Wave-turbulence interaction in shallow water numerical models: asymptotic limits, and subgrid interactions

Submitted by Andrew John Wheadon, to the University of Exeter
as a thesis for the degree of Doctor of Philosophy in Mathematics, February 2018.

Signature:

Abstract

The ability to directly simulate all atmospheric motion is currently well beyond the limits of the computers available to us. As such techniques must be developed that accurately model important processes in an affordable manner.

Large-scale balanced motion is well understood, but as affordable resolution increases, models are able to resolve scales where large-scale turbulence and small-scale waves are important. This requires a new set of techniques that respect the interactions between these different kinds of motion.

In this thesis we look at two ways of assessing the accuracy of models capable of representing the scales at which these interactions occur. The first approach uses asymptotic limit solutions to derive a set of terms whose scale is known. These terms can then be evaluated as the model approaches a relevant asymptotic regime, and a 'good' model should reproduce the expected rate of scaling.

We apply this method of asymptotic limit solutions to an Eulerian and a Lagrangian shallow water model. The former is based upon ENDGame, the model currently in use at the Met Office, and the latter is based upon a candidate model from GungHo which is seeking a replacement for ENDGame. In addition, the Eulerian model is evaluated with both small and large timesteps and the results confirm the ability of the semi-implicit scheme to retain accuracy at large timesteps. Errors in the higher-order diagnostics used in this section highlight the need to make these analytic diagnostics consistent with the discretisations of the model in question.

The second method involves looking at the exchanges of energy in a spectral shallow water model in order to inform the design of subgrid models. By running a high-resolution simulation and truncating the energy at a certain wavenumber, comparing the result to a run without truncation shows the contribution of the scales below the truncation limit. We extend this by separating the total energy into separate components that may be truncated and evaluated individually in order to give a more complete picture of energy exchanges at the subgrid scale.

Contents

Part I

Fundamental ideas

Chapter 1

Introduction

The atmosphere is a vastly complex system of interconnected mechanics with a huge number of potential inputs and variables. An understanding of the atmosphere is relevant to many areas of society, from prediction of extreme weather events to studying the impact of climate change, from mapping the water cycle to describing the weather of other planets.

Phenomena in the atmosphere occur across a wide range of scales, from planet-wide motion patterns down to hurricanes, to weather fronts and localised cloud patterns, examples of which can be seen in figure 1.1. Finally the Kolmogorov length scale is reached at around 1mm where turbulent energy is dissipated into heat. The picture is further complicated by the fact that these different scales can affect one another.

To understand the atmosphere as a whole, we would need to understand the dynamics operating at each scale. We would then be able to construct a model that could take a set of data as input and produce predictions about the future state of the system or interactions between different components. Many of the processes involved have been well studied and the fundamental set of physical laws governing mass, momentum and energy (known as the *primitive equations*) underpin all modern numerical models of the atmosphere [14]. Unfortunately our knowledge is still limited, and perhaps always will be, so a fully accurate model lies beyond our grasp.

An ideal model would allow all scales to be simulated. However, the finest scale that is computationally feasible in production models is typically around 1km and in most cases is much coarser. Clearly this will be insufficient for representing motion at the Kolmogorov scale, and for this reason some form of additional subgrid model or implicit parametrisation is required to simulate the effect of the unresolved scales upon the resolved ones.

It is evident that a balance must be struck between accuracy and cost when designing a model. By focussing attention on specific processes, we may simplify

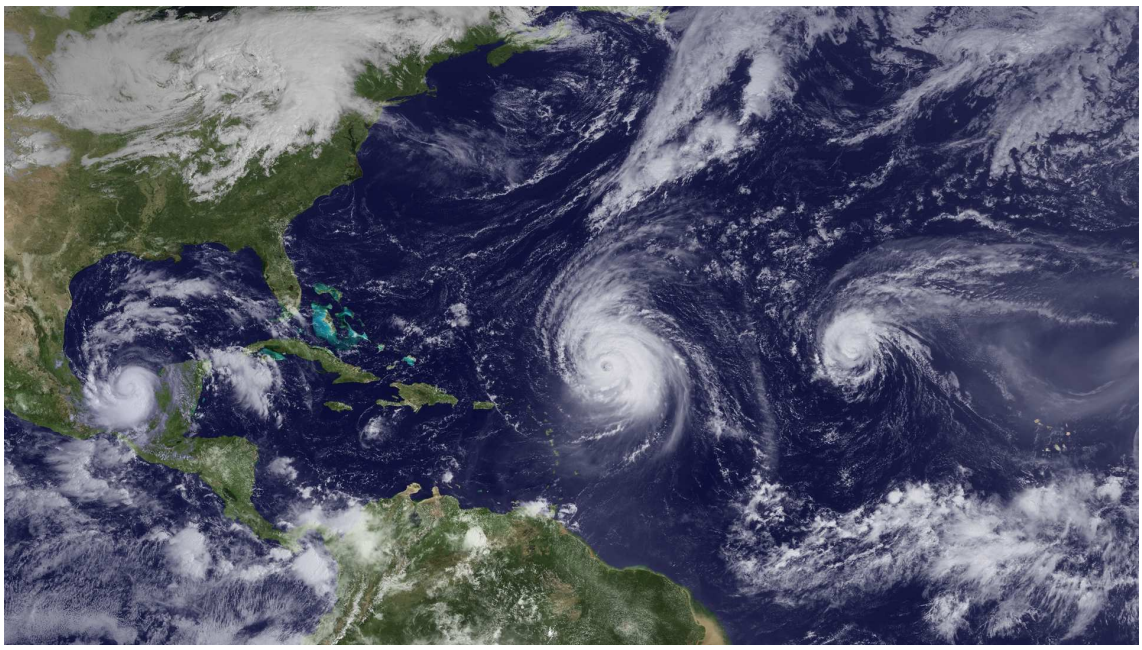

Figure 1.1: Hurricanes Karl, Igor and Julia, September 2010 (NOAA).

the model by identifying those parts of the model that have the greatest effect on such processes and retaining them while less important effects may be approximated or ignored. For this we need a way of measuring the relative importance of factors within a system.

At a given scale certain assumptions may be made about the characteristics of the flow in order to enable practical modelling. For example, at global scales the atmosphere is strongly stratified and the planet's rotational effects are important. The balance between these forces is given by the Burger number

$$B = \frac{NH}{fL}.$$

N is the Brunt–Väisälä frequency, the frequency at which a parcel of air will oscillate when displaced vertically within a non-turbulent fluid. H and L are typical vertical and horizontal length scales respectively. The Coriolis parameter f is a measure of the locally vertical component of vorticity at a latitude φ, such that $f = 2\Omega \sin \varphi$ where Ω is the rotation rate of earth. At global scales on earth, $N \sim 10^{-2} s^{-1}$, $f \sim 10^{-4} s^{-1}$ and $H/L \sim 10^{-3}$ [54]. This gives a Burger number of approximately 0.1 and hence the fluid motion is dominated by vortical motion. The Burger number may also be expressed in terms of the Rossby (Ro) and Froude (Fr) numbers:

$$B = \frac{Ro}{Fr}, \text{ where } Fr = \frac{U}{NH} \text{ and } Ro = \frac{U}{fL}.$$

The Froude number is a measure of the ratio of the inertial to gravity forces in a flow. In three-dimensional flow it may also be used to quantify the strength of stratification [27]. The Rossby number may be thought of as a measure of how

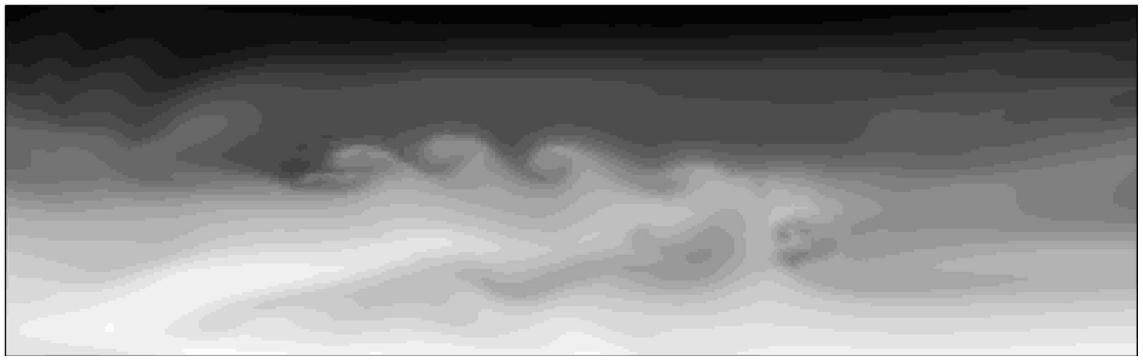

Figure 1.2: Kelvin-Helmholtz formations shown in the density field of a strongly stratified flow [37].

significant the effects of rotation are within a flow.

If the Rossby number is small, a flow is said to be near geostrophic balance, where the Coriolis force is balanced by the force resulting from differences in pressure (known as pressure-gradient force). Thus an approximation of geostrophic turbulence such as the quasigeostrophic (QG) approximation may be used, whereby the horizontal wind and advection are replaced by their geostrophic counterparts. As we approach the mesoscales (around 1000km in the atmosphere), the Burger number approaches unity, and the effects of stratification become more important, as can be seen in figure 1.2.

These nondimensional numbers then are a useful way of classifying the behaviour of a system. By understanding how the significance of different processes varies with these numbers we may define certain asymptotic regimes. Depending on the values chosen, some terms in the governing equations will become very small. These small terms may be moved to the right hand side of the equation and hence define the "imbalance" in the terms on the left hand side. Scale analysis can show how these imbalance terms are expected to scale as relevant asymptotic limits are approached. A model should respect these asymptotic limit solutions without explicit reliance on resolution or the accuracy of subgrid models. For a detailed introduction to this method see Cullen [8], as well as examples for the semigeostrophic approximation for flow over a ridge [10] and a demonstration that the semigeostrophic solution to the Eady frontogenesis problem converges to the full Euler solution when Ro is small [11]. These works were extended by Visram, Cotter and Cullen [53] who showed that the expected convergence rate can be achieved despite the presence of discontinuities and past the point of frontal collapse.

A method such as that above is a useful test for a model, but the information it can tell us about the actual processes involved is somewhat limited, especially if

we are concerned with those terms that are very small. At such small scales waves become important, and at even smaller scales both waves and vortices break down into turbulence. Finally, at the smallest scales viscous dissipation occurs. It is logical that behaviour at two different scales may not necessarily conform to the same assumptions (geostrophic balance is not valid at small scales, for example), and so at unresolvable scales we may have important processes happening (such as action to remove energy from a system) that bear no resemblance to those that can be directly resolved.

The transfer from scales at which waves dominate to those at which turbulence dominates implies a scale at which the waves and turbulence interact. The dynamics at this "macroturbulence" scale are an area of much interest (see e.g. [35]) and in two-dimensional turbulence it is related to the topic of energy cascades.

Scalar interactions are not simply related to spatial scales: at large scales the atmosphere acts as a two-time-scale system, whereby its dynamics may be separated into slowly-evolving balanced flow and fast inertia-gravity waves. The separation between these two types of motion may be estimated by the Rossby/Froude numbers, small Ro or Fr indicating large separation and vice versa. Although these two time scales are sometimes so significantly separated as to almost decouple from one another, interaction between the balanced and unbalanced flow provides a mechanism for spontaneously generating fast waves [52].

The effects of temporal and spatial scales may be understood by comparison between models running at different resolutions. For example, models running with large time-steps may not capture the fast motion that models with a smaller time-step can. In a similar manner high spatial resolution simulations may capture more processes than low-resolution models can. It would be useful then to have a robust method of examining how the numerical solution of a system changes as more processes are resolved.

