A three-component model for the coupled evolution of near-inertial waves, quasi-geostrophic flow and the near-inertial second harmonic

G. L. Wagner[1],† and W. R. Young[2]

[1]Department of Mechanical and Aerospace Engineering, University of California San Diego,
La Jolla, CA 90293-0411, USA

[2]Scripps Institution of Oceanography, University of California San Diego, La Jolla,
CA 90293-0213, USA

We derive an asymptotic model that describes the nonlinear coupled evolution of (i) near-inertial waves (NIWs), (ii) balanced quasi-geostrophic flow and (iii) near-inertial second harmonic waves with frequency near $2f_0$, where f_0 is the local inertial frequency. This 'three-component' model extends the two-component model derived by Xie & Vanneste (*J. Fluid Mech.*, vol. 774, 2015, pp. 143–169) to include interactions between near-inertial and $2f_0$ waves. Both models possess two conservation laws which together imply that oceanic NIWs forced by winds, tides or flow over bathymetry can extract energy from quasi-geostrophic flows. A second and separate implication of the three-component model is that quasi-geostrophic flow catalyses a loss of NIW energy to freely propagating waves with near-$2f_0$ frequency that propagate rapidly to depth and transfer energy back to the NIW field at very small vertical scales. The upshot of near-$2f_0$ generation is a two-step mechanism whereby quasi-geostrophic flow catalyses a nonlinear transfer of near-inertial energy to the small scales of wave breaking and diapycnal mixing. A comparison of numerical solutions with both Boussinesq and three-component models for a two-dimensional initial value problem reveals strengths and weaknesses of the model while demonstrating the extraction of quasi-geostrophic energy and production of small vertical scales.

Key words: internal waves, ocean processes, quasi-geostrophic flows

---

1. Introduction

Near-inertial waves (NIWs) are inertia–gravity waves in rotating and stratified fluids with frequencies near the local inertial frequency, f_0. In the oceans of Earth, an almost-universal strong density stratification means that NIWs have very small aspect ratios, large vertical shears and the lowest of internal wave frequencies. Partly because of their small aspect ratios and low frequencies, oceanic NIWs are generated by such

† Email address for correspondence: glwagner@ucsd.edu

diverse processes as fluctuating winds and flow over topography, contain roughly half of the total internal wave kinetic energy and are a main contributor to diapycnal mixing (Ferrari & Wunsch 2009).

The weak dispersion and slow propagation of NIWs exposes them to strong interaction with balanced quasi-geostrophic (QG) flows. A basic introduction to NIW propagation through non-uniform balanced flows is given by the ray theories of Mooers (1975) and Kunze (1985), which show that near-inertial energy is attracted to regions of negative balanced vorticity and expelled from regions of positive vorticity. A more general theory valid both for ray-like NIW propagation and scattering by smaller-scale balanced flows was developed by Young & Jelloul (1997, YBJ hereafter). YBJ linearized the Boussinesq equations around a prescribed background flow and exploited weak near-inertial dispersion to develop a two-time asymptotic expansion that isolates the slow evolution of NIWs. The resulting YBJ NIW equation, which is similar to (1.8) below, describes the weakly dispersive propagation of β-plane NIWs though advecting and refracting balanced flows of near-arbitrary spatial structure.

The YBJ NIW equation successfully describes many aspects of near-inertial propagation through realistic balanced flows (Klein & Smith 2001; Klein, Smith & Lapeyre 2004; Danioux, Klein & Rivière 2008), but ignores nonlinear finite-amplitude NIW dynamics and their corresponding feedback onto the balanced flow. In pursuit of a richer theory describing the coupled evolution of NIWs and balanced flows, Xie & Vanneste (2015, XV hereafter) derived a generalized-Lagrangian-mean model which joins the YBJ NIW equation to the QG equations. Like Bühler & McIntyre (1998) and Wagner & Young (2015) and as in (1.7) below, in the XV model an NIW-induced balanced flow takes part in advecting QG potential vorticity and thus in the evolution of QG flow.

1.1. *The $2f_0$ harmonic and motivation for a three-component model*

Both YBJ and XV lack a conspicuous aspect of NIW evolution observed in the kinetic energy frequency spectra of the Ocean Storms Experiment (D'Asaro *et al.* 1995), the observations of Niwa & Hibiya (1999) and in the simulations of NIW–QG interaction by Danioux *et al.* (2008): the nonlinear generation of internal waves with frequency $2f_0$. While these $2f_0$ waves have little horizontal kinetic energy relative to their parent NIWs, they can dominate pressure fields and contribute appreciably to vertical velocity fields and isopycnal displacements. Remarkably, $2f_0$ generation and subsequent horizontal radiation can remove energy from spatially compact regions of NIW–QG interaction, as discussed below and illustrated in figure 1. A primary motivation for this paper is the derivation of a more complete set of equations that contains the essential elements of YBJ and XV while including $2f_0$ waves. This derivation yields a model with three components: NIW velocity, balanced QG potential vorticity and the amplitude of the $2f_0$ pressure.

To motivate the three-component model, we consider an initial value problem in the Boussinesq equations in which a surface-concentrated NIW interacts with a balanced barotropic jet in two dimensions (x, z). We use a constant inertial frequency $f_0 = 10^{-4}$ s^{-1} and buoyancy frequency $N = 2 \times 10^{-3}$ s^{-1} associated with a stable background buoyancy profile. The velocity field is $\boldsymbol{u} = u\hat{\boldsymbol{x}} + v\hat{\boldsymbol{y}} + w\hat{\boldsymbol{z}}$ and the dynamic buoyancy perturbation from the background is b, so that the available potential energy density is $b^2/2N^2$. The initial v is a barotropic jet in geostrophic and hydrostatic balance flowing along the axis of y, while the initial u is a surface-concentrated,
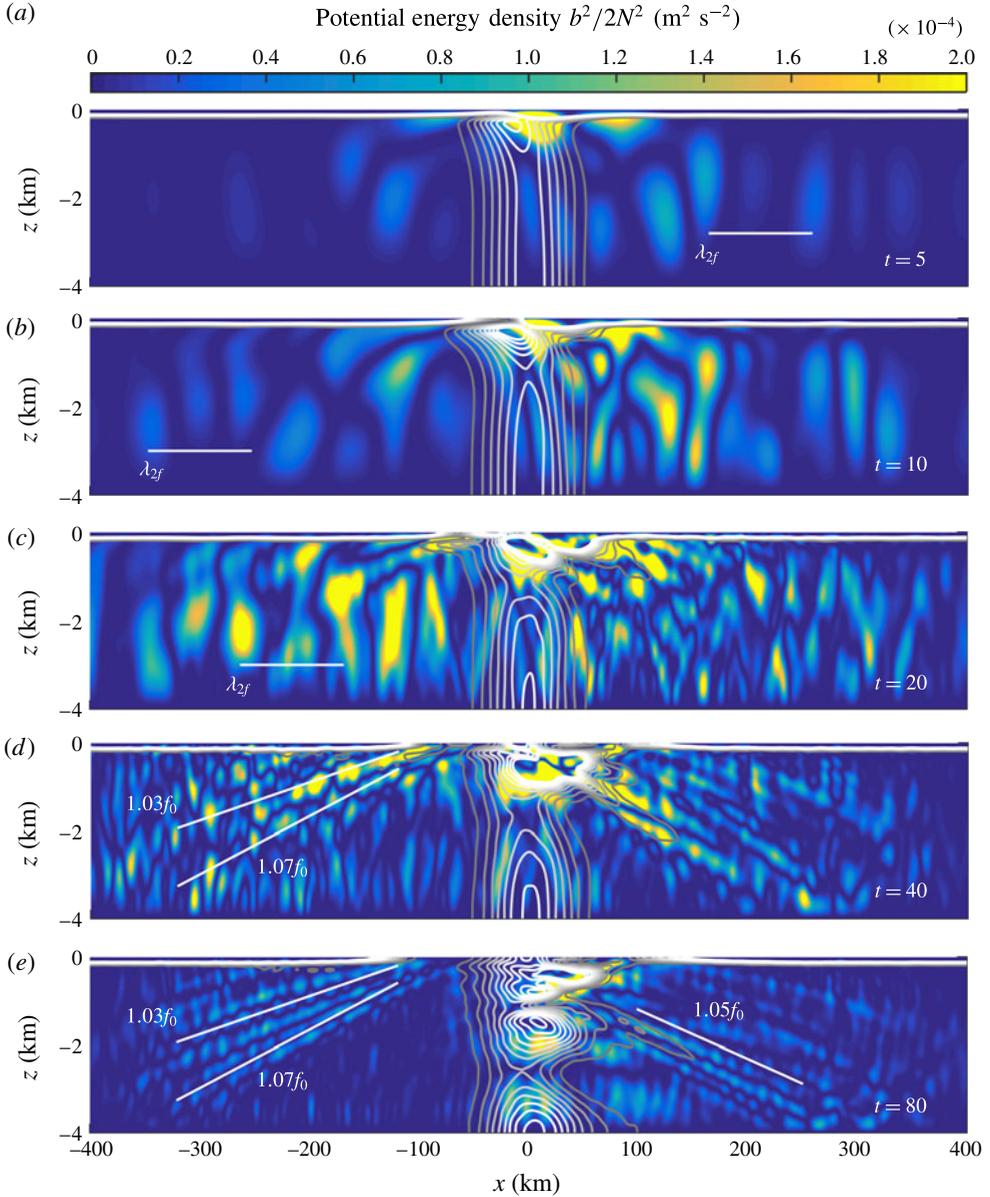
FIGURE 1. (Colour online) The numerical solution of a two-dimensional Boussinesq initial value problem involving the interaction between a barotropic jet and a surface-concentrated NIW. The shading shows the potential energy density $b^2/2N^2$ and the contours show the kinetic energy density at 10 levels between 0.01 and 0.1 m^2 s^{-2} at $t = 5$–$t = 80$ inertial periods. The horizontal line on the $t = 5$, 10 and 20 snapshots shows the wavelength of a vertical mode-one $2f_0$ frequency internal wave. The slanting lines on the $t = 40$ and $t = 80$ snapshots show the characteristic propagation angles of NIWs with the indicated frequencies. The initial v and u are given in (1.1) and (1.2), where $V_1 = 0.4$ m s^{-1}, $L = 40$ km, $U = 0.8$ m s^{-1} and $h = 100$ m. Only the central 800 km of a 1200 km computational domain is shown. The solution is relaxed to a surface-concentrated inertial oscillation in sponge layers 150 km thick on the edges of the domain.

horizontally uniform and unbalanced flow which develops into an NIW. The balanced jet has the Gaussian profile

$$v(x, z, 0) = V_0 + V_1 \exp(-x^2/2L^2), \tag{1.1}$$

where V_0 is defined so that $v(x, z, 0)$ has zero horizontal average and thus no unbalanced component. The initial u is horizontally uniform and concentrated in a layer of depth h:

$$u(x, z, 0) = U_0 \exp(-z^2/2h^2). \tag{1.2}$$

The initial buoyancy b and vertical velocity w are zero. We solve this initial value problem in the Boussinesq equations using the spectral model of Winters, MacKinnon & Mills (2004) with 1536 Fourier modes in x and 768 sine/cosine modes in z.

If the jet in (1.1) were not present, the initial condition in u would develop into a horizontally uniform non-propagating perpetual inertial oscillation. Instead, refraction by the imposed jet injects small horizontal scales of size $\sim L$ into the NIW field, thus inducing near-inertial vertical propagation and catalysing radiation of low-mode $2f_0$ internal waves. The development of this process is illustrated in figure 1, which shows snapshots of the potential energy density at $t = 5$, 10, 20, 40 and 80 inertial periods. The kinetic energy density is indicated by 10 overlain contours between 0.01 and 0.1 m^2 s^{-2}. Throughout the simulation, the kinetic energy remains localized in the surface layer and in the near field of the barotropic jet. Bulges in the kinetic energy appearing at $t = 20$–80 inertial periods reveal the progress of vertical NIW propagation and show how NIW energy is refractively focused into the region of negative vorticity (Balmforth, Smith & Young 1998; Lee & Niiler 1998; Balmforth & Young 1999).

The vertical propagation of NIW kinetic energy is attended by an evolving potential energy field. Its most conspicuous aspect is a signal that extends for the full domain depth and radiates horizontally from the region of jet–NIW interaction. In the early stages, the potential energy signal has vertical mode-one structure. The horizontal line on the panels at $t = 5$, 10 and 20 inertial periods indicates the horizontal wavelength $\lambda_{2f} = 2NH/\sqrt{3}f_0 = 92.4$ km of a mode-one $2f_0$ frequency internal wave. Remarkably, while this $2f_0$ signal is generated by nonlinear NIW self-interaction in a small region, it rapidly radiates to fill a much larger volume without significant NIW activity (Danioux *et al.* 2008; Danioux & Klein 2008).

In addition to the low- and intermediate-mode $2f_0$ signal, narrow beams of potential energy radiate downwards and outwards from the centre of the domain. These beams are NIWs propagating at the characteristic angles indicated by the slanting lines on the snapshots at $t = 40$ and $t = 80$ inertial periods. The beams are produced by a scattering interaction between the surface-concentrated NIW and the jet. The rightward radiating beams are NIWs escaping the region of negative jet vorticity.

The two-dimensional NIW–jet interaction is thus characterized by at least three distinct phenomena: trapping of near-inertial energy in regions of negative balanced vorticity, beam-like radiation of near-inertial energy and emission of $2f_0$ waves. We use a multiple space- and time-scale expansion of the Boussinesq equations to construct a three-component model describing all of these processes.

1.2. *Summary of the three-component model*

In the three-component model, the horizontal velocity is

$$u + iv \stackrel{\text{def}}{=} e^{-if_0t} \, \mathsf{L} A + (-\partial_y + i\partial_x)\psi + \cdots, \tag{1.3}$$

where $A(x, y, z, t)$ is the NIW envelope and $\psi(x, y, z, t)$ is the QG streamfunction. The differential operator L in (1.3) is defined below in (1.7), and the dots on the right-hand side of (1.3) stand for additional contributions to the horizontal velocity: NIW harmonics, Stokes corrections and ageostrophic flow. The pressure field is

$$p = f_0 \psi + \frac{\mathrm{i} f_0}{2} [\mathrm{e}^{-\mathrm{i} f_0 t} (\partial_x - \mathrm{i} \partial_y) A + \mathrm{e}^{-2\mathrm{i} f_0 t} 2B] + \text{c.c.} + \cdots, \tag{1.4}$$

where $B(x, y, z, t)$ is the $2f_0$ wave envelope, 'c.c.' stands for 'complex conjugate' and the dots indicate unimportant high-order corrections. The vertical velocity w is

$$w = -\frac{f_0^2}{2N^2} [\mathrm{e}^{-\mathrm{i} f_0 t} (\partial_x - \mathrm{i} \partial_y) A_z + \mathrm{e}^{-2\mathrm{i} f_0 t} 4B_z] + \text{c.c.} \tag{1.5}$$

The $2f_0$ contribution in B features prominently in the vertical velocity field, despite its small contribution to the horizontal velocity.

