J. Fluid Mech. (2016), *vol.* 788, *pp.* 5–41.
doi:10.1017/jfm.2015.700

The role of mixed-layer instabilities in submesoscale turbulence

Jörn Callies[1,†], Glenn Flierl[1], Raffaele Ferrari[1] and Baylor Fox-Kemper[2]

[1]Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

[2]Department of Earth, Environmental and Planetary Sciences, Brown University, 324 Brook Street, Providence, RI 02912, USA

(Received 23 April 2015; revised 23 September 2015; accepted 24 November 2015; first published online 22 December 2015)

Upper-ocean turbulence at scales smaller than the mesoscale is believed to exchange surface and thermocline waters, which plays an important role in both physical and biogeochemical budgets. But what energizes this submesoscale turbulence remains a topic of debate. Two mechanisms have been proposed: mesoscale-driven surface frontogenesis and baroclinic mixed-layer instabilities. The goal here is to understand the differences between the dynamics of these two mechanisms, using a simple quasi-geostrophic model. The essence of mesoscale-driven surface frontogenesis is captured by the well-known surface quasi-geostrophic model, which describes the sharpening of surface buoyancy gradients and the subsequent breakup in secondary roll-up instabilities. We formulate a similarly archetypical Eady-like model of submesoscale turbulence induced by mixed-layer instabilities. The model captures the scale and structure of this baroclinic instability in the mixed layer. A wide range of scales are energized through a turbulent inverse cascade of kinetic energy that is fuelled by the submesoscale mixed-layer instability. Major differences to mesoscale-driven surface frontogenesis are that mixed-layer instabilities energize the entire depth of the mixed layer and produce larger vertical velocities. The distribution of energy across scales and in the vertical produced by our simple model of mixed-layer instabilities compares favourably to observations of energetic wintertime submesoscale flows, suggesting that it captures the leading-order balanced dynamics of these flows. The dynamics described here in an oceanographic context have potential applications to other geophysical fluids with layers of different stratifications.

Key words: geophysical and geological flows, ocean circulation, quasi-geostrophic flows

1. Introduction

The upper ocean is host to energetic flows at scales smaller than the order 100 km mesoscale eddies. Sharp surface fronts associated with strong along-front currents are generated in high-resolution numerical simulations (e.g. Capet *et al.* 2008*b*; Klein *et al.* 2008) and are observed in the wintertime midlatitude ocean (Shcherbina *et al.* 2013; Callies *et al.* 2015). These submesoscale flows at scales 1–100 km are associated with large vertical fluxes of both physical and biogeochemical tracers that

†Email address for correspondence: joernc@mit.edu

have been argued to regulate the oceanic heat and carbon uptake in global warming scenarios (Capet *et al.* 2008*b*; Klein & Lapeyre 2009; Ferrari 2011; Lévy *et al.* 2012; Mahadevan 2014). Current global ocean models do not resolve submesoscale flows, so these fluxes must be represented by parameterizations that should be based on physical understanding.

Despite the attention received by submesoscale flows in both the theoretical and observational oceanographic communities, the dynamics that generates them remains a topic of debate. Two mechanisms have been proposed: mesoscale-driven surface frontogenesis (Lapeyre & Klein 2006; Roullet *et al.* 2012) and baroclinic mixed-layer instabilities (Boccaletti, Ferrari & Fox-Kemper 2007). It is important to understand the differences between these two mechanisms, because they produce – as we shall see – distinct submesoscale flow characteristics and vertical fluxes.

The essential physics of mesoscale-driven surface frontogenesis can be understood with quasi-geostrophic (QG) dynamics (Stone 1966*a*). A mesoscale strain field sharpens lateral buoyancy gradients at the surface more effectively than in the interior of the ocean. An ageostrophic circulation develops in response to the increasing lateral buoyancy gradient, as described by the omega equation (e.g. Hoskins, Draghici & Davies 1978). In the interior, this circulation acts to weaken the lateral buoyancy gradient: light water downwells on the dense side and dense water upwells on the light side of the gradient. At the surface, however, the vertical velocity must vanish and the ageostrophic circulation cannot counteract the increase in lateral buoyancy gradient – the mesoscale strain field is left unopposed to create strong submesoscale surface fronts.

Mixed-layer instabilities, on the other hand, can energize submesoscale flows by releasing available potential energy stored in large- and mesoscale buoyancy gradients in the surface mixed layer. The weak stratification in deep wintertime mixed layers allows baroclinically unstable modes to rapidly amplify (Haine & Marshall 1998). Much like deep mesoscale modes in the ocean interior (e.g. Gill, Green & Simmons 1974), these mixed-layer modes slide dense water under light water, but their horizontal scale is 1–10 km and they grow on time scales of order 1 day (Boccaletti *et al.* 2007).

The presence of a seasonal cycle in submesoscale turbulence suggests that baroclinic mixed-layer instabilities are an important aspect of upper-ocean dynamics. Both modelling (Mensa *et al.* 2013; Sasaki *et al.* 2014) and observations (Callies *et al.* 2015) show that submesoscale turbulence is energized in winter and suppressed in summer. Mixed-layer instabilities are expected to undergo a strong seasonal cycle, following the seasonal cycle of the mixed-layer depth and the associated mixed-layer potential energy. Mesoscale-driven surface frontogenesis, on the other hand, is not expected to vary seasonally, because mesoscale eddies do not undergo a strong seasonal cycle (cf. Qiu 1999; Qiu & Chen 2004). A full understanding of how submesoscale turbulence is energized by baroclinic mixed-layer instabilities, however, is not as well established as that of mesoscale-driven surface frontogenesis.

The simplest model capturing the essence of mesoscale-driven surface frontogenesis is the surface QG model (Blumen 1978; Held *et al.* 1995). It assumes an infinitely deep ocean with constant stratification and vanishing interior QG potential vorticity (PV),

$$q = \nabla^2 \psi + \frac{\partial}{\partial z}\left(\frac{f^2}{N^2}\frac{\partial \psi}{\partial z}\right) = 0, \tag{1.1}$$

where q is the PV, ψ is the geostrophic streamfunction, f is the (constant) Coriolis frequency and N is the buoyancy frequency. The streamfunction is related to the

horizontal flow by $\boldsymbol{u} = (-\partial\psi/\partial y,\ \partial\psi/\partial x)$. The evolution of the flow is completely determined by the lateral advection of buoyancy at the surface, which represents the no-normal-flow boundary condition (i.e. the vertical advection of the background stratification vanishes),

$$\frac{\partial b}{\partial t} + \mathrm{J}(\psi, b) = 0, \tag{1.2}$$

where $b = f\partial\psi/\partial z$ is buoyancy and

$$\mathrm{J}(\psi, b) = \frac{\partial\psi}{\partial x}\frac{\partial b}{\partial y} - \frac{\partial\psi}{\partial y}\frac{\partial b}{\partial x} \tag{1.3}$$

is the Jacobian operator. The surface buoyancy supplies the boundary condition for the elliptic problem (1.1). Straining by mesoscale eddies creates sharp buoyancy gradients associated with strong flows at the surface. Filamentary instabilities eventually lead to fully-turbulent dynamics (Held *et al.* 1995). Kolmogorov-like dimensional arguments (Kolmogorov 1941) predict that the fully-turbulent surface kinetic and potential energy spectra scale like $K_{k_h} = P_{k_h} \sim {k_h}^{-5/3}$ in a submesoscale inertial range in which surface potential energy b^2/N^2 is cascaded to small scales (Blumen 1978) – k_h is the horizontal wavenumber. This prediction implies that the gradients of velocity and buoyancy (i.e. fronts), whose spectra scale like ${k_h}^2 K_{k_h}$ and ${k_h}^2 P_{k_h}$, are stronger at small submesoscales (large k_h) than at large submesoscales (small k_h). The forward cascade of surface potential energy occurs in conjunction with an inverse cascade of surface kinetic energy fed by the release of potential energy through slumping fronts (Capet *et al.* 2008*a*). The submesoscale energy generated by surface QG turbulence is surface trapped: modes decay exponentially in the vertical, with small-scale modes decaying more rapidly than large-scale modes (e.g. Scott 2006). (Heuristic extensions of surface QG ideas have been developed to infer interior flows from surface properties (Lapeyre & Klein 2006). These extensions are diagnostic in nature and do not attempt to describe the evolution of the flow or make predictions for submesoscale energy levels. We therefore limit our discussion of surface QG dynamics to the case with zero interior PV.)

If non-QG dynamics is taken into account, ageostrophic advection of buoyancy further accelerates frontogenesis and potentially leads to frontal collapse, the formation of true discontinuities in buoyancy (Hoskins & Bretherton 1972). In this case, the submesoscale energy spectrum is modified to $E_{k_h} \sim {k_h}^{-2}$ (Boyd 1992) and the decay in the vertical direction becomes less rapid (Badin 2012). An additional modification of the dynamics by non-QG effects is that the release of potential energy leads to near-surface restratification (Hakim, Snyder & Muraki 2002; Lapeyre, Klein & Hua 2006). This effect is neglected in QG dynamics, where stratification is fixed. Non-QG surface frontogenesis also induces a finite forward flux of kinetic from small submesoscales (order 1 km) to dissipation scales, which is not present in QG dynamics (Capet *et al.* 2008*b*,*c*; Klein *et al.* 2008; Molemaker, McWilliams & Capet 2010). Despite these omissions, surface QG turbulence predicts many of the characteristics found in primitive equation simulations of mesoscale-driven surface frontogenesis (e.g. Klein *et al.* 2008). Our working hypothesis is thus that surface QG turbulence adequately captures the leading-order dynamics of the balanced flow in mesoscale-driven surface frontogenesis.

The upper ocean does not have a nearly constant PV, however, contrary to what is assumed in surface QG turbulence. Instead, a weakly-stratified mixed layer typically overlies a strongly-stratified thermocline (figure 1). There is a sharp step-like increase

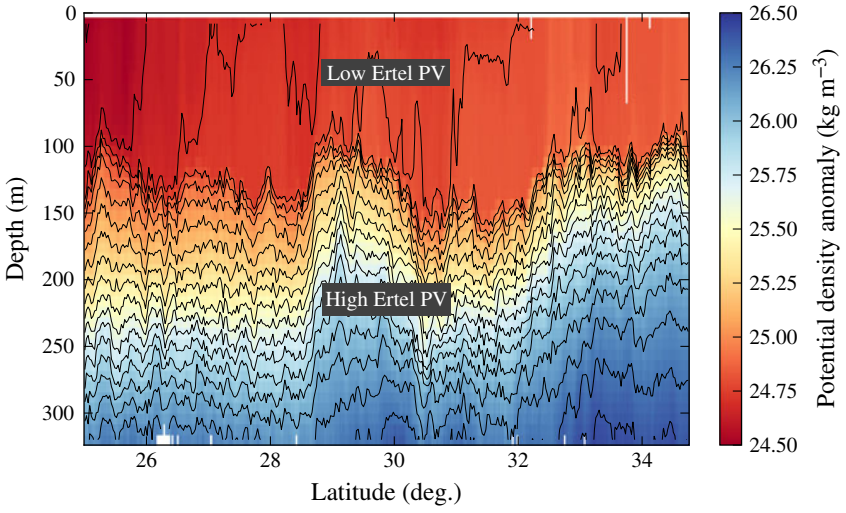

FIGURE 1. Potential temperature section from the wintertime eastern subtropical North Pacific. The data were obtained using a conductivity–temperature–depth sensor towed on a sawtooth profile along 140° W as part of the Spice experiment on 28 January–2 February, 1997. For more information on the data, see Ferrari & Rudnick (2000).

in stratification at the base of the mixed layer, corresponding to a step-like increase in background PV. This PV jump is dynamically important, because it supports edge waves that have the potential to interact with surface edge waves and thus produce a baroclinic instability in the mixed layer (e.g. Haine & Marshall 1998). This linear instability is to leading order captured by an Eady model with a rigid interface at the base of the mixed layer (Eady 1949). Corrections due to ageostrophic effects and a moveable interface at the mixed-layer base can be computed (Stone 1966*b*; Boccaletti *et al.* 2007), but for typical wintertime conditions, the instability scale and growth rate are qualitatively captured by Eady's QG model. When baroclinic mixed-layer instabilities grow to finite amplitude, turbulent scale interactions distribute energy across scales. Because of the rotational constraint, they transfer energy preferentially to larger scales. This nonlinear dynamics has been studied in idealized mixed-layer models, where baroclinic mixed-layer instabilities grow on a prescribed front (e.g. Boccaletti *et al.* 2007; Fox-Kemper, Ferrari & Hallberg 2008).

In the real ocean, baroclinic mixed-layer instabilities occur in the presence of an energetic mesoscale eddy field, so mixed-layer modes can grow on mesoscale buoyancy gradients and can be sheared by mesoscale strain fields. Realistic submesoscale-permitting models capture this dynamics (e.g. Mensa *et al.* 2013; Sasaki *et al.* 2014), but the models' complexity makes it hard to distil the essence of the dynamics and even these simulations only marginally resolve many submesoscale phenomena. In this article, we explore submesoscale dynamics by formulating a QG model that allows both mixed layer and thermocline instabilities. This simple model of submesoscale turbulence energized by baroclinic mixed-layer instabilities captures salient features of wintertime observations of submesoscale flows. If the mixed layer in this model is eliminated, the submesoscale dynamics reverts to surface QG turbulence, which allows a straightforward comparison of the two mechanisms that can energize submesoscale turbulence.

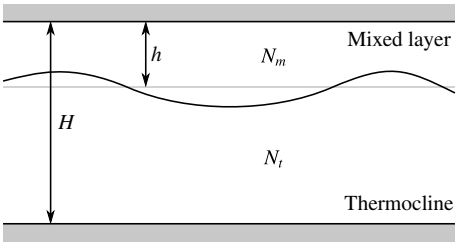

FIGURE 2. Schematic of the model set-up in a vertical–horizontal plane. There are rigid surfaces at $z = 0$ and $z = -H$ and a deformable interface at the mean depth $z = -h$, separating layers of constant stratification, N_m in the mixed layer and N_t in the thermocline.

We use QG scaling to formulate the dynamics of the model, which requires small Rossby and Froude numbers (e.g. Pedlosky 1987). Typical mesoscale Rossby and Froude numbers are on the order 0.1 and increase slowly with wavenumber if the submesoscales are energetic, reaching order 1 at scales of order 1 km (Callies *et al.* 2015). While the QG approximation does not apply anymore at these small scales, the QG system can be expected to capture the leading-order dynamics over the 10–100 km range.

A major limitation of QG dynamics in representing mixed-layer instabilities is that QG scaling does not allow for restratification. The weak mixed-layer stratification will be fixed in our QG model, whereas in reality there is a competition between restratification and atmospherically forced mixed-layer turbulence, which tends to keep the mixed layer deep and unstratified. The assumption is that the slower balanced dynamics described by the QG model develops on top of this background state, which is maintained by fast small-scale turbulence. A full description of mixed-layer dynamics will eventually need to consider the interplay of the fast and slow dynamics, a topic we hope to address in a future study (cf. Hamlington *et al.* 2014). We take up the discussion of how other non-QG effects may alter the dynamics toward the end of the article.

As described above, our inquiry into the dynamics of a weakly-stratified mixed layer coupled to a strongly-stratified thermocline is motivated by the study of the submesoscale upper ocean. The model we present and the dynamics we describe, however, have relevance for the atmosphere as well, where a weakly-stratified troposphere is capped by a strongly-stratified stratosphere (e.g. Eady 1949). The dynamics may also apply to the atmospheres of gas giants (e.g. Seiff *et al.* 1998) or other geophysical fluids that have layers of different stratification.

We formulate the model and give some physical intuition for its behaviour in § 2. In § 3, we investigate the linear dynamics of the model to understand its stability properties. This linear dynamics is suggestive of the fully-nonlinear turbulent dynamics that we address in § 4, where we analyse the energy spectra and fluxes for cases with and without baroclinic mixed-layer instabilities. We compare the results to observations in § 5 and conclude in § 6.

2. Model formulation

Consider two layers with constant stratification and constant mean shear on an f-plane, so that each layer has constant PV (figure 2). The upper layer represents the

mixed layer, which has a mean depth h, stratification N_m, and mean zonal shear Λ_m that is in thermal wind balance with the mean meridional buoyancy gradient $-f\Lambda_m$. The lower layer represents the thermocline and has stratification N_t and mean zonal shear Λ_t that is in thermal wind balance with the mean meridional buoyancy gradient $-f\Lambda_t$. The total depth is H. The layers are coupled through a deformable interface; flat rigid boundaries are assumed at both the surface and the bottom. The presence of a rigid bottom at the base of the thermocline is not realistic, but we will show that the bottom layer still captures the key thermocline physics relevant to our study. A weakly-stratified abyssal layer could be included but is omitted for simplicity, because it does not significantly affect the surface submesoscale dynamics of interest here. The approximation that the stratification is discontinuous at the base of the mixed layer is appropriate at horizontal scales larger than the deformation radius Nd/f associated with the transition depth d between the mixed layer and the thermocline (Smith & Bernard 2013). The transition at the base of the mixed layer is typically quite sharp (figure 1), so this deformation radius is much smaller than the submesoscales we are interested in here.

