An anisotropic subgrid-scale parameterization for large-eddy simulations of stratified turbulence

Sina Khani[*]

Applied Physics Laboratory and School of Oceanography, University of Washington, Seattle, WA, USA

Michael L. Waite

Department of Applied Mathematics, University of Waterloo, Waterloo, ON, Canada

[*]*Corresponding author address:* Applied Physics Laboratory, University of Washington

1013 NE 40th Street, Seattle, WA 98105

E-mail: skhani@uw.edu

ABSTRACT

Subgrid scale (SGS) parameterizations in atmosphere and ocean models are often defined independently in the horizontal and vertical directions because the grid spacing is not the same in these directions (anisotropic grids). In this paper, we introduce a new anisotropic SGS model in large-eddy simulations (LES) of stratified turbulence based on horizontal filtering of the equations of motion. Unlike the common horizontal SGS parameterizations in atmosphere and ocean models, the vertical derivatives of the horizontal SGS fluxes are included in our anisotropic SGS scheme, and therefore the horizontal and vertical SGS dissipation mechanisms are not disconnected in the newly developed model. Our model is tested with two vertical grid spacings and various horizontal grid spacing where the horizontal grid spacing is comparatively larger than that in the vertical. Our anisotropic LES model can successfully reproduce the results of direct numerical simulations, while the computational cost is significantly reduced in the LES. We suggest the new anisotropic SGS model as an alternative to current SGS parameterizations in atmosphere and ocean models, in which the schemes for horizontal and vertical scales are often not connected. The new SGS scheme may improve the dissipative performance of atmosphere and ocean models without adding any backscatter or other energizing terms at small horizontal scales.

2

1. Introduction

Large-eddy simulation (LES) is a useful numerical approach for simulations of geophysical turbulence at the small-scale end of the atmospheric mesoscale, oceanic submesoscale, and smaller scales, where there is forward kinetic energy transfer from large to small horizontal scales. These scale ranges are broadly characterized by strong stratification and weak rotation, i.e. stratified turbulence (e.g. Riley & Lindborg 2008). It has been shown that fundamental characteristics of stratified turbulence that are seen in direct numerical simulation (DNS), such as a $-5/3$ spectral slope in the horizontal wavenumber energy spectra (Waite & Bartello 2004; Lindborg 2006; Brethouwer et al. 2007), layered structures with Kelvin-Helmholtz (KH) instabilities (Bartello & Tobias 2013; Khani & Waite 2016), non-local horizontal energy transfer from large scales to the buoyancy scale (Waite 2011; Khani & Waite 2013), and small or negative local Richardson number associated with overturning (Waite & Bartello 2004; Bartello & Tobias 2013), can be captured by LES if the buoyancy scale $L_b = 2\pi u_{rms}/N$ is sufficiently well resolved (i.e. $\Delta < L_b$; see Khani & Waite 2014, 2015). Here, u_{rms}, N and Δ are the root-mean-square (r.m.s.) velocity, buoyancy frequency and grid spacing, respectively. These and other LES (e.g. Siegel & Domaradzki 1994; Kang et al. 2003; Paoli et al. 2014) used isotropic grid spacing in the horizontal and vertical directions, i.e. $\Delta_h = \Delta_z$. However, coarser and anisotropic grid spacings are usually employed in atmosphere and ocean simulations, which require a different subgrid approach.

In large-scale atmosphere and ocean models, it is not possible to explicitly resolve the buoyancy scale L_b in the horizontal direction, due to limits on computation. For example, the horizontal grid spacing Δ_h in global weather prediction models is often around 10 km or larger, while L_b is on the order of 1 km in the atmosphere (see e.g. Augier & Lindborg 2013; Brune & Becker 2013; Schaefer-Rolffs & Becker 2018). Therefore, it is not computationally feasible to use isotropic grid

spacing to resolve L_b, and therefore different grid spacings in the horizontal and vertical directions (i.e. anisotropic grids) are usually employed. Atmosphere and ocean models typically use finer grid spacing in the vertical direction than in the horizontal (i.e. $\Delta_z < \Delta_h$), and the question of sufficient vertical resolution for capturing the $-5/3$ power law in the horizontal wavenumber energy spectrum has been an active area of discussion (see e.g. Brune & Becker 2013; Augier & Lindborg 2013; Waite 2016; Schaefer-Rolffs & Becker 2018; Skamarock et al. 2019). Generally, most atmosphere and ocean models use different dissipation schemes in the horizontal and vertical directions because of the large difference in horizontal and vertical grid spacings. These dissipation schemes are typically independent of one another: for example, models may use the horizontal Smagorinsky subgrid-scale (SGS) model for horizontal mixing and a vertical stability-dependent eddy viscosity, possibly as part of the boundary layer scheme, for vertical mixing (see e.g. Skamarock et al. 2008; Griffies & Hallberg 2000, for the Weather Research and Forecasting (WRF) model and the Modular Ocean Model (MOM), respectively).

In this paper, we use homogeneous stratified turbulence as an idealized problem in which to investigate the consequences of using disconnected horizontal and vertical SGS dissipation schemes in the limit of high vertical resolution. We develop an anisotropic scheme for LES of stratified turbulence based on a scale analysis of the SGS momentum and potential temperature fluxes in stratified turbulence. Initially, we set Δ_z very small (as in DNS) in our LES runs to evaluate the dependence of our new anisotropic scheme on the horizontal grid spacing Δ_h by comparison with a more typical SGS scheme, in which the horizontal and vertical dissipations are treated separately. Next, we study the effects of vertical resolutions in our new scheme. The rest of this paper is composed as follows: the governing equations and mathematical formulations are given in section 2. Section 3 presents the methodology and numerical setup. Results are shown and discussed in section 4, followed by conclusions in section 5.

4

2. Governing equations

The governing equations of motion under the Boussinesq approximation with uniform stratification can be written in the following non-dimensional form (as in e.g. Khani & Waite 2013)

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} = -\nabla p + \frac{1}{Fr_\ell{}^2}\theta \mathbf{e}_z + \frac{1}{Re_\ell}\nabla^2 \mathbf{u}, \tag{1}$$

$$\nabla \cdot \mathbf{u} = 0, \tag{2}$$

$$\frac{\partial \theta}{\partial t} + \mathbf{u} \cdot \nabla \theta + w = \frac{1}{Re_\ell Pr}\nabla^2 \theta, \tag{3}$$

where $\mathbf{u} = (u, v, w)$ is the velocity vector, θ and p are the potential temperature and pressure perturbations, respectively; and $Re_\ell = \mathcal{U}\ell/\nu$, $Fr_\ell = \mathcal{U}/(\ell N)$, and $Pr = \nu/\kappa$ are the Reynolds, Froude and Prandtl numbers, respectively. Here, the velocity and length scales are taken to be unity (i.e. $\mathcal{U} \equiv 1$ and $\ell \equiv 1$), and ν and κ are the molecular viscosity and diffusivity, respectively. In LES, the flow variables are filtered using a filtering operator G. For example, the filtered velocity field $\bar{\mathbf{u}}$ is defined as follows (e.g. Pope 2000)

$$\bar{\mathbf{u}}(\mathbf{x}, t) = \int_D \mathbf{u}(\mathbf{x} + \mathbf{r}, t)G(\mathbf{r})\mathrm{d}\mathbf{r}, \tag{4}$$

where $\mathbf{x} = (x, y, z)$ are the Cartesian coordinates and D is the spatial domain. Applying the filtering operator G to the equations of motions (1-3) is straightforward except for the nonlinear terms, which lead to the subgrid scale (SGS) momentum stress

$$\tau_{ij} = \overline{u_i u_j} - \bar{u}_i \bar{u}_j, \tag{5}$$

and SGS potential temperature flux

$$h_j = \overline{\theta u_j} - \bar{\theta}\bar{u}_j, \tag{6}$$

which are not known in terms of the filtered variables and must be parameterized in LES. In summary, the filtered Navier-Stokes equations under the Boussinesq approximation can be written

5

as

$$\frac{\partial \bar{u}_i}{\partial t} + \frac{\partial}{\partial x_j}(\bar{u}_i \bar{u}_j) = -\frac{\partial \bar{p}}{\partial x_i} + \frac{1}{Fr_\ell{}^2}\bar{\theta}\mathbf{e}_z + \frac{1}{Re_\ell}\frac{\partial^2 \bar{u}_i}{\partial x_j \partial x_j} - \frac{\partial \tau_{ij}}{\partial x_j}, \tag{7}$$

$$\frac{\partial \bar{u}_i}{\partial x_i} = 0, \tag{8}$$

$$\frac{\partial \bar{\theta}}{\partial t} + \frac{\partial}{\partial x_j}(\bar{\theta}\bar{u}_j) + \bar{w} = \frac{1}{Re_\ell Pr}\frac{\partial \bar{\theta}}{\partial x_j \partial x_j} - \frac{\partial h_j}{\partial x_j}. \tag{9}$$

Before introducing SGS models for the momentum and potential temperature fluxes, it is useful to investigate the SGS momentum stress τ_{ij} using Taylor series and the definition of the filtering operator. A similar procedure is also done for the potential temperature SGS flux h_j.

We can expand the velocity field $\mathbf{u}(\mathbf{x}+\mathbf{r})$ using a Taylor series at a given point $\mathbf{x}$ in $\mathbf{r}$, which is on the order of the filter width Δ (see e.g. Pope 2000; Meneveau & Katz 2000; Khani & Porté-Agel 2017a,b)

$$u_i(\mathbf{x}+\mathbf{r}) = u_i(\mathbf{x}) + \frac{\partial u_i(\mathbf{x})}{\partial x_k}r_k + \frac{1}{2}\frac{\partial^2 u_i(\mathbf{x})}{\partial x_k \partial x_l}r_k r_l + \frac{1}{6}\frac{\partial^3 u_i(\mathbf{x})}{\partial x_m \partial x_k \partial x_l}r_m r_k r_l + \dots \tag{10}$$

Using this expansion, we can find the following nonlinear approximation for the SGS stress tensor (see appendix A for the details)

$$\tau_{ij}(\mathbf{x}) \approx \frac{\partial u_i(\mathbf{x})}{\partial x_l}\frac{\partial u_j(\mathbf{x})}{\partial x_k}\int_D r_k r_l G(\mathbf{r})\,d\mathbf{r}. \tag{11}$$

The SGS stress τ_{ij} depends on the filtering function G and the integral over the associated domain. For example, if G is an isotropic Gaussian function with variance $\Delta^2/12$, equation (11) yields

$$\tau_{ij}(\mathbf{x}) \approx \frac{\Delta^2}{12}\frac{\partial u_i(\mathbf{x})}{\partial x_k}\frac{\partial u_j(\mathbf{x})}{\partial x_k}. \tag{12}$$

Horizontal SGS mixing parameterizes the effects of small unresolved horizontal scales. As a result, it can be investigated by applying a filter to horizontal scales only, i.e. $G = G(r_x, r_y)$. In this case, the dummy indices l and k in equation (11) will span $\{1,2\}$ and equation (11) will not include

6

z derivatives. Therefore, using an isotropic horizontal Gaussian filter, equation (12) becomes

$$\tau_{ij}(\mathbf{x}) \approx \frac{\Delta^2}{12} \left(\frac{\partial u_i(\mathbf{x})}{\partial x} \frac{\partial u_j(\mathbf{x})}{\partial x} + \frac{\partial u_i(\mathbf{x})}{\partial y} \frac{\partial u_j(\mathbf{x})}{\partial y} \right), \quad i,j = 1,2,3. \tag{13}$$

The vertical components of the SGS stress τ_{ij} (i.e. τ_{13}, τ_{23} and τ_{33}) are not zero because the horizontal derivatives of vertical motions (i.e. $\partial w/\partial x$ and $\partial w/\partial y$) are non-zero.