The system consists of three equations: a wave-averaged QG potential vorticity equation, the NIW equation and a '$2f_0$ equation' governing the evolution of $2f_0$ waves. The wave-averaged potential vorticity equation is

$$q_t + \mathrm{J}(\psi, q) = 0, \tag{1.6}$$

where the potential vorticity is

$$q = \underbrace{\left(\partial_x^2 + \partial_y^2 + \underbrace{\partial_z \frac{f_0^2}{N^2} \partial_z}_{\overset{\text{def}}{=} \text{L}} \right)}_{\overset{\text{def}}{=} \Delta} \psi + \beta y + \frac{\mathrm{i}}{2f_0} \mathrm{J}(\mathrm{L} A^*, \mathrm{L} A) + \frac{1}{4f_0} \Delta |\mathrm{L} A|^2. \tag{1.7}$$

In (1.6) and (1.7) the operator $\mathrm{J}(a, b) = a_x b_y - a_y b_x$ is the Jacobian, the inertial of Coriolis frequency is $f = f_0 + \beta y$ and $N(z)$ is the depth-dependent buoyancy frequency associated with strong background stratification. The two rightmost terms in (1.7) are quadratic NIW contributions to the wave-averaged potential vorticity. We assume that the $2f_0$ waves are too weak to contribute appreciably to the potential vorticity. The evolution of the NIW field is described by a generalization of the YBJ equation,

$$\mathrm{L} A_t + \frac{\mathrm{i}}{2} f_0 \Delta A + \mathrm{J}(\psi, \mathrm{L} A) + \mathrm{i} \mathrm{L} A \left(\frac{1}{2} \Delta \psi + \beta y \right) + \frac{1}{2} \mathrm{L} A^* (\partial_x + \mathrm{i} \partial_y)^2 B = 0. \tag{1.8}$$

Equation (1.8) accounts for NIW dispersion and group propagation, horizontal advection by balanced flows, refraction by balanced flows and non-uniform planetary vorticity, and nonlinear NIW–$2f_0$ interaction. The NIW–$2f_0$ interaction term on the right end of (1.8) is identical to the term introduced by Young, Tsang & Balmforth (2008) into the YBJ equation to analyse near-inertial parametric subharmonic instability (PSI); in that work, the NIW–$2f_0$ interaction was implicated in the production of very small NIW vertical scales.

The evolution of the $2f_0$ amplitude B is described by

$$(\Delta + 13\mathrm{L}) B_t + 4\mathrm{i} f_0 (\Delta - 3\mathrm{L}) B = -\tfrac{3}{2} (\partial_x - \mathrm{i} \partial_y)^2 (\mathrm{L} A)^2. \tag{1.9}$$

Equation (1.9) describes dispersion and propagation of $2f_0$ waves and energy transfer from NIWs via the source term on the right.

The three-component model, comprising (1.6)–(1.9), describes the coupled evolution of NIWs, QG flow and near-$2f_0$ internal waves. Like the XV system, the three-component model conserves two integral quantities: 'wave action' and 'coupled energy'. The wave action is the sum of the NIW kinetic energy and the total energy of freely propagating near-$2f_0$ waves. The coupled energy is the sum of the total balanced energy, the near-inertial potential energy, an NIW–β interaction term and terms associated with the NIW–$2f_0$ interaction.

A striking implication of both the XV and the three-component models is that NIWs can extract energy from balanced QG flow. This follows from the separation of two conservation laws for wave action and coupled energy, which together imply that an increase in NIW potential energy during NIW–flow interaction comes at the expense of balanced energy. Balanced flow thus loses energy when interacting with NIWs that consist almost entirely of kinetic energy, and NIW–QG interaction forms a link between large-scale balanced energy, the energy contained in the internal wave field, and wave breaking and diapycnal mixing. XV refer to this wave–mean interaction as 'stimulated loss of balance' to distinguish it from spontaneous loss of balance (Vanneste 2013), emphasizing that it requires externally forced waves to 'stimulate' further production of wave energy at the expense of balanced energy. Unlike spontaneous wave generation, stimulated wave generation is a potentially significant energy sink for nearly balanced flows with small Rossby numbers.

Next, in §2, we introduce the Boussinesq equations and special scaling assumptions, and define the multiple time and multiple vertical scales required to meet solvability conditions in the asymptotic derivation. In §3 we expand the Boussinesq equations in wave amplitude, deriving the NIW equation as well as the $2f_0$ equation governing the evolution of the $2f_0$ harmonic. In §4 we apply the wave-averaged contribution to the QG potential vorticity found by Wagner & Young (2015) to the near-inertial case. In §5 we heuristically revise the formal theory derived in §§3 and 4 to arrive at the implementable model of (1.6)–(1.9). In §6 we derive two conserved integral quantities from (1.6)–(1.9). In §7 we compare numerical solutions of a two-dimensional initial value problem in both Boussinesq and three-component models, and in §8 we assess the physical implications of the solutions. We conclude with a discussion of the significance and implications of the model in §9.

2. The Boussinesq equations

We use the Boussinesq approximation by defining a constant reference density ρ_0 and decomposing the total density into

$$\rho(\boldsymbol{x}, t) = \rho_0 \left[1 + g^{-1} \int_z^0 N^2(z')\,\mathrm{d}z' - g^{-1} b(\boldsymbol{x}, t) \right], \tag{2.1}$$

where $\boldsymbol{x} = (x, y, z)$ is position and t is time. In (2.1), the resting buoyancy profile is the vertical integral of the buoyancy frequency $N^2(z)$, and $b(\boldsymbol{x}, t)$ is the dynamic buoyancy perturbation from background. Pressure is decomposed similarly into $-\rho_0 g z + \rho_0 P(z) + \rho_0 p(\boldsymbol{x}, t)$, where $P(z)$ is the resting hydrostatic pressure and the dynamic component of pressure is $\rho_0 p$.

Using these definitions, the Boussinesq equations are

$$u_t + \boldsymbol{u} \cdot \nabla u - f v + p_x = 0, \tag{2.2}$$
$$v_t + \boldsymbol{u} \cdot \nabla v + f u + p_y = 0, \tag{2.3}$$

$$w_t + \boldsymbol{u} \cdot \nabla w + p_z = b, \tag{2.4}$$

$$b_t + \boldsymbol{u} \cdot \nabla b + wN^2 = 0, \tag{2.5}$$

$$u_x + v_y + w_z = 0, \tag{2.6}$$

where $\boldsymbol{u} = u\hat{\boldsymbol{x}} + v\hat{\boldsymbol{y}} + w\hat{\boldsymbol{z}}$ is the fluid velocity. Subscripts with respect to (x, y, z) or t denote partial derivatives. We use the β-plane approximation by introducing $f(y) = f_0 + \beta y$, where f_0 is the local inertial frequency and β is its latitudinal variation.

2.1. *Non-dimensionalization and scaling assumptions*

We set the asymptotic reduction in motion by non-dimensionalizing the Boussinesq equations (2.2)–(2.6). We choose a spatial scaling that isolates NIWs at leading order and a velocity scaling that ensures that the back-rotated velocity and the QG potential vorticity share the same evolutionary time scale. Specifically, this requires that NIW dispersion acts on the same time scale as advection and refraction by the balanced flow. We use a single horizontal length scale, L, and denote the scale of the near-inertial horizontal velocity with $\tilde{U}$. The NIW 'amplitude parameter'

$$\epsilon \overset{\text{def}}{=} \frac{\tilde{U}}{f_0 L} \tag{2.7}$$

is crucial: $\epsilon \ll 1$ implies that nonlinear terms are small so that the NIW field is governed by linear dynamics to leading order.

The amplitude and importance of nonlinearity to balanced-flow evolution are measured by the Rossby number. We assume that the balanced flow is weak relative to the NIWs and that $\bar{U} = \epsilon \tilde{U}$, where $\bar{U}$ is the characteristic velocity of the balanced flow. Under this scaling assumption the Rossby number is

$$Ro \overset{\text{def}}{=} \frac{\bar{U}}{f_0 L} = \epsilon^2. \tag{2.8}$$

The NIW amplitude parameter and Rossby number have superficial similarity but different physical interpretations. The NIW amplitude can be interpreted as the ratio between the length scale L and the radius of particle orbits in an inertial circle, $\tilde{U}/f_0$. The Rossby number, on the other hand, is the ratio of the rotation time scale $1/f_0$ and advective time scale $L/\bar{U} = (\epsilon^2 f_0)^{-1}$. The NIW envelope and the balanced flow co-evolve on the slow time scale $(\epsilon^2 f_0)^{-1}$.

We denote the vertical scale of the NIWs by $\tilde{H}$, and we use $\tilde{U}$, L and $\tilde{H}$ to non-dimensionalize the horizontal and vertical velocities,

$$(u, v) = \tilde{U}\,(\check{u}, \check{v}), \quad w = \frac{\tilde{H}\tilde{U}}{L}\,\check{w}, \tag{2.9a,b}$$

where non-dimensional variables are distinguished by $\check{}$. Introducing f_0, the local Coriolis or inertial frequency, and N_0, the characteristic magnitude of $N(z)$, we non-dimensionalize the buoyancy field with

$$b = (\tilde{H}N_0^2\tilde{U}/f_0 L)\,\check{b}. \tag{2.10}$$

We adopt a geostrophic scaling for the pressure such that

$$p = f_0 L\bar{U}\check{p} = \epsilon f_0 L\tilde{U}\check{p}. \tag{2.11}$$

The inertial frequency is scaled so that

$$\check{f} = 1 + \epsilon^2 \check{\beta} \check{y}, \tag{2.12}$$

where $(x, y) = L(\check{x}, \check{y})$ and

$$\check{\beta} = \frac{\beta L^2}{\bar{U}}. \tag{2.13}$$

Finally, we define an aspect ratio

$$\alpha \stackrel{\text{def}}{=} \epsilon \frac{N_0}{f_0}. \tag{2.14}$$

By assuming $\alpha = O(1)$, we imply that $f_0/N_0 = O(\epsilon)$ and justify the hydrostatic approximation in the vertical momentum equation at all relevant orders in the perturbation theory.

2.2. *Multiple scales: time and space*

To describe both internal waves and slowly evolving balanced flow, we use the two-time method with a 'fast' time $\tilde{t} = f_0 t$ and a 'slow' time $\bar{t} = \epsilon^2 f_0 t$. Time derivatives are mapped according to

$$\partial_t \mapsto f_0(\partial_{\tilde{t}} + \epsilon^2 \partial_{\bar{t}}). \tag{2.15}$$

We use an Eulerian time average denoted with an overbar and defined as

$$\bar{\phi}(\boldsymbol{x}, \bar{t}) \stackrel{\text{def}}{=} \frac{1}{T} \int_{t-T/2}^{t+T/2} \phi(\boldsymbol{x}, t') \, dt', \quad \text{where} \quad \frac{1}{f_0} \ll T \ll \frac{L}{\bar{U}}, \tag{2.16}$$

to separate fast and slow flow components. Thus, any field ϕ can be represented as

$$\phi = \bar{\phi} + \tilde{\phi}, \tag{2.17}$$

where $\bar{\phi}$ is the slowly evolving time-mean part and $\tilde{\phi}$ is the wavy part with $\overline{\tilde{\phi}} = 0$.

A multiple-vertical-scale expansion in the vertical is motivated by the disparity in aspect ratio between NIWs, and both observed $2f_0$ scales as well as standard QG flow. Denoting the vertical scale of the NIWs by $\tilde{H}$, the internal wave dispersion relation implies that internal waves are near-inertial when the Burger number of the wave is small, or when

$$\left(\frac{N_0 \tilde{H}}{f_0 L} \right)^2 \ll 1. \tag{2.18}$$

On the other hand, the standard QG equations assume that the Burger number of the balanced flow is of order unity. We use this requirement to define the vertical scale of the balanced flow, $\bar{H}$, as

$$\bar{H} \stackrel{\text{def}}{=} \frac{f_0 L}{N_0}. \tag{2.19}$$

We make the scaling assumption that

$$\tilde{H} = \epsilon \bar{H}. \tag{2.20}$$

This prescription for $\tilde{H}$ unifies the slow NIW dispersion time scale with the balanced-flow advection time scale. To capture both vertical scales in the expansion we split the vertical coordinate into a fast component, $\tilde{z}$, and a slow component, $\bar{z}$. Under this two-scale splitting, vertical derivatives become

$$\partial_z \mapsto \tilde{H}^{-1}(\partial_{\tilde{z}} + \epsilon \, \partial_{\bar{z}}). \tag{2.21}$$

The vertical-scale splitting requires the introduction of a vertical average, which we define as

$$\hat{\phi} = \frac{1}{H'} \int_{\bar{z}-H'/2}^{\bar{z}+H'/2} \phi \, \mathrm{d}\tilde{z}', \quad \text{where } \tilde{H} \ll H' \ll \bar{H}. \tag{2.22}$$

The increase in complexity incurred by the multiple space-scale expansion is justified by a systematic explanation of the prominence and impact of the $2f_0$ harmonic on NIW evolution.

2.3. *Complexifed non-dimensionalized equations*

The derivation is greatly simplified by defining the complex horizontal coordinate and velocity field,

$$s \stackrel{\text{def}}{=} x + \mathrm{i}y \quad \text{and} \quad \mathcal{U} \stackrel{\text{def}}{=} u + \mathrm{i}v. \tag{2.23a,b}$$

Spatial derivatives are expressed in terms of s and s^* via

$$\partial_s = \tfrac{1}{2}(\partial_x - \mathrm{i}\partial_y), \quad \partial_{s^*} = \tfrac{1}{2}(\partial_x + \mathrm{i}\partial_y). \tag{2.24a,b}$$

These imply that $\Delta = 4\partial_s\partial_{s^*}$, and that

$$u_x + v_y = \mathcal{U}_s + \mathcal{U}_{s^*}^* \quad \text{and} \quad v_x - u_y = \mathrm{i}\mathcal{U}_{s^*}^* - \mathrm{i}\mathcal{U}_s. \tag{2.25a,b}$$

Using the scaling assumptions outlined above, and dropping decorations on non-dimensional variables, the complexified non-dimensional Boussinesq equations become

$$\mathcal{U}_{\tilde{t}} + \mathrm{i}\mathcal{U} = -\epsilon(\boldsymbol{u} \cdot \nabla\mathcal{U} + 2\,p_{s^*}) - \epsilon^2(\mathcal{U}_{\tilde{t}} + w\mathcal{U}_{\bar{z}} + \mathrm{i}\beta y\mathcal{U}), \tag{2.26}$$

$$p_{\bar{z}} = \epsilon\,(b - p_{\bar{z}}) - \epsilon^2\alpha^{-2}[w_{\tilde{t}} + \epsilon(\boldsymbol{u} \cdot \nabla w) + \epsilon^2(ww_{\bar{z}} + w_{\tilde{t}})], \tag{2.27}$$

$$b_{\tilde{t}} + wN^2 = -\epsilon\,\boldsymbol{u} \cdot \nabla b - \epsilon^2(b_{\tilde{t}} + wb_{\bar{z}}), \tag{2.28}$$

$$\mathcal{U}_s + \mathcal{U}_{s^*}^* + w_{\bar{z}} = -\epsilon\,w_{\bar{z}}. \tag{2.29}$$

The bracketed terms in (2.27) are included for completeness, but never appear in the theory that follows. In terms of complex velocity the advection operators in (2.26) and (2.28) are

$$\boldsymbol{u} \cdot \nabla = \mathcal{U}\partial_s + \mathcal{U}^*\partial_{s^*} + w\partial_{\bar{z}}. \tag{2.30}$$

The system in (2.26)–(2.29) is the basis for our asymptotic derivation.

3. The NIW equation

The NIW equation is derived by developing a perturbation expansion of (2.26)–(2.29) for $\epsilon \ll 1$. We begin by expanding $\boldsymbol{u}$, b and p each in a series in ϵ. For example, the complex velocity $\mathcal{U}$ has the expansion

$$\mathcal{U} = \mathcal{U}_0 + \epsilon\,\mathcal{U}_1 + \epsilon^2\,\mathcal{U}_2 + \cdots. \tag{3.1}$$

We develop (2.26)–(2.29) order by order in ϵ. For clarity, we express our results in dimensional variables, although the non-dimensional forms are indispensable for distinguishing each order in the development.

