coefficients can be found in Appendix A.

3. Numerical Experiments

The parameterizations developed in the foregoing section are compared in the context of a quasigeostrophic double-gyre model. The domain is a square midlatitude basin of width 3,072 km and depth 4 km. Simulations are run with grid sizes of 24, 16, 12, 7, and 4 km, where the 7 km run uses a grid of 449×449 points, including boundary values; there are six layers ranging from the 385 m thick top layer to the 2270 m thick bottom layer. Shevchenko and Berloff (2017) found that there is a significant difference between three and six layers in a very similar QG model, but found the six layer results to be very similar to results with twelve layers. Simulations are spun up from rest for 10 years (it takes about 5 years for the kinetic energy to reach statistical equilibrium). Once spun up, the results are compared based on the final 10 years of each simulation. The time mean streamfunction and spatial pattern of kinetic energy are computed online. Full details on the model configuration and numerical methods can be found in Appendix A.

Figure 1 shows the behavior of a simulation with constant biharmonic diffusion of vorticity and no backscatter at 4 km resolution. The upper left panel shows a snapshot of upper-layer QG PV. As usual for a double-gyre model, the most prominent feature is an eastward jet separating from the western boundary currents, along with its adjacent recirculation zones. In addition to the jet and the associated QG PV front between the gyres, the entire domain is populated with mesoscale vortices. The upper right panel shows the time-mean streamfunction in the top layer. The time-mean jet is not purely zonal, and the time-mean recirculation zones that flank the jet are evident. The lower left panel shows the root-mean-square current speed, i.e. the square root of the time mean of $u^2 + v^2$ in the upper layer. Although eddies are present throughout the domain, most of the kinetic energy is found

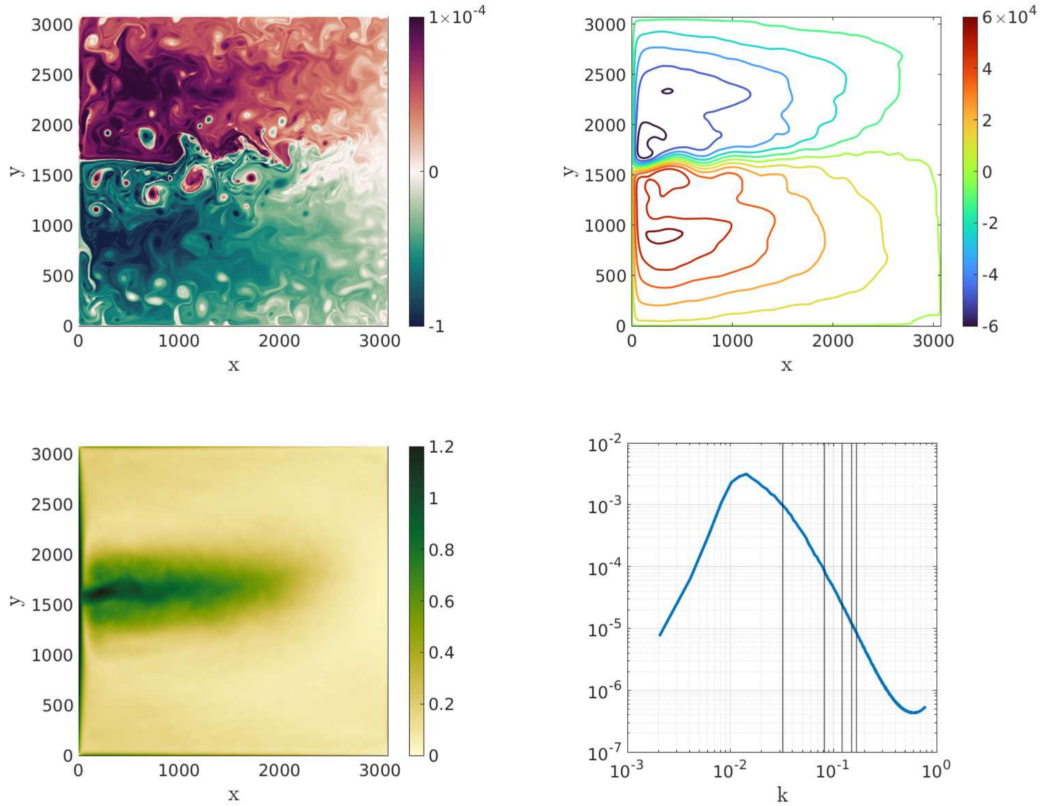

Figure 1: Results from a simulation at 4 km resolution with constant biharmonic diffusion of vorticity. Upper left: A snapshot of QG PV in the upper layer (units s^{-1}). Upper right: Contours of time-mean streamfunction ψ in the upper layer; the contour interval is 1.2×10^4 m^2/s. Lower left: Square root of the time mean of $u^2 + v^2$ in the upper layer (units m/s). Lower right: Time-mean eddy kinetic energy spectrum of the upper layer; the five vertical lines show the wavenumbers associated with the baroclinic deformation radii. Units for the x and y axes are km, while units of the k axis are km^{-1}.

in the jet and in the eddies that separate from it. The lower right panel shows the kinetic energy spectrum of the deviations of the upper layer from its time mean. The five baroclinic deformation radii, from 31 km to 6 km, are shown as vertical lines in the figure. The kinetic energy spectrum peaks at a radius (one over the wavenumber) of 70 km. The bump at the tail of the spectrum suggests that a larger viscous coefficient could have been used, but the effect is small, as seen in the smoothness of the upper left panel. Shevchenko and Berloff (2015, 2017) provide an in-depth discussion of the dynamics of a similarly-configured QG model.

Twelve schemes are compared: The six new schemes described in the preceding section, plus four nonlinear viscosities without backscatter (harmonic and biharmonic versions of 2D and QG Leith), plus a constant biharmonic viscosity. The constant biharmonic viscosity was set at each resolution to be near the maximum coefficient produced by the biharmonic 2D Leith scheme (2DL4) at that resolution. The final scheme uses the traditional 2D Leith scaling for the biharmonic dissipation, and adds backscatter via negative viscosity with a coefficient scaled so that 100% of the energy dissipated by the biharmonic term is backscattered by the negative viscosity term.

Rather than pick a single scheme to generate a reference simulation at the highest resolution of 4 km, all the schemes are run at all the resolutions. The following subsection describes the differences between the methods at the highest resolution (4 km), and the following subsections describe how the results vary with resolution for each scheme.

3.1. Comparison at High Resolution

The total energy dissipation rate is a combination of dissipation due to the frictional bottom boundary layer and the combined effect of the horizontal viscous dissipation and backscatter. At high resolution one expects the total energy dissipation to be dominated by the frictional component. The twelve schemes compared here all have small viscous fractions of the total energy dissipation rate, but there remain significant differences between the schemes. These differences are shown in Figure 2, which also shows the total kinetic energy for each scheme.