Similarly, for the SGS flux term $h_j(\mathbf{x}) = \overline{\theta(\mathbf{x}) u_j(\mathbf{x})} - \bar{\theta}(\mathbf{x}) \bar{u}_j(\mathbf{x})$, we can write (see appendix B for more information)

$$h_j(\mathbf{x}) \approx \frac{\partial \theta(\mathbf{x})}{\partial x_l} \frac{\partial u_j(\mathbf{x})}{\partial x_k} \int_D r_k r_l G(\mathbf{r}) \, d\mathbf{r}. \tag{14}$$

Again, assuming a horizontal Gaussian filter function G, we get

$$h_j(\mathbf{x}) \approx \frac{\Delta^2}{12} \left(\frac{\partial \theta(\mathbf{x})}{\partial x} \frac{\partial u_j(\mathbf{x})}{\partial x} + \frac{\partial \theta(\mathbf{x})}{\partial y} \frac{\partial u_j(\mathbf{x})}{\partial y} \right), \tag{15}$$

where the vertical component of the SGS potential temperature flux h_j is non-zero because $\partial w/\partial x$ and $\partial w/\partial y$ are non-zero [similar to equation (13)]. Note that we do not use equations (13) and (15) as a parameterization, but rather as a guide to determine what terms in τ_{ij} and h_j should be retained and parameterized. Overall, from equations (13) and (15) it is clear that the vertical component of the SGS flux is not zero even when the focus is on only unresolved horizontal scales, i.e. with a purely horizontal filtering operator G.

We use scale analysis to estimate the size of the various terms in equations (13) and (15) in geophysical simulations. Let l_h and l_z be the horizontal and vertical scales, respectively, where $l_z \ll l_h$. In this case, the horizontal and vertical components of the SGS tensor τ_{ij} can be scaled as (recall that τ_{ij} is symmetric)

$$\tau_{ij} \sim \begin{cases} \Delta^2 \left(\frac{U}{l_h} \right)^2 & \text{for} \quad i,j = 1,2 \\ \Delta^2 \left(\frac{U}{l_h} \right)^2 \left(\frac{l_z}{l_h} \right) & \text{for} \quad i = 1,2 \text{ and } j = 3 \\ \Delta^2 \left(\frac{U}{l_h} \right)^2 \left(\frac{l_z}{l_h} \right)^2 & \text{for} \quad i,j = 3, \end{cases} \tag{16}$$

7

where $\sim$ denotes order of magnitude, U is the horizontal velocity scale, and we have used the continuity equation to scale the vertical velocity as $l_z U / l_h$ (as in e.g. Riley & Lelong 2000). Using a similar scale analysis, the SGS stress tensor divergence in equation (7) can be scaled as

$$
\frac{\partial \tau_{ij}}{\partial x_j} \sim
\begin{cases}
\frac{\Delta^2}{l_h} \left(\frac{U}{l_h} \right)^2 & \text{for} \quad i, j = 1, 2 \\[2mm]
\frac{\Delta^2}{l_h} \left(\frac{U}{l_h} \right)^2 & \text{for} \quad i = 1, 2 \text{ and } j = 3 \\[2mm]
\frac{\Delta^2}{l_h} \left(\frac{U}{l_h} \right)^2 \left(\frac{l_z}{l_h} \right) & \text{for} \quad i, j = 3.
\end{cases}
\tag{17}
$$

The horizontal derivatives of horizontal stress ($i, j = 1, 2$) and vertical derivatives of SGS stresses with $i = 1, 2$ and $j = 3$ are of the same order of magnitude, and therefore the latter terms are not negligible in comparison with the former when a horizontal filter function is employed. Yet these terms, $\partial \tau_{13}/\partial x_3$ and $\partial \tau_{23}/\partial x_3$, are not included in purely horizontal mixing schemes (e.g. horizontal Smagorinsky in WRF; see Skamarock et al. 2008). Similarly, we can scale the SGS potential temperature flux divergence $\partial h_j/\partial x_j$ as

$$
\frac{\partial h_j}{\partial x_j} \sim
\begin{cases}
\frac{\Delta^2}{l_h} \left(\frac{\Theta U}{l_h} \right)^2 & \text{for} \quad j = 1, 2 \\[2mm]
\frac{\Delta^2}{l_h} \left(\frac{\Theta U}{l_h} \right)^2 & \text{for} \quad j = 3
\end{cases}
\tag{18}
$$

where Θ is the potential temperature scale. Again, both horizontal and vertical derivatives of h_j are of the same order of magnitude.

The SGS term $\partial \tau_{ij}/\partial x_j$ includes the following terms in the x, y and z directions, respectively

$$
\left(\frac{\partial \tau_{11}}{\partial x} + \frac{\partial \tau_{12}}{\partial y} + \frac{\partial \tau_{13}}{\partial z} \right), \qquad x\text{-direction} \tag{19}
$$

$$
\left(\frac{\partial \tau_{12}}{\partial x} + \frac{\partial \tau_{22}}{\partial y} + \frac{\partial \tau_{23}}{\partial z} \right), \qquad y\text{-direction} \tag{20}
$$

$$
\left(\frac{\partial \tau_{13}}{\partial x} + \frac{\partial \tau_{23}}{\partial y} \right), \qquad z\text{-direction} \tag{21}
$$

where only the term $\partial \tau_{33}/\partial z$ is negligible [see equation (17)]. Also, the potential temperature flux $\partial h_j/\partial x_j$ includes the following terms

$$\left(\frac{\partial h_1}{\partial x} + \frac{\partial h_2}{\partial y} + \frac{\partial h_3}{\partial z} \right), \tag{22}$$

where all terms are important [see equation (18)]. In later sections, we will perform LES runs with anisotropic dissipation following equations (19) to (22), and compare the results with DNS, and classic LES where the horizontal and vertical dissipation schemes are not connected (i.e. the vertical components τ_{13}, τ_{23} and h_3 are omitted).

3. Methodology

We consider a domain with periodic boundary conditions. The horizontal side length is $\mathcal{L}_h = 2\pi$, and the vertical side length, depending on the simulation, is $\mathcal{L}_v = 2\pi$ or π. Decaying stratified turbulence is simulated: we have scaled the velocity amplitude to set an initial energy of 0.13 with random phases over a spherical wavevector shell $k_i - 0.5 < k \leq k_i + 0.5$, where $k = \sqrt{k_x^2 + k_y^2 + k_z^2}$ is the total wavenumber and $k_i = 3$ is the initial wavenumber (a similar initial condition is used in Bartello & Tobias 2013). The initial potential temperature fluctuation is zero. The buoyancy frequency is $N = 2.1$ or 4.2, both of which ensure small initial Froude numbers in our simulations (i.e. $Fr_\ell = 1/N$). Grid sizes vary from $n_x = n_y = 240$ to 960. For the DNS, $n_z = n_x$ (for $\mathcal{L}_v = 2\pi$) or $n_z = n_x/2$ (for $\mathcal{L}_v = \pi$), where n_i is the number of grid points in the $i = x$, y and z directions. Spatial derivatives are discretized using the spectral transform method, and the two-thirds rule (Orszag 1971) is employed to eliminate aliasing errors, which leads to an effective grid spacing $\Delta_h = 3\mathcal{L}_h/(2n_x)$ and $\Delta_z = 3\mathcal{L}_v/(2n_z)$. The third-order Adams-Bashforth scheme is employed for the time stepping of all terms in the transport equations, except the molecular dissipation terms, which are treated with a Crank-Nicolson approach (see e.g. Durran 2010).

9

A new anisotropic[1] LES method, in which the vertical derivatives of the SGS stress and flux are retained as shown in equations (19-22), is tested. We employ the dynamic Smagorinsky SGS model because it has the best performance in comparison with other SGS parameterizations in LES of stratified turbulence (see Khani & Waite 2014, 2015). The eddy viscosity and diffusivity terms in the anisotropic dynamic Smagorinsky model are given as follows

$$\frac{\partial}{\partial x}\left(2\bar{S}\bar{s}_{11}\right) + \frac{\partial}{\partial y}\left(2\bar{S}\bar{s}_{12}\right) + \frac{\partial}{\partial z}\left(2\bar{S}\bar{s}_{13}\right), \qquad x\text{-direction momentum Eq.} \qquad (23)$$

$$\frac{\partial}{\partial x}\left(2\bar{S}\bar{s}_{12}\right) + \frac{\partial}{\partial y}\left(2\bar{S}\bar{s}_{22}\right) + \frac{\partial}{\partial z}\left(2\bar{S}\bar{s}_{23}\right), \qquad y\text{-direction momentum Eq.} \qquad (24)$$

$$\frac{\partial}{\partial x}\left(2\bar{S}\bar{s}_{13}\right) + \frac{\partial}{\partial y}\left(2\bar{S}\bar{s}_{23}\right), \qquad z\text{-direction momentum Eq.} \qquad (25)$$

and

$$\frac{\partial}{\partial x}\left(\frac{\bar{S}}{Pr_t}\frac{\partial\bar{\theta}}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{\bar{S}}{Pr_t}\frac{\partial\bar{\theta}}{\partial y}\right) + \frac{\partial}{\partial z}\left(\frac{\bar{S}}{Pr_t}\frac{\partial\bar{\theta}}{\partial z}\right), \qquad \text{potential temperature equation} \qquad (26)$$

where

$$\bar{s}_{ij} = (1/2)\left(\partial\bar{u}_i/\partial x_j + \partial\bar{u}_j/\partial x_i\right), \qquad (27)$$

is the rate-of-strain tensor, $\bar{S} = -c_s\Delta_h^2\left(2\bar{s}_{ij}\bar{s}_{ij}\right)^{1/2}$ is proportional to the total strain rate, c_s is the time and space dependent Smagorinsky coefficient, which is calculated using an isotropic explicit coarse filter scale $\tilde{\Delta} = 2\Delta_h$ (negative c_s values are clipped, see Khani & Waite 2015); the turbulent Prandtl number $Pr_t = 1$. Since the anisotropic LES approach is based on horizontal-only filtering, these simulations use high (DNS) resolution in the vertical.