3.1. *Leading order: NIWs*

The leading-order terms in (2.26)–(2.29) are

$$\mathcal{U}_{0\tilde{t}} + \mathrm{i}f_0\mathcal{U}_0 = 0, \tag{3.2}$$

$$p_{0\tilde{z}} = 0, \tag{3.3}$$

$$b_{0\tilde{t}} + w_0 N^2 = 0, \tag{3.4}$$

$$\mathcal{U}_{0s} + \mathcal{U}_{0s^*}^* + w_{0\tilde{z}} = 0. \tag{3.5}$$

We write the solution to the horizontal momentum equation (3.2) in terms of an NIW envelope M or A,

$$\mathcal{U}_0 = \mathrm{e}^{-\mathrm{i}f_0\tilde{t}}\, M_{\tilde{z}} = \mathrm{e}^{-\mathrm{i}f_0\tilde{t}}\, \tilde{\mathrm{L}}A, \tag{3.6}$$

where $\tilde{\mathrm{L}}$ is a second-order differential operator,

$$\tilde{\mathrm{L}} \overset{\text{def}}{=} \partial_{\tilde{z}}\left(\frac{f_0^2}{N^2}\partial_{\tilde{z}}\right). \tag{3.7}$$

Both $A(x, y, \tilde{z}, \bar{z}, \bar{t})$ and $M = (f_0^2/N^2)A_{\tilde{z}}$ prove useful for confronting the algebra that ensues. The representation in (3.6) ensures that the leading-order horizontal velocity is inertial over short times; small deviations in wave field frequency from f_0 are captured by the dependence of M or A on the slow time $\bar{t}$. The construction in (3.6) also implies that the vertical average of the NIW horizontal velocity is zero at this order.

With the representation in (3.6), we can integrate the continuity equation (3.5) over the fast vertical coordinate $\tilde{z}$ to yield

$$w_0 = -\mathrm{e}^{-\mathrm{i}f_0\tilde{t}} M_s - \mathrm{e}^{\mathrm{i}f_0\tilde{t}} M_{s^*}^* + \hat{w}_0, \tag{3.8}$$

where the $\tilde{z}$-independent function of integration $\hat{w}_0(x, y, \bar{z}, \tilde{t}, \bar{t})$ is necessary to ensure solvability of the perturbation expansion at next order. If $\hat{w}_0$ is not included in (3.8), then the $O(\epsilon)$ velocity field cannot satisfy continuity and the boundary conditions. At $O(\epsilon)$ in (2.26)–(2.29), we find that $\hat{w}_0$ oscillates on the fast time scale with frequency $2f_0$ and is forced nonlinearly by NIW horizontal self-advection.

The leading-order buoyancy b_0 follows from integration of the buoyancy equation (3.4) using w_0 in (3.8),

$$b_0 = \mathrm{i}f_0(\mathrm{e}^{-\mathrm{i}f_0\tilde{t}} A_{\tilde{z}s} - \mathrm{e}^{\mathrm{i}f_0\tilde{t}} A_{\tilde{z}s^*}^*) + \hat{b}_0, \tag{3.9}$$

where as in (3.8) we include the function of integration $\hat{b}_0(x, y, \bar{z}, \tilde{t}, \bar{t})$. The vertical momentum equation (3.3) implies that the leading-order pressure p_0 does not depend on the fast vertical scale $\tilde{z}$, or that

$$p_0 = \hat{p}_0. \tag{3.10}$$

The leading-order pressure p_0 is eventually determined by (3.28) and (3.29) below and oscillates on a fast time scale with frequency $2f_0$. An important feature eventually revealed by this expansion is that the large-vertical-scale $2f_0$ fields $\hat{w}_0$, $\hat{b}_0$ and $\hat{p}_0$ appear at leading order in (3.8), (3.9) and (3.28). The magnitudes of these $2f_0$ fields follow from the two assumptions that the $2f_0$ horizontal velocity and vertical scales are respectively $O(\epsilon)$ and $O(\epsilon^{-1})$ compared with their near-inertial counterparts. This scaling produces a description of NIW–$2f_0$ interaction that neglects the complicating effect of the mean flow ψ on $2f_0$ evolution.

3.2. *First order: wave-averaged geostrophic balance and $2f_0$ harmonic*

The $O(\epsilon)$ terms in (2.26)–(2.29) are

$$\mathcal{U}_{1\tilde{t}} + \mathrm{i}f_0\mathcal{U}_1 = -2p_{0s^*} - \boldsymbol{u}_0 \cdot \boldsymbol{\nabla}\mathcal{U}_0, \tag{3.11}$$

$$p_{1\tilde{z}} = b_0 - p_{0\tilde{z}}, \tag{3.12}$$

$$b_{1\tilde{t}} + w_1N^2 = -\boldsymbol{u}_0 \cdot \boldsymbol{\nabla}b_0, \tag{3.13}$$

$$\mathcal{U}_{1s} + \mathcal{U}_{1s^*}^* + w_{1\tilde{z}} = -w_{0\tilde{z}}. \tag{3.14}$$

These equations describe wave-averaged geostrophic balance and the nonlinearly forced $2f_0$ harmonic.

3.2.1. *Wave-averaged geostrophic balance*

The time average of (3.11)–(3.14) yields the wave-averaged geostrophic balance conditions. These balance conditions are similar to those in Wagner & Young (2015) except that the restriction to NIWs means that there is no Stokes pressure contribution. We show this explicitly by first noting that the nonlinear term on the right-hand side of (3.11) is

$$\boldsymbol{u}_0 \cdot \boldsymbol{\nabla}\mathcal{U}_0 = \mathcal{J}_0 + \mathrm{e}^{-2\mathrm{i}f_0\tilde{t}}\,\mathcal{J}_2 + \mathrm{e}^{-\mathrm{i}f_0\tilde{t}}\,\hat{w}_0 M_{\tilde{z}\tilde{z}}, \tag{3.15}$$

where $\mathcal{J}_0$ and $\mathcal{J}_2$ are Jacobians defined by

$$\mathcal{J}_0 \stackrel{\text{def}}{=} \frac{\partial(M^*, M_{\tilde{z}})}{\partial(\tilde{z}, s^*)} \quad \text{and} \quad \mathcal{J}_2 \stackrel{\text{def}}{=} \frac{\partial(M, M_{\tilde{z}})}{\partial(\tilde{z}, s)}. \tag{3.16a,b}$$

Next, we observe that the horizontal and vertical Stokes drifts, $\mathcal{U}^S$ and w^S, are defined by

$$\mathcal{U}^S \stackrel{\text{def}}{=} \overline{\boldsymbol{\xi}_0 \cdot \boldsymbol{\nabla}\mathcal{U}_0} \quad \text{and} \quad w^S \stackrel{\text{def}}{=} \overline{\boldsymbol{\xi}_0 \cdot \boldsymbol{\nabla}w_0}, \tag{3.17a,b}$$

where $\boldsymbol{\xi}_0 = \xi_0\hat{\boldsymbol{x}} + \eta_0\hat{\boldsymbol{y}} + \zeta_0\hat{\boldsymbol{z}}$ is the wave particle displacement, defined via $\boldsymbol{\xi}_{0\tilde{t}} = \boldsymbol{u}_0$ and $\bar{\boldsymbol{\xi}}_0 = 0$. A direct calculation shows that

$$\mathrm{i}f_0\mathcal{U}^S = \overline{\boldsymbol{u}_0 \cdot \boldsymbol{\nabla}\mathcal{U}_0} = \mathcal{J}_0. \tag{3.18}$$

A similar calculation for the vertical Stokes drift w^S shows that

$$\mathrm{i}f_0 w^S = \frac{\mathrm{i}f_0}{N^2}\overline{\boldsymbol{u}_0 \cdot \boldsymbol{\nabla}b_0} = \mathcal{K}_0^* - \mathcal{K}_0, \tag{3.19}$$

where $\mathcal{K}_0$ is the Jacobian

$$\mathcal{K}_0 \stackrel{\text{def}}{=} \frac{\partial(M^*, M_s)}{\partial(\tilde{z}, s^*)}. \tag{3.20}$$

The identity $\mathcal{J}_{0s} + \mathcal{J}_{0s^*}^* + \mathcal{K}_{0z}^* - \mathcal{K}_{0z} = 0$ implies that the three-dimensional Stokes velocity in (3.18) and (3.19) is non-divergent.

Defining the QG streamfunction as

$$\psi \stackrel{\text{def}}{=} \frac{\bar{p}_0}{f_0}, \tag{3.21}$$

we use the expressions for the Stokes velocities in (3.18) and (3.19) to write the time averages of (3.11) and (3.13) as

$$\bar{\mathcal{U}}_1 + \mathcal{U}^S = 2\mathrm{i}\psi_{s^*}, \tag{3.22}$$

$$\bar{w}_1 + w^S = 0. \tag{3.23}$$

Equation (3.22) is the wave-averaged geostrophic balance condition for QG flow evolution in a field of strong NIWs. This balance condition lacks the Stokes pressure correction term that appears in the more general balance condition expressed by (4.38) in Wagner & Young (2015). From the leading-order vertical momentum equations (3.3), the pressure p_0, and therefore ψ, does not depend on the fast vertical coordinate $\tilde{z}$.

3.2.2. *The $2f_0$ harmonic*

Using the two-time decomposition in (2.17), we write the wavy parts of the first-order equations (3.11), (3.12) and (3.14),

$$\tilde{\mathcal{U}}_{1\tilde{t}} + \mathrm{i}f_0\tilde{\mathcal{U}}_1 + 2\tilde{p}_{0s*} = -\mathrm{e}^{-2\mathrm{i}f_0\tilde{t}}\,\mathcal{J}_2 - \mathrm{e}^{-\mathrm{i}f_0\tilde{t}}\,\hat{w}_0 M_{\tilde{z}\tilde{z}}, \tag{3.24}$$

$$\tilde{p}_{1\tilde{z}} = \tilde{b}_0 - \tilde{p}_{0\bar{z}}, \tag{3.25}$$

$$\hat{b}_{0\tilde{t}} + \hat{w}_0 N^2 = 0, \tag{3.26}$$

$$\tilde{\mathcal{U}}_{1s} + \tilde{\mathcal{U}}_{1s*}^* + \tilde{w}_{1\tilde{z}} = -\tilde{w}_{0\bar{z}}, \tag{3.27}$$

where with (3.26) we include the vertically averaged leading-order buoyancy equation. It is (3.26), rather than the wavy part of (3.13), that describes the part of the $2f_0$ buoyancy field with large vertical scale. It should be noted that the final term on the right-hand side of (3.24) is not resonant because $\hat{w}_0$ oscillates with $2f_0$ frequency.

The system in (3.24)–(3.27) provides a complete description of the $2f_0$ harmonic of the NIW field. Importantly, part of this $2f_0$ harmonic response does not depend on the fast vertical coordinate $\tilde{z}$. To isolate the slowly vertically varying part of the $2f_0$ harmonic we average (3.24)–(3.27) over $\tilde{z}$ and wrangle the resulting system into a single equation. We leave the details to appendix A and note the final result. Using the notation

$$\tilde{p}_0 = \mathrm{i}f_0[\mathrm{e}^{-2\mathrm{i}f_0\tilde{t}}B(x, y, \bar{z}, \bar{t}) - \mathrm{e}^{2\mathrm{i}f_0\tilde{t}}B^*(x, y, \bar{z}, \bar{t})], \tag{3.28}$$

we find that B solves

$$\mathrm{i}f_0(\Delta - 3\bar{\mathrm{L}})B = -\tfrac{3}{2}\partial_s^2\,\widehat{M_{\tilde{z}}^2}. \tag{3.29}$$

The operator $\bar{\mathrm{L}}$ is a second-order differential operator analogous to $\tilde{\mathrm{L}}$ but defined in terms of the slow vertical scale $\bar{z}$,

$$\bar{\mathrm{L}} \stackrel{\mathrm{def}}{=} \partial_{\bar{z}}\frac{f_0^2}{N^2}\partial_{\bar{z}}. \tag{3.30}$$

The '$2f_0$ equation' in (3.29) describes forced $2f_0$ oscillations with a much larger vertical scale than the near-inertial fields. Because of this vertical-scale discrepancy, the vertical velocity of the $2f_0$ harmonic appears alongside the NIW vertical velocity at leading order in (3.8). As it stands, however, (3.29) cannot describe freely propagating $2f_0$ waves and thus cannot describe the waves that produce the prominent potential energy signal in figure 1. We modify (3.29) heuristically in §5 to describe the freely propagating parts of B, which roughly satisfy the $2f_0$ dispersion relation and thus obey $\Delta B \approx 3\mathrm{L}\,B$.

Continuing with the derivation of the NIW evolution equation, we use the expression for $\tilde{p}_0$ in (3.28) to integrate (3.24) for $\tilde{\mathcal{U}}_1$. The full $\mathcal{U}_1$ field is

$$\mathcal{U}_1 = 2\mathrm{i}\psi_{s*} + \mathrm{i}f_0^{-1}\mathcal{J}_0 + \mathrm{e}^{-2\mathrm{i}f_0\tilde{t}}(2B_{s*} - \mathrm{i}f_0^{-1}\mathcal{J}_2) + \tfrac{2}{3}\mathrm{e}^{2\mathrm{i}f_0\tilde{t}}B_{s*}^* + \cdots, \tag{3.31}$$

where the dots indicate terms proportional to $e^{-3if_0\tilde{t}}$ and $e^{if_0\tilde{t}}$. Finally, we find $p_1 - \hat{p}_1$ by subtracting the vertical average from (3.25), using (3.9) and integrating in $\tilde{z}$ to find

$$p_1 - \hat{p}_1 = if_0(A_s e^{-if_0\tilde{t}} - A_{s^*}^* e^{if_0\tilde{t}}). \tag{3.32}$$

We now have $\mathcal{U}_1$, $\hat{w}_0$ and p_1, and are ready to proceed to the second-order system.

3.3. *Second order: an NIW amplitude evolution equation*

The $O(\epsilon^2)$ terms in the horizontal momentum equation (2.26) are

$$\mathcal{U}_{2\tilde{t}} + if_0\mathcal{U}_2 = -\boldsymbol{u}_0 \cdot \nabla\mathcal{U}_1 - \boldsymbol{u}_1 \cdot \nabla\mathcal{U}_0 - \mathcal{U}_{0\bar{t}} - i\beta y\mathcal{U}_0 - 2p_{1s^*} - w_0\mathcal{U}_{0\tilde{z}}. \tag{3.33}$$

Here, we finally apply the solvability condition arising from the introduction of multiple time scales. The solvability condition prevents the disordering of terms that would result from secular growth in $\mathcal{U}_2$: we isolate resonant forcing terms on the right of (3.33) and set them collectively to zero. The amplitude equation yielded by this procedure governs the dependence of the NIW envelope A on the slow time $\bar{t}$. We note that the vertical average of (3.33) has no resonant terms.

