J. Fluid Mech. (2008), *vol.* 596, *pp.* 201–219.
doi:10.1017/S0022112007009421 Printed in the United Kingdom

Balance in non-hydrostatic rotating stratified turbulence

WILLIAM J. McKIVER AND DAVID G. DRITSCHEL

School of Mathematics and Statistics, University of St Andrews, St Andrews, UK

(Received 7 February 2006 and in revised form 1 October 2007)

It is now well established that two distinct types of motion occur in geophysical turbulence: slow motions associated with potential vorticity advection and fast oscillations due to inertia–gravity waves (or acoustic waves). Many studies have theorized the existence of a flow for which the entire motion is controlled by the potential vorticity (or one 'master variable') – this is known as balance. In real geophysical flows, deviations from balance in the form of inertia–gravity waves or 'imbalance' have often been found to be small. Here we examine the extent to which balance holds in rotating stratified turbulence which is nearly balanced initially.

Using the non-hydrostatic fluid dynamical equations under the Boussinesq approximation, we analyse properties of rotating stratified turbulence spanning a range of Rossby numbers ($Ro \equiv |\zeta|_{\max}/f$) and the frequency ratios ($c \equiv N/f$) where ζ is the relative vertical vorticity, f is the Coriolis frequency and N is the buoyancy frequency. Using a recently introduced diagnostic procedure, called 'optimal potential vorticity balance', we extract the balanced part of the flow in the simulations and assess how the degree of imbalance varies with the above parameters.

We also introduce a new and more efficient procedure, building upon a quasi-geostrophic scaling analysis of the complete non-hydrostatic equations. This 'nonlinear quasi-geostrophic balance' procedure expands the equations of motion to second order in Rossby number but retains the exact (unexpanded) definition of potential vorticity. This proves crucial for obtaining an accurate estimate of balanced motions. In the analysis of rotating stratified turbulence at $Ro \lesssim 1$ and $N/f \gg 1$, this procedure captures a significantly greater fraction of the underlying balance than standard (linear) quasi-geostrophic balance (which is based on the linearized equations about a state of rest). Nonlinear quasi-geostrophic balance also compares well with optimal potential vorticity balance, which captures the greatest fraction of the underlying balance overall.

More fundamentally, the results of these analyses indicate that balance dominates in carefully initialized simulations of freely decaying rotating stratified turbulence up to $O(1)$ Rossby numbers when $N/f \gg 1$. The fluid motion exhibits important quasi-geostrophic features with, in particular, typical height-to-width scale ratios remaining comparable to f/N.

1. Introduction

Fluid motion in the Earth's atmosphere and oceans, and in other planetary atmospheres, is subject to strong constraints arising from rotation and stable density stratification. Over intermediate to large scales rotation and stratification induce layerwise two-dimensional motion, with relatively weak vertical velocities. Moreover, the predominant 'balances' in the equations of motion are between the Coriolis

acceleration and the horizontal pressure gradient (geostrophic), and between the buoyancy and the vertical pressure gradient (hydrostatic). That is, the other terms in the equations tend to be of secondary importance. This state of affairs is known generally as 'balance', and can extend much further than the simple geostrophic and hydrostatic (collectively 'quasi-geostrophic') balance just described (Hoskins, McIntyre & Robertson 1985).

A 'balanced flow' refers to any flow controlled solely by a single 'master variable' (see Charney 1948; Warn *et al.* 1995; Ford, McIntyre & Norton 2000 and references therein). Here, as in Dritschel & Viúdez (2003) and Dritschel & Viúdez (2007), we take the master variable to be the 'potential vorticity' (PV), a materially conserved scalar quantity in the absence of friction and diabatic heating. Then, the balanced flow can be recovered explicitly from the potential vorticity through a process called 'PV inversion' (Hoskins *et al.* 1985). In PV inversion, all of the dynamical and thermodynamical fields are recovered by solving *diagnostic* 'balance relations' (such as geostrophic and hydrostatic) and using the definition of potential vorticity or an approximation thereof, as in 'quasi-geostrophic' theory (cf. Gill 1982).

Balance has been found to be an excellent approximation in rapidly rotating flows having a small 'Rossby number' $Ro \equiv |\zeta|_{\max}/f \ll 1$ and in strongly stratified flows having a small Froude number $Fr \equiv |\boldsymbol{\omega}_h|_{\max}/N \ll 1$ (where ζ and $\boldsymbol{\omega}_h$ are the vertical and horizontal vorticity components). Moreover, stratification should be 'stronger' than rotation, in the sense that $f/N \ll 1$, as observed in the Earth's atmosphere and oceans over a broad range of scales. A consequence of this balance is that vertical motions tend to be much weaker than horizontal motions (Viúdez & Dritschel 2003). Nevertheless, departures from balance appear inevitable. Inertia–gravity waves (IGWs) – what we call the 'imbalance' – can be spontaneously emitted, for instance in response to evolving vortical features, small-scale convection, or interactions with topography (Ford *et al.* 2000; Vanneste & Yavneh 2004; Lane *et al.* 2004; Viúdez & Dritschel 2006; Waite & Bartello 2006).

Measuring this imbalance is tricky, and subtle. What constitutes 'imbalance' depends on one's definition of balance, specifically the choice of the 'balance conditions'. And, alas, there is an infinite multiplicity of these (cf. Mohebalhojeh & Dritschel 2001 in the simplest shallow-water context). Balance appears to lack a precise mathematical definition. Hence, there is no rigorous, or unambiguous, way of separating a general nonlinear flow into its balanced and unbalanced parts.

Recently however, a new procedure called 'Optimal PV (OPV) balance' (Viúdez & Dritschel 2004) was introduced to overcome part of this ambiguity by eliminating the need to choose balance conditions. This procedure obtains balance – or more correctly minimizes the imbalance – by artificially ramping up the PV anomaly carried by each fluid particle over a prescribed time period Δ, starting from a state of rest. The full dynamical equations are used, except for this modification of PV. The practical difficulty is that one must find, iteratively, the 'base configuration' of fluid particles that evolves into the 'target configuration' by the end of the ramp (see § 3.2 below). If the ramping is done sufficiently slowly (for Δ exceeding about 3 inertial periods $T_{ip} = 2\pi/f$), the state of motion at the end of the ramp period exhibits minimal IGW activity (Viúdez & Dritschel 2004, Dritschel & Viúdez 2007). The procedure is not perfect as the 'balance' depends on a parameter Δ, but the procedure does not enforce any specific balance conditions. As a result, some of the resulting balance contains a small amount of imbalance, which cannot be reduced further (see Dritschel & Viúdez 2007 for further details). Nevertheless, OPV balance is a robust procedure permitting one to assess the degree of balance and imbalance in a wide class of flows.

OPV balance is one of many procedures available for this purpose. Previous work, based on the hydrostatic equations (which are already semi-balanced), includes the 'nonlinear normal mode initialization' procedure of Machenhauer (1977), Baer (1977), Baer & Tribbia (1977) and Leith (1980), the 'slaving' procedure of Warn *et al.* (1995) and Bokhove (1997), and the 'QG^{+1}' model of Muraki, Snyder & Rotunno (1999), to name but a few (see also Vallis 1996 and references therein). Here we introduce yet another procedure, which however is based on the reformulation of the non-hydrostatic equations introduced by Dritschel & Viúdez (2003). That reformulation takes PV and the two horizontal components of the ageostrophic vorticity to be the prognostic variables, replacing velocity and buoyancy. This reformulation makes explicit the separation between the leading-order hydrostatic–geostrophic balance and the departure therefrom. Here, the PV represents the balanced part of the flow and the ageostrophic vorticity represents the leading-order 'imbalance', that is the departure from thermal wind balance. Moreover, in this reformulation, the governing equations directly simplify to the quasi-geostrophic (QG) equations in the limit of small Rossby number $Ro \ll 1$.

The new balance procedure, called 'nonlinear quasi-geostrophic' (NQG) balance, is in many ways similar to previous ones: two time derivatives (here on the ageostrophic horizontal vorticity) are removed from the equations, giving rise to a pair of diagnostic balance relations (see Bolin 1955, 1956 and Charney 1955, 1962). This filters inertia–gravity waves (IGWs). What is novel in this context is the use of the unapproximated form of the (Rossby–Ertel) PV. There is no need to approximate PV to eliminate IGWs. Of course, since the balance relations are themselves only accurate to $O(Ro^2)$, one might argue that it is consistent to expand the definition of PV to $O(Ro^2)$, as in the QG^{+1} procedure of Muraki *et al.* (1999). However, the numerical evidence below shows that there is a distinct advantage in using the unapproximated form of the PV, as has been repeatedly demonstrated in the shallow-water context (McIntyre & Norton 2000; Ford *et al.* 2000; Mohebalhojeh & Dritschel 2000, 2001, 2004; Mohebalhojeh 2002).

In the present work, both OPV and NQG balance, together with linear QG balance, are compared in simulations of complex rotating stratified turbulence across a range of Rossby and Froude numbers. The simulations are initiated from a QG one at a mature state of evolution. Specifically, the initial conditions for the non-hydrostatic simulations are found by a PV-ramping procedure like that carried out in OPV balance, to start from a nearly balanced flow. In time, IGWs are excited, and here we quantify the extent to which the above balance procedures can detect them.

The plan of the paper is as follows. In § 2 we describe the problem set-up, recall the reformulation of the governing equations, and develop the NQG balance procedure. Details of the numerical methods and initialization are next provided in § 3. Results are presented in § 4, first by illustrating the evolution of PV, then by analysing the full, balanced and unbalanced components of various dynamical fields. Here, the various balance procedures are compared. The paper finishes with concluding remarks in § 5.