The assumption of a uniform PV within the two layers greatly simplifies the dynamics. PV conservation within the layers is trivial, as in the classic Eady (1949) problem. The flow in the interior of the layers is obtained by solving (1.1), with the boundary conditions supplied by the distribution of buoyancy at the surface and bottom and by matching conditions at the interface between the mixed layer and the thermocline.

In QG, the buoyancy anomaly b is governed by the horizontal advection of buoyancy anomalies by the geostrophic flow and by the vertical advection of the background buoyancy field,

$$\frac{\partial b}{\partial t} + \mathrm{J}(\psi, b) + wN^2 = 0. \tag{2.1}$$

At the surface and bottom, where the vertical velocity w vanishes, buoyancy anomalies are conserved under horizontal advection and (2.1) reduces to (1.2). To ensure that pressure is continuous at the interface, we require that the streamfunction ψ is continuous. Mass conservation requires that the vertical velocity w also is continuous. These conditions are applied at $z = -h$, consistent with QG scaling. The conservation equations for buoyancy just above the interface at $z = -h$,

$$\frac{\partial b^+}{\partial t} + \mathrm{J}(\psi_1, b^+) + wN_m^2 = 0, \quad b^+ = f\frac{\partial \psi}{\partial z}(-h^+), \tag{2.2}$$

and just below the interface,

$$\frac{\partial b^-}{\partial t} + \mathrm{J}(\psi_1, b^-) + wN_t^2 = 0, \quad b^- = f\frac{\partial \psi}{\partial z}(-h^-), \tag{2.3}$$

can then be combined to eliminate w, where ψ_1 denotes the streamfunction at $z = -h$. This gives a conservation law for the quantity

$$\theta_1 = f\left(\frac{b^+}{N_m^2} - \frac{b^-}{N_t^2}\right), \tag{2.4}$$

which is simply advected by the horizontal flow at the interface,

$$\frac{\partial \theta_1}{\partial t} + \mathrm{J}(\psi_1, \theta_1) = 0. \tag{2.5}$$

It should be noted that this does not ensure that buoyancy is continuous at $z=-h$. Instead, there is an implied interface displacement and buoyancy is continuous across the displaced interface. The displacement is small and, consistent with QG scaling, the matching conditions are applied at $z=-h$.

The quantity θ_1 is nothing but the integrated PV associated with the interface displacement, as can be verified by integrating

$$q=\nabla^2\psi+f\frac{\partial}{\partial z}\left(\frac{b}{N^2}\right) \tag{2.6}$$

across the interface. (The relative vorticity term vanishes because ψ is continuous across the interface.) While there are no PV anomalies within the two layers, the displacement of the interface between the layers induces a PV anomaly that, according to (2.5) and consistent with QG dynamics, is advected by the geostrophic flow at $z=-h$. The conservation equation (2.5) has been used to study the dynamics of the tropopause, which is similarly an interface between the weakly stratified fluid in the troposphere and the strongly stratified fluid in the stratosphere (Eady 1949; Rivest, Davis & Farrell 1992; Juckes 1994; Held *et al.* 1995).

The two-layer model can equivalently be interpreted as consisting of three PV sheets:

$$q=\theta_0\delta(z)+\theta_1\delta(z+h)+\theta_2\delta(z+H), \tag{2.7}$$

where δ is Dirac's delta function and $\theta_0=-fb/N_m^2$ at $z=0$, $\theta_2=fb/N_t^2$ at $z=-H$ and θ_1 at $z=-h$ is given in (2.4). PV is advected by the geostrophic flow, so

$$\frac{\partial\theta_j}{\partial t}+\mathrm{J}(\psi_j,\theta_j)=0, \tag{2.8}$$

where $j=0,\,1,\,2$ and ψ_j is the streamfunction at the level corresponding to θ_j. This formulation is simply an extension of Bretherton's (1966) representation of boundary conditions to include an interior PV sheet due to the deflection of an interface between layers of different stratification.

Note that even though θ_1 is only advected by the geostrophic flow, this does not imply that $w=0$ at the interface, much like the fact that interior PV anomalies in the QG system are only advected by the geostrophic flow does not imply that $w=0$. The vertical velocity is implicit in the dynamics and can be solved for using the omega equation (e.g. Hoskins *et al.* 1978).

To complete the dynamics, we require an inversion relation that allows us to obtain the streamfunctions ψ_j from the conserved quantities θ_j. For simplicity, we consider a doubly-periodic domain and express the inversion relation as a linear equation for Fourier coefficients of the variables θ_j and ψ_j:

$$\hat{\boldsymbol{\theta}}=\boldsymbol{L}\hat{\boldsymbol{\psi}},\quad \boldsymbol{\theta}=(\theta_0,\theta_1,\theta_2)^{\mathrm{T}},\quad \boldsymbol{\psi}=(\psi_0,\psi_1,\psi_2)^{\mathrm{T}}. \tag{2.9a–c}$$

where Fourier transforms are denoted by carets. The matrix $\boldsymbol{L}$, which depends on the zonal and meridional wavenumbers k and l, is determined by solving

$$-{k_h}^2\hat{\psi}+\frac{\partial}{\partial z}\left(\frac{f^2}{N^2}\frac{\partial\hat{\psi}}{\partial z}\right)=0 \tag{2.10}$$

in each layer, where $k_h=(k^2+l^2)^{1/2}$ is the magnitude of the horizontal wavenumber vector. The first column of $\boldsymbol{L}$ is determined by setting $\hat{\boldsymbol{\psi}}=(1,\,0,\,0)^{\mathrm{T}}$, solving (2.10)

Parameter	Symbol	Value
Mixed-layer depth	h	100 m
Total depth	H	500 m
Mixed-layer stratification	N_m	2×10^{-3} s^{-1}
Thermocline stratification	N_t	8×10^{-3} s^{-1}
Mixed-layer shear	Λ_m	10^{-4} s^{-1}
Thermocline shear	Λ_t	10^{-4} s^{-1}
Coriolis frequency	f	10^{-4} s^{-1}
Domain size	a	500 km
Numerical resolution	Δx	$\sim$1 km

TABLE 1. Parameters used throughout this article unless otherwise noted. These are typical of the wintertime midlatitude ocean.

for $\hat{\psi}(z)$ and subsequently calculating θ_0, θ_1 and θ_2. Repeating for $\hat{\boldsymbol{\psi}} = (0, 1, 0)^{\mathrm{T}}$ and $\hat{\boldsymbol{\psi}} = (0, 0, 1)^{\mathrm{T}}$ gives

$$\boldsymbol{L} = fk_h \begin{pmatrix} -\dfrac{\coth \mu_m}{N_m} & \dfrac{\operatorname{csch} \mu_m}{N_m} & 0 \\ \dfrac{\operatorname{csch} \mu_m}{N_m} & -\dfrac{\coth \mu_m}{N_m} - \dfrac{\coth \mu_t}{N_t} & \dfrac{\operatorname{csch} \mu_t}{N_t} \\ 0 & \dfrac{\operatorname{csch} \mu_t}{N_t} & -\dfrac{\coth \mu_t}{N_t} \end{pmatrix}, \tag{2.11}$$

where $\mu_m = N_m k_h h/f$ and $\mu_t = N_t k_h (H - h)/f$ are non-dimensional wavenumbers. This 3×3 matrix can easily be inverted.

This model can be generalized to an arbitrary number of layers of constant stratification and shear, which may be a useful way to approximate more realistic stratification and shear profiles. This is discussed in appendix A. The model can also be extended to include a density jump at the interface, as is sometimes present at the base of the mixed layer. The formulation is given in appendix B. Here we consider only the case of a continuous density profile, which is simpler and captures the essential physics of the submesoscale ocean.

To build intuition for the dynamics of the model, we illustrate the vertical structure of the flow associated with anomalies of θ_j at the surface, the interface and the bottom. Here and throughout the article, we use the parameters given in table 1, which are typical of the wintertime midlatitude ocean. (The values listed imply a mixed-layer Richardson number of $N_m^2/\Lambda_m^2 = 400$, which is larger than the order-1 Richardson numbers typically considered (e.g. Boccaletti *et al.* 2007). The relatively large Richardson number is the result of a relatively weak shear, which is chosen such that realistic energy levels are reached in the nonlinear simulations described below. QG dynamics overestimates the baroclinic growth rate for small Richardson numbers (Stone 1966*b*), which would result in unrealistically strong mixed-layer instabilities if a larger shear was chosen. It should also be noted that the leading-order QG dynamics can be rescaled to different Richardson numbers.) At the largest scales, for θ_j anomalies with wavelength $\lambda = 1000$ km or $k_h \ll f/N_t H$, the flow is nearly depth independent, irrespective of which level the anomaly is at (figure 3*a*). Around the thermocline deformation radius, at $\lambda = 100$ km or $k_h \sim f/N_t H \ll f/N_m h$, flow anomalies

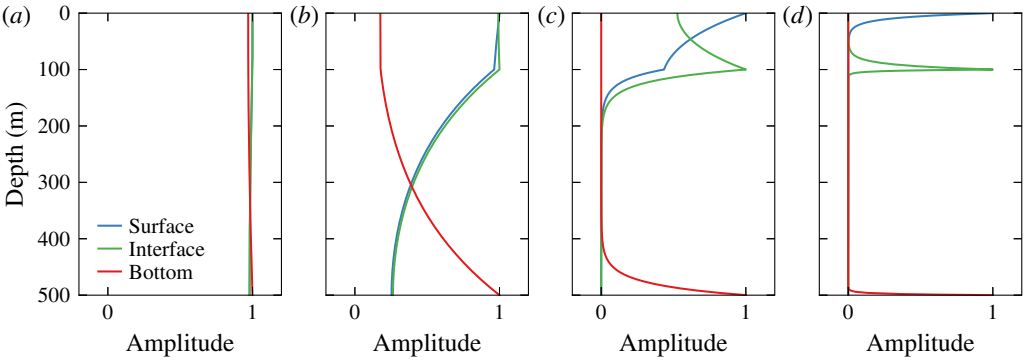

FIGURE 3. Vertical structure of streamfunction amplitude associated with anomalies of θ_0 (surface), θ_1 (interface) and θ_2 (bottom). Shown are the vertical profiles for θ_j anomalies with different horizontal wavenumbers $k_h = 2\pi/\lambda$. The wavelength λ is given in the respective panel title: (*a*) 1000 km wavelength; (*b*) 100 km wavelength; (*c*) 10 km wavelength; (*d*) 1 km wavelength.

significantly decay in the thermocline, while the flow is nearly uniform across the mixed layer (figure 3*b*). Surface (θ_0) and interface (θ_1) anomalies still induce significant flow at the bottom and *vice versa*. Around the mixed-layer deformation radius, at $\lambda = 10$ km or $k_h \sim f/N_m h \gg f/N_t H$, surface ($\theta_0$) and interface ($\theta_1$) anomalies induce very little flow at the bottom and *vice versa* (figure 3*c*). The flow now varies significantly across the mixed layer, but surface anomalies (θ_0) still induce significant flow at the interface and *vice versa*. At $\lambda = 1$ km or $k_h \gg f/N_m h$, all levels are decoupled: θ_j anomalies on any one of the levels induce very little flow at the other levels (figure 3*d*).

The dependence of the vertical flow structure on the horizontal scale of the anomalies illuminates the qualitative dynamics of the model. At the largest scales, the flow is essentially depth independent and follows two-dimensional dynamics. At scales $k_h \sim f/N_t H$, around the thermocline deformation radius, surface or interface anomalies can interact with bottom anomalies, allowing phase locking and a thermocline instability. At scales $k_h \sim f/N_m h$, around the mixed-layer deformation radius, surface and interface anomalies can interact, enabling an instability in the mixed layer. Bottom anomalies, on the other hand, are decoupled, so there is no thermocline instability at these scales. At the smallest scales, all three levels are independent and follow surface QG dynamics.

3. Linear stability analysis

We now analyse the linear stability of the model formulated above. This linear analysis reveals the nature of the instabilities that fuel the nonlinear turbulence, which we describe in the next section.

Blumen (1979) analysed short-wave instabilities in the atmosphere using a model consisting of two coupled constant-PV layers. He performed a linear stability analysis equivalent to what will be presented here. For completeness, we repeat the analysis in the context of upper ocean dynamics to emphasize the aspects most relevant for the nonlinear regime.

We consider the linear stability of normal-mode perturbations to a zonal flow with constant vertical shear Λ_m in the mixed layer and Λ_t in the thermocline (figure 4).

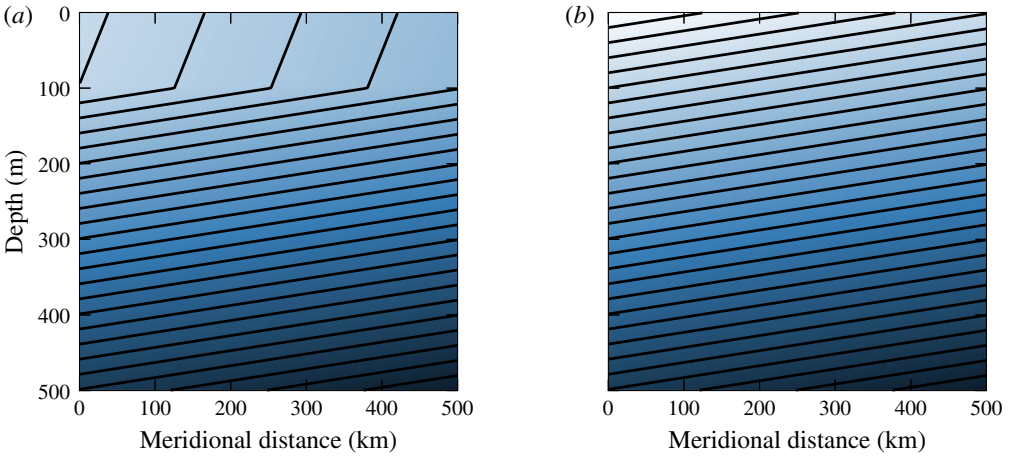

FIGURE 4. Mean buoyancy structure in the meridional–vertical plane for (*a*) the full model and (*b*) the thermocline-only model. The contours show isopycnals; light shading indicates more buoyant fluid.

The linearized conservation equation for the perturbations from this mean state, written in Fourier space, is

$$\frac{\partial \hat{\boldsymbol{\theta}}}{\partial t} + \mathrm{i}k\boldsymbol{U}\hat{\boldsymbol{\theta}} + \mathrm{i}k\boldsymbol{\Gamma}\hat{\boldsymbol{\psi}} = 0, \tag{3.1}$$

where the mean zonal flows and mean meridional PV gradients at the respective levels are represented by the diagonal elements of the matrices $\boldsymbol{U}$ and $\boldsymbol{\Gamma}$:

$$\boldsymbol{U} = \mathrm{diag}(0, -\Lambda_m h, -\Lambda_m h - \Lambda_t(H-h)), \tag{3.2}$$

$$\boldsymbol{\Gamma} = \mathrm{diag}(f^2\Lambda_m/N_m^2, -f^2\Lambda_m/N_m^2 + f^2\Lambda_t/N_t^2, -f^2\Lambda_t/N_t^2). \tag{3.3}$$

The system is Galilean invariant, so we are free to set the mean zonal flow to zero at the surface. Using the inversion relation (2.9*a*), we can replace the $\hat{\boldsymbol{\psi}}$ and obtain an equation for the $\hat{\boldsymbol{\theta}}$ coefficients only,

$$\frac{\partial \hat{\boldsymbol{\theta}}}{\partial t} + \mathrm{i}k\boldsymbol{U}\hat{\boldsymbol{\theta}} + \mathrm{i}k\boldsymbol{\Gamma}\boldsymbol{L}^{-1}\hat{\boldsymbol{\theta}} = 0. \tag{3.4}$$

Substituting $\hat{\boldsymbol{\theta}} = \tilde{\boldsymbol{\theta}}\mathrm{e}^{-\mathrm{i}\omega t}$, with complex frequency ω, turns this equation for $\hat{\boldsymbol{\theta}}$ into the eigenvalue problem

$$(\boldsymbol{U} + \boldsymbol{\Gamma}\boldsymbol{L}^{-1})\tilde{\boldsymbol{\theta}} = c\tilde{\boldsymbol{\theta}}, \tag{3.5}$$

where the eigenvalue is $c = \omega/k$. The real part of c is the zonal phase speed; the imaginary part gives the growth rate $\sigma = k\,\mathrm{Im}c$.

Being a third-order system, (3.5) can be solved analytically, but the solutions are rather complicated and give little useful insight. We instead explore the characteristics of the solutions numerically for the set of parameters given in table 1. We then explain the stability properties and parameter dependencies by considering special limit cases.

