

## An Interpretation of Baroclinic Initial Value Problems: Results for Simple Basic States with Nonzero Interior PV Gradients

HYLKE DE VRIES, JOHN METHVEN, AND THOMAS H. A. FRAME

*Department of Meteorology, University of Reading, Reading, United Kingdom*

BRIAN J. HOSKINS

*Department of Meteorology, University of Reading, Reading, and Grantham Institute for Climate Change, Imperial College, London, United Kingdom*

(Manuscript received 27 February 2008, in final form 8 October 2008)

### ABSTRACT

In the Eady model, where the meridional potential vorticity (PV) gradient is zero, perturbation energy growth can be partitioned cleanly into three mechanisms: (i) shear instability, (ii) resonance, and (iii) the Orr mechanism. Shear instability involves two-way interaction between Rossby edge waves on the ground and lid, resonance occurs as interior PV anomalies excite the edge waves, and the Orr mechanism involves only interior PV anomalies. These mechanisms have distinct implications for the structural and temporal linear evolution of perturbations.

Here, a new framework is developed in which the same mechanisms can be distinguished for growth on basic states with nonzero interior PV gradients. It is further shown that the evolution from quite general initial conditions can be accurately described (peak error in perturbation total energy typically less than 10%) by a reduced system that involves only three Rossby wave components. Two of these are counter-propagating Rossby waves—that is, generalizations of the Rossby edge waves when the interior PV gradient is nonzero—whereas the other component depends on the structure of the initial condition and its PV is advected passively with the shear flow. In the cases considered, the three-component model outperforms approximate solutions based on truncating a modal or singular vector basis.

### 1. Introduction

Zonal flows may support exponentially amplifying growing normal modes (gNMs). There are two well-known necessary conditions for baroclinic or barotropic shear instability. The first of these is that the meridional potential vorticity (PV) gradient (denoted  $\bar{q}_y$  for short) changes sign at least once within the domain (Charney and Stern 1962). The second necessary condition is that the zonal flow is positively correlated with  $\bar{q}_y$  (Fjørtoft 1950). However, transient energy growth may occur even if the basic state does not support gNMs, as was shown for instance by Orr (1907). The simplest illustration of what is now generally known as the *Orr mechanism* is the evolution of an initially upshear tilted plane wave in PV for an unbounded flow with constant

shear and  $\bar{q}_y = 0$ . As time increases, the shear acts to reduce the tilt by rotating the phase lines until they are aligned perpendicular to the flow. This has been described as the “Venetian blind effect.” As the PV structure untilts, cancellation between opposite-signed PV anomalies in the inversion to obtain velocity reduces with time. The energy of the perturbation rapidly increases, peaks when the phase lines are perpendicular, and then decays algebraically as  $t^{-2}$  as they tilt downshear (Orr 1907). Yamagata (1976) and Boyd (1983) showed that the basic Orr mechanism is unchanged if  $\bar{q}_y$  is a constant ( $\beta$ ); the PV pattern remains of constant amplitude but retrogresses westward at a rate given by the instantaneous Rossby wave dispersion relation.

The objective of this paper is to obtain a comprehensive understanding of the transient and asymptotic linear development from more general initial conditions (ICs) in basic states with nonzero  $\bar{q}_y$  in which both mechanisms above are active. ICs that will be considered are restricted to be horizontally wavelike but to have a

---

*Corresponding author address:* Hylke de Vries, Department of Meteorology, University of Reading, P.O. Box 243, Earley Gate, Reading RG6 6BB, United Kingdom.  
E-mail: h.devries@reading.ac.uk

particular vertical structure, such as a PV monopole, dipole, or tripole. The aim is to find an approximate representation of the dynamics that involves the least number of components (specifically, three). Rossby wave components are considered because their propagation and interaction properties depend only on their relative phases and amplitudes, such that their evolution can be anticipated without explicit calculation.

Other studies relevant to the paper are now briefly discussed. Farrell (1982) extended Orr's work and argued that the growth of extratropical weather systems could be dominated by finite-time "non-modal" growth. The fastest-growing perturbation, as measured by a chosen norm over a finite time window, can be calculated as the leading singular vector (SV) of the linear dynamics; Farrell referred to it as an "optimal perturbation" (Farrell 1984, 1989; Farrell and Ioannou 1996). Several authors have focused on understanding the evolution of the optimal perturbations, but attention has been restricted mostly to the Eady (1949) model in which  $\bar{q}_y$  is zero in the interior (O'Brien 1992; Mukougawa and Ikeda 1994; Fischer 1998; Morgan 2001; Morgan and Chen 2002; De Vries and Opsteegh 2005). Most growth at early times results from the Orr mechanism because the SVs have PV tilted upshear at initial time. Badger and Hoskins (2001) considered an isolated vortex as the IC, which can be represented by a positive PV anomaly sandwiched between negative PV anomalies above and below. In this case, the Orr mechanism again results in reduced cancellation between oppositely signed PV anomalies. For this reason, they described the process as "PV unshielding." Hakim (2000) studied cyclogenesis initiated by an isolated upper-level PV disturbance and concluded that normal-mode growth was dominant for the development.

As a step toward understanding more general initial value problems, Dirren and Davies (2004) considered horizontally wavelike ICs in the Eady (1949) model, with a  $\delta$ -function PV structure in height. The streamfunction obtained by inverting this PV structure was referred to as a kernel Rossby wave (KRW) by Heifetz and Methven (2005). Although its PV is concentrated into a  $\delta$  function, the flow it induces extends far away. Therefore, the KRW creates new PV anomalies by meridional advection at any location where  $\bar{q}_y$  is nonzero. Because  $\bar{q}_y \neq 0$  only at the boundaries of the Eady model, the KRW excites Rossby edge waves. In the interior, any PV structure is "passively" advected by the basic-state zonal flow. Based on the partitioning into Rossby wave components (KRWs in the interior and two Rossby edge waves at the boundaries), De Vries and Opsteegh (2007b) defined three distinct growth mechanisms:

- (I) *Shear instability* arising from coupling between two counterpropagating Rossby waves (CRWs), as described by Bretherton (1966) and Hoskins et al. (1985). In the Eady model, the CRWs are edge waves on the ground and rigid lid; their evolution can be described by simple ordinary differential equations (ODEs; Davies and Bishop 1994). The growth is exponential because the interaction is two way: wave 1 induces flow that increases the PV anomalies of wave 2, and this in turn increases its induced flow that acts upon wave 1.
- (II) *Resonance* arising from transient excitation of an edge wave (CRW) by an interior KRW. Now the CRW amplifies linearly in time because the interaction is one way: although the KRW can advect the boundary PV gradient at the boundaries, there is no interior PV gradient for the edge waves to advect; thus, they cannot make the KRW grow.
- (III) *Orr mechanism* arising from the differential advection of interior PV structures by the shear flow and the superposition of their induced velocity fields. Transient energy growth occurs through this mechanism although PV anomaly growth does not.

This paper investigates two generalizations of this analysis: the effects of positive PV gradient  $\beta$  throughout the flow interior (section 4) and the evolution from more general initial conditions (section 5). Two idealized basic states are considered: the Boussinesq version of the Charney (1947) model (Kuo 1973), which is unbounded above, and the Green (1960) model, which imposes a rigid lid. Both models have a basic-state zonal wind with uniform shear and  $\bar{q}_y = \beta$  in the interior.

The presence of a nonzero interior PV gradient implies that PV anomalies can be created at any location, making analysis more difficult than in the  $f$ -plane setting. Heifetz et al. (2004a) showed that the CRW interpretation of shear instability (I) generalizes to zonally symmetric flows with  $\bar{q}_y \neq 0$ , provided that the Rossby edge waves are replaced by CRW structures with interior as well as boundary PV. They are obtained by a linear combination of a gNM and its complex conjugate, the decaying normal mode (dNM). However, because the two CRW structures are obtained from just the gNM-dNM pair, it is clear that they cannot provide a complete description of the problem. They do describe correctly the linear asymptotic evolution, but a consideration of the continuous spectrum (CS) is generally required to understand the general initial value problem. It is also the CS that is intimately connected to the resonance and Orr mechanisms.

The continuum modes (CMs), forming the CS, are neutral normal modes (NMs) with a steering level within

the domain of the flow (Case 1960; Pedlosky 1964; Burger 1962). On the  $f$  plane, they have a PV  $\delta$  function at their steering level (Chang 1992; Jenkner and Ehrendorfer 2006). This PV  $\delta$  function is simply advected by the flow but induces a meridional wind at the ground. CMs therefore also have a boundary PV anomaly, the amplitude of which is determined by matching the Rossby wave propagation rate with the steering-level zonal flow (De Vries and Opsteegh 2007a). When there is an interior PV gradient, the CMs have PV at all interior levels but the importance of the PV near the steering level and boundary PV anomalies persists (appendix A).

As already stated, the aim of this paper is to represent the complete (transient and asymptotic) linear evolution using as few wave components as possible. Because generally a large portion of the CS is required to represent the initial evolution, wave components are sought that capture most aspects of the CS evolution without explicitly solving for the CS itself. The problem therefore involves identifying suitable components and deriving equations for their evolution.

The paper is structured as follows: In section 2a, a model is detailed that uses KRWs to describe quasi-geostrophic (QG) dynamics of zonally wavelike perturbations. This model provides the solution for the “full model” evolution. The relationship between PV anomalies and meridional air parcel displacements is made explicit in section 3. This step turns out to be crucial in the interpretation of baroclinic initial value problems. The projection of perturbations onto Rossby wave components is formulated in section 4a. Because normal modes or Rossby wave components will have overlapping PV structure when  $\bar{q}_y \neq 0$ , section 4 considers the simplest initial value problem in which the initial PV structure is an isolated PV  $\delta$  function (KRW). Clearly, the PV of an isolated KRW cannot be sheared apart, but the question is how PV elsewhere is excited through meridional advection. More general ICs with distributed PV structure (but still monochromatic waves) are investigated in sections 5 and 6. In section 7, the three-component model resulting from the analysis is compared with decomposition in terms of the normal-mode or singular-vector bases. Conclusions follow in section 8, together with a discussion of the limitations and possible extensions of the theory.

## 2. Rossby wave components

### a. Model basis constructed using kernel Rossby waves

The quasigeostrophic evolution of two-dimensional perturbations in absence of diabatic heating or friction

is described by the linearized QG potential vorticity conservation equation:

$$\frac{\partial q}{\partial t} + \bar{u} \frac{\partial q}{\partial x} + v \frac{\partial \bar{q}}{\partial y} = 0, \quad (1)$$

where  $q$  is the perturbation QG potential vorticity,  $v = \partial\psi/\partial x$  is the geostrophic meridional wind, and  $\psi$  is the perturbation streamfunction. Basic-state quantities carry a bar. In the following sections we mostly consider the Charney (1947) model with constant density (Kuo 1973) (zonal flow  $\bar{u} = \Lambda z$  with constant buoyancy frequency  $N^2$  and no upper boundary). Northern Hemispheric conditions are assumed and the zonal wavelength of perturbations is set to that of the fastest-growing NM. The equations are nondimensionalized using a height scale  $H = (f_0/N)^2 \Lambda / \beta \sim 15$  km and horizontal scale  $L = (N/f_0)H$ , with other parameters and scalings as in Heifetz et al. (2004b).

For a single wavenumber  $k$ , the differential operator relating  $q$  to  $\psi$  can be inverted by means of a Green function:

$$\psi(z) = \int_0^\infty G(z, h) q(h) dh. \quad (2)$$

To render  $G(z, h)$  unique, a consideration of the boundary conditions is required. By using  $\partial G/\partial z = 0$  at rigid lids [or  $G \rightarrow 0$  at infinity if no upper boundary is present, such as in the Charney (1947) model],  $G(z, h)$  represents the streamfunction that can be associated with a PV  $\delta$  function at  $z = h$  (and zero potential temperature at rigid lids). This was called a *kernel Rossby wave* by Heifetz and Methven (2005). Figure 1 shows a KRW at  $z = H$  in the Charney (1947) model.