1.1 Background

Questions regarding the transfer of energy across different scales have been the subject of much research for a long time, and without understanding the effects of these energy transfers we impose a limitation to the accuracy our models can achieve.

In 1920, Richardson [36] describes how instabilities in a large-scale flow lead to energy transfer to eddies of smaller scales which in turn transfer their energy downscale until viscosity effects become important and the energy is dissipated. Vortex stretching is an important element of this process as it creates eddies and

energy at small scales.

These ideas allow one to calculate the energy spectral density $\mathcal{E}(k)$ for turbulent flow using dimensional analysis. Such spectra are specified in the inertial range, so called because it exists between the large and small scales, where inertial terms dominate and the effects of forcing and dissipation are negligible. This was achieved for the energy cascade in a general sense by Kolmogorov [23], and Obukhov phrased it in the form $\mathcal{E}(k) = \mathcal{K}_\varepsilon \varepsilon^{2/3} k^{-5/3}$ [32], where ε is the energy flux, k is the horizontal wavenumber and $\mathcal{K}_\varepsilon$ is a dimensionless constant found to be around 1.5 in 3D turbulence. An analagous spectrum may be given in terms of the 2D enstrophy as $\mathcal{E}(k) = \mathcal{K}_\eta \eta^{2/3} k^{-3}$, where η is the enstrophy cascade rate.

In 1985 an analysis of over 6000 commercial aircraft flights, performed by Nastrom and Gage, demonstrated a -5/3 spectral slope in the atmospheric mesoscale at wavelengths in the range 2.6 to 300-400km and noted that at larger scales the slopes approach -3 [31]. By accounting for the conversion of kinetic to potential energy, Lumley [29] was able to derive a spectrum proportional to k^{-3} at low wavenumbers $k \ll k_b$ that transitioned to the $k^{-5/3}$ form at higher wavenumbers $k \gg k_b$, where k_b is the buoyancy or Ozmidov wavenumber

$$k_b = \left(\frac{N^3}{\varepsilon} \right)^{1/2}.$$

This result has been confirmed in a large eddy simulation [5]. By considering the (small) length scales at which viscosity effects become important, we may also derive the wavenumber at which dissipation effects take hold (e.g. [50]), given as

$$k_v = \left(\frac{\varepsilon}{\nu^3} \right)^{1/4},$$

where ν is the kinematic viscosity. This is also known as the Kolmogorov wavenumber. These results give a clear indication of how the nature of a flow differs with scale.

While the synoptic k^{-3} spectrum is generally agreed to be dominated by balanced motion [7], explanations for the origin of these spectra is an area of continuing interest and debate. They are generally attributed to internal wave processes, and accurate numerical reproductions of them have been the subject of several papers (e.g. [40], [49] and [41]), and there is an implication that the mesoscale $k^{-5/3}$ is dominated by unbalanced motion. Waite and Bartello [55] perform a set of simulations of stratified turbulence forced by vortical motion at a range of Froude Numbers that show reasonable reproduction of the observations for weak stratification. In the limit of strong stratification however, the results from the simulations do not resemble the observations at all. The paper also notes that k_b introduces a scale below which stratification is unimportant and isotropic three-dimensional turbulence exists. This scale is unresolved in the majority of their

simulations. This analysis was continued in the case of stratified turbulence generated by internal waves [56]. Once again, the results are inconsistent with most observations. In both of the papers, the authors suggest that higher resolutions may be sufficient to reproduce the observed energy spectra. The effect of rotation on stratified turbulence has been investigated by keeping the Froude number small and varying the Rossby number [57]. The limit exists between the QG limit and the non-rotating case, and is therefore in the region where the stratified turbulence spectrum ($k^{-5/3}$) transitions to the QG spectrum (k^{-3}). They conclude that the transition from stratified to QG turbulence occurs around the macroscale, velocity-based Rossby number $Ro_u = 0.4$ and the microscale, vorticity-based Rossby number $Ro_\omega = 3$. Further, it is suggested at high resolutions the vertical grid spacing should be sufficiently fine so as to resolve overturning, which occurs at $H \sim U/N$, where H is the vertical length scale, U is the root mean square velocity and N is the Brunt-Väisälä frequency. In the atmosphere this corresponds a vertical length scale of $O(1)$km.

In three-dimensional turbulence, energy is cascaded downscale as described above. However, in some respects the atmosphere and ocean behave as two-dimensional, which results in some very different dynamics. In two dimensional turbulence, the vortex stretching term vanishes on average and this leads to an infinite family of conserved moments of vorticity, of which a particularly important one is the enstrophy. In two-dimensional flow, we find that energy actually transfers upscale, and it is enstrophy that cascades downscale. Waite and Bartello note that the transition from stratified to QG turbulence is manifested by such an inverse cascade [56]. This fundamental difference has a number of implications: Energy must now be removed at large scales, and enstrophy at small scales. This can lead to the formation of two separate inertial ranges for the energy and enstrophy. Vallis [50] calculates the energy spectrum of the enstrophy inertial range as $\mathcal{E}(k) = \mathcal{K}_\eta \eta^{2/3} k^{-3}$ and the energy inertial range as $\mathcal{E}(k) = \mathcal{K}_\varepsilon \varepsilon^{2/3} k^{-5/3}$ (notice the similarity to the 3D case, except here the energy cascade is upscale). In addition, he also defines the frictional wavenumber

$$k_r = \left(\frac{r^3}{\varepsilon} \right)^{1/2},$$

where r can be thought of as a drag coefficient. At scales larger than k_r^{-1} frictional effects become important. The eddy turnover time derived from the enstrophy inertial range is found to be independent of length scale, in contrast to the situation in three-dimensional flow.

Another important example of the interactions between different scales within two-dimensional flow (though not limited to 2D) is the spontaneous emission of inertia-gravity waves (broadly, a gravity wave large enough to be affected by Co-

riolis force, henceforth referred to as IGWs) from vortical flow. This concept is linked with ideas of balanced flow and the slow manifold ([26], [28]), that being an invariant manifold free from IGWs. The existence of such a slow manifold is investigated in Ford, McIntyre and Norton [17] in a regime equivalent to the barotropic vorticity regime we will discuss later. By the method of matched asymptotic expansions, the authors are able to obtain a mathematical description of the spontaneous emissions up to $O(Fr^4)$, where the Froude number is small. This result contradicts the basis of the slow manifold and it is concluded that, "it seems practically certain that the entity traditionally called the slow manifold - whose practical usefulness is not in question - is not, in fact, a manifold. We therefore suggest...that this entity might be referred to as the slow quasimanifold.". This is similar to the findings of Thomas [43], which concluded that in the quasigeostrophic case with initially unbalanced data interactions between the fast and slow dynamics make the existence of a slow manifold impossible in the long-time limit.

Vanneste has investigated spontaneous wave emission in the small Rossby number regime in [51] and [52]. It is demonstrated that a hierarchy of slow quasimanifolds can be used to approximate balanced motion up to $O(Ro^N)$ where N is arbitrarily large - though the accuracy is limited due to the nonexistence of a true slow manifold. The result of this is that even perfectly balanced initial conditions will generate IGWs, but that they will be exponentially weak. This work is an interesting comparison to Ford et al. [17], where in the small Froude number regime there is not a timescale separation at every scale which ultimately leads to stronger wave generation, with a power law scaling for the wave amplitudes rather than exponential. The implication is that significant IGW generation can only take place when the Rossby number is not small.

Concepts like energy cascades and spontaneously generated waves involve interactions between processes at different scales. It is not possible, currently, to directly resolve the smallest of these scales and so they fall into the 'subgrid scale'. Numerical models approximate these subgrid terms as best as they can, but they may struggle to fully emulate the true nature of the processes as turbulence remains an unsolved problem. One approach is Implicit Large Eddy Simulation, an example of which is the work by Kent, Thuburn and Wood [22], who looked at a case of two-dimensional turbulence using the barotropic vorticity equation. One of the results of this paper was that while the models replicated the downscale enstrophy cascade, none could capture the upscale energy transfer from unresolved to resolved scales.

The examples presented here indicate how large the topic of scale interactions is. Even as our knowledge increases, we find contradictions between theory

and observations and the fundamental mechanics driving these processes are still somewhat unknown. As the resolvable limit for our models increases, more and more we will be beyond the balance-dominated part of the spectrum and into the wave-dominated. Not only do we need a better understanding of how processes interact when both waves and turbulence are important, but our subgrid models also need to improve.

In this thesis we will examine the topic of wave-turbulence interactions in two ways:

By scale analysis of terms in a shallow water model, we derive diagnostics that are expected to scale in a predictable way as the model approaches an asymptotic limit of the governing equations, such as the QG limit. This method will be applied to two different shallow water models. The general technique used here is quite broadly applicable, and provides a relatively simple way of assessing whether models respect asymptotic limits.

Further, we build upon the work by Kent et al. by looking at a spectral shallow water model and decomposing the energy into various components. By truncating the energy below a certain wavenumber and comparing to the untruncated energy spectrum, we can form a picture of what the contribution of the scales below the truncation wavenumber is. The decomposition allows us to go further by truncating specific parts of the energy spectrum (say, the energy related to divergence) and looking at the effect upon other components. By this method we not only can learn how these components interact with each other, but we can use this information to aid in the construction of more effective subgrid models.

1.2 Thesis Outline

This thesis is structured as follows:

Chapter 2 provides an introduction to the use of asymptotic limit solutions in determining the accuracy of models. The shallow water equations are introduced and a number of basic features of the equations are outlined. Particular attention is given to the ability to classify a flow in terms of its Rossby (Ro) and Froude (Fr) numbers. Some simple examples are given that highlight some considerations regarding the differences between analytic solutions and the discretised form the model will use.

Chapter 3 provides more detail on the experimental setup. The 2D shallow water equations are given for each of the models (one Eulerian, one semi-Lagrangian) we will be looking at, and a number of diagnostics are discussed.

We also introduce a possible means of quantifying gravity wave activity in a flow.

The results for the semi-Lagrangian finite difference model, ENDGame, are presented in chapter 4. This includes field maps showing the evolution of the flow, time series of a number of diagnostic quantities discussed in chapter 3, and convergence diagrams that attempt to show if the model converges to the asymptotic limit solution at the expected order. Where this is found not to be the case, an attempted explanation is provided.

The same structure holds for chapter 5, which looks at an Eulerian finite element model developed as part of the GungHo project. As the model is perhaps not as intuitive in its structure, extra detail is provided as to its design. Results for the time series and convergence are presented and discussed as before.

Chapter 6 summarises the results of part II and identifies areas that warrant further study.

Chapter 7 provides the introduction to Part III with a brief overview of turbulent cascade theory. Aspects of this chapter will be helpful in setting up the experiment and also interpreting its results.

Chapter 8 looks at a number of ways in which we can define and decompose energy in a spectral model. This chapter also provides an overview of the scheme that was used in part III.

Further details on the experimental setup of part III are provided in chapter 9. This includes specifics of the forcing, dissipation and process for running the model.

The results are provided in chapter 10, broken down by regime. There is discussion both of the characteristics of the energy spectra and energy tendency diagnostics, and also of the effects that truncation has on each of the tendency diagnostics.

Chapter 11 discusses key findings from the previous chapter and identifies important energy transfers. These results are compared to previous work and some thought is given to further ways in which this effort could be extended.

Chapter 12 summarises the thesis and justifies the work we have done.