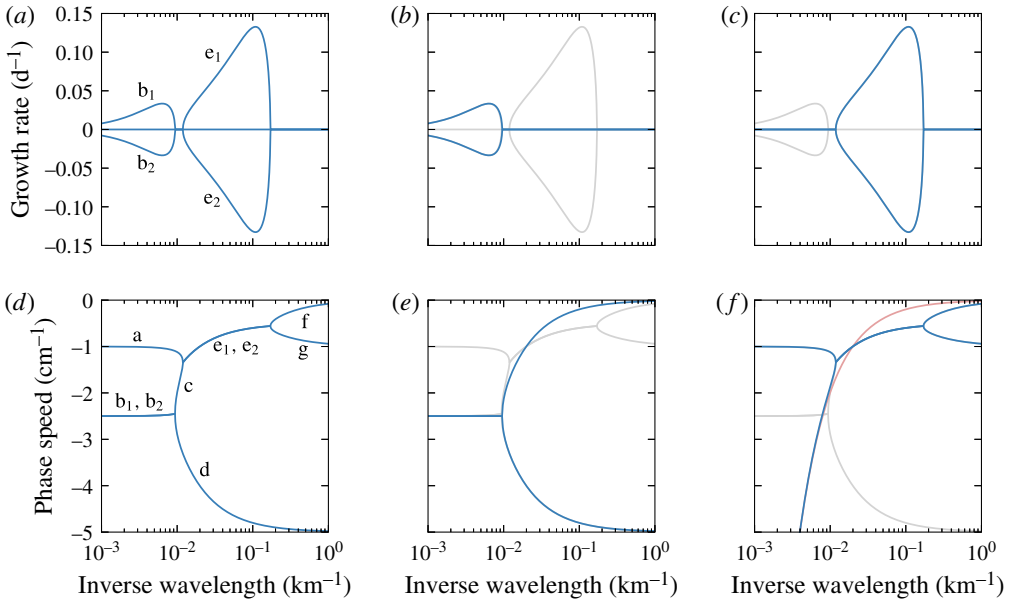

FIGURE 5. Linear stability analysis of the model equations. (*a*) Growth rates and (*d*) phase speeds of the full model, (*b*) growth rates and (*e*) phase speeds of the thermocline-only model, (*c*) growth rates and (*f*) phase speeds of the mixed-layer-only model. Growth rates and phase speeds are shown in blue; the growth rates and phase speeds of the full model are overlaid for reference in grey. The phase speed of a surface edge wave is given in faint red in (*f*).

3.1. *Full model*

We start by considering the growth rate σ as a function of horizontal wavenumber. The eigenvalue problem (3.5) only depends on $k_h = (k^2 + l^2)^{1/2}$, so the eigenvalue c is a function of k_h only. For a given k_h, the maximum growth rate $\sigma = k\,\mathrm{Im}c$ hence occurs at $l = 0$. We therefore only consider disturbances with no meridional dependence.

Plotting the growth rate σ as a function of zonal wavenumber k reveals that there are two lobes of instability: one at the mesoscale and one at the submesoscale (figure 5*a*, branches 'b' and 'e'). The maximum growth rates occur at zonal wavelengths of approximately 160 km (mesoscale) and 9 km (submesoscale). The two lobes can overlap, for example if the mixed layer is deeper or if a density jump at the base of the mixed layer is included (not shown). The submesoscale instability has a peak growth rate much larger than the mesoscale instability. The growth rates are similar to what Boccaletti *et al.* (2007) found in a linear QG stability analysis of a realistic mean state of the wintertime eastern subtropical North Pacific. The magnitudes are slightly smaller here, because the shear is slightly weaker. But the similarity of the instabilities supports that this model, despite being highly idealized, captures the essential physics of mesoscale and submesoscale instabilities. Whether it also captures the essential physics in the nonlinear regime will be discussed in §§ 4–6.

The mesoscale and submesoscale instabilities have very different vertical structures, as also noted by Boccaletti *et al.* (2007). The perturbation streamfunctions – derived from the eigenvectors of (3.5) – show that the fastest-growing mesoscale mode is

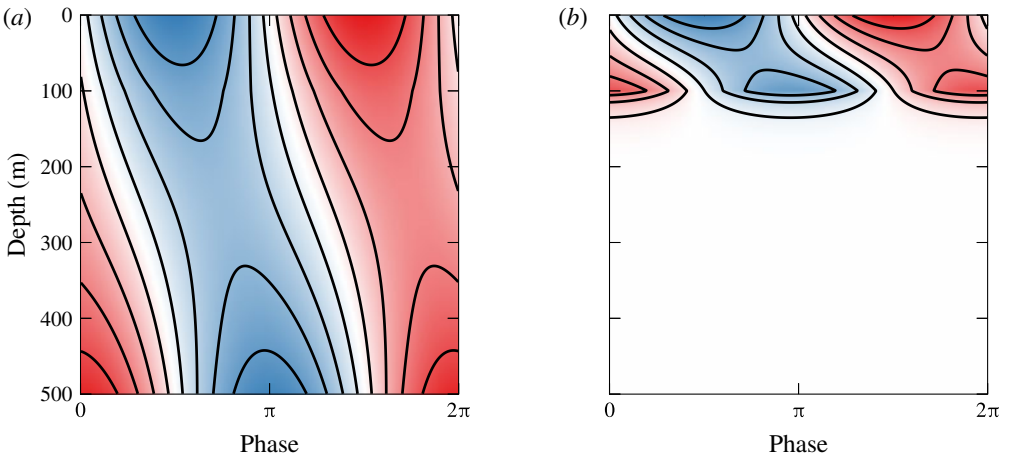

FIGURE 6. Perturbation streamfunction of the most unstable (*a*) mesoscale (160 km) and (*b*) submesoscale (9 km) modes of the full model. Red and blue shading represents positive and negative values, respectively.

deep and spans the entire water column (figure 6*a*), whereas the fastest-growing submesoscale mode is almost completely confined to the mixed layer, with only weak penetration into the thermocline below 100 m depth (figure 6*b*). Both modes exhibit the familiar pattern of baroclinically unstable modes with streamfunction perturbations tilted into the shear, as is necessary to extract potential energy from the mean flow.

Further insight into the dynamics of the model can be gained by considering the propagation speeds of the linear modes in conjunction with their growth rates (figure 5*a*,*d*). Being a third-order system, the model has three normal modes at each wavenumber. In both lobes of instability, the growing modes are conjugate to decaying modes, which have the same phase speeds. This is the familiar phase locking of counter-propagating waves in baroclinic instability (branches 'b' and 'e'). In these unstable wavenumber ranges, there is an additional neutral mode (branches 'a' and 'd'). At wavenumbers with no instability, all three modes have distinct phase speeds – no phase locking occurs. We will discuss the dynamics of the various branches by considering approximations to the full model.

3.2. *Thermocline only*

We start by examining the deep mesoscale instability of the full model. As discussed in the model formulation (§ 2), the mesoscale modes are deep and only slightly modified by the presence of the mixed layer. We can understand the mesoscale instability by eliminating the mixed layer altogether and considering a thermocline-only model (figure 4*b*). This amounts to setting $h = 0$ (or $N_m = N_t$ and $\Lambda_m = \Lambda_t$) in the full model. In this limit, the model reduces to one layer with the dynamics controlled by buoyancy advection at the surface and bottom only – the classic Eady (1949) model. The inversion matrix (2.11) reduces to

$$\boldsymbol{L} = f k_h \begin{pmatrix} -\dfrac{\coth \mu_t}{N_t} & \dfrac{\operatorname{csch} \mu_t}{N_t} \\ \dfrac{\operatorname{csch} \mu_t}{N_t} & -\dfrac{\coth \mu_t}{N_t} \end{pmatrix}, \tag{3.6}$$

and the matrices representing the mean flow are

$$\boldsymbol{U} = \mathrm{diag}(0, -\Lambda_t H), \tag{3.7}$$

$$\boldsymbol{\Gamma} = \mathrm{diag}(f^2 \Lambda_t / N_t^2, -f^2 \Lambda_t / N_t^2). \tag{3.8}$$

Solving the eigenvalue problem (3.5) with these matrices, we find the eigenvalues

$$c = -\frac{\Lambda_t H}{2} \pm \frac{\mathrm{i} \Lambda_t H}{\mu_m} \left(\mu_t \coth \mu_t - 1 - \frac{\mu_t^2}{4} \right)^{1/2}, \tag{3.9}$$

where $\mu_t = N_t k_h H / f$ is the non-dimensional wavenumber (Eady 1949).

The solution (3.9) shows that this thermocline-only model has a baroclinic instability near the thermocline deformation radius $N_t H/f$. The maximum growth rate $\sigma = 0.31 f \Lambda_t / N_t$ occurs at $\mu_t = 1.6$ and $l = 0$, which corresponds to a zonal wavelength $\lambda = 3.9 N_t H/f$. The growth curve for this thermocline-only model approximates the mesoscale lobe of the full model very well (figure 5*b*). The short-wave cutoff in the Eady model at $\lambda = 2.6 N_t H/f$ nearly coincides with the short-wave cutoff of the mesoscale instability in the full model. The phase speed of the phase-locked waves $-\Lambda_t H/2$ very nearly matches the phase speed of the unstable mesoscale mode of the full model (figure 5*e*). The split at the short-wave cutoff into surface and bottom modes also features in the full model. In the thermocline-only model, these surface and bottom modes are very nearly Eady edge waves that do not sense the other boundary. The bottom mode of the thermocline-only model very nearly matches that of the full model (branch 'd'). The surface mode of the thermocline-only model traces out branch 'c' of the full model, but then the full model transitions to dynamics associated with the mixed layer that are not present in the thermocline-only model.

This comparison shows that the mesoscale instability of the full model very nearly follows Eady dynamics. The presence of the mixed layer only modifies the characteristics of the instability slightly. At submesoscales, on the other hand, the thermocline-only model has surface QG dynamics, as opposed to the mixed-layer dynamics of the full model.

3.3. *Mixed layer only*

Turning our attention to the submesoscale instability, we note that the submesoscale instability peaks around the deformation radius of the mixed layer $N_m h/f$. Based on the discussion in the model formulation (§ 2) and the vertical structure of this instability (figure 6*b*), we anticipate that this instability arises from the interaction between anomalies at the surface and at the interface between mixed layer and thermocline.

In a first attempt to isolate the submesoscale instability, we disregard the possibility that surface and interface anomalies induce flow in the thermocline and assume a rigid bottom at the base of the mixed layer. This reduces the full model to an Eady model for the mixed layer, which is the limit of infinite thermocline stratification. This Eady model reasonably approximates the location and magnitude of the peak growth rate with $\lambda = 3.9 N_m h/f = 8$ km and $\sigma = 0.31 f \Lambda_m / N_m = 0.13$ d^{-1}. The Eady model captures the short-wave cutoff of the full model, but misses the long-wave cutoff. This suggests that the fastest-growing mode approximately follows Eady dynamics as if the thermocline acted like a rigid bottom, but also that larger-scale modes are significantly modified by reaching into the thermocline.

All features of the submesoscale instability are captured if flow in the thermocline is allowed. To still isolate the submesoscale instability, we consider again the layered model but let the thermocline be infinitely deep. That eliminates bottom edge waves, so no mesoscale instability occurs. Eady (1949) considered the upside-down atmospheric analogue to this system, relaxing the assumption that a rigid lid is placed at the tropopause.

Our system with no bottom again reduces to two variables; the inversion matrix is

$$\boldsymbol{L} = fk_h \begin{pmatrix} -\dfrac{\coth \mu_m}{N_m} & \dfrac{\operatorname{csch} \mu_m}{N_m} \\ \dfrac{\operatorname{csch} \mu_m}{N_m} & -\dfrac{\coth \mu_m}{N_m} - \dfrac{1}{N_t} \end{pmatrix} \tag{3.10}$$

and the mean flow is represented by

$$\boldsymbol{U} = \operatorname{diag}(0, -\Lambda_m h), \tag{3.11}$$

$$\boldsymbol{\Gamma} = \operatorname{diag}(f^2 \Lambda_m / N_m^2, -f^2 \Lambda_m / N_m^2 + f^2 \Lambda_t / N_t^2). \tag{3.12}$$

In the case $\Lambda = \Lambda_m = \Lambda_t$, the solution to the eigenvalue problem (3.5) is

$$c = -\frac{\Lambda h}{2}\left(1 + \frac{\alpha}{\mu_m}\right) \pm \frac{\mathrm{i}\Lambda h}{\mu_m}\left[\frac{(1-\alpha^2)(\mu_m - \tanh \mu_m)}{\tanh \mu_m + \alpha} - \frac{1}{4}(\mu_m - \alpha)^2\right]^{1/2}, \tag{3.13}$$

where $\alpha = N_m/N_t$ (Eady 1949; Blumen 1979). This converges to the classic Eady solution if $\alpha \ll 1$ and $\alpha \ll \mu_m$, which is equivalent to $N_t \gg N_m$ and $k_h \gg f/N_t h$. This shows that large thermocline stratification acts like a rigid bottom, but only for scales that are not too large, as alluded to above. Modes of large horizontal scale penetrate into the thermocline and their dynamics is altered.

The growth rates and phase speeds of this reduced model very nearly match the growth rates and phase speeds of the full model at scales smaller than about 100 km (figure 5*c*,*f*). This model now captures the long-wave cutoff of the submesoscale instability. At large scales, where $\mu_m \ll \alpha$ and $\mu_m \ll 1$ or equivalently $k_h \ll f/N_t h$ and $k_h \ll f/N_m h$, the dynamics split into modes that are barotropic and baroclinic in the mixed layer. The barotropic mode behaves like a surface edge wave, which has a phase speed $-f\Lambda/N_t k_h$ and does not sense the mixed layer (figure 5*f*). The baroclinic mode is baroclinic in the mixed layer and remains shallow for large scales – its critical level is the base of the mixed layer and its phase speed is $-\Lambda h$. The vastly different phase speeds of these two modes prevent phase locking, so no instability occurs at large scales. This stabilization is analogous to that by the β-effect (Phillips 1954; Lindzen 1994; Vallis 2006). Note that no tilt in the interface is required for this long-wave cutoff (cf. Boccaletti *et al.* 2007). For the unstable modes, the reduced model with no bottom also captures the deepening of the critical level as the scale increases, $-\Lambda h(1 + f/N_t k_h h)/2$, which is due to the increasing penetration of the unstable mode into the thermocline.

The location of the long-wave cutoff in this constant-shear case depends on the ratio N_m/N_t. In the more general case $\Lambda_m \neq \Lambda_t$, it also depends on the ratio Λ_m/Λ_t. No long-wave cutoff occurs if $\Lambda_t = 0$, as found by Rivest *et al.* (1992), who considered the atmospheric case with no shear in the stratosphere. There is also no long-wave cutoff if $N_m/N_t \to 0$, which is the Eady limit. The instability itself requires a reversal of the PV gradient, so the condition for instability is $\Lambda_m/N_m^2 > \Lambda_t/N_t^2$. This condition is typically satisfied in the ocean, because the thermocline stratification is much larger than the mixed-layer stratification and horizontal buoyancy gradients are typically of the same order in the mixed layer as in the thermocline, if not larger.

3.4. *Summary*

We are now in a position to understand all branches in the phase speed diagram of the full model. Branch 'a' is a mode that is baroclinic in the mixed layer and does not penetrate much into the thermocline. It does not sense the bottom. Branch 'b_1' is the unstable branch corresponding to the Eady-like thermocline instability; branch 'b_2' is the conjugate decaying branch. Branch 'c' is a mode that is nearly barotropic in the mixed layer and behaves like a surface edge wave in the thermocline. It does not interact much with the bottom. Branch 'd' is a bottom edge wave that is independent of the surface and interface. Branch 'e_1' is the unstable branch corresponding to the mixed-layer instability; branch 'e_2' is the conjugate decaying branch. The instability is significantly modified by the modes' penetration into the thermocline, but the scale and growth rate of the most unstable mode still scale with the mixed-layer deformation radius and the Eady growth rate. Branches 'f' and 'g' are edge waves propagating on the surface and the interface that do not interact with any of the other edge waves.

4. Nonlinear dynamics

We now turn to the nonlinear dynamics that arises when perturbations are amplified by the instabilities and grow to finite amplitude. We solve numerically the full nonlinear equations

$$\frac{\partial \boldsymbol{\theta}}{\partial t} + \boldsymbol{U}\frac{\partial \boldsymbol{\theta}}{\partial x} + \boldsymbol{\Gamma}\frac{\partial \boldsymbol{\psi}}{\partial x} + \mathrm{J}(\boldsymbol{\psi}, \boldsymbol{\theta}) = r\nabla^{-2}\boldsymbol{\theta} - \nu(-\nabla^2)^n\boldsymbol{\theta}, \tag{4.1}$$

where the Jacobian operator is understood to act element-wise:

$$\mathrm{J}(\boldsymbol{\psi}, \boldsymbol{\theta}) = (\mathrm{J}(\psi_0, \theta_0), \mathrm{J}(\psi_1, \theta_1), \mathrm{J}(\psi_2, \theta_2))^{\mathrm{T}}. \tag{4.2}$$

These are the evolution equations for perturbations from the prescribed mean zonal flow, which appears in form of the diagonal matrices $\boldsymbol{U}$ and $\boldsymbol{\Gamma}$. We consider flows that are doubly periodic in the perturbations, so no modification of the prescribed mean can occur. We introduce hypoviscosity with coefficient r, which provides a drag to remove energy from large scales, and hyperviscosity with coefficient ν and order n, which helps ensure numerical stability and absorbs enstrophy at small scales. Hypoviscosity is a convenient but somewhat unphysical choice. We introduce it to halt the inverse cascade and allow for mesoscale equilibration. Hypoviscosity appears in the dynamical equations for the conserved quantities at the surface, interface and bottom, but it can be thought of as acting throughout the layers. If applied to buoyancy and momentum, it does not affect PV and PV conservation within the layers remains trivial. Linear drag cannot prevent an inverse cascade to the domain scale without significantly damping the instabilities.