Following De Vries and Opsteegh (2007b), the vertically discretized version of (1) can be written as

$$\dot{\mathbf{q}}(t) = \mathbf{A} \cdot \mathbf{q}(t), \quad \mathbf{A} = -ik(\mathbf{U} + \mathbf{Qy} \cdot \mathbf{G} \cdot \mathbf{\Delta}), \quad (3)$$

where  $[\mathbf{U}]_{ij} = \delta_{ij} \bar{u}(z_i)$ ,  $[\mathbf{Qy}]_{ij} = \delta_{ij} \bar{q}_y(z_i)$ ,  $\mathbf{G}_{ij} = G(z_i, z_j)$ , and  $\delta_{ij}$  is the Kronecker symbol. The Bretherton (1966) approach is used to treat the boundary conditions. As a result, the matrix  $\mathbf{\Delta} = \text{diag}(1, \Delta z, \Delta z, \dots)$ , where  $\Delta z$  is the level spacing, is required to properly weight the boundary PV delta function. Any profile  $\bar{u}(z)$  and  $\bar{q}_y$  can be used in (3), but one has to keep in mind that  $G(z, h)$  depends nontrivially on the buoyancy frequency profile. Throughout this paper, the full model solutions are obtained by finding  $G$  analytically for each basic state considered and integrating (3) numerically.

### b. Counterpropagating Rossby waves

As discussed in the introduction, initial value problems in the Eady (1949) model are compactly described

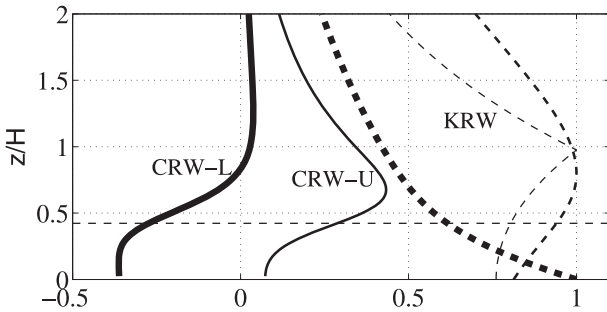


FIG. 1. Amplitude–height profile of streamfunction for a KRW located at  $z = H = (f_0/N)^2 \Lambda/\beta$  (section 2a) in the Charney model (for zonal wavelength equal to that of the fastest-growing normal mode). Also shown are CRW structures (section 2b) obtained using the orthogonality method. Solid lines represent PV, dashed lines  $-\psi$ . The horizontal dashed line denotes steering level of the gNM ( $z_s/H = 0.47$ ). All structures have been normalized to have unit maximum streamfunction.

by partitioning into Rossby wave components. Because  $\bar{q}_y = 0$  in the interior, all interior PV is simply advected by the zonal flow. However, there is a PV gradient at the boundaries, which allows creation of boundary PV anomalies also known as Rossby edge waves. The edge waves induce meridional winds with untilted vertical structure in quadrature with their boundary PV wave. This enables them to propagate counter to the shear flow and in opposite directions to one another, but also to influence the PV distribution at the opposite boundary such that they may phase-lock in a configuration of mutual amplification (gNM).

Heifetz et al. (2004a) showed that when  $\bar{q}_y \neq 0$  throughout the domain, baroclinic instability can still be viewed in terms of two untilted wave components that amplify mutually as they approach a phase-locked configuration. These two wave components are called CRWs and evolve just as the edge waves of the Eady model. A pair of CRW structures can be obtained from each gNM structure using wave activity orthogonality<sup>1</sup> following Heifetz et al. (2004a), as summarized in appendix B. Figure 1 shows the PV and streamfunction structure of the CRWs obtained from the fastest-growing NM of the constant density Charney model.

<sup>1</sup> The orthogonality approach was motivated by Held (1985), who showed that all neutral NMs (e.g., the CMs) are orthogonal to one another in terms of pseudomomentum and pseudoenergy. They are also orthogonal to the gNM and dNM but the latter two, however, are not orthogonal to one another. The orthogonality method described in Heifetz et al. (2004a) combines the gNM and dNM into two new untilted structures (CRWs) that are orthogonal with respect to pseudomomentum and a term appearing in the pseudoenergy.

The upper CRW’s PV is small at the ground and has a maximum at  $z_u/H \sim 0.70$ , well above the steering level of the gNM at  $z_s/H \sim 0.47$ . The lower CRW has the positive boundary PV of the gNM and negative PV above. Note how the streamfunction of the upper CRW is almost uniform for heights  $0 < z/H < 1$ , whereas the streamfunction of the lower CRW decays almost exponentially from the ground (in fact, it decays more rapidly with height than a pure boundary edge wave would decay because of the presence of the negative PV just above the positive surface buoyancy anomaly).

### 3. Relating PV anomalies and air parcel displacements

This section considers the relationship between PV anomalies and air parcel motions that will underpin the theory derived in later sections. When the flow is adiabatic and frictionless, PV is materially conserved. Therefore, disturbances that grow from infinitesimal amplitude on a zonally symmetric basic state must have PV anomalies that are obtained solely through meridional advection of the basic-state PV contours. The rate of change of the meridional coordinate of an air parcel  $\eta$  is related to the meridional flow by

$$\frac{\partial \eta}{\partial t} + \bar{u} \frac{\partial \eta}{\partial x} = v. \tag{4}$$

Eliminating  $v$  from (1) using the kinematic relation (4) gives the usual relation for PV anomalies associated with displacements of PV contours from the zonally symmetric initial state:

$$q_d = -\eta \frac{\partial \bar{q}}{\partial y}. \tag{5}$$

The discrete normal modes are solutions to this problem and therefore have PV anomalies described entirely by  $q_d$ . The same holds for CRWs because they can be obtained by linear superposition of a gNM and dNM. However, general ICs may include PV that did not arise from displacements of a zonally symmetric state, such that

$$q = q_d + q_p. \tag{6}$$

Substituting this partition into (1) gives

$$-\frac{\partial \bar{q}}{\partial y} \left( \frac{\partial \eta}{\partial t} + \bar{u} \frac{\partial \eta}{\partial x} - v \right) = \frac{\partial q_p}{\partial t} + \bar{u} \frac{\partial q_p}{\partial x}. \tag{7}$$

The left-hand side is zero for any  $\bar{q}_y$  because of (4); therefore, the right-hand side shows that  $q_p$  is advected as though passive, and  $q_p$  will henceforth be referred to

as passive PV. If the PV inversion operator is linear, as in QG dynamics, the meridional wind can also be partitioned with the PV components,  $v = v_d + v_p$ . The left-hand side of (7) then gives

$$\frac{\partial q_d}{\partial t} + \bar{u} \frac{\partial q_d}{\partial x} + v_d \frac{\partial \bar{q}}{\partial y} = -v_p \frac{\partial \bar{q}}{\partial y}, \quad (8)$$

which makes clear that although  $q_p$  is advected like a passive tracer, the wind it induces advects air parcels and forces the displacement PV anomalies. It is therefore dynamically active. Wherever  $\bar{q}_y = 0$ , it is clear that  $q_d = 0$  and that any initial PV is described by the  $q_p$  component. However, where  $\bar{q}_y \neq 0$  there is freedom to partition between the displacement and passive components. The partition has implications for the interpretation of perturbation evolution (8), as discussed further below.

#### 4. A simple baroclinic initial value problem

Consider an experiment in which the Charney basic state is perturbed by a single KRW at level  $H = (f_0/N)^2 \Lambda/\beta$  with zonal wavenumber  $k$ . The bottom row in Fig. 2 shows the evolution of the perturbation streamfunction as obtained by integrating the discretized<sup>2</sup> evolution Eq. (3). Note that the panels in Fig. 2 start at  $t = 1$ , where the IC corresponds to  $t = 0$ . Qualitatively, the evolution shows a gradual transition from the barotropic initial upper-level disturbance to a westward-tilted structure resembling the gNM. The perturbation is viewed from a frame of reference moving with the phase speed of the gNM.

In the following sections the evolution is analyzed by performing a simple two-step partitioning. First, the full perturbation PV  $q$  is partitioned into  $q = q_d + q_p$  (see section 3). In the second step,  $q_d$  is projected onto CRW structures. The results (section 4b) motivate the construction of a reduced model in section 4c. Finally, the evolution is summarized using phase-space diagrams in section 4d.

##### a. Partitioning into Rossby wave components

The positive PV anomaly of the KRW is positioned at level  $z = H$  and centered on  $x = 0$  at initial time. It induces poleward flow at  $kx = \pi/2$  and equatorward flow at  $kx = -\pi/2$ . Recall that the relative vorticity  $\xi = -k^2\psi$  so the negative streamfunction ( $\psi$ ) is collocated with positive vorticity. The top row in Fig. 2 shows

the streamfunction attributable to the PV at  $z = H$ , the level of the initial PV spike. Clearly, it propagates with a speed equal to the local flow  $\bar{u}(H)$ , which is faster than the gNM.

##### 1) PARTITIONING THE INITIAL PERTURBATION PV

An untilted PV wave propagates relative to the flow at the speed  $-\bar{q}_y v/(kq)$  (Heifetz et al. 2004a). Therefore, an isolated interior KRW cannot propagate relative to the flow at its level, even if the PV gradient is nonzero, because the wind associated with the KRW Green function is finite whereas the PV  $\delta$  function is an infinite spike. Therefore, it is reasonable to assign the initial PV entirely to the “passive component”—that is,  $q_p(0) = q(0)$ ;  $q_d(0) = 0$ —with the knowledge that it will simply drift with the flow. All other Rossby wave components, which are associated with  $q_d$ , have zero amplitude initially.

##### 2) PROJECTING $q_d$ ONTO CRWS

The next consideration is how to extract the CRW contributions from the displacement PV  $q_d$ . It is convenient to use pseudomomentum brackets to define these projections because the CRWs are orthogonal to one another and all continuum modes with respect to these brackets (Held 1985). At each instant the amplitude  $a_j$  and phase  $\epsilon_j$  of CRW- $j$  ( $j = 1, 2$  for the lower and upper CRW respectively) are obtained by projecting  $q_d$  onto the CRW structure function for meridional displacements  $\eta_j$ :

$$a_j(t) e^{i\epsilon_j(t)} = - \frac{\{ \eta_j, q_d(t) \}}{\left\{ \eta_j, \frac{\partial \bar{q}}{\partial y} \eta_j \right\}}. \quad (9)$$

The bracket notation  $\{, \}$  is defined in (B1). Their sum is

$$q_{\text{crws}} = a_1 e^{i\epsilon_1} q_1 + a_2 e^{i\epsilon_2} q_2. \quad (10)$$

The displacement PV remaining after subtracting the projection on to the CRW pair

$$q_n = q_d - q_{\text{crws}} \quad (11)$$

must be orthogonal to the CRWs and therefore could be written entirely as a superposition of continuum modes. The term *residual* PV will be used to describe

$$q_r = q - q_{\text{crws}} (= q_p + q_n), \quad (12)$$

which represents the portion of the total PV perturbation that has not been partitioned into the CRWs. The next section illustrates how the perturbation evolution

<sup>2</sup> The vertical domain extends from  $z = 0$  (ground) to  $z = 5H$  and is discretized into 500 layers. The Green function is obtained analytically assuming it is unbounded above  $5H$ .