We construct the amplitude equation piece by piece, starting at the far-right end of (3.33) and proceeding to the left. The final term $w_0\mathcal{U}_{0\tilde{z}}$ in (3.33) has no parts proportional to $e^{-if_0\tilde{t}}$ and so does not contribute to the amplitude equation. The next three terms from the left side of (3.33) are

$$\mathcal{U}_{0\bar{t}} + i\beta y\mathcal{U}_0 + 2\partial_{s^*}(p_1 - \hat{p}_1) = e^{-if_0\tilde{t}}(\tilde{L}A_{\bar{t}} + i\beta y\tilde{L}A + 2if_0A_{ss^*}) + \text{NRT}, \tag{3.34}$$

where NRT stands for 'non-resonant terms'. Next in line is

$$(\boldsymbol{u}_1 \cdot \nabla)\mathcal{U}_0 = e^{-if_0\tilde{t}}\left[2i\frac{\partial(\psi, M_{\tilde{z}})}{\partial(s^*, s)} - \mathcal{U}^S M_{\tilde{z}s} - \mathcal{U}^{S*}M_{\tilde{z}s^*} - w^S M_{\tilde{z}\tilde{z}}\right] + \text{NRT}. \tag{3.35}$$

It should be noted that to find (3.35) we need only consider the time-mean velocity $\bar{\boldsymbol{u}}_1$, since $\mathcal{U}_0$ is proportional to $e^{-if_0\tilde{t}}$. The first term on the right of (3.33), involving the zero-order advection of the first-order velocity, is the most complicated. Carefully compiling the terms, we find

$$(\boldsymbol{u}_0 \cdot \nabla)\mathcal{U}_1 = e^{-if_0\tilde{t}}(M_{\tilde{z}}\mathcal{U}_{1s} - M_s\mathcal{U}_{1\tilde{z}}) + e^{if_0\tilde{t}}(M_{\tilde{z}}^*\mathcal{U}_{1s^*} - M_{s^*}^*\mathcal{U}_{1\tilde{z}}) + \hat{w}_0\mathcal{U}_{1\tilde{z}}, \tag{3.36}$$

$$= e^{-if_0\tilde{t}}\left[\frac{\partial(M, \bar{\mathcal{U}}_1)}{\partial(\tilde{z}, s)} + \frac{i}{f_0}\frac{\partial(\mathcal{J}_2, M^*)}{\partial(\tilde{z}, s^*)} + 2M_{\tilde{z}}^*B_{s^*s^*}\right] + \text{NRT}. \tag{3.37}$$

Adding (3.35) to (3.37) yields

$$(\boldsymbol{u}_1 \cdot \nabla)\mathcal{U}_0 + (\boldsymbol{u}_0 \cdot \nabla)\mathcal{U}_1 = e^{-if_0\tilde{t}}\left[2i\frac{\partial(\psi, M_{\tilde{z}})}{\partial(s^*, s)} + 2i\psi_{ss^*}M_{\tilde{z}} + 2M_{\tilde{z}}^*B_{s^*s^*}\right] + \text{NRT}. \tag{3.38}$$

The absence of terms cubic in M is a remarkable aspect of (3.38): all 16 cubic M terms in (3.35) and (3.37) conspire in collective cancellation. This simplification was previously noted by Falkovich, Kuznetsov & Medvedev (1994) and Zeitlin, Reznik & Ben Jelloul (2003), and is the reason why no cubic terms appear in XV.

It is thus notable that our expansion identifies a surviving 'honorary' cubic term, proportional to $M_{\tilde{z}}^*B_{s^*s^*}$, in (3.38). This new term results from the interaction of NIWs with both forced and freely propagating $2f_0$ waves. The requirement for $2f_0$ fields

arises when the first-order continuity equation (3.27) is averaged over the small NIW vertical scale: if the vertical average of $\mathcal{U}_{1s} + \mathcal{U}_{1s^*}^*$ is non-zero, for example, then continuity can only be satisfied if $\mathcal{U}_1$ is permitted its own independent evolution. This solvability issue is addressed by introducing the $2f_0$ fields $\hat{w}_0$, $\hat{b}_0$ and $\hat{p}_0$ at leading order in (3.8) and (3.9). This is the non-obvious step that ultimately produces the new term in (3.38).

The amplitude equation is then obtained from the sum of (3.34) and (3.38). In Cartesian coordinates and in terms of A, the amplitude equation is

$$\tilde{L}A_{\tilde{t}} + \frac{\mathrm{i}f_0}{2}\Delta A + \mathrm{J}(\psi, \tilde{L}A) + \mathrm{i}\tilde{L}A\left(\frac{1}{2}\Delta\psi + \beta y\right) + \frac{1}{2}\tilde{L}A^*(\partial_x + \mathrm{i}\partial_y)^2 B = 0. \qquad (3.39)$$

The amplitude equation (3.39) is the YBJ equation except for the final $2f_0$ interaction term on the left-hand side. The $2f_0$ interaction term is identical to the one found by Young *et al.* (2008) in their analysis of energy transfer from prescribed $2f_0$ motions to NIWs by PSI.

4. The NIW-averaged available potential vorticity

Wave-averaged QG flow is governed by a wave-averaged potential vorticity equation (Bühler & McIntyre 1998; Wagner & Young 2015),

$$q_{\tilde{t}} + \mathrm{J}(\psi, q) = 0, \qquad (4.1)$$

where ψ is defined through the balance condition (3.22) and q is the wave-averaged available potential vorticity. Wagner & Young (2015) give a number of expressions for q. Here, we use

$$q \stackrel{\text{def}}{=} (\Delta + \mathrm{L})\psi + \beta y + \underbrace{\overline{\mathrm{J}(u_0, \xi_0)} + \overline{\mathrm{J}(v_0, \eta_0)} + f_0\overline{\mathrm{J}(\xi_0, \eta_0)} + \tfrac{1}{2}\overline{(\overline{\xi_{0i}\xi_{0j}})_{,ij}}}_{\stackrel{\text{def}}{=}q^w}. \qquad (4.2)$$

In (4.2) we define the wave contribution to the available potential vorticity, q^w, in terms of the leading-order wave particle displacement $\boldsymbol{\xi}_0 = \xi_0\hat{\boldsymbol{x}} + \eta_0\hat{\boldsymbol{y}} + \zeta_0\hat{\boldsymbol{z}}$ defined through $\boldsymbol{\xi}_{0\tilde{t}} = \boldsymbol{u}_0$.

In the present multiple-scale theory, ψ and q are both time-averaged and vertically averaged quantities. Consistency then demands that q^w in (4.2) be vertically averaged as well. With the leading-order wave expressions (3.6) and (3.8) and using M, a bit of algebra leads to

$$q^w = -\frac{1}{f_0}(M_{ss^*}^* M_{\tilde{z}\tilde{z}} - 2M_{\tilde{z}s^*}^* M_{\tilde{z}s}^* + M_{ss^*}M_{\tilde{z}\tilde{z}}^*). \qquad (4.3)$$

This is the expression for q^w found by XV.

We then take the vertical average of q^w, which yields a number of representations via integration by parts in $\tilde{z}$, such as

$$\widehat{q^w} = -\frac{1}{f_0}(\widehat{M_{ss^*}^* M_{\tilde{z}\tilde{z}}} - 2\widehat{M_{\tilde{z}s^*}^* M_{\tilde{z}s}^*} + \widehat{M_{ss^*}M_{\tilde{z}\tilde{z}}^*}), \qquad (4.4)$$

$$= \frac{\mathrm{i}}{2f_0}\mathrm{J}(\widehat{M_{\tilde{z}}^*, M_{\tilde{z}}}) + \frac{1}{4f_0}\Delta\,\widehat{|M_{\tilde{z}}|^2}. \qquad (4.5)$$

We take the second form, in (4.5), which is the form needed to furnish the three-component model in (1.6)–(1.9) with a coupled wave–mean-energy conservation law.

5. Remodelling

The formally derived model comprising (3.39), (4.2), (4.5) and (3.29) is a first draft that we heuristically revise to obtain the simpler and well-posed system in (1.6)–(1.9). This remodelling addresses two concerns with the multiple-scale formulation. First, the multiple vertical scales and vertical averages in (4.5) and (3.29) complicate computations. Second, the $2f_0$ equation in (3.29) cannot be solved when the NIW field and its $2f_0$ harmonic interact resonantly, which occurs when the nonlinear forcing on the right-hand side of (3.29) has spectral components in the null space of the operator $\Delta - 3L$ on the left.

To address the first concern we reconsolidate vertical scales and eliminate vertical averaging from equations (4.5) and (3.29). Although this modification admits spurious small vertical scales into ψ and B, these small-scale parts of ψ and B contain little energy due to the 'self-averaging' Helmholtzian inversions that determine ψ and B through (4.2) and (3.29). In particular, energy is transferred most effectively to B at resonant or near-resonant spectral components of (3.29) and (5.3). These spectral components lie close to the $2f_0$ dispersion relation and have large vertical scale relative to NIWs.

After consolidation of scales and dismissal of vertical averages the potential vorticity is given in terms of ψ and wave-averaged properties as

$$q = (\Delta + L)\psi + \frac{i}{2f_0} J(LA^*, LA) + \frac{1}{4f_0}\Delta|LA|^2, \tag{5.1}$$

and the NIW equation is

$$LA_t + \frac{i}{2}f_0\Delta A + J(\psi, LA) + iLA\left(\frac{1}{2}\Delta\psi + \beta y\right) + \frac{1}{2}LA^*(\partial_x + i\partial_y)^2 B = 0. \tag{5.2}$$

In (5.2) above, $L = \partial_z f_0^2/N^2 \partial_z$ is the operator originally defined in (1.7) in terms of the single vertical scale z. The evolution of q in (5.1) is governed by the potential vorticity equation in (1.6).

The second issue regarding the non-invertibility of $\Delta - 3L$ and the description of freely propagating $2f_0$ waves is addressed by applying the map $\partial_t \mapsto -2if_0 + \partial_{\bar{t}}$ to (3.24)–(3.27) prior to deriving (3.29). This procedure installs a time derivative in the $2f_0$ evolution equation (3.29) and fixes its resonance problem. We leave the details for appendix A and report the resulting modified $2f_0$ equation:

$$(\Delta + 13L)B_t + 4if_0(\Delta - 3L)B = -\tfrac{3}{2}(\partial_x - i\partial_y)^2(LA)^2. \tag{5.3}$$

Non-dimensionalizing (5.3) in the manner of §2.1 reveals that $(\Delta + 13L)B_t$ is ϵ^2 smaller than the rest of (5.3). The small term $(\Delta + 13L)B_t$ becomes important under conditions of near resonance when $4if_0(\Delta - 3L)B$ is relatively small.

The expansion of ∂_t used to derive (5.3) is an application of the 'method of reconstitution' (Roberts 1985). Reconstitution successfully improves many asymptotic expansions, including the Navier–Stokes equations. Here, reconstitution of $2f_0$ dynamics by addition of the high-order term $(\Delta + 13L)B_t$ empowers (5.3) to describe freely propagating $2f_0$ waves.

6. Conservation laws

Like XV, we find that the wave-averaged system (1.6)–(1.9) conserves two integral quantities, which we call 'wave action' and 'coupled energy'.

6.1. *Wave action*

The first conservation law follows from the wave equations (1.8) and (1.9). We multiply (1.8) with $(1/2)LA^*$, add the complex conjugate and integrate over the domain. Using the $2f_0$ equation (5.3), and a liberal application of integration by parts, we find

$$\frac{d}{dt} \int \frac{1}{2}|LA|^2 + \frac{1}{6}|\nabla B|^2 + \frac{13f_0^2}{6N^2}|B_z|^2 \, dV = 0. \tag{6.1}$$

The appearance of the B terms in this first conservation law is a consequence of the time derivative in the $2f_0$ equation (5.3) and corresponds to the total energy in the freely propagating part of the near-$2f_0$ wave field. The conservation law (6.1) thus implies that resonant generation of propagating $2f_0$ waves extracts near-inertial kinetic energy.

6.2. *Coupled energy*

The second conserved quantity is a wave–mean coupled energy. We derive the associated conservation law by multiplying the potential vorticity equation (4.1) with ψ and integrating over the domain. The Jacobian term $\psi J(\psi, q)$ can be written as an exact derivative and integrates to zero. Applying integration by parts, we are left with

$$\frac{d\mathscr{E}_\psi}{dt} = \int \psi q_t^w \, dV, \tag{6.2}$$

where q^w is the wave potential vorticity defined in (4.2) and

$$\mathscr{E}_\psi \overset{\text{def}}{=} \int \frac{1}{2}|\nabla \psi|^2 + \frac{1}{2}\frac{f_0^2}{N^2}\psi_z^2 \, dV \tag{6.3}$$

is the total balanced QG energy. Next, we multiply (1.8) by $iLA_t^*/2f_0$, add the complex conjugate and integrate over the domain to obtain

$$\frac{d\mathscr{E}_f}{dt} = -\int \psi q_t^w \, dV - \frac{i}{2f_0} \int B^* \partial_t \partial_s^2 (LA)^2 - B\partial_t \partial_{s*}^2 (LA^*)^2 \, dV, \tag{6.4}$$

where

$$\mathscr{E}_f \overset{\text{def}}{=} \int \frac{f_0^2}{4N^2}|\nabla A_z|^2 + \frac{\beta y}{2f_0}|LA|^2 \, dV \tag{6.5}$$

is the sum of the NIW potential energy and an action-like term associated with the β-effect. The first term on the right of (6.4) corresponds to the term on the right of (6.2) and will cancel when these equations are added. Substitution of the $2f_0$ equation (1.9) and its complex conjugate into the second integral on the right of (6.4) followed by persistent integration by parts produces

$$\frac{d\mathscr{E}_{2f}}{dt} = \frac{i}{2f_0} \int B^* \partial_t \partial_s^2 (LA)^2 - B \partial_t \partial_{s*}^2 (LA^*)^2 \, dV, \tag{6.6}$$

where

$$\mathscr{E}_{2f} \overset{\text{def}}{=} \int \frac{i}{12f_0}[B(\Delta + 13L)B_t^* - B^*(\Delta + 13L)B_t] - \frac{1}{3}|\nabla B|^2 + \frac{f_0^2}{N^2}|B_z|^2 \, dV, \tag{6.7}$$

$$= \int \frac{i}{8f_0}[B^*(\partial_x - i\partial_y)^2(LA)^2 - B(\partial_x + i\partial_y)^2(LA^*)^2] + \frac{1}{3}|\nabla B|^2 - \frac{f_0^2}{N^2}|B_z|^2 \, dV. \tag{6.8}$$

An additional substitution of (1.9) and its complex conjugate transforms equation (6.7) into (6.8). The conservation of coupled energy emerges from the combination of (6.2), (6.4) and (6.6),

$$\frac{d}{dt}(\mathscr{E}_\psi + \mathscr{E}_f + \mathscr{E}_{2f}) = 0. \tag{6.9}$$

The conservation law in (6.9) is identical to XV's conservation law except for the addition of the $\mathscr{E}_{2f}$ term. Equation (6.9) is the analogue of a conservation law found by Danioux, Vanneste & Bühler (2015) that relates the evolution of NIW potential energy to advection and refraction by steady geostrophic flows.

A thought experiment due to XV illuminates an important implication of (6.9). Envision the rapid and stormy deposition of a horizontally extensive surface-concentrated current in a region of geostrophic turbulence. As the storm passes, the unbalanced current develops into a surface-concentrated NIW that is almost horizontally uniform, and therefore has little potential energy, and thus $\mathscr{E}_f \approx 0$. Over subsequent inertial periods, NIW refraction and advection by the geostrophic flow drives the creation of near-inertial horizontal scales and potential energy, which leads to vertical propagation of a refracted NIW signal and the generation of $2f_0$ internal waves. Because the wave action in (6.1) and coupled energy in (6.9) are distinct and independent conservation laws, the total NIW and $2f_0$ wave energy increases in this process at the sole expense of energy in the geostrophic flow. The role of $\mathscr{E}_{2f}$ in (6.9) is unfortunately obscure in this thought experiment, although the diagnosis of (6.9) presented in figure 7 below shows that the effect is minor in some cases.

7. Comparison of the three-component model and the Boussinesq equations

To build confidence in the heuristic and asymptotic approximations used to develop the three-component model, we compare numerical solutions of a two-dimensional initial value problem in the three-component model and the Boussinesq equations. The initial problem is similar to that shown in figure 1, in which a surface-concentrated NIW interacts with a barotropic balanced velocity field. In addition to solutions intended for direct comparison, we compute solutions to a two-component model without $2f_0$ dynamics, and a three-component model with the PSI-like part of the NIW–$2f_0$ interaction removed. The physical implications of the numerical solutions are discussed in § 8.