For comparison, classic LES and DNS runs are also performed. In the classic LES runs, the vertical dissipation scale is resolved with high (DNS) vertical resolution, and therefore the vertical derivative terms in equations (23-26) are neglected in the classic horizontal SGS parameterization. In the DNS runs, high resolution is used in all directions and no SGS model is included. The DNS

---

[1]Here, "anisotropic" refers to a SGS closure based on horizontal filtering only.

resolution of the Kolmogorov scale L_d is $k_{max}/k_d \gtrsim 0.67$, which is in line with the criterion given by Moin & Mahesh (1998). Here, k_{max} is the maximum wavenumber and k_d is the Kolmogorov wavenumber, calculated over the time interval of maximum kinetic energy dissipation. Our setup with small Δ_z (DNS) is not standard in atmosphere and ocean simulations – indeed, the vertical grids in such models do not ever resolve the Kolmogorov scale – but it is a sensible experiment in which to investigate the parameterization of subgrid horizontal scales. In current SGS models, which we call here classic LES, the limit of $\Delta_z \to 0$ would shut off the requirement for a vertical mixing scheme; here we show that simulations with the classic LES model will fail in such a limit, for example, the kinetic energy spectra are under-dissipated (see below), thereby illustrating a problem with the classic approach.

Simulation results are averaged over a time interval around which the kinetic energy dissipation rate ε is maximum. The r.m.s. velocity $u_{rms} = \sqrt{\langle E(t) \rangle}$, where the angle bracket $\langle \cdot \rangle$ denotes time-averaging and $E(t)$ is the domain-averaged kinetic energy. The buoyancy scale $L_b = 2\pi u_{rms}/N$, the Ozmidov scale $L_o = 2\pi \left(\varepsilon/N^3 \right)^{1/2}$, and the Kolmogorov scale $L_d = 2\pi \left(v^3/\varepsilon \right)^{1/4}$. In our LES, unless there is no eddy viscosity/diffusivity terms (i.e. in cases with high vertical resolution Δ_z), molecular viscosity and diffusivity are switched off. Table 1 shows a list of averaged variables and parameters for the DNS and anisotropic LES runs[2]. In the rest of this paper, simulation names start with 'D' for DNS, 'LA' for anisotropic LES with the same vertical resolution as DNS but coarser horizontal resolution, and 'LAV' for anisotropic LES with the half vertical resolution compared with 'LA' runs (see table 1). For LAV runs, the Kolmogorov scale is not resolved in the vertical direction, and there is no extra vertical SGS parameterization (only the anisotropic SGS model is employed). The classic LES simulations, which are labeled with 'LC' runs (see the caption of figure 5), are not reported in table 1 because these runs are under-dissipated (see below).

---

[2]Simulations in table 1 are named by their initial Reynolds numbers Re_ℓ and buoyancy frequencies $N (= 1/Fr_\ell)$.

4. Results and discussion

a. Overview of simulations

Figure 1 shows the time evolution of the total, kinetic and potential energy in panels (a-d), respectively, for the DNS and anisotropic LES runs. The total energy is almost constant up to approximately $t = 5$, and then it decays due to the onset of turbulence (see below). The time series of the kinetic and potential energy (KE and PE) show oscillations, mainly before turbulence decay, due to buoyancy exchanges between KE and PE, since only KE is present at $t = 0$ (figures 1c,d). The oscillation time scale is related to the frequency of the gravity waves excited by the initial conditions, and therefore becomes smaller in the case with stronger stratification (figures 1c,d). The anisotropic LES runs correctly capture the energy oscillations and onset of dissipation from the DNS runs, although the anisotropic LES cases with coarser horizontal resolution (i.e. cases LA18N2b and LA22N4b) slightly underestimate the total energy level after the occurrence of turbulence, which is due to larger eddy dissipation in these coarse LES simulations. Moreover, the onset of turbulence happens earlier in the anisotropic LES cases with the lowest horizontal resolution (see green and magenta dash-dot lines in figures 1a,b).

The kinetic energy dissipation rate ε for cases with the initial Reynolds number $Re_\ell = 18000$, buoyancy frequency $N = 2.1$, and $Re_\ell = 22200$, $N = 4.2$ are displayed in figures 2(a,b), respectively. The solid black line in figure 2(a) shows ε for the DNS run, to which we compare the ε in the anisotropic LES cases (red dash and green dash-dot lines). A similar comparison is provided in figure 2(b), in which the solid gray line shows ε for the DNS case, and the blue dash and magenta dash-dot lines show ε in the anisotropic LES runs. The kinetic energy dissipation rate has a maximum around $t = 7$ and $t = 8$ in figures 2(a,b), respectively, for the DNS runs. These maxima give the approximate time at which turbulence onset occurs. The onset time for turbulence is

12

relatively well estimated by the anisotropic LES cases with finer horizontal resolutions (i.e. cases LA18N2a and LA22N4a), although the magnitudes of ε at the maximum times are a little higher in these LES cases in comparison with the DNS. If the horizontal grid spacing decreases further in the anisotropic LES cases (LA18N2b and LA22N4b), turbulence onset occurs earlier while the maximum ε values would be around or a little smaller than the corresponding DNS runs (solid versus dash-dot line in figure 2). Overall, the anisotropic LES runs show larger ε at early times compared to the DNS cases, but differences between the kinetic energy dissipation rates in the DNS and anisotropic LES cases are smaller after the onset of turbulence (figure 2).

Figures 3 and 4 show the y-component of vorticity on the x-z plane at $y = 0.25$ and $t = 15$ for weak and strong stratification cases, respectively (panel a for DNS and panels b,c for anisotropic LES). In the DNS, the vorticity snapshot shows layers, KH instabilities, and regions of more isotropic small-scale turbulence (see, for example, regions around $z = 2$ and $x = [0 - 3]$, or $z \approx 0.8$ and $x = [5 - 6]$ in figure 4a). Layering is more pronounced in the simulation with larger stratification, which has smaller Re_b and is therefore more influenced by dissipation (as shown in figure 4a, and also see Brethouwer et al. 2007; Bartello & Tobias 2013; Khani & Waite 2014). For example, figure 3(b) depicts many regions with small-scale isotropic turbulence, while layered structures are more visible in figure 4(b) where stratification is increased. If we further decrease the horizontal resolution in the anisotropic LES, similar large-scale structures are generally seen in both the weak and strong stratification cases. Interestingly, the anisotropic LES runs with larger grid spacing (i.e. coarser resolution compared to DNS) reproduce similar structures that are seen in the DNS runs with smaller grid spacing. The horizontal layers are much more pronounced in these low resolution simulations due to significantly larger horizontal dissipation because Δ_h is much larger here, which reduces the transition to small-scale isotropy (figures 3c and 4c).

13

b. Kinetic energy spectra

The horizontal and vertical wavenumber kinetic energy spectra for the DNS, anisotropic LES,

and classic LES runs are shown in figure 5 (panels a,b show simulations with $Re_\ell = 18000$ and

$N = 2.1$, and panels c,d show simulations with $Re_\ell = 22200$ and $N = 4.2$). The spectra are av-

eraged over a time interval $\Delta t = 4$ around the maximum kinetic energy dissipation rate. The

high-resolution anisotropic LES (LA18N2a and LA22N4a cases) show almost identical vertical

wavenumber kinetic energy spectra to those for DNS (red and blue dashed versus black and gray

solid lines in figures 5b,d). This trend may not be unexpected since both the DNS and high-

resolution anisotropic LES have the same vertical resolution, but the horizontal resolutions are

different. The horizontal wavenumber kinetic energy spectra of the anisotropic LES and DNS are

also very similar with higher horizontal resolution in the LES (i.e. $\Delta_h^{LA} = 2\Delta_h^D$; red and blue dashed

versus black and gray solid lines in figures 5a,b). If we further reduce the horizontal resolution in

the anisotropic LES to $4\Delta_h^D$, we still can obtain reasonable results at large scales, although both

the vertical and horizontal wavenumber kinetic energy spectra are less energetic at smaller scales

due to the larger eddy dissipation at small horizontal scales (see green and magenta dash-dotted

lines in figure 5). Overall, the anisotropic LES model is able to capture the inertial subrange of

stratified turbulence similar to the DNS, with significantly less computational cost. Nevertheless,

the kinetic energy spectra in the coarse horizontal-resolution cases (LA18N2b and LA22N4b) are

more steep at small resolved scales.

The coarser resolution anisotropic LES cases have more dissipation at small vertical scales, as

evidenced by the steeper vertical spectra, in comparison with DNS (or high-resolution anisotropic

LES), despite the fact that they have the same Δ_z. This behavior suggests that horizontal resolution

can have a significant impact on the resolution of small vertical scales in LES of stratified turbu-

lence, and that the dissipation mechanisms in the horizontal and vertical directions are actually connected.

If the vertical derivatives of the SGS stress and flux are neglected in our LES runs (i.e. classic LES), the impact on the kinetic energy spectra are significant. Indeed, the spectra are under-dissipated; there is insufficient small-scale dissipation and, as a result, energy accumulates around the smallest resolved scales in both horizontal and vertical wavenumber spectra (see cyan lines in figure 5). The vertical derivatives of the SGS flux, which are missing in the classic LES runs, therefore play an important role in removing energy from small horizontal and vertical scales. Neglecting these terms can lead to unrealistic results, even with fine (DNS) grid spacings in the vertical. Overall, the results of this section show that the scale analyses in equations (16) and (17), which lead to the anisotropic LES parameterizations that are shown by (19-22), are confirmed using numerical simulations.