We integrate these equations on a 500 km $\times$ 500 km domain using a fully dealiased pseudo-spectral code with a resolution 512×512. The time derivatives are discretized using a fourth-order Runge–Kutta scheme. The hypoviscosity coefficient is $r = 10^{-16}$ m^{-2} s^{-1}; the hyperviscosity is of order $n = 10$ and the coefficient is $\nu = 2.5 \times 10^{46}$ m^{20} s^{-1}. All calculations are initialized with white noise of small amplitude in θ_j.

Before considering the combined effect of mesoscale and mixed-layer instabilities, we first consider them separately. We start with the thermocline-only model, which allows only the mesoscale thermocline instability while submesoscale flows follow surface QG dynamics. We subsequently contrast this case with the mixed-layer-only model, which allows only the submesoscale mixed-layer instability. We finally consider the full model, in which both instabilities occur.

4.1. *Thermocline only*

We start by studying surface QG turbulence generated by mesoscale eddies, one of the proposed mechanisms to energize submesoscale flows. Surface QG flows cannot themselves extract energy from the mean flow, so they must be forced at the mesoscale. Instead of prescribing external forcing (e.g. Pierrehumbert, Held & Swanson 1994; Scott 2006), we use the thermocline-only Eady model to generate mesoscale eddies that in turn generate the submesoscale flows. As we saw in the linear stability analysis, the dynamics of surface buoyancy anomalies in this model decouple from the bottom at scales smaller than the thermocline deformation radius, so flows very nearly follow surface QG dynamics at the submesoscales, i.e. the scales smaller than the thermocline deformation radius. The mesoscale instability is an obviously crude representation of the real mesoscale instability, with no interior PV gradients and the presence of an artificial rigid interface at the base of the thermocline. But the instability does generate mesoscale eddies of roughly the right scale, which is sufficient to drive the submesoscale surface QG flows (cf. Roullet *et al.* 2012).

Since the dissipative terms are weak in the linear regime, the instability grows until it reaches finite amplitude, when the nonlinear terms become important. Secondary instabilities set in and the flow quickly evolves into a fully turbulent regime. The perturbations grow in scale until they reach a scale where hypoviscosity is significant. Thereby, the flow comes into statistical equilibrium, which is the time period considered in what follows.

A snapshot from the equilibrated state exhibits a patchy surface buoyancy field with strong buoyancy gradients (figure 7*b*). The largest eddies are about 200 km in scale. The strongest coherent vortices have a scale of about 50 km. Smaller-scale vortices are present, but are weaker the smaller the scale. They result from a roll-up instability that features prominently in the evolution of the flow (Held *et al.* 1995).

As typical for turbulent flows, a continuum of scales is energized. This is quantified by the kinetic and potential energy spectra in statistical equilibrium, $\langle K_{k,l}\rangle$ and $\langle P_{k,l}\rangle$, which are defined by

$$K_{k,l} = \frac{1}{2}(|\hat{u}|^2 + |\hat{v}|^2), \quad P_{k,l} = \frac{1}{2}\frac{|\hat{b}|^2}{N^2}. \tag{4.3a,b}$$

The angle brackets denote an average in time, performed over the statistical equilibrium, and u and v denote the leading-order zonal and meridional geostrophic velocity components. Isotropic spectra $\langle K_{k_h}\rangle$ and $\langle P_{k_h}\rangle$ are computed by averaging $\langle K_{k,l}\rangle$ and $\langle P_{k,l}\rangle$ over circles of constant k_h in wavenumber space – the statistics are very nearly isotropic.

The surface spectra of both kinetic and potential energy peak at a wavelength of approximately 200 km and fall off roughly like $\langle K_{k_h}\rangle \sim \langle P_{k_h}\rangle \sim k_h^{-5/3}$ (figure 8), as predicted by surface QG turbulence theory for scales smaller than the scales at which mesoscale instabilities inject energy into the system (Blumen 1978). Since small-scale modes decay more rapidly in the vertical than large-scale modes, the spectra are steeper in the interior (e.g. Scott 2006). At 100 m depth, the mesoscale energy levels are similar to those at the surface, but submesoscale energy levels are much lower.

A useful diagnostic of turbulent dynamics is the spectral energy budget (e.g. Larichev & Held 1995; Roullet *et al.* 2012). While the dynamics is completely determined by the advection of conserved quantities at the surface and bottom, we first consider the energy budget over the entire depth range. We will take into account the reduced nature of these models below, where we present a vertically integrated energy budget for the mixed-layer-only case.

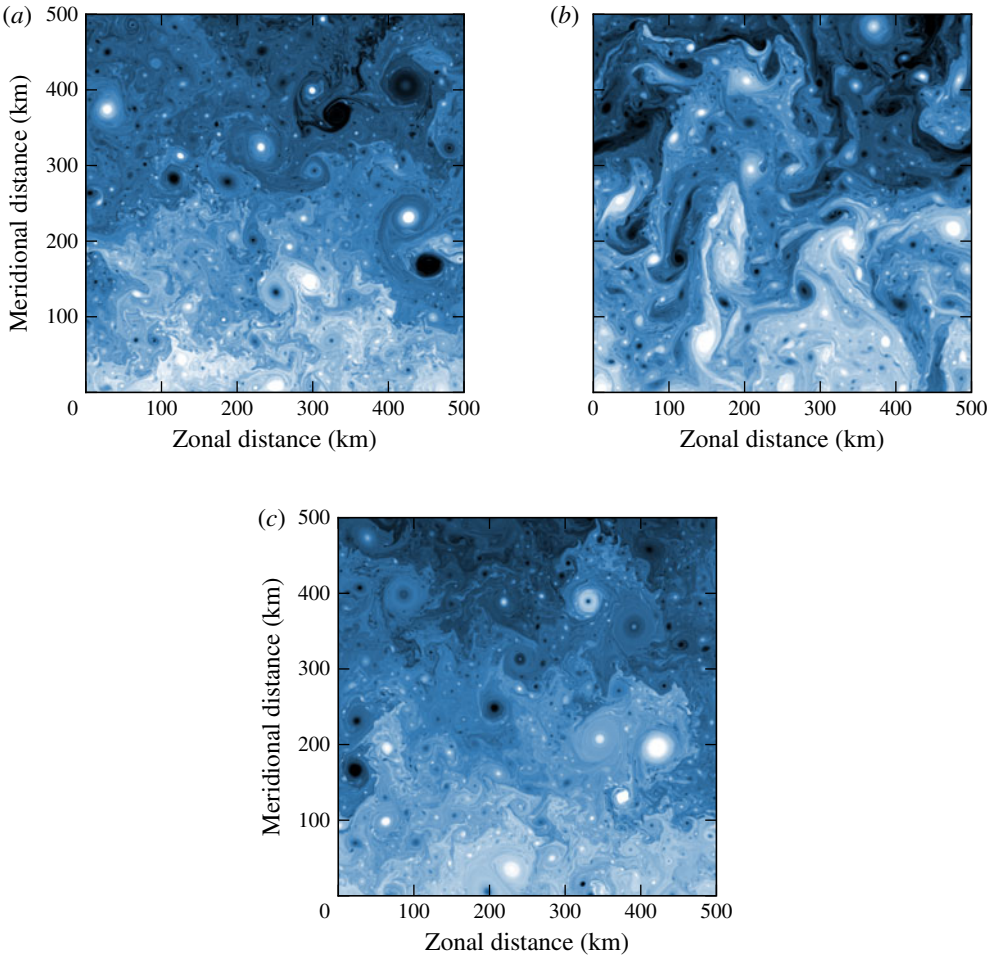

FIGURE 7. Snapshots of surface buoyancy (mean plus anomalies) from the equilibrated states of the (*a*) the full model, (*b*) the thermocline-only model, and (*c*) the mixed-layer-only model. The colour scale extends from white (more buoyant) through blue to black (less buoyant) and extends between $\pm f\Lambda a$.

The equations for the spectral perturbation potential and kinetic energies are

$$\frac{\partial P_{k,l}}{\partial t} = \mathrm{Re}\left[\frac{f\Lambda}{N^2}\hat{v}^*\hat{b} - \hat{w}^*\hat{b} - \frac{1}{N^2}\hat{b}^*\hat{\mathrm{J}}(\psi, b)\right] - (rk_h{}^{-2} + \nu k_h{}^{2n})P_{k,l} \tag{4.4}$$

$$\frac{\partial K_{k,l}}{\partial t} = \mathrm{Re}\left[-f\frac{\partial}{\partial z}(\hat{w}^*\hat{\psi}) + \hat{w}^*\hat{b} + \hat{\psi}^*\hat{\mathrm{J}}(\psi, \nabla^2\psi)\right] - (rk_h{}^{-2} + \nu k_h{}^{2n})K_{k,l}, \tag{4.5}$$

where the asterisks denote complex conjugates and Re denotes taking the real part. The Fourier transforms in the square bracket are all understood to be evaluated at the wavenumbers k and l. The first term on the right-hand side of the potential energy equation represents the extraction of potential energy from the mean flow. The second term represents the conversion from potential to kinetic energy. This term appears as a source term in the kinetic energy budget. The third term in the potential

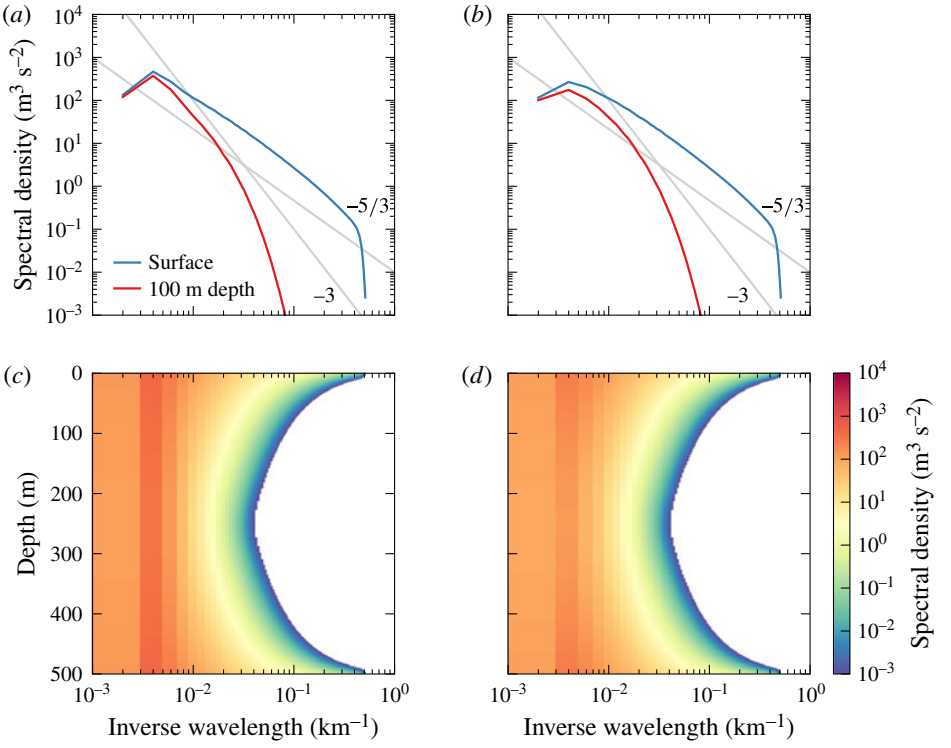

FIGURE 8. Wavenumber spectra of kinetic and potential energy from the thermocline-only simulation. (*a*) Kinetic and (*b*) potential energy spectra at the surface and 100 m depth, spectral density of (*c*) kinetic and (*d*) potential energy in the wavenumber–depth plane. In (*c*,*d*), no values below 10^{-3} m^3 s^{-2} are shown. Reference lines with slopes -3 and $-5/3$ are shown in grey.

energy budget represents spectral transfer by triadic interactions. The sum of this term over all wavenumbers vanishes. An equivalent spectral transfer term appears in the kinetic energy budget (third term). Kinetic energy can also be distributed vertically by pressure fluxes, represented by the first term in the kinetic energy budget. The vertical integral of this term vanishes. The viscosity terms act as sinks for both potential and kinetic energy – hypoviscosity acting at large scales, hyperviscosity at small scales. We present these budgets averaged azimuthally in wavenumber space and over time.

The extraction of potential energy from the mean is dominated by the largest, most energetic eddies (figure 9*a*). The extraction is independent of depth, because $q = 0$ and therefore

$$0 = \mathrm{Re}\hat{v}^*\hat{q} = \mathrm{Re}\frac{\partial}{\partial z}\left(\frac{f}{N^2}\hat{v}^*\hat{b}\right), \tag{4.6}$$

where it was used that the term involving advection of relative vorticity vanishes. Potential energy is transferred downscale by triadic interactions and deposited near the deformation radius as well as in wedges near the surface and the bottom that reach to much smaller scales (figure 9*b*). Where potential energy is deposited by scale interactions, it is converted into kinetic energy (figure 9*c*). Near the mesoscale deformation radius, this conversion is due to the mesoscale instability that produces vertical buoyancy fluxes. In the wedges near the surface and bottom, the conversion

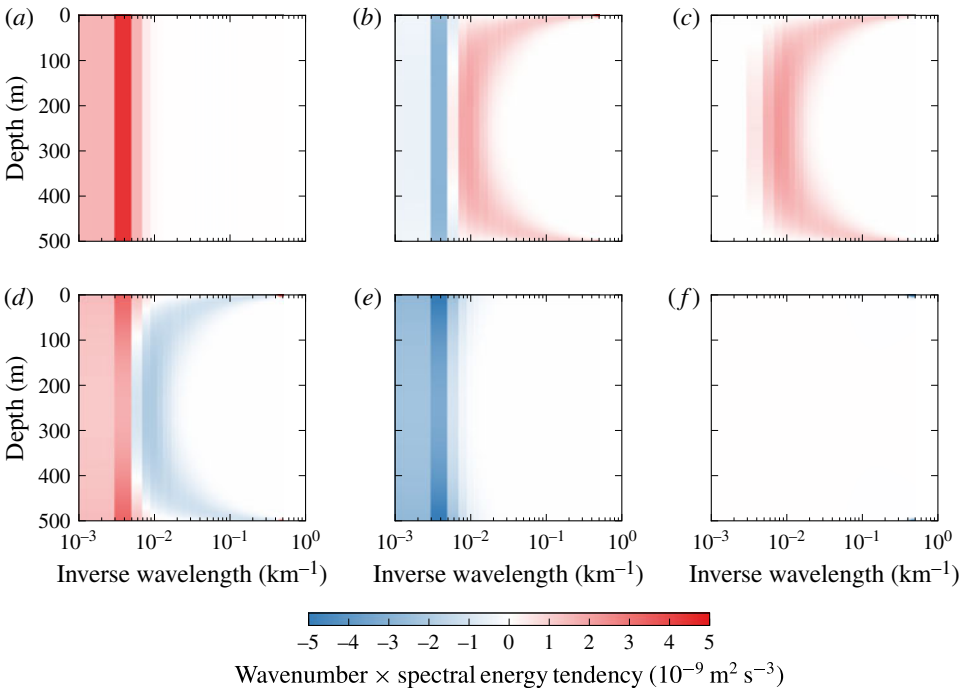

FIGURE 9. Spectral energy budget for the thermocline-only simulation. The terms are (*a*) potential energy extraction from the mean, (*b*) spectral potential energy flux divergence, (*c*) potential to kinetic energy conversion, (*d*) kinetic energy flux divergence, including spectral flux and pressure flux, (*e*) hypoviscosity on both kinetic and potential energy, and (*f*) hyperviscosity on both kinetic and potential energy. All terms are multiplied by the wavenumber to compensate for logarithmic shrinking.

is due to frontogenesis and secondary instabilities present in the surface QG cascades, which occur independently at the surface and the bottom (Roullet *et al.* 2012). The kinetic energy thus created is transferred back to large scales (figure 9*d*). The bulk of the energy is dissipated through hypoviscosity at the scales of the largest, most energetic, eddies (figure 9*e*). The energy dissipation through hyperviscosity is small, which reflects the fundamental property of geostrophic turbulence that energy is trapped at large scales and viscous energy dissipation vanishes in the limit of infinite resolution and zero (hyper-)viscosity (Kraichnan 1967; Charney 1971).

4.2. *Mixed layer only*

We now turn our attention to the nonlinear dynamics of the submesoscale mixed-layer instability and compare its turbulent dynamics to the surface QG turbulence of the thermocline-only case. We study the case with an infinitely deep thermocline, which allows an accurate representation of the submesoscale instability, while eliminating the thermocline instability (figure 5*c*,*f*).