2. Problem formulation

2.1. *The non-hydrostatic model*

We consider an inviscid incompressible rotating stratified fluid under the Boussinesq approximation. The latter requires that the density ρ varies weakly from a mean background value, ρ_0, i.e.

$$\rho(\boldsymbol{x}, t) = \rho_0 + \varrho_z z + \rho'(\boldsymbol{x}, t) \tag{2.1}$$

where $\varrho_z z$ is the mean linear density ($\varrho_z < 0$ is a constant) and $\rho'(\boldsymbol{x}, t)$ is the anomalous density. Neglecting terms of $O(\mu^2)$ where $\mu = (\rho - \rho_0)/\rho_0$, the non-hydrostatic (NH) equations become

$$\dot{\boldsymbol{u}} + f\boldsymbol{k} \times \boldsymbol{u} = -\rho_0^{-1}\nabla\Phi + b\boldsymbol{k}, \tag{2.2a}$$

$$\dot{b} + N^2 w = 0, \tag{2.2b}$$

$$\nabla \cdot \boldsymbol{u} = 0, \tag{2.2c}$$

where $\boldsymbol{u} \equiv (u, v, w)$ is the three-dimensional velocity field, $\dot{(\ )} = \mathrm{D}(\ )/\mathrm{D}t = (\ )_t + \boldsymbol{u}\cdot\nabla(\ )$ denotes the material time derivative (in the rotating frame), Φ is the geopotential, $b \equiv -g\rho'/\rho_0$ is the buoyancy, g is acceleration due to gravity, and $\boldsymbol{k}$ denotes the vertical unit vector. N is the mean buoyancy frequency defined by $N^2 \equiv -g\varrho_z/\rho_0$. Here we take the Coriolis (f) and buoyancy (N) frequencies to be constant.

As in Dritschel & Viúdez (2003) we recast the above set of equations in terms of variables which represent the leading-order hydrostatic–geostrophic balance and the departure therefrom in flows with $Ro \ll 1$ and $f/N \ll 1$. The balanced evolution is controlled by the (dimensionless) PV,

$$\Pi \equiv (\boldsymbol{k} + \boldsymbol{\omega}/f)\cdot(\boldsymbol{k} + \nabla b/N^2) = 1 + \frac{\zeta}{f} + \frac{b_z}{N^2} + \frac{\boldsymbol{\omega}\cdot\nabla b}{fN^2}, \tag{2.3}$$

where $\boldsymbol{\omega} = \nabla \times \boldsymbol{u}$ is the vorticity. The leading-order imbalance is represented by the horizontal part of the vector $\mathscr{A}$, defined by

$$\mathscr{A} = \boldsymbol{\omega}/f + \nabla b/f^2\,. \tag{2.4}$$

The horizontal part of $\mathscr{A}$, denoted $\mathscr{A}_h$, is the dimensionless ageostrophic vorticity. While $\mathscr{A}_h$ is here the leading-order imbalance, it generally contains $O(Ro^2)$ balanced motions (see § 2.2 below). The main point is that $\mathscr{A}_h$ does not contain any balanced motions at $O(Ro)$.

In terms of Π and $\mathscr{A}_h$, the flow evolution is governed by conservation of PV, or of the PV anomaly $\varpi \equiv \Pi - 1$,

$$\dot{\varpi} = 0, \tag{2.5}$$

and by the horizontal ageostrophic vorticity equation,

$$\dot{\mathscr{A}}_h = -f\boldsymbol{k} \times \mathscr{A}_h + (1 - c^2)\nabla_h w + f^{-1}\boldsymbol{\omega}\cdot\nabla\boldsymbol{u}_h + c^2\nabla_h\boldsymbol{u}\cdot\nabla\mathscr{D}, \tag{2.6}$$

where $\mathscr{D} = -b/N^2$ is the isopycnal displacement and $c \equiv N/f$ (see Dritschel & Viúdez 2003 for details).

The original 'primitive' variables $(\boldsymbol{u}, b)$ are recovered by inverting the definitions of ϖ and $\mathscr{A}_h$. This is facilitated by introducing a vector potential $\boldsymbol{\varphi} \equiv (\varphi, \psi, \phi)$, such that

$$\mathscr{A} = \nabla^2\boldsymbol{\varphi}\,. \tag{2.7}$$

Then taking the divergence of (2.4) and inverting the Laplacian leads to

$$\mathscr{D} = -c^{-2}\nabla\cdot\boldsymbol{\varphi} \tag{2.8}$$

whereas taking the curl of (2.4) gives

$$\boldsymbol{u} = -f\nabla \times \boldsymbol{\varphi} \tag{2.9}$$

(assuming all fields have zero domain average and ignoring boundaries). Thus $\boldsymbol{u}$ and $b = -N^2\mathscr{D}$, which are needed to evolve ϖ and $\mathscr{A}_h$, all derive from a single vector potential $\boldsymbol{\varphi}$. This reduction occurs because $\nabla\cdot\boldsymbol{u} = 0$. The horizontal components of $\boldsymbol{\varphi}$

are obtained from

$$\boldsymbol{\varphi}_h = \nabla^{-2}\mathscr{A}_h \tag{2.10}$$

whereas the vertical component is obtained from a double Monge–Ampère equation,

$$\mathscr{L}_{qg}\phi = \varpi + (1-c^{-2})\Theta_z - c^{-2}\mathscr{N}(\boldsymbol{\varphi}) \tag{2.11}$$

where $\mathscr{L}_{qg}$ is the QG operator defined by

$$\mathscr{L}_{qg}\phi \equiv \phi_{xx} + \phi_{yy} + c^{-2}\phi_{zz}, \tag{2.12}$$

$$\Theta \equiv \nabla_h \cdot \boldsymbol{\varphi}_h, \tag{2.13}$$

and

$$\mathscr{N}(\boldsymbol{\varphi}) \equiv \nabla(\nabla\cdot\boldsymbol{\varphi})\cdot[\nabla^2\boldsymbol{\varphi} - \nabla(\nabla\cdot\boldsymbol{\varphi})]. \tag{2.14}$$

The main complication of this reformulated set of equations is the above Monge–Ampère equation, arising from the explicit use of PV. Nevertheless, the set of equations can be discretized numerically and solved efficiently. The PV is represented as contours on isopycnal surfaces (surfaces of constant density) and evolved using Contour Advection (Dritschel & Ambaum 1997; Dritschel & Viúdez 2003).

2.2. *Nonlinear QG balance*

In this section, we apply quasi-geostrophic (QG) scaling to the reformulated non-hydrostatic (NH) equations to obtain ageostrophic balanced fields at $O(Ro^2)$. The characteristic space and time scales are

$$x, y \sim L, \quad z \sim H, \quad t \sim (\epsilon f)^{-1}, \tag{2.15}$$

where here we have introduced $\epsilon \equiv |\varpi|_{\max}$ as a PV-based Rossby number, and where $H/L \sim f/N \equiv c^{-1}$. In addition, the characteristic horizontal and vertical velocity scales are

$$\boldsymbol{u}_h \sim \epsilon f L, \qquad w \sim \epsilon^2 f L/c \tag{2.16}$$

while the characteristic isopycnal displacement scale is

$$\mathscr{D} \sim \epsilon L/c. \tag{2.17}$$

Note that this scaling is equivalent (apart from the definition of Rossby number) to that used in Muraki *et al.* (1999, hereafter referred to as 'MSR99'). However, for clarity, we do not apply this scaling to non-dimensionalize the equations.

The reformulated equations make it straightforward to generalize QG balance to include $O(\epsilon^2)$ terms. First of all note that the dimensionless ageostrophic vorticity $\mathscr{A}_h = O(\epsilon^2)$, because the flow is assumed to be in hydrostatic–geostrophic or 'thermal wind' balance at leading order. We can express this by expanding the vector potential $\boldsymbol{\varphi} \equiv (\varphi, \psi, \phi)$ in ϵ,

$$\boldsymbol{\varphi} = \boldsymbol{\varphi}_1 + \boldsymbol{\varphi}_2 + O(\epsilon^3) \tag{2.18}$$

taking $\boldsymbol{\varphi}_1 = O(\epsilon)$, $\boldsymbol{\varphi}_2 = O(\epsilon^2)$, and by taking $\varphi_1 = \psi_1 = 0$. Inserting this expansion into the NH equations, at $O(\epsilon)$ we obtain the standard (linearly–balanced) QG equations

$$\mathscr{L}_{qg}\phi_1 = \varpi, \tag{2.19a}$$

$$\dot{\varpi} = 0, \tag{2.19b}$$

$$u_1 = -f\phi_{1y}, \qquad v_1 = f\phi_{1x}, \tag{2.19c}$$

for constant f and N. At $O(\epsilon^2)$, there is an infinite multiplicity of conditions which can be imposed to close the system of equations (see Mohebalhojeh 2002, in the related shallow-water context). This multiplicity arises because we use the exact definition of PV in lieu of (2.19a) (see § 2.3 below). We thus have to add another equation to the system to close it. Here, we require that the linear part of the PV vanish at $O(\epsilon^2)$, i.e. $\zeta_2/f + b_{2z}/N^2 = 0$, cf. (2.3). We have tried a variety of other conditions, such as $\phi_2 = 0$, but we have found that this condition captures the greatest proportion of the balance (consistent with the findings in Mohebalhojeh 2002 in the shallow-water context). With this condition, equations (2.6) and (2.11) reduce to

$$c^2 \mathscr{L}_{qg} \boldsymbol{\varphi}_2 + (1 - c^2) \nabla \Theta_2 = \mathscr{S} \tag{2.20}$$

where $\Theta_2 \equiv \nabla_h \cdot \boldsymbol{\varphi}_{h2}$ and the vector $\mathscr{S}$ is defined by $\mathscr{S} \equiv (\mathscr{S}_\varphi, \mathscr{S}_\psi, \mathscr{S}_\phi)$ where

$$\mathscr{S}_\varphi = -2J_{yz}(\phi_{1x}, \phi_{1y}), \tag{2.21a}$$

$$\mathscr{S}_\psi = -2J_{zx}(\phi_{1x}, \phi_{1y}), \tag{2.21b}$$

$$\mathscr{S}_\phi = 0, \tag{2.21c}$$

and where $J_{yz}(a, b) = a_y b_z - a_z b_y$ denotes the Jacobian ($J_{zx}(a, b)$ is analogous). The vertical component of (2.20) together with (2.21c) expresses the condition $\zeta_2/f + b_{2z}/N^2 = 0$.