If we further reduce the horizontal resolution in the anisotropic LES model, the results become under-dissipated when $\Delta_h/\Delta_z > 4$ (not shown). This trend suggest that the ratio Δ_h/Δ_z can also play a role in the anisotropic LES runs. To investigate this point further, we consider a series of additional anisotropic LES runs in the case with $Re_\ell = 18000$ and $N = 2.1$, for which the vertical grid spacing is double the vertical grid spacing of the DNS runs, with different horizontal resolution (these runs are labeled with 'LAV'). Figure 6 shows the horizontal and vertical wavenumber kinetic energy spectra with $\Delta_z^{LAV} = 2\Delta_z^D$, where cases b, c, d show simulations with different horizontal resolutions: i.e. $\Delta_{h,b}^{LAV} = 1/2\Delta_{h,c}^{LAV} = 2/3\Delta_{h,d}^{LAV} = \Delta_{h,b}^{LA}$ (see table 1). We should mention here that the coarse horizontal resolution in cases c,d gave under-dissipated results with high vertical resolution Δ_z^D (not shown), but interestingly, these cases are not under-dissipated when coarser vertical grid spacing $2\Delta_z^D$ is used (figure 6). Anisotropic LES with coarser vertical resolution (LAV) and DNS curves show very similar kinetic energy spectra when k_h and k_v are $\lesssim 20$. For larger

15

wavenumbers, however, LAV runs are much more dissipative than the DNS. This behavior is expected because both horizontal and vertical grids are larger than those in DNS; for example, the curve with dash-dot green line in figure 6 is for a simulation with $\Delta_z^{LAV} = 2\Delta_z^D$ and $\Delta_h^{LAV} = 8\Delta_h^D$. In this case, we can significantly save on computational resources since this anisotropic LES run is ≈ 300 times cheaper than the DNS run because of having larger grid spacing and time step for running simulations. From DNS to anisotropic LES, we can successfully reduce computational costs remarkably while the accuracy of results are still high.

Overall, we found that retaining the vertical derivatives of the SGS fluxes as given by equations (19-22) is a key in our anisotropic LES scheme to reproduce DNS results. Also, our results show that by increasing the ratio of horizontal to vertical grid spacing in the anisotropic LES model, we can help to prohibit under-dissipative LES results. Nevertheless, this latter parameter setup needs to be further investigated in the realistic atmosphere and ocean models with our new horizontal dissipation scheme along with an appropriate vertical mixing scheme. Implementing this parameterization in atmosphere and ocean models would require geometrical adjustments for spherical coordinates (some geometrical modifications for diffusion coefficients in spherical geometry have been introduced, as in e.g. Gordon & Stern 1982; Smagorinsky 1993; Becker & Burkhardt 2007).

c. Mixing efficiency

Mixing efficiency is a key parameter in atmospheric sciences and physical oceanography, where breaking internal waves in stratified shear layers and diapycnal mixing in the upper ocean are significantly influenced by the efficiency of turbulent mixing (see e.g. Riley & Lelong 2000; Gregg et al. 2018). The irreversible mixing efficiency γ_i is defined as the ratio of the potential energy dissipation to the total dissipation rates $\varepsilon_p / (\varepsilon + \varepsilon_p)$ (Winters & D'Asaro 1996; Caulfield & Peltier 2000). This quantity has been extended to be used in LES with SGS dissipation rates (Khani

16

310 2018). The SGS mixing efficiency γ_i depends on the turbulent Prandtl number Pr_t as follows

$$\gamma_i \sim \frac{1}{1+2Pr_t}, \tag{28}$$

311 implying $\gamma_i \approx 1/3$ in stratified turbulence with $Pr_t = 1$ (Khani 2018).

312 Figure 7 shows the irreversible mixing efficiency γ_i versus the resolution of the Ozmidov scale

313 L_o in the horizontal direction, for DNS and anisotropic LES approaches. As expected, the ratio

314 L_o/Δ_h is larger in DNS compared to that in LES. Nevertheless, values of γ_i in LES overlap well

315 with those from DNS and are in line with the theoretical estimate of $1/3$ for LES of stratified

316 turbulence (figure 7). This agreement is due to the resolution of the Ozmidov scale L_o in LES runs

317 (see Khani (2018) for more information). Noteworthy, unlike in the simulations in Khani (2018),

318 where only large-scale vortical modes were initially excited, here we excite large horizontal and

319 vertical motions, which results in more efficient energy exchange between KE and PE through the

320 buoyancy fluxes. In this case, γ_i is slightly larger than $1/3$ for both DNS and LES runs (figure 7).

5. Conclusions

322 A new anisotropic SGS model in LES of stratified turbulence is introduced. The new scheme

323 uses coarse grid spacing in the horizontal direction, and also retains the vertical derivatives of

324 horizontal motions in the eddy dissipation terms, which are omitted in the classic LES approach

325 for horizontal dissipation. Therefore, our new model maintains anisotropy in the resolution, and

326 the connection between the horizontal and vertical motions in the eddy dissipation. The new

327 anisotropic SGS parameterization is tested in LES of decaying stratified turbulence, and the results

328 are compared with those from DNS: the time series of total energy and kinetic energy dissipation

329 rate, vorticity field, horizontal and vertical wavenumber spectra, and mixing efficiency are fairly

17

well reproduced in the new LES scheme similar to those in DNS, while the computational cost is largely decreased in LES.

It has been shown that if we neglect the horizontal derivatives of vertical motions in our eddy dissipation terms, our results will be under-dissipated at small scales. We think a similar story should exist in atmosphere and ocean models (e.g. Skamarock et al. 2008; Griffies & Hallberg 2000), where the vertical derivatives of SGS fluxes are neglected by horizontal mixing schemes. As a result, we hypothesize that the horizontal eddy dissipation parameters may sometimes be artificially increased in atmosphere and ocean models to ensure model convergence since the zonal and meridional SGS eddy fluxes do not include fluxes from vertical motions. This unrealistically enhanced horizontal eddy dissipation can affect the results of atmosphere and ocean models, and may be compensated by adding an energizing term in the form of a stochastic or negative Laplacian backscatter (as in Mana & Zanna 2014; Jansen & Held 2014) to the equations of motion, in order to improve the performance of these models. Our work suggests that if we keep the neglected terms in the horizontal eddy dissipation scheme, the model performance may be improved without adding any additional energizing terms to the zonal and meridional momentum equations. Nevertheless, this suggestion has to be tested in large-scale atmosphere and ocean models since the flow regime in such models, even at the grid scale, is affected by rotating, unlike the stratified turbulence considered here. In addition, as model resolutions continue to increase, grid-scale motions in such models will become closer to the stratified turbulence regime, and our findings will become increasingly relevant.

In atmosphere and ocean models, different types of SGS eddy viscosity and diffusivity parameterizations can be used. In addition to the Smagorinsky model, a common SGS model in atmosphere and ocean simulations is the turbulent kinetic energy (TKE) model, where the term $\bar{S}$ in equations (23-26) is replaced by $K_{tke} = c_k \mathscr{L}_m \bar{e}^{1/2}$. Here, c_k is a constant coefficient, $\mathscr{L}_m$ is a

18

mixing length that is usually proportional to the grid spacing, and $\bar{e}$ is the SGS turbulence kinetic energy that is computed by solving a transport equation for the TKE budget (see e.g. Sommeria 1976; Deardorff 1980; Schumann 1991; Kaltenbach et al. 1994). Our anisotropic horizontal-filtering framework can be easily adapted to such models. Furthermore, our approach can also be used in more complicated atmosphere and ocean models where, for example, water vapor or other scalars are considered. For these cases, we would need to perform scale analysis for the SGS fluxes, which are proportional to the gradient of corresponding quantities, and include the vertical derivatives of quantities in the horizontally-filtered transport equations. Meanwhile, we should evaluate the performance of our anisotropic SGS parameterization in atmosphere and ocean simulations, which have much lower vertical resolution than what we consider here, with a hierarchy of vertical subgrid schemes. We will consider simulations with and without current vertical SGS parameterizations in atmosphere and ocean models. Also, depending on the ratio of Δ_h/Δ_z, we can develop a new SGS scheme based on a vertical-filtering approach. For future work, we plan to test the performance of our new anisotropic SGS parameterization in atmosphere and ocean models considering these modifications and complexities.

19

<div align="center">

378 APPENDIX A

379 A nonlinear approximation for the SGS stress τ_{ij}

</div>

380 Using the Taylor series expansion of the velocity field $\mathbf{u}$, the nonlinear tensor $u_i(\mathbf{x}+\mathbf{r})u_j(\mathbf{x}+\mathbf{r})$

381 can also be expanded, keeping up to cubic terms in $\mathbf{r}$, as[3]

$$
\begin{aligned}
u_i(\mathbf{x}+\mathbf{r})u_j(\mathbf{x}+\mathbf{r}) = {} & u_i(\mathbf{x})u_j(\mathbf{x}) + \frac{\partial[u_i(\mathbf{x})u_j(\mathbf{x})]}{\partial x_k}r_k + \frac{\partial u_i(\mathbf{x})}{\partial x_k}\frac{\partial u_j(\mathbf{x})}{\partial x_l}r_k r_l \\
& + \frac{1}{2}u_j(\mathbf{x})\frac{\partial^2 u_i(\mathbf{x})}{\partial x_k \partial x_l}r_k r_l + \frac{1}{2}u_i(\mathbf{x})\frac{\partial^2 u_j(\mathbf{x})}{\partial x_k \partial x_l}r_k r_l \\
& + \frac{1}{2}\frac{\partial u_i(\mathbf{x})}{\partial x_m}\frac{\partial^2 u_j(\mathbf{x})}{\partial x_k \partial x_l}r_m r_k r_l + \frac{1}{2}\frac{\partial u_j(\mathbf{x})}{\partial x_m}\frac{\partial^2 u_i(\mathbf{x})}{\partial x_k \partial x_l}r_m r_k r_l \\
& + \frac{1}{6}u_i(\mathbf{x})\frac{\partial^3 u_j(\mathbf{x})}{\partial x_m \partial x_k \partial x_l}r_m r_k r_l + \frac{1}{6}u_j(\mathbf{x})\frac{\partial^3 u_i(\mathbf{x})}{\partial x_m \partial x_k \partial x_l}r_m r_k r_l + \dots
\end{aligned}
\tag{A1}
$$

382 We can apply the filter function $G(\mathbf{r})$ to equations (10) and (A1), respectively, and integrate over

383 the domain D in order to find the filtered variables. For velocities $u_i(\mathbf{x})$ and $u_j(\mathbf{x})$, keeping up to

384 the cubic terms in the Taylor series, we have

$$
\begin{aligned}
\bar{u}_i(\mathbf{x}) \approx {} & u_i(\mathbf{x}) + \frac{\partial u_i(\mathbf{x})}{\partial x_k}\int_D r_k G(\mathbf{r})\,d\mathbf{r} + \frac{1}{2}\frac{\partial^2 u_i(\mathbf{x})}{\partial x_k \partial x_l}\int_D r_k r_l G(\mathbf{r})\,d\mathbf{r} \\
& + \frac{1}{6}\frac{\partial^3 u_i(\mathbf{x})}{\partial x_m \partial x_k \partial x_l}\int_D r_k r_l r_m G(\mathbf{r})\,d\mathbf{r},
\end{aligned}
\tag{A2}
$$

385

$$
\begin{aligned}
\bar{u}_j(\mathbf{x}) \approx {} & u_j(\mathbf{x}) + \frac{\partial u_j(\mathbf{x})}{\partial x_k}\int_D r_k G(\mathbf{r})\,d\mathbf{r} + \frac{1}{2}\frac{\partial^2 u_j(\mathbf{x})}{\partial x_k \partial x_l}\int_D r_k r_l G(\mathbf{r})\,d\mathbf{r} \\
& + \frac{1}{6}\frac{\partial^3 u_j(\mathbf{x})}{\partial x_m \partial x_k \partial x_l}\int_D r_k r_l r_m G(\mathbf{r})\,d\mathbf{r}.
\end{aligned}
\tag{A3}
$$

---

[3]For simplicity, partial derivatives with respect to the variable $r_k|_{r_k=0}$ are written as partial derivatives with respect to x_k because term $\mathbf{x}+\mathbf{r}$
reduces to $\mathbf{x}$ when $\mathbf{r}=0$ (as in e.g. Pope 2000).