The submesoscale instability grows to finite amplitude and the flow becomes turbulent. There is a turbulent spin-up phase, in which the eddies, which are initially of the size of the instability, grow larger until they reach a statistical equilibrium with hypoviscosity. The flow is host of numerous coherent vortices embedded in a filamentary sea with strong buoyancy gradients (figure 7*c*). A snapshot of surface

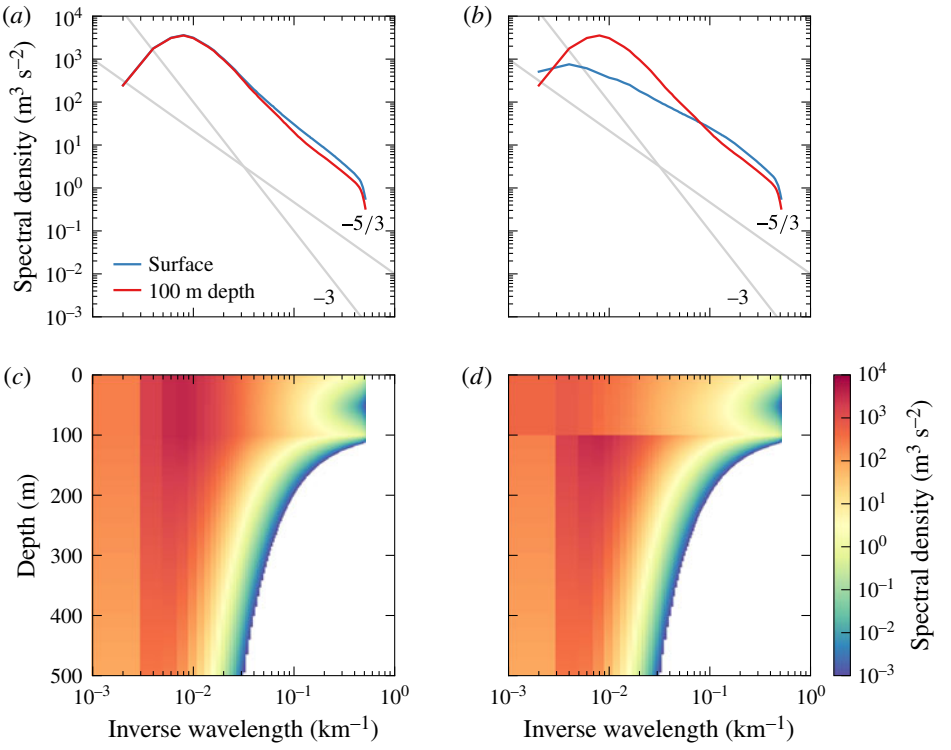

FIGURE 10. Wavenumber spectra of kinetic and potential energy from the mixed-layer-only simulation. (*a*) Kinetic and (*b*) potential energy spectra at the surface and 100 m depth (just below mixed-layer base), spectral density of (*c*) kinetic and (*d*) potential energy in the wavenumber–depth plane. In (*c*,*d*), no values below 10^{-3} m^3 s^{-2} are shown. Reference lines with slopes -3 and $-5/3$ are shown in grey.

buoyancy appears quite different from the thermocline-only case, but this visual difference is due mostly to the smaller size of the most energetic eddies.

The energy spectra reflect the nearly frontal structure at the surface (figure 10). The kinetic energy spectra fall off slightly more steeply than $\langle K_{k_h} \rangle \sim {k_h}^{-5/3}$ in the scale range of the linear instability and like $\langle K_{k_h} \rangle \sim {k_h}^{-5/3}$ at scales smaller than the linear short-wave cutoff, both at the surface and at the base of the mixed layer at 100 m depth (figure 10*a*). The mixed-layer instabilities energize the entire depth of the mixed layer. This is in sharp contrast to the thermocline-only simulation, in which surface QG turbulence energizes a thin wedge close to the surface only.

The equilibrated flow in the mixed-layer-only case is much more energetic than in the thermocline-only case and more energetic than is realistic. While the equilibration by hypoviscosity is unrealistic, we will see that the enhanced energy levels are due to more efficient extraction of mean potential energy in the weakly-stratified mixed layer, which is a dynamical property of the system that does not depend on how the flow is equilibrated. We will discuss possible reasons for these unrealistically high energy levels in § 5.

Below the base of the mixed layer, the potential energy spectra are the same as the kinetic energy spectra (figure 10*c*). In the mixed layer, the potential energy spectra are significantly flatter than the kinetic energy spectra. This is in contrast

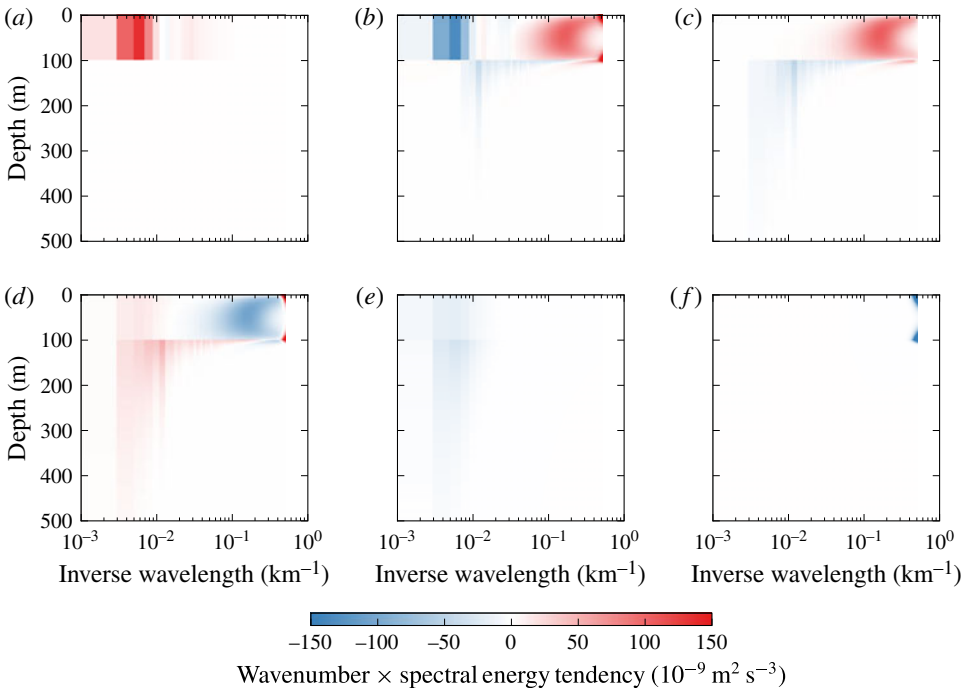

FIGURE 11. Spectral energy budget for the mixed-layer-only simulation. The terms are (*a*) potential energy extraction from the mean, (*b*) spectral potential energy flux divergence, (*c*) potential to kinetic energy conversion, (*d*) kinetic energy flux divergence, including spectral flux and pressure flux, (*e*) hypoviscosity on both kinetic and potential energy and (*f*) hyperviscosity on both kinetic and potential energy. All terms are multiplied by the wavenumber to compensate for logarithmic shrinking.

to observations that show rough equipartition between kinetic and potential energy (Callies & Ferrari 2013; Callies *et al.* 2015). We currently do not understand the reason for this difference.

The vertical structure of energy shows that the mixed-layer instabilities also energize the thermocline below (figure 10*b*,*d*). At the instability scale, the flow does not reach much into the thermocline. But as the horizontal scale of the flow increases, so does the vertical scale. The flow exhibits the familiar property of geostrophic turbulence and barotropizes as it increases its horizontal scale (Charney 1971; Smith & Vallis 2001).

The energy transfer into the thermocline is best examined through the spectral energy budget (figure 11). Potential energy is extracted at the scale of the largest, most energetic eddies, but the extraction is confined to the mixed layer (figure 11*a*). Potential energy is transferred from the extraction scale to the scale of the mixed-layer instability (figure 11*b*). The mixed-layer instability converts potential energy into kinetic energy in the mixed layer, at the instability scale (figure 11*c*). The kinetic energy created by the instability undergoes an inverse cascade, in which energy is not only transferred to large horizontal scales, but also vertically into the thermocline (figure 11*d*). The deposition of kinetic energy at the scale of the largest eddies is well distributed across the mixed layer and upper thermocline. The vertical distribution of the energy sink through hypoviscosity confirms that the flow extends below the mixed layer at the scale of the largest eddies, where hypoviscosity acts (figure 11*e*).

Hyperviscosity acts only at the smallest resolved scales (figure 11*f*). While small, it does affect the other terms in the budget. We do not discuss its effects any further, because they are expected to disappear if the resolution is increased and the hyperviscosity coefficient decreased.

These energy pathways are reminiscent of the phenomenology of two-layer baroclinic turbulence. The turbulent dynamics of a two-layer system can be understood in terms of a dual cascade (Rhines 1977; Salmon 1978). Baroclinic energy is extracted from the mean at the scale of the largest, most energetic, eddies. The barotropic flow dominates at these scales and transfers the baroclinic energy downscale. The baroclinic mode behaves like a passive tracer at these scales. Around the deformation radius, the instability converts baroclinic energy into barotropic energy. The barotropic energy then enters an inverse cascade, which is arrested at some scale by drag or hypoviscosity. The forward cascade of baroclinic energy is compensated by the inverse cascade of barotropic energy, such that no spectral transfer of total energy occurs. This is consistent with the phenomenology that all sources and sinks of total energy occur at the scale of the largest, most energetic eddies – no energy is dissipated at small scales.

Can the turbulent dynamics induced by mixed-layer instabilities be understood in similar terms? To pursue the analogy, we must first introduce a modal decomposition of the eddy energy. In our system, the vertically integrated total energy can be written entirely in terms of the quantities at the surface and interface:

$$E_{k,l} = -\tfrac{1}{2}\hat{\boldsymbol{\psi}}^{\dagger}\hat{\boldsymbol{\theta}} = -\tfrac{1}{2}\hat{\boldsymbol{\psi}}^{\dagger}\boldsymbol{L}\hat{\boldsymbol{\psi}}, \tag{4.7}$$

where the conjugate transpose is denoted with a dagger. Since $\boldsymbol{L}$ is real and symmetric, it can be diagonalized through a unitary matrix $\boldsymbol{S}$,

$$\boldsymbol{L} = \boldsymbol{S}^{\dagger}\boldsymbol{D}\boldsymbol{S}, \tag{4.8}$$

where $\boldsymbol{D}$ is diagonal and consists of the real eigenvalues of $\boldsymbol{L}$, $\boldsymbol{D}_{ij} = \lambda_j$. The energy can now be written as

$$E_{k,l} = -\frac{1}{2}(\boldsymbol{S}\hat{\boldsymbol{\psi}})^{\dagger}\boldsymbol{D}(\boldsymbol{S}\hat{\boldsymbol{\psi}}) = -\frac{1}{2}\sum_j \lambda_j |(\boldsymbol{S}\hat{\boldsymbol{\psi}})_j|^2. \tag{4.9}$$

This defines the modes $(\boldsymbol{S}\hat{\boldsymbol{\psi}})_j$ that are orthogonal with respect to the energy norm, i.e. the energy can be partitioned into contributions $E^j_{k,l}$ from these modes. The structure of the modes depends on wavenumber, because $\boldsymbol{L}$ and therefore $\boldsymbol{S}$ does.

For the mixed-layer-only case, with $\boldsymbol{L}$ given by (3.10), the eigenvalues of $\boldsymbol{L}$ are

$$\lambda_{0,1} = fk_h\left(\frac{\coth\mu_m}{N_m} + \frac{1}{2N_t} \pm \sqrt{\frac{\operatorname{csch}^2\mu_m}{N_m^2} + \frac{1}{4N_t^2}}\right) \tag{4.10}$$

and we obtain the eigenvectors as the columns of $\boldsymbol{S}$,

$$\boldsymbol{S} = \begin{pmatrix} \dfrac{1}{\left[1+\left(\cosh\mu_m + \dfrac{N_m\lambda_0}{fk_h}\sinh\mu_m\right)^2\right]^{1/2}} & \dfrac{1}{\left[1+\left(\cosh\mu_m + \dfrac{N_m\lambda_1}{fk_h}\sinh\mu_m\right)^2\right]^{1/2}} \\ \dfrac{\cosh\mu_m + \dfrac{N_m\lambda_0}{fk_h}}{\left[1+\left(\cosh\mu_m + \dfrac{N_m\lambda_0}{fk_h}\sinh\mu_m\right)^2\right]^{1/2}} & \dfrac{\cosh\mu_m + \dfrac{N_m\lambda_1}{fk_h}}{\left[1+\left(\cosh\mu_m + \dfrac{N_m\lambda_1}{fk_h}\sinh\mu_m\right)^2\right]^{1/2}} \end{pmatrix}. \tag{4.11}$$

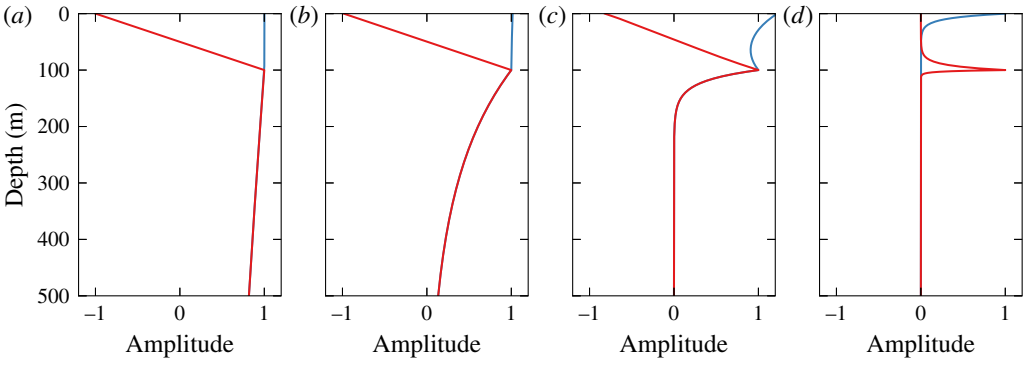

FIGURE 12. Vertical structure of the streamfunction corresponding to orthogonal modes in the mixed-layer-only case for different wavenumbers $k_h = 2\pi/\lambda$, with the wavelength λ given in the panel titles. For (*a–c*), the modes are normalized to unity at the interface at 100 m depth; for panel (*d*), the modes are normalized to have a maximum value of unity. Mode 0 is shown in blue, mode 1 in red. In (*a–c*), the two modes coincide below the interface. (*a*) 1000 km wavelength; (*b*) 100 km wavelength; (*c*) 10 km wavelength; (*d*) 1 km wavelength.

For large scales, $k_h \ll f/N_t h$ or $\mu_m \ll N_m/N_t$, this reduces to

$$\lambda_0 = -\frac{f k_h}{2N_t}, \quad \lambda_1 = -\frac{2f^2}{N_m^2 h}, \tag{4.12a,b}$$

and simply

$$\boldsymbol{S} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{4.13}$$

Equation (4.13) indicates that at large scales, the first mode is barotropic in the mixed layer. It behaves like a surface QG mode penetrating into the thermocline. The streamfunction is proportional to k_h times the conserved quantity (Held *et al.* 1995),

$$(\boldsymbol{S}\hat{\boldsymbol{\psi}})_0 = -\frac{f k_h}{2N_t}(\boldsymbol{S}\hat{\boldsymbol{\theta}})_0. \tag{4.14}$$

The second mode at large scales is baroclinic in the mixed layer. The relation between the streamfunction and the conserved quantity is

$$(\boldsymbol{S}\hat{\boldsymbol{\psi}})_1 = -\frac{2f^2}{N_m^2 h}(\boldsymbol{S}\hat{\boldsymbol{\theta}})_1, \tag{4.15}$$

which is independent of k_h, as expected for a baroclinic mode. These are the same modes as those found in the linear stability analysis for large scales (figure 5*f*).

This description of the orthogonal modes as barotropic and baroclinic mixed-layer modes only applies at large scales. At smaller scales, the modes have a more complicated vertical structure (figure 12). At scales smaller than the mixed-layer deformation radius, they morph into modes that are decoupled and localized in the vertical at the surface and at the interface. But for the cascade dynamics to be discussed, the mode structure at large scales is what is most important.