The kinetic energies produced by each scheme are different; the largest difference amounts to 14%. There are also differences between the viscous percentages produced by the schemes. In the harmonic QG Leith scheme (QGL2) 12.6% of the total dissipation rate is associated with the parameterization. Note that in all of the QG schemes (cf. the 2D schemes), viscous

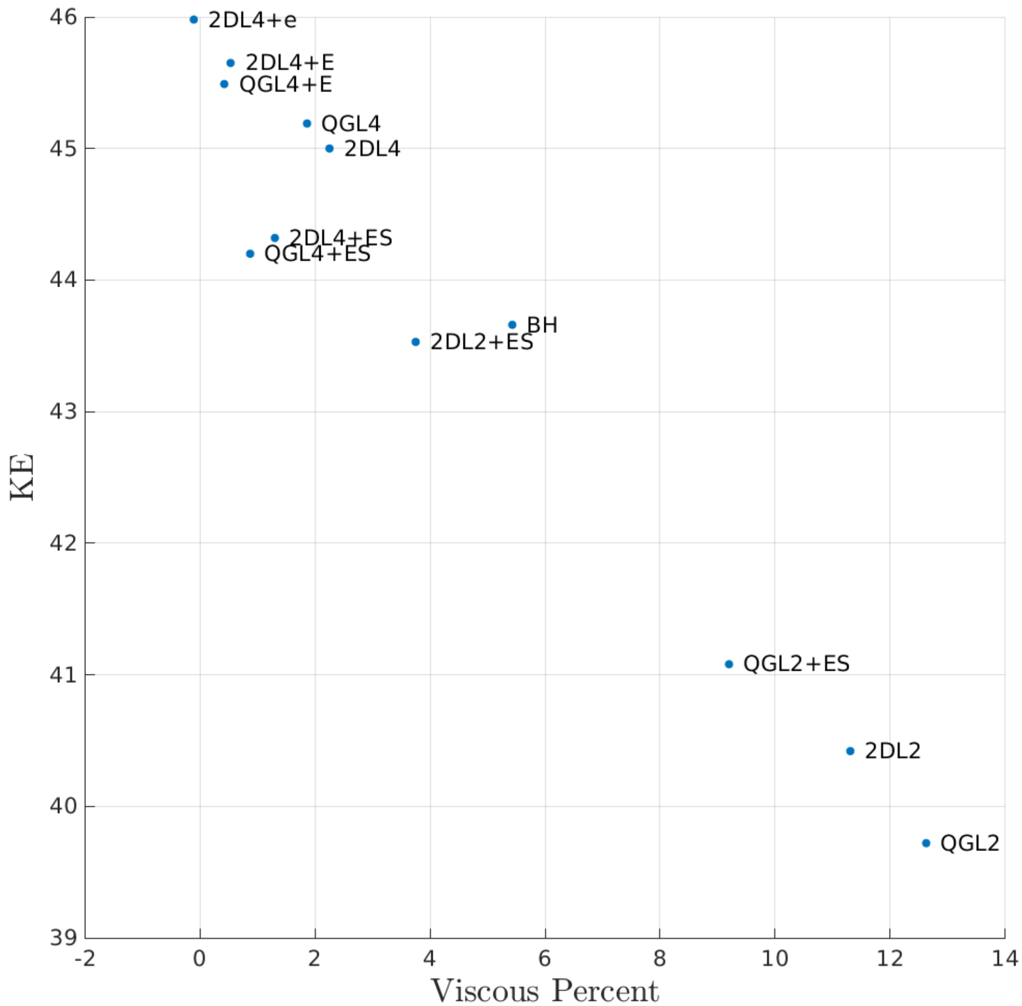

Figure 2: Results for each of the 12 schemes at 4 km resolution. The horizontal axis measures the percent of the total energy dissipation rate that is attributable to the parameterization. The vertical axis shows the total kinetic energy, measured in units of $10^{13}\times$ m^5 s^{-2} (i.e. $(u^2+v^2)/2$ integrated over the volume).

dissipation of APE is included in the viscous percent of the dissipation rate for total energy. In the 2DL4+e scheme the backscatter is slightly stronger than the biharmonic dissipation, resulting in a viscous percentage of -0.11%. The time-mean states produced by the 12 schemes are all very similar at 4 km resolution.

The next section presents the way that each scheme varies with resolution. The differences at 4 km resolution underscore the importance of comparing each scheme to its own reference at 4 km.

3.2. Varying Resolution: Enstrophy Flux

Since the parameterizations developed here are all intended to apply when the grid scale lies within an enstrophy cascade range, the goal of this subsection is to establish that such a range exists. Figure 3 shows the spectral potential enstrophy flux in the top layer, computed from the results of the constant-coefficient biharmonic simulations at all resolutions. A positive flux (i.e. towards small scales) is found all resolutions, though the flux amplitude increases as the resolution improves. The flux profile is not flat, i.e. independent of k, for any range of wavenumbers, as one might expect in a purely inertial range. The potential reasons for this include the fact that the flow is markedly inhomogeneous, with different cascade rates in different parts of the domain, and the fact that enstrophy is not necessarily injected to the system at a single length scale, but can be associated with higher-mode baroclinic instabilities allowed by the use of 6 layers. The increase in the magnitude of the flux from 24 km through 12 km grids is incremental, and is followed by large jumps as the grid size reduces to 7 and 4 km.

3.3. Varying Resolution: Kinetic Energy

We begin by evaluating how the total kinetic energy for each scheme converges as the resolution is varied. Figure 4 shows the ratio of the total KE at a given resolution to the total KE at 4 km resolution for each scheme. The schemes are differentiated in three ways: 2D schemes have open markers while QG schemes have filled markers; schemes with harmonic dissipation are marked with circles while schemes with biharmonic dissipation are marked with squares; schemes without backscatter use solid lines, schemes with deterministic backscatter use dashed lines, and schemes with stochastic backscatter use dash-dotted lines. The results from the constant biharmonic scheme are marked with triangles and the 2DL4+e scheme is marked by a diamond. In the discussion that follows the phrases 'higher KE' and 'lower

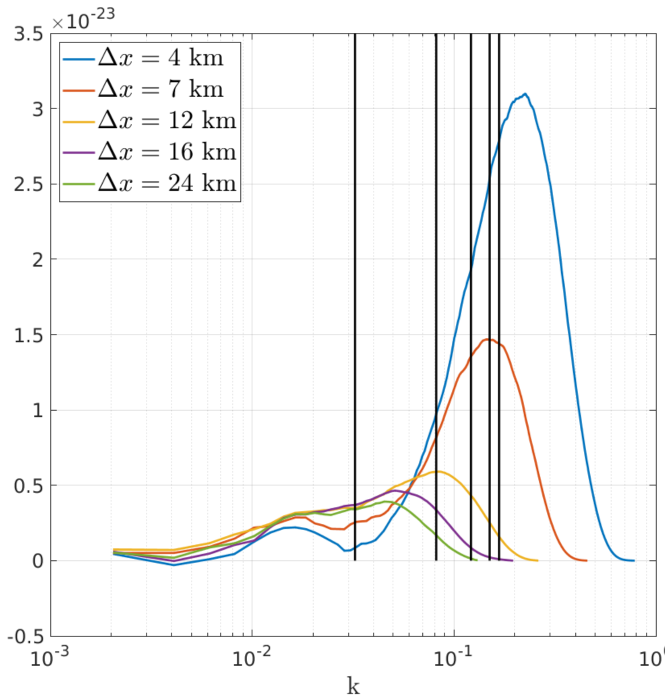

Figure 3: Spectral potential enstrophy flux in the top layer for the constant-coefficient biharmonic scheme at all resolutions. The five vertical lines show the wavenumbers associated with the baroclinic deformation radii. The horizontal axis unit is km^{-1}.