Similarly for $u_i(\mathbf{x})\,u_j(\mathbf{x})$, keeping up to cubic terms, we have

$$
\begin{aligned}
\overline{u_i(\mathbf{x})\,u_j(\mathbf{x})} \approx\; & u_i(\mathbf{x})\,u_j(\mathbf{x}) + \frac{\partial[u_i(\mathbf{x})\,u_j(\mathbf{x})]}{\partial x_k}\int_D r_k G(\mathbf{r})\,d\mathbf{r} \\
& + \frac{\partial u_i(\mathbf{x})}{\partial x_l}\frac{\partial u_j(\mathbf{x})}{\partial x_k}\int_D r_k r_l G(\mathbf{r})\,d\mathbf{r} \\
& + \frac{1}{2}u_i(\mathbf{x})\frac{\partial^2 u_j(\mathbf{x})}{\partial x_k \partial x_l}\int_D r_k r_l G(\mathbf{r})\,d\mathbf{r} + \frac{1}{2}u_j(\mathbf{x})\frac{\partial^2 u_i(\mathbf{x})}{\partial x_k \partial x_l}\int_D r_k r_l G(\mathbf{r})\,d\mathbf{r} \\
& + \frac{1}{2}\frac{\partial u_i(\mathbf{x})}{\partial x_m}\frac{\partial^2 u_j(\mathbf{x})}{\partial x_k \partial x_l}\int_D r_k r_l r_m G(\mathbf{r})\,d\mathbf{r} + \frac{1}{2}\frac{\partial u_j(\mathbf{x})}{\partial x_m}\frac{\partial^2 u_i(\mathbf{x})}{\partial x_k \partial x_l}\int_D r_k r_l r_m G(\mathbf{r})\,d\mathbf{r} \\
& + \frac{1}{6}u_i(\mathbf{x})\frac{\partial^3 u_j(\mathbf{x})}{\partial x_m \partial x_k \partial x_l}\int_D r_k r_l r_m G(\mathbf{r})\,d\mathbf{r} + \frac{1}{6}u_j(\mathbf{x})\frac{\partial^3 u_i(\mathbf{x})}{\partial x_m \partial x_k \partial x_l}\int_D r_k r_l r_m G(\mathbf{r})\,d\mathbf{r}.
\end{aligned}
\tag{A4}
$$

Using equations (A2) and (A3), we can also approximate the nonlinear filtered product $\bar{u}_i(\mathbf{x})\,\bar{u}_j(\mathbf{x})$, keeping up to cubic terms, as follows

$$
\begin{aligned}
\bar{u}_i(\mathbf{x})\,\bar{u}_j(\mathbf{x}) \approx\; & u_i(\mathbf{x})\,u_j(\mathbf{x}) + \frac{\partial[u_i(\mathbf{x})\,u_j(\mathbf{x})]}{\partial x_k}\int_D r_k G(\mathbf{r})\,d\mathbf{r} \\
& + \left(\frac{\partial u_i(\mathbf{x})}{\partial x_l}\int_D r_l G(\mathbf{r})\,d\mathbf{r}\right)\left(\frac{\partial u_j(\mathbf{x})}{\partial x_k}\int_D r_k G(\mathbf{r})\,d\mathbf{r}\right) \\
& + \frac{1}{2}u_i(\mathbf{x})\frac{\partial^2 u_j(\mathbf{x})}{\partial x_k \partial x_l}\int_D r_k r_l G(\mathbf{r})\,d\mathbf{r} + \frac{1}{2}u_j(\mathbf{x})\frac{\partial^2 u_i(\mathbf{x})}{\partial x_k \partial x_l}\int_D r_k r_l G(\mathbf{r})\,d\mathbf{r} \\
& + \frac{1}{2}\left(\frac{\partial u_i(\mathbf{x})}{\partial x_m}\int_D r_m G(\mathbf{r})\,d\mathbf{r}\right)\left(\frac{\partial^2 u_j(\mathbf{x})}{\partial x_k \partial x_l}\int_D r_k r_l G(\mathbf{r})\,d\mathbf{r}\right) \\
& + \frac{1}{2}\left(\frac{\partial u_j(\mathbf{x})}{\partial x_m}\int_D r_m G(\mathbf{r})\,d\mathbf{r}\right)\left(\frac{\partial^2 u_i(\mathbf{x})}{\partial x_k \partial x_l}\int_D r_k r_l G(\mathbf{r})\,d\mathbf{r}\right) \\
& + \frac{1}{6}u_i(\mathbf{x})\frac{\partial^3 u_j(\mathbf{x})}{\partial x_m \partial x_k \partial x_l}\int_D r_k r_l r_m G(\mathbf{r})\,d\mathbf{r} + \frac{1}{6}u_j(\mathbf{x})\frac{\partial^3 u_i(\mathbf{x})}{\partial x_m \partial x_k \partial x_l}\int_D r_k r_l r_m G(\mathbf{r})\,d\mathbf{r}.
\end{aligned}
\tag{A5}
$$

Subtracting equation (A5) from equation (A4) results in a mathematical formulation for the approximate SGS stress $\tau_{ij}(\mathbf{x}) = \overline{u_i(\mathbf{x})\,u_j(\mathbf{x})} - \bar{u}_i(\mathbf{x})\,\bar{u}_j(\mathbf{x})$, written as

$$
\tau_{ij}(\mathbf{x}) = \frac{\partial u_i(\mathbf{x})}{\partial x_l}\frac{\partial u_j(\mathbf{x})}{\partial x_k}\int_D r_k r_l G(\mathbf{r})\,d\mathbf{r} + \mathcal{O}(\Delta^4),
\tag{A6}
$$

where we have assumed that the odd moments of the filter function $G(\mathbf{r})$ are zero, and $\mathcal{O}(\Delta^4)$ shows the order of leading error in this approximation, since the fourth order moments of $G(\mathbf{r})$ are proportional to Δ^4.

21

395   **A nonlinear approximation for the SGS flux h_j**

396   We can expand the potential temperature field θ using the Taylor series at a given point $\mathbf{x}$ in $\mathbf{r}$,

397   which is of the order filter width Δ (again we keep up to cubic terms)

$$\theta\left(\mathbf{x}+\mathbf{r}\right) = \theta\left(\mathbf{x}\right) + \frac{\partial\theta\left(\mathbf{x}\right)}{\partial x_k}r_k + \frac{1}{2}\frac{\partial^2\theta\left(\mathbf{x}\right)}{\partial x_k\partial x_l}r_kr_l + \frac{1}{6}\frac{\partial^3\theta\left(\mathbf{x}\right)}{\partial x_m\partial x_k\partial x_l}r_mr_kr_l + \dots \tag{B1}$$

398   If we employ the filter function $G\left(\mathbf{r}\right)$ to the equation (B1), and keep up to cubic terms, we obtain

$$\begin{aligned}\bar{\theta}\left(\mathbf{x}\right) &\approx \theta\left(\mathbf{x}\right) + \frac{\partial\theta\left(\mathbf{x}\right)}{\partial x_k}\int_D r_kG\left(\mathbf{r}\right)d\mathbf{r} + \frac{1}{2}\frac{\partial^2\theta\left(\mathbf{x}\right)}{\partial x_k\partial x_l}\int_D r_kr_lG\left(\mathbf{r}\right)d\mathbf{r} \\ &+ \frac{1}{6}\frac{\partial^3\theta\left(\mathbf{x}\right)}{\partial x_m\partial x_k\partial x_l}\int_D r_mr_kr_lG\left(\mathbf{r}\right)d\mathbf{r}.\end{aligned} \tag{B2}$$

399   Similarly, we can expand the SGS potential temperature flux $\overline{\theta\left(\mathbf{x}\right)u_j\left(\mathbf{x}\right)}$ and resolved potential