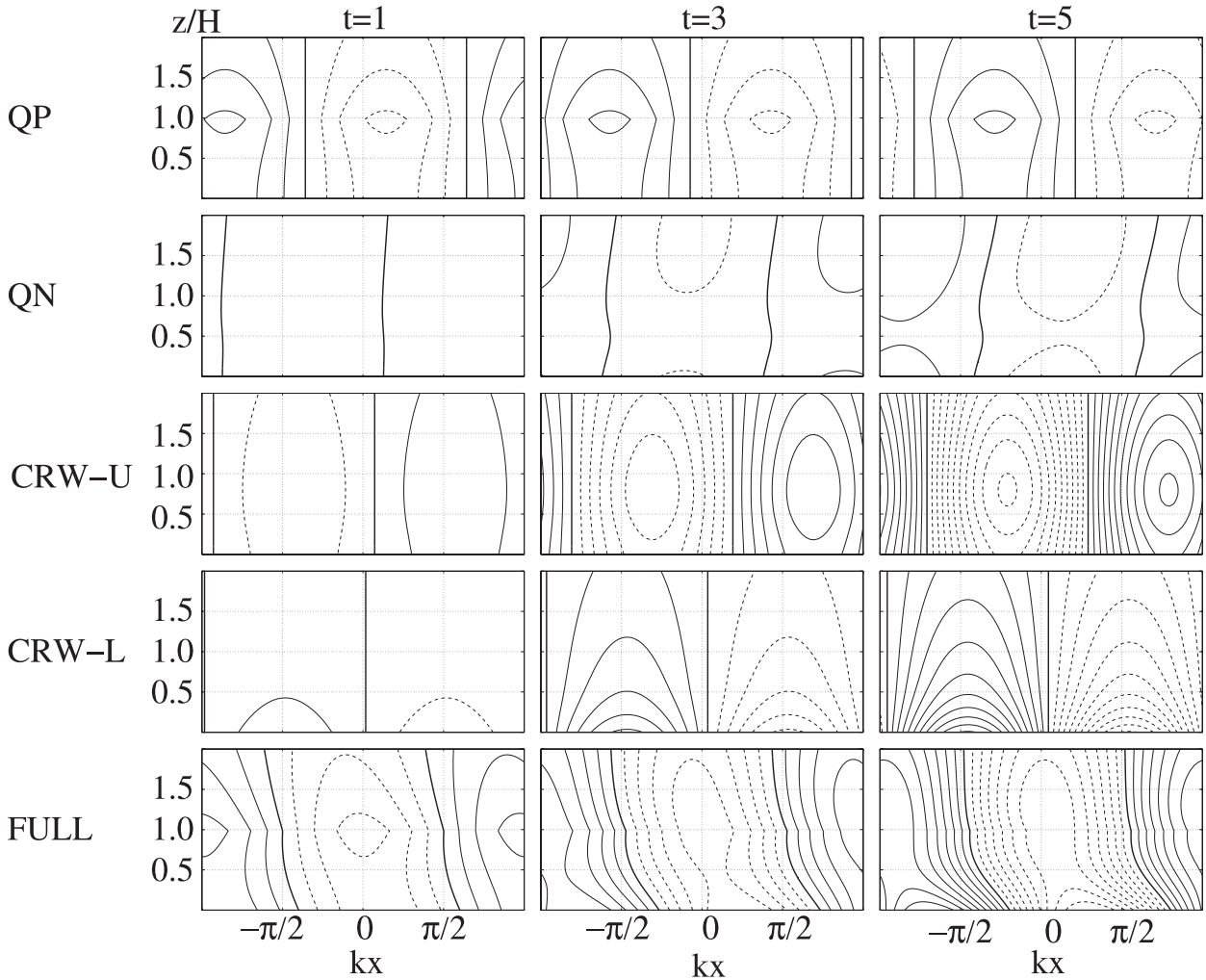


FIG. 2. Streamfunction evolution in terms of Rossby wave components starting from the initial condition with a KRW at  $z = H$ , viewed from a frame moving with the normal-mode phase speed:  $t =$  (left) 1, (middle) 3, and (right) 5. (from top to bottom) The KRW at  $z = H$  ( $q_p$ ), neglected PV ( $q_n = q - q_{crws} - q_p$ ), upper CRW, lower CRW, and total streamfunction. One nondimensional time unit corresponds to approximately 10 h. The same contour interval is used in all panels; negative contours are dashed.

can be readily understood in terms of the Rossby wave components just described and their interaction.

*b. Evolution of Rossby wave components*

The first column of Fig. 2 shows the structures after 1 nondimensional time unit ( $\sim 10$  h). Advection of the +ve interior basic-state PV by the initial KRW excites the upper CRW such that its +ve PV anomaly ( $-ve \psi$ ) is located where the flow is equatorward (shifted  $\pi/2$  to west of the KRW). Advection of the  $-ve$  boundary basic-state PV excites the lower CRW such that its +ve boundary PV anomaly ( $-ve \psi$ ) is located where the flow is poleward (shifted  $\pi/2$  to east of the KRW). Note how each CRW has an untilted structure. The CRWs are excited  $\pi$  out of phase with one another so that their interaction hinders the self-propagation of each CRW

counter to the flow at its home base [see Heifetz et al. (2004a) for the physical arguments relating propagation and growth behavior to the phase difference between CRWs]. The shear overwhelms propagation and reduces the phase difference between the CRWs, which will move into a westward-tilting configuration where mutual amplification can occur. Eventually they reach the phase-locked configuration of the gNM.

Note how the streamfunction associated with  $q_n$  (11) retains a small amplitude at all stages in the evolution. This shows that the excitation of continuum modes by  $q_p$  in this case is very weak throughout the evolution,<sup>3</sup>

<sup>3</sup> On the  $f$  plane this component would be identically zero at all times.

motivating the construction of a reduced set based on the two CRWs and the passive PV only.

### c. Three-wave evolution equations

Dirren and Davies (2004) obtained three-wave equations describing exactly the evolution of one interior KRW and the two edge waves (CRWs) of the Eady model. Based on the observation that also on the  $\beta$  plane the initial KRW and the two CRWs appear to dominate the perturbation evolution (Fig. 2), equations are constructed that describe the evolution of the three components, assuming  $q_n = 0$  (11) at all times:

$$\begin{aligned} \dot{a}_1 &= \gamma_{12}a_2 \sin \epsilon_{12} + \gamma_{13}a_3 \sin \epsilon_{13}, \\ \epsilon_1 &= -kc_{11} + \gamma_{12} \frac{a_2}{a_1} c_{12} \cos \epsilon_{12} + \gamma_{13} \frac{a_3}{a_1} c_{13} \cos \epsilon_{13}, \\ \dot{a}_2 &= \gamma_{21}a_1 \sin \epsilon_{21} + \gamma_{23}a_3 \sin \epsilon_{23}, \\ \epsilon_2 &= -kc_{22} + \gamma_{21} \frac{a_1}{a_2} c_{21} \cos \epsilon_{21} + \gamma_{23} \frac{a_3}{a_2} c_{23} \cos \epsilon_{23}, \\ \dot{a}_3 &= 0, \quad \epsilon_3 = -k\bar{u}_3, \end{aligned} \quad (13)$$

where subscripts 1, 2, and 3 are used to denote the lower and upper CRW and the KRW, respectively, and  $\epsilon_{ij} = \epsilon_i - \epsilon_j$  denotes the phase difference between two wave components. The factor  $c_{ii} = \bar{U}_i - \gamma_{ii}/k$  can be interpreted as the “self-propagation speed” of wave  $i$  ( $\bar{U}_i = \mathcal{W}_{ii}/\mathcal{P}_{ii}$ ; see appendix B). The interaction coefficients  $\gamma_{ij}$  are given by the global integrals (B2).

The key difference between (13) and their  $f$ -plane counterparts is that waves 1 and 2 correspond to CRWs with distributed PV structures and only wave 3 corresponds to a PV delta function (KRW). In the limit of increasing resolution ( $\Delta z \rightarrow 0$ ), wave 3 has infinite PV at its home base but finite meridional wind; it follows that  $\gamma_{3j} = 0$ . This in turn implies that wave 3 has a constant PV amplitude ( $\dot{a}_3 = 0$ ) and cannot propagate relative to the basic flow at its level ( $\epsilon_3 = -k\bar{u}_3$ ) as on the  $f$  plane. CRW interaction is described by the factors  $\gamma_{12}$  and  $\gamma_{21}$ . The KRW influences the growth and propagation of the CRWs through the wind field it induces ( $\gamma_{13}$  and  $\gamma_{23}$ ). When  $a_3 = 0$ , the equations reduce to the two-wave system of Heifetz et al. (2004a).

#### 1) COMPARING THE THREE-WAVE TO THE FULL MODEL

The CRW amplitude and phases obtained using the three-wave Eqs. (13) are *identical* to those obtained using the full model integration and the pseudomomentum projection (9) of  $q_d$  at each instant. This result is at first sight surprising. The explanation relates to the orthogonality properties of NMs (Held 1985). The CRW pair is simply a linear transformation of the unstable NM complex conjugate pair. These two NMs are

orthogonal to all the remaining neutral modes. Therefore, the part of the displacement PV ( $q_d$ ) that does not project onto the two CRWs ( $q_n$ ) must be orthogonal to them under the pseudomomentum bracket and therefore cannot have any influence on the evolution of the amplitudes and phases of the CRWs. In a later section we will show that the three-wave approximation is also able to predict the perturbation total energy, typically to within 10%.

#### 2) COMPARING THE THREE-WAVE TO THE TWO-WAVE MODEL

An alternative breakdown of the problem is to use a modal decomposition of the total PV perturbation and consider the evolution of the resulting normal modes. Because a CRW pair is obtained from an unstable NM complex conjugate pair and the unstable modes are orthogonal to all the continuum modes, the CRWs evolve from their initial projection following the two-wave equations [obtained by setting  $a_3 = 0$  in (13)]—that is, as if all other PV were nonexistent.

Figure 3 contrasts the evolution of CRW amplitudes and phases obtained by integrating the three-wave equations [starting from  $q_p(0) = q(0)$  and zero CRW amplitudes] with the evolution obtained by projecting the initial  $\delta$  function  $q(0)$  onto the CRWs and integrating the two-wave equations (equivalent to modal decomposition of the full solution at every instant). Clearly the CRW evolution from this modal perspective is markedly different. It would not be a good representation of the full model evolution without superposing the continuum modes. This aspect is explored more fully in section 7.

### d. CRW solutions when forced by a KRW

#### 1) PHASE-SPACE TRAJECTORIES

Because the Rossby wave components (KRW and CRWs) are untilted, the entire evolution is readily summarized in a phase-space diagram. In Fig. 4, the angle in polar coordinates shows the phase difference between any wave component and the lower CRW. The radius represents the ratio of surface meridional velocity amplitude of each component relative to the lower CRW, rescaled using the ratio of the CRW winds in the gNM configuration  $\hat{r}_{\text{NM}}$ ,

$$r = (4/\pi) \arctan(\hat{r}/\hat{r}_{\text{NM}}), \quad (14)$$

such that  $r = [0,1,2]$  corresponds to  $\hat{r} = [0, \hat{r}_{\text{NM}}, \infty]$ . We now discuss Fig. 4a. At the initial time only the KRW at  $z = H$  is present, but it immediately excites CRW-L  $\pi/2$  to the east. Therefore, the phase-space trajectory describing the evolution of the KRW relative to CRW-L

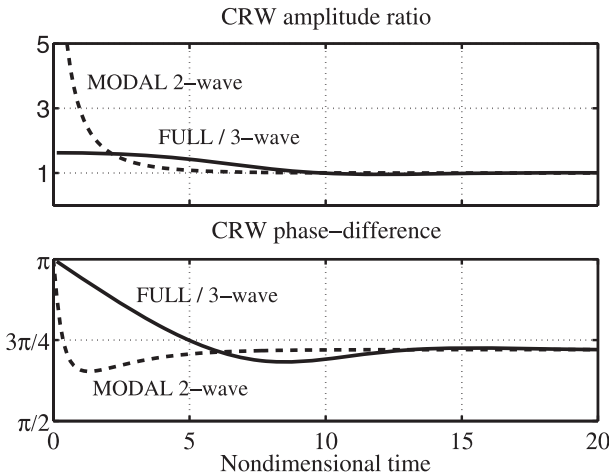


FIG. 3. The CRW (top) amplitude ratio and (bottom) phase difference, as obtained from the three-wave equations (identical to results from the full model projecting  $q - q_p$  onto CRWs) or the two-wave equations (dashed) as obtained using modal decomposition of the total perturbation  $q$  in which the CRWs evolve independently of the continuous spectrum.

must start at  $(r, \epsilon) = (2, \pi/2)$ . The amplitude of the KRW is constant but CRW-L grows; thus,  $r$  decreases with time. Their phase difference continually decreases so that the trajectory spirals clockwise. This spiral occurs mainly because the KRW is located at a height where the basic flow speed greatly exceeds  $c_{ii}$  and  $\bar{U}_i$  of the CRWs. As a result, the CRWs cannot keep pace with the KRW and the latter continues to pass over the CRWs. As discussed before, CRW-U is initially excited  $\pi$  out of phase with CRW-L, which is a hindering configuration. For any  $t > 0$ , the KRW also affects the CRW propagation: it hinders the propagation of CRW-

U while being of less influence on the propagation of CRW-L. The net result is that the phase difference between the CRWs rapidly decreases and approaches the gNM configuration  $\epsilon_+$  after approximately 5 time units (about 48 h), at which point their amplitude ratio  $r \sim 1$  (Note that the phase difference slightly overshoots for  $t > 5$  as a result of the by then small but non-negligible effects of the KRW).

Note that surface wind contribution of  $q_n$  is not zero (see curve labeled “N” in Fig. 4a). Figure 2 shows that the streamfunction associated with  $q_n$  remains almost untilted and approximately in quadrature with the initial KRW ( $q_p$ ), confirming that  $q_n$  is continually excited through advection by the KRW at  $z = H$ . Initially both the lower CRW and  $q_n$  have similar growth rates because they are being excited predominantly by the resonance mechanism (labeled II in the introduction) so the “N” curve maintains constant radius. Eventually, however,  $q_n$  is overwhelmed by the growth of the lower CRW due to the shear instability mechanism (I).