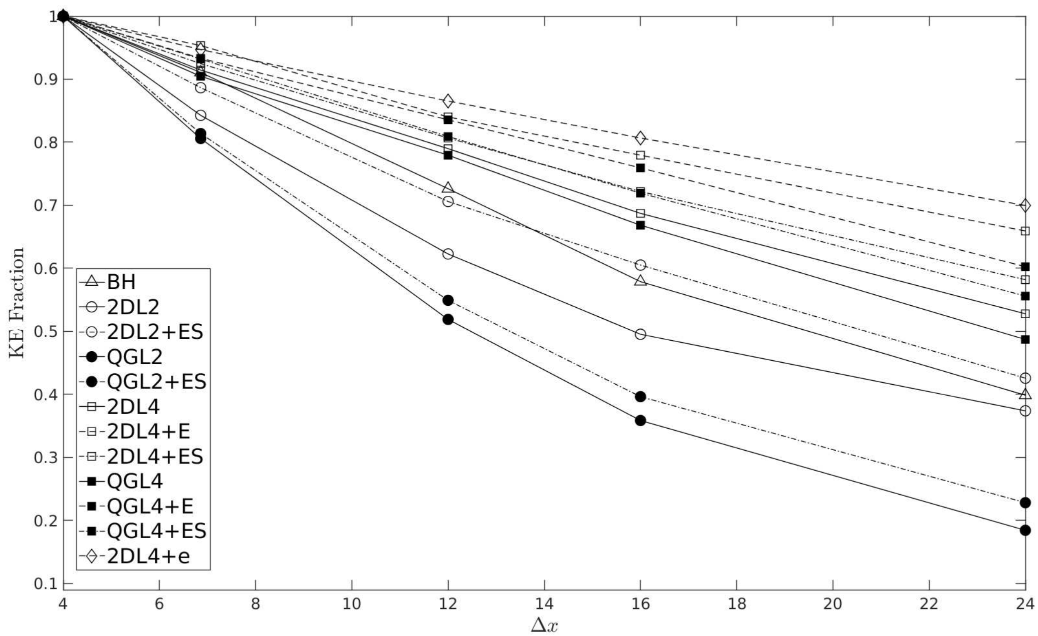

Figure 4: Convergence of total KE as a function of resolution. 2D schemes have open markers while QG schemes have filled markers; schemes with harmonic dissipation are marked with circles while schemes with biharmonic dissipation are marked with squares; schemes without backscatter use solid lines, schemes with deterministic backscatter use dashed lines, and schemes with stochastic backscatter use dash-dotted lines. The horizontal axis unit is km.

KE' should be understood in a relative sense, i.e. 'scheme A has higher KE than scheme B' means that at a given resolution the KE from scheme A is closer to its 4 km value than the KE from scheme B is to its 4 km value.

First compare the schemes without backscatter (solid lines). Both the harmonic 2D and QG Leith schemes generate lower KE than the constant-coefficient biharmonic scheme, which itself generates lower KE than the biharmonic 2D and QG Leith schemes. For both harmonic and biharmonic schemes the QG Leith scheme produces less KE than the 2D Leith scheme.

It is not surprising that the biharmonic Leith schemes have higher KE than the constant-coefficient biharmonic scheme, because the harmonic coefficients ν_4 produced by the Leith scaling are lower than the constant coefficient over most of the domain. It is somewhat surprising that the harmonic Leith schemes have lower KE than the constant-coefficient biharmonic scheme. This is presumably attributable to the fact that energy dissipation in the harmonic schemes is less scale-selective than the biharmonic schemes.

It is also worth noting that 16 km resolution corresponds to 1.94 grid points per deformation radius, and 2 grid points per deformation radius is sometimes considered to be a rule-of-thumb for resolving mesoscale eddies. Nevertheless, in these experiments the total KE at 16 km resolution is far from the value at 4 km. The worst scheme in this regard is harmonic QG Leith (QGL2), which at 16 km resolution has only 36% of its limiting value at 4 km resolution. The best scheme in this regard is biharmonic 2D Leith (2DL4), which at 16 km resolution has 69% of its limiting value at 4 km resolution.

At 24 km resolution, which would be considered eddy-permitting for a deformation radius of 31 km, there is wide variation among the schemes. The harmonic QG Leith scheme in particular does very poorly – significantly worse than even than the harmonic 2D Leith scheme. This difference is presumably due to the fact that QG Leith dissipates both APE and KE while 2D Leith dissipates only KE, and also to the fact that the harmonic viscosity is less scale-selective and thus more dissipative than the biharmonic viscosity.

Next compare the effect of adding backscatter to a Leith scheme. In all cases the KE increases. The scheme which stands to gain the most from backscatter is harmonic QG Leith, but the backscattering version (QGL2+ES) still has less KE than the harmonic 2D Leith scheme (2DL2). The addition of stochastic backscatter to the harmonic Leith schemes helps both of them, but in neither case is it able to bring them to parity with the non-backscattering biharmonic Leith schemes.

Considering the addition of backscatter to the biharmonic Leith schemes, deterministic backscatter is more effective than stochastic backscatter at raising the total KE, though the difference is not large, and decreases as resolution improves. Biharmonic 2D Leith with deterministic backscatter already has 66% of the limiting value at 4 km resolution, which, though too low, is a huge improvement compared to the worst method (QGL2) which at 24 km resolution has only 18% of the limiting KE. This underscores the huge impact that the choice of parameterization can have at eddy-permitting resolution.

Next note that the 2DL4+e scheme has the highest KE overall. Recall that this scheme uses a traditional biharmonic 2D Leith scaling for ν_4 and then adds on deterministic backscatter with a coefficient chosen so that the backscatter exactly cancels the energy dissipation from the biharmonic term. This method is only slightly better than the 2DL4+E scheme, which adjusts the biharmonic coefficient ν_4 to account for the the backscatter of enstrophy.

Next note that the biharmonic Leith schemes all have significantly higher KE than their harmonic counterparts. At 7 km resolution all of the biharmonic schemes have KE within 10% of their value at 4 km resolution, whereas the harmonic schemes are still significantly lower. The 7 km resolution has better than four grid points per deformation radius and nearly two points per *second* baroclinic deformation radius, yet the schemes with harmonic viscosity, even when backscatter is included, still have only 80 to 90% of their KE at 4 km resolution.

Finally, note that it is not fair to compare KE across resolutions because the total KE at 4 km resolution includes KE from scales that are simply not present on the lower resolution grids. To examine the importance of this effect, the flow field from the constant-coefficient biharmonic simulation at 4 km resolution was filtered using the Taper filter described by Grooms et al. (2021) and implemented in the Python package `gcm-filters` (Loose et al., 2022) across a range of filter scales out to a filter scale of 96 km. At a filter scale of 24 km the KE of the filtered field was still 99.6% of the total KE, and at a filter scale of 96 km the KE of the filtered field was still 90.2% of the total KE. The conclusion is that the lack of energy in the low-resolution simulations is almost entirely due to incorrect representation of the scales that can be represented at those resolutions and not to the absence of KE