400   temperature flux $\bar{\theta}\left(\mathbf{x}\right)\bar{u}_j\left(\mathbf{x}\right)$, keeping up to cubic terms, in the following forms

$$\begin{aligned}\overline{\theta\left(\mathbf{x}\right)u_j\left(\mathbf{x}\right)} &\approx \theta\left(\mathbf{x}\right)u_j\left(\mathbf{x}\right) + \frac{\partial[\theta\left(\mathbf{x}\right)u_j\left(\mathbf{x}\right)]}{\partial x_k}\int_D r_kG\left(\mathbf{r}\right)d\mathbf{r} \\ &+ \frac{\partial\theta\left(\mathbf{x}\right)}{\partial x_l}\frac{\partial u_j\left(\mathbf{x}\right)}{\partial x_k}\int_D r_kr_lG\left(\mathbf{r}\right)d\mathbf{r} \\ &+ \frac{1}{2}\theta\left(\mathbf{x}\right)\frac{\partial^2u_j\left(\mathbf{x}\right)}{\partial x_k\partial x_l}\int_D r_kr_lG\left(\mathbf{r}\right)d\mathbf{r} + \frac{1}{2}u_j\left(\mathbf{x}\right)\frac{\partial^2\theta\left(\mathbf{x}\right)}{\partial x_k\partial x_l}\int_D r_kr_lG\left(\mathbf{r}\right)d\mathbf{r} \\ &+ \frac{1}{2}\frac{\partial\theta\left(\mathbf{x}\right)}{\partial x_m}\frac{\partial^2u_j\left(\mathbf{x}\right)}{\partial x_k\partial x_l}\int_D r_kr_lr_mG\left(\mathbf{r}\right)d\mathbf{r} + \frac{1}{2}\frac{\partial u_j\left(\mathbf{x}\right)}{\partial x_m}\frac{\partial^2\theta\left(\mathbf{x}\right)}{\partial x_k\partial x_l}\int_D r_kr_lr_mG\left(\mathbf{r}\right)d\mathbf{r} \\ &+ \frac{1}{6}\theta\left(\mathbf{x}\right)\frac{\partial^3u_j\left(\mathbf{x}\right)}{\partial x_m\partial x_k\partial x_l}\int_D r_kr_lr_mG\left(\mathbf{r}\right)d\mathbf{r} + \frac{1}{6}u_j\left(\mathbf{x}\right)\frac{\partial^3\theta\left(\mathbf{x}\right)}{\partial x_m\partial x_k\partial x_l}\int_D r_kr_lr_mG\left(\mathbf{r}\right)d\mathbf{r},\end{aligned} \tag{B3}$$

and

$$
\begin{aligned}
\bar{\theta}\left(\mathbf{x}\right)\bar{u}_j\left(\mathbf{x}\right) \approx{} & \theta\left(\mathbf{x}\right)u_j\left(\mathbf{x}\right) + \frac{\partial\left[\theta\left(\mathbf{x}\right)u_j\left(\mathbf{x}\right)\right]}{\partial x_k}\int_D r_k G\left(\mathbf{r}\right)d\mathbf{r} \\
& + \left(\frac{\partial\theta\left(\mathbf{x}\right)}{\partial x_l}\int_D r_l G\left(\mathbf{r}\right)d\mathbf{r}\right)\left(\frac{\partial u_j\left(\mathbf{x}\right)}{\partial x_k}\int_D r_k G\left(\mathbf{r}\right)d\mathbf{r}\right) \\
& + \frac{1}{2}\theta\left(\mathbf{x}\right)\frac{\partial^2 u_j\left(\mathbf{x}\right)}{\partial x_k \partial x_l}\int_D r_k r_l G\left(\mathbf{r}\right)d\mathbf{r} + \frac{1}{2}u_j\left(\mathbf{x}\right)\frac{\partial^2\theta\left(\mathbf{x}\right)}{\partial x_k \partial x_l}\int_D r_k r_l G\left(\mathbf{r}\right)d\mathbf{r} \\
& + \frac{1}{2}\left(\frac{\partial\theta\left(\mathbf{x}\right)}{\partial x_m}\int_D r_m G\left(\mathbf{r}\right)d\mathbf{r}\right)\left(\frac{\partial^2 u_j\left(\mathbf{x}\right)}{\partial x_k \partial x_l}\int_D r_k r_l G\left(\mathbf{r}\right)d\mathbf{r}\right) \\
& + \frac{1}{2}\left(\frac{\partial u_j\left(\mathbf{x}\right)}{\partial x_m}\int_D r_m G\left(\mathbf{r}\right)d\mathbf{r}\right)\left(\frac{\partial^2\theta\left(\mathbf{x}\right)}{\partial x_k \partial x_l}\int_D r_k r_l G\left(\mathbf{r}\right)d\mathbf{r}\right) \\
& + \frac{1}{6}\theta(\mathbf{x})\frac{\partial^3 u_j(\mathbf{x})}{\partial x_m \partial x_k \partial x_l}\int_D r_k r_l r_m G\left(\mathbf{r}\right)d\mathbf{r} + \frac{1}{6}u_j(\mathbf{x})\frac{\partial^3\theta(\mathbf{x})}{\partial x_m \partial x_k \partial x_l}\int_D r_k r_l r_m G\left(\mathbf{r}\right)d\mathbf{r}.
\end{aligned}
\tag{B4}
$$

Therefore, the SGS potential temperature flux term $h_j\left(\mathbf{x}\right) = \overline{\theta\left(\mathbf{x}\right)u_j\left(\mathbf{x}\right)} - \bar{\theta}\left(\mathbf{x}\right)\bar{u}_j\left(\mathbf{x}\right)$ is

$$
h_j\left(\mathbf{x}\right) = \frac{\partial\theta\left(\mathbf{x}\right)}{\partial x_l}\frac{\partial u_j\left(\mathbf{x}\right)}{\partial x_k}\int_D r_k r_l G\left(\mathbf{r}\right)d\mathbf{r} + \mathcal{O}(\Delta^4),
\tag{B5}
$$

where we have again assumed that the odd moments of the filter function $G\left(\mathbf{r}\right)$ are zero.

References

Augier P. & Lindborg E. 2013 A New Formulation of the Spectral Energy Budget of the Atmosphere, with Application to Two High-Resolution General Circulation Models. *J. Atmos. Sci.* **70**, 2293-2308.

Bartello P. & Tobias M. 2013 Sensitivity of stratified turbulence to the buoyancy Reynolds number. *J. Fluid Mech.* **725**, 1-22.

Becker E. & Burkhardt U. 2007 Nonlinear horizontal diffusion for GCMs. *Mon. Weather Rev.* **135**, 1439-1454.

Billant P. & Chomaz J.-M. 2001 Self-similarity of strongly stratified inviscid flows. *Phys. Fluids* **13**(6), 1645-1651.

23

Brethouwer G., Billant P., Lindborg E. & Chomaz J.-M. 2007 Scaling analysis and simulation of strongly stratified turbulent flows. *J. Fluid Mech.* **585**, 343-368.

Brune S. & Becker E. 2013 Indications of Stratified Turbulence in a Mechanistic GCM. *J. Atmos. Sci.* **70**, 231-247.

Caulfield C. P. & Peltier W. R. 2000 The anatomy of the mixing transition in homogeneous and stratified free shear layers. *J. Fluid Mech.* **413**, 1-47.

Deardorff J. W. 1980 Stratocumulus-capped mixed layers derived from a three-dimensional model. *Boundary-Layer Meteorol.* **18**, 495-527.

Durran D. R. 2010 Numerical Methods for Fluid Dynamics with Application to Geophysics. *Springer.*

Gordon C. T. & Stern W. F. 1982 A description of the GFDL global spectral model. *Mon. Weather Rev.* **110**, 625-644.

Gregg M. C., D'Asaro E. A., Riley J. J. & Kunze E. 2018 Mixing efficiency in the Ocean. *Annu. Rev. Mar. Sci.* **10**, 443-473.

Griffies S. M. & Hallberg R. 2000 Biharmonic friction with a Smagorinsky-like viscosity for use in large-scale eddy-permitting ocean models. *Mon. Weather Rev.* **128**, 2935-2946.

Griffies S. M., Harrison M. J., Pacanowski R. C. & Rosati A. 2004 A technical guide to MOM4, GFDL ocean group technical report No.5. *NOAA/Geophysical Fluid Dynamics Laboratory* (www.gfdl.noaa.gov), 60-66.

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

24

Kaltenbach H.-J., Gerz T. & Schumann U. 1994 Large-eddy simulation of homogeneous turbulence and diffusion in stably stratified shear flow. *J. Fluid Mech.* **280**, 1-40.

Kang H. S., Chester S. & Meneveau C. 2003 Decaying turbulence in an active-grid-generated flow and comparisons with large-eddy simulation. *J. Fluid Mech.* **480**, 129-160.

Khani S. & Waite M. L. 2013 Effective eddy viscosity in stratified turbulence. *J. Turbul.* **14**(7), 49-70.

Khani S. & Waite M. L. 2014 Buoyancy scale effects in large-eddy simulations of stratified turbulence. *J. Fluid Mech.* **754**, 75-97.

Khani S. & Waite M. L. 2015 Large eddy simulations of stratified turbulence: the dynamic Smagorinsky model. *J. Fluid Mech.* **773**, 327-344.

Khani S. & Waite M. L. 2016 Backscatter in stratified turbulence. *Eur. J. Mech. B/Fluids* **60**, 1-12.

Khani S. & Porté-Agel F. 2017a Evaluation of non-eddy viscosity subgrid-scale models in stratified turbulence using direct numerical simulations. *Eur. J. Mech. B/Fluids* **65**, 168-178.

Khani S. & Porté-Agel F. 2017b A modulated-gradient parametrization for the large-eddy simulation of the atmospheric boundary layer using the Weather Research and Forecasting model. *Boundary-Layer Meteorol.* **165**(3), 385-404.

Khani S. 2018 Mixing efficiency in large-eddy simulations of stratified turbulence. *J. Fluid Mech.* **849**, 373-394.

Lindborg E. 2006 The energy cascade in a strongly stratified fluid. *J. Fluid Mech.* **550**, 207-242.

Mana P. P. & Zanna L. 2014 Toward a stochastic parameterization of ocean mesoscale eddies. *Ocean Modelling* **79**, 1-20.

25

Meneveau C., & Katz J. 2000 Scale-invariance and turbulence models for large-eddy simulation. *Ann. Rev. Fluid Mech.* **32**, 1-32.

Moin P. & Mahesh K. 1998 Direct numerical simulation: a tool in turbulence research. *Rev. Fluid Mech.* **30**, 539-578.

Orszag S. A. 1971 On the elimination of aliasing in finite-difference schemes by filtering high-wavenumber components. *J. Atmos. Sci.* **28**, 1074-1074.

Paoli R., Thouron O., Escobar J., Picot J. & Cariolle D. 2014 High-resolution large-eddy simulations of stably stratified flows: application to subkilometer-scale turbulence in the upper troposphere-lower stratosphere . *Atmos. Chem. Phys.* **14**, 5037-5055.

Pope S. B. 2000 Turbulent Flows. *Cambridge University Press*, Cambridge.

Riley J. J. & Lelong M.-P. 2000 Fluid motions in the presence of strong stable stratification. *Annu. Rev. Fluid Mech.* **32**, 613-657.

Riley J. J. & Lindborg E. 2008 Stratified turbulence: a possible interpretation of some geophysical turbulence measurements. *J. Atmos. Sci.* **65**, 2416-2424.

Schaefer-Rolffs U. & Becker E. 2018 Scale-invariant formulation of momentum diffusion for high-resolution atmospheric circulation models. *Mon. Weather Rev.* **146**, 1045-1062.

Schumann U. 1991 Subgrid length-scale for large-eddy simulation of stratified turbulence. *Theoret. Comput. Fluid Dynamics* **2**, 279-290.