We can now consider the energy budget of these modes. We start from the vertically integrated spectral energy budget, written in terms of the conserved quantities and

corresponding streamfunctions at the surface and interface:

$$\frac{\partial E_{k,l}}{\partial t} = -\mathrm{Re}\hat{\boldsymbol{\psi}}^{\dagger}\frac{\partial \hat{\boldsymbol{\theta}}}{\partial t}. \tag{4.16}$$

Using the unitary matrix $\boldsymbol{S}$, we can rewrite this as

$$\frac{\partial E_{k,l}}{\partial t} = -\mathrm{Re}\,(\boldsymbol{S}\hat{\boldsymbol{\psi}})^{\dagger}\frac{\partial}{\partial t}(\boldsymbol{S}\hat{\boldsymbol{\theta}}) = -\mathrm{Re}\sum_j (\boldsymbol{S}\hat{\boldsymbol{\psi}})^*_j \frac{\partial}{\partial t}(\boldsymbol{S}\hat{\boldsymbol{\theta}})_j \tag{4.17}$$

and thus split the energy budget into its modal components:

$$\frac{\partial E^j_{k,l}}{\partial t} = -\mathrm{Re}\,(\boldsymbol{S}\hat{\boldsymbol{\psi}})^*_j \frac{\partial}{\partial t}(\boldsymbol{S}\hat{\boldsymbol{\theta}})_j. \tag{4.18}$$

The terms on the right-hand side of this budget can be obtained by substituting in the spectral form of the evolution equation (4.1). To separate out the advective interactions of the modes with themselves and with each other, we further expand the nonlinear terms in (4.1), using the distributive property of the Jacobian operator, into

$$\mathrm{J}(\boldsymbol{\psi},\boldsymbol{\theta}) = \mathrm{J}(\boldsymbol{\psi}^0,\boldsymbol{\theta}^0) + \mathrm{J}(\boldsymbol{\psi}^0,\boldsymbol{\theta}^1) + \mathrm{J}(\boldsymbol{\psi}^1,\boldsymbol{\theta}^0) + \mathrm{J}(\boldsymbol{\psi}^1,\boldsymbol{\theta}^1). \tag{4.19}$$

Here we split the vectors holding the streamfunction and the conserved quantities at the surface and interface into their modal components,

$$\hat{\boldsymbol{\psi}}^j = \boldsymbol{S}^{\dagger}\boldsymbol{P}_j\boldsymbol{S}\hat{\boldsymbol{\psi}}, \quad \hat{\boldsymbol{\theta}}^j = \boldsymbol{S}^{\dagger}\boldsymbol{P}_j\boldsymbol{S}\hat{\boldsymbol{\theta}}, \quad \boldsymbol{\psi} = \sum_j \boldsymbol{\psi}^j, \quad \boldsymbol{\theta} = \sum_j \boldsymbol{\theta}^j, \tag{4.20a–d}$$

where $\boldsymbol{P}_j$ are the projections onto the respective modes,

$$\boldsymbol{P}_0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \boldsymbol{P}_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \tag{4.21a,b}$$

The first term in (4.19) represents the advection of the barotropic mode by the barotropic mode, to use the naming convention introduced above. The second term represents the advection of the baroclinic mode by the barotropic mode, and so on.

In terms of the orthogonal modes, the energy budget is very similar to that of a baroclinic two-layer system (figure 13, cf. Larichev & Held 1995). The extraction of potential energy from the mean flow is concentrated at the scale of the largest, most energetic, eddies and creates mostly baroclinic energy (figure 13*b*). The dominant sink is by hypodiffusion, which also acts on the largest, most energetic, eddies. Barotropic energy dominates at these scales, so hypodiffusion takes out mostly barotropic energy (figure 13*a*). The transfer of energy from the baroclinic mode to the barotropic mode occurs through a dual cascade in the submesoscale range. The baroclinic and barotropic energy components are cascaded in opposite directions so as to yield a vanishing spectral transfer of total energy. Baroclinic energy is transferred down to the instability scale, achieved by the advection of the baroclinic mode by the barotropic mode (figure 13*b*). The energy deposited around the instability scale is transferred to the barotropic mode by interactions between the two modes, which represents baroclinic instability (figure 13*b*). This energy enters the barotropic budget rather less localized in wavenumber space (figure 13*a*). An upscale spectral transfer of barotropic

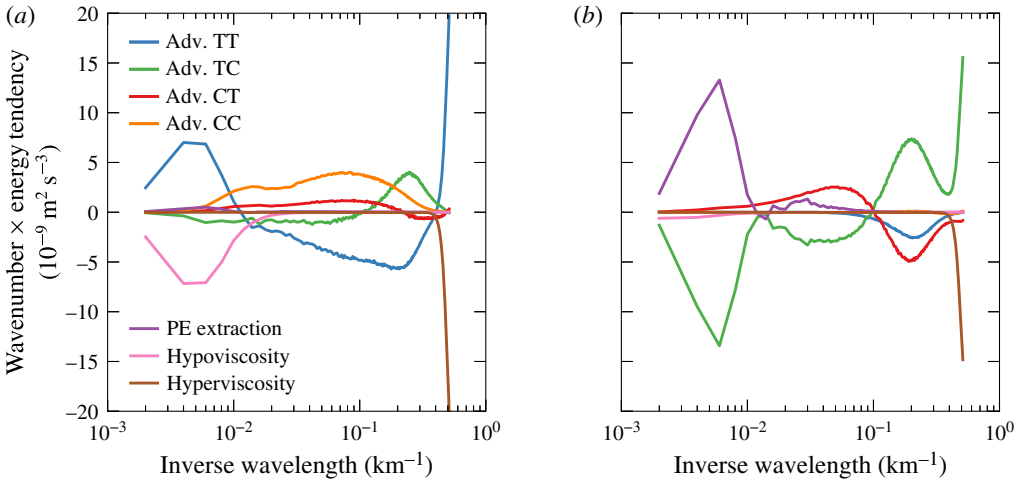

FIGURE 13. Modal energy budget for the mixed-layer-only case. The advective terms correspond to the contributions from the four terms in (4.19). The energy tendencies are multiplied by wavenumber to compensate for logarithmic shrinking. (*a*) Barotropic budget; (*b*) baroclinic budget.

energy closes the budget, taking energy from the instability scale to the scale of the largest, most energetic, eddies, where hypodiffusion acts (figure 13*a*). Energy loss by hyperdiffusion again enters the budget, but is neglected in this discussion, because it is an artefact of finite resolution.

This model thus exhibits a dual cascade analogous to the classic two-layer system. Baroclinic energy is transferred downscale through advection by the barotropic mode, baroclinic instability converts baroclinic into barotropic energy, and barotropic energy is transferred back upscale in an inverse cascade. The difference is that the barotropic mode at large scales behaves like a surface QG mode, instead of a truly barotropic or two-dimensional mode. The inverse cascade is therefore expected to yield a $\langle E_{k_h} \rangle \sim k_h^{-1}$ surface energy spectrum (Blumen 1978), which we find to emerge if the inertial range is wide enough (not shown). More importantly, the surface-QG-like behaviour implies that in the inverse cascade, energy is transferred to successively larger vertical scales. This provides a pathway for mixed-layer instabilities to energize the thermocline below.

4.3. *Full model*

We now consider the case with both mesoscale and submesoscale instabilities present. This full model allows us to address how mesoscale thermocline instabilities modify the energy cycle induced by submesoscale mixed-layer instabilities. Furthermore, we pursue a one-to-one comparison between surface-QG dynamics and the dynamics modified by mixed-layer instabilities, with a focus on vertical energy distribution and vertical velocities.

The linear growth rate of the mixed-layer instability is much larger than that of the thermocline instability (figure 5*a*), so during the initial transient of the nonlinear simulations the mixed-layer instability grows to finite amplitude first. The evolution in the mixed layer is very similar to that of the mixed-layer-only case: the eddies grow in size until they come into statistical equilibrium with hypoviscosity.

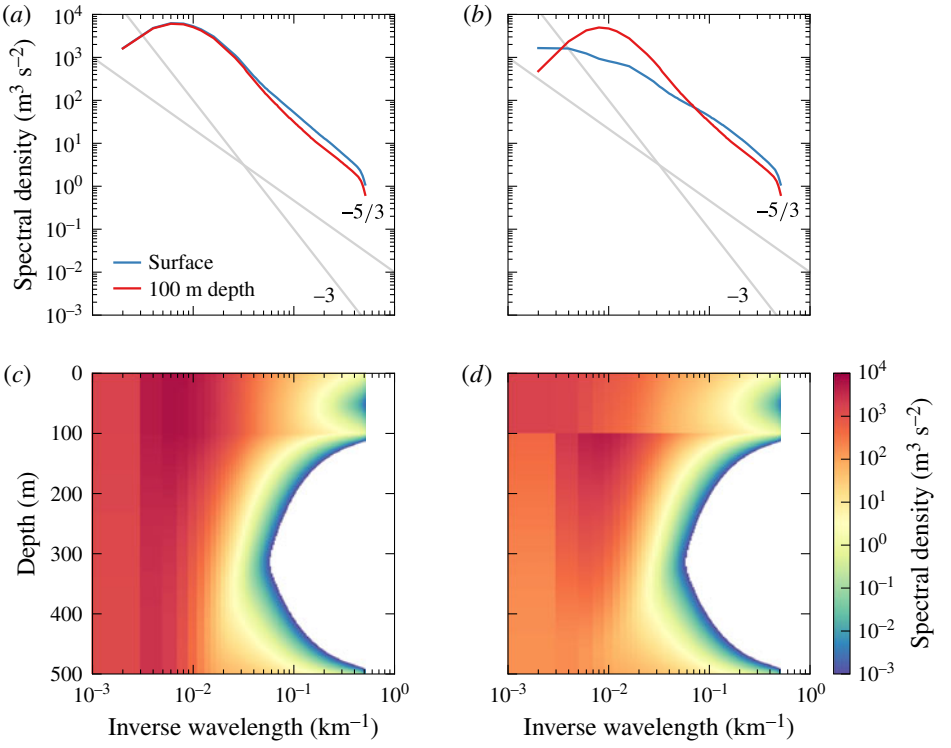

FIGURE 14. Wavenumber spectra of kinetic and potential energy from the full model simulation. (*a*) Kinetic and (*b*) potential energy spectra at the surface and 100 m depth (just below mixed-layer base), spectral density of (*c*) kinetic and (*d*) potential energy in the wavenumber–depth plane. In (*c*,*d*), no values below 10^{-3} m^3 s^{-2} are shown. Reference lines with slopes -3 and $-5/3$ are shown in grey.

The equilibrated state of the full model is also very similar to that of the mixed-layer-only case in the mixed layer and upper thermocline (figure 14). The energy levels and spectra at the surface and the base of the mixed layer are very similar. Near the bottom, a wedge in wavenumber–depth space is energized in the full model, just like in the thermocline-only case. This is due to surface QG turbulence at the bottom level.

The energy budget is similar to the mixed-layer-only case (figure 15). The main energy pathway is again extraction of potential energy in the mixed layer, transfer to the mixed-layer instability scale, conversion to kinetic energy, transfer back to large scales and into the thermocline and dissipation by hypoviscosity. There is additional energy extraction in the thermocline, but that is weak compared to the extraction in the mixed layer. The dominant dynamics is therefore that described for the mixed-layer-only case. Interaction with the bottom level is possible, but of secondary importance in the parameter regime of relevance.

A different picture emerges when the horizontal buoyancy gradient and the associated geostrophic shear in the mixed layer is (unrealistically) reduced. We choose the mixed-layer shear such that the growth rates of the two instabilities are comparable, which from Eady scaling occurs if $\Lambda_m/N_m = \Lambda_t/N_t$ and we set $\Lambda_m = 2.5 \times 10^{-5}$ s^{-1}. The horizontal scales of the instabilities and the overall structure of the dispersion curves are the same as those in figure 5(*a*,*d*).

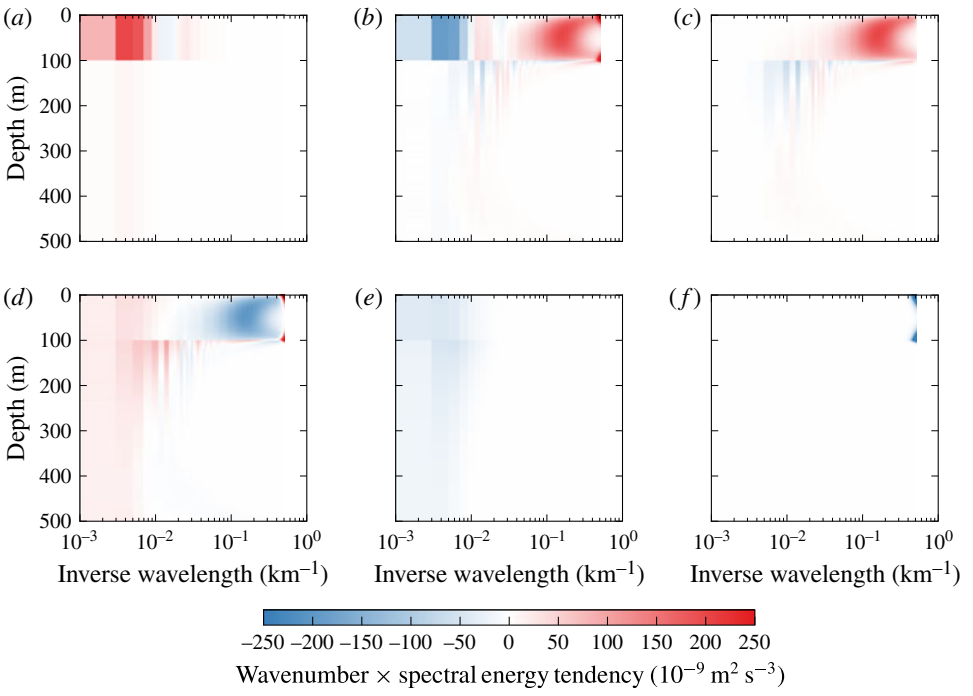

FIGURE 15. Spectral energy budget for the full model simulation. The terms are (*a*) potential energy extraction from the mean, (*b*) spectral potential energy flux divergence, (*c*) potential to kinetic energy conversion, (*d*) kinetic energy flux divergence, including spectral flux and pressure flux, (*e*) hypoviscosity on both kinetic and potential energy, and (*f*) hyperviscosity on both kinetic and potential energy. All terms are multiplied by the wavenumber to compensate for logarithmic shrinking.

This system with reduced mixed-layer shear equilibrates to much lower energy levels than the constant-shear case. The energy levels are comparable to the thermocline-only case and thus allow a one-to-one comparison of the dynamics with and without a mixed layer. The vertical structure of energy in this case with a mixed layer is quite different from the thermocline-only case, because the energy pathway enabled by the mixed-layer instability is still present – mixed-layer instabilities are, on average, not suppressed by the mesoscale strain field (cf. Bishop 1993*a*,*b*; Spall 1997; McWilliams, Molemaker & Olafsdottir 2009). The mixed-layer instability, while not significantly increasing the mesoscale energy levels, does energize the mixed layer at the submesoscales.

This difference between mixed-layer dynamics and surface QG turbulence is also reflected in the vertical velocities that are produced by the instabilities (figure 16). While the available potential energies are the same and the resulting surface energy levels comparable between this reduced mixed-layer shear case and the thermocline-only case, there are much larger vertical velocities in the presence of a mixed layer. These enhanced vertical velocities extend significantly below the base of the mixed layer. The largest vertical velocities are located near fronts in the filamentary sea (figure 17). Coherent vortices, while associated with large buoyancy gradients, induce relatively weak vertical motion. The large vertical velocities appear instead to be associated with the filamentary structure generated by mixed-layer instabilities.

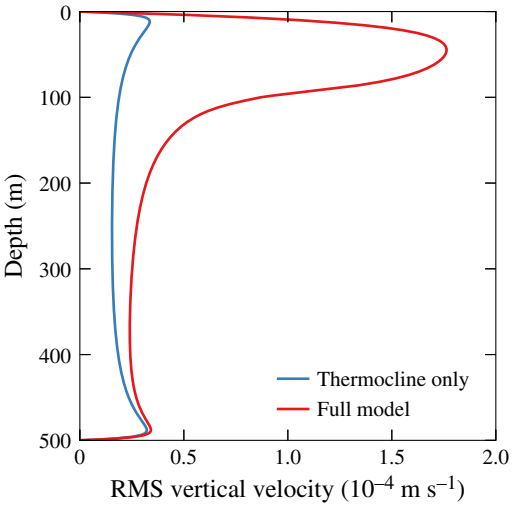

FIGURE 16. Profiles of root-mean-square vertical velocity for the thermocline-only simulation and the full model simulation with reduced mixed-layer shear.

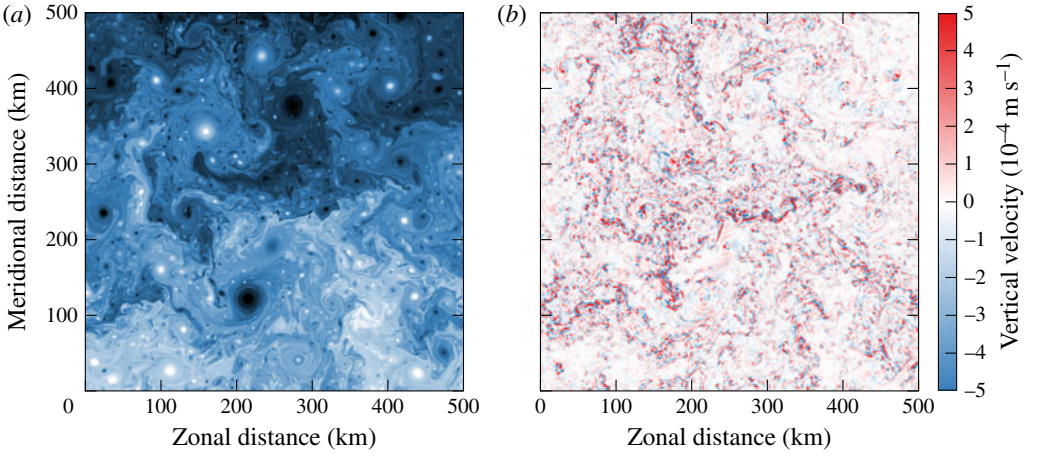

FIGURE 17. Concurrent surface buoyancy and vertical velocity snapshots from the full model simulation with reduced mixed-layer shear: (*a*) surface buoyancy equivalent to figure 7 and (*b*) vertical velocity at 47 m, the depth of the maximum root mean square vertical velocity.