2) GROWTH MECHANISMS

On an  $f$  plane without a lid (e.g., the semi-infinite Eady model), the initial KRW would lead to a transient linear amplification of the boundary PV anomaly through the resonance mechanism (II). Because the lower CRW is dominated by the presence of its boundary contribution, the resonance-like growth occurring in the  $f$ -plane model remains approximately valid on the  $\beta$  plane, the edge wave being replaced by the lower CRW. On the  $\beta$  plane, an upper CRW is also excited just as an edge wave would be on the lid of the Eady model. As the phase difference between the CRWs reduces from  $\pi$ ,

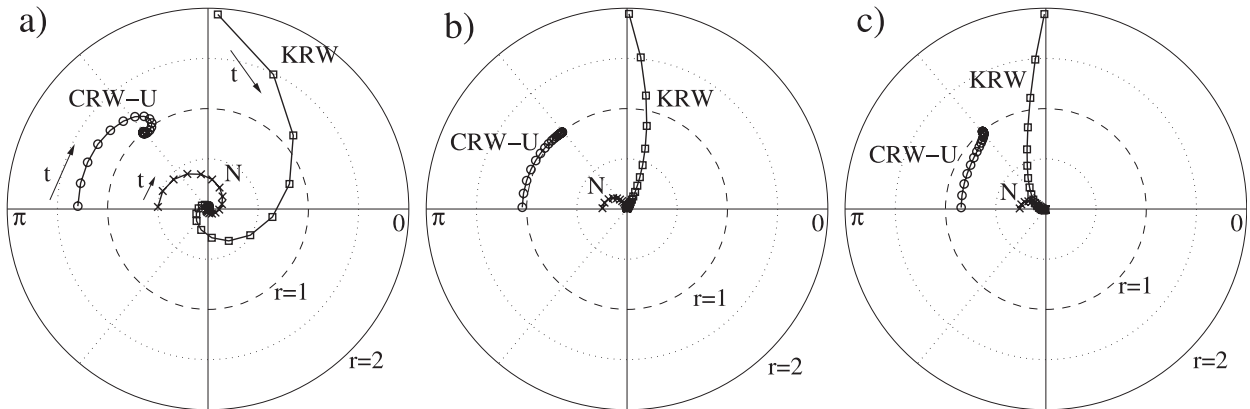


FIG. 4. Phase-space trajectories of the Rossby wave components from initial conditions with a KRW at (a)  $z = H$ , (b)  $z = z_s$  (the steering level of the gNM), and (c)  $z = z_s/2$ . The angle denotes the phase difference between surface wind,  $v(z = 0)$ , of a component and  $v(z = 0)$  of CRW-L. The radius denotes the ratio between  $|v(z = 0)|$  of a component and  $|v(z = 0)|$  of CRW-L, rescaled by (14); N refers to the neglected PV  $q_n$ , which typically has a tilt—the other components are untilted. The dots along trajectories are separated by 1 nondimensional time unit, starting at  $t = 1/8$ .

they can grow mutually through the shear instability mechanism (I). Eventually this dominates, but the resonance mechanism (II) can be important for a surprisingly long time.

To help provide more quantitative insight into the growth mechanisms, Fig. 5 displays the contributions of the KRW and other CRW to the growth rate of CRW-U and CRW-L. These contributions are computed easily from the amplitude equations of the three-wave system. Qualitatively, the evolution of the two CRWs is very similar. Because both have zero amplitude at initial time, their growth rate (thin lines) is very large for  $t < 2$  and gradually approaches the normal-mode value. The early growth of both CRW-U and CRW-L is caused by the KRW (dashed line). This is the equivalent of the resonant growth on the  $f$  plane. The growth of the CRWs due to their mutual reinforcing interaction is initially zero (they are generated  $\pi$  out of phase) and steadily increases toward the gNM value (full lines). As explained before [see (11)],  $q_n$  cannot contribute to the evolution of the CRWs because it is orthogonal to them.

### 3) VARYING THE HEIGHT OF THE INITIAL KRW

The phase-space evolution is qualitatively similar as the level of the initial perturbation is varied. Figures 4b,c show the phase-space evolution for an initial KRW located near the steering level  $z_s$  of the gNM and near  $z_s/2$  respectively.

In the steering-level case, the KRW is advected naturally with the phase speed of the gNM. Therefore, the phase difference between the KRW and lower and upper CRWs barely changes. The phase difference  $\pi/2$  is optimum for increasing the amplitude of the lower CRW by the resonance mechanism (II). The phase difference between the CRWs decreases, but at a (nearly) constant amplitude ratio, somewhat similar to the “synchronous” initial value problem discussed in Heifetz et al. (2004a). By comparing the separation of the dots in the CRW-U trajectories of Fig. 4a to those of Fig. 4b, it becomes apparent that the CRWs approach phase-locking less rapidly if the KRW is located at the steering level. This is because the KRW in the steering level case hardly influences the phase propagation of the CRWs. When the initial KRW is below the steering level (Fig. 4c), it is advected more slowly than the propagation rate of the lower CRW and cannot keep up with the emerging mode. As a result, the phase difference between the KRW and lower CRW increases with time (anticlockwise spiral). Note that CRW-U is now excited with smaller amplitude than CRW-L, reflecting the low altitude of the initial perturbation.

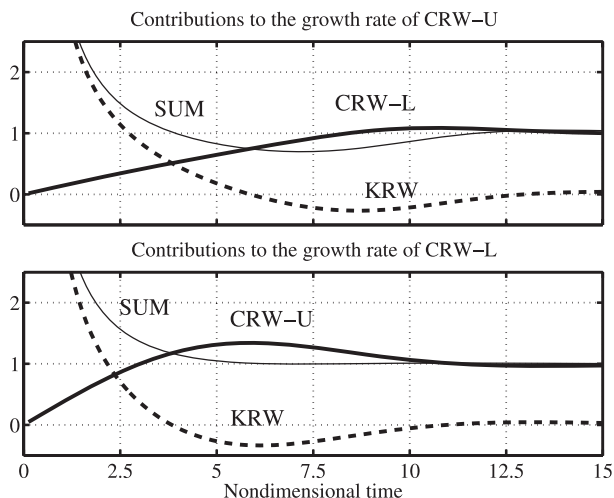


FIG. 5. Contributions to the growth rate of (top) CRW-U and (bottom) CRW-L arising from interaction with the KRW or other CRW, starting from initial conditions with PV only at  $z = H$ . The ordinate has been rescaled using the growth rate of the gNM. The different contributions have been labeled in the figure.

## 5. General initial conditions and the PAR-PV approximation

The last section showed that when the IC consists of a simple PV  $\delta$  function, the component associated with subsequent excitation of the continuous spectrum ( $q_n$ ) remains small compared to the other wave components at all stages in the evolution. Moreover it was shown that if  $q_n$  is ignored altogether, simple ODEs describe the entire evolution. In this section we hypothesize that more general initial value problems can also be described to a good approximation in terms of the coupled evolution of three Rossby wave components. Two of these components are the CRWs that are readily obtained from the gNM and the third is the passive PV component  $q_p$ . The above scheme assumes that the residual PV  $q_r = q_p$  (or equivalently that  $q_n = 0$  at all times); it will be referred to as the *passively advected residual PV* (PAR-PV) approximation.

### a. CRW evolution equations under the PAR-PV approximation

Once the passive PV component  $q_p$  has been identified (more about this in section 5b), its evolution is given simply by advection with the basic-state zonal flow. The winds induced by this passive component create meridional displacements (cf. section 3) and therefore excite the CRWs. The PAR-PV approximation neglects any excitation of CMs by  $q_p$ . Equations for the CRW amplitude and phase evolution follow by taking

the pseudomomentum bracket (B1)  $\eta_i(\partial\bar{q}/\partial y) \times$  (8) and making use of the orthogonality relations to get

$$\dot{\alpha} = \mathbf{L}\alpha + \mathcal{F}(t), \quad \mathbf{L} = -ik \begin{pmatrix} c_{11} & -\gamma_{12}/k \\ -\gamma_{21}/k & c_{22} \end{pmatrix}, \quad (15)$$

where  $\alpha = (\alpha_1, \alpha_2)^T$  and  $\alpha_j = a_j \exp(i\epsilon_j)$  denote the CRW amplitudes and phases. The components of the ‘‘forcing’’  $\mathcal{F}(t)$  are given by

$$\mathcal{F}_j(t) = -i \frac{\left\{ \eta_j, \frac{\partial \bar{q}}{\partial y} v_p(t) \right\}}{\left\{ \eta_j, \frac{\partial \bar{q}}{\partial y} \eta_j \right\}},$$

where  $\{, \}$  is defined in (B1) and  $v_p$  is the meridional wind associated with  $q_p$ ; that is,  $v_p(z, t) = ik \int G(z, z') q_p(z', t) dz'$ , with  $q_p(z, t) = q_p(z, 0) \exp[-ik\bar{u}(z)t]$ . Note that (15) is a generalization of (13) and could be written in similar form by separating it into real and imaginary components. The formal solution of (15) is

$$\alpha(t) = \mathbf{P}(t, 0)\alpha(0) + \int_0^t \mathbf{P}(t', 0)\mathcal{F}(t') dt', \quad (16)$$

where  $\mathbf{P}(t, t_0) = \exp[\mathbf{L}(t - t_0)]$  is the propagator. In general (16) cannot be integrated analytically and a numerical method must be employed. Nevertheless, the phase-space trajectories of the CRWs have a predictable shape (as illustrated for instance in Fig. 4), and on the basis of the Rossby wave properties the qualitative features of the general initial value problem can be anticipated in many cases without explicit calculation.

*b. Partitioning the initial perturbation PV*

The single KRW used as an initial condition in section 4 was not able to propagate relative to the basic flow, making it reasonable to assign all initial PV to the passive component:  $q_p(0) = q(0)$ . However, when  $\bar{q}_y \neq 0$ , it is not generally possible to distinguish the PV anomalies that are associated with meridional displacements of basic-state PV contours  $q_d$ , and those that are not,  $q_p$ . However, the partitioning influences the phase-space trajectories deduced for the CRWs, the relative contributions of different growth mechanisms, and the accuracy of the PAR-PV approximation, as seen in section 4. The approach taken here is to vary the partition between  $q_d$  and  $q_p$  systematically and seek the initial partition that minimizes the error made by the PAR-PV approximation.

One extreme is to assign the complete IC to  $q_p$ . This would be exact if the initial perturbation exists where  $\bar{q}_y = 0$  and is also appropriate when the IC is given by a single KRW, as in section 4. In this case the initial CRW

amplitudes must be zero. The opposite extreme is to project the full initial perturbation onto the CRWs (as with modal decomposition), which can only be associated with displacement PV. Intermediate ICs are represented by a single tuning parameter  $s \in [0, 1]$ , which varies the amount of projection of the IC onto the CRWs:

$$q_d(0) = s[A_1 q_1 + A_2 q_2],$$

where  $A_{1,2}$  are the projection coefficients of  $q(0)$  [NB: not  $q_d(0)$ ] onto the CRWs. Above it is implicitly assumed that all the PV that is not partitioned into the CRWs at initial time contributes to the passive PV component so that  $q_p(0) = q(0) - q_d(0)$ . Here  $s = 0$  implies  $q_p(0) = q(0)$  and  $s = 1$  implies subtracting the entire CRW projections from  $q(0)$ . Note that the  $s = 1$  assumption does not imply the same CRW evolution as obtained by modal decomposition; under the PAR-PV approximation,  $q_p(0) = q(0) - q_{\text{crws}}(0)$  still forces the CRWs following (15), whereas the CRWs in a modal decomposition follow the two-wave equations. The optimal partitioning of the initial condition is given by the value of  $s$  for which the error in the subsequent perturbation solution made by the PAR-PV approximation is a minimum. The schematic in Fig. 6 summarizes the essential aspects of the PAR-PV approximation.

Another, alternative partition is to vary the initial CRW amplitudes and phases such that  $|q_i(0)|^2$  is minimum, where  $|\cdot|$  indicates some norm (for instance, total energy). Although it is certainly possible to devise more sophisticated methods for determining an ‘‘optimal’’ partitioning,<sup>4</sup> these in general require a prior estimate of the full system’s evolution.

**6. Testing the PAR-PV approximation**

In this section the PAR-PV approximation is tested for more general ICs. As a measure of the accuracy of the PAR-PV approximation, we calculate perturbation total energy

$$E = -\frac{1}{2} [\{\psi, q\} + \{q, \psi\}] \quad (17)$$

of the reduced model and compare it with the result for the full model (3). The PAR-PV approximation gives exact results in the long-time limit when the gNM emerges as dominant. However, at intermediate times it deviates from the full model and the maximum energy contained in the neglected component  $q_n$  is measured.