Siegel D. A. & Domaradzki J. A. 1994 Large-eddy simulation of stably stratified turbulence. *J. Phys. Oceanogr.* **24**, 2353-2386.

26

476 Skamarock W. C., Snyder C., Klemp J. B. & Park S.-H. 2019 Vertical resolution requirements in

477 atmospheric simulation. *Mon. Weather Rev.* **147**, 2641-2656.

478 Skamarock W. C. & coauthors 2008 A description of the advanced research WRF version 3. *Tech.*

479 *Rep. NCAR/TN-475+STR*, National Center for Atmospheric Research, Boulder, CO, USA.

480 Sommeria G. 1976 Three-dimensional simulation of turbulence processes in an undisturbed trade

481 wind boundary layer. *J. Atmos. Sci.* **33**, 216-241.

482 Smagorinsky J. 1993 Some historical remarks on the use of nonlinear viscosities. *Large Eddy*

483 *Simulation of Complex Engineering and Geophysical Flows*, B. Galperin and St. A. Orszag,

484 Ed., *Cambridge University Press*, 336.

485 Waite M. L. 2011 Stratified turbulence at the buoyancy scale. *Phys. Fluids* **23**, 066602.

486 doi:10.1063/1.3599699.

487 Waite M. L. 2016 Dependence of model energy spectra on vertical resolution. *Mon. Weather Rev.*

488 **144**, 1407-1421.

489 Waite M. L. & Bartello P. 2004 Stratified turbulence dominated by vortical motion. *J. Fluid Mech.*

490 **517**, 281-303.

491 Winters K. B. & D'Asaro E. A. 1996 Diascalar flux and the rate of fluid mixing. *J. Fluid Mech.*

492 **317**, 179-193.

LIST OF TABLES

DNS	N	$n_{x,y}$	n_z	$L_{x,y}$	L_z	$\langle\varepsilon\rangle$	$\langle\varepsilon_p\rangle$	$\langle E(t)\rangle$	Re_b	Fr_h	k_b	k_o
D18N2	2.1	960	960	2π	2π	7.42×10^{-3}	4.22×10^{-3}	0.074	30.3	0.048	7.7	35.3
D22N4	4.2	960	480	2π	π	4.93×10^{-3}	2.67×10^{-3}	0.076	6.2	0.015	15.2	122.6

LES	N	$n_{x,y}$	n_z	$L_{x,y}$	L_z	$\langle\varepsilon\rangle$	$\langle\varepsilon_p\rangle$	$\langle E(t)\rangle$	k_c	Fr_h	k_b	k_o
LA18N2a	2.1	480	960	2π	2π	8.34×10^{-3}	4.29×10^{-3}	0.074	158	0.054	7.7	33.3
LA18N2b	2.1	240	960	2π	2π	6.94×10^{-3}	3.77×10^{-3}	0.082	78	0.040	7.4	36.5
LA22N4a	4.2	480	480	2π	π	5.75×10^{-3}	3.23×10^{-3}	0.079	158	0.017	15.0	113.5
LA22N4b	4.2	240	480	2π	π	5.21×10^{-3}	3.15×10^{-3}	0.082	78	0.015	14.7	119.2
LAV18N2b	2.1	240	480	2π	2π	6.69×10^{-3}	3.75×10^{-3}	0.081	78	0.036	7.1	37.2
LAV18N2c	2.1	120	480	2π	2π	7.10×10^{-3}	3.86×10^{-3}	0.078	38	0.043	7.5	36.1
LAV18N2d	2.1	160	480	2π	2π	6.26×10^{-3}	3.41×10^{-3}	0.083	52	0.036	7.3	38.5

TABLE 1. List of numerical simulations with DNS and LES.
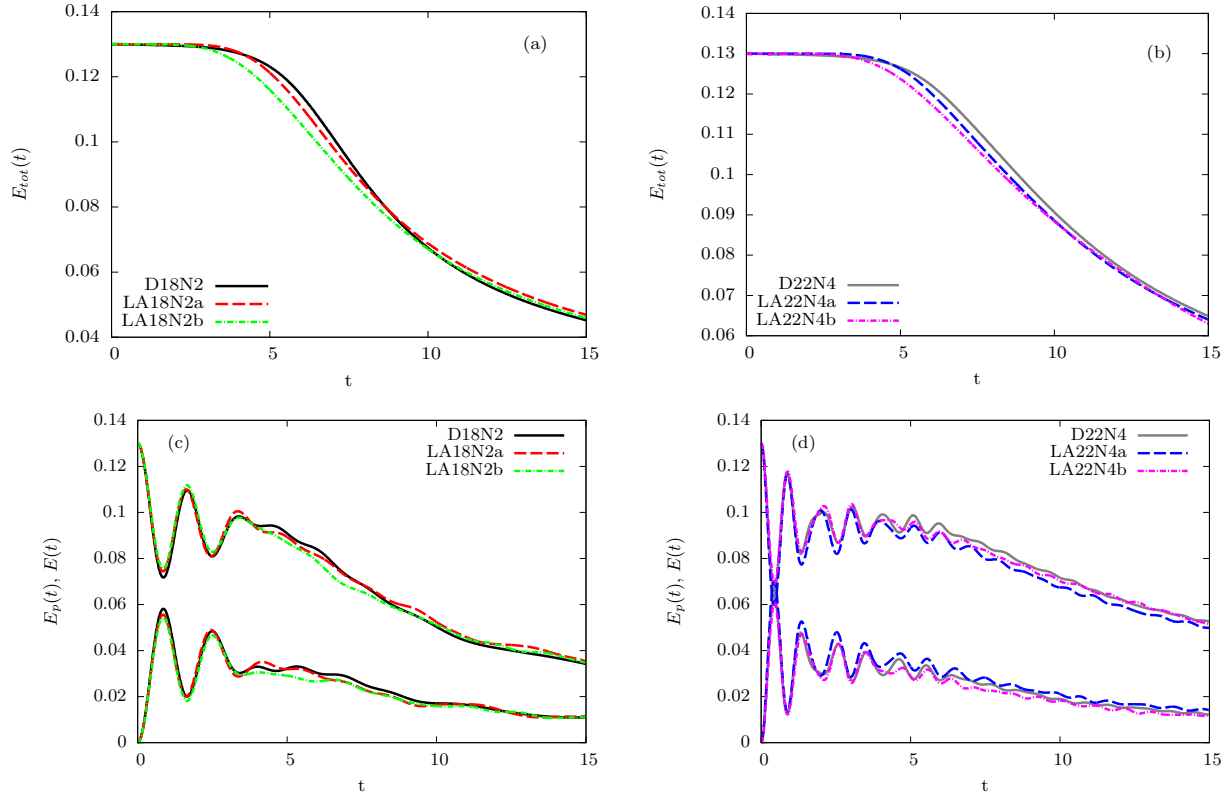
LIST OF FIGURES