In a triply periodic domain, these 'Nonlinear QG' (NQG) equations are solved most readily in spectral space to give

$$\hat{\varphi}_2 = \hat{\mathscr{F}}[c^2 \hat{\mathscr{S}}_{\varphi xx} + \hat{\mathscr{S}}_{\varphi yy} + \hat{\mathscr{S}}_{\varphi zz} + (c^2 - 1)\hat{\mathscr{S}}_{\psi xy}], \tag{2.22a}$$

$$\hat{\psi}_2 = \hat{\mathscr{F}}[c^2 \hat{\mathscr{S}}_{\psi yy} + \hat{\mathscr{S}}_{\psi zz} + \hat{\mathscr{S}}_{\psi xx} + (c^2 - 1)\hat{\mathscr{S}}_{\varphi xy}], \tag{2.22b}$$

$$\hat{\phi}_2 = \hat{\mathscr{F}} K^2 (1 - c^2)(\hat{\varphi}_{2xz} + \hat{\psi}_{2yz}), \tag{2.22c}$$

where $\hat{\chi}$ denotes the spectral transform of a field χ, K is the three-dimensional wavenumber defined by $K^2 \equiv k^2 + l^2 + m^2$, and $\hat{\mathscr{F}} = 1/[K^2(c^2(k^2 + l^2) + m^2)]$. Thus knowing the $O(\epsilon)$ fields satisfying (2.19), we can solve for the next order using (2.22).

The analysis outlined above is similar to that carried out in MSR99, except that they effectively expand the definition of PV (see § 2.3 below) and consequently have a non-zero source term $\mathscr{S}_\phi$. Also, in their approach they use a Helmholtz decomposition with the potentials (F, G, Φ) related to the potentials we use by $F = -fc\psi$, $G = fc\varphi$ and $\Phi = f\phi$. Furthermore they work in an isotropic domain, thus going from their procedure to ours $\nabla^2 \to c^2 \mathscr{L}_{qg}$. There are extra terms involving Θ_2 in our equations which do not appear in the QG^{+1} equations. These terms arise because of differences in the definition of the buoyancy. They define the buoyancy as $b = N^2(\varphi_x + \psi_y) + f^2 \phi_z$ whereas we take $b = f^2 \nabla \cdot \boldsymbol{\varphi}$. It can be shown with some manipulation that both the NQG and QG^{+1} potentials yield the same diagnostic relation for the velocity field, i.e.

$$c^2 \mathscr{L}_{qg} \boldsymbol{u}_2 = -f \nabla \times \mathscr{S} \tag{2.23}$$

(though with a different source $\mathscr{S}$). However, our choice of buoyancy has the advantage of allowing us to write the equations in a coordinate-independent form.

2.3. *Iterative method*

The above analysis is incomplete if we insist on using the exact, unexpanded definition of PV. In this case, we replace the QG PV inversion equation (2.19a) by the exact, nonlinear Monge–Ampère equation (2.11). However, (2.11) provides ϕ, not ϕ_1. The latter is recovered by the difference $\phi_1 = \phi - \phi_2$ (correct to $O(\epsilon^2)$), where ϕ_2 is found

from the vertical component of (2.20) or equivalently from (2.22*c*). This implies that the equations at $O(\epsilon)$ (2.19*a*) and $O(\epsilon^2)$ (2.20) are in fact coupled nonlinearly.

In practice, we solve these coupled equations by iteration. For the first iteration, $n=1$, we set both the first- and second-order fields to zero ($\boldsymbol{\varphi}^{n-1}=\boldsymbol{\varphi}^0=\boldsymbol{\varphi}_1^0+\boldsymbol{\varphi}_2^0=0$) and then proceed as follows:

(1) invert the PV (2.11) using $\boldsymbol{\varphi}^{n-1}$ in $\mathscr{N}(\boldsymbol{\varphi})$ and in Θ to obtain ϕ^n;
(2) obtain an approximation to ϕ_1^n from $\phi_1^n=\phi^n-\phi_2^{n-1}$;
(3) use ϕ_1^n in (2.21) and (2.22) to obtain $\boldsymbol{\varphi}_2^n$;
(4) return to (1) if $|\phi_1^n+\phi_2^n-\phi^n|$ is greater than a prescribed tolerance.

Hence, starting with ϖ at the time to be diagnosed, step (1) provides ϕ, which in the first instance is equal to ϕ_1 (since both the first- and second-order fields are initially zero). From ϕ_1, all the next-order fields can be calculated from step (3). Using these fields in (1) again gives a better approximation to the balanced fields to $O(\epsilon^2)$. Then (2) gives an improved approximation to ϕ_1 and this is used in (3) to calculate the next-order fields. In practice, we iterate until the pointwise difference in ϕ^n and $\phi_1^n+\phi_2^n$ is less than 10^{-7}. Note that, upon convergence, $\phi_1+\phi_2=\phi$.

In MSR99 (see also Rotunno, Muraki & Snyder 2000) they apply a similar technique in their QG^{+1} balance model, but in effect only carry out one iteration of the above cycle followed by step (1) again. They first set the fields to zero and do the inversion for ϕ_1 as in (1) and (2) above. Then, they use ϕ_1 in (3) to compute the horizontal part of $\boldsymbol{\varphi}_2^n$, for $n=1$. Finally, they repeat step (1) to obtain ϕ to $O(\epsilon^2)$, except they do not include $\boldsymbol{\varphi}_2^n$ in the right-hand side of (2.11) since this would contribute at $O(\epsilon^3)$. Other differences arise from our inclusion of non-hydrostatic terms, which are expected to be small for $c\gg 1$.

In §4.1 we apply the NQG procedure using different numbers of iterations. Substantial gains in accuracy are found with repeated iteration.

3. Simulations

3.1. *Initialization and parameter settings*

Several simulations were conducted to examine the nature of balance in rotating stratified turbulence at different Rossby numbers. We first conducted a QG simulation (corresponding to $\epsilon\to 0$) which began with an 'isotropic' PV field (after stretching the vertical coordinate by N/f) consisting of spheres of QG PV $q=\pm 4\pi$. This implies a characteristic vortex rotation period of $T_{eddy}=1$ (note: the Rossby number does not enter explicitly in the standard QG equations). Initially, an equal number and volume of cyclonic and anticyclonic vortices were placed randomly (without overlapping) in the domain. Their sizes were picked from a frequently observed power-law number density distribution (Reinaud, Dritschel & Koudella 2003). Figure 1(*a*) shows the initial configuration with 500 vortices of each sign. The ratio between the volume of the largest vortex and the volume of the smallest vortex is 20.

The QG and NH simulations were carried out using identical parameter settings in a triply periodic cube (in stretched coordinates x, y and Nz/f), with a basic grid resolution of $128\times 128\times 128$. Four times as many layers were used to represent the PV on isopycnal surfaces, and for consistency a grid four times finer in each horizontal direction was used in converting the PV contours to gridded PV values, needed in the rest of the algorithm. Contour surgery, which limits the growth in complexity of the contours, was applied at a twentieth of the horizontal grid resolution used to

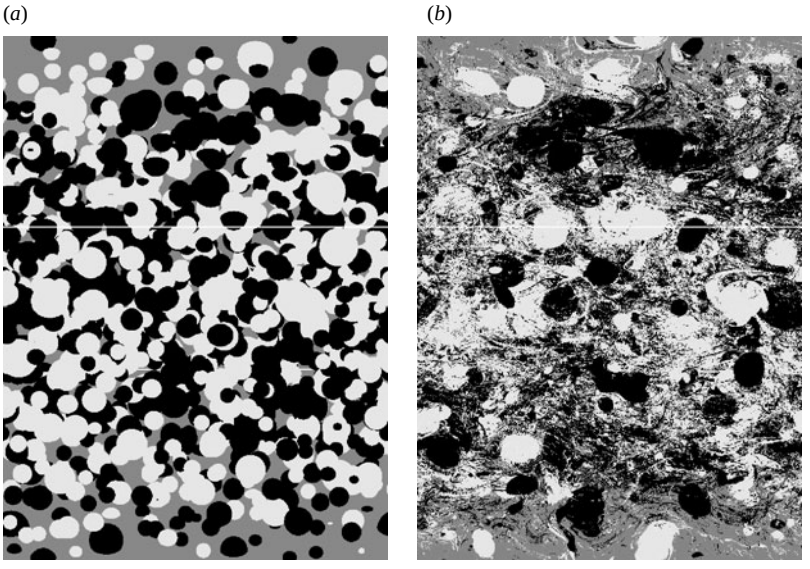

FIGURE 1. Distribution of QG PV anomaly q at times (*a*) $t=0$ and (*b*) $t=40$. The view is orthographic, at an angle of 60° from the vertical, and from the (y, z)-plane. From this view we can see the top and front faces of the domain and a white line indicates where these faces intersect. Cyclonic vortices are lightly shaded while anticyclonic vortices are darkly shaded.

represent $\boldsymbol{\varphi}$, $\boldsymbol{u}$, etc. (see Dritschel & Viúdez 2003 for further details of these standard parameter settings).