<sup>4</sup> One such method is to choose the initial CRW projection coefficients such that they minimize the error measure over a finite time window, somewhat similar to a four-dimensional variational data assimilation (4D-VAR) approach.

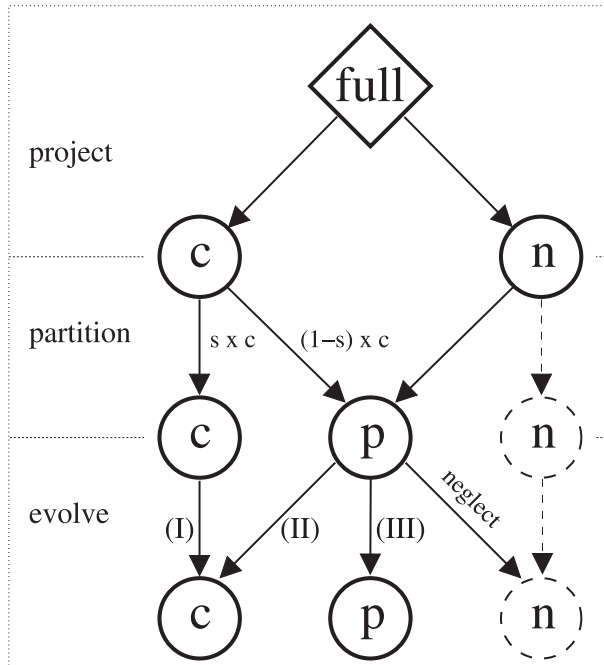


FIG. 6. Schematic overview of the PAR-PV approximation approach. In the first stage (“project”),  $q$  (full) is projected onto  $q_{\text{crws}}$  ( $c$ ) and  $q_n$  ( $n$ ) using the pseudomomentum brackets. In the second stage (“partition”), a fraction  $s \cdot q_{\text{crws}}$  ( $0 \leq s \leq 1$ ) is retained in  $q_{\text{crws}}$ , the fraction  $(1 - s)q_{\text{crws}}$  is assigned to  $q_p$  ( $p$ ), and all  $q_n$  ( $n$ ) is assigned to  $q_p$  ( $p$ ). At this stage no approximation has been made. In the third stage (“evolve”),  $q_{\text{crws}}$  ( $c$ ) and  $q_p$  ( $p$ ) are evolved in time. The PAR-PV approximation neglects the excitation of  $q_n$  by  $q_p$  (arrow from  $p$  to  $n$ ). Note that the components  $q_{\text{crws}}$  ( $c$ ) and  $q_n$  ( $n$ ) are not connected by arrows because they are orthogonal. Note that using a model approach,  $n$  would evolve independently of  $c$  (dashed arrow) and  $p$  would not be considered. The roman capitals refer to the growth mechanisms mentioned in the introduction (I for shear instability, II for resonance, and III for Orr mechanism).

### a. PV dipole

The first illustration is for a PV dipole situated between  $z \sim 0.2H$  and  $z \sim 0.5H$  (central location near the steering level of the gNM). Such structures could be created by diabatic PV production associated with latent heat release, as argued for example by Badger and Hoskins (2001). In the midtroposphere, ascent of moist air results in condensation and latent heat release. The diabatic PV tendency is proportional to the vertical heating gradient and is therefore positive below and negative above the heating.

As with the single KRW, for simplicity the IC is assigned entirely to  $q_p$ , which is subsequently advected passively (it will be verified a posteriori that this is a reasonable choice). Its evolution is shown in the top row of Fig. 7. At every instant  $q_d = q - q_p$  is projected onto the two CRW structures using the pseudomomentum brackets (9). The component  $q_n$  follows again simply

from (11). It is immediately obvious that the component  $q_n$  plays only a minor role in the entire time evolution, showing that the PAR-PV approximation ( $q_n \equiv 0$ ) can be used, just as in section 4, to track in a compact way the most important features and to understand their interaction. In this case, rapid initial growth occurs via the Orr mechanism, related to the tilting of  $q_p$  with the shear, followed by the resonance mechanism that excites the two CRWs; finally, the shear instability mechanism takes over as the CRWs interact. The Orr mechanism and resonance were together described as “growth by PV unshielding” by Badger and Hoskins (2001).

Note that PV anomalies are affected by meridional advection at all levels, as shown for instance by Robinson (1989), but the accuracy of the PAR-PV approximation in this case shows that the new displacement PV anomalies project almost entirely onto the CRWs rather than the CMs.

### b. Other disturbances

Similar results are obtained for other ICs, such as PV monopoles, the tripole IC shown in Methven and De Vries (2008), and different basic states such as the Green (1960) model (Eady model with  $\bar{q}_y = \beta$ ). In all experiments the wavenumber was set equal to that of the fastest-growing NM.<sup>5</sup> In each simulation the maximum energy fraction  $E(q_n)/E$  (in percent) was monitored using the various ways to partition the IC described in section 5b. This is a measure of the error associated with the PAR-PV approximation. Table 1 summarizes the results.

The first thing to notice is that for the majority of the cases the percentage errors are (much) smaller than 10%. Although not shown here, the maximum error usually occurs early in the time evolution, for  $t \leq 4$ . The PAR-PV approximation generally performs less well for the Charney model than for the Green model. Focusing on the columns for the optimal partitioning (labeled  $s = \text{opt}$ ), it is possible to see that the errors typically increase with the altitude of the initial perturbation in the Charney model. Both effects occur because the higher initial perturbations project less well onto the CRWs. Furthermore, the meridional winds that the passive PV induces do not match well onto any superposition of the CRWs, so a larger component of displacement PV must project onto the CMs and therefore onto  $q_n$ . The existence of the rigid lid of the Green model implies that almost any perturbation gives rise to boundary PV

<sup>5</sup> For the Green model, we located a rigid lid at  $H = 10$  km. The fastest-growing NM for  $\beta = 1.6 \times 10^{-11} \text{ s}^{-1}$  then has  $kL \sim 1.6$  ( $L = NH/f_0$ ) and a wavelength of about 4000 km).

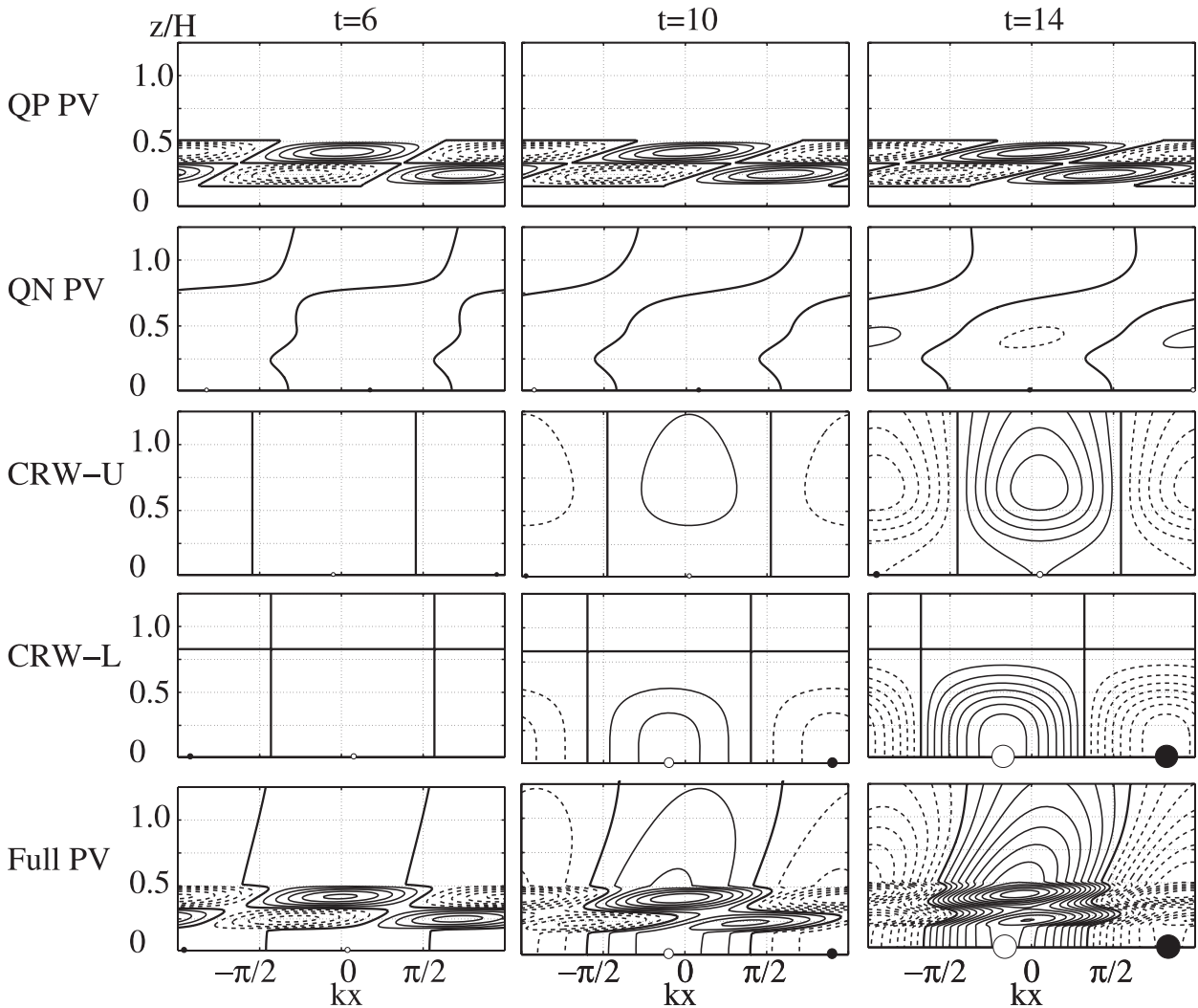


FIG. 7. Zonal cross sections of the PV (dashed contours for negative) starting from a PV dipole, shown at times  $t = 6, 10, 14$ . Circles show maxima (filled) and minima (open) in boundary PV, with radius indicating amplitude.

anomalies, which then project onto the upper CRW. In contrast, the Charney model is unbounded and so initial perturbations far above the maximum amplitude in the CRW structures can have small projections onto both CRWs. In the Green model, the PAR-PV approximation worsens if the tropopause height (upper rigid lid) is raised from an altitude of 10 to, say, 20 km or if the wavelength is decreased (not shown here). Other basic states have been investigated, and the PAR-PV approximation works best if the positive interior PV gradient is concentrated near a particular level representing the tropopause. The rigid lid is just the limit of the change in static stability across the tropopause tending to infinity.

Apart from monopole and dipole perturbations, other initial disturbances have been studied such as the optimal perturbation described by the leading singular

vector attaining greatest total energy growth over a finite time window. In case of the Green model, it is found that the evolution of the first SV is explained to within a few percent for any optimization time by the PAR-PV approximation. This result implies that the dominant SVs can be understood almost completely in terms of the three mechanisms mentioned in the introduction: (I) shearing over the initial PV passively, (II) exciting the CRWs via resonance and (III) CRW interaction at long times.

*c. Sensitivity to the initial partitioning*

Table 1 confirms that the choice  $s = 0$  (see section 5b for a definition of  $s$ ), used for obtaining the results of the PV dipole (Fig. 7), was reasonable but not the best possible option. For most ICs discussed in the table, the best results

TABLE 1. Maximum occurring energy of the component neglected by the PAR-PV approximation  $E(q_n)/E$  (in percent, rounded to one decimal) for various initial conditions in the Green and Charney model. For the Green model, the rigid lid is at  $H = 10$  km ( $kL \sim 1.6$ ), and 101 levels were used to discretize the vertical. For the Charney model, the domain extends to  $z = 5H$  [ $H = (f_0/N_0)^2 \Lambda/\beta$ ], and  $N = 301$  levels were used. The various columns show the different partitionings of  $q(0)$  as discussed in the main text. In the column  $s = \text{opt}$ , the  $s$  parameter (section 5b) is varied between 0 and 1 in steps of 0.1. The column header “alt” indicates the partitioning method that maximizes the initial energy associated with the CRWs (equivalent to minimizing the initial energy in the remainder  $q_r$ ). Specifications of the ICs:  $z_c/H$  denotes the center of the perturbation and  $w/H$  denotes its vertical extent (not relevant for the KRW experiments).