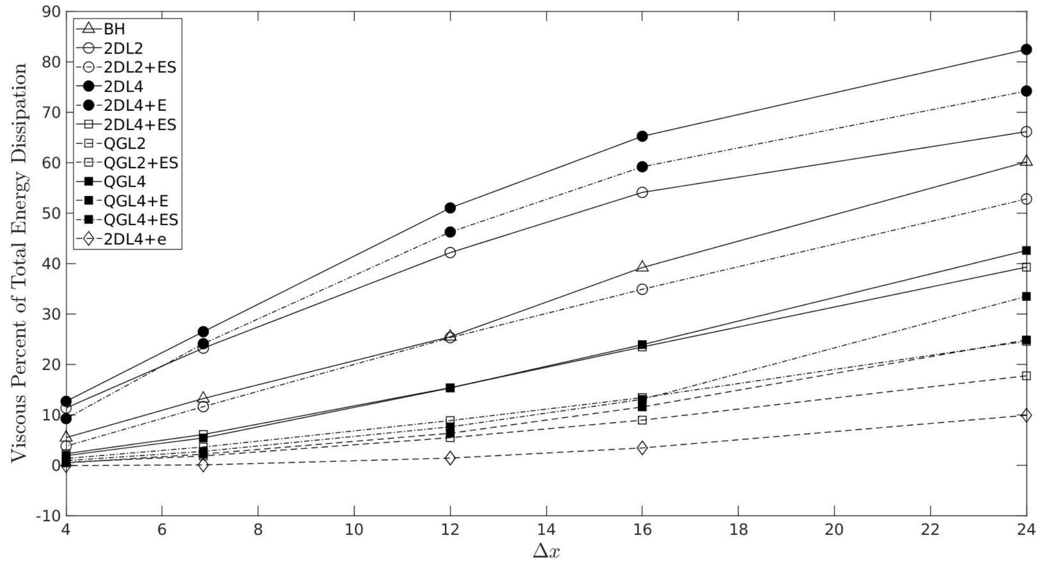

Figure 5: Percentage of the net dissipation rate for total energy that is attributable to the combined effect of viscosity and backscatter, shown for all schemes as a function of grid scale. 2D schemes have open markers while QG schemes have filled markers; schemes with harmonic dissipation are marked with circles while schemes with biharmonic dissipation are marked with squares; schemes without backscatter use solid lines, schemes with deterministic backscatter use dashed lines, and schemes with stochastic backscatter use dash-dotted lines. The horizontal axis unit is km.

from scales that cannot be represented at those resolutions.

3.4. Energy Dissipation Rates

This section compares how the dissipation rates vary across resolution for each scheme. The total energy dissipation rate must match the total energy generation rate, and the latter is set by shape of the upper layer streamfunction. To wit, the time-mean energy generation rate is

$$\mathrm{E}_{\mathrm{gen}} = -H_1 \int_A \overline{\psi}_1 F(x,y)\mathrm{d}A \tag{58}$$

where the integral is over the horizontal domain and $F(x,y)$ is related to the wind stress curl (see Appendix A). Across all twelve schemes and all five resolutions the time-mean energy generation rate varies only a little, and therefore the net dissipation rate of total energy also varies only a little. (The terms 'gross dissipation' and 'net dissipation' are used to indicate whether backscatter is included in the sum ['net'] or not ['gross']; the term 'total' refers to the sum of KE and APE.) We therefore begin by comparing how much of the total energy dissipation rate comes from the net effect of viscosity and backscatter across all schemes and all resolutions. The results are shown in Fig. 5.

The results in Fig. 5 mimic the results of Fig. 4 for KE: the harmonic schemes have higher viscous percentages than the biharmonic schemes; the addition of backscatter generally reduces the viscous percent; and the 2D schemes have lower viscous percent than their QG counterparts. A striking aspect of Fig. 5 is that the backscatter schemes, which were designed to have zero net viscous dissipation, do not, in most cases, achieve that design goal. Though one could increase the backscatter by increasing the values of the coefficients c_K or c_P, it is of interest to discuss why setting $c_K = c_P = 1$ does not, in practice, lead to an exact cancellation of backscatter and dissipation rates.

The deterministic backscatter schemes have a built-in limiter that prevents the backscatter coefficient from growing so large that it violates scale separation between backscatter and dissipation (see the discussion in section 2.1 around the constraint 11); this limiter could partially account for the failure of the backscatter to completely cancel the viscous dissipation. Another culprit is the smoothing that is applied to both the backscatter and dissipation coefficients before use, which breaks the exact link between the backscatter and dissipation rates.

The deterministic schemes perform better than the stochastic ones in the sense that the deterministic schemes do a better job of canceling the viscous dissipation. Like the deterministic schemes, the stochastic schemes use smoothing of the backscatter and dissipation coefficients. The stochastic schemes also suffer from two limitations that are different from the deterministic schemes. First, the mean backscatter rate is computed under the assumption that the amplitude is depth-independent, which is not correct in practice. This relates to the difficulty, discussed in section 2.3, in cleanly separating the KE and APE backscatter rates in a QG code, which would be far more straightforward in a primitive-equation model. Second, as noted in section 2.3, the spatial structure of the backscatter used here generates a forcing spectrum that is peaked at a length scale approximately equal to twice the grid scale. This may not be as well separated from the dissipation range as the deterministic backscatter, whose spectrum is depends on the resolved kinetic energy spectrum. The peak of the stochastic backscatter

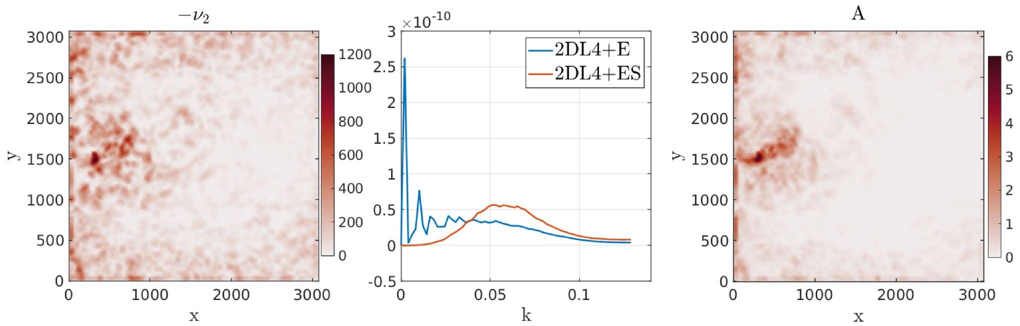

Figure 6: Left: A snapshot of the negative viscous coefficient from the 2DL4+E model (units $\mathrm{m^2\ s^{-1}}$). Center: The backscatter spectrum produced by the 2DL4+E model (blue) and by the 2DL4+ES model (red). Right: A snapshot of the backscatter amplitude (units $\mathrm{m\ s^{-3/2}}$). Axis units in the left and right panels are km; axis units in the center panel are $\mathrm{km^{-1}}$.

spectrum could be shifted to larger scales, and this could be accomplished by smoothing the noise many more times; the practical drawback, besides the computational expense of the smoothing, is that in a parallel primitive-equation implementation many smoothing cycles would require lots of slow communication between parallel processes (halo updates). Optimal performance in a parallel primitive-equation model may require a stochastic pattern generator that does not rely on smoothing spatially-white noise.

To illustrate the differences between the backscatter spectrum produced by the deterministic and stochastic schemes, we show in the middle panel of Fig. 6 the backscatter spectrum produced by the 2DL4+E and 2DL4+ES schemes at 24 km resolution. The stochastic backscatter spectrum has a broad peak at about half the smallest representable wavenumber, i.e. about twice the grid scale, while the deterministic backscatter spectrum is broader and has more backscatter at larger scales. A peculiarity of the deterministic backscatter spectrum is the spike at scales near the box size; this is presumably a result of aliasing arising from the product of the backscatter coefficient ν_2 and the vorticity, though it may also result from the action of the negative viscosity on the large-scale gyre circulation. For reference, the left and right panels of Fig. 6 show the negative viscous coefficient $-\nu_2$ and the backscatter amplitude A. These have broadly similar features, with a concentration of backscatter near the separation of the western boundary currents. The deterministic backscatter spectrum does not go exactly to zero at half the smallest representable wavenumber because the theoretical guarantee is applies to the combined backscatter and dissipation spectrum, and is based on the assumption that ν_2 and ν_4 are constant; nevertheless, the deterministic backscatter remains weak at small scales.