Chapter 2

Numerical considerations

One of the key aims of this thesis is to examine the use of asymptotic limit solutions as a method for assessing the accuracy of numerical models. As it is necessary to validate the accuracy of proposed operational models, a logical approach is to compare the results of the operational model to that of a simplified model at a relevant asymptotic limit. This is because these solutions have known accuracy at affordable resolutions [8]. As such, it will be helpful to have some context for the effects that the choice of model can have on such solutions.

In this chapter we will introduce some basic ideas that will be helpful in Part II, and demonstrate some of the consequences of this theory with a simple one-dimensional system.

The first part of this project will be to construct a one-dimensional shallow water model that will be used to demonstrate with simple examples some fundamental ideas and issues with the development of numerical schemes, and how they might fail to respect the asymptotic limit solutions. We will then look at a selection of current numerical schemes and examine their behaviour under certain parameter regimes of low Rossby (Ro) and low Froude (Fr) numbers. These results will help us to interpret more complex simulations in later parts of the project. Of particular interest is the ability of these schemes to capture both balance and spontaneous wave generation.

2.1 The governing equations

The standard equations for fluid dynamics (see [58] for example) are given by the momentum equation

$$\rho \frac{D\mathbf{u}}{Dt} + \nabla p - \rho \mathbf{F} = 0, \tag{2.1.1}$$

and the mass (or continuity) equation

$$\frac{D\rho}{Dt} + \rho\nabla \cdot \mathbf{u} = 0, \tag{2.1.2}$$

where ρ is fluid density, $\mathbf{u} = (u, v, w)$ is the three-dimensional velocity vector, p is the pressure and $\mathbf{F}$ represents additional forces such as gravity or viscosity. The Lagrangian form is used here where

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla.$$

These equations are typically combined with the first law of thermodynamics and the ideal gas law, which have been omitted here for simplicity.

2.1.1 Large horizontal scales

The shallow water equations are valid in cases where the horizontal scales are much greater than the vertical scale. In such a case the mass equation implies that vertical velocities are small and thus when gravitational force is considered the vertical momentum equation can be approximated by hydrostatic balance:

$$\frac{\partial p}{\partial z} = \rho g. \tag{2.1.3}$$

2.1.2 Incompressibility

The shallow water equations by design apply to a single layer of fluid that is unaffected by stratification, thus the density must be constant. As the depth of the fluid layer is likely to be insignificant in comparison to the radius of the planet, we can also assume that gravity is constant, and we can integrate this over the fluid depth and substitute into the horizontal momentum equation to give

$$\frac{D\mathbf{u}}{Dt} + \nabla_H \Phi + f\mathbf{k} \times \mathbf{u}_H = 0, \tag{2.1.4}$$

where $\Phi = gh$ is the geopotential with $h = h(x, y, t)$, the height of the fluid and g the acceleration due to gravity. The gradient operator and velocity are now horizontal such that $\nabla_H = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right)$ and $\mathbf{u}_H = (u, v)$. Coriolis forces have also been included.

Constant density transforms the mass equation to

$$\nabla \cdot \mathbf{u} = 0. \tag{2.1.5}$$

We can write this as

$$\frac{\partial w}{\partial z} = -\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}, \tag{2.1.6}$$

and integrate over the fluid depth to give

$$\frac{Dh}{Dt} + h\nabla \cdot \mathbf{u} = 0 \qquad \text{or} \qquad \frac{D\Phi}{Dt} + \Phi \nabla \cdot_H \mathbf{u} = 0. \tag{2.1.7}$$

Divergence is also now horizontal so $\nabla \cdot_H = \left(\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right)$. (2.1.4) and (2.1.7) together are the shallow water equations. By stacking multiple layers of shallow water on top of each other with different densities the effects of stratification may also be modelled.

Note: As the rest of this paper deals with two dimensional equations, the horizontal subscripts will be dropped for convenience. Additionally, some generalisation to account for topography has been omitted above, as the examples in this thesis do not take it into account.

2.2 Non-dimensional shallow water equations

In general, the components of the governing equations have some magnitude that is locally consistent in time. By estimating these magnitudes we can gain a lot of information from the equations such as the relative significance of each term or how the equations themselves scale as individual components scale. The process for doing this is known as nondimensionalisation and involves expressing terms as a product of a magnitude and a dimensionless variable of order one as below.

The shallow water equations in one dimension are:

$$\frac{Du}{Dt} - fv + \frac{\partial \Phi}{\partial x} = 0, \tag{2.2.1}$$

$$\frac{Dv}{Dt} + fu = 0, \tag{2.2.2}$$

$$\frac{D\Phi}{Dt} + \Phi \frac{\partial u}{\partial x} = 0. \tag{2.2.3}$$

Rotation is allowed through the Coriolis parameter f, which we take to be constant. We assume flat bottom topography.

We can non-dimensionalise these equations by defining L, a typical length scale, and U, a typical velocity scale. If time scales advectively then we may write

$$x = L\hat{x}, \quad (u, v) = U(\hat{u}, \hat{v}), \quad t = \frac{L}{U}\hat{t}. \tag{2.2.4}$$

where hats indicate the variable is nondimensional.

We also express Φ as the sum of a constant and a perturbation $\Phi = \Phi_0 + \Phi'$ (here $\Phi_0 = gh_0$ where h_0 is the constant mean depth and we let $\Phi' = fUL\hat{\Phi}'$).

Thus, after rearrangement, we obtain our non-dimensionalised equations:

$$\frac{U}{fL}\frac{D\hat{u}}{D\hat{t}} - \hat{v} + \frac{\partial\hat{\Phi}'}{\partial\hat{x}} = 0, \tag{2.2.5}$$

$$\frac{U}{fL}\frac{D\hat{v}}{D\hat{t}} + \hat{u} = 0, \tag{2.2.6}$$

$$\frac{U^2}{\Phi_0}\frac{D\hat{\Phi}}{D\hat{t}} + \left(\frac{U}{fL} + \frac{U^2}{\Phi_0}\hat{\Phi}'\right)\frac{\partial\hat{u}}{\partial\hat{x}} = 0. \tag{2.2.7}$$

The factors $\frac{U}{fL}$ and $\frac{U^2}{\Phi_0}$ are the Rossby number (Ro) and the Froude number squared (Fr^2), respectively. These dimensionless numbers will be useful when we come to define our asymptotic limit regimes in the two-dimensional case.

2.3 The dispersion relation

The dispersion relation gives the relationship between the angular frequency and wavenumber of a wave within a system. The dispersion relation provides a method of describing the speed of the wave in terms of the *phase velocity* (the speed at which a wave of a given frequency propagates) and the *group velocity* (the speed at which the envelope of the wave propagates).

In Eulerian form, the shallow water equations (2.2.5-7) are

$$\frac{U}{fL}\left(\frac{\partial\hat{u}}{\partial\hat{t}} + \hat{u}\frac{\partial\hat{u}}{\partial\hat{x}}\right) - \hat{v} + \frac{\partial\hat{\Phi}'}{\partial\hat{x}} = 0, \tag{2.3.1}$$

$$\frac{U}{fL}\left(\frac{\partial\hat{v}}{\partial\hat{t}} + \hat{u}\frac{\partial\hat{v}}{\partial\hat{x}}\right) + \hat{u} = 0, \tag{2.3.2}$$

$$\frac{U^2}{\Phi_0}\left(\frac{\partial\hat{\Phi}'}{\partial\hat{t}} + \hat{u}\frac{\partial\hat{\Phi}'}{\partial\hat{x}}\right) + \frac{U}{fL}\frac{\partial\hat{u}}{\partial\hat{x}} + \frac{U^2}{\Phi_0}\hat{\Phi}'\frac{\partial\hat{u}}{\partial\hat{x}} = 0. \tag{2.3.3}$$

Nonlinear terms have been indicated in red.

To obtain the dispersion relation we linearise the shallow water equations around a state of rest such that $(\hat{u},\hat{v}) = (u',v')$. The perturbation quantities are assumed small, and hence nonlinear terms will be an order smaller. Removing these smallest terms, our equations become

$$Ro\frac{\partial\hat{u}'}{\partial\hat{t}} - \hat{v}' + \frac{\partial\hat{\Phi}'}{\partial\hat{x}} = 0, \tag{2.3.4}$$

$$Ro\frac{\partial\hat{v}'}{\partial\hat{t}} + \hat{u}' = 0, \tag{2.3.5}$$

$$Fr^2\frac{\partial\hat{\Phi}'}{\partial\hat{t}} + Ro\frac{\partial\hat{u}'}{\partial\hat{x}'} = 0. \tag{2.3.6}$$

If we seek wavelike solutions such that $(\hat{u},\hat{v},\hat{\Phi}) = (\tilde{u},\tilde{v},\tilde{\Phi})\exp^{i(\hat{k}\hat{x}-\hat{\omega}\hat{t})}$ our system

of equations yields

$$-iRo\,\hat{\omega}\tilde{u} - \tilde{v} + i\,\hat{k}\tilde{\Phi} = 0, \tag{2.3.7}$$

$$-iRo\,\hat{\omega}\tilde{v} + \tilde{u} = 0, \tag{2.3.8}$$

$$-iFr^2\,\hat{\omega}\tilde{\Phi} + iRo\hat{k}\tilde{u} = 0. \tag{2.3.9}$$

We write this system in matrix form

$$\begin{pmatrix} -iRo\,\hat{\omega} & -1 & i\,\hat{k} \\ 1 & -iRo\,\hat{\omega} & 0 \\ iRo\hat{k} & 0 & -iFr^2\,\hat{\omega} \end{pmatrix} \begin{pmatrix} \tilde{u} \\ \tilde{v} \\ \tilde{\Phi} \end{pmatrix} = 0. \tag{2.3.10}$$

This system is solved if the determinant is zero, and we thus have

$$\hat{\omega}(\hat{\omega}^2 - Ro^{-2} - Fr^{-2}\hat{k}^2) = 0. \tag{2.3.11}$$

There are two kinds of solution to (2.3.11). The first is $\hat{\omega} = 0$, which corresponds to time-independent (steady) flow in geostrophic balance. The second group of solutions satisfy $\hat{\omega}^2 = Ro^{-2} + Fr^{-2}\hat{k}^2$ and correspond to inertia-gravity waves.

2.4 The "slow equation"

Part II of this thesis will deal with a method for analysing the accuracy of numerical models. Following the arguments in Cullen [8], by returning to equations (2.2.5-7) we demonstrate a useful tool known as the slow equation with a simple example.

Expressing the equations in terms of the divergence $\Delta = \frac{\partial \hat{u}}{\partial \hat{x}}$ we can write the system as

$$\frac{\partial \hat{u}}{\partial \hat{t}} + \hat{u}\Delta - \frac{\hat{v}}{Ro} + \frac{\partial \hat{\Phi}'}{\partial \hat{x}} = 0, \tag{2.4.1}$$

$$\frac{\partial \hat{v}}{\partial \hat{t}} + \hat{u}\frac{\partial \hat{v}}{\partial \hat{x}} + \frac{\hat{u}}{Ro} = 0, \tag{2.4.2}$$

$$\frac{\partial \hat{\Phi}'}{\partial \hat{t}} + \hat{u}\frac{\partial \hat{\Phi}'}{\partial \hat{x}} + \left(\frac{1}{Fr^2} + \hat{\Phi}'\right)\Delta = 0. \tag{2.4.3}$$

Taking the $\hat{x}$-derivative followed by the time derivative of (2.4.1), and moving non-linear terms to the right-hand side, we have

$$\frac{\partial^2 \Delta}{\partial \hat{t}^2} - \frac{1}{Ro}\frac{\partial^2 \hat{v}}{\partial \hat{x}\partial \hat{t}} + \frac{\partial^3 \hat{\Phi}'}{\partial \hat{x}^2 \partial \hat{t}} = A_1. \tag{2.4.4}$$

A_1 here represents all the remaining terms. The same applies to A_2, A_3 etc. below.