Structure		Maximum instantaneous value of $E(q_n)/E$ (%)							
$z_c/H$	$w/H$	Green model				Charney model			
		$s = 0$	$s = 1$	$s = \text{opt}$	alt	$s = 0$	$s = 1$	$s = \text{opt}$	alt
KRW									
0		0.4	0.0	0.0 ( $s = 0.9$ )	0.0	4.3	1.4	0.6 ( $s = 0.7$ )	0.5
1/4		0.0	0.3	0.0 ( $s = 0.1$ )	0.1	2.3	3.4	0.8 ( $s = 0.5$ )	1.2
1/2		0.1	1.4	0.1 ( $s = 0.0$ )	0.2	1.9	7.5	0.9 ( $s = 0.4$ )	1.9
3/4		0.6	0.5	0.3 ( $s = 0.5$ )	0.4	3.1	6.2	1.7 ( $s = 0.4$ )	2.4
1		3.0	0.1	0.1 ( $s = 1.0$ )	0.0	6.8	3.8	3.0 ( $s = 0.7$ )	2.4
Monopole									
1/4	1/2	0.1	0.4	0.0 ( $s = 0.2$ )	0.1	2.4	3.8	0.7 ( $s = 0.5$ )	1.2
1/2	1/2	0.1	1.1	0.1 ( $s = 0.1$ )	0.2	2.1	7.0	0.9 ( $s = 0.5$ )	2.0
3/4	1/2	0.6	0.5	0.2 ( $s = 0.5$ )	0.3	3.3	5.9	1.7 ( $s = 0.5$ )	2.4
1	1/2					7.5	3.6	3.1 ( $s = 0.8$ )	2.4
1/2	3/4	0.1	0.8	0.1 ( $s = 0.1$ )	0.2	2.3	6.3	0.9 ( $s = 0.5$ )	3.0
1/2	1	0.1	0.6	0.1 ( $s = 0.2$ )	0.2	2.5	5.5	1.0 ( $s = 0.5$ )	1.9
Dipole									
1/4	1/2	0.5	4.9	0.4 ( $s = 0.2$ )	0.3	10.6	6.8	3.2 ( $s = 0.6$ )	1.1
1/2	1/2	0.5	1.3	0.4 ( $s = 0.2$ )	0.4	3.1	20.1	1.0 ( $s = 0.3$ )	0.9
3/4	1/2	4.3	6.0	4.3 ( $s = 0.0$ )	3.0	4.3	13.4	3.5 ( $s = 0.2$ )	2.8
1	1/2					9.8	16.4	9.8 ( $s = 0.0$ )	9.9
1/2	3/4	0.8	1.3	0.4 ( $s = 0.4$ )	0.4	4.0	16.9	1.1 ( $s = 0.4$ )	1.1
1/2	1	1.2	0.9	0.4 ( $s = 0.5$ )	0.4	5.3	12.4	1.2 ( $s = 0.4$ )	1.2
Tripole									
1/2	1/2	0.4	10.7	0.4 ( $s = 0.0$ )	0.5	1.3	10.2	1.3 ( $s = 0.0$ )	1.0
Other									
TESV1, $t_{\text{opt}} = 5$		0.2	1.8	0.2 ( $s = 0.0$ )	0.2	12.2	12.1	9.1 ( $s = 0.8$ )	9.9
TESV1, $t_{\text{opt}} = 10$		0.1	2.2	0.1 ( $s = 0.0$ )	0.2	3.2	13.9	1.7 ( $s = 0.3$ )	2.2

were obtained when the initial CRW projection coefficients are determined either by the constraint of minimum energy contained in  $q_r(0)$  (listed under “alt” in the table), or for the optimal  $s$ -partitioning (listed under  $s = \text{opt}$ ). The latter technique however requires knowledge of the full evolution, which is a disadvantage.

## 7. Comparing the PAR-PV approximation to other reductions of perturbation evolution

The accuracy of the PAR-PV approximation is compared to results obtained with different truncated systems with a similar number of degrees of freedom. The results are given in Fig. 8 for two cases in the Charney model (the PV dipole of Fig. 7 and a PV monopole at the same location) and will be discussed below.

### a. Modal decomposition

First, the results of the PAR-PV approximation are compared to modal decomposition. If we let  $\mathbf{E}$  denote

the matrix of normal-mode PV eigenvectors  $\hat{e}_i$ , the IC decomposes as

$$\mathbf{q}(0) = \mathbf{E}\mathbf{b} \Rightarrow \mathbf{b} = \mathbf{E}^{-1}\mathbf{q}(0),$$

where  $\mathbf{b}$  is the projection coefficient vector in the modal basis. The modal coefficients evolve as  $b_i(t) = b_i(0) \exp(-ikc_i t)$ , where  $c_i$  is the complex phase speed of normal mode  $i$ . The simplest truncated model based on modal decomposition sets to zero all coefficients of  $\mathbf{b}$  not associated with the gNM and dNM. The resulting system evolves according to the two-wave equations, as already discussed [given by setting  $a_3 = 0$  in (13)]. This is the dotted line in Fig. 8. It is seen that by retaining no effects from the continuous spectrum at all, the initial total energy is overestimated in case of the monopole by more than 200%. A slight underestimation is seen in case of the dipole. In the latter case, however, the error rapidly increases after  $t = 0$  and it takes a very long time (more than 15 time units) for the error to become small

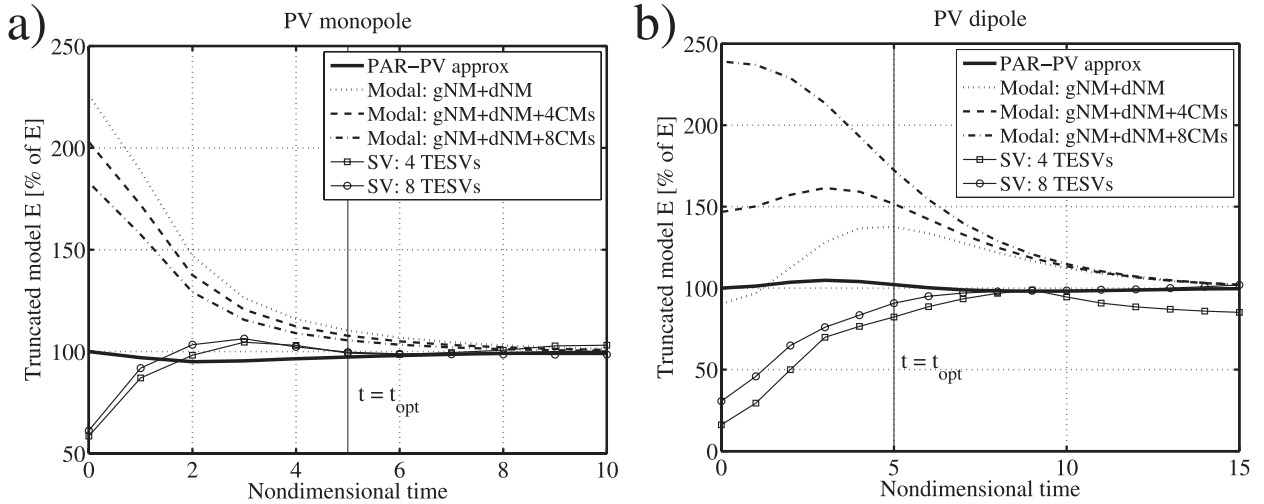


FIG. 8. Total energy ( $E$ ) vs time as obtained with different truncations of the full model (see labels) and measured relative to  $E$  of the full model integration. Initial conditions are (a) a PV monopole and (b) a PV dipole, as in Fig. 7. In the modal and SV cases, the initial condition has been projected onto only a subset of the full space (see text). The SVs were computed with an optimization time of 5 time units.

compared to the energy of the full model. The final time energy must converge to the full model as the gNM emerges as dominant. The gNM and dNM are orthogonal to the CMs (Held 1985) and therefore evolve independently, so truncation of the number of CMs included has no impact on the final gNM amplitude.

With the aim of better representing the short-time evolution, a limited number of CMs were included. The criterion for selecting these CMs is based on comparing the contribution of each CM to the initial total energy, which can be computed as

$$E_i = -\frac{1}{2} \{ \psi, b_i \hat{e}_i \} + \text{c.c.}, \quad (18)$$

where c.c. denotes complex conjugate. The energy of the full system (17) is recovered by adding the contributions from all normal modes  $E = \sum_i E_i$ . The initial energy contributions  $E_i$  are ordered and only  $n$  CMs ( $n = 4, 8$ ) with the largest  $|E_i|$  are retained.<sup>6</sup> In case of the monopole, it is seen that by including more CMs the approximation gets better. However, to obtain total energy errors as low as with the PAR-PV approximation, many CMs (for this resolution more than 75) need to be included. In contrast to the results for the monopole, which improved as more CMs were included, in case of the dipole the total energy error gets (much) larger if one includes only a limited number of CMs.

<sup>6</sup> Note that the normal modes are *not* orthogonal with respect to the energy norm (Held 1985) and therefore that this ordering varies over time.

Similar to the monopole case, many CMs (more than 50) are needed to obtain total energy errors as low as with the PAR-PV approximation.

### b. Singular-value decomposition

A second class of truncated model is constructed by computing the leading singular vectors of the model. The SVs are computed with standard techniques (e.g., Borges and Hartmann 1992), using a total energy norm and an optimization time of 5 time units. In contrast to the modal method, for which the selection criterion was based on contribution to the initial energy, here the first  $n$  (with  $n = 4, 8$ ) fastest-growing SVs are retained. The initial perturbation is subsequently projected on to this subset and the model is run forward again. For both the PV monopole and dipole (Fig. 8) now the initial energy is much smaller, reflecting the fact that the initial perturbation projects poorly on the leading SVs but also that the SVs amplify much more rapidly than the full perturbation. Unlike the modal approach, the SV approach is not guaranteed to give the correct total energy at final time. The gNM projects on many SVs. Therefore, by neglecting some of them we will not obtain the correct asymptotic answer. In the case of the dipole, the error at  $t = 5$ , the optimization time, is 20% if four SVs are used and 10% if eight SVs are used. Clearly, even the SV approach with eight SVs performs considerably worse compared to the PAR-PV approximation.

Note that the criterion for selecting the SVs was based on the ordering of the singular values. It can be verified that to achieve the same accuracy as the PAR-PV

approximation for explaining the PV monopole for instance, one needs to include more than 100 SVs. A method that reduces this large number of SVs involves first computing all SVs and subsequently ordering them in terms of their energy projection onto the initial perturbation [e.g., similar to (18)]. If that method is used, it is found that the current resolution requires about 25 SVs to achieve the same accuracy.

### c. Structural differences at initial time

In both modal and SV decomposition, the greatest differences occur at initial time. This is most apparent when the structures are examined. Figure 9 shows the dipole IC and its representation in different truncated bases. First note that the wind field of the dipole has greater amplitude below than above. This arises from the proximity of the lower boundary and its effects on the PV inversion.

When projected onto the gNM and dNM alone (Fig. 9b), it is clear that the structure of the upper CRW dominates the interior PV so that the top half of the dipole is spread across a great depth. The lower CRW is also present, introducing boundary PV anomalies even though none are present in the IC. As discussed before, the evolution of this initial structure is described exactly by the two-wave equations but is not a good description of the perturbation evolution until the gNM dominates ( $t > 10$ ).

The addition of four CMs to the gNM and dNM (Fig. 9c) localizes the initial PV anomaly so that it more closely resembles the top half of the dipole. This is achieved by the positive and negative spike structure about the steering level of each CM (shown in Fig. 10). The initial dipole projects most strongly onto the CMs that have phase speeds close to that of the gNM. However, as shown in Fig. 8, the total energy of this perturbation is far in excess of the dipole itself, reflecting the fact that this IC will grow more slowly than the dipole to achieve the same amplitude at final time.

The initial SV structures have PV that tilts upshear (Fig. 9d), optimizing finite time energy growth by the Orr mechanism. However, this structure obviously grows much faster than the dipole (Fig. 8) and bears little resemblance structurally. This is in agreement with Snyder and Hakim (2005; see also Hakim 2000), who note that generally a large number of SVs are required to simulate cyclogenesis from realistic initial conditions.

## 8. Conclusions

A new framework has been developed to simplify baroclinic initial value problems for small-amplitude perturbations to zonal shear flows. In all the experi-

ments examined, the full perturbation could be described to a good approximation by only three components: two counterpropagating Rossby wave structures (which are untilted) and a passive PV component  $q_p$  that simply tilts with the shear as though it were a passive tracer. The passive PV is dynamically active in that it excites the CRW structures continuously, as described by a simple system of coupled ODEs [section 5, Eqs. (15)]. Its initial distribution is assumed to equal the residual PV obtained after subtracting the initial perturbation associated with the CRWs. Therefore, the method is referred to as the *passively advected residual PV* (PAR-PV) approximation. It implicitly assumes that all PV anomalies generated through meridional advection project only onto the CRW structures; it neglects the generation of continuum modes by  $q_p$ . On the  $f$  plane, or more generally if  $\bar{q}_y$  vanishes in the interior (e.g., Lindzen 1994), CMs cannot be excited by advection and the PAR-PV approach gives exact results. However, if  $\bar{q}_y \neq 0$  in the interior, it constitutes an approximation to the full evolution.