7.1. *A surface-concentrated NIW in random barotropic flow*

The initial value problem involves the interaction of a surface-concentrated NIW with random barotropic balanced flow. The two-dimensional physical domain is bounded by rigid lids in z with height $H = 4$ km and is periodic in x with width $L = 800$ km. The stratification is uniform, with buoyancy frequency $N = 2 \times 10^{-3}$ s^{-1}, and the inertial frequency is $f_0 = 10^{-4}$ s$^{-1} = N/20$, with $\beta = 0$. As for the problem considered in § 1, the NIW initial condition is

$$L A(x, z, 0) = u(x, z, 0) = U_0 \exp(-z^2/2h^2), \tag{7.1}$$

with $h = 100$ m. We consider initial NIW surface velocities of $U_0 = 0.4$, 0.2 and 0.1 m s^{-1}.

The initial balanced v velocity is

$$\psi_x(x, z, 0) = v(x, z, 0) = V_0 \sum_{n=4}^{14} \left(\frac{k_4}{k_n}\right)^2 \cos(k_n x + \phi_n), \tag{7.2}$$
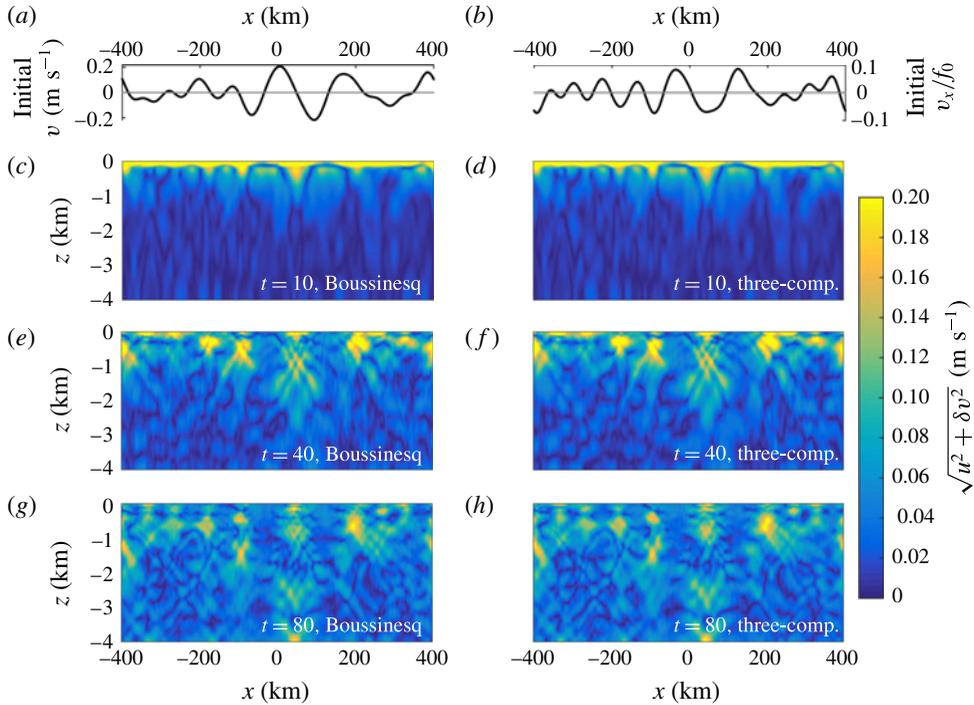
FIGURE 2. (Colour online) Comparison of wave speed in numerical solutions to the three-component and Boussinesq models. (a,b) The x dependence of the initial balanced velocity v (a) and the balanced vorticity normalized by f_0, v_x/f_0 (b). $(c–h)$ The wave speed defined in (7.8) at $t = 10$, 40 and 80 inertial periods in the Boussinesq model (c,e,g) and the three-component model (d,f,h). The initial NIW surface velocity is $U_0 = 0.4$ m s^{-1}.

where $k_n \stackrel{\text{def}}{=} 2\pi n/L$ and the ϕ_n are random phases between 0 and 2π for each component of the geostrophic flow. We choose $V_0 = 0.1$ m s^{-1} for all simulations. This produces a maximum velocity of $\max(v) \approx 0.2$ m s^{-1}, a maximum Rossby number of $\max(v_x)/f_0 \approx 0.1$ and a root-mean-square Rossby number of r.m.s.$(v_x)/f_0 \approx 0.05$. The balanced flow and its associated vorticity field are plotted in figures 2(a,b)–4(a,b).

The numerical solutions we report are listed in table 1. We choose simulation parameters both for ease of numerical integration and for consistency with oceanic scenarios. In particular, $V_0 = 0.1$ m s^{-1} leads to 'reasonable' NIW vertical propagation in 10 or 20 inertial periods; smaller values of V_0 lead to slower vertical propagation. It should be noted, however, that $V_0 = 0.1$ m s^{-1} violates assumptions made in § 2 to justify the asymptotic derivation of the three-component model: the expansion assumes that $\bar{U}/\tilde{U} = \tilde{H}/\bar{H} = \epsilon \ll 1$. In the following simulations, $\tilde{H}/\bar{H} = \pi h/H = 0.08 \ll 1$ while $\bar{U}/\tilde{U} = V_0/U_0$ ranges between 0.25 and 1, so that the balanced flow seems to be too strong given that the non-dimensionalization in § 2.1 assumes $\tilde{H}/\bar{H} \sim \bar{U}/\tilde{U} = \epsilon \ll 1$. Nevertheless, we find generally good agreement between the predictions of the three-component model and full Boussinesq solutions. Other calculations with V_0 reduced to 0.025 m s^{-1} not shown here also validate the three-component model.

U_0 (m s^{-1})	Resolution	Model(s)	Notes
0.4	1× and 2×	Boussinesq and three-component model	
0.4	1×	Two-component model with $B \mapsto 0$	'no $2f_0$'
0.2	1×	Boussinesq and three-component model	
0.2	1×	Two-component model with $B \mapsto 0$	'no $2f_0$'
0.2	1×	Equations (7.3), (7.5) and (7.4) with $B_{xx}LA^* \mapsto 0$	'no PSI'
0.1	1×	Boussinesq and three-component model	

TABLE 1. Parameters and models for the numerical simulations reported in §§ 7 and 8. A resolution of '1×' is $nx \times nz = 1024 \times 2048$, and '2×' is twice that. In all runs $\psi_x^q = v(t=0)$ is given by (7.2) with $V_0 = 0.1$ m s^{-1}.

We note finally that that our two-dimensional computations do not test all aspects of the three-component model: with $\partial_y = 0$ the advective terms $J(\psi, LA)$ and $J(\psi, q)$ vanish and the available potential vorticity (Wagner & Young 2015) q does not evolve from its initial distribution. The two-dimensional simulations reported here are interesting primarily as a test of the nonlinear NIW–$2f$ interaction and the associated production of small vertical scales. Probing of the advective terms requires either three-dimensional solutions like those in Danioux *et al.* (2008) or solutions of the two-dimensional model introduced in § 6.2 of XV.

7.2. *Methods*

In two dimensions, the APV equation (1.6) reduces to $q_t = 0$ and implies that q is constant. We decompose the balanced streamfunction into $\psi(x, z, t) = \psi^q(x, z) + \psi^w(x, z, t)$, where

$$(\partial_x^2 + L)\psi^q = q \quad \text{and} \quad (\partial_x^2 + L)\psi^w = -\frac{1}{4f_0}\partial_x^2|LA|^2. \qquad (7.3a,b)$$

Like q, ψ^q is constant in time and is determined by the initial condition. Because LA is initially uniform, we have $\psi^q = \psi(t=0)$ and thus $\psi_{xx}^q = v_x(t=0)$. With this decomposition the two-dimensional three-component system becomes

$$LA_t + \frac{i}{2}f_0 A_{xx} + \frac{i}{2}(\psi_{xx}^q + \psi_{xx}^w)LA + \frac{1}{2}B_{xx}LA^* = -D(LA), \qquad (7.4)$$

$$(\partial_x^2 + 13L)B_t + 4if_0(\partial_x^2 - 3L)B + \frac{3}{2}\partial_x^2(LA)^2 = -D(LB), \qquad (7.5)$$

where the linear 'hyperdiffusion' operator

$$D \stackrel{\text{def}}{=} \nu \left[\left(\frac{\delta x}{\delta z}\right)^2 \partial_x^2 + \partial_z^2 \right]^8 \qquad (7.6)$$

helps to ensure numerical stability. In (7.6) ν is the hyperviscosity and $(\delta x, \delta z)$ is the resolution in (x, z). We set $\nu = 10^6$ m^{16} s^{-1} for all simulations reported here and find that the fractional energy lost to dissipation is negligible.

Equations (7.3)–(7.5) are solved with a pseudospectral method by decomposing A and B into the constant-N vertical modes $\cos(n\pi z/H)$ in z, and Fourier modes in x.

Fast Fourier transforms are used for vertical and horizontal modal projections. Time integration of (7.4) and (7.5) is performed with the exponential time differencing method described by Cox & Matthews (2002), Kassam & Trefethen (2005) and Grooms & Julien (2011). The exponential time differencing method is crucial for integrating (7.4) and (7.5) efficiently due to the stiffness generated by both hyperdiffusion at small scales and the term $(\partial_x^2 - 3L)B$ in (7.5) at large scales.

A subtlety of the vertical mode decomposition of A emerges because we set the barotropic vertically uniform modes of A to zero. For the vertically uniform and x-dependent parts of A, this helps to ensure that w_0 in (3.8) vanishes at the top and bottom boundaries. However, the vertically and horizontally uniform part of A does not have this constraint: this 'domain mode' corresponds to a pure inertial oscillation with no spatial structure or evolution that exactly solves the volume-integrated Boussinesq equations in (2.2)–(2.6). The domain mode might be regarded as a covert and constant 'fourth component' that can be eliminated by posing an initial condition with no net momentum. In the results presented here the complex velocity of the domain mode is $e^{-if_0 t}(HL)^{-1} \int LA(x, z, 0) \, dx \, dz$, where $LA(x, z, 0)$ is the initial NIW amplitude given in (7.1). Because the domain mode is included in our Boussinesq solutions, we add it to the horizontal velocity of the numerical solution of (7.4) and (7.5) shown in figures 2 and 8. The domain mode makes a small but discernible impact on the solution at $t = 10$ in figure 2.

The non-hydrostatic Boussinesq equations in (2.2)–(2.6) with $\beta = 0$ are solved with the model of Winters *et al.* (2004), which employs a pseudospectral method with Fourier horizontal modes, sine vertical modes for w, b, cosine vertical modes for u, v, and an integrating factor method with a third-order Adams–Bashforth scheme for time stepping.

We use the same order of hyperdiffusion for the three-component and Boussinesq models. Non-exhaustive trial and error indicates that our three-component code is stable with time steps at least 10 times larger than those demanded by Winters' Boussinesq model. The simulations reported here use 1024 Fourier modes in x and 2048 vertical cosine modes in z. To test the dependence on resolution, we ran simulations with double the resolution for $U_0 = 0.4$ m s^{-1} in both the Boussinesq and three-component models. The results are almost identical for the two resolutions.

7.3. *Points of comparison*

We use the horizontal velocity, vertical velocity and domain-integrated vertical kinetic energy to compare the Boussinesq and three-component models. Because v is initially balanced, the unbalanced part of v is approximately isolated with

$$\delta v(x, z, t) = v(x, z, t) - v(x, z, 0). \tag{7.7}$$

On the other hand, u is unbalanced because $p_y = 0$. We thus define the unbalanced horizontal 'wave speed' as

$$\text{wave speed} \stackrel{\text{def}}{=} \sqrt{u^2 + \delta v^2}. \tag{7.8}$$

The wave speed in (7.8) includes NIW and $2f_0$ components as well as a much smaller wave-induced mean component. In figure 2, we compare the wave speed from the Boussinesq solution with $|\tilde{\mathcal{U}}| \approx |\mathcal{U}_0 + \tilde{\mathcal{U}}_1|$ diagnosed from the three-component solution, where $\mathcal{U}_0 = e^{-if_0 t}LA$ and $\tilde{\mathcal{U}}_1$ is the wavy part of (3.31). The comparison is
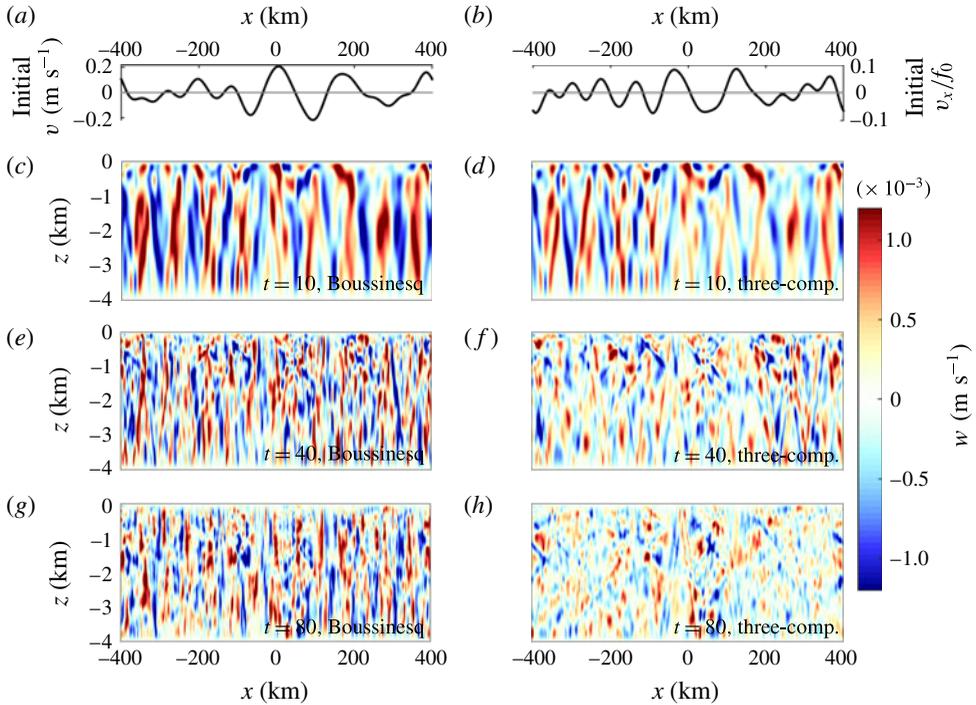
FIGURE 3. (Colour online) Comparison of the vertical velocity in numerical solutions with the Boussinesq and three-component models. (*a,b*) The *x* dependence of the initial *v* (*a*) and v_x/f_0 (*b*). (*c–h*) The vertical velocity *w* at $t = 10$, 40 and 80 inertial periods for the Boussinesq model (*c,e,g*) and the three-component model (*d,f,h*). The initial NIW surface velocity is $U_0 = 0.4$ m s^{-1}.

made at $t = 10$, 40 and 80 inertial periods. The initial NIW magnitude in figure 2 is $U_0 = 0.4$ m s^{-1}, and the initial balanced barotropic *v* and local Rossby number v_x/f_0 are plotted in (*a*) and (*b*).

The wave speed shown in figure 2 indicates good agreement between the three-component model and the Boussinesq equations. A close inspection of the fields is required to discern differences that arise between the two models at late times. It is our consistent experience that the wave speed field is well estimated by the three-component model for the two-dimensional initial value problems examined here; we therefore focus the following discussion on the more interesting and worst-case comparison of vertical velocity.