The enhancement of vertical velocities in the presence of mixed-layer instabilities can be understood by considering the omega equation (Hoskins *et al.* 1978):

$$N^2\nabla^2 w + f^2\frac{\partial^2 w}{\partial z^2} = -2\nabla\cdot\boldsymbol{Q}, \tag{4.22}$$

where

$$\boldsymbol{Q} = \left(\frac{\partial \boldsymbol{u}}{\partial x}\cdot\nabla b, \frac{\partial \boldsymbol{u}}{\partial y}\cdot\nabla b\right). \tag{4.23}$$

The vertical velocity can be written as a convolution of the Green's function of (4.22) with the forcing term on the right of (4.22). While the forcing term is not changing

much between the cases with and without the mixed layer, the Green's functions do. The reduced stratification in the mixed layer enhances the response to the forcing term there (Thomas, Tandon & Mahadevan 2008). The enhancement of vertical velocities in our simulation with a mixed layer can thus be attributed to the reduced stratification. Such a reduction of stratification is always associated with a mixed-layer instability, however, so the enhanced vertical velocities are inextricably linked to mixed-layer instabilities.

The forcing term on the right of (4.22) is of the same order in our two cases, because by design the mean states have the same available potential energy and similar submesoscale energy levels are produced. It should be kept in mind, however, that the forcing term does likely increase in the real ocean when mixed layers become deep and mixed-layer instabilities energize the submesoscale range. The more energetic submesoscale turbulence in the wintertime mixed layer is expected to be associated with stronger submesoscale strains than are present in summer (Callies *et al.* 2015). In the wintertime mixed layer, vertical velocities are then enhanced by both a decreased stratification and an increased forcing term on the right of (4.22). Mixed-layer instabilities in the real ocean thus most likely drive an even more dramatic increase in vertical velocities in winter than is present in our simulations with and without mixed layer.

The root-mean-square vertical velocities (figure 16) are similar in structure to those found in primitive equation models (Capet *et al.* 2008*b*). A careful comparison is necessary to establish whether the QG dynamics described here reproduces the magnitude of the vertical velocities or whether non-QG effects significantly enhance or reduce them (cf. Mahadevan & Tandon 2006). Such a comparison is beyond the scope of this article.

The interface between mixed layer and thermocline is located at $z = -h + \eta$, where the interface displacement η is determined by requiring the total buoyancy field to be continuous at the interface,

$$N_m^2 \eta + b(-h^+) = N_t^2 \eta + b(-h^-), \tag{4.24}$$

where the total buoyancy was linearized around $z = -h$, consistent with QG scaling. The interface is material in the sense that

$$\frac{\partial \eta}{\partial t} + \mathrm{J}(\psi, \eta) = w. \tag{4.25}$$

To leading order, there is therefore no exchange of fluid between mixed layer and thermocline. If non-QG effects are taken into account, however, thermocline waters can be folded into the mixed layer (Garner, Nakamura & Held 1992), where atmospherically forced small-scale turbulence can transform them into mixed-layer waters. It thus seems likely that mixed-layer instabilities enhance the exchange between mixed layer and thermocline, but future work will have to investigate what sets the rate of exchange.

5. Comparison to observations

A prominent feature of observed (Callies *et al.* 2015) and modelled (Mensa *et al.* 2013; Sasaki *et al.* 2014) submesoscale flows is their seasonal modulation: they are much more energetic in winter than in summer. This is consistent with an energization of the submesoscale by baroclinic mixed-layer instabilities. Given that these instabilities grow on time scales of order 1 day, they can quickly release large amounts of available potential energy stored in lateral buoyancy gradients in deep winter mixed layers and energize the submesoscale range. Mixed-layer instabilities are

instead weak in summer, when mixed layers are shallow and little potential energy is available for release. Mesoscale-driven surface frontogenesis, on the other hand, is not expected to drive a seasonal cycle in submesoscale turbulence, because the mesoscale eddies that generate submesoscale filaments through frontogenesis are about as strong in winter as they are in summer (or even slightly stronger in summer, Qiu (1999), Qiu & Chen (2004) and Sasaki *et al.* (2014)).

Submesoscale flows observed in the wintertime Gulf Stream region (Callies *et al.* 2015) are energetic throughout the deep mixed layer and decay rapidly below. The energy spectra roll off roughly like k_h^{-2} in the mixed layer and transition to roughly k_h^{-3} below. This spectral and vertical distribution of energy resembles that produced by our simple model of baroclinic mixed-layer instabilities, which similarly has spectra that roll off roughly like k_h^{-2} in the mixed layer and more steeply in the thermocline (figure 14). The equilibrated turbulent flow produced by baroclinic mixed-layer instabilities in this model is thus qualitatively consistent with the energetic wintertime submesoscale flows observed in the Gulf Stream region. In summer, on the other hand, submesoscale flows in the Gulf Stream region are weak. The spectral roll off is a rapid k_h^{-3}, even close to the surface. This lack of submesoscale energy is expected from the lack of energy input from baroclinic mixed-layer instabilities.

There is so far no observational evidence for submesoscale flows that are governed by mesoscale-driven surface frontogenesis as described by surface QG turbulence. The spectral slope of surface kinetic energy may be consistent with the predictions of surface QG theory (e.g. Le Traon *et al.* 2008), but the subsurface structure is not. The observations from the Gulf Stream region (Callies *et al.* 2015) show that the spectral and vertical distribution of submesoscale energy is different from that produced by surface QG turbulence (figure 8), in both summer and winter. In winter, submesoscale flows are observed to be energetic throughout the mixed layer – not just in a thin surface layer as predicted by surface QG turbulence. In summer, when baroclinic mixed-layer instabilities are not active and mesoscale-driven surface frontogenesis could dominate, submesoscale flows are observed to be weak – there is no surface-trapped enhancement as predicted by surface QG turbulence. This suggests that these weak summertime submesoscale flows are instead dominated by deep modes generated by mesoscale thermocline instabilities. Observations from Drake Passage also show no signature of mesoscale-driven surface frontogenesis (Rocha *et al.* 2015). Whether it dominates the energization of submesoscale flows elsewhere remains an open question. It may be expected to dominate outside the major current systems, where thermocline instabilities depend on surface buoyancy gradients and can more effectively drive a surface QG cascade (Charney 1947; Tulloch *et al.* 2011; Roullet *et al.* 2012), but observations are lacking. Balanced submesoscale flows can also be masked by internal waves, especially where mesoscale eddies – and consequently any balanced submesoscale flows – are weak (Richman *et al.* 2012; Callies & Ferrari 2013; Bühler, Callies & Ferrari 2014; Rocha *et al.* 2015).

Spectra that are observed to fall off like k_h^{-3} in the seasonal thermocline in summer and in the permanent thermocline throughout the year (Callies *et al.* 2015) cannot be reproduced by the model formulated in this article. These steep spectra are likely the result of deep Phillips-type instabilities (Phillips 1954) and the potential enstrophy cascade of interior quasi-geostrophic turbulence (Charney 1971). Such a cascade is not present in our model, which collapses the interior PV gradients into delta sheets at the interface between mixed layer and thermocline and at the rigid bottom at the base of the thermocline. Our model does capture, however, the steepening of energy spectra below the winter mixed layer – it only lacks the appropriate thermocline dynamics to yield k_h^{-3} spectra.

The cascade dynamics of our simple QG representation of baroclinic mixed-layer instabilities shows that the mesoscale can effectively be energized by baroclinic mixed-layer instabilities. This would suggest that not only submesoscale but also mesoscale eddies are more energetic in winter than in summer. Altimetry observations (Qiu 1999; Qiu & Chen 2004) and realistic model simulations (Sasaki *et al.* 2014), however, suggest that there is not a large modulation of mesoscale energy levels – mesoscale eddies are even slightly stronger in summer than in winter. This lack of wintertime energization of the mesoscale may result from a coupling with unbalanced motions that is not captured by QG dynamics or from an equilibration process that is longer than a season.

The QG model makes the baroclinic mixed-layer instabilities very effective in energizing the entire turbulent flow, because there is no forward energy cascade resulting in small-scale energy dissipation. If non-QG effects were allowed, a fraction of the energy extracted from the mean in the mixed layer would be dissipated at small scales (Capet *et al.* 2008*d*). This energy leak to small scales is likely as Rossby and Froude numbers become order 1 at scales of order 1 km. The possibility of an energy leak to small scales was demonstrated by Molemaker *et al.* (2010), who studied an Eady instability with $Ro = Fr = 0.5$, using the full Boussinesq equations. While much of the energy extracted from the mean is still trapped at large scales, as predicted by QG dynamics, some is lost to dissipation at small scales. A small leak of energy in the instability may make a big difference in the cascade dynamics, because that energy is not transferred back to mesoscales, where it would further enhance the extraction of potential energy from the mean. Such an effect could be parameterized in our QG model, but is beyond the scope of this study.

An additional energy sink occurs if the balanced flow interacts with ageostrophic instabilities in the mixed layer. In the presence of geostrophic shear, convective motions forced by the atmosphere are the result of symmetric rather than gravitational instabilities and are slantwise rather than upright (e.g. Emanuel 1994; Haine & Marshall 1998; Thomas & Lee 2005). Symmetric instabilities can extract kinetic energy from the geostrophic shear, so they can drain energy from the balanced flow and increase dissipation (Taylor & Ferrari 2010; Thomas *et al.* 2013). This is another way to render the inverse cascade less effective.

Additional sinks for energy in balanced flows in the mixed layer can be the interaction with externally forced near-inertial waves (e.g. Whitt & Thomas 2015; Xie & Vanneste 2015) and the interaction with surface gravity waves (McWilliams & Fox-Kemper 2013; Hamlington *et al.* 2014). Both may drain enough energy out of the balanced flow to prevent an effective inverse cascade of submesoscale kinetic energy to mesoscales.

QG dynamics further do not allow for a feedback of eddies on the mean stratification of the mixed layer (e.g. Fox-Kemper *et al.* 2008), so that no restratification can occur. In the real ocean, the restratification through baroclinic mixed-layer instabilities is opposed by convective and mechanical mixing that is driven by atmospheric forcing. The fixed mean stratification in the QG model assumes that this forced vertical mixing is in balance with restratification. It is unclear whether and how vertical mixing alters the energy budget of the balanced flow in the mixed layer. It is also unclear what the effect of an unsteady atmospheric forcing is, which upsets the balance between vertical mixing and restratification.

Another possible explanation for the lack of a wintertime energization of the mesoscale by baroclinic mixed-layer instabilities is that it takes a few months for the kinetic energy injected by submesoscale mixed-layer instabilities to arrive at the

mesoscale (Sasaki *et al.* 2014). This could also explain the (weak) summer maximum in mesoscale energy (Qiu 1999; Qiu & Chen 2004). This time dependence is not addressed in our simple model of baroclinic mixed-layer instabilities, in which we prescribe perpetual winter conditions. If the time scale of turbulent equilibration is not much shorter than the seasonal time scale, it is likely that submesoscale – and possibly mesoscale – flows are not in statistical equilibrium. It remains to be investigated how a seasonally modulated mixed layer modifies the inverse cascade and the energization of the mesoscale by baroclinic mixed-layer instabilities.

Realistic high-resolution simulations of the type employed by Shcherbina *et al.* (2013), Sasaki *et al.* (2014), Gula, Molemaker & McWilliams (2015) may be able to address these caveats and bridge the gap between observations and the idealized QG dynamics described here. These primitive equation and Boussinesq simulations can provide insights into how non-QG effects modify the submesoscale dynamics induced by mixed-layer instabilities. An exploration of non-QG effects in more idealized set-ups may also prove useful. Primitive equation and Boussinesq simulations can also address how non-QG effects modify mesoscale-driven surface frontogenesis and allow an estimate of the importance of corrections to surface QG turbulence (cf. Hakim *et al.* 2002; Capet *et al.* 2008*c*; Klein *et al.* 2008; Badin 2012; Roullet *et al.* 2012), providing another stepping stone for understanding observations.

6. Conclusions

The simple model formulated in this article sharpens our understanding of how baroclinic mixed-layer instabilities can energize submesoscale turbulence and how this mechanism differs from mesoscale-driven surface frontogenesis. Our analysis suggests that the presence of a mixed layer has a profound effect on submesoscale turbulence. Lateral buoyancy gradients, combined with the low stratification in the mixed layer, provide a large amount of available potential energy that can be extracted by baroclinic instabilities in the mixed layer. The extraction of available potential energy from the large-scale mean is dominated by mesoscale eddies, but potential energy is subsequently transferred downscale to the deformation radius of the mixed layer, where baroclinic instability converts it into kinetic energy. In the QG dynamics considered here, the energy lost to small scales is negligible. The entire energy extracted from the mean in the mixed layer is converted to kinetic energy around the deformation radius of the mixed layer and subsequently transferred back to larger scales in an inverse cascade that also energizes the thermocline below. Through this process, baroclinic mixed-layer instabilities can energize the submesoscale range and even the mesoscale eddy field.

This turbulent dynamics follows a dual cascade similar to that present in two-layer QG flow (Rhines 1977; Salmon 1978; Larichev & Held 1995) and in idealized continuously stratified QG flows (Smith & Vallis 2002) – but with mixed-layer modes. The energy in the baroclinic mode, which is baroclinic in the mixed layer and does not reach much into the thermocline, is transferred downscale through advection by the barotropic mode, which is barotropic in the mixed layer and decays surface-QG-like in the thermocline. Around the deformation radius of the mixed layer, baroclinic instability transforms baroclinic into barotropic energy, which then enters an inverse cascade.

The dynamics resulting from baroclinic mixed-layer instabilities substantially differs from mesoscale-driven surface frontogenesis, as described by surface QG turbulence forced by a mesoscale eddy field, which is often invoked to explain

energetic submesoscales. Surface QG turbulence can only energize a thin surface layer. Mixed-layer instabilities, instead, energize the entire depth of the mixed layer. Vertical velocities are drastically enhanced in the presence of baroclinic mixed-layer instabilities compared to surface QG flows of similar energy levels.

The enhancement of submesoscale energy throughout the mixed layer and the decay below its base, as generated by baroclinic mixed-layer instabilities, are consistent with wintertime observations from the Gulf Stream region (Callies *et al.* 2015). These observations, as well as models (Mensa *et al.* 2013; Sasaki *et al.* 2014), also show that submesoscale flows are most energetic in winter, when baroclinic mixed-layer instabilities are active. This evidence points to the importance of baroclinic mixed-layer instabilities in energizing the submesoscale.

It remains to be investigated how deviations from QG dynamics affect submesoscale flows when a mixed layer is present. The restratification by mixed-layer instabilities, the formation of buoyancy discontinuities, ageostrophic instabilities and forced mixed-layer turbulence all have the potential to modify the leading-order balanced dynamics described here. For example, it remains unclear how much of the submesoscale kinetic energy generated by baroclinic mixed-layer instabilities is cascaded to mesoscales, how vertical velocities are modified by non-QG effects and how the enhancement of vertical velocities by mixed-layer instabilities translates into an exchange of fluid between mixed layer and thermocline.

It is hoped that the model and dynamics discussed here in an oceanographic context are of interest in a broader geophysical fluid dynamics context. Atmospheres often display layers of different stratification, which likely induce similar dynamics. These have partly been explored for the terrestrial atmosphere and it seems likely that similar dynamics also occur in the atmospheres of other planets.

Acknowledgements

Much of this work was conducted at the Geophysical Fluid Dynamics Summer School at the Woods Hole Oceanographic Institution. We gratefully acknowledge the NSF's support of this program and the fruitful interactions we had with staff and visitors.

Appendix A

A QG system of n layers of constant PV, of thickness h_j and stratification N_j, consists of $n+1$ conserved quantities that are advected by the geostrophic flow at their respective levels. Compared to the two-layer model considered in the main text, additional interface quantities analogous to θ_1 are present. The linear operator in the inversion relation (2.9*a*) has tridiagonal structure:

$$\boldsymbol{L} = fk_h \begin{pmatrix} -\dfrac{\coth\mu_0}{N_0} & \dfrac{\operatorname{csch}\mu_0}{N_0} & & & \\ \dfrac{\operatorname{csch}\mu_0}{N_0} & -\dfrac{\coth\mu_0}{N_0} - \dfrac{\coth\mu_1}{N_1} & \dfrac{\operatorname{csch}\mu_1}{N_1} & & \\ & \ddots & \ddots & \ddots & \\ & & \dfrac{\operatorname{csch}\mu_{n-1}}{N_{n-1}} & -\dfrac{\coth\mu_{n-1}}{N_{n-1}} - \dfrac{\coth\mu_n}{N_n} & \dfrac{\operatorname{csch}\mu_n}{N_n} \\ & & & \dfrac{\operatorname{csch}\mu_n}{N_n} & -\dfrac{\coth\mu_n}{N_n} \end{pmatrix}, \tag{A 1}$$

where $\mu_j = N_j k_h h_j / f$. It may be more efficient to solve the inversion relation numerically instead of calculating the inverse of this matrix, which will generally be full.