The underpinning for the approximation stems from the relationship between PV anomalies and air parcel displacements and the orthogonality properties of NMs with respect to pseudomomentum (Held 1985). For the calculation of NM structures, it is assumed that all PV anomalies arise adiabatically from meridional displacement of basic-state PV contours (5). However, initial conditions can in general contain PV unrelated to displacements. It was shown in section 3 that this component,  $q_p$ , must evolve as if passive. The displacement PV projecting onto the CRWs must also be orthogonal to  $q_n$  (11), its projection onto the CMs. Therefore, the evolution of the CRWs is unaffected by  $q_n$  and hence the trajectories of CRWs in phase space under the PAR-PV approximation (Fig. 4) are exactly the same as from the full model.

The PAR-PV approximation enables partition of growth into three distinct mechanisms, just as for the simpler situation on the  $f$  plane (see the introduction). The interaction between the CRWs represents the shear instability mechanism (I). The excitation of CRWs by  $q_p$  represents the resonance mechanism (II), whereas the simple evolution of  $q_p$  (passive advection by the zonal flow) takes into account the Orr mechanism (III). The approximate set therefore quantifies the PV thinking concept (Hoskins et al. 1985) in time-dependent problems.

The PAR-PV approximation was found to be accurate in terms of energy to within a few percent for the initial value problems considered (see Table 1). Also, in the problems studied (section 7), the PAR-PV approximation provided a much more efficient and accurate framework for analysis than a reduction based on a modal decomposition (followed by truncation to a limited number of CMs) or singular-value decomposition

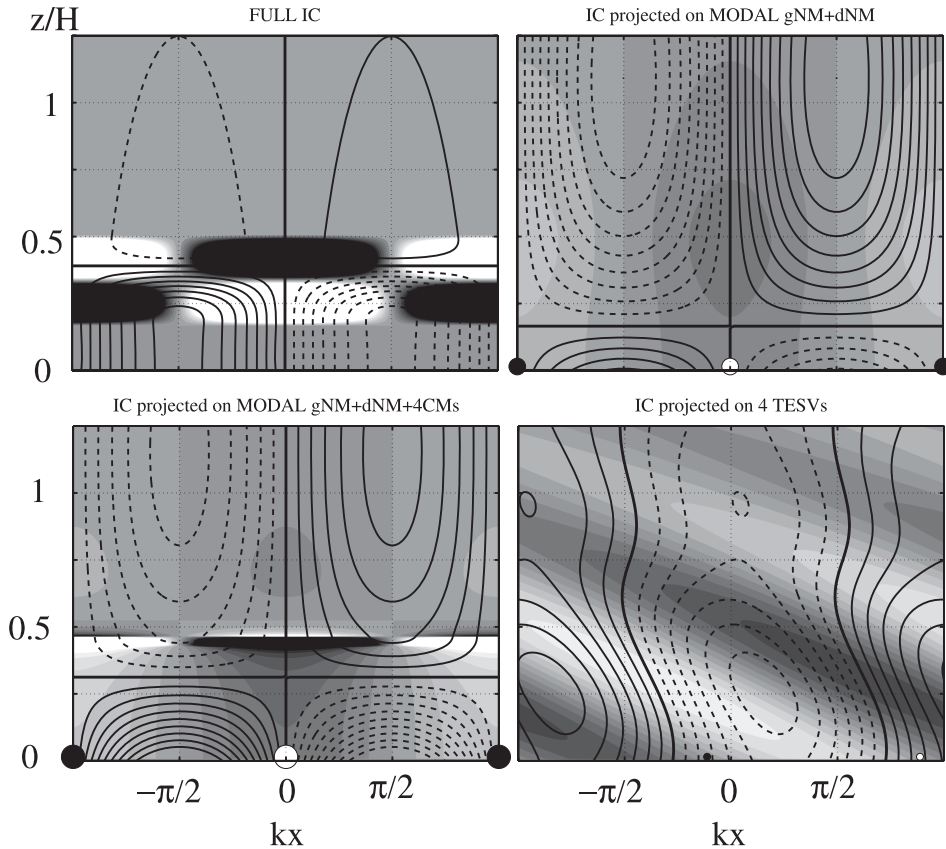


FIG. 9. Zonal–height cross sections showing the initial PV (shading; dark means positive) and  $v$  (contours; dashed means southward) for the (top left) PV dipole initial condition of Fig. 7 and (remaining panels) its representation following decomposition on to the modal or SV basis and truncation to a few components. Circles show boundary PV as in Fig. 7.

(followed by truncation to a few singular vectors). The PAR-PV approximation is also attractive because there is no need to calculate as many CM or SV structures as there are degrees of freedom in the (numerical) model. For a specified zonal wavenumber, it is only necessary to calculate the small subset of growing NMs (there is only one for the simplified states studied). In addition, the structure of CMs, obtained by eigenanalysis, is much more sensitive to model resolution than the growing NMs because the CM PV has finescale structure at their steering levels (see appendix A). Finally, the PAR-PV approximation accurately describes the evolution of individual singular vectors and thus provides interpretation of their growth, initially dominated by the Orr mechanism, followed by resonant growth and shear instability.

The PAR-PV approximation was shown to yield accurate results for the simple unstable basic states examined, starting from physically motivated ICs. Further investigation is needed to assess the relevance of the PAR-PV approximation for more general initial conditions and basic states. However, some limitations were

recognized on the basis of Table 1. The approximation worked better in the Green model in which a rigid lid is imposed on the Charney model. In the unbounded Charney model, the accuracy lessened as the initial perturbations were moved well above the CRW structures. However, in the atmosphere PV gradients are concentrated across the tropopause, and it is anticipated that the PAR-PV approximation would be appropriate for initial perturbations in the troposphere and not work well for those far above the tropopause. Far from boundaries (assuming uniform  $\bar{q}_y$ ), the PAR-PV approximation is likely to be invalid.

An example, mentioned in the introduction, is the Orr problem on the unbounded  $\beta$  plane; a monochromatic plane wave perturbation to a zonal flow with uniform shear. Because  $\bar{q}_y$  is uniform there are no unstable NMs or alternatively CRWs. The analytic solution presented in Yamagata (1976) and Boyd (1983) shows that the PV pattern shears over exactly as it would on an  $f$  plane but that the whole pattern retrogresses westward at a rate given by the usual Rossby wave dispersion relation,

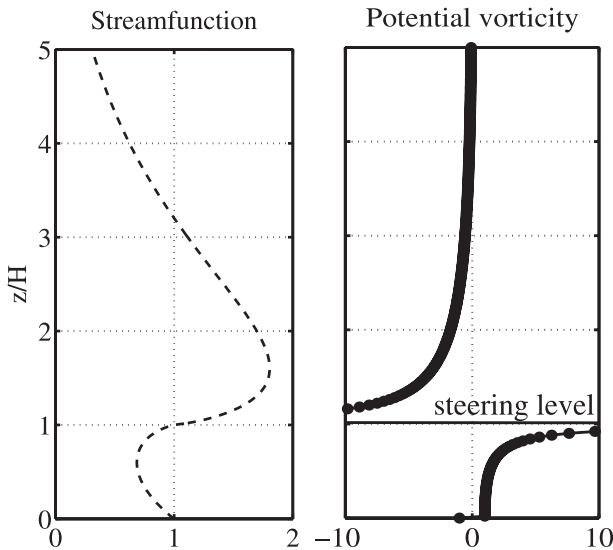


FIG. 10. Continuum mode in the Charney model with a steering level at  $z = H$ : (left)  $\psi$  and (right) PV (values larger than 10 are not shown).

allowing for the change in vertical wavenumber as the Venetian blind pattern tilts over. This result is readily understood if we consider KRWs. Consider the meridional velocity induced by a KRW at location  $z_0$  and the same for neighboring KRWs at  $z_0 - \Delta z$  and  $z_0 + \Delta z$ . The flow induced by the pair of flanking kernels cancels exactly at the PV maxima (along  $z_0$ ) for any  $\Delta z$  because of the symmetry of the tilted plane wave. Indeed, the meridional wind along  $z_0$  is in quadrature with the PV; therefore, the PV wave remains neutral but propagates relative to the flow, such that  $q(z, t) = q_0 \exp[i\epsilon(z, t)]$ . It is then easy to show that the instantaneous phase speed of each KRW, defined as  $c \equiv -\dot{\epsilon}/k$ , is given by<sup>7</sup>  $c(z, t) = \bar{u} - \beta \int G(z, z') q(z') dz' = \bar{u} - \beta/[k^2 + m^2(t)]$ , with  $m(t) = m_0 + k\Delta t$  being the vertical wavenumber ( $m_0$  is fixed by the tilt at  $t = 0$ ). As a result of the retrogression, the PAR-PV approximation fails in this example, although clearly it can be interpreted using Rossby wave components.

The weakest point of the PAR-PV approximation is that there is no unique way to partition an arbitrary initial condition into PV associated with displacements from a zonal reference state  $q_d$  and the remainder  $q_p$ . However, the choice of partition does influence the accuracy of the results when comparing the evolution of the reduced system under the PAR-PV approximation to the full model solution (see Table 1). Regions of

concentrated PV gradients (including the ground) in the basic state are crucial. If an initial perturbation is focused near the ground or the tropopause, it is reasonable to suppose that it will have a large projection onto the CRWs. In contrast, vertically confined PV anomalies in the midtroposphere are likely to have a small projection on CRWs and better results would probably be obtained by assuming that the initial perturbation projected mainly onto passive PV. This argument would carry more weight if it was also known that the PV anomaly had recently been generated diabatically, for example by latent heat release, for then it could not have its origins in the adiabatic displacement of basic-state PV contours.

The theory presented here is adiabatic and therefore would only be applicable if a PV anomaly had been generated but the diabatic process had halted. However, if latent heat release continues then the theory must be modified to account for the continuous PV anomaly generation. For example, this could be done using a parameterization assuming saturation in the ascending regions of a baroclinic wave and dry descent (e.g., Emanuel et al. 1987; Montgomery and Farrell 1991; Moore and Montgomery 2004) or by using a baroclinic wave conditional instability of the second kind (CISK) approach (Mak 1982; Snyder and Lindzen 1991; Parker and Thorpe 1995). One way of incorporating latent heat release effects into the PAR-PV approach would be to let them enter as a time-dependent source term in the zonally advected passive PV component. Moore et al. (2008) have recently shown an example of diabatic Rossby vortex in which the early stages of cyclogenesis appear to be described by the cooperative interaction between PV generated diabatically (by saturated ascent and condensation in the warm sector of a low level cyclone) and the potential temperature anomalies at the ground. Further research is required to determine whether this type of situation could be described quantitatively in terms of a few PV components governed by a reduced system of evolution equations in the spirit of this study.

There are many obstacles remaining before the theory outlined here could be applied in a diagnosis of atmospheric case studies. Extending the PAR-PV approximation to unstable zonal jets on a spherical domain should be possible following Methven et al. (2005). An important extension would be to include many zonal wavenumbers such that an analysis of local structures, wave packets, and downstream development becomes possible. Another critical issue is to what extent the PAR-PV approach is a good approximation at finite amplitude. This last issue is complicated by the thorny problem of partitioning a real flow into a “perturbation” and “reference state” component.

<sup>7</sup> Using  $G(z, z') = -1/(2k)\exp(-klz - z'l)$ , where the vertical coordinate  $z$  has been rescaled to  $Nz_*/f_0$ , for the Green function of the unbounded problem.

*Acknowledgments.* The authors acknowledge the National Environmental Research Council (NERC, Grant NE/D011507/1) for making the research possible. HdV acknowledges the Netherlands Organisation for Scientific Research (NWO) for a Rubicon Grant. JM is grateful for a Research Council UK Academic Fellowship. Thanks to two reviewers, especially John Nielsen-Gammon, for their constructive criticism that led to improvements in the article.