The vertical velocities in the Boussinesq and three-component solutions are compared in figures 3 and 4 for initial NIW magnitudes of $U_0 = 0.4$ and 0.2 m s^{-1}. Vertical velocity is plotted from (*c,d*) to (*g,h*) at $t = 10$, 40 and 80 inertial periods. For both cases, the agreement is good at $t = 10$ inertial periods but degrades progressively thereafter. Conspicuous aspects of the Boussinesq solution absent from the three-component solution are features with small horizontal scales and steep characteristic angles. These features are especially prominent in figure 3 for the most nonlinear case with $U_0 = 0.4$ m s^{-1} at $t = 40$ and 80 inertial periods.

We dissect this failure of the three-component model in figure 5, which compares vertical kinetic energy (VKE) spectra between the three-component and Boussinesq
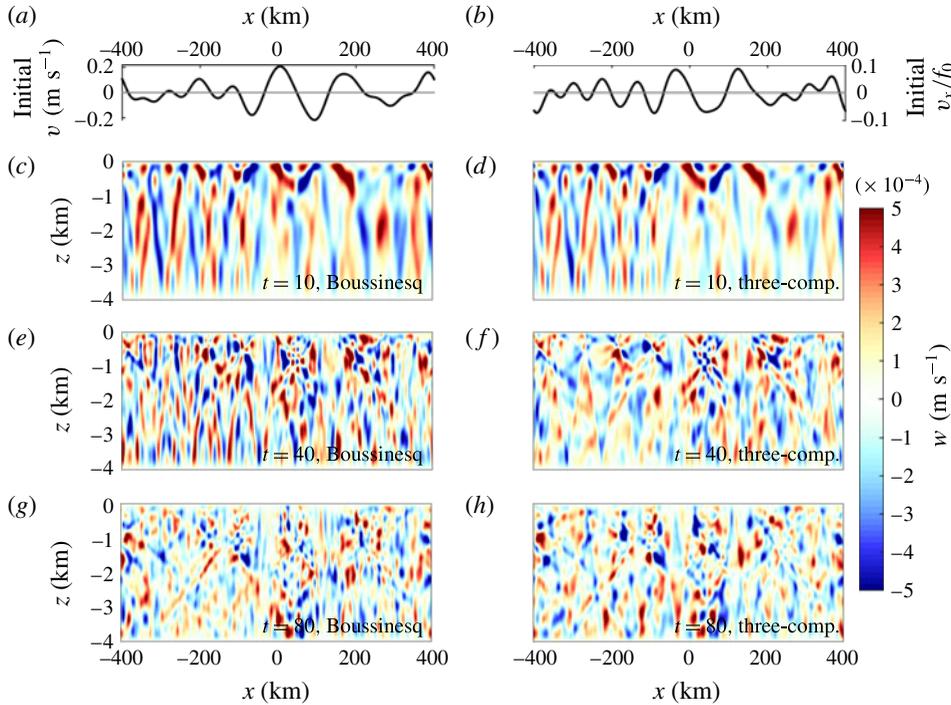
FIGURE 4. (Colour online) Like figure 3 but with initial NIW surface velocity $U_0 = 0.2$ m s^{-1}. The agreement between the Boussinesq and three-component models is better than for $U_0 = 0.4$ m s^{-1}.

models for $U_0 = 0.4$ m s^{-1} at $t = 10$ and 40 inertial periods. The five lines indicate internal wave frequencies based on the linear dispersion relation; proceeding clockwise from the vertical axes these frequencies are $1.01f_0$, $1.08f_0$, $2f_0$, $3f_0$ and $4f_0$, with the dashed line corresponding to $2f_0$. The dynamics are clear: in the Boussinesq simulations, substantial VKE leaks into higher harmonic frequencies $3f_0$ and $4f_0$. By $t = 40$ inertial periods, the fraction of VKE contained in frequencies greater than $2.8f_0$ is 49 %. This transfer of VKE to higher harmonics decreases with U_0: for $U_0 = 0.2$ and 0.1 m s^{-1}, the fraction is 10 % and just over 1 % respectively at $t = 40$ inertial periods.

The effect of the energy transfer to NIW harmonics on total VKE is demonstrated in figure 6, which shows the evolution of total VKE, $\int w^2/2 \, dx \, dz$, for (a) $U_0 = 0.4$ m s^{-1} and (b) $U_0 = 0.2$ m s^{-1}. Four models are considered: Boussinesq (solid lines), three-component model (dashed lines), a two-component model that neglects $2f_0$ (dash-dotted lines), and a modification of the three-component model with PSI suppressed by removing $B_{xx}LA^*$ from the NIW equation (7.4) (dotted line, figure 6(b) only). The three-component model underestimates the amplitude of the VKE, having 54 % of the Boussinesq solution at $t = 40$ inertial periods and 43 % of the total at $t = 80$ inertial periods. The 'extra' Boussinesq VKE is thus similar to that contained in frequencies greater than $2.8f_0$, implying that it originates in a transfer of horizontal NIW kinetic energy to high NIW harmonics not accounted for in the three-component model. It is not surprising that this transfer to $3f_0$ and $4f_0$ is strongest in the most nonlinear case with $U_0 = 0.4$ m s^{-1}.
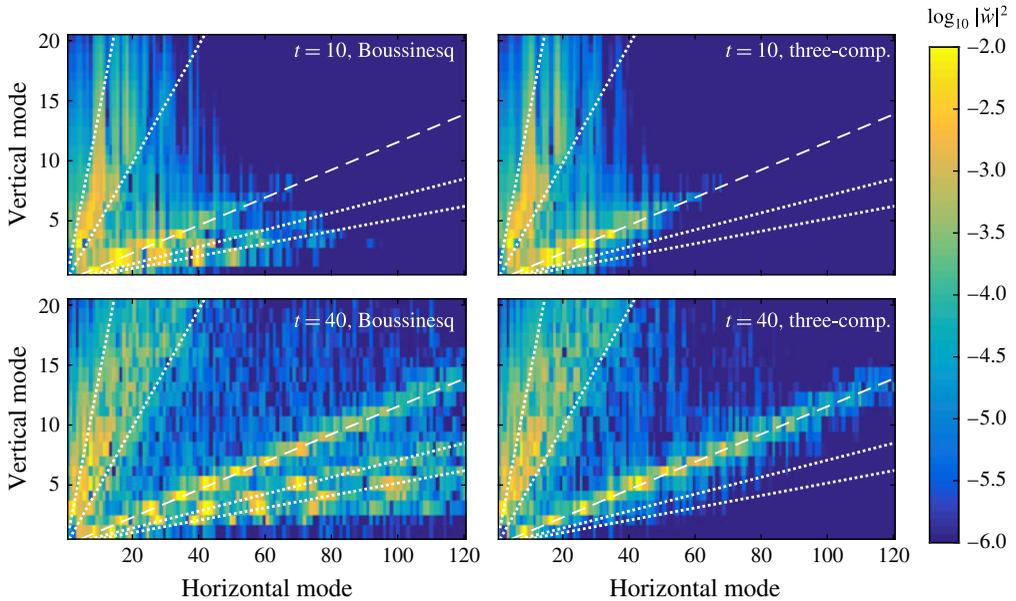
FIGURE 5. (Colour online) Snapshots of VKE spectral components, $|\check{w}|^2$, where $\check{w}$ denotes the Fourier and vertical mode transform of w, for $U_0 = 0.4$ m s^{-1} at $t = 10$ and 40 inertial periods. The spectral components are normalized by the total Boussinesq VKE, and horizontal modes include energy from both positive and negative horizontal wavenumbers. The five lines show the linear dispersion relation for five internal wave frequencies; proceeding clockwise from the vertical axes these frequencies are $1.01f_0$, $1.08f_0$, $2f_0$, $3f_0$ and $4f_0$, with the dashed line corresponding to $2f_0$. By $t = 40$ inertial periods, 49 % of the Boussinesq VKE is in frequencies higher than $2.8f_0$.

For the case $U_0 = 0.2$ m s^{-1} the three-component model correctly estimates the amplitude, but not the phase of the VKE. Unsurprisingly, given the impact of NIW–harmonic interactions on the VKE, the two-component solutions with $B \mapsto 0$ and thus no $2f_0$ cannot capture the evolution of the VKE for either $U_0 = 0.4$ or 0.2 m s^{-1}. In figure 6(b), the suppression of PSI leads to an unrealistic accumulation of VKE in $2f_0$ motions starting at approximately $t = 20$ inertial periods. This indicates that the transfer of energy from $2f_0$ back to NIWs must be accounted for to accurately capture VKE evolution.

7.4. *Summary*

The comparison presented in this section shows that the three-component model well describes NIW evolution and nonlinear NIW–$2f_0$ interaction. That the three-component model describes NIW evolution in the cases shown here is not too surprising, since it is probably driven by a linearized YBJ-type flow-induced refraction. The success of the three-component model in describing NIW–$2f_0$ interaction is more surprising and vindicates the heuristic derivation of $2f_0$ dynamics. On the other hand, the model grossly underestimates the vertical velocity magnitude when the NIWs are strong, which follows from the neglect of NIW harmonics higher than $2f_0$. We stress that this two-dimensional comparison cannot test whether the three-component model correctly captures the impact of NIWs on balanced flow evolution.
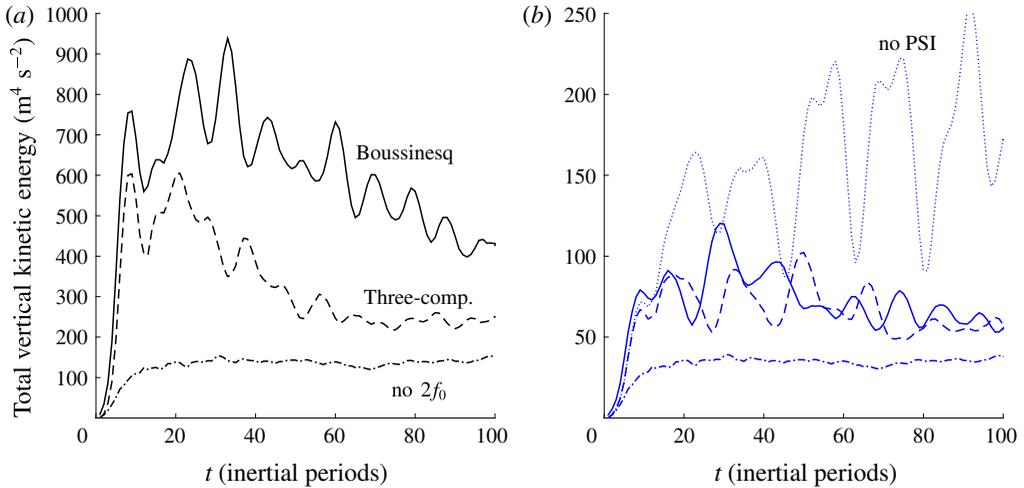
FIGURE 6. (Colour online) Evolution of total VKE, $\int w^2/2 \, \mathrm{d}x \, \mathrm{d}z$, for (*a*) $U_0 = 0.4$ m s^{-1} and (*b*) $U_0 = 0.2$ m s^{-1}. Results are diagnosed from the Boussinesq model (solid lines), the three-component model (dashed lines), a 'no $2f_0$' two-component model with $B \mapsto 0$ (dash-dotted lines) and a 'no PSI' three-component model with the term $B_{xx}LA^*$ removed from the NIW equation (7.4) (dotted line in (*b*) only). Black colours are used for (*a*) $U_0 = 0.4$ m s^{-1} and blue colours for (*b*) $U_0 = 0.2$ m s^{-1} here and in figures 7 and 9.
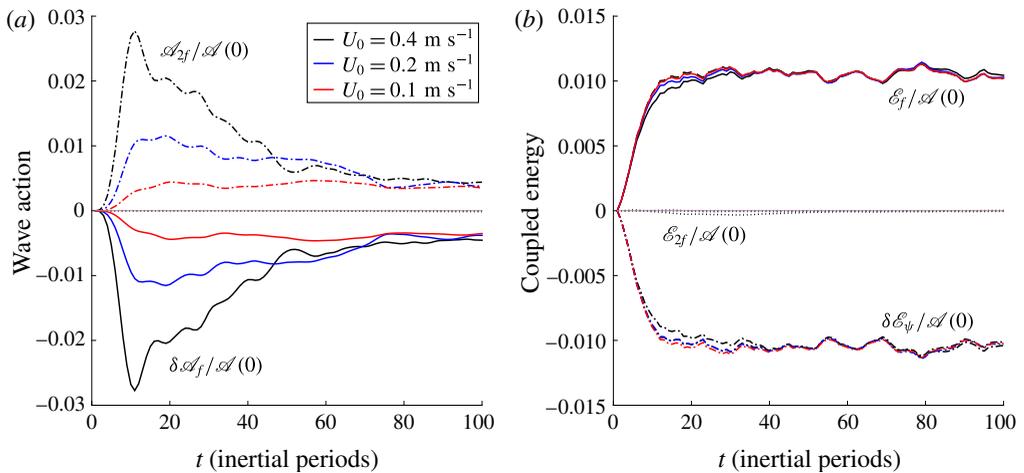

FIGURE 7. (Colour online) The evolution of (*a*) wave action and (*b*) coupled energy in the three-component system with initial NIW velocity in (7.1) and $U_0 = 0.4$, 0.2 and 0.1 m s^{-1}, and initial balanced velocity in (7.2) with $V_0 = 0.1$ m s^{-1}, as shown in figures 2–4.

8. Energy transfer and production of small vertical scales

In this section we continue to explore the initial value problem of §7 by looking at the energy transfer between the three flow components and the surprising role played by $2f_0$ in the evolution of the smallest vertical scales.
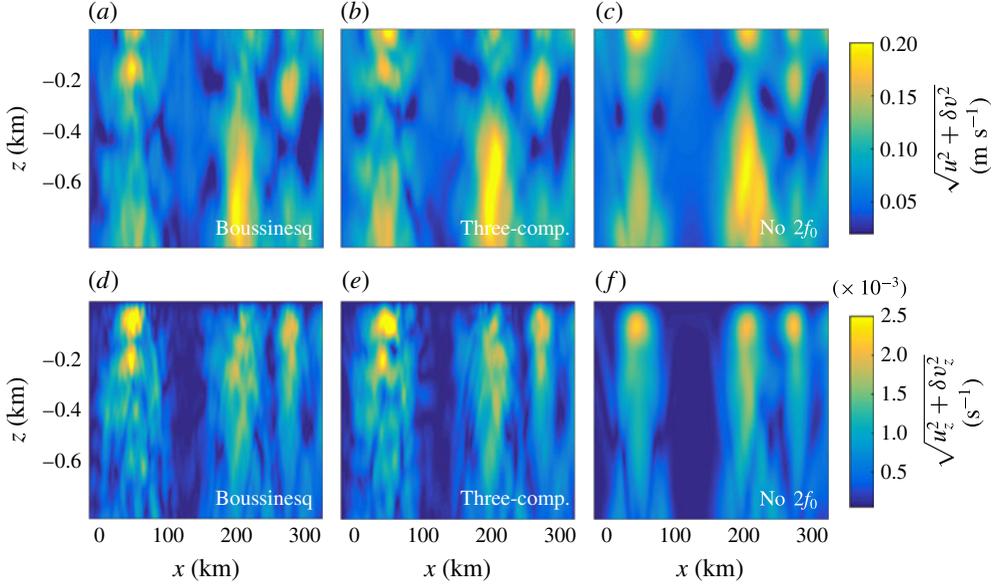
FIGURE 8. (Colour online) Comparison of velocity magnitude and shear magnitude between the Boussinesq equations, the three-component model and a two-component model with $B \mapsto 0$ and thus no $2f_0$ waves. The snapshots are taken at $t = 80$ inertial periods and comprise a portion of the full domain shown in figure 2.