The QG simulation was run for 60 time units, well into the decaying stage (peak contour complexity occurs around 20 time units). The QG PV field q at time $t=40$ (see figure 1*b*) was then used to initialize the simulations using the NH equations at finite Rossby number (henceforth we redefine the $t=40$ state in the QG simulation as the initial state, i.e. '$t=0$'). We set the PV anomaly $\varpi = \Pi - 1$ by linearly scaling the QG PV anomaly, i.e. $\varpi = \epsilon q/4\pi$, for a given 'Rossby number' ϵ. To obtain nearly balanced initial values for the ageostrophic horizontal vorticity $\mathscr{A}_h$, we used a time ramping procedure (Dritschel & Viúdez 2003) like in OPV balance. Starting from a state of rest, the PV anomaly was slowly ramped up from 0 to ϖ over 10 inertial periods, $T_{ip} = 2\pi/f$, while integrating the full dynamical equations for $\mathscr{A}_h$, cf. (2.6). However, unlike in OPV balance (see below), for simplicity we fixed the PV contours (or fluid particles) while ramping.

The NH simulations were then integrated forwards over the equivalent of 20 QG time units, in six cases: for two frequency ratio values $c \equiv N/f = 10$ and 100, and for three PV-based Rossby numbers $\epsilon = 0.25$, 0.5 and 0.75. To ensure that the inertia–gravity waves, having frequencies between f and N, are well resolved in time, we used an explicit third-order Adams–Bashforth time-stepping procedure with a time step $\Delta t = 0.1 T_{buoy}$, where $T_{buoy} = 2\pi/N$ is the buoyancy period. Also, to control the generation of grid-scale noise during the time integration, a weak bi-harmonic hyperdiffusion was added to the $\mathscr{A}_h$ tendencies. The maximum damping rate (on the highest wavenumber in spectral space) was taken to be $1 + 10\epsilon^4$ per inertial period. As demonstrated in Dritschel & Viúdez (2003), this damping rate is much less than that required in a pure pseudo-spectral simulation.

3.2. *Balance diagnosis procedures*

After conducting the simulations, 'balance' was diagnosed using a variety of procedures: linear QG, OPV, and NQG balance. These are compared below in § 4.3.

Linear QG balance corresponds to the first-order QG equations (2.19) alongside the ω-equation (the vertical component of equation (2.23)) for the vertical velocity field w (Waite & Bartello 2006; Viúdez & Dritschel 2003).

OPV balance is based on a simpler procedure first introduced by Dritschel & Viúdez (2003) and used here to set up the initial fields. Briefly, all dynamical fields are first set to zero ($\boldsymbol{\varphi} = 0, \varpi = 0$); then, the NH equations are evolved over a fictitious ramp time τ, during which the PV anomaly is multiplied by a ramp function $T(\tau)$ which varies smoothly from 0 to 1 from the beginning ($\tau = 0$) to the end ($\tau = \Delta$) of the ramp period. The ramp function used is $T(\tau) \equiv \frac{1}{2}\,[1 - \cos \pi\tau/\Delta]$. The PV is not advected by the flow, but held fixed in space. While this procedure tends to generate nearly balanced fields, fixing the PV in space is inconsistent with material conservation of PV. OPV balance (Viúdez & Dritschel 2004) relaxes this constraint. However, one must then find the positions of fluid particles (or material contours) which end up, after ramping, in the 'correct' positions, i.e. corresponding to the PV field being diagnosed. Practically, one has to make a guess for this 'base configuration' at the beginning of the ramp, then iterate until convergence. Iterations correspond to backward and forward time integrations in which the IGWs are removed at the beginning of the ramp and PV is restored at the end of the ramp. The balanced flow obtained is a solution of the IGW-permitting dynamics in which the amplitude of IGWs is minimal. In previous studies, OPV balance has been found to be an effective means of identifying and quantifying IGWs (Viúdez & Dritschel 2006; Dritschel & Viúdez 2007). Here we apply this procedure using 5 loops, 3 forward and 2 backward integrations, over a ramp period of $\Delta = 5T_{ip}$. The results are not sensitive to Δ for $\Delta > 3T_{ip}$ (see Dritschel & Viúdez 2007).

NQG balance solves equations (2.20) to directly obtain balanced fields to $O(\epsilon^2)$. We apply this procedure using the iterative technique introduced in § 2.3. The NQG procedure is much more efficient than the OPV procedure (≈ 200 times faster), as OPV balance relies on a repeated time integration of the NH equations over the ramp period. However, in principle, OPV balance is not limited to a particular order of accuracy, since OPV balance does not impose balance relations. We expect it to be the most 'accurate' overall, in the sense of ascribing the greatest proportion of a flow to balanced dynamics. Of course, in a flow dominated by IGWs, OPV balance may not be significantly more accurate than simpler procedures. But that is not the situation being considered here. Our aim is to understand the degree to which the PV alone may control – through some underlying balance – fluid motion in rotating stratified turbulence.

4. Results

4.1. *Potential vorticity evolution*

A comparison of the PV anomaly at the equivalent time of $t = 5$ QG time units in the QG simulation and the NH simulations with $\epsilon = 0.25$, 0.5 and 0.75 is presented in figure 2, with $c = 10$ in the top row and $c = 100$ in the bottom row. Note that, in the simulations with $\epsilon = 0.25$, this time corresponds to 400 buoyancy periods or 40 inertial periods, plenty of time for exciting inertia–gravity waves. There is remarkably close agreement in the PV fields at this time, even though $Ro_{max} = 0.88$ and $Fr_{max} = 0.49$ in

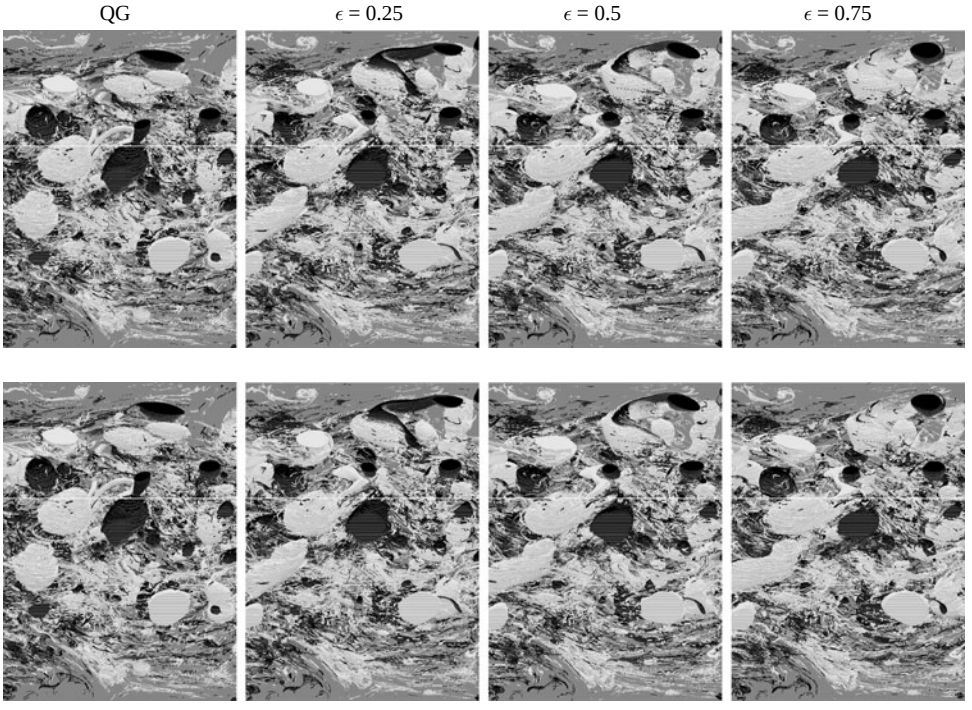

FIGURE 2. Comparison of the PV anomaly fields ϖ for various PV-based Rossby numbers ϵ as labelled, at 5 QG time units. The top row is for $c = 10$ and the bottom row is for $c = 100$. The view and shading is as in figure 1, but only the inner eighth of the domain is shown.

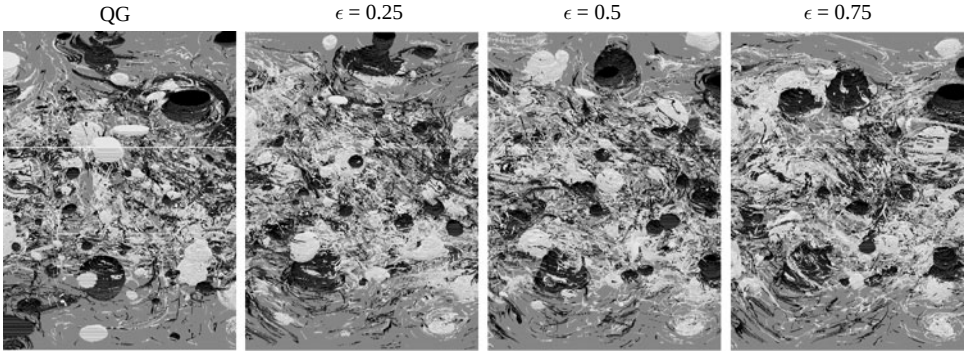

FIGURE 3. Comparison of the QG and NH simulations with $c = 10$ at time $t = 20$. The view and shading is as in figure 2.

the case with $\epsilon = 0.75$. Also, there is no significant dependence on the frequency ratio $c = N/f$ as can be seen by comparing images in the top and bottom rows.