The $\hat{x}$-derivative of (2.4.2), again moving non-linear terms to the RHS, is

$$\frac{\partial^2 \hat{v}}{\partial \hat{t} \partial \hat{x}} + \frac{\Delta}{Ro} = A_2, \tag{2.4.5}$$

and the second $\hat{x}$-derivative of (2.4.3) is

$$\frac{\partial^3 \hat{\Phi}'}{\partial \hat{x}^2 \partial \hat{t}} + \frac{1}{Fr^2} \frac{\partial^2 \Delta}{\partial \hat{x}^2} = A_3. \tag{2.4.6}$$

Substituting (2.3.5-6) into (2.4.4) yields

$$\frac{\partial^2 \Delta}{\partial \hat{t}^2} + \left(\frac{1}{Ro^2} - \frac{1}{Fr^2} \frac{\partial^2}{\partial \hat{x}^2} \right) \Delta = A_4, \tag{2.4.7}$$

which follows the same general form

$$\frac{\partial^2 \Delta}{\partial t^2} + \mathbf{L}\Delta = A \tag{2.4.8}$$

that we see in equation (2.25) of Cullen. If we seek wavelike solutions as before for Δ then (2.4.7) becomes

$$-\hat{\omega}^2 + Ro^{-2} + Fr^{-2}\hat{k}^2 = A_5. \tag{2.4.9}$$

If we set the nonlinear and forcing terms contained on the RHS to zero, then (2.4.9) becomes the analogue of the dispersion relation we found in the previous section (ignoring the $\hat{\omega} = 0$ solution).

If the $\mathbf{L}\Delta$ and A terms in (2.4.8) are of similar magnitude, and the frequency of the waves is large compared to those in A then (2.4.9) can be approximated by the "slow equation":

$$(Ro^{-2} + Fr^{-2}\hat{k}^2)\Delta = A. \tag{2.4.10}$$

This occurs where either Ro or Fr are small and so ω is much larger than U/L. Later in this thesis we will see some of the problems that can occur if **L** is not accurately represented in a numerical model.

From (2.4.10), we can deduce system behaviour for particular asymptotic regimes. If the Rossby number is small ($Ro = \varepsilon$, where ε is a small parameter), then this will correspond to strongly rotating flow in geostrophic balance. Depending on the scale of the Froude number in relation to the Rossby number, this could also describe the quasigeostrophic limit ($Ro = Fr = \varepsilon$) or the semigeostrophic limit ($Ro = Fr^2 = \varepsilon$). If the Froude number is small ($Fr = \varepsilon$), then the system is non-divergent and resembles the barotropic vorticity equation.

2.5 Discrete dispersion relation

It is useful to demonstrate the effect that the choice of grid and discretisation will have on the numerical solutions our model calculates, and more specifically how

well they capture the dynamics of certain asymptotic regimes briefly described above. In general, wavelike solutions can only be sought in the linear system, so we look at the dimensional, linearised system of equations given by

$$\frac{\partial u'}{\partial t} - fv' + \frac{\partial \Phi'}{\partial x} = 0, \tag{2.5.1}$$

$$\frac{\partial v'}{\partial t} + fu' = 0, \tag{2.5.2}$$

$$\frac{\partial \Phi'}{\partial t} + \Phi_0 \frac{\partial u'}{\partial x} = 0. \tag{2.5.3}$$

The dispersion relation for gravity waves in this system is given by

$$\omega^2 = f^2 + \Phi_0 k^2, \tag{2.5.4}$$

or the trivial solution

$$\omega = 0, \tag{2.5.5}$$

as before.

For illustration, we have chosen to spatially discretise our system using the Arakawa A- and C-grids as described in Arakawa and Lamb [2]. The A-grid discretisation may be thought of as the most 'natural' distribution of variables, however there may be advantages to staggering the placements.

In his discussion of asymptotic regimes, Cullen [8] notes that the behaviour of the numerical solution will vary depending both on the regime and the chosen grid. For this reason it is difficult to call any one grid superior.

The C-grid is one of the most common (if not *the* most common) grids in oceanic and atmospheric modelling, such as in the ARW core of the WRF model used by the National Centre for Atmospheric Research (NCAR) [42], the Met Office's model [12], ICON (versions of which are used by the Leibniz-Institute of Atmospheric Physics (IAP) [19], and the Max-Planck Institute of Meteorology (MPI-M) and the German Weather Service (DWD) [20]) and more.

Time discretisation is handled here via a Crank-Nicolson scheme similar to the UM. This scheme is absolutely stable and is second-order accurate in time.

2.5.1 A-grid

On the A-grid all variables are evaluated at the same grid points, as shown in figure 2.1. Our system becomes

$$\delta_t u_j - f\overline{v_j}^t + \delta_x \overline{\Phi_j}^t = 0, \tag{2.5.6}$$

$$\delta_t v_j + f\overline{u_j}^t = 0, \tag{2.5.7}$$

$$\delta_t \Phi_j + \Phi_0 \delta_x \overline{u_j}^t = 0. \tag{2.5.8}$$

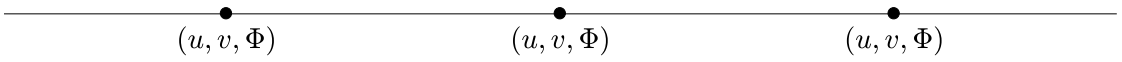

Figure 2.1: 1D visualisation of the Arakawa A-grid.

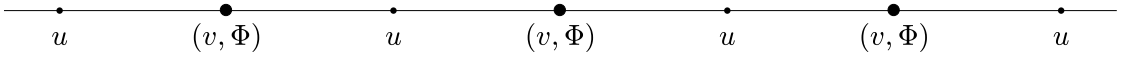

Figure 2.2: 1D visualisation of the Arakawa C-grid.

Here $\delta_t u_j = (u_j^{n+1} - u_j^n)/\Delta t$, $\delta_x u_j = (u_{j+1} - u_{j-1})/2\Delta x$ and $\overline{u}^t = (u^{n+1} + u^n)/2$.

Seeking wavelike solutions such that $(u_j^n, v_j^n, \Phi_j^n) = (\tilde{u}, \tilde{v}, \tilde{\Phi}) \exp^{i(kx_j - \omega t_n)} = (\tilde{u}, \tilde{v}, \tilde{\Phi}) \exp^{i(kj\Delta x - \omega n \Delta t)}$, and shifting the time index by half a step to simplify our calculations, we can write our system of equations as

$$-i\tilde{\omega}u - fv + i\tilde{k}_A \Phi = 0, \tag{2.5.9}$$

$$-i\tilde{\omega}v + fu = 0, \tag{2.5.10}$$

$$-i\tilde{\omega}\Phi + \Phi_0 i\tilde{k}_A u = 0, \tag{2.5.11}$$

where $\tilde{\omega} = 2\tan(\omega\Delta t/2)/\Delta t$ and $\tilde{k}_A = \sin(k\Delta x)/\Delta x$. Note that sub- and superscripts have been dropped as all variables are evaluated at the same points in time and space. Equations (2.5.9-11) yield the relation

$$\tilde{\omega}^2 = f^2 + \Phi_0 \tilde{k}_A^2 \qquad \text{or} \qquad \tilde{\omega} = 0. \tag{2.5.12}$$

This dispersion relation is the discretised analogue of (2.5.4).

2.5.2 C-grid

Inspection shows that in using centred differences in space, we effectively create two separate sub-grids whose solutions can become decoupled. Staggering the evaluation points of the variables will avoid this. It also improves dispersion in the scheme by effectively halving the 'space step'. The one-dimensional Arakawa C-grid is a staggered grid in which v and Φ are kept at the original grid points and u is evaluated between these points, as shown in figure 2.2. On this grid, the system of equations is

$$\delta_t u_{j+1/2} - f\overline{v_{j+1/2}}^{x,t} + \delta_x \overline{\Phi_{j+1/2}}^t = 0, \tag{2.5.13}$$

$$\delta_t v_j + f\overline{u_j}^{x,t} = 0, \tag{2.5.14}$$

$$\delta_t \Phi_j + \Phi_0 \delta_x \overline{u_j}^t = 0, \tag{2.5.15}$$

where here $\delta_x u_j = (u_{j+1/2} - u_{j-1/2})/\Delta x$ and $\overline{u_j}^x = (u_{j+1/2} + u_{j-1/2})/2$.

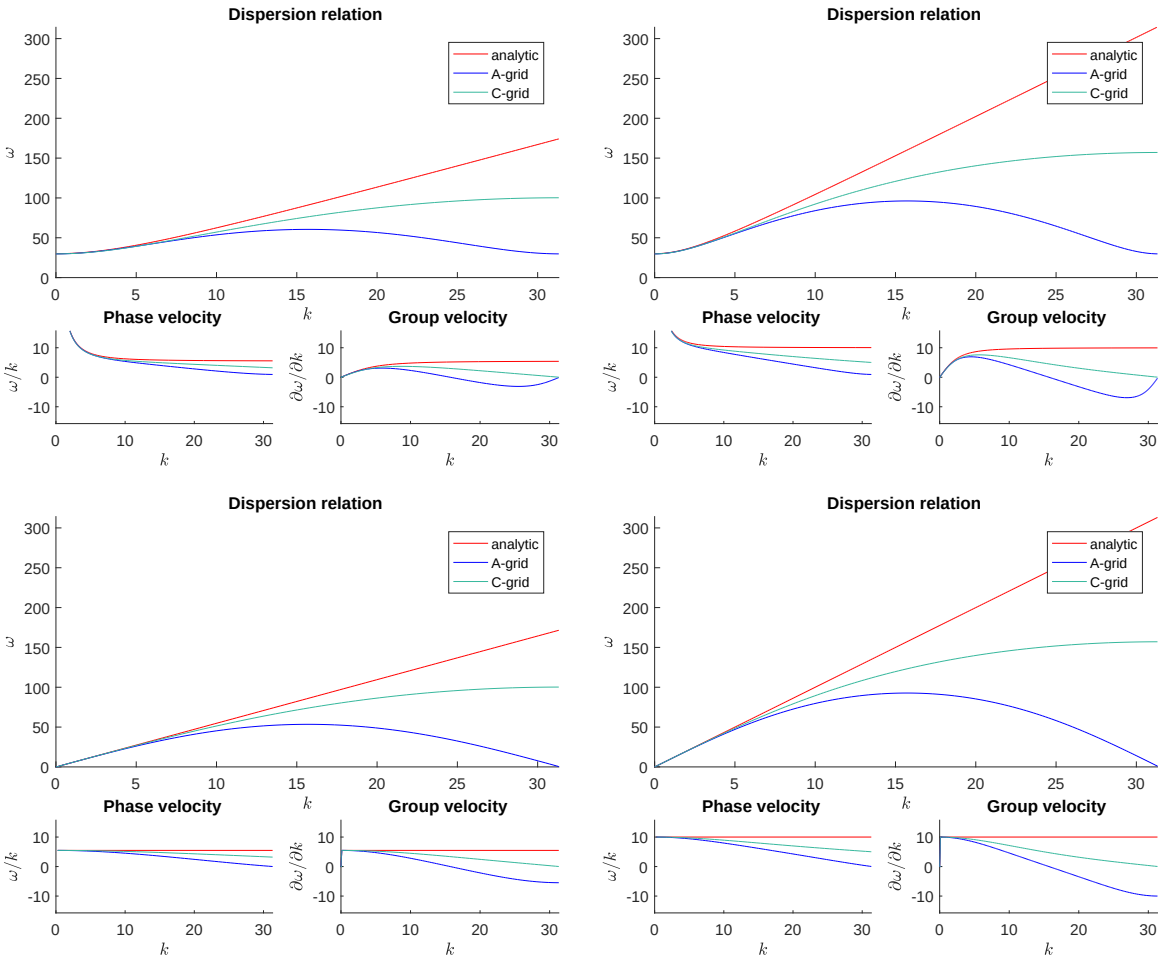

Figure 2.3: Dispersion relation, along with the phase and group velocity. In each figure $\Delta x = 0.1, \Delta t = 0.001$, and so $\omega \approx f$ at $k = 0$. Top left: $f = 30, \Phi_0 = 30$. Top right: $f = 30, \Phi_0 = 100$. Bottom left: $f = 0.0001, \Phi_0 = 30$. Bottom right: $f = 0.0001, \Phi_0 = 100$. Both schemes give increasingly poor estimates to the analytic solution as the nyquist limit is approached. The A-grid solutions demonstrate two wavenumber responses to the same frequency forcing, and group velocity becomes negative.