8.1. *Energy transfer between flow components*

The two conserved quantities in the three-component model are wave action and coupled energy defined in (6.1) and (6.9) and plotted in figure 7(*a*,*b*). Figure 7(*a*) illustrates the transfer between NIW kinetic energy and the total energy of the $2f_0$ field, defined respectively as

$$\mathscr{A}_f = \int \frac{1}{2}|\mathrm{L}A|^2 \,\mathrm{d}x\,\mathrm{d}z \quad \text{and} \quad \mathscr{A}_{2f} = \int \frac{1}{6}|B_x|^2 + \frac{13f_0^2}{6N^2}|B_z|^2 \,\mathrm{d}x\,\mathrm{d}z. \tag{8.1a,b}$$

Figure 7(*a*) shows the components of wave action change $\delta\mathscr{A}_f(t) \overset{\text{def}}{=} \mathscr{A}_f(t) - \mathscr{A}(0)$ and $\mathscr{A}_{2f}$. Figure 7(*a*) also shows the very small change in total wave action $\delta\mathscr{A} = \delta\mathscr{A}_f + \mathscr{A}_{2f}$ due to hyper-dissipation with dotted lines. All curves are normalized by the initial wave action $\mathscr{A}(0)$, which is equal to the kinetic energy in the near-inertial initial condition. Three cases corresponding to different initial amplitudes of the NIW are shown: $U_0 = 0.1, 0.2$ and 0.4 m s^{-1} in red, blue and black. The action transferred from $\mathscr{A}_f$ to $\mathscr{A}_{2f}$ increases initially to a maximum value and thereafter decays to a constant asymptotic value as $t \to \infty$. Although the short-term maximum transfer increases with the initial NIW amplitude U_0, the fraction as $t \to \infty$ is independent of U_0 and indicates that less than 1 % of the near-inertial action is ultimately transferred to the $2f_0$ field.

Figure 7(*b*) shows the evolution of $\delta\mathscr{E}_\psi(t) \overset{\text{def}}{=} \mathscr{E}_\psi(t) - \mathscr{E}(0)$, $\mathscr{E}_f$ and $\mathscr{E}_{2f}$ following the definitions in (6.3), (6.5) and (6.7) respectively. All energies are normalized by the initial near-inertial kinetic energy $\mathscr{A}(0)$, thus revealing an uncanny correspondence between cases: the energy transferred from balanced flow to NIWs is a constant fraction of the initial NIW kinetic energy, $\mathscr{A}_f(0)$.

8.2. *The* $2f_0$ *motions are a stepping stone to small vertical scales*

The evolution of $\mathscr{A}_{2f}$ in figure 7(a) is unspectacular and suggests that NIW–$2f_0$ interaction is not important because at most a mere 3 % of the initial NIW kinetic energy is transferred to $2f_0$ when $U_0 = 0.4$ m s^{-1}. Yet the possibility for a PSI-type energy transfer from $2f_0$ to NIW hints that the inclusion of $2f_0$ and nonlinear NIW–$2f_0$ interaction may be necessary to capture the production of small NIW vertical scales.

We isolate the effect of this process by computing a 'no $2f_0$' solution of (7.3) and (7.4). In this solution we set $B \mapsto 0$, thus removing $2f_0$ waves and the $2f_0$-mediated transfer of energy. Figure 8 gives a qualitative impression of the results, where the wave speed (a–c) and wave shear magnitude (d–f) are plotted for three model solutions with $U_0 = 0.4$ m s^{-1}: Boussinesq (a,d), three-component model (b,e) and the two-component 'no $2f_0$' solution of (7.3) and (7.4) with $B \mapsto 0$ (c,f). Both the Boussinesq and the three-component results have small scales in the vertical velocity which are lacking when $2f_0$ is removed, and thus must be created by nonlinear NIW–$2f_0$ interaction. Without $2f_0$ the magnitude of the vertical shear is also underestimated near $(x, z) = (-0.1, 40)$ km. At the same time, the overall flow structure agrees between the three models.

A more quantitative estimate of small vertical scales is provided by the metric $Ri^\dagger(t)$, which measures the smallest Richardson numbers and thus the potential for wave breaking and mixing were such processes resolved. Here, $Ri^\dagger$ is defined as the average of the smallest 0.1 % of Richardson numbers:

$$Ri^\dagger(t) \stackrel{\text{def}}{=} \text{mean [smallest 0.1 % of } Ri \text{ values]}, \quad \text{where} \quad Ri \stackrel{\text{def}}{=} \frac{N^2 + b_z}{u_z^2 + v_z^2}. \quad (8.2)$$

The evolution of $Ri^\dagger$ normalized by its initial value is shown in figure 9 for the cases $U_0 = 0.4$ and 0.2 m s^{-1}. Results are compared between the Boussinesq, three-component and two-component models. The comparison reveals that small values of $Ri^\dagger$ and thus small vertical scales are produced by at least two distinct physical mechanisms: $Ri^\dagger$ first decreases to a minimum value between $t = 15$ and 20 inertial periods and rises gradually thereafter. The early-time agreement between all five cases means that $Ri^\dagger$ is controlled by refraction of the NIW field by balanced flow during this first stage.

However, at approximately $t = 40$ inertial periods, the results diverge and $Ri^\dagger$ is smaller for $U_0 = 0.4$ m s^{-1} in both the Boussinesq and the three-component model. It is conspicuous that in the two-component model with $U_0 = 0.4$ m s^{-1}, $Ri^\dagger$ is overestimated and stays close to the more linear $U_0 = 0.2$ m s^{-1} results. At this stage, the smallness of $Ri^\dagger$ and thus small NIW vertical scales in the Boussinesq and three-component models must be controlled by nonlinear NIW–$2f_0$ interaction. Strikingly, and despite the fact that they contain little instantaneous energy, $2f_0$ motions provide a crucial stepping stone through which NIW energy is transferred to small vertical scales. The surprisingly accurate description of this process by the three-component model suggests that it is controlled by the interaction of relatively large-vertical-scale $2f_0$ motions with small-scale NIWs, which figures 2 and 3 show to be well captured by the three-component model.

9. Discussion

We have developed a three-component model for the coupled evolution of NIWs, QG flow and internal waves with frequency near $2f_0$. The three-component model
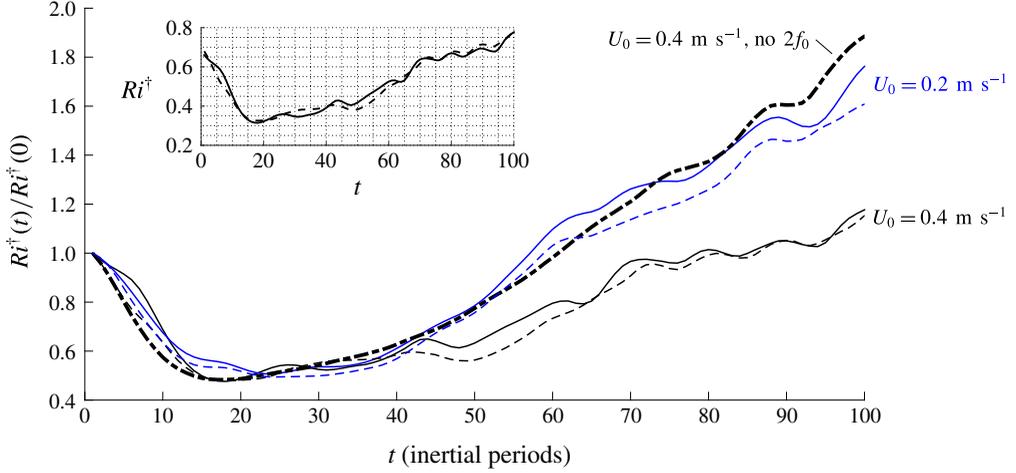
FIGURE 9. (Colour online) The evolution of $Ri^{\dagger}$ defined in (8.2) and normalized by its initial value for the cases $U_0 = 0.4$ m s^{-1} (black) and $U_0 = 0.2$ m s^{-1} (blue or grey) and in three models: Boussinesq (solid lines), three-component model (dashed lines) and the two-component model with $B \mapsto 0$ and thus no $2f_0$ (dashed-dotted lines). The inset shows the numerical values of $Ri^{\dagger}$ approaching the critical value $Ri = 1/4$ for the case $U_0 = 0.4$ m s^{-1} in the Boussinesq and three-component models. The metric $Ri^{\dagger}$ is a measure of the smallest vertical scales in the flow, whose evolution cannot be captured without $2f_0$.

adds $2f_0$ dynamics to the two-component NIW–QG model derived by Xie & Vanneste (2015), and thereby describes the prominent $2f_0$ vertical velocities and production of small NIW vertical scales that numerical solutions of the Boussinesq equations show to be important features of the coupled evolution of NIWs and balanced flow.

A striking prediction of both the three-component model and XV's two-component model is that forced oceanic NIWs extract energy from large-scale balanced flows. Because this requires externally forced internal waves, XV call this mechanism 'stimulated loss of balance', distinguishing it from the spontaneous loss of balance that occurs without external forcing of waves. Stimulated loss of balance acts even in small-Rossby-number flows, and our numerical solutions suggest that energy transfer to NIWs increases with the strength of the externally forced waves. Gertz & Straub (2009) showed that stimulated loss of balance can result in a forward energy cascade and dissipation of wind-driven gyres in an unstratified thin-aspect-ratio fluid. The significance of stimulated loss of balance in real ocean flows is uncertain.

The three-component model connects the $2f_0$ generation mechanism identified by Danioux & Klein (2008) with the YBJ-based near-inertial PSI mechanism of Young *et al.* (2008). The form of the NIW–$2f_0$ coupling implies a two-step cycle for NIW energy. First, advection and refraction by balanced flow catalyses transfer of NIW energy to $2f_0$ waves. These newly produced $2f_0$ waves have large, often depth-spanning, vertical scales and propagate rapidly in the horizontal. Second, a PSI-like interaction transfers energy from $2f_0$ waves back to the NIW field at very small vertical scales. This two-step process provides a path from the large scales of NIW forcing to the small scales of wave breaking and mixing. Advection and refraction of NIWs by non-uniform QG flows leads to relatively small NIW horizontal scales and thus plays a catalytic role in activating this path. Interestingly,

the rapid horizontal and vertical propagation of the nascent $2f_0$ waves can excite small-scale NIWs in regions remote from the initial NIW forcing. Two-dimensional numerical solutions of both the three-component and Boussinesq models give tentative confirmation of this mechanism.

The numerical comparison with the Boussinesq equations shows that the strengths of the three-component model include its description of NIW refraction by balanced flow, and prediction of both the phase and the amplitude of growing $2f_0$ waves at short times. A weakness of the three-component model is the underestimation of the vertical velocity and vertical kinetic energy under increasingly nonlinear conditions due to its neglect of $3f_0$ and $4f_0$ frequency NIW harmonics. Despite this shortcoming, the three-component model captures with surprising accuracy the long-time evolution of the very smallest NIW vertical scales that result from nonlinear NIW–$2f_0$ interaction.

The numerical comparison primarily tests the accuracy of NIW–$2f_0$ dynamics in the three-component model in a regime where refraction by APV-induced balanced flow controls the large-scale NIW evolution. The magnitude of APV and the APV-induced flow mean that our comparison does not isolate the existence and impact of balanced flow induced by quadratic NIW terms in (1.7). In addition, because APV cannot evolve from its initial condition in our two-dimensional scenario, the comparison cannot explore dynamic NIW–QG interaction. A three-dimensional comparison of three-component and Boussinesq dynamics is required to define the regimes of validity of the three-component model in more realistic scenarios and to unravel the effects of NIWs and their wave-induced balanced flow on the evolution of oceanic QG motion.

The applicability of the three-component model to a particular part of the ocean can be assessed using kinetic energy frequency spectra derived from long-term mooring observations of horizontal velocity. Where non-wave flows of NIW scale have small Rossby number, the three-component model well approximates the dynamics of any motion with Eulerian frequencies near f_0. In flows with relative vorticities near or greater than f_0, or under conditions of active wave breaking, the relevance of the three-component model is uncertain. The ubiquitous appearance of a spectral peak at f_0 combined with the belief that large NIW-scale vortical flows are predominantly balanced (Ferrari & Wunsch 2009) hints at, but does not confirm, the potentially broad applicability of the three-component model. Such confirmation requires further observations, such as the difficult simultaneous observation of large-scale balanced vorticity and storm-driven NIW evolution. The applicability of the three-component model to real flows is of consequence because prediction of the climatic evolution of diapycnal mixing likely requires a firm understanding of NIW physics – a link between the large and small scales of oceanic motion.

Appendix A. The $2f_0$ equation

In this appendix, we outline the asymptotic and heuristic steps that lead to the $2f_0$ equation in (1.9) and (5.3).

A.1. *The $2f_0$ frequency forcing at first order*

The vertically averaged time-fluctuating part of the first-order Boussinesq system in (3.24)–(3.27) is

$$\widehat{\widetilde{\mathcal{U}}_{1\,\tilde{t}}} + \mathrm{i}f_0 \widehat{\widetilde{\mathcal{U}}_1} + 2\widetilde{p}_{0s*} = -\mathrm{e}^{-2\mathrm{i}f_0\tilde{t}}\,\widehat{\mathcal{J}}_2, \tag{A 1}$$

$$\widetilde{p}_{0\tilde{z}\tilde{t}} + \widehat{\widetilde{w}_0}\,N^2 = 0, \tag{A 2}$$

$$\widehat{\widetilde{\mathcal{U}}_{1s}} + \widehat{\widetilde{\mathcal{U}}_{1\,s*}^*} + \widehat{\widetilde{w}_{0\tilde{z}}} = 0, \tag{A 3}$$

where the wavy part of the leading-order pressure $\widetilde{p}_0$ does not depend on the fast scale $\tilde{z}$. The system above describes hydrostatic internal waves of general aspect ratio driven by the $2f_0$ forcing on the right of (A 1).

A bit of wrangling with (A 1)–(A 3) leads to a single equation for the wavy part of the leading-order pressure field:

$$\partial_{\tilde{t}}[\,\partial_{\tilde{t}}^2\,\bar{\mathrm{L}} + f_0^2(\Delta + \bar{\mathrm{L}})]\widetilde{p}_0 = 3\mathrm{i}f_0^3(\mathrm{e}^{-2\mathrm{i}f_0\tilde{t}}\,\widehat{\mathcal{J}}_{2s} - \mathrm{e}^{2\mathrm{i}f_0\tilde{t}}\,\widehat{\mathcal{J}}_{2s*}^*). \tag{A 4}$$

Equation (A 4) is the hydrostatic internal wave equation forced at frequency $2f_0$. Writing $\widetilde{p}_0$ as

$$\widetilde{p}_0 = \mathrm{i}f_0[\mathrm{e}^{-2\mathrm{i}f_0\tilde{t}}B(x, y, \bar{z}, \bar{t}) - \mathrm{e}^{2\mathrm{i}f_0\tilde{t}}B^*(x, y, \bar{z}, \bar{t})], \tag{A 5}$$

and noting that (3.16*a,b*) implies

$$\widehat{\mathcal{J}}_2 = \partial_s \widehat{M_{\tilde{z}}^2}, \tag{A 6}$$

we find that B satisfies

$$\mathrm{i}f_0(\Delta - 3\bar{\mathrm{L}})B = -\tfrac{3}{2}\partial_s^2 \widehat{M_{\tilde{z}}^2}. \tag{A 7}$$

A.2. *Resonant and near-resonant NIW–$2f_0$ interaction*

Equation (A 7) describes forced oscillations with frequency $2f_0$. It cannot describe the resonant and near-resonant generation and free propagation of $2f_0$ internal waves. Near-resonant generation can be understood by projecting (A 7) onto vertical modes $h_n(z)$ which satisfy

$$\mathrm{L}h_n + \kappa_n^2 h_n = 0 \quad \text{and} \quad h'_n = 0 \quad \text{at top and bottom}, \tag{A 8a,b}$$

where the eigenvalue κ_n is the Rossby deformation wavenumber of mode n. If we look for solutions of the form $B \sim \mathrm{e}^{\mathrm{i}kx+\mathrm{i}\ell y}h_n(z)$, we find that (A 7) cannot be solved when

$$k^2 + \ell^2 = 3\kappa_n^2. \tag{A 9}$$

These combinations, which are circular slices of (k, ℓ) space at each vertical mode, are the wavenumber combinations that satisfy the linear internal wave dispersion relation at frequency $2f_0$. Freely propagating $2f_0$ internal waves are generated when the NIW forcing $\partial_s^2 \widehat{M_{\tilde{z}}^2}$ has non-zero spectral content near these wavenumber combinations. The generality of near-resonant $2f_0$ generation in NIW–balanced-flow interaction is evident from the results in figure 1 and the simulations in Danioux *et al.* (2008).