FIG. 1. Time series of total energy (top), and the kinetic and potential energy (bottom) for DNS and anisotropic LES runs with weak (a,c) and strong (b,d) stratification. Potential energy curves in panels (c,d) are those start from zero.
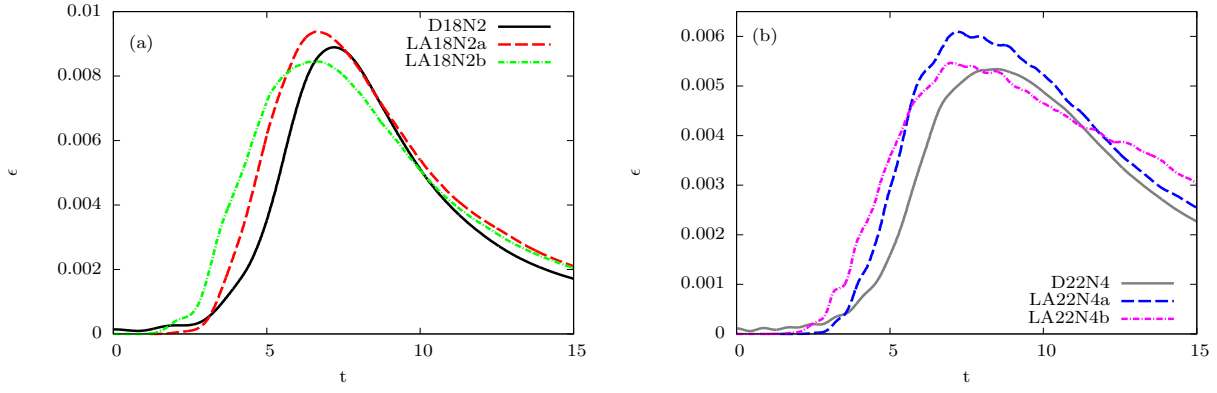
FIG. 2. Time series of the kinetic energy dissipation rate for DNS and anisotropic LES runs with weak (a) and strong (b) stratification.
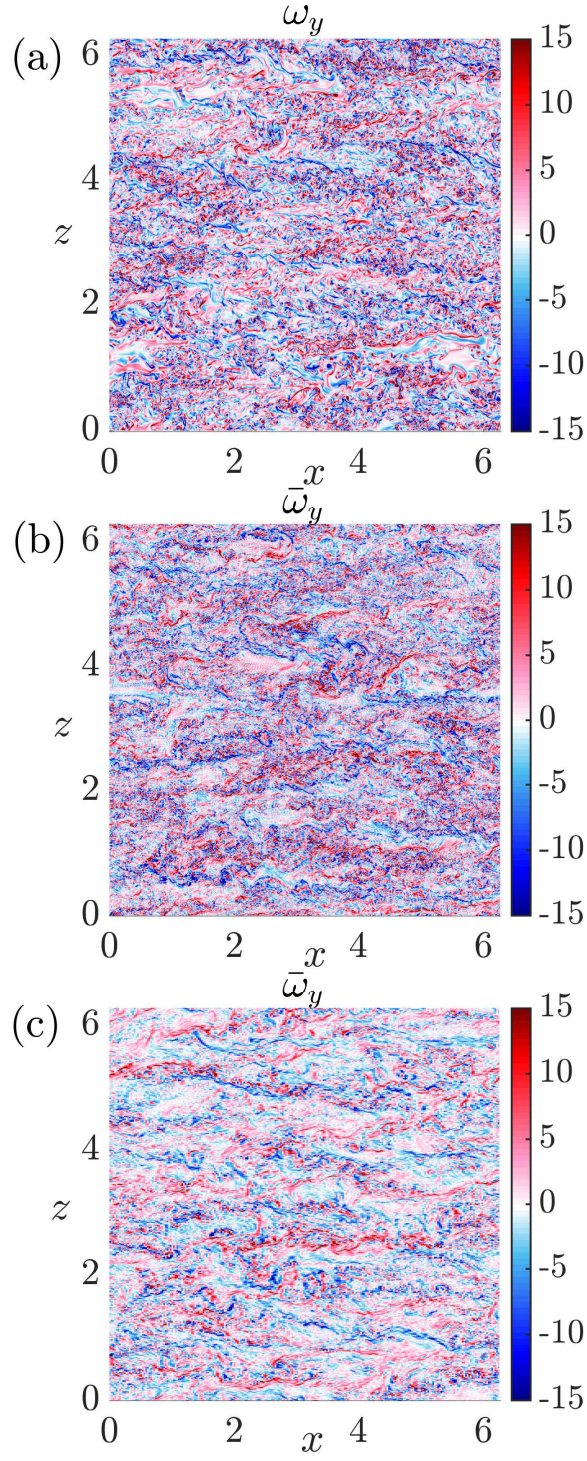
FIG. 3. Vorticity field in y direction on the x-z plane at $y = 0.25$ and $t = 15$ for the case with $Re_\ell = 18000$ and $N = 2.1$: (a) DNS, (b) high horizontal-resolution and (c) low horizontal-resolution anisotropic LES.
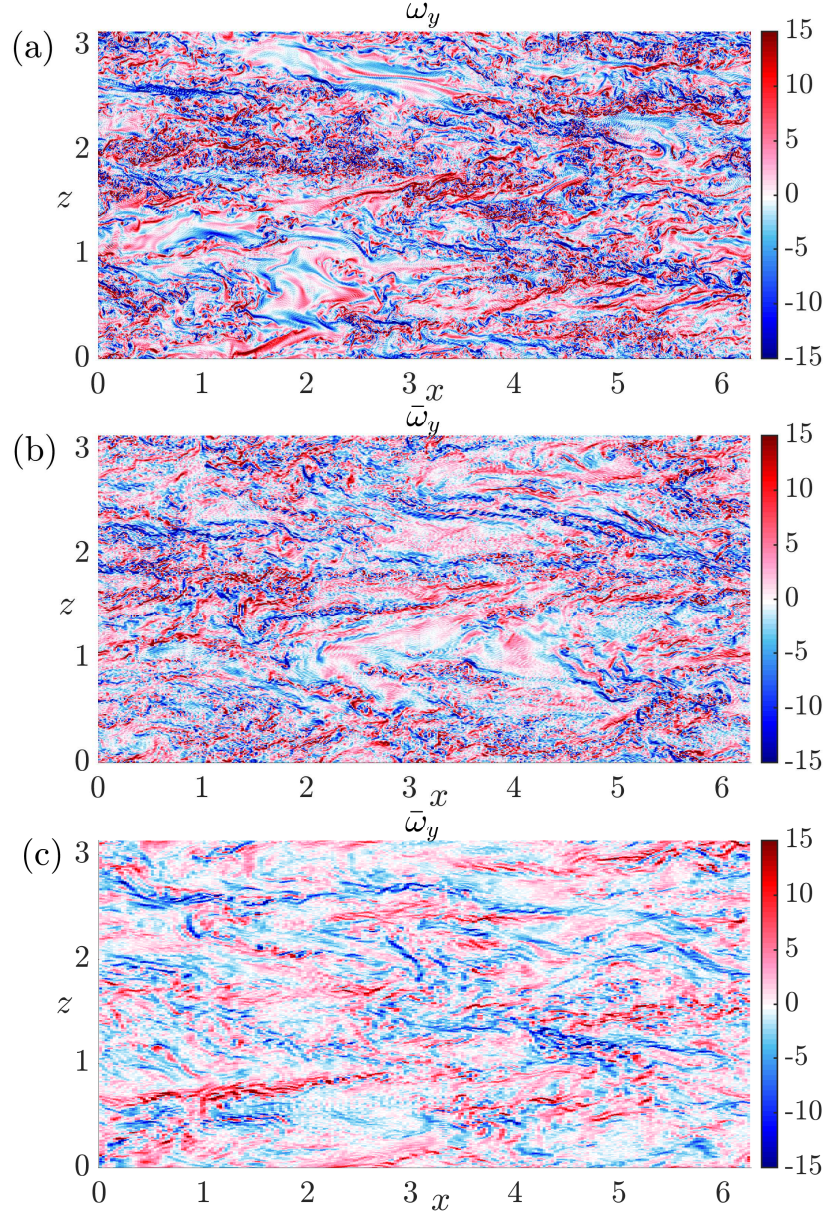
FIG. 4. Vorticity field in y direction on the x-z plane at $y = 0.25$ and $t = 15$ for the case with $Re_\ell = 22200$ and $N = 4.2$: (a) DNS, (b) high horizontal-resolution anisotropic LES, and (c) low horizontal-resolution anisotropic LES.
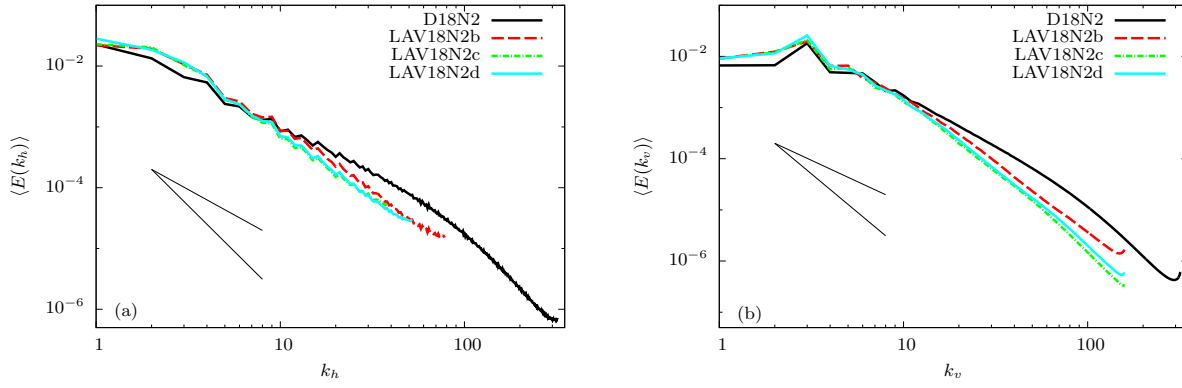
35
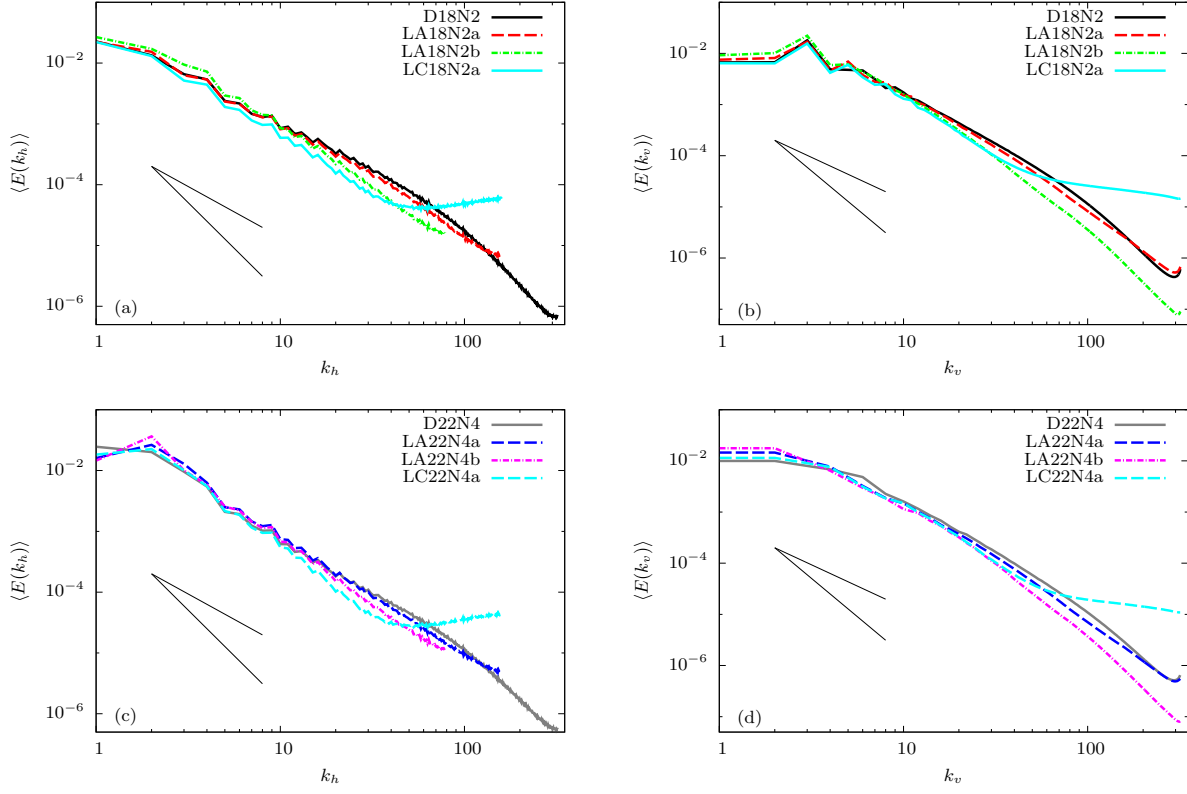
FIG. 6. The time-averaged horizontal (left) and vertical (right) wavenumber kinetic energy spectra for DNS and vertically-reduced resolution anisotropic LES for the weak stratification case. Time averaging is performed over a window ($\Delta t = 4$) around the maximum kinetic energy dissipation rate. The solid black line segments show $-5/3$ and -3 slopes.
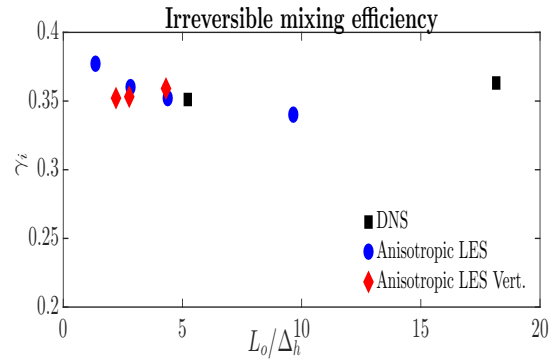
36

FIG. 7. Irreversible mixing efficiency γ_i versus the ratio L_o/Δ_h for DNS and anisotropic LES runs. LES and LES vert. refer to those anisotropic LES runs with high and low vertical resolutions (LA and LAV), respectively.