Seeking solutions as before, we achieve

$$-i\tilde{\omega}u - \tilde{f}v + i\tilde{k}_C\Phi = 0, \tag{2.5.16}$$

$$-i\tilde{\omega}v + \tilde{f}u = 0, \tag{2.5.17}$$

$$-i\tilde{\omega}\Phi + \Phi_0 i\tilde{k}_C u = 0, \tag{2.5.18}$$

with $\tilde{k}_C = 2\sin(k\Delta x/2)/\Delta x$ and $\tilde{f} = f\cos(k\Delta x/2)$. This yields the relation

$$\tilde{\omega}^2 = \tilde{f}^2 + \Phi_0\tilde{k}_C^2 \qquad \text{or} \qquad \tilde{\omega} = 0. \tag{2.5.19}$$

2.5.3 Further investigation

Figure 2.3 demonstrates the differences in the dispersion relations for the two grids against the analytic solution. In all schemes it can be seen that at $k = 0, \tilde{\omega} = f$, and for small Δt this is a good approximation to $\omega = f$. For low wavenumber/angular frequency both schemes are very close to the true solution. As the wavenumber approaches the Nyquist limit $k = \pi/\Delta x$ both numerical solutions increasingly underestimate the true solution and the frequency in the A-grid scheme actually falls back to $\omega = f$ after $k = \pi/2\Delta x$. This results in there being two wavenumber responses to a single frequency forcing, and the higher wavenumber response will have negative group velocity.

To further investigate the behaviour of these discrete dispersion relationships, it is helpful to write the full equations for each grid:

$$\text{A:} \quad \frac{4\tan^2(\omega\Delta t/2)}{\Delta t^2} = f^2 + \Phi_0 \frac{\sin^2(k\Delta x)}{\Delta x^2} \tag{2.5.20}$$

$$\text{C:} \quad \frac{4\tan^2(\omega\Delta t/2)}{\Delta t^2} = f^2\cos^2(k\Delta x/2) + \Phi_0 \frac{4\sin^2(k\Delta x/2)}{\Delta x^2} \tag{2.5.21}$$

If k is small, then both grids reduce to approximately

$$2\tan(\omega\Delta t/2) = \pm f\Delta t. \tag{2.5.22}$$

The solution closest to zero is

$$\omega = \frac{2}{\Delta t}\tan^{-1}\left(\frac{f\Delta t}{2}\right). \tag{2.5.23}$$

If x is small then $\tan^{-1}(x) \approx x$ and so we may claim that if k and $f\Delta t$ are small then $\omega \approx f$.

Similarly, if we look at k as we approach the Nyquist limit we arrive at the following:

$$\text{A:} \quad \frac{4\tan^2(\omega\Delta t/2)}{\Delta t^2} = f^2 \tag{2.5.24}$$

$$\text{C:} \quad \frac{4\tan^2(\omega\Delta t/2)}{\Delta t^2} = \frac{4\Phi_0}{\Delta x^2} \tag{2.5.25}$$

The A-grid equation will have the same solution as above, indicating that we expect $\omega \approx f$ at the Nyquist limit. The C-grid solution is

$$\omega = \frac{2}{\Delta t}\tan^{-1}\left(\frac{\sqrt{\Phi_0}\Delta t}{\Delta x}\right). \tag{2.5.26}$$

Note the appearance of $\frac{\sqrt{\Phi_0}\Delta t}{\Delta x}$. This quantity is known as the *wave Courant number* - a measure of how well-resolved gravity waves will be in a model. In many schemes it is a requirement to keep this value low (explicit time integration schemes typically require this to be less than one) for the scheme to remain stable.

This and the *advective Courant number* will appear a number of times in this thesis.

Based on the above, it is reasonable to assume that this number is small, and so

$$\omega \approx \frac{2\sqrt{\Phi_0}}{\Delta x} \tag{2.5.27}$$

at the Nyquist limit. The analytic solution here is

$$\omega = \sqrt{f^2 + \frac{\pi^2 \Phi_0}{\Delta x^2}}. \tag{2.5.28}$$

This is not very close to the C-grid solution, and indeed there is no clever manipulation of variables that would allow them to be close. The best we can do is look to minimise (2.5.28), which occurs when $f\Delta x/\sqrt{\Phi_0}$ is minimised. This value $f\Delta x/\sqrt{\Phi_0}$ represents how well a length scale known as the Rossby Radius of Deformation ($\sqrt{\Phi_0}/f$) is resolved. This number too will occur elsewhere in this thesis.

Clearly as there is no way to satisfactorily approximate the dispersion relation near the Nyquist limit (at least with the simple examples presented here) in order to properly represent high-wavenumber motion it is evident that the limit should therefore be as high as possible. i.e. Δx should be kept low. However, this is not a very useful restriction at all as in practise we will always run with the smallest affordable Δx.

Minimising Δx then may enable us to maintain accuracy at larger scales, but this accuracy will always be lost near the resolution limit - smaller Δx merely moves that limit further away, as it were. Effort must then be spent to ensure that key properties of a system, such as balance and timescale separations, are maintained rather than focussing simply on the accuracy of results.

2.6 Semi-implicit time stepping

Given a semi-implicit discretisation of the slow equation

$$\delta_{tt}\Delta + \overline{\mathbf{L}\Delta}^{tt} = \overline{A}^{tt}, \tag{2.6.1}$$

if we assume that $A \sim e^{i\nu t}$ and replace **L** by its eigenvalue ω^2, then there is a solution proportional to $e^{i\nu t}$

$$-\tan^2\left(\frac{\nu\Delta t}{2}\right)\Delta + \Delta t^2\omega^2\Delta = A\Delta t^2. \tag{2.6.2}$$

If A is slowly varying compared to both the evaluation of **L** and the time step, then

$$|\omega\Delta t| \gg \left|\tan\left(\frac{\nu\Delta t}{2}\right)\right|, \tag{2.6.3}$$

and so the solution reduces to

$$\overline{\mathbf{L}\Delta}^{tt} = \overline{A}^{tt} \implies \mathbf{L}\Delta = A, \tag{2.6.4}$$

and the response to forcing is correct.

However, for free waves such that

$$\delta_{tt}\Delta + \overline{\mathbf{L}\Delta}^{tt} = 0, \tag{2.6.5}$$

the frequency is reduced to

$$\omega\frac{\cos\omega\Delta t}{\sin\omega\Delta t} \ll \omega, \tag{2.6.6}$$

and hence such waves will be slowed, which affects the group velocity and possibly the adjustment to balance.

2.7 Numerical examples

We have seen some of the ways in which discretisation may fail to accurately reproduce the dispersion relation. We can demonstrate the effects of this in a more visual way through some simple numerical examples.

We must discretise our chosen shallow water model in order to investigate it fully. We look at the linear and nonlinear equations on both the A- and C-grid for a total of four cases. In order to solve the equations resulting from the Crank-Nicolson time stepping it is standard practise in such numerical models to substitute the momentum equations into the mass equation to form an expression for the 'next' time step purely in terms of the current one. In this simple 1D case, however, the u equation contains both v and Φ terms and the Φ equation contains no v contribution and so would first require v to be expressed in terms of u. It is much more convenient then to substitute the Φ and v equations into the u equation. When moving on to higher dimensional cases it will be necessary to solve for the Φ as usual.

2.7.1 Linear Equations

A-grid

The linear discretised system on the A-grid may be given in the form

$$u_j^{n+1} = u_j^n + \Delta t(f\overline{v_j}^t - \delta_x\overline{\Phi_j}^t), \tag{2.7.1}$$

$$v_j^{n+1} = v_j^n - f\Delta t\overline{u_j}^t, \tag{2.7.2}$$

$$\Phi_j^{n+1} = \Phi_j^n - \Phi_0\Delta t\delta_x\overline{u_j}^t. \tag{2.7.3}$$

We can substitute (2.7.2) and (2.7.3) into (2.7.1) and rearrange to obtain the expression

$$\begin{aligned}\left(1+\left(\frac{f\Delta t}{2}\right)^2\right)u_j^{n+1}-\Phi_0\left(\frac{\Delta t}{4\Delta x}\right)^2\left(u_{j+2}^{n+1}-2u_j^{n+1}+u_{j-2}^{n+1}\right)=\\ \left(1-\left(\frac{f\Delta t}{2}\right)^2\right)u_j^{n}+\Phi_0\left(\frac{\Delta t}{4\Delta x}\right)^2\left(u_{j+2}^{n}-2u_j^{n}+u_{j-2}^{n}\right)+\\ f\Delta t v_j^n-\frac{\Delta t}{2\Delta x}(\Phi_{j+1}^n-\Phi_{j-1}^n).\end{aligned} \tag{2.7.4}$$

The left hand side is a Helmholtz equation of the form $\nabla^2 A-\kappa^2 A$ and can be solved using a tridiagonal matrix algorithm.

C-grid

On the C-grid, the linear system is

$$u_{j+1/2}^{n+1}=u_{j+1/2}^n+\Delta t(f\overline{v_{j+1/2}}^{x,t}-\delta_x\overline{\Phi_{j+1/2}}^t), \tag{2.7.5}$$

$$v_j^{n+1}=v_j^n-f\Delta t\overline{u_j}^{x,t}, \tag{2.7.6}$$

$$\Phi_j^{n+1}=\Phi_j^n-\Phi_0\Delta t\delta_x\overline{u_j}^t. \tag{2.7.7}$$

Substituting (2.7.6) and (2.7.7) into (2.7.5) yields

$$\begin{aligned}u_{j+1/2}^{n+1}+\left(\frac{f\Delta t}{4}\right)^2(u_{j+3/2}^{n+1}+2u_{j+1/2}^{n+1}+u_{j-1/2}^{n+1})\\ -\Phi_0\left(\frac{\Delta t}{2\Delta x}\right)^2(u_{j+3/2}^{n+1}-2u_{j+1/2}^{n+1}+u_{j-1/2}^{n+1})\\ =u_{j+1/2}^n+\frac{f\Delta t}{2}(v_{j+1}^n+v_j^n)-\left(\frac{f\Delta t}{4}\right)^2(u_{j+3/2}^n+2u_{j+1/2}^n+u_{j-1/2}^n)\\ -\frac{\Delta t}{\Delta x}(\Phi_{j+1}^n-\Phi_j^n)+\Phi_0\left(\frac{\Delta t}{2\Delta x}\right)^2(u_{j+3/2}^n-2u_{j+1/2}^n+u_{j-1/2}^n).\end{aligned} \tag{2.7.8}$$

If we compare this to (2.7.4) we can see that the general form has remained the same, though several terms have changed: the v term has been replaced by its space-averaged form, and one of the u terms has been replaced by a 'doubly space-averaged' form that comes from the space-averaged u term being substituted from the space-averaged v term. In addition, as the central differences are taken over a step half the size, several Δx terms have been correspondingly halved. The same applies to the δ_x term, which uses points a half-space away to evaluate rather than the full space used in the A-grid.