## APPENDIX A

### Continuum Modes for a Nonzero PV Gradient

Here we deduce the PV structure of continuum modes when  $\bar{q}_y$  is nonzero. CMs are (neutral) NMs of the linear dynamical equations that propagate with the zonal flow at a particular level  $h$ . When  $\bar{q}_y$  vanishes in the interior, CM has a PV  $\delta$  function at its steering level, as well as a nontrivial boundary PV contribution (see, e.g., De Vries and Opsteegh 2007a and references therein). The CM structure is necessarily different if  $\bar{q}_y = \beta$ . To make sense as a physical mode, the CM streamfunction must be continuous at  $z_s$ . It is generally found to be also nonzero at  $z_s$ , implying that the CM has nonzero meridional wind at  $z_s$ . On the  $\beta$  plane this meridional wind field immediately would start to advect any finite-amplitude PV anomaly, thereby violating the constraint of uniform phase speed. The only way to retain a uniform phase speed is if the PV becomes infinite at  $z_s$ , resembling the PV  $\delta$  function of the  $f$ -plane theory. On the  $\beta$  plane, however, the CM PV is generally also nonzero at all other interior levels. The above rationale is readily verified by substituting the normal-mode approach into (1). This gives

$$q(z) = -\beta \frac{\phi(z)}{\bar{u}(z) - c},$$

which indeed becomes singular for real  $c$  at a steering level  $z_s$ , where  $\bar{u}(z_s) - c \equiv 0$ . The above expression further shows that the PV changes sign across the steering level. If one for simplicity assumes that  $\phi(z)$  takes one sign throughout the domain, the PV structure of the CM can be anticipated. The numerically calculated structure of a CM is shown in Fig. 10.

## APPENDIX B

### CRW Structures Obtained Using Orthogonality

The meridional displacement structure of the lower and upper CRW is denoted by  $\eta_1(z)$  and  $\eta_2(z)$ , respec-

tively. Given  $\eta_i$ , the PV, meridional wind, and streamfunction structures are given by  $q_i = -\eta_i \bar{q}_y$ ,  $v_i = -k\psi_i$ , and  $\psi_i = \int G(z, z') q_i(z') dz'$ , respectively. The CRW structures are obtained from the gNM and dNM by imposing two constraints. Using the inner product

$$\{p, q\} \equiv \frac{1}{2} \iint p^* \bar{q} \bar{\rho} dy dz, \quad (B1)$$

the (globally conserved) total pseudomomentum is proportional to  $\mathcal{P} = \text{Re}\{\eta, \bar{q}_y \eta\}$ . The CRWs are made orthogonal with respect to the pseudomomentum bracket  $\mathcal{P}_{ij} \equiv \{\eta_i, \bar{q}_y \eta_j\}$ , such that  $\mathcal{P}_{ij} = 0$  for  $i \neq j$ . The CRWs are further made orthogonal with respect to the wind-weighted term appearing in the expression for pseudoenergy:  $\mathcal{W}_{ij} \equiv -\{\eta_i, \bar{u} \bar{q}_y \eta_j\}$ , such that  $\mathcal{W}_{ij} = 0$  for  $i \neq j$ . Because the CRW structures are real and untilted by definition, the factors  $\mathcal{P}_{ij}$  and  $\mathcal{W}_{ij}$  are real. By convention the CRW amplitudes are normalized such that they have equal and opposite pseudomomenta,  $\mathcal{P}_{11} = -\mathcal{P}_{22}$  (Heifetz et al. 2004a). The coefficients describing the propagation and interaction characteristics of the CRWs are given by

$$\gamma_{ij} = -\frac{\{\eta_i, \bar{q}_y v_j\}}{\{\eta_i, \bar{q}_y \eta_i\}}. \quad (B2)$$

Pseudomomentum density is defined as  $\frac{1}{2}[\bar{\rho}|\eta|^2 \bar{q}_y]$ . Being proportional to the PV gradient, it is typically positive in the interior and has a negative lower boundary contribution. It can be shown that the CRWs have the extreme values of pseudomomentum that can be obtained by combining the gNM and dNM subject to the orthogonality constraint  $\mathcal{W}_{ij} = 0$  for  $i \neq j$  (C. Bishop 2000, personal communication). Thus the lower CRW incorporates most of the gNM's boundary wave activity associated with negative pseudomomentum ( $\mathcal{P}_{11} < 0$ ), whereas the upper CRW concentrates where the gNM's pseudomomentum density is most positive ( $\mathcal{P}_{22} > 0$ ). All these conditions are satisfied by the edge waves of the Eady model. CRWs are generalizations of these components for basic states with nonzero PV gradients. See Heifetz et al. (2004a,b) for more detail.

## REFERENCES

- Badger, J., and B. J. Hoskins, 2001: Simple initial value problems and mechanisms for baroclinic growth. *J. Atmos. Sci.*, **58**, 38–49.
- Borges, M. D., and D. L. Hartmann, 1992: Barotropic instability and optimal perturbations of observed nonzonal flows. *J. Atmos. Sci.*, **49**, 335–353.
- Boyd, J. P., 1983: The continuous spectrum of linear Couette flow with the beta effect. *J. Atmos. Sci.*, **40**, 2304–2308.
- Bretherton, F. P., 1966: Baroclinic instability and the short wavelength cut-off in terms of potential vorticity. *Quart. J. Roy. Meteor. Soc.*, **92**, 335–345.

- Burger, A. P., 1962: On the non-existence of critical wavelengths in a continuous baroclinic stability problem. *J. Atmos. Sci.*, **19**, 31–38.
- Case, K. M., 1960: Stability of inviscid plane Couette flow. *Phys. Fluids*, **3**, 143–148.
- Chang, E. K. M., 1992: Resonating neutral modes of the Eady model. *J. Atmos. Sci.*, **49**, 2452–2463.
- Charney, J. G., 1947: The dynamics of long waves in a baroclinic westerly current. *J. Meteor.*, **4**, 135–162.
- , and M. E. Stern, 1962: On the stability of internal baroclinic jets in a rotating atmosphere. *J. Atmos. Sci.*, **19**, 159–172.
- Davies, H. C., and C. H. Bishop, 1994: Eady edge waves and rapid development. *J. Atmos. Sci.*, **51**, 1930–1946.
- De Vries, H., and J. D. Opsteegh, 2005: Optimal perturbations in the Eady model: Resonance versus PV unshielding. *J. Atmos. Sci.*, **62**, 492–505.
- , and —, 2007a: Interpretation of discrete and continuum modes in a two-layer Eady model. *Tellus*, **59A**, 182–197.
- , and —, 2007b: Resonance in optimal perturbation evolution. Part I: Two-layer Eady model. *J. Atmos. Sci.*, **64**, 673–694.
- Dirren, S., and H. C. Davies, 2004: Combined dynamics of boundary and interior perturbations in the Eady setting. *J. Atmos. Sci.*, **61**, 1549–1565.
- Eady, E. T., 1949: Long waves and cyclone waves. *Tellus*, **1**, 33–52.
- Emanuel, K. A., M. Fantini, and A. J. Thorpe, 1987: Baroclinic instability in an environment of small stability to slantwise moist convection. Part I: Two-dimensional models. *J. Atmos. Sci.*, **44**, 1559–1573.
- Farrell, B. F., 1982: The initial growth of disturbances in a baroclinic flow. *J. Atmos. Sci.*, **39**, 1663–1686.
- , 1984: Modal and non-modal baroclinic waves. *J. Atmos. Sci.*, **41**, 668–673.
- , 1989: Optimal excitation of baroclinic waves. *J. Atmos. Sci.*, **46**, 1193–1206.
- , and P. J. Ioannou, 1996: Generalized stability theory. Part I: Autonomous operators. *J. Atmos. Sci.*, **53**, 2025–2040.
- Fischer, C., 1998: Linear amplification and error growth in the 2D Eady problem with uniform potential vorticity. *J. Atmos. Sci.*, **55**, 3363–3380.
- Fjørtoft, R., 1950: Application of integral theorems in deriving criteria of stability for laminar flows and for the baroclinic circular vortex. *Geophys. Publ.*, **17**, 1–52.
- Green, J. S. A., 1960: A problem in baroclinic stability. *Quart. J. Roy. Meteor. Soc.*, **86**, 237–251.
- Hakim, G. J., 2000: Role of nonmodal growth and nonlinearity in cyclogenesis initial-value problems. *J. Atmos. Sci.*, **57**, 2951–2967.
- Heifetz, E., and J. Methven, 2005: Relating optimal growth to counterpropagating Rossby waves in shear instability. *Phys. Fluids*, **17**, 064107. doi:10.1063/1.1937064.
- , C. H. Bishop, B. J. Hoskins, and J. Methven, 2004a: The counter-propagating Rossby-wave perspective on baroclinic instability. I: Mathematical basis. *Quart. J. Roy. Meteor. Soc.*, **130**, 211–231.
- , J. Methven, B. J. Hoskins, and C. H. Bishop, 2004b: The counter-propagating Rossby-wave perspective on baroclinic instability. II: Application to the Charney model. *Quart. J. Roy. Meteor. Soc.*, **130**, 233–258.
- Held, I. M., 1985: Pseudomomentum and the orthogonality of modes in shear flows. *J. Atmos. Sci.*, **42**, 2280–2288.
- Hoskins, B. J., M. E. McIntyre, and A. W. Robertson, 1985: On the use and significance of isentropic potential vorticity maps. *Quart. J. Roy. Meteor. Soc.*, **111**, 877–946.
- Jenkner, J., and M. Ehrendorfer, 2006: Resonant continuum modes in the Eady model with rigid lid. *J. Atmos. Sci.*, **63**, 765–773.
- Kuo, H. L., 1973: Dynamics of quasi-geostrophic flows and instability theory. *Adv. Appl. Mech.*, **13**, 247–330.
- Lindzen, R. S., 1994: The Eady problem for a basic state with zero PV gradient but  $\beta \neq 0$ . *J. Atmos. Sci.*, **51**, 3221–3226.
- Mak, M., 1982: On moist quasi-geostrophic baroclinic instability. *J. Atmos. Sci.*, **39**, 2028–2037.
- Methven, J., and H. de Vries, 2008: Comment on “Piecewise potential vorticity inversion: Elementary tests.” *J. Atmos. Sci.*, **65**, 3003–3008.
- , E. Heifetz, B. J. Hoskins, and C. H. Bishop, 2005: The counter-propagating Rossby-wave perspective on baroclinic instability. Part III: Primitive-equation disturbances on the sphere. *Quart. J. Roy. Meteor. Soc.*, **131**, 1393–1424.
- Montgomery, M. T., and B. F. Farrell, 1991: Moist surface frontogenesis associated with interior potential vorticity anomalies in a semigeostrophic model. *J. Atmos. Sci.*, **48**, 343–367.
- Moore, R. W., and M. T. Montgomery, 2004: Reexamining the dynamics of short-scale, diabatic Rossby waves and their role in midlatitude moist cyclogenesis. *J. Atmos. Sci.*, **61**, 754–768.
- , —, and H. C. Davies, 2008: The integral role of a diabatic Rossby vortex in a heavy snowfall event. *Mon. Wea. Rev.*, **136**, 1878–1898.
- Morgan, M. C., 2001: A potential vorticity and wave activity diagnosis of optimal perturbation evolution. *J. Atmos. Sci.*, **58**, 2518–2544.
- , and C. C. Chen, 2002: Diagnosis of optimal perturbation evolution in the Eady model. *J. Atmos. Sci.*, **59**, 169–185.
- Mukougawa, H., and T. Ikeda, 1994: Optimal excitation of baroclinic waves in the Eady model. *J. Meteor. Soc. Japan*, **72**, 499–513.
- O’Brien, E., 1992: Optimal growth rates in the quasigeostrophic initial value problem. *J. Atmos. Sci.*, **49**, 1557–1570.
- Orr, W., 1907: The stability or instability of the steady motions of a perfect liquid and of a viscous liquid. *Proc. Roy. Irish Acad.*, **27**, 9–138.
- Parker, D. J., and A. J. Thorpe, 1995: Conditional convective heating in a baroclinic atmosphere: A model of convective frontogenesis. *J. Atmos. Sci.*, **52**, 1699–1711.
- Pedlosky, J., 1964: An initial value problem in the theory of baroclinic instability. *Tellus*, **16**, 12–17.
- Robinson, W. A., 1989: On the structure of potential vorticity in baroclinic instability. *Tellus*, **41**, 275–284.
- Snyder, C., and R. S. Lindzen, 1991: Quasi-geostrophic wave-CISK in an unbounded baroclinic shear. *J. Atmos. Sci.*, **48**, 76–86.
- , and G. J. Hakim, 2005: Cyclogenetic perturbations and analysis errors decomposed into singular vectors. *J. Atmos. Sci.*, **62**, 2234–2247.
- Yamagata, T., 1976: On trajectories of Rossby wave-packets released in a lateral shear flow. *J. Oceanogr. Soc. Japan*, **32**, 162–168.