In figure 3 we compare the PV anomaly at the later time of $t = 20$ QG time units for the case $c = 10$. By this time, significant differences have developed, particularly between the QG and the high-Rossby-number cases. Most of these differences arise from balanced, ageostrophic motions, according to the results presented in § 4.3. These ageostrophic motions modify the flow field induced by the vortices, and hence the interactions between the vortices. Surprisingly, there is still some visible agreement between the simulations, particularly in the size and shape of certain vortex structures

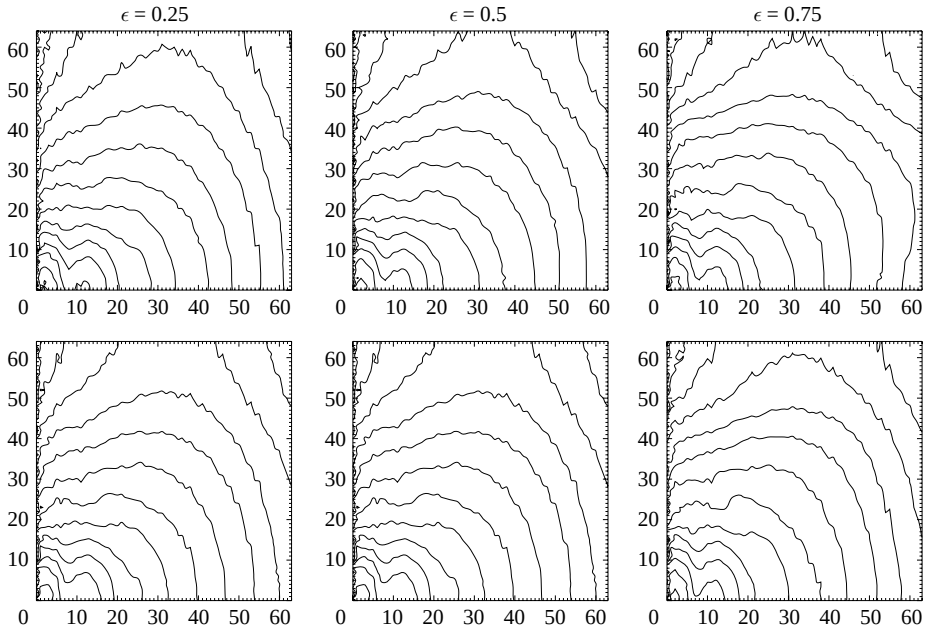

FIGURE 4. Comparison of time-averaged kinetic energy spectra $E(k_h, k_z)$ for the simulations with $\epsilon = 0.25$ (left column), $\epsilon = 0.5$ (middle column) and $\epsilon = 0.75$ (right column), and with $c = 10$ (top row) and $c = 100$ (bottom row). The horizontal wavenumber k_h is scaled by f/N. The time average is taken between $t = 10$ to $t = 20$ QG time units.

(such as the dark anticyclonic one slightly left of centre near the bottom of the domain). Perhaps most significantly, the ageostrophic motions do not lead to visible, identifiable differences in the spatial structure of turbulence. The vortices appear to have remained roughly isotropic, in the sense of retaining vertical to horizontal aspect ratios H/L comparable to f/N. This can be quantified by examining kinetic energy spectra. In figure 4 we plot the time-averaged spectra obtained from all six simulations, where k_h is the horizontal wavenumber scaled by f/N and k_z is the vertical wavenumber. The roughly circular shapes of the iso-contours confirm that the fluid motion is approximately isotropic in the x, y and Nz/f coordinate space. If the fluid motion were isotropic in the unscaled x, y and z coordinate space, the iso-contours of kinetic energy in figure 4 would be elliptical, with aspect ratios of 10 and 100 in the top and bottom rows respectively.

The fact that the spectra for $c = 100$ are almost identical to those for $c = 10$ strongly supports the common practice in atmosphere/ocean modelling of using a much finer vertical grid spacing than the horizontal grid spacing (here f/N finer). This scaling is necessary to accurately represent balanced motions and, arguably, unbalanced motions as well.

4.2. *Vertical velocity and isopycnal displacement*

The vertical velocity field w is extraordinarily weak in a wide range of atmospheric and oceanic motions, and is often taken to be a measure of IGW activity (see Viúdez & Dritschel 2003). However, there can be a significant balanced component (i.e. arising from the PV), particularly for small Ro and Fr, and when starting close to a state of minimal IGW activity.

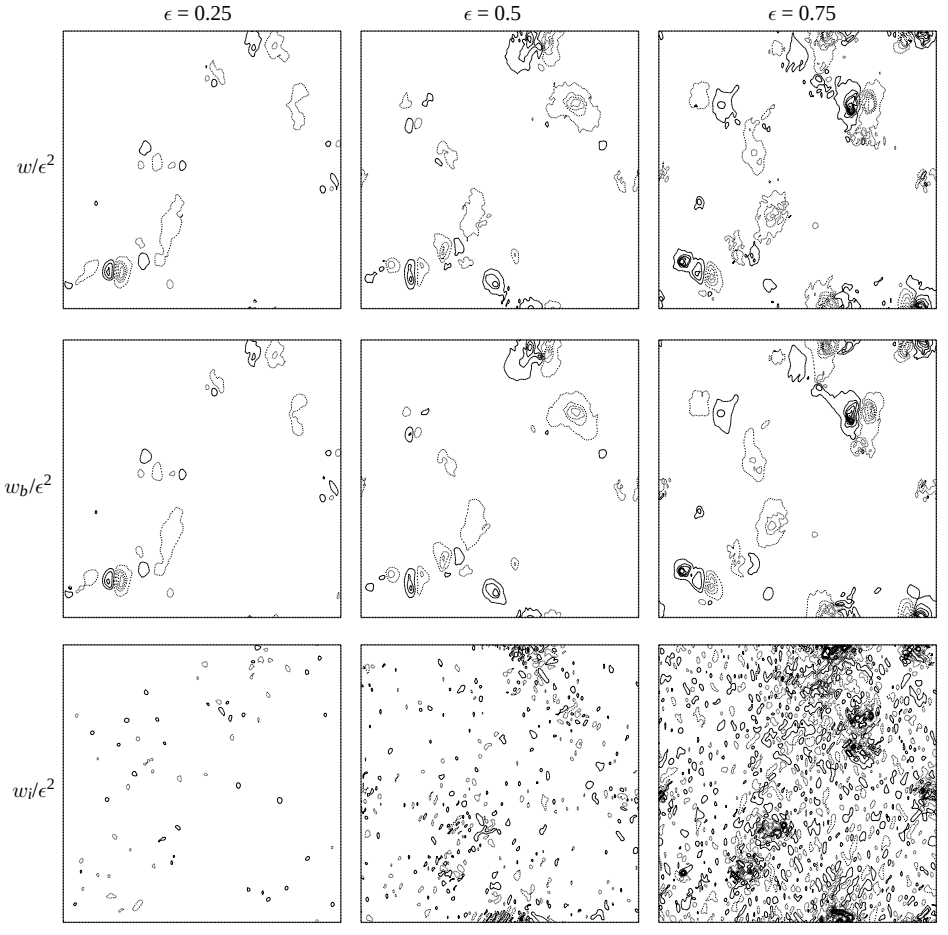

FIGURE 5. Comparison of the full (top row), balanced (middle row) and unbalanced (bottom row) components of the vertical velocity field (in a $y=0$ cross-section) at 20 QG time units, for $c=10$, and for the PV-based Rossby numbers indicated. The plotted contours have values $\pm\Delta/2$, $\pm 3\Delta/2, \ldots$, where Δ is the contour interval (negative contours have dashed lines, positive contours have solid lines; the zero value is omitted). For the full and balanced cases the contour interval is $\Delta = 0.0008$. The unbalanced contour interval is 1/4th of the balanced contour intervals.

Using the OPV balance procedure we diagnose the balanced vertical velocity field in the NH simulations at the final time $t=20$. In figure 5 we show the full w, balanced w_b and unbalanced $w_i = w - w_b$ components of w for the three simulations with $c=10$. Here, we show only a $y=0$ (vertical) cross-section, but this is typical of other cross-sections. Note that we have chosen the contour interval to be proportional to ϵ^2 to allow a comparison of the three Rossby numbers. The contour interval for w_i is four times smaller than that used for w and w_b. For the two smallest Rossby number cases there is close agreement in the largest magnitude structures in w, suggesting these are balanced motions linked to the PV. There is little or no imbalance in the smallest Rossby number case, though, as expected, the imbalance grows with increasing Rossby number. The structures which appear in the unbalanced field for the $\epsilon = 0.75$ case are mostly small scale and most likely to be interfering patterns of inertia–gravity waves. Most significantly, the unbalanced field is much smaller than

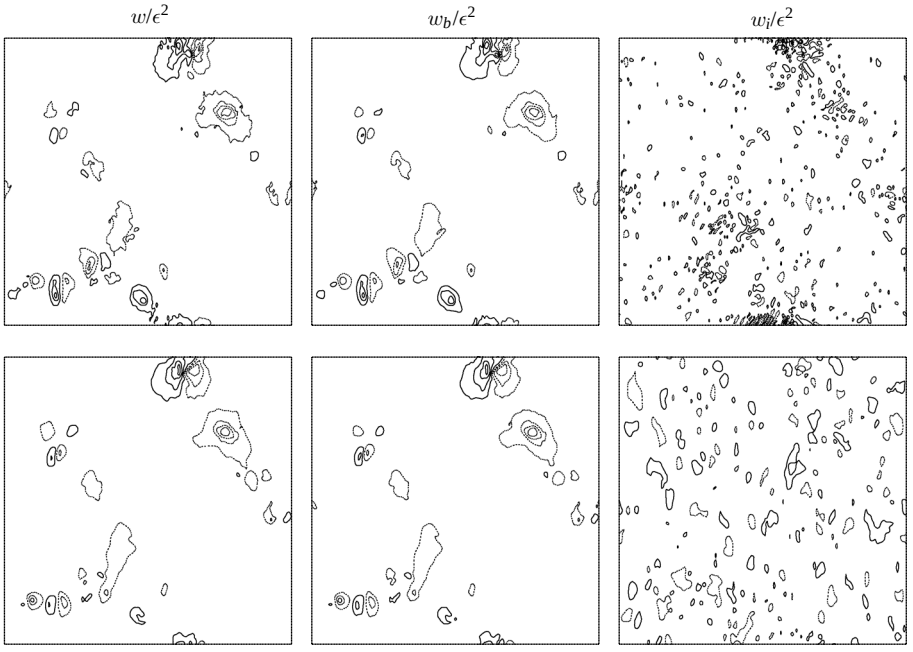

FIGURE 6. Comparison of the full (left column), balanced (middle column) and unbalanced (right column) components of the vertical velocity field at 20 QG time units in a $y=0$ cross-section for the cases $c=10$ (top row) and $c=100$ (bottom row), all with $\epsilon=0.5$. The contour intervals for the full, balanced and unbalanced fields are Δ, Δ and $\Delta/4$ where $\Delta=0.008/c$.

the balanced field, indicating that the flow has remained close to balance over the entire evolution.