To dig more deeply into the behavior of the backscatter schemes, we show in Fig. 7 the breakdown of the total dissipation budget for all backscattering schemes at 24 km resolution, where the differences are greatest. All of the schemes achieve similar gross dissipation rates, but the deterministic schemes (the rightmost three columns) have somewhat more gross dissipation than the stochastic schemes. The deterministic schemes offset this increase in gross dissipation by a corresponding increase in backscatter rates so that, per Fig. 5, the net viscous percentage achieved by the deterministic schemes is slightly smaller than their stochastic counterparts.

Comparing the backscattering QG schemes to their 2D counterparts shows broad similarity. One significant difference in the biharmonic schemes is that while the gross dissipation of total energy in the QGL4+E and QGL4+ES schemes is similar to the gross dissipation of total energy in the 2DL4+E and 2DL4+ES schemes, the QG schemes achieve that total by dissipating less KE and more APE. The harmonic QG scheme (QGL2+ES) both dissipates and backscatters at higher rates than the harmonic 2D scheme (2DL2+ES). The rate of APE backscatter in the QGL2+ES scheme is simply not sufficient to match the very large rate of APE dissipation, with serious deleterious consequences for the overall KE level: QGL2+ES has by far the lowest KE of any backscattering scheme at 24 km resolution.

Finally note that although 2DL4+e, which does not adjust the 2D Leith scaling to account for enstrophy backscatter, has the lowest viscous fraction of any scheme across all resolutions, it achieves this by both dissipating and backscattering more energy than 2DL4+E. The goal of the backscatter is to correct spurious energy dissipation. Spurious energy dissipation is a model error, and the schemes developed here introduce a new model error associated with non-physical backscatter to compensate for the original model error of spurious energy dissipation. The 2DL4+e scheme, despite resulting in a slightly better total energy level and viscous percentage of dissipation, achieves this through a combination of two large and compensating model errors: it generates significantly more spurious energy dissipation than the 2DL4 or 2DL4+E schemes; it happens to also be able to correct this spurious dissipation slightly more efficiently than the 2DL4+E scheme, at least in the results reported here.

Similarly, the 2DL4+e scheme both dissipates and backscatters more enstrophy than the 2DL4+E scheme. For example, at 12 km resolution the 2DL4 scheme leads to a net enstrophy dissipation rate integrated over the top layer of 9.1×10^4 m^3 s^{-3}. (Henceforth all enstrophy rates will be given in units of 10^4 m^3 s^{-3}.) The 2DL4+e and 2DL4+E schemes generate slightly higher net enstrophy dissipation rates of 10.3 and 10.4, respectively. However, the 2DL4+e scheme achieves this net enstrophy dissipation rate by a combination of a gross dissipation rate of 15.3 and a backscatter rate of 5.1, while the 2DL4+E scheme has lower gross enstrophy dissipation and backscatter rates of 13.9 and 3.5, respectively.

Finally, the 2DL4+e scheme, despite having the same formula for the biharmonic coefficient ν_4 as the 2DL4 scheme, generates significantly larger values of ν_4 than the 2DL4 scheme: at 12 km resolution the 2DL4 scheme generates ν_4 whose median value over the top layer is 1.11×10^9 m^4 s^{-1}, while the 2DL4+e scheme generates a median value of 1.35×10^9 m^4 s^{-1}, which is 22% higher. The median value of the biharmonic coefficient ν_4 over the top layer is also larger in the 2DL4+e scheme than in the 2DL4+E scheme: At 4 km resolution the value is 14% higher and the ratio increases across resolutions until at 24 km resolution the value is 43% higher. In contrast, the 2DL4+E scheme produces values of ν_4 that are close to those generated by the 2DL4 scheme across all resolutions. While in this idealized QG model there do not appear to be severe consequences for combining high rates of dissipation and backscatter, it remains to be seen whether this will prove unstable in a more comprehensive primitive-equation model.

3.5. Sensitivity to c_K

This section explores the sensitivity of the 2DL4+E scheme to the coefficient c_K, which is the fraction of the rate of kinetic energy dissipation by the biharmonic term that is backscattered by the harmonic term. At 16 km resolution, the value of c_K was varied from 0 (which reduces to the 2DL4 scheme) to 1. Figure 8 shows that the rates of kinetic energy dissipation from friction and from the biharmonic term increase monotonically with c_K. The rate of kinetic energy backscatter also increases monotonically with c_K, but faster than the rate at which viscous dissipation increases so that the net rate of dissipation that results from the combined backscatter and dissipation terms reduces monotonically with c_K, as does the viscous percentage of total dissipation. The kinetic energy level increases monotonically with c_K (not shown).

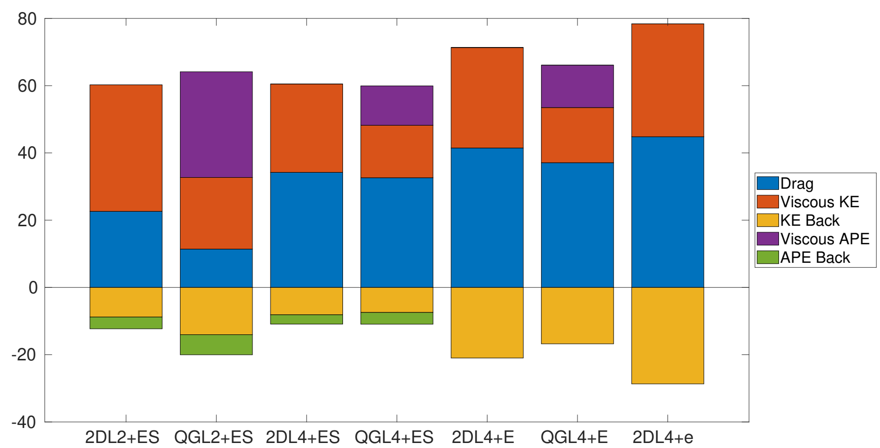

Figure 7: Dissipation and backscatter rates for all of the backscattering models at 24 km resolution. Blue: Frictional KE dissipation rate; Red: viscous KE dissipation rate; Purple: viscous APE dissipation rate; Yellow: KE backscatter rate; Green: APE backscatter rate. The units are $10^6\times$ m^5 s^{-3} (volume integral of rates of change of energy).

3.6. Sensitivity to Υ

This section explores the sensitivity of the 2DL4+E scheme to the coefficient Υ. The coefficient Υ is the ratio of the viscous dissipation scale for enstrophy to the grid scale. Raised to the sixth power, it controls the magnitude of the biharmonic dissipation coefficient. Since the backscatter coefficient is simply $-m\nu_4$, it is clear that the factor of Υ^6 simultaneously controls the amplitude of the backscatter and dissipation.