As resonant generation is generic, we seek to describe it by modifying (A 7). In particular, we need a term proportional to $B_{\tilde{t}}$ in (A 7) in order to describe

time-dependent B generation and free near-$2f_0$ propagation. We achieve this by applying the map

$$\partial_t \mapsto -2\mathrm{i}f_0 + \partial_{\bar{t}} \tag{A 10}$$

to (A 4) and re-deriving the $2f_0$ equation.

The scaling in §2.1 implies that $\partial_{\bar{t}}$ is ϵ^2 smaller than $2f_0$; thus, in applying (A 10) to (A 4) we ignore the even smaller $O(\epsilon^4)$ terms. Introducing (A 5) into the result then yields

$$-(\Delta - 11\bar{\mathsf{L}})B_{\bar{t}} + 2\mathrm{i}f_0(\Delta - 3\bar{\mathsf{L}})B = -3\partial_s^2\widehat{M_{\bar{z}}^2}. \tag{A 11}$$

The leftmost term is ϵ^2 smaller than $(\Delta - 3\bar{\mathsf{L}})B$ and only becomes important when $(\Delta - 3\bar{\mathsf{L}})B \approx 0$. Moreover, the addition of any multiple of $(\Delta - 3\bar{\mathsf{L}})B_{\bar{t}}$ does not reduce the 'accuracy' of the approximation in (A 11).

We exploit this ambiguity to improve the already approximate form of (A 11). Consider the exact vertical mode-n dispersion relation for linear hydrostatic internal waves,

$$\Sigma = \pm f_0\sqrt{1 + \frac{k^2}{\kappa_n^2}}, \tag{A 12}$$

where $\Sigma(k, \kappa_n)$ is the hydrostatic internal wave frequency, k is the horizontal wavenumber and κ_n is the horizontal wavenumber of the nth vertical mode. The Taylor expansion of the positive root of Σ around $k = \sqrt{3}\kappa_n$ with κ_n fixed is

$$\Sigma = 2f_0 + \frac{\sqrt{3}f_0}{2\kappa_n}(k - \sqrt{3}\kappa_n) + \frac{f_0}{16\kappa_n^2}(k - \sqrt{3}\kappa_n)^2 + \cdots. \tag{A 13}$$

On the other hand, the approximate dispersion relation implied by (A 11) is found by linearizing (A 11), projecting it onto vertical modes and proposing $B \sim \mathrm{e}^{\mathrm{i}kx - \mathrm{i}\sigma\bar{t}}$ so that the frequency of B is $2f_0 + \sigma$. Algebra reveals that $\Sigma_k = \sigma_k$ when $k = \sqrt{3}\kappa_n$ and $\Sigma = 2f_0$. As a consequence, the $2f_0$ approximation in (A 11) produces the correct group velocity.

This feature is preserved under the addition of any multiple of $(\Delta - 3\bar{\mathsf{L}})B_{\bar{t}}$ to (A 11). We use this freedom to increase the accuracy of $2f_0$ linear dispersion in the three-component model. Subtracting $(9/2)(\Delta - 3\bar{\mathsf{L}})B_{\bar{t}}$ from (A 11), we obtain

$$(\Delta + 13\bar{\mathsf{L}})B_{\bar{t}} + 4\mathrm{i}f_0(\Delta - 3\bar{\mathsf{L}})B = -6\partial_s^2\widehat{M_{\bar{z}}^2}. \tag{A 14}$$

The approximate dispersion relation implied by (A 14) is

$$\sigma = 4f_0\frac{k^2 - 3\kappa_n^2}{k^2 + 13\kappa_n^2}, \tag{A 15}$$

which yields $\sigma_{kk} = \Sigma_{kk}$ and means that (A 14) produces the correct near-$2f_0$ group velocity over a range of wavenumbers. Figure 10 compares the exact dispersion relation with the approximate dispersion relations for both the $2f_0$ harmonic component as well as the NIW component, demonstrating the accuracy of our 'Padé' approximation to the $2f_0$ dispersion relation. We use (A 14) to model the $2f_0$ component of flow in the three-component system. It should be noted also that such a 'Padé' approximation can be applied in the same manner to the NIW equation.
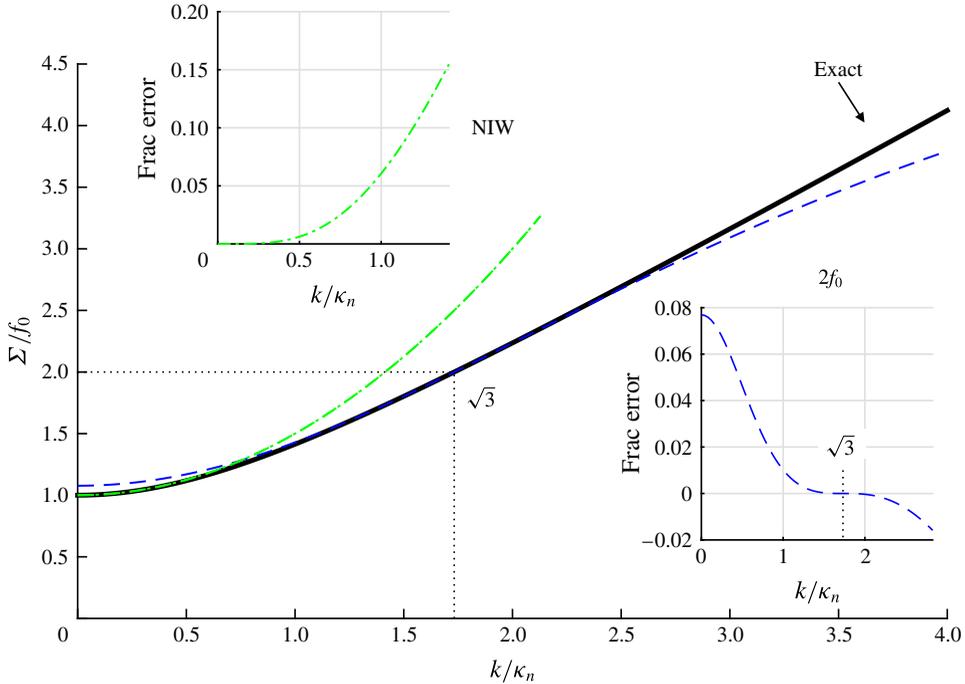
FIGURE 10. (Colour online) Comparison between the exact hydrostatic internal wave dispersion relation and the approximate linear dispersion relations in the three-component model. The thick black line traces the exact hydrostatic internal wave dispersion relation. The green dash-dotted line is the approximate dispersion relation for the NIW component, $f_0(1 + k^2/2\kappa_n^2)$, obtained by linearizing (1.8). The blue dashed line is the approximate dispersion relation for the $2f_0$ component implied by (1.9) and (A 14) and given by $2f_0 + \sigma$ in (A 15). The insets show the fractional errors of the approximate NIW and $2f_0$ dispersion relations.

A.3. *Expressions for $\mathcal{U}_1$ and $\hat{w}_0$*

With $\tilde{p}_0$ defined through B, we can calculate $\tilde{\mathcal{U}}_1$. The vertically averaged horizontal momentum equation is

$$\widehat{\tilde{\mathcal{U}}}_{1\tilde{t}} + \mathrm{i}f_0\widehat{\tilde{\mathcal{U}}}_1 = -2\tilde{p}_{0s*} - \mathrm{e}^{-2\mathrm{i}f_0\tilde{t}}\,\widehat{\mathcal{J}}_2, \tag{A 16}$$

$$= -2\mathrm{i}f_0\mathrm{e}^{-2\mathrm{i}f_0\tilde{t}}B_{s*} + 2\mathrm{i}f_0\mathrm{e}^{2\mathrm{i}f_0\tilde{t}}B_{s*}^* - \mathrm{e}^{-2\mathrm{i}f_0\tilde{t}}\partial_s\widehat{M}_z^2, \tag{A 17}$$

which means that

$$\widehat{\tilde{\mathcal{U}}}_1 = \mathrm{e}^{-2\mathrm{i}f_0\tilde{t}}(2B_{s*} - \mathrm{i}f_0^{-1}\partial_s\widehat{M}_z^2) + \tfrac{2}{3}\mathrm{e}^{2\mathrm{i}f_0\tilde{t}}B_{s*}^*. \tag{A 18}$$

The vertically averaged vertical velocity $\hat{w}_0$ is obtained from (A 2),

$$\hat{w}_0 = -\frac{2f_0^2}{N^2}(\mathrm{e}^{-2\mathrm{i}f_0\tilde{t}}B_{\bar{z}} + \mathrm{e}^{2\mathrm{i}f_0\tilde{t}}B_{\bar{z}}^*). \tag{A 19}$$

With $\hat{w}_0$ we can obtain the full expression for $\tilde{\mathcal{U}}_1$ by solving (3.24), which yields

$$\tilde{\mathcal{U}}_1 = \mathrm{e}^{-2\mathrm{i}f_0\tilde{t}}(2B_{s*} - \mathrm{i}f_0^{-1}\mathcal{J}_2) + \frac{2}{3}\mathrm{e}^{2\mathrm{i}f_0\tilde{t}}B_{s*}^* + \frac{f_0}{N^2}M_{\bar{z}\bar{z}}(\mathrm{e}^{-3\mathrm{i}f_0\tilde{t}}B_{\bar{z}} - \mathrm{e}^{\mathrm{i}f_0\tilde{t}}B_{\bar{z}}^*). \tag{A 20}$$

REFERENCES

BALMFORTH, N. J., SMITH, S. L. & YOUNG, W. R. 1998 Enhanced dispersion of near-inertial waves in an idealized geostrophic flow. *J. Mar. Res.* **56** (1), 1–40.

BALMFORTH, N. J. & YOUNG, W. R. 1999 Radiative damping of near-inertial oscillations in the mixed layer. *J. Mar. Res.* **57** (4), 561–584.

BÜHLER, O. & MCINTYRE, M. E. 1998 On non-dissipative wave–mean interactions in the atmosphere or oceans. *J. Fluid Mech.* **354**, 609–646.

COX, S. M. & MATTHEWS, P. C. 2002 Exponential time differencing for stiff systems. *J. Comput. Phys.* **176** (2), 430–455.

DANIOUX, E. & KLEIN, P. 2008 A resonance mechanism leading to wind-forced motions with a $2f$ frequency. *J. Phys. Oceanogr.* **38** (10), 2322–2329.

DANIOUX, E., KLEIN, P. & RIVIÈRE, P. 2008 Propagation of wind energy into the deep ocean through a fully turbulent mesoscale eddy field. *J. Phys. Oceanogr.* **38** (10), 2224–2241.

DANIOUX, E., VANNESTE, J. & BÜHLER, O. 2015 On the concentration of near-inertial waves in anticyclones. *J. Fluid Mech.* **773**, R2.

D'ASARO, E. A., ERIKSEN, C. C., LEVINE, M. D., PAULSON, C. A., NIILER, P. P. & VAN MEURS, P. 1995 Upper-ocean inertial currents forced by a strong storm. Part I: data and comparisons with linear theory. *J. Phys. Oceanogr.* **25**, 2909–2936.

FALKOVICH, G., KUZNETSOV, E. & MEDVEDEV, S. 1994 Nonlinear interaction between long inertio-gravity and Rossby waves. *Nonlinear Process. Geophys.* **1**, 168–171.

FERRARI, R. & WUNSCH, C. 2009 Ocean circulation kinetic energy: reservoirs, sources, and sinks. *Annu. Rev. Fluid Mech.* **41**, 253–282.

GERTZ, A. & STRAUB, D. N. 2009 Near-inertial oscillations and the damping of midlatitude gyres: a modeling study. *J. Phys. Oceanogr.* **39** (9), 2338–2350.

GROOMS, I. & JULIEN, K. 2011 Linearly implicit methods for nonlinear PDEs with linear dispersion and dissipation. *J. Comput. Phys.* **230** (9), 3630–3650.

KASSAM, A.-K. & TREFETHEN, L. N. 2005 Fourth-order time-stepping for stiff pdes. *SIAM J. Sci. Comput.* **26** (4), 1214–1233.

KLEIN, P. & SMITH, S. L. 2001 Horizontal dispersion of near-inertial oscillations in a turbulent mesoscale eddy field. *J. Mar. Res.* **59**, 697–723.

KLEIN, P., SMITH, S. L. & LAPEYRE, G. 2004 Organization of near-inertial energy by an eddy field. *Q. J. R. Meteorol. Soc.* **130**, 1153–1166.

KUNZE, E. 1985 Near inertial wave propagation in geostrophic shear. *J. Phys. Oceanogr.* **15**, 544–565.

LEE, D.-K. & NIILER, P. P. 1998 The inertial chimney: the near-inertial energy drainage from the ocean surface to the deep layer. *J. Geophys. Res.* **103** (C4), 7579–7591.

MOOERS, C. N. K. 1975 Several effects of a baroclinic current on the cross-stream propagation of inertial-internal waves. *Geophys. Fluid Dyn.* **6**, 245–275.

NIWA, Y. & HIBIYA, T. 1999 Response of the deep ocean internal wave field to traveling midlatitude storms as observed in long-term current measurements. *J. Geophys. Res.* **104** (C5), 10981–10989.

ROBERTS, A. J. 1985 An introduction to the technique of reconstitution. *SIAM J. Math. Anal.* **16** (6), 1243–1257.

VANNESTE, J. 2013 Balance and spontaneous wave generation in geophysical flows. *Annu. Rev. Fluid Mech.* **45**, 147–172.

WAGNER, G. L. & YOUNG, W. R. 2015 Available potential vorticity and wave-averaged quasi-geostrophic flow. *J. Fluid Mech.* **785**, 401–424.

WINTERS, K. B., MACKINNON, J. A. & MILLS, B. 2004 A spectral model for process studies of rotating, density-stratified flows. *J. Atmos. Ocean. Technol.* **21** (1), 69–94.

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.

YOUNG, W. R. & BEN JELLOUL, M. 1997 Propagation of near-inertial oscillations through a geostrophic flow. *J. Mar. Res.* **55** (4), 735–766.

YOUNG, W. R., TSANG, Y.-K. & BALMFORTH, N. J. 2008 Near-inertial parametric subharmonic instability. *J. Fluid Mech.* **607**, 25–49.

ZEITLIN, V., REZNIK, G. M. & BEN JELLOUL, M. 2003 Nonlinear theory of geostrophic adjustment. Part 2. Two-layer and continuously stratified primitive equations. *J. Fluid Mech.* **491**, 207–228.