In figure 6 we compare the full, balanced and unbalanced components of the vertical velocity field for the case $\epsilon=0.5$ at time $t=20$, for both $c=10$ and $c=100$. The $c=100$ cases are very similar to the $c=10$ cases. As with the PV field, varying c does not qualitatively change the dynamics: the structure of the vertical velocity field is the same. However the magnitude of the velocity field does vary, inversely proportionally to c, consistent with QG scaling – see (2.16). Also there is less small-scale structure for $c=100$, suggesting that there is less IGW activity as c increases. In fact, this is true only for earlier times in the evolution, as shown in figure 7, illustrating the evolution of the r.m.s. imbalance in w as a percentage of the total r.m.s. vertical velocity. Overall, the percentage of imbalance is reduced for the $c=100$ case, but the difference may not be significant for these two (large) values of c.

We turn next to the isopycnal displacement field $\mathscr{D}=-b/N^2$, together with its balanced $\mathscr{D}_b$ and unbalanced components $\mathscr{D}_i$, in figure 8. Here, images for all three Rossby numbers and $c=10$ are displayed in the same format as used for w in figure 5. Note that we have scaled $\mathscr{D}$ by ϵ, as suggested by the QG analysis, see (2.16). This analysis is supported here by the close similarity of $\mathscr{D}$ or $\mathscr{D}_b$ across Rossby number. Dipolar structures are prominent in these fields (in a vertical cross-section). Such structures are consistent with PV inversion and balance: cyclonic PV anomalies squeeze the isopycnals locally, while anticyclonic PV anomalies spread them out. Compared to the vertical velocity field w in figure 5, the displacement field $\mathscr{D}$ is substantially smoother, and is evidently much more balanced – the contour interval

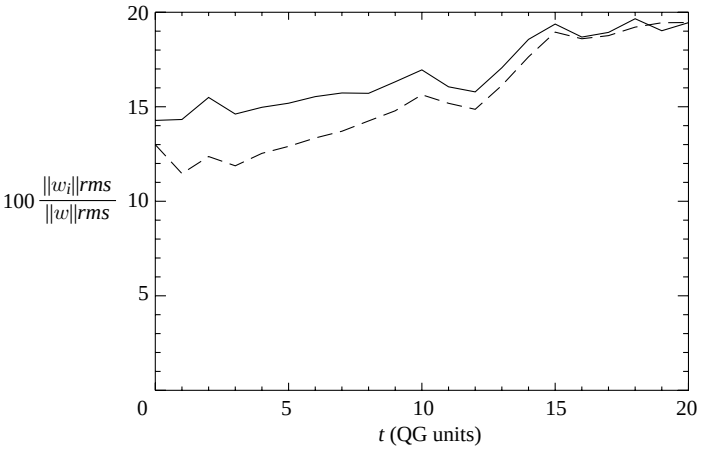

FIGURE 7. Time evolution of the r.m.s. imbalance in w as a percentage of the total for $c = 10$ (solid line) and $c = 100$ (dashed line) when $\epsilon = 0.5$.

used here for $\mathscr{D}_i$ is only 1/50th of that used for $\mathscr{D}$ and $\mathscr{D}_b$. Quantitative results presented next demonstrate that $\mathscr{D}$, as well as the horizontal velocity $\boldsymbol{u}_h$, are much more balanced than w throughout the evolution.

4.3. *Unbalanced evolution and balance procedure comparison*

In this section we quantify the imbalance in various fields as a function of time and using various balance procedures. All results shown are for $c = N/f = 10$. In figure 9 we show the percentage of r.m.s. imbalance in the fields u (top row), w (middle row) and $\mathscr{D}$ (bottom row), obtained using OPV (bold solid line), NQG (thin solid line) and linear QG balance (dashed line) for the three PV-based Rossby numbers ϵ. As expected, the percentage of imbalance increases with ϵ, and the greatest percentage of imbalance occurs in the vertical velocity field w. After the initial adjustment (due to the imperfect initialization of the simulations), the percentage of imbalance in the fields does not change substantially. There is a weak upward trend in $\%w_i$, but a nearly level evolution of $\%u_i$ and $\%\mathscr{D}_i$. This small growth in $\%w_i$ contributes negligibly to the overall imbalance. For instance, in the energy norm $E_i = \langle |\boldsymbol{u}_i|^2 + N^2\mathscr{D}_i^2\rangle^{1/2}/\langle |\boldsymbol{u}|^2 + N^2\mathscr{D}^2\rangle^{1/2}$ (where $\langle . \rangle$ denotes the domain average), there is no visible trace of w, see figure 10. Hence, these flows remain close to balance throughout the evolution – even for $O(1)$ Rossby numbers.

Next we examine how the various balance procedures compare. Overall, OPV balance is most accurate, in the sense that it ascribes the greatest proportion of the flow to balance, and thereby the smallest proportion to imbalance (see discussion at the end of § 3.2). In this sense, linear QG balance is least accurate. Substantial improvements can be obtained by taking into account second-order corrections in Rossby number as in QG^{+1} balance and in NQG balance. We emphasize that such improvements are likely to be greatest for flows which are carefully initialized and which do not contain large-amplitude IGWs. From this and previous work (Dritschel & Viúdez 2007), it appears that careful initialization is sufficient to keep IGWs weak, at least for PV-based Rossby numbers $\epsilon < 1$ and for $f/N \ll 1$.

Comparing NQG and OPV balance, it is clear that while NQG balance is less accurate it is still effective at capturing a significant proportion of the underlying balance. There are only minor differences in OPV and NQG balance for $\%u_i$ and $\%D_i$; significant differences occur only for $\%w_i$.

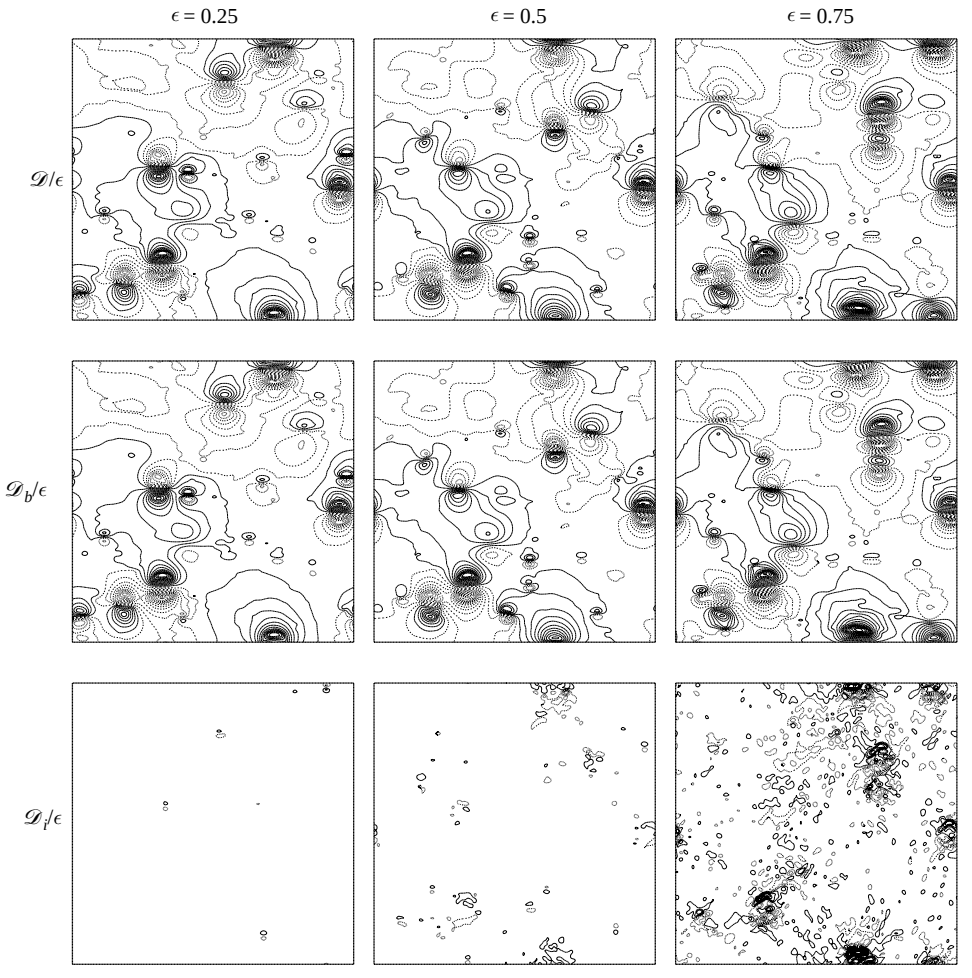

FIGURE 8. Comparison of the full (top), balanced (middle) and unbalanced (bottom) components of the displacement field $\mathscr{D}$ (in a $y = 0$ cross-section) at 20 QG time units for $c = 10$ and for the PV-based Rossby numbers indicated. The contour intervals for the full and balanced fields are $\Delta = 0.008$. Here, the unbalanced contour interval is 1/50th of the balanced contour intervals.