To investigate the effect of Υ on the dynamics, simulations with the 2DL4+E scheme were run at 16 km resolution with the following values of Υ^6: 1, $(1.2)^6 \approx 3$, $(1.4)^6 \approx 7.5$, $(1.5)^6 \approx 11.4$, $(1.6)^6 \approx 16.8$, and $(1.7)^6 \approx 24$. The time-mean eddy kinetic energy spectra for each of these simulations are shown in Fig. 9. Differences are seen mainly in the small-scale end of the KE spectrum. At small Υ there is too much energy accumulation at small scales, as the parameterization is not strong enough to dissipate enstrophy near the grid scale. Once Υ reaches 1.4 the KE spectrum rolls off more smoothly at the grid scale. At values of Υ larger than 1.5 the total KE begins to degrade, and the viscous percentage of total dissipation also begins to rise (not shown). For this reason $\Upsilon = 1.5$ has been used in all the 2D Leith schemes. Qualitatively similar behavior in the QG schemes led to a choice

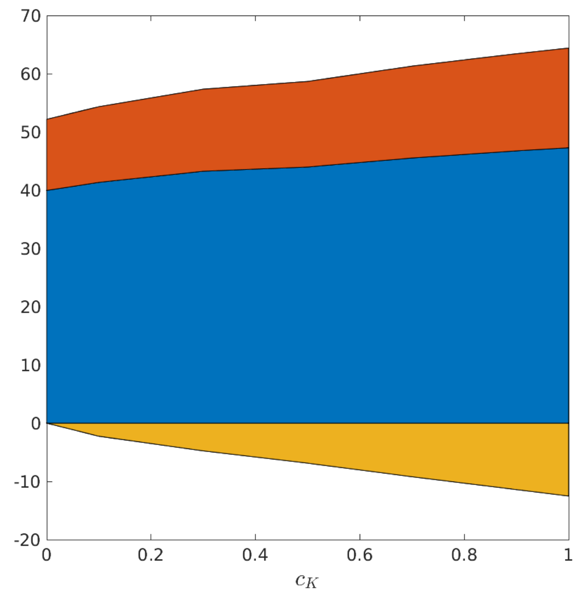

Figure 8: Kinetic energy dissipation and backscatter rates for the 2DL4+E scheme at 16 km resolution as a function of c_K. Blue: Frictional KE dissipation rate; Red: viscous KE dissipation rate; Yellow: KE backscatter rate. The units are $10^6\times$ m^5 s^{-3} (volume integral of rates of change of energy).

of $\Upsilon = 1.3$. The lower value in the QG schemes results from the fact that the QG schemes are more dissipative than their 2D counterparts at the same value of Υ because they dissipate APE in addition to KE.

4. Conclusions

Leith-scaled nonlinear viscosities are promising parameterizations for eddying ocean models (Fox-Kemper and Menemenlis, 2008; Bachman et al., 2017; Pearson et al., 2017). These parameterizations can still dissipate too much energy, similar to their constant-coefficient counterparts, but less severely. One way to rectify excess dissipation by a viscous or diffusive closure is to backscatter some of the dissipated energy. Backscatter parameterizations based on this idea go back to Shutts (2005) and are well-developed in an atmospheric context (Frederiksen and Kepert, 2006; Berner et al., 2008, 2009); parameterizations based on this idea were introduced in ocean modeling by Jansen and Held (2014) and Storto and Andriopoulos (2021). This paper combines these ideas – arresting the forward enstrophy cascade on the grid, and backscattering dissipated energy – to develop six new energetically-constrained Leith parameterizations. The six parameterizations are divided in two groups of three: 2D-Leith parameterizations that are based on enstrophy and dissipate KE, and QG-Leith parameterizations that are based on potential enstrophy and dissipate both KE and APE. Within each group there are two kinds of backscatter: deterministic backscatter using a negative viscosity and stochastic backscatter. Stochastic backscatter can be paired with either harmonic or biharmonic viscosity, but the deterministic backscatter schemes are only paired with biharmonic viscosity. The schemes are compared with each other and with non-backscattering Leith schemes in a six-layer QG double-gyre model across a range of resolutions from 4 km to 24 km.

Although the QG schemes seem to have better theoretical motivation, at least insofar as the QG approximation is a better approximation than 2D dynamics for ocean mesoscales, the QG schemes did not perform as well as their 2D counterparts in the tests reported here: They dissipate too much energy and result in simulations with less total KE, especially at lower eddy-permitting resolutions. The addition of backscatter uniformly improves the results, especially at the lower resolutions. The schemes with harmonic viscosity are uniformly worse than their biharmonic counterparts.

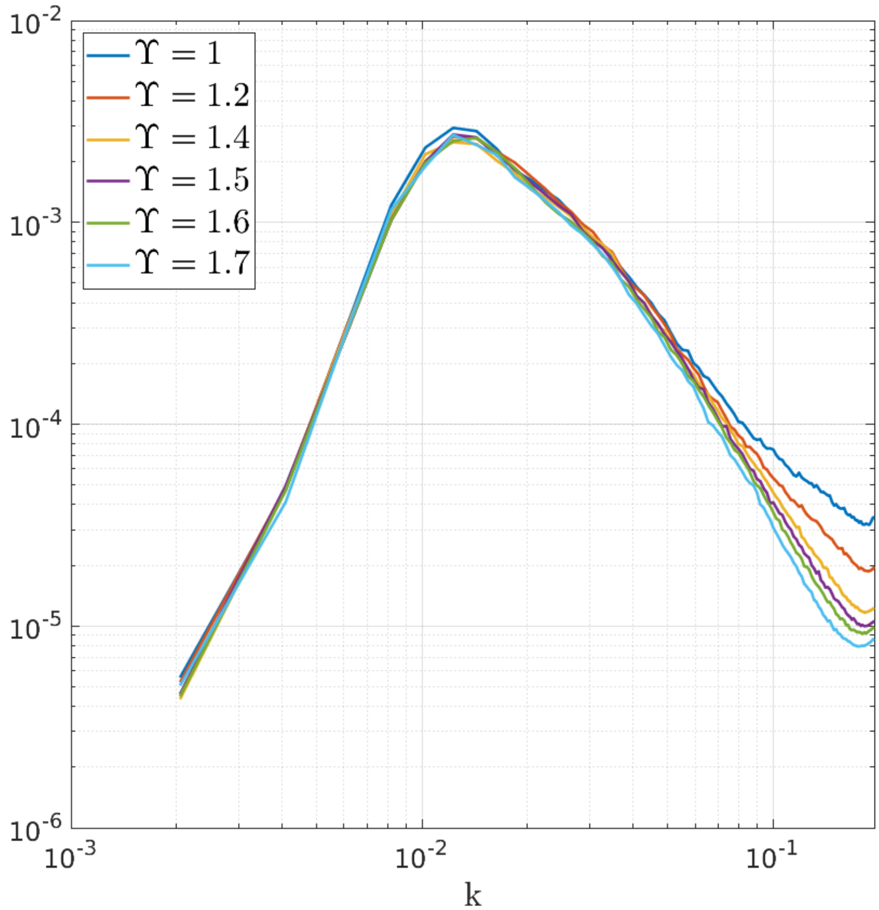

Figure 9: Time-mean eddy kinetic energy spectra from the 2DL4+E model at 16 km resolution for varying values of Υ. Units of the k axis are km^{-1}.

Both stochastic and deterministic backscatter perform well, though the deterministic schemes have an edge in terms of total KE and backscatter efficiency.

Extrapolating towards global ocean models, one strongly expects the addition of backscatter to lead to improvements compared to non-backscattering Leith schemes, even at resolutions that might be considered eddy-resolving. One also expects biharmonic Leith schemes to perform significantly better than harmonic ones, especially at lower resolutions. One does not expect these schemes to perform well at non-eddying resolutions, or even at the lowest eddy-permitting resolutions (e.g. $1/3^\circ$) because they are based on ideas about the forward enstrophy cascade, and need such a cascade to be represented on the resolved scales. Backscatter remains important in the inverse cascade range (Loose et al., 2023), but one expects different parameterizations to be used to represent backscatter on this range of scales (e.g. Bachman, 2019; Jansen et al., 2019).