One can also include a PV gradient due to differential rotation. This can be done using a trick described by Lindzen (1994): instead of using linear shear and constant stratification in the layers, one can use parabolic shear or a modified stratification profile, which allows cancellation of the contribution from the β-effect and retaining constant PV within the layers. The PV gradient due to β is then included in the PV sheets at the interfaces.

Appendix B

If there is a buoyancy jump g' at the interface, the matching conditions must be modified. To ensure a continuous pressure at the interface at $z = -h + \eta$, we require

$$\psi(-h^+) - \psi(-h^-) = -\frac{g'}{f}\eta. \tag{B 1}$$

Here, η is the perturbation of the interface between the constant-PV layers. The buoyancy equations (2.2) and (2.3) can be combined with the kinematic condition

$$w = \frac{\partial \eta}{\partial t} + \mathrm{J}(\psi, \eta), \tag{B 2}$$

applied at $z = -h^+$ and $z = -h^-$, to give

$$\frac{\partial \theta_1}{\partial t} + \mathrm{J}(\psi_1, \theta_1) = 0, \quad \frac{\partial \theta_2}{\partial t} + \mathrm{J}(\psi_2, \theta_2) = 0, \tag{B 3a,b}$$

where

$$\theta_1 = f\frac{b(-h^+)}{N_m^2} + f\eta, \quad \theta_2 = f\frac{b(-h^-)}{N_t^2} + f\eta, \quad \psi_1 = \psi(-h^+), \quad \psi_2 = \psi(-h^-). \tag{B 4a–d}$$

Together with the conservation of surface and bottom buoyancy,

$$\theta_0 = -f\frac{b(0)}{N_m^2}, \quad \theta_3 = f\frac{b(-H)}{N_t^2}, \tag{B 5a,b}$$

and the inversion relation obtained by solving (2.10) with the matching conditions above, the model is complete. It now consists of four conserved quantities.

REFERENCES

BADIN, G. 2012 Surface semi-geostrophic dynamics in the ocean. *Geophys. Astrophys. Fluid Dyn.* **107** (5), 526–540.

BISHOP, C. H. 1993*a* On the behaviour of baroclinic waves undergoing horizontal deformation. I: The 'RT' phase diagram. *Q. J. R. Meteorol. Soc.* **119** (510), 221–240.

BISHOP, C. H. 1993*b* On the behaviour of baroclinic waves undergoing horizontal deformation. II: Error-bound amplification and Rossby wave diagnostics. *Q. J. R. Meteorol. Soc.* **119** (510), 241–267.

BLUMEN, W. 1978 Uniform potential vorticity flow: Part I. Theory of wave interactions and two-dimensional turbulence. *J. Atmos. Sci.* **35** (5), 774–783.

BLUMEN, W. 1979 On short-wave baroclinic instability. *J. Atmos. Sci.* **36** (10), 1925–1933.

BOCCALETTI, G., FERRARI, R. & FOX-KEMPER, B. 2007 Mixed layer instabilities and restratification. *J. Phys. Oceanogr.* **37** (9), 2228–2250.

BOYD, J. P. 1992 The energy spectrum of fronts: time evolution of shocks in Burgers' equation. *J. Atmos. Sci.* **49** (2), 128–139.

BRETHERTON, F. P. 1966 Critical layer instability in baroclinic flows. *Q. J. R. Meteorol. Soc.* **92** (393), 325–334.

BÜHLER, O., CALLIES, J. & FERRARI, R. 2014 Wave–vortex decomposition of one-dimensional ship-track data. *J. Fluid Mech.* **756**, 1007–1026.

CALLIES, J. & FERRARI, R. 2013 Interpreting energy and tracer spectra of upper-ocean turbulence in the submesoscale range (1–200 km). *J. Phys. Oceanogr.* **43** (11), 2456–2474.

CALLIES, J., FERRARI, R., KLYMAK, J. M. & GULA, J. 2015 Seasonality in submesoscale turbulence. *Nat. Commun.* **6**, 6862.

CAPET, X., KLEIN, P., HUA, B. L., LAPEYRE, G. & MCWILLIAMS, J. C. 2008*a* Surface kinetic energy transfer in surface quasi-geostrophic flows. *J. Fluid Mech.* **604**, 165–174.

CAPET, X., MCWILLIAMS, J. C., MOLEMAKER, M. J. & SHCHEPETKIN, A. F. 2008*b* Mesoscale to submesoscale transition in the California current system. Part I: flow structure, eddy flux, and observational tests. *J. Phys. Oceanogr.* **38** (1), 29–43.

CAPET, X., MCWILLIAMS, J. C., MOLEMAKER, M. J. & SHCHEPETKIN, A. F. 2008*c* Mesoscale to submesoscale transition in the California current system. Part II: frontal processes. *J. Phys. Oceanogr.* **38** (1), 44–64.

CAPET, X., MCWILLIAMS, J. C., MOLEMAKER, M. J. & SHCHEPETKIN, A. F. 2008*d* Mesoscale to submesoscale transition in the California current system. Part III: energy balance and flux. *J. Phys. Oceanogr.* **38** (10), 2256–2269.

CHARNEY, J. G. 1947 The dynamics of long waves in a baroclinic westerly current. *J. Meteorol.* **4** (5), 135–162.

CHARNEY, J. G. 1971 Geostrophic turbulence. *J. Atmos. Sci.* **28** (6), 1087–1095.

EADY, E. T. 1949 Long waves and cyclone waves. *Tellus* **1** (3), 33–52.

EMANUEL, K. A. 1994 *Atmospheric Convection*. Oxford University Press.

FERRARI, R. 2011 A frontal challenge for climate models. *Science* **332** (6027), 316–317.

FERRARI, R. & RUDNICK, D. L. 2000 Thermohaline variability in the upper ocean. *J. Geophys. Res.* **105** (C7), 16857–16883.

FOX-KEMPER, B., FERRARI, R. & HALLBERG, R. W. 2008 Parameterization of mixed layer eddies. Part I: theory and diagnosis. *J. Phys. Oceanogr.* **38** (6), 1145–1165.

GARNER, S. T., NAKAMURA, N. & HELD, I. M. 1992 Nonlinear equilibration of two-dimensional Eady waves: a new perspective. *J. Atmos. Sci.* **49** (21), 1984–1996.

GILL, A. E., GREEN, J. S. A. & SIMMONS, A. J. 1974 Energy partition in the large-scale ocean circulation and the production of mid-ocean eddies. *Deep-Sea Res.* **21** (7), 499–528.

GULA, J., MOLEMAKER, M. J. & MCWILLIAMS, J. C. 2015 Gulf stream dynamics along the southeastern U.S. seaboard. *J. Phys. Oceanogr.* **45** (3), 690–715.

HAINE, T. W. N. & MARSHALL, J. 1998 Gravitational, symmetric, and baroclinic instability of the ocean mixed layer. *J. Phys. Oceanogr.* **28** (4), 634–658.

HAKIM, G. J., SNYDER, C. & MURAKI, D. J. 2002 A new surface model for cyclone–anticyclone asymmetry. *J. Atmos. Sci.* **59** (16), 2405–2420.

HAMLINGTON, P. E., VAN ROEKEL, L. P., FOX-KEMPER, B., JULIEN, K. & CHINI, G. P. 2014 Langmuir–submesoscale interactions: descriptive analysis of multiscale frontal spindown simulations. *J. Phys. Oceanogr.* **44** (9), 2249–2272.

HELD, I. M., PIERREHUMBERT, R. T., GARNER, S. T. & SWANSON, K. L. 1995 Surface quasi-geostrophic dynamics. *J. Fluid Mech.* **282**, 1–20.

HOSKINS, B. J. & BRETHERTON, F. P. 1972 Atmospheric frontogenesis models: mathematical formulation and solution. *J. Atmos. Sci.* **29** (1), 11–37.

HOSKINS, B. J., DRAGHICI, I. & DAVIES, H. C. 1978 A new look at the ω-equation. *Q. J. R. Meteorol. Soc.* **104** (439), 31–38.

JUCKES, M. 1994 Quasigeostrophic dynamics of the tropopause. *J. Atmos. Sci.* **51** (19), 2756–2768.

KLEIN, P., HUA, B. L., LAPEYRE, G., CAPET, X., LE GENTIL, S. & SASAKI, H. 2008 Upper ocean turbulence from high-resolution 3D simulations. *J. Phys. Oceanogr.* **38** (8), 1748–1763.

KLEIN, P. & LAPEYRE, G. 2009 The oceanic vertical pump induced by mesoscale and submesoscale turbulence. *Annu. Rev. Mater. Sci.* **1**, 351–375.

KOLMOGOROV, A. N. 1941 The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. *Dokl. Akad. Nauk SSSR* **30**, 301–305; (in Russian).

KRAICHNAN, R. H. 1967 Inertial ranges in two-dimensional turbulence. *Phys. Fluids* **10** (7), 1417–1423.

LAPEYRE, G. & KLEIN, P. 2006 Dynamics of the upper oceanic layers in terms of surface quasigeostrophy theory. *J. Phys. Oceanogr.* **36** (2), 165–176.

LAPEYRE, G., KLEIN, P. & HUA, B. L. 2006 Oceanic restratification forced by surface frontogenesis. *J. Phys. Oceanogr.* **36** (8), 1577–1590.

LARICHEV, V. D. & HELD, I. M. 1995 Eddy amplitudes and fluxes in a homogeneous model of fully developed baroclinic instability. *J. Phys. Oceanogr.* **25** (10), 2285–2297.

LE TRAON, P.-Y., KLEIN, P., HUA, B. L. & DIBARBOURE, G. 2008 Do altimeter wavenumber spectra agree with the interior or surface quasigeostrophic theory? *J. Phys. Oceanogr.* **38** (5), 1137–1142.

LÉVY, M., IOVINO, D., RESPLANDY, L., KLEIN, P., MADEC, G., TRÉGUIER, A.-M., MASSON, S. & TAKAHASHI, K. 2012 Large-scale impacts of submesoscale dynamics on phytoplankton: local and remote effects. *Ocean Model.* **43–44**, 77–93.

LINDZEN, R. S. 1994 The Eady problem for a basic state with zero PV gradient but $\beta \neq 0$. *J. Atmos. Sci.* **51** (22), 3221–3226.

MAHADEVAN, A. 2014 Eddy effects on biogeochemistry. *Nature* **506**, 168–169.

MAHADEVAN, A. & TANDON, A. 2006 An analysis of mechanisms for submesoscale vertical motion at ocean fronts. *Ocean Model.* **14** (3–4), 241–256.

MCWILLIAMS, J. C. & FOX-KEMPER, B. 2013 Oceanic wave-balanced surface fronts and filaments. *J. Fluid Mech.* **730**, 464–490.

MCWILLIAMS, J. C., MOLEMAKER, M. J. & OLAFSDOTTIR, E. I. 2009 Linear fluctuation growth during frontogenesis. *J. Phys. Oceanogr.* **39** (12), 3111–3129.

MENSA, J. A., GARRAFFO, Z., GRIFFA, A., ÖZGÖKMEN, T. M., HAZA, A. & VENEZIANI, M. 2013 Seasonality of the submesoscale dynamics in the Gulf stream region. *Ocean Dyn.* **63** (8), 923–941.

MOLEMAKER, M. J., MCWILLIAMS, J. C. & CAPET, X. 2010 Balanced and unbalanced routes to dissipation in an equilibrated Eady flow. *J. Fluid Mech.* **654** (2010), 35–63.

PEDLOSKY, J. 1987 *Geophysical Fluid Dynamics*, 2nd edn. Springer.

PHILLIPS, N. A. 1954 Energy transformations and meridional circulations associated with simple baroclinic waves in a two-level, quasi-geostrophic model. *Tellus* **6** (3), 273–286.

PIERREHUMBERT, R. T., HELD, I. M. & SWANSON, K. L. 1994 Spectra of local and nonlocal two-dimensional turbulence. *Chaos, Solitons Fractals* **4** (6), 1111–1116.

QIU, B. 1999 Seasonal eddy field modulation of the north pacific subtropical countercurrent: TOPEX/Poseidon observations and theory. *J. Phys. Oceanogr.* **29** (10), 2471–2486.

QIU, B. & CHEN, S. 2004 Seasonal modulations in the eddy field of the south pacific ocean. *J. Phys. Oceanogr.* **34** (7), 1515–1527.

RHINES, P. B. 1977 The dynamics of unsteady currents. In *Sea* (ed. E. Goldberg), vol. VI, pp. 189–318. Wiley.

RICHMAN, J. G., ARBIC, B. K., SHRIVER, J. F., METZGER, E. J. & WALLCRAFT, A. J. 2012 Inferring dynamics from the wavenumber spectra of an eddying global ocean model with embedded tides. *J. Geophys. Res.* **117** (C12), C12012.

RIVEST, C., DAVIS, C. A. & FARRELL, B. F. 1992 Upper-tropospheric synoptic-scale waves. Part I: maintenance as Eady normal modes. *J. Atmos. Sci.* **49** (22), 2108–2119.

ROCHA, C. B., CHERESKIN, T. K., GILLE, S. T. & MENEMENLIS, D. 2015 Mesoscale to submesoscale wavenumber spectra in Drake passage. *J. Phys. Oceanogr.* (in press).

ROULLET, G., MCWILLIAMS, J. C., CAPET, X. & MOLEMAKER, M. J. 2012 Properties of steady geostrophic turbulence with isopycnal outcropping. *J. Phys. Oceanogr.* **42** (1), 18–38.

SALMON, R. 1978 Two-layer quasi-geostrophic turbulence in a simple special case. *Geophys. Astrophys. Fluid Dyn.* **10** (1), 25–52.

SASAKI, H., KLEIN, P., QIU, B. & SASAI, Y. 2014 Impact of oceanic-scale interactions on the seasonal modulation of ocean dynamics by the atmosphere. *Nat. Commun.* **5**, 5636.

SCOTT, R. K. 2006 Local and nonlocal advection of a passive scalar. *Phys. Fluids* **18** (11).

SEIFF, A., KIRK, D. B., KNIGHT, T. C. D., YOUNG, R. E., MIHALOV, J. D., YOUNG, L. A., MILOS, F. S., SCHUBERT, G., BLANCHARD, R. C. & ATKINSON, D. 1998 Thermal structure of Jupiter's atmosphere near the edge of a 5 μm hot spot in the north equatorial belt. *J. Geophys. Res. Planets* **103** (103), E10.

SHCHERBINA, A. Y., D'ASARO, E. A., LEE, C. M., KLYMAK, J. M., MOLEMAKER, M. J. & MCWILLIAMS, J. C. 2013 Statistics of vertical vorticity, divergence, and strain in a developed submesoscale turbulence field. *Geophys. Res. Lett.* **40** (17), 4706–4711.

SMITH, K. S. & BERNARD, E. 2013 Geostrophic turbulence near rapid changes in stratification. *Phys. Fluids* **25** (4), 046601.

SMITH, K. S. & VALLIS, G. K. 2001 The scales and equilibration of midocean eddies: freely evolving flow. *J. Phys. Oceanogr.* **31** (2), 554–571.

SMITH, K. S. & VALLIS, G. K. 2002 The scales and equilibration of midocean eddies: forced–dissipative flow. *J. Phys. Oceanogr.* **32** (6), 1699–1720.

SPALL, M. A. 1997 Baroclinic jets in confluent flow. *J. Phys. Oceanogr.* **27** (6), 1054–1071.

STONE, P. H. 1966*a* Frontogenesis by horizontal wind deformation fields. *J. Atmos. Sci.* **23** (5), 455–465.

STONE, P. H. 1966*b* On non-geostrophic baroclinic stability. *J. Atmos. Sci.* **23** (4), 390–400.

TAYLOR, J. R. & FERRARI, R. 2010 Buoyancy and wind-driven convection at mixed layer density fronts. *J. Phys. Oceanogr.* **40** (6), 1222–1242.

THOMAS, L. N. & LEE, C. M. 2005 Intensification of ocean fronts by down-front winds. *J. Phys. Oceanogr.* **35** (6), 1086–1102.

THOMAS, L. N., TANDON, A. & MAHADEVAN, A. 2008 Submesoscale processes and dynamics. In *Ocean Modeling in an Eddying Regime*, pp. 17–38. American Geophysical Union.

THOMAS, L. N., TAYLOR, J. R., FERRARI, R. & JOYCE, T. M. 2013 Symmetric instability in the Gulf stream. *Deep-Sea Res. II* **91**, 96–110.

TULLOCH, R., MARSHALL, J., HILL, C. & SMITH, K. S. 2011 Scales, growth rates, and spectral fluxes of baroclinic instability in the ocean. *J. Phys. Oceanogr.* **41** (6), 1057–1076.

VALLIS, G. K. 2006 *Atmospheric and Oceanic Fluid Dynamics*. Cambridge University Press.

WHITT, D. B. & THOMAS, L. N. 2015 Resonant generation and energetics of wind-forced near-inertial motions in a geostrophic flow. *J. Phys. Oceanogr.* **45** (1), 181–208.

XIE, J.-H. & VANNESTE, J. 2015 A generalised-Lagrangian-mean model of the interactions between near-inertial waves and mean flow. *J. Fluid Mech.* **774**, 143–169.