Finally, it is instructive to examine how the use of the iteration in NQG balance improves the estimate of balance. In figure 11 we plot the time-averaged r.m.s. percentage of imbalance of u, w and $\mathscr{D}$ versus the number of iterations used. As we increase the number of iterations, the percentage of imbalance decreases markedly in all cases, saturating around 4 or 5 iterations. Even just a few iterations results in a large percentage reduction in imbalance, by as much as 20% in some cases (for instance, $\%u_i$ reduces from 25% to 1.5% for $\epsilon = 0.75$). Hence, iteration improves accuracy. The significance of this improvement is further illustrated in figure 12 for the displacement field at the final time and at the highest Rossby number. In the NQG procedure with only two iterations, the 'imbalance' appears rather to be mostly the balance it has failed to capture, here ageostrophic corrections associated with the PV anomalies. The converged NQG procedure is much closer to the OPV procedure but

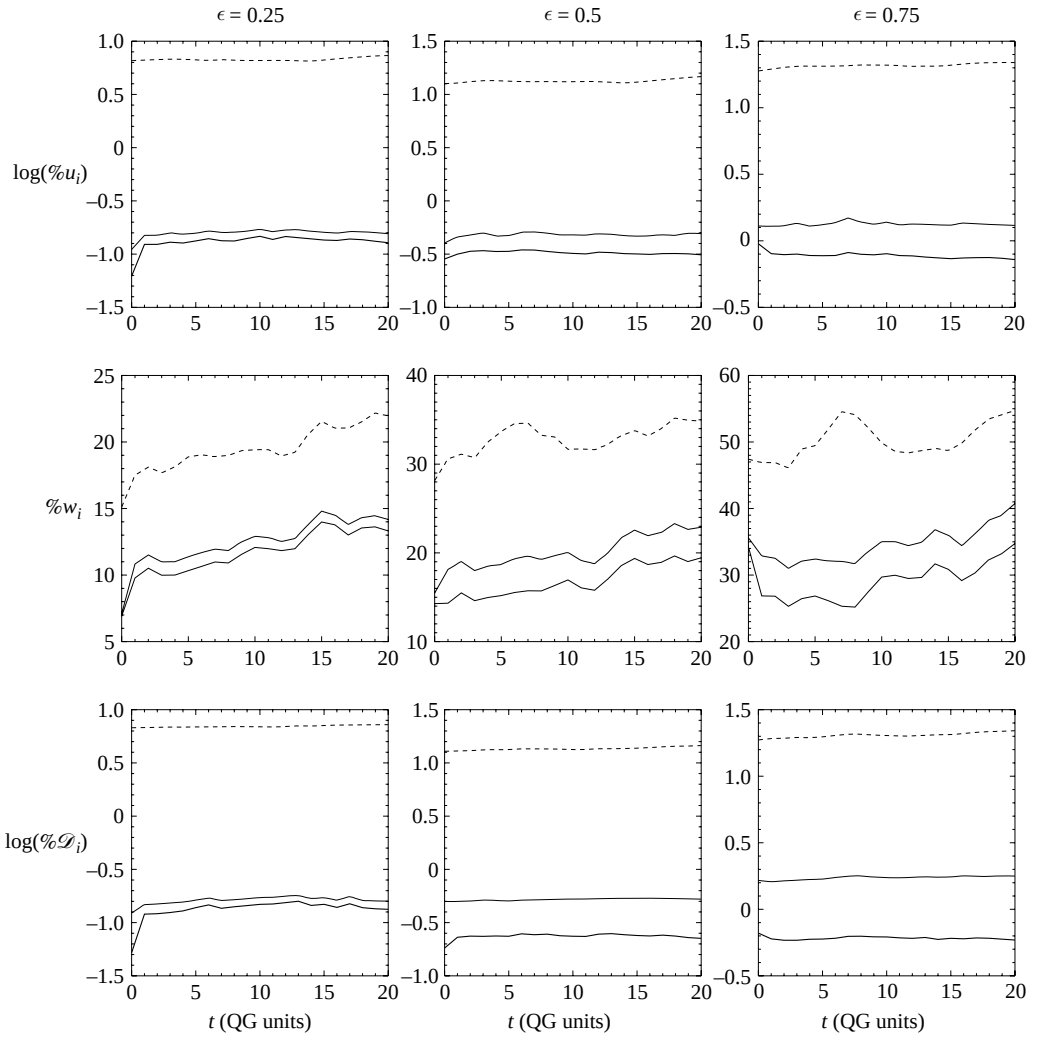

FIGURE 9. The r.m.s. imbalance as a percentage of the total fields obtained using OPV balance (bold solid line), NQG balance (thin solid line) and linear QG balance (dashed line) for u (top row), w (middle row) and $\mathscr{D}$ (bottom row), and for $\epsilon = 0.25$ (left column), $\epsilon = 0.5$ (middle column) and $\epsilon = 0.75$ (right column). The percentage of imbalance $\%\chi_i$ is defined by $100\ ||\chi_i||_{rms}/||\chi||_{rms}$ for a field χ. Note that logarithmic scales (to the base 10) are used for $\%u_i$ and $\%\mathscr{D}_i$ to distinguish the OPV and NQG curves.

still shows a weak remnant of apparently balanced motions. These results underscore the importance of using the exact PV in defining balance.

5. Discussion and conclusions

In this study we have examined the nature of 'balance', the control exerted by potential vorticity, in turbulent rotating stratified flows at Rossby numbers $Ro \lesssim 1$. Overall we find that such flows remain close to balance if carefully initialized, even when $Ro = O(1)$ (see figure 10). For small Rossby numbers, the flows are described well by the quasi-geostrophic model.

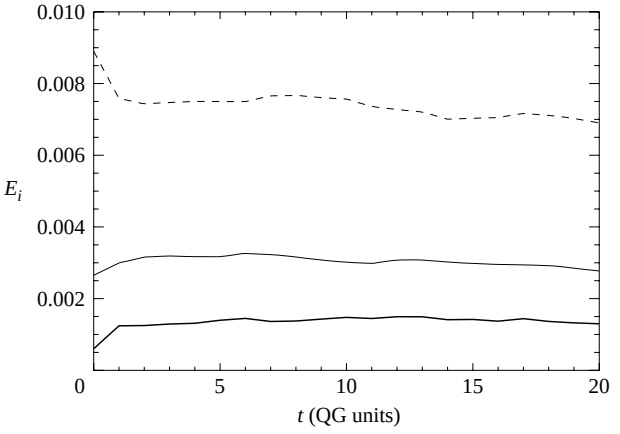

FIGURE 10. Evolution of the imbalance energy norm, E_i, as determined from OPV balance for $\epsilon = 0.25$ (bold solid line), $\epsilon = 0.5$ (thin solid line) and $\epsilon = 0.75$ (dashed line).

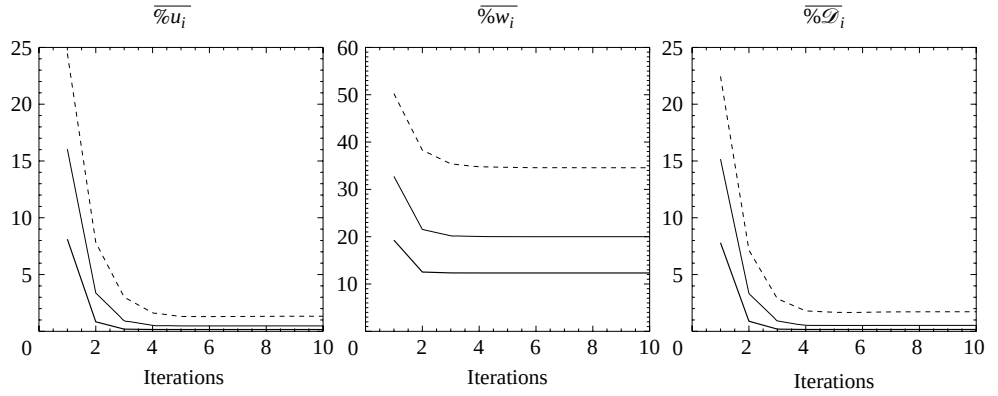

FIGURE 11. Time-averaged r.m.s. percentage of imbalance of u, w and $\mathscr{D}$ versus the number of iterations used in the NQG balance procedure. The three curves are for $\epsilon = 0.25$ (bold solid line), $\epsilon = 0.5$ (thin solid line) and $\epsilon = 0.75$ (dashed line).

We have found little dependence on the frequency ratio N/f, for $N/f \geqslant 10$, apart from expected changes in the scales of various fields. IGW activity reduces slightly when going from $N/f = 10$ to $N/f = 100$. Reducing N/f below 10 is likely to enhance both the balanced and unbalanced vertical velocities, and values of N/f of order unity appear to result in strong IGW emission, of the type not seen in the present study (W. Dewar, personal communication 2006). This remains to be investigated systematically however.