It will be of interest to compare the performance of the new backscattering Leith schemes to other backscatter schemes for ocean models. Many of these are based on a prognostic budget for subgrid-scale kinetic energy (e.g. Jansen et al., 2015; Klöwer et al., 2018; Jansen et al., 2019; Juricke et al., 2019, 2020b); others include the GM+E parameterization of Bachman (2019), the Stochastic Kinetic Energy Backscatter Scheme (SKEBS) developed by Storto and Andriopoulos (2021) on the basis of the atmospheric scheme of the same name developed by Berner et al. (2009), the kinematic scheme of Juricke et al. (2020a), and the stochastic transport schemes based on the work of Mémin (2014) and Holm (2015). It will also be of interest to modify the schemes so that they backscatter not only the energy dissipated by the Leith closure, but also the potential energy removed from resolved scales by the mixed-layer eddy parameterization of Fox-Kemper et al. (2008). Such a modification might enable an ocean model that does not resolve submesoscales to nevertheless capture the seasonal variations in mesoscale kinetic energy that are driven by seasonal fluctuations in the rates of submesoscale eddy activity (Qiu et al., 2014; Callies et al., 2015; Dong et al., 2020; Steinberg et al., 2022).

Acknowledgements

This research was supported by the NSF OCE grant 1912332. The author thanks two anonymous reviewers for their suggestions, which improved the clarity of the presentation.

Appendix A. QG Model Configuration and Numerics

In each layer the QG PV evolution equation is

$$\partial_t q_n + J[\psi_n, q_n] + \beta v_n = \delta_{1n} F(x,y) - \delta_{6n} r\omega_6 + \mathcal{B}_n + \mathcal{D}_n \tag{A.1}$$

where δ_{ij} is the Kronecker delta, $J[a,b] = (\partial_x a)(\partial_y b) - (\partial_x b)(\partial_y a)$ is the advection term, $\mathcal{B}_n$ is the backscatter term, and $\mathcal{D}_n$ is the horizontal dissipation term. The layer index $n = 1, \ldots, 6$ starts at the top. The form of the backscatter and dissipation terms depends on the parameterization, as detailed in section 2; boundary conditions on the harmonic and biharmonic operators are stress-free. The beta parameter is $\beta = 2 \times 10^{-11}$ m^{-1} s^{-1} and the drag coefficient is $r = 2.2 \times 10^{-7}$ s^{-1}. The forcing F is similar to the asymmetric double-gyre pattern used by Porta Mana and Zanna (2014)

$$F(x,y) = \begin{cases} -\tau_0 \frac{2\pi}{0.9 L_y} \sin\left(\frac{\pi y}{g(x)}\right) \\ \tau_0 \frac{2\pi}{0.9 L_y} \sin\left(\frac{\pi(y - g(x))}{L_y - g(x)}\right) \end{cases} \tag{A.2}$$

where τ_0 is the amplitude of the wind stress forcing, $L_x = L_y = 3{,}072$ km, and

$$g(x) = \frac{L_x}{2} + 0.2\left(x - \frac{L_x}{2}\right). \tag{A.3}$$

The QG PV q_n and streamfunction ψ_n are related by the system of equations

$$q_1 = \nabla^2\psi_1 + \frac{f_0^2}{H_1}\left(\frac{\psi_2 - \psi_1}{g'_1}\right) - \frac{f_0^2}{gH_1}\psi_1 \tag{A.4}$$

$$\text{for } 2 \le n \le 5,\ q_n = \nabla^2\psi_n + \frac{f_0^2}{H_n}\left(\frac{\psi_{n-1} - \psi_n}{g'_{n-1}} - \frac{\psi_n - \psi_{n+1}}{g'_n}\right) \tag{A.5}$$

$$q_6 = \nabla^2\psi_6 + \frac{f_0^2}{H_6}\left(\frac{\psi_5 - \psi_6}{g'_5}\right) \tag{A.6}$$

The Coriolis parameter is $f_0 = 10^{-4}$ s^{-1}. The layer thicknesses are (top to bottom) 385, 289, 269, 305, 482, and 2270 m. The reduced gravities are (top top bottom) 0.0041, 0.0049, 0.0048, 0.0038, and 0.0017 m s^{-2}. The boundary conditions on the QG PV inversion from q to ψ are mass-conserving, i.e. the value of ψ_n on the boundary is set so that $\int_A \psi_n \mathrm{d}A = 0$ where the integral is over the horizontal domain. The baroclinic deformation radii are 30.93, 12.25, 8.24, 6.66, and 5.98 km.

The numerical discretization follows Nadeau and Straub (2009). The discretization is via second-order finite differences, and the nonlinear terms are discretized using the energy and enstrophy conserving method of Arakawa (1966). The QG PV inversion is accomplished by converting to a system of six independent 2D screened-Poisson equations (one for each vertical mode), each of which is solved using a multigrid V(2,2) cycle with red/black Gauss-Seidel smoothing. The time coordinate is discretized using the third-order Adams-Bashforth scheme.

The grid sizes are $\Delta x = 4$, 6.86, 12, 16, and 24 km. The time steps at these resolutions are $\Delta t = 400$, 400, 600, 900, and 1200 s. The maximum biharmonic viscous coefficient is set to $\nu_4^{\max} = \Delta x^4/(320\Delta t)$, and the maximum harmonic coefficient is set to $\nu_2^{\max} = \Delta x^2/(80\Delta t)$. The minimum biharmonic viscous coefficient at each resolution (smallest to largest) is $\nu_4^{\min} = 5 \times 10^6$, 10^7, 1.5×10^8, 6.5×10^8, and 5×10^9 m^4 s^{-1}. The minimum harmonic coefficient at each resolution is $\nu_2^{\min} = 0.15$, 1, 5, 10, and 35 m^2 s^{-1}. In the runs with constant-coefficient biharmonic viscosity, the coefficients are $\nu_4 = 5 \times 10^8$, 3×10^9, 2×10^{10}, 7×10^{10}, and 3×10^{11}.

References

Arakawa, A., 1966. Computational design for long-term numerical integration of the equations of fluid motion: Two-dimensional incompressible flow. Part I. J. Comput. Phys. 1, 119–143.

Bachman, S.D., 2019. The GM+E closure: A framework for coupling backscatter with the Gent and McWilliams parameterization. Ocean Model. 136, 85–106.

Bachman, S.D., Fox-Kemper, B., Pearson, B., 2017. A scale-aware subgrid model for quasi-geostrophic turbulence. J. Geophys. Res.-Oceans 122, 1529–1554.

Barham, W., Bachman, S., Grooms, I., 2018. Some effects of horizontal discretization on linear baroclinic and symmetric instabilities. Ocean Model. 125, 106–116.

Barham, W., Grooms, I., 2019. Exact instantaneous optimals in the non-geostrophic eady problem and the detrimental effects of discretization. Theor. Comp. Fluid Dyn. 33, 125–139.

Berloff, P., 2018. Dynamically consistent parameterization of mesoscale eddies. Part III: Deterministic approach. Ocean Model. 127, 1–15.

Berner, J., Doblas-Reyes, F., Palmer, T., Shutts, G., Weisheimer, A., 2008. Impact of a quasi-stochastic cellular automaton backscatter scheme on the systematic error and seasonal prediction skill of a global climate model. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 366, 2559–2577.

Berner, J., Shutts, G., Leutbecher, M., Palmer, T., 2009. A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in the ECMWF ensemble prediction system. J. Atmos. Sci. 66, 603–626.

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

Böning, C.W., Budich, R.G., 1992. Eddy dynamics in a primitive equation model: Sensitivity to horizontal resolution and friction. J. Phys. Ocean. 22, 361–381.