2.7.2 Nonlinear equations

A-grid

We define a nonlinear discretised system on the A-grid, and express it as

$$u_j^{n+1} = u_j^n + \Delta t(f\overline{v_j}^t - \delta_x\overline{\Phi_j}^t - [u_j\delta_x u_j]^{n+1/2}), \tag{2.7.9}$$

$$v_j^{n+1} = v_j^n - \Delta t(f\overline{u_j}^t + [u_j\delta_x v_j]^{n+1/2}), \tag{2.7.10}$$

$$\Phi_j^{n+1} = \Phi_j^n - \Delta t(\Phi_0\delta_x\overline{u_j}^t - [\Phi_0\delta_x u_j - \delta_x\Phi_j u_j]^{n+1/2}), \tag{2.7.11}$$

with $a^{n+1/2} = a^n$ for the first time step and $\frac{1}{2}(3a^n - a^{n-1})$ for all subsequent time steps. Comparing to the linear system (2.6.1-3) you can see that the linear terms are the same and that there is now the addition of nonlinear terms extrapolated in time. Excluding the first step, (2.7.10) and (2.7.11) may be substituted into (2.7.9) as before to obtain

$$\begin{aligned}\left(1+\left(\frac{f\Delta t}{2}\right)^2\right)u_j^{n+1} - \Phi_0\left(\frac{\Delta t}{4\Delta x}\right)^2(u_{j+2}^{n+1} - 2u_j^{n+1} + u_{j-2}^{n+1}) = \\ \left(1-\left(\frac{f\Delta t}{2}\right)^2\right)u_j^n + f\Delta t v_j^n - \frac{\Delta t}{2\Delta x}(\Phi_{j+1}^n - \Phi_{j-1}^n) \\ +\Phi_0\left(\frac{\Delta t}{4\Delta x}\right)^2(u_{j+2}^n - 2u_j^n + u_{j-2}^n) - \frac{f\Delta t^2}{4\Delta x}[u_j(v_{j+1} - v_{j-1})]^{n+1/2} \\ -\frac{\Delta t^2}{8\Delta x^2}[\Phi_0(u_{j+2} - 2u_j + u_{j-2}) - (\Phi u)_{j+2} + 2(\Phi u)_j - (\Phi u)_{j-2}]^{n+1/2} \\ -\frac{\Delta t}{2\Delta x}[u_j(u_{j+1} - u_{j-1})]^{n+1/2}.\end{aligned} \tag{2.7.12}$$

Only the linear part of this equation remains in Crank-Nicolson form, and so is easily solveable.

C-grid

On the C-grid, the nonlinear system is

$$\begin{aligned}u_{j+1/2}^{n+1} = u_{j+1/2}^n + \Delta t(f\overline{v_{j+1/2}}^{x,t} - \delta_x\overline{\Phi_{j+1/2}}^t) \\ -\Delta t[u_{j+1/2}\partial_x u_j]^{n+1/2},\end{aligned} \tag{2.7.13}$$

$$v_j^{n+1} = v_j^n - f\Delta t\overline{u_j}^{x,t} - \Delta t[\overline{u_j}^t\partial_x v_{j+1/2}]^{n+1/2}, \tag{2.7.14}$$

$$\Phi_j^{n+1} = \Phi_j^n - \Phi_0\Delta t\delta_x\overline{u_j}^t + \Delta t[\Phi_0\partial_x u_j - \partial_x\Phi_j\overline{u_j}^x]^{n+1/2}. \tag{2.7.15}$$

As above, we can substitute (2.7.14) and (2.7.15) into (2.7.13) such that

$$\begin{aligned}
u^{n+1}_{j+1/2} + \left(\frac{f\Delta t}{4}\right)^2 (u^{n+1}_{j+3/2} + 2u^{n+1}_{j+1/2} + u^{n+1}_{j-1/2}) \\
-\Phi_0 \left(\frac{\Delta t}{2\Delta x}\right)^2 (u^{n+1}_{j+3/2} - 2u^{n+1}_{j+1/2} + u^{n+1}_{j-1/2}) \\
= u^n_{j+1/2} - \left(\frac{f\Delta t}{4}\right)^2 (u_{j+3/2} + 2u^n_{j+1/2} + u^n_{j-1/2}) \\
+\Phi_0 \left(\frac{\Delta t}{2\Delta x}\right)^2 (u^n_{j+3/2} - 2u^n_{j+1/2} + u^n_{j-1/2}) + \frac{f\Delta t}{2}(v^n_{j+1} + v^n_j) \\
-\frac{\Delta t}{\Delta x}(\Phi^n_{j+1} - \Phi^n_j) - \frac{f\Delta t^2}{16\Delta x}[(u_{j+3/2} + u_{j+1/2})(v_{j+2} - v_j) \\
+(u_{j+1/2} + u_{j-1/2})(v_{j+1} - v_{j-1})]^{n+1/2} \\
-\frac{\Delta t^2}{2\Delta x^2}[\Phi_0(u_{j+3/2} - 2u_{j+1/2} + u_{j-1/2}) - \frac{1}{4}(\Phi_{j+2}(u_{j+5/2} + u_{j+3/2}) \\
-\Phi_{j+1}(u_{j+3/2} + u_{j+1/2}) - \Phi_j(u_{j+1/2} + u_{j-1/2}) \\
+\Phi_{j-1}(u_{j-1/2} + u_{j-3/2})]^{n+1/2} - \frac{\Delta t}{2\Delta x}[u_{j+1/2}(u_{j+3/2} - u_{j-1/2})]^{n+1/2}.
\end{aligned} \tag{2.7.16}$$

2.8 Response to simple forcing

To demonstrate the response of our schemes to simple forcing, we begin with an initial state $(u, v, \Phi) = (0, 0, \Phi_0)$. We directly force a single Φ point in the centre of the domain at frequency ω, and observe the response.

Figure 2.4 demonstrates the schemes' solutions for low frequency forcing. The analytic plot indicates a left- and a right-moving wave with wavenumber $k \approx \pi$ starting at the central point. For all points behind the wavefront the numerical schemes should match this solution as closely as possible. It can be seen that both the A- and C-grid schemes' solutions show a wavenumber response close to the analytic solution. Figure 2.5 shows the dispersion relation(s) for this case, and at $\omega = 10\pi$ the different relations are very close. The A-grid dispersion relation has two solutions though, and so the response to forcing appears as a superposition of two different wavenumbers, π and $\frac{\pi}{\Delta x} - \pi \approx 19\pi$. The amplitude of the numerical solutions is reduced if the size of the time step used is too large.

At higher frequency forcing, figure 2.5 shows that the analytic and numerical solutions are no longer close to each other. We see the A-grid solution break down - the solution is still a superposition of two wavenumbers, but neither is particularly close to the analytic solution. While the C-grid's solution is closer, we can see that it no longer aligns with the analytic solution due to the lower wavenumber response from the discrete dispersion relation.

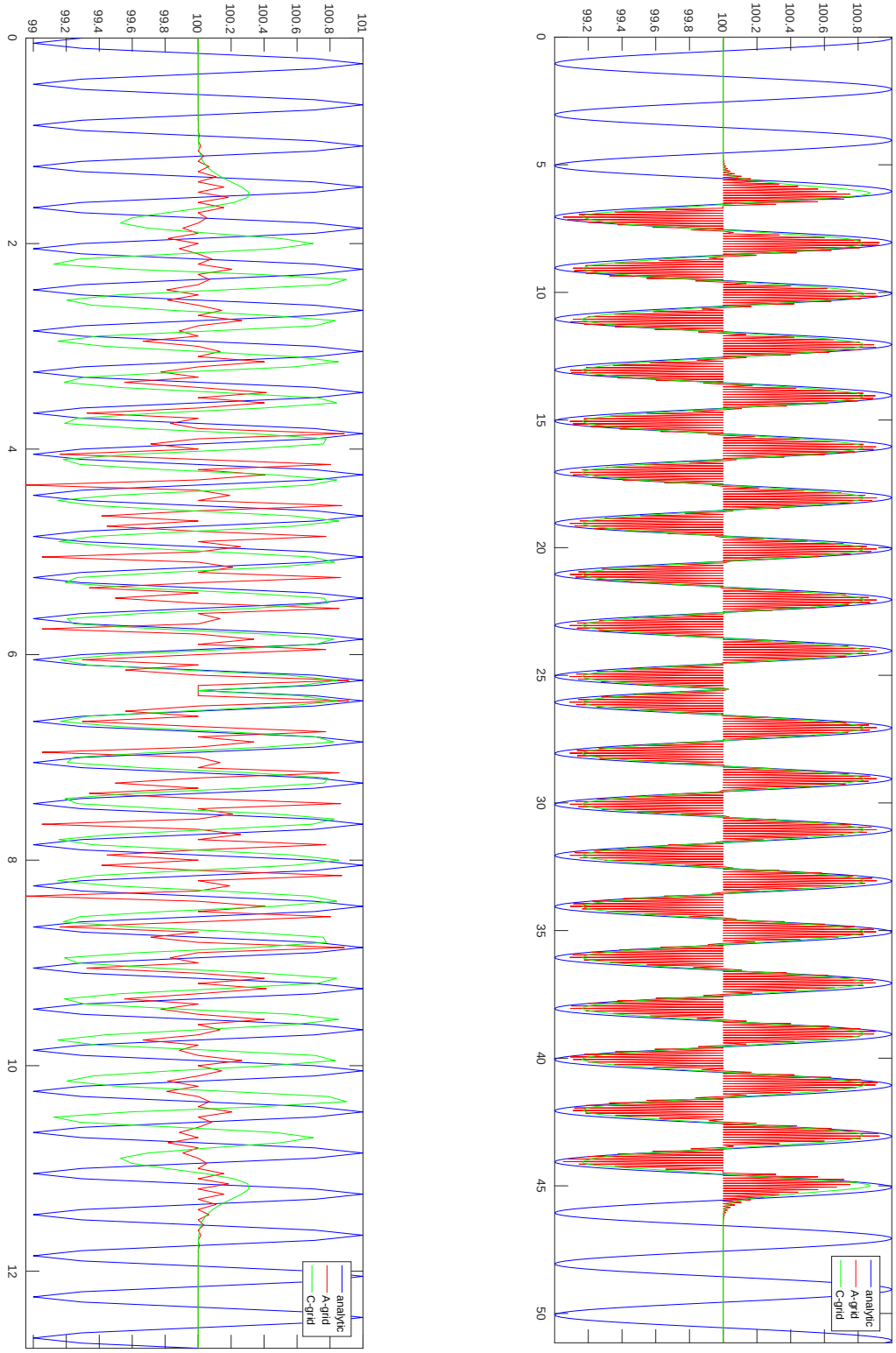

Figure 2.4: Right: response to forcing $\omega = 10\pi$ at time $t = 2$. As f is small, the analytic wavenumber is $k \approx \pi$. The C-grid and lower-wavenumber response of the A-grid solution lie close to the analtic solution, indicating that the group velocity is well approximated.
Left: response to forcing $\omega = 50\pi$ at time $t = 0.5$. The analytic wavenumber is $k \approx 5\pi$. Here the group velocity approximations are poorer.
In both examples $\Delta x = 0.05, \Delta t = 0.001, f = 0.0001, \Phi_0 = 100$.

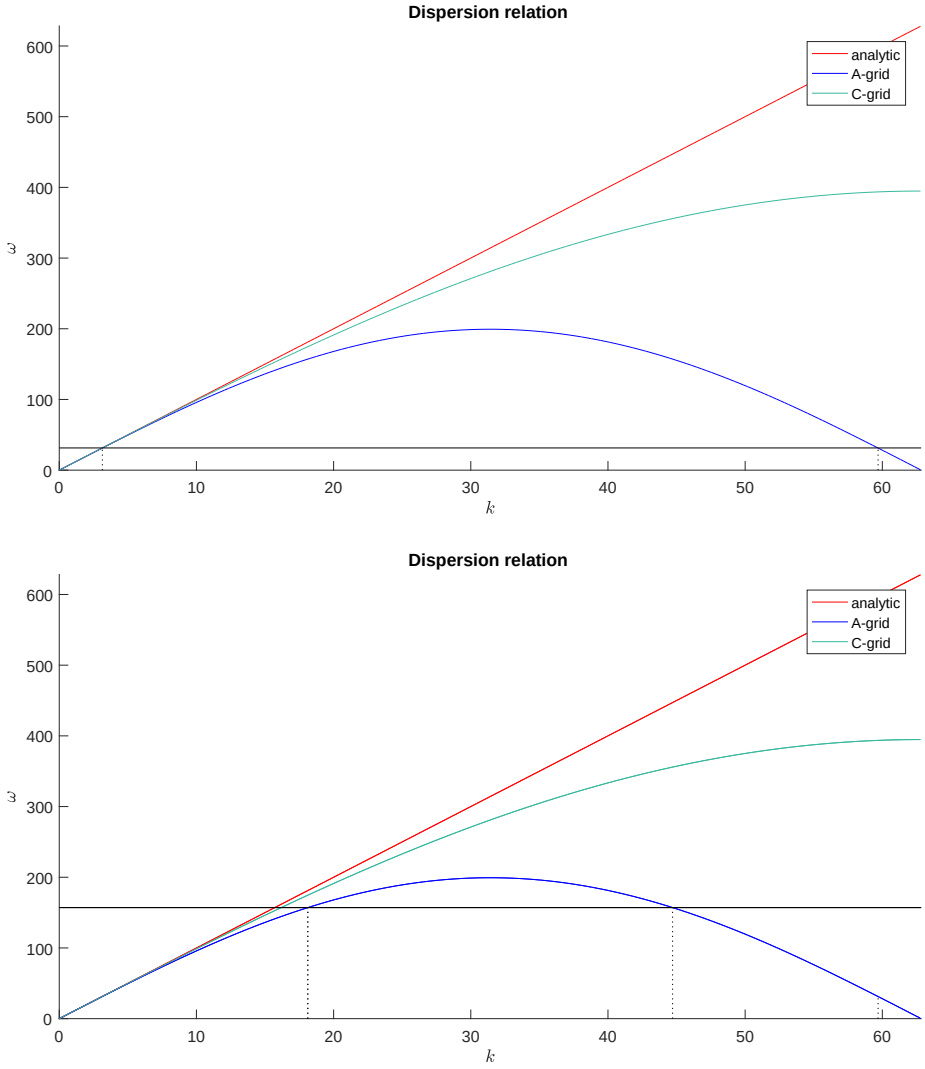

Figure 2.5: Top: dispersion relation showing response to forcing $\omega = 10\pi$.
Bottom: dispersion relation showing response to forcing $\omega = 50\pi$.
Dotted lines indicate the wavenumber responses on the A-grid. As in figure 2.4
$\Delta x = 0.05, \Delta t = 0.001, f = 0.0001, \Phi_0 = 100$

2.9 The adjustment problem

To test the response of these schemes to an initial unbalanced state, we look at the adjustment problem - the process by which a system in an initially unbalanced state evolves towards balance.

Using equations (2.5.1-3), we can consider the vorticity $\xi = \frac{\partial v}{\partial x}$ and take the x-derivative of (2.5.2) to give

$$\frac{\partial \xi}{\partial t} + f\frac{\partial u}{\partial x} = 0. \tag{2.9.1}$$

Substituting this into the continuity equation yields the linearised potential vorticity equation

$$\frac{\partial}{\partial t}\left(\frac{\xi}{f} - \frac{\Phi'}{\Phi_0}\right) = 0, \tag{2.9.2}$$

which can be integrated in time to give

$$Q = \frac{\xi}{\Phi_0} - \frac{f\Phi'}{\Phi_0^2} = \text{constant}, \tag{2.9.3}$$

indicating that it is independent of t. We can also take the time derivative of (2.5.3) and substitute in the x-derivative of (2.5.1) we get an equation in terms of Φ' and ξ

$$\frac{\partial^2 \Phi'}{\partial t^2} - \Phi_0 \frac{\partial^2 \Phi'}{\partial x^2} + \Phi_0 f\xi = 0, \tag{2.9.4}$$

and substituting in (2.9.3) we get

$$\frac{\partial^2 \Phi'}{\partial t^2} - \Phi_0 \frac{\partial^2 \Phi'}{\partial x^2} + f^2\Phi' = -f\Phi_0^2 Q. \tag{2.9.5}$$

2.9.1 The steady state

If there is a steady state solution to our system of equations it will, by definition, be time-independent. From our linearised momentum equations, we can see that a steady state of this system is thus one of geostrophic balance:

$$fu = 0, \quad fv = \frac{\partial \Phi'}{\partial x}. \tag{2.9.6}$$

If there is no rotation ($f = 0$) then the steady state is trivial - a flat surface. In the rotating case, the steady state will also have $u = \frac{\partial u}{\partial x} = 0$. This problem is degenerate, and so information from the potential vorticity is required to solve the steady state.

If we require $\frac{\partial}{\partial t} \equiv 0$ in essense we are seeking the long-term state of the system, that being the steady state after the waves caused by the initial condition have passed and the system has adjusted to geostrophic equilibrium. Under this condition (2.9.5) becomes

$$\frac{1}{f\Phi_0}\frac{\partial^2 \Phi'}{\partial x^2} - \frac{f}{\Phi_0^2}\Phi' = Q. \tag{2.9.7}$$

This suggests that using this and (2.9.6) we can infer the final steady state from the initial conditions without needing to calculate the flow evolution.

We can rearrange (2.9.7) as

$$\left(f^2 - \Phi_0 \frac{\partial^2}{\partial x^2}\right)\Phi' = -f\Phi_0^2 Q. \tag{2.9.8}$$

$$\tilde{\mathbf{L}}\Phi' = \tilde{A} \tag{2.9.9}$$

The operator $\tilde{\mathbf{L}}$ looks a lot like the **L** operator in the 'slow equation' (2.4.8). Indeed, if we use (2.5.3) to express (2.9.5) in terms of the divergence then we recover

$$\frac{\partial^2 \Delta}{\partial t^2} + \left(f^2 - \Phi_0 \frac{\partial^2}{\partial x^2}\right)\Delta = 0, \tag{2.9.10}$$

which is the dimensional form of (2.4.7) (with the right-hand side set to zero, as this example is linear). This operator **L** (or $\tilde{\mathbf{L}}$) then is important both in capturing the steady state but also in the adjustment process itself.

Thus the (numerical) adjustment problem can be seen as the model attempting to approximate the operator $\tilde{\mathbf{L}}$.

Example

We define the initial condition $\Phi(x, t=0) = \Phi_0 + \Phi_0' = \Phi_0 + e^{-\frac{(x-x_0)^2}{2\sigma^2}}\sin(kx)$, a Gaussian envelope in Φ, and $u, v = 0$. If we use periodic boundary conditions, the domain must be sufficiently large to examine the evolution of our systems.

The initial potential vorticity is given by

$$Q = -\frac{f\Phi_0'}{\Phi_0^2}, \tag{2.9.11}$$

and therefore (2.9.7) can be written

$$\frac{\partial^2 \Phi'}{\partial x^2} - \frac{f^2}{\Phi_0}\Phi' = -\frac{f^2}{\Phi_0}\left(e^{-\frac{(x-x_0)^2}{2\sigma^2}}\sin(kx)\right). \tag{2.9.12}$$

Solving this analytically is not convenient without simplification. If we require k and σ to be sufficiently large, then when seeking solutions of the form $Ce^{-\frac{(x-x_0)^2}{2\sigma^2}}\sin(kx)$,

we find that terms with coefficients of $\frac{1}{\sigma^2}$ are significantly smaller than terms with coefficients of k^2, and so can be neglected such that

$$\frac{\partial^2}{\partial x^2}\left(Ce^{-\frac{(x-x_0)^2}{2\sigma^2}}\sin(kx)\right) \approx -Ck^2e^{-\frac{(x-x_0)^2}{2\sigma^2}}\sin(kx). \tag{2.9.13}$$

Substituting this into (2.9.12) gives a value of $C = \frac{f^2}{k^2\Phi_0+f^2}$. The same result can be achieved by initially assuming a sinusoidal form and applying the envelope shape after solving.

We choose the wavenumber $k = \frac{\pi}{2\Delta x}$ such that we can observe the A-grid solution exhibiting group velocity zero. Additionally it would be beneficial if C, and by extension f, is large enough for the steady state to be clearly seen in our figures!

Solutions

As with the discrete dispersion relations calculated in earlier in this chapter, some terms in the numerical balanced state solutions will have to be estimates. Investigation reveals that we may use the same discretisation as before, such that the estimate of C using the A-grid scheme is

$$C_A = \frac{f^2}{\tilde{k}_A^2\Phi_0 + f^2}, \tag{2.9.14}$$

and the estimate of C on the C-grid is

$$C_C = \frac{\tilde{f}^2}{\tilde{k}_C^2\Phi_0 + \tilde{f}^2}. \tag{2.9.15}$$

As a reminder, $\tilde{k}_A = \sin(k\Delta x)/\Delta x$, $\tilde{k}_C = 2\sin(k\Delta x/2)/\Delta x$ and $\tilde{f} = f\cos(k\Delta x/2)$. Figure 2.6 compares the analytic C with the two discrete approximations for varying Δx. We can see that the A-grid approximation will over-estimate and the C-grid approximation will under-estimate the true C as $k\Delta x$ increases. Again we see that the approximation of terms is better at small $k\Delta x$, but as we previously discussed, this is not much of a restriction.

Figure 2.7 shows some examples of the adjustment problem. The top figure shows the analytic steady state alongside the C-grid solution. This steady state is 'left behind' after wave fronts propagate to the left and right (not shown). It is clearly seen that the C-grid solution is of lower amplitude than the analytic solution. In the A-grid solution group velocity is zero, and so the wave fronts do not move. The system oscillates in place and no adjustment process happens, thus no steady state is obtained. This highlights an inherent limitation of the A-grid, in that there will always be some $k\Delta x$ for which the group velocity is zero.

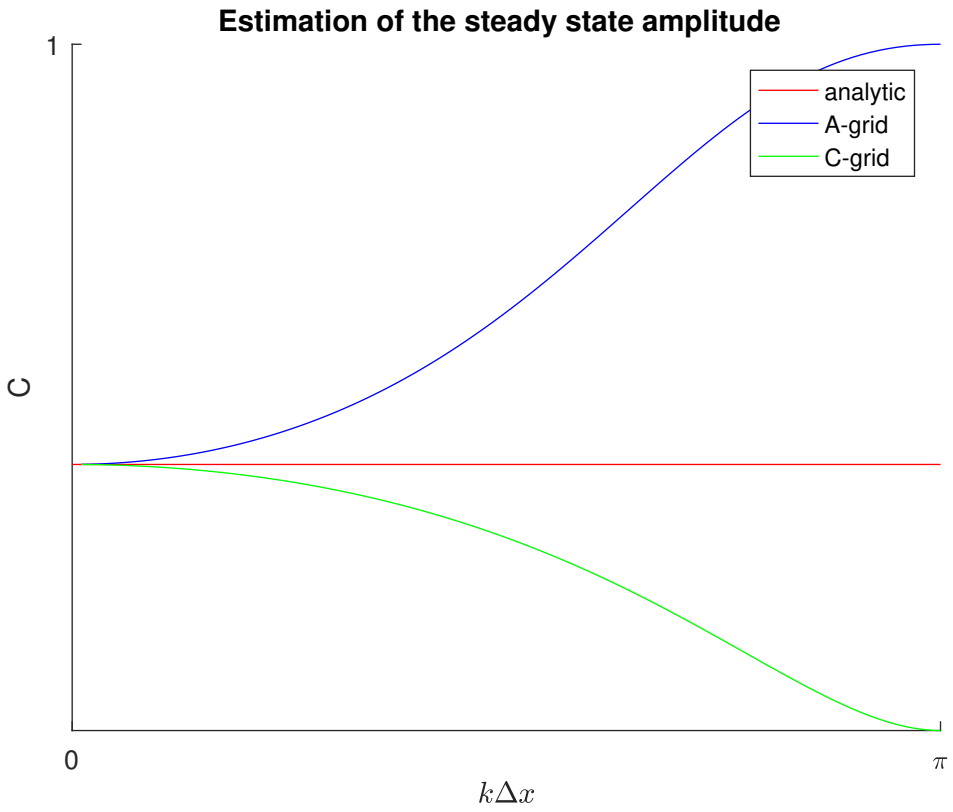

Figure 2.6: Approximations of the steady state amplitude. The A-grid solution overestimates this amplitude, where the C-grid solution underestimates it.

2.10 Summary

We have introduced the shallow water equations and some of their key features. We showed how the dispersion relation for shallow water waves may be found, and the connection between the dispersion relation and the "slow equation". Two different grids are discussed and we demonstrate the discretisation process for each grid as well as showing how successful these grids are at approximating the dispersion relation based on different parameters. This highlights some of the problems inherent in these grids. The semi-implicit time integration is briefly discussed in preparation for part II. A somewhat complex example is given in the form of the adjustment problem, and focus on the steady state approximation from each grid.

It is seen that the operator **L** is important both for the approximation of the steady state but also for the representation of waves. Thus, discretisation errors in **L** will have a number of consequences: they lead to an incorrect wave response to forcing; they affect the adjustment process from an unbalanced state; and they affect the balanced state itself via $\tilde{\mathbf{L}}\Phi'$, which is closely related to $\mathbf{L}$.

These examples are intended to give an idea of how models can fail to represent both unbalanced and balanced motion. From our analysis, several of these

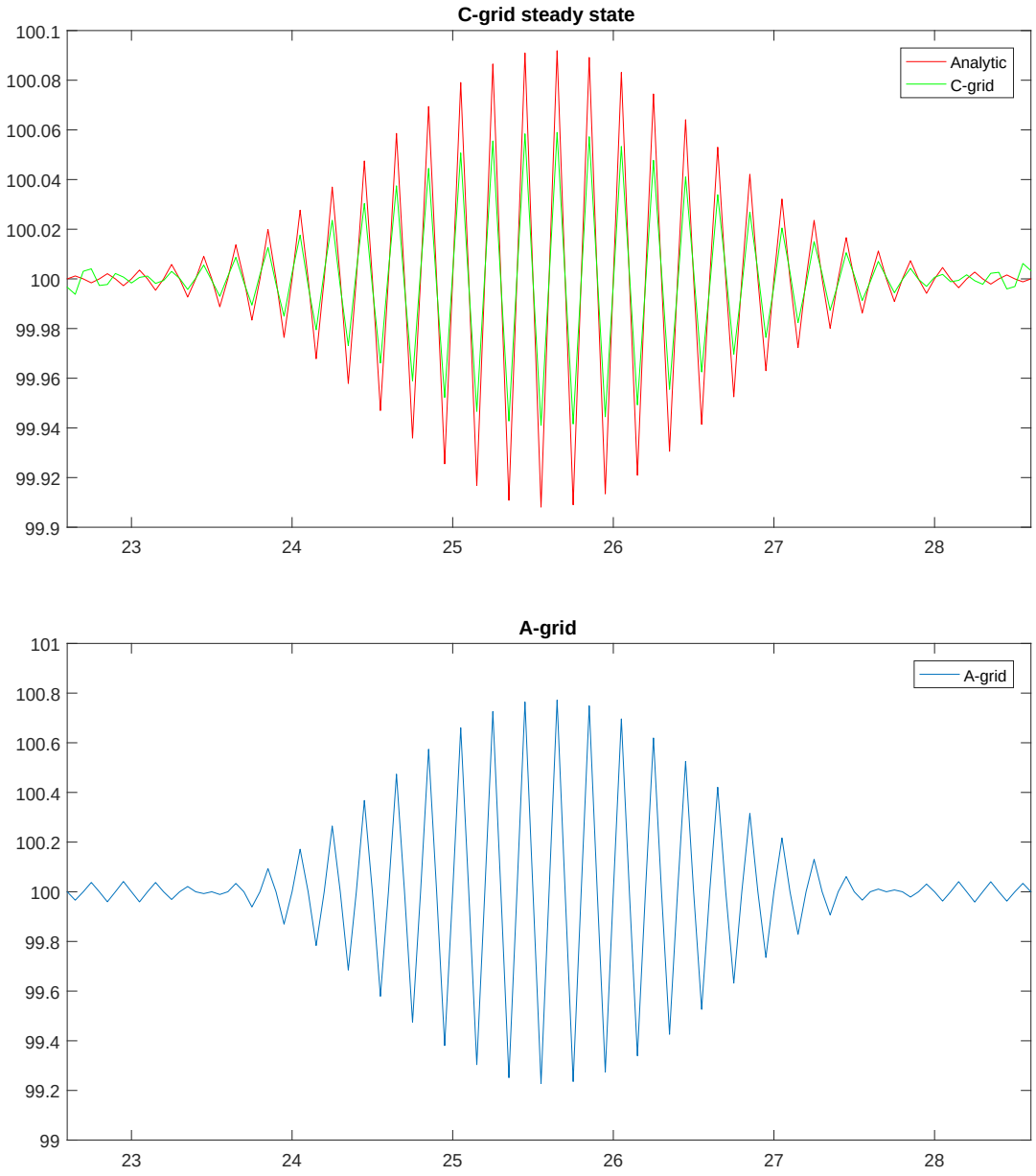

Figure 2.7: Numerical solutions to the adjustement problem at $t = 200$ on the C-grid (top) and A-grid (bottom). Note that the C-grid has achieved a steady state, while the A-grid is instead oscillating in place. Also bear in mind the difference in scale.

In both cases $f = 100$, $\Phi_0 = 100$, $\Delta x = 0.05$, $k\Delta x = \pi/2$.

examples seem as though results may only be improved via minimisation of Δx. While this will lead to a minimisation of errors, those errors are not removed but rather shifted downscale. There will thus always be some scale at which this leads to inaccurate results. As we discussed in the introduction of chapter 1, it is believed that all motion down to the Kolmogorov length scale is significant. Currently then, we have no hope of obtaining a sufficiently small Δx such that we may directly resolve all important motion within a system. Our efforts then should be instead on maintaining certain key properties of the flow. Part II of this thesis will look at a method in which we derive equations whose values are expected to scale in a certain way as a number of important flow regimes are approached. This method may be used to assess the accuracy of models in a way that is less dependent on spatial or temporal resolution.

Part II

Asymptotic limits

Chapter 3

Experimental design

3.1 Introduction

The previous part of this thesis demonstrated some simple ways in which errors can be introduced to a solution at the basic level. Clearly, some degree of error is unavoidable in complex models, and there is no way to exactly quantify the difference between the model and the true solution. So practical efforts to improve such models may be focussed on reducing rather than eliminating these errors. The method of *Asymptotic Limit Solutions* [8] is one method of assessing the errors (and by extension the accuracy) of a model. In this chapter we will introduce four asymptotic limit regimes and define imbalance equations whose scaling we can evaluate as we approach these regimes.

We start from the two-dimensional shallow water equations in the form

$$Ro\frac{D\hat{u}}{D\hat{t}} - \hat{v} + \frac{\partial\hat{\Phi}'}{\partial\hat{x}} = 0, \tag{3.1.1}$$

$$Ro\frac{D\hat{v}}{D\hat{t}} + \hat{u} + \frac{\partial\hat{\Phi}'}{\partial\hat{y}} = 0, \tag{3.1.2}$$

$$Fr^2\frac{D\hat{\Phi}'}{D\hat{t}} + \left(Ro + Fr^2\hat{\Phi}'\right)\left(\frac{\partial\hat{u}}{\partial\hat{x}} + \frac{\partial\hat{v}}{\partial\hat{y}}\right) = 0, \tag{3.1.3}$$

here expressed in terms of the Rossby (Ro) and Froude (Fr) number. Although several simplifications of the Navier-Stokes equations have occurred, this system cannot be solved analytically. However, if the Ro or Fr parameters were to approach zero, the system would become simpler, and 'true' solutions in such a regime might be both obtainable and affordable to compute. This is the basic idea behind the use of asymptotic limits solutions - by careful choice of how terms in the governing equations scale, we may know the accuracy of the asymptotic limit solution and hence we can measure the accuracy of the approximation to this solution provided by a given model. Following on from this, as we approach these limit regimes the system will be closer to the appropriate limit solution and we

would expect our model to converge towards this solution as well. We can thus derive predictions of the scaling behaviour in the system as we approach these limit regimes. In our case, should the numerical solutions to the full shallow water system match the analytic scaling then we may feel that the model is valid in that regime.

We look at four regimes, defined in terms of the Rossby and Froude numbers. These are as follows:

3.1.1 The strong rotation regime

$Ro = \varepsilon \to 0$, Fr finite. Unsurprisingly, in this regime rotational Coriolis forces dominate and the flow is in geostrophic balance. In this regime there is strong scale separation between the slow and fast motion. A three dimensional description of this regime may be found in Wingate, Embid, Holmes-Cerfon and Taylor [60]. In this strong rotation limit the shallow water equations become

$$-\hat{v} + \frac{\partial \hat{\Phi}'}{\partial \hat{x}} = O(\varepsilon), \tag{3.1.4}$$

$$\hat{u} + \frac{\partial \hat{\Phi}'}{\partial \hat{y}} = O(\varepsilon), \tag{3.1.5}$$

$$Fr^2 \frac{D\hat{\Phi}'}{D\hat{t}} + Fr^2 \hat{\Phi}' \left(\frac{\partial \hat{u}}{\partial \hat{x}} + \frac{\partial \hat{v}}{\partial \hat{y}} \right) = O(\varepsilon). \tag{3.1.6}$$

Redimensionalising the momentum equations and separating the momentum into geostrophic and ageostrophic parts such that $\mathbf{u} = \mathbf{u}_g + \mathbf{u}_a$ suggests that in this regime (and indeed all small Ro regimes) the ageostrophic momentum $\mathbf{u}_a = O(U\varepsilon)$. As the geostrophic momentum $\mathbf{u}_g = O(U)$ this is a simple example of the kind of scale separations that will occur as the balanced limit $\varepsilon \to 0$ is approached.

3.1.2 The barotropic vorticity regime

$Fr = \varepsilon \to 0$