We have developed a new procedure for diagnosing balance, 'nonlinear quasi-geostrophic (NQG) balance', based on a quasi-geostrophic scaling of the non-hydrostatic equations. This approach is broadly similar to the Nonlinear Normal Mode Initialization procedure (Machenhauer 1977; Baer 1977; Baer & Tribbia 1977; Leith 1980) and more closely-related to the QG^{+1} balance procedure of MSR99 derived for the hydrostatic equations. The crucial difference is that NQG balance does not expand the definition of potential vorticity. We have shown that this significantly improves the estimate of balance in rotating stratified flows, a result which echoes previous findings in the shallow-water context (McIntyre & Norton

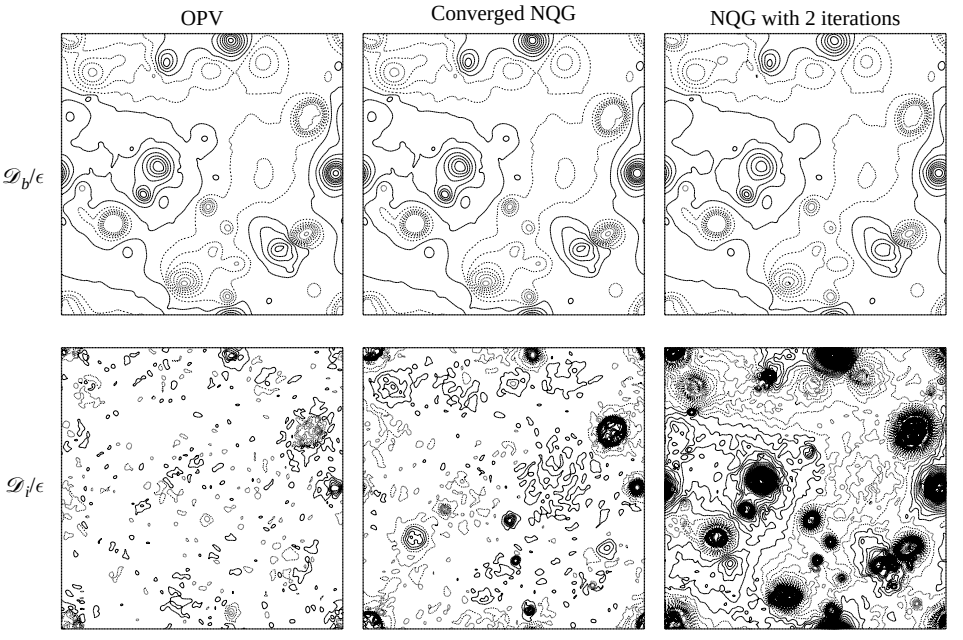

FIGURE 12. Comparison of the balanced (top row) and unbalanced (bottom row) isopycnal displacement obtained using OPV (left column), converged NQG (middle column) and NQG with 2 iterations (right column) — now in a $z=0$ cross-section—at $t=20$ time units for $\epsilon=0.75$. The balanced contour interval is $\Delta=0.008$ and the unbalanced coutour interval is $\Delta/50$.

2000; Ford *et al.* 2000; Mohebalhojeh & Dritschel 2000, 2001, 2004; Mohebalhojeh 2002).

Comparisons of OPV and NQG balance reveal that OPV balance is the most accurate, capturing the greatest proportion of the balanced component of the full dynamics (or equivalently attributing the smallest fraction of the flow to imbalance). However NQG balance is only marginally less accurate, particularly with respect to horizontal velocity and buoyancy, at considerably less numerical cost. Moreover, the use of iteration to preserve the exact definition of PV substantially improves the accuracy of NQG balance. The accuracy and efficiency of NQG balance suggests that it could be useful not only as a diagnostic procedure but also as a balanced flow evolution model, like QG, except employing the Rossby–Ertel PV as the prognostic variable, i.e. solving (2.19*b*) in conjunction with the iterative method described in § 2.3. This is currently being explored.

Support for this research has come from the UK Engineering and Physical Sciences Research Council (grant number XEP294). We wish to thank Dr Ali Mohebalhojeh, Dr Jean Reinaud and three anonymous referees for their helpful comments.

REFERENCES

BAER, F. 1977 Adjustment of initial conditions required to suppress gravity oscillations in nonlinear flows. *Beitr. Phys. Atmos.* **50**, 350–366.

BAER, F. & TRIBBIA, J. J. 1977 On complete filtering of gravity modes through nonlinear initialization. *Mon. Wea. Rev.* **105**, 1536–1539.

BOKHOVE, O. 1997 Slaving principles, balanced dynamics, and the Hydrostatic Boussinesq Equations. *J. Atmos. Sci.* **54**, 1662–1674.

Bolin, B. 1955 Numerical forecasting with the barotropic model. *Tellus* **7**, 27–49.

Bolin, B. 1956 An improved barotropic model and some aspects of using the balance equations for three-dimensional flow. *Tellus* **8**, 61–75.

Charney, J. G. 1948 On the scale of atmospheric motions. *Geofys. Publ.* **17**(2), 3–17.

Charney, J. G. 1955 The use of the primitive equations of motion in numerical prediction. *Tellus* **7**, 22–26.

Charney, J. G. 1962 Integration of the primitive and balance equations. *Proc. Intl Symp. on Numerical Weather Predic.* (ed. S. Syono), pp. 131–152. Meteorological Society of Japan.

Dritschel, D. G. & Ambaum, M. H. P. 1997 A contour-advective semi-lagrangian algorithm for the simulation of fine-scale conservative fields. *Q. J. R. Met. Soc.* **123**, 1097–1130.

Dritschel, D. G. & Viúdez, Á. 2003 A balanced approach to modelling rotating stably stratified geophysical flows. *J. Fluid Mech.* **488**, 123–150.

Dritschel, D. G. & Viúdez, Á. 2007 The persistence of balance in geophysical flows. *J. Fluid Mech.* **570**, 365–383.

Ford, R., McIntyre, M. E. & Norton, W. A. 2000 Balance and the slow quasimanifold: Some explicit results. *J. Atmos. Sci.* **57**, 1236–1254.

Gill, A. E. 1982 *Atmosphere-Ocean Dynamics*. Academic.

Hoskins, B. J., McIntyre, M. E. & Robertson, A. W. 1985 On the use and significance of isentropic potential-vorticity maps. *Q. J. R. Met. Soc.* **111**, 877–946.

Lane, T. P., Doyle, J. D., Plougonven, R., Shapiro, M. A. & Sharman, R. D. 2004 Observations and numerical simulations of inertia–gravity waves and shearing instabilities in the vicinity of a jet stream. *J. Atmos. Sci.* **61**, 2692–2706.

Leith, C. E. 1980 Nonlinear normal mode initialization and quasi-geostrophic theory. *J. Atmos. Sci.* **37**, 958–968.

Machenhauer, B. 1977 On the dynamics of gravity oscillations in a shallow water model, with application to normal mode initialization. *Beitr. Phys. Atmos.* **50**, 253–271.

McIntyre, M. E. & Norton, W. A. 2000 Potential vorticity inversion on a hemisphere. *J. Atmos. Sci.* **57**, 1214–1235, Corrigendum **58**, 949.

Mohebalhojeh, A. R. 2002 On shallow-water potential-vorticity inversion by Rossby-number expansions. *Q. J. R. Met. Soc.* **128**, 679–694.

Mohebalhojeh, A. R. & Dritschel, D. G. 2000 On the representation of gravity waves in numerical models of the shallow water equations. *Q. J. R. Met. Soc.* **126**, 669–688.

Mohebalhojeh, A. R. & Dritschel, D. G. 2001 Hierarchies of balance conditions for the f-plane shallow water equations. *J. Atmos. Sci.* **58**, 2411–2426.

Mohebalhojeh, A. R. & Dritschel, D. G. 2004 Contour-advective semi-Lagrangian algorithms for many-layer primitive equation models. *Q. J. R. Met. Soc.* **130**, 347–364.

Muraki, D. J., Snyder, C. & Rotunno, R. 1999 The next-order corrections to quasigeostrophic theory. *J. Atmos. Sci.* **56**, 1547–1560.

Reinaud, J., Dritschel, D. G. & Koudella, C. R. 2003 The shape of vortices in quasi-geostrophic turbulence. *J. Fluid Mech.* **474**, 175–191.

Rotunno, R., Muraki D. J. & Snyder, C. 2000 Unstable baroclinic waves beyond quasigeostrophic theory. *J. Atmos. Sci.* **57**, 3285–3295.

Vallis, G. K. 1996 Potential vorticity inversion and balanced equations of motion for rotating and stratified flows. *Q. J. R. Met. Soc.* **122**, 291–322.

Vanneste, J. & Yavneh, I. 2004 Exponentially small inertia–gravity waves and the breakdown of quasi-geostrophic balance. *J. Atmos. Sci.* **61**, 211–223.

Viúdez, Á. & Dritschel, D. G. 2003 Vertical velocity in mesoscale geophysical flows. *J. Fluid Mech.* **483**, 199–223.

Viúdez, Á. & Dritschel, D. G. 2004 Optimal potential vorticity balance of geophysical flows. *J. Fluid Mech.* **521**, 343–352.

Viúdez, Á. & Dritschel, D. G. 2006 Spontaneous generation of inertia–gravity wave packets by balanced geophysical flows. *J. Fluid Mech.* **553**, 107–117.

Waite, M. L. & Bartello, P. 2006 The transition from geostrophic to stratified turbulence. *J. Fluid Mech.* **568**, 89–108.

Warn, T, Bokhove, O., Shepherd, T. G. & Vallis, G. K. 1995 Rossby number expansions, slaving principles, and balance dynamics. *Q. J. R. Met. Soc.* **121**, 723–739.