Callies, J., Ferrari, R., Klymak, J.M., Gula, J., 2015. Seasonality in submesoscale turbulence. Nature communications 6, 6862.

Charney, J.G., 1971. Geostrophic turbulence. J. Atmos. Sci. 28, 1087–1095.

Dong, J., Fox-Kemper, B., Zhang, H., Dong, C., 2020. The seasonality of submesoscale energy production, content, and cascade. Geo. Res. Lett. 47, e2020GL087388.

Fox-Kemper, B., Ferrari, R., Hallberg, R., 2008. Parameterization of mixed layer eddies. part i: Theory and diagnosis. J. Phys. Ocean. 38, 1145–1165.

Fox-Kemper, B., Menemenlis, D., 2008. Can Large Eddy Simulation Techniques Improve Mesoscale Rich Ocean Models?. American Geophysical Union (AGU). pp. 319–337. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/177GM19, doi:https://doi.org/10.1029/177GM19, arXiv:https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/177GM19.

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

Frederiksen, J.S., Kepert, S.M., 2006. Dynamical subgrid-scale parameterizations from direct numerical simulations. J. Atmos. Sci. 63, 3006–3019.

Grooms, I., 2016. A Gaussian-product stochastic Gent-McWilliams parameterization. Ocean Model. 106, 27–43.

Grooms, I., Lee, Y., Majda, A.J., 2015. Numerical schemes for stochastic backscatter in the inverse cascade of quasigeostrophic turbulence. Multiscale Modeling & Simulation 13, 1001–1021.

Grooms, I., Loose, N., Abernathey, R., Steinberg, J., Bachman, S.D., Marques, G., Guillaumin, A.P., Yankovsky, E., 2021. Diffusion-based smoothers for spatial filtering of gridded geophysical data. J. Adv. Model. Earth Syst. 13, e2021MS002552.

Holm, D.D., 2015. Variational principles for stochastic fluid dynamics. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 471, 20140963.

Jansen, M.F., Adcroft, A., Khani, S., Kong, H., 2019. Toward an energetically consistent, resolution aware parameterization of ocean mesoscale eddies. J. Adv. Model. Earth Syst. 11, 2844–2860.

Jansen, M.F., Held, I.M., 2014. Parameterizing subgrid-scale eddy effects using energetically consistent backscatter. Ocean Model. 80, 36–48.

Jansen, M.F., Held, I.M., Adcroft, A., Hallberg, R., 2015. Energy budget-based backscatter in an eddy permitting primitive equation model. Ocean Model. 94, 15–26.

Juricke, S., Danilov, S., Koldunov, N., Oliver, M., Sein, D., Sidorenko, D., Wang, Q., 2020a. A kinematic kinetic energy backscatter parametrization: From implementation to global ocean simulations. J. Adv. Model. Earth Syst. 12, e2020MS002175.

Juricke, S., Danilov, S., Koldunov, N., Oliver, M., Sidorenko, D., 2020b. Ocean kinetic energy backscatter parametrization on unstructured grids:

Impact on global eddy-permitting simulations. J. Adv. Model. Earth Syst. 12, e2019MS001855.

Juricke, S., Danilov, S., Kutsenko, A., Oliver, M., 2019. Ocean kinetic energy backscatter parametrizations on unstructured grids: Impact on mesoscale turbulence in a channel. Ocean Model. 138, 51–67.

Kitsios, V., Frederiksen, J.S., Zidikheri, M.J., 2013. Scaling laws for parameterisations of subgrid eddy–eddy interactions in simulations of oceanic circulations. Ocean Modelling 68, 88–105.

Klöwer, M., Jansen, M.F., Claus, M., Greatbatch, R.J., Thomsen, S., 2018. Energy budget-based backscatter in a shallow water model of a double gyre basin. Ocean Model. 132, 1–11.

Kraichnan, R.H., 1976. Eddy viscosity in two and three dimensions. J. Atmos. Sci. 33, 1521–1536.

Leith, C., 1996. Stochastic models of chaotic systems. Physica D 98, 481–491.

Leith, C.E., 1990. Stochastic backscatter in a subgrid-scale model: Plane shear mixing layer. Physics of Fluids A: Fluid Dynamics (1989-1993) 2, 297–299.

Loose, N., Abernathey, R., Grooms, I., Busecke, J., Guillaumin, A., Yankovsky, E., Marques, G., Steinberg, J., Ross, A.S., Khatri, H., Bachman, S., Zanna, L., Martin, P., 2022. Gcm-filters: A Python package for diffusion-based spatial filtering of gridded data. Journal of Open Source Software 7, 3947. URL: `https://doi.org/10.21105/joss.03947`, doi:`10.21105/joss.03947`.

Loose, N., Bachman, S., Grooms, I., Jansen, M., 2023. Diagnosing scale-dependent energy cycles in a high-resolution isopycnal ocean model. J. Phys. Ocean. 53, 157–176.

Mason, P.J., Thomson, D.J., 1992. Stochastic backscatter in large-eddy simulations of boundary layers. J. Fluid Mech. 242, 51–78.

Mémin, E., 2014. Fluid flow dynamics under location uncertainty. Geophys. Astro. Fluid 108, 119–146.

Nadeau, L.P., Straub, D.N., 2009. Basin and channel contributions to a model Antarctic circumpolar current. J. Phys. Ocean. 39, 986–1002.

Pearson, B., Fox-Kemper, B., Bachman, S., Bryan, F., 2017. Evaluation of scale-aware subgrid mesoscale eddy models in a global eddy-rich model. Ocean Model. 115, 42–58.

Porta Mana, P., Zanna, L., 2014. Toward a stochastic parameterization of ocean mesoscale eddies. Ocean Model. 79, 1–20.

Qiu, B., Chen, S., Klein, P., Sasaki, H., Sasai, Y., 2014. Seasonal mesoscale and submesoscale eddy variability along the north pacific subtropical countercurrent. J. Phys. Ocean. 44, 3079–3098.

Schumann, U., 1995. 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, 293–318.

Semtner, A., Mintz, Y., 1977. Numerical simulation of the gulf stream and mid-ocean eddies. J. Phys. Ocean. 7, 208–230.

Shevchenko, I., Berloff, P., 2015. Multi-layer quasi-geostrophic ocean dynamics in eddy-resolving regimes. Ocean Model. 94, 1–14.

Shevchenko, I., Berloff, P., 2017. On the roles of baroclinic modes in eddy-resolving midlatitude ocean dynamics. Ocean Model. 111, 55–65.

Shutts, G., 2005. A kinetic energy backscatter algorithm for use in ensemble prediction systems. Q. J. Roy. Meteor. Soc. 131, 3079–3102.

Smagorinsky, J., 1963. General circulation experiments with the primitive equations: I. the basic experiment. Mon. Weather Rev. 91, 99–164.

Steinberg, J.M., Cole, S.T., Drushka, K., Abernathey, R.P., 2022. Seasonality of the mesoscale inverse cascade as inferred from global scale-dependent eddy energy observations. J. Phys. Ocean. 52, 1677–1691.

Storto, A., Andriopoulos, P., 2021. A new stochastic ocean physics package and its application to hybrid-covariance data assimilation. Q. J. Roy. Meteor. Soc. 147, 1691–1725.

Sukoriansky, S., Chekhlov, A., Orszag, S.A., Galperin, B., Staroselsky, I., 1996. Large eddy simulation of two-dimensional isotropic turbulence. J. Sci. Comput. 11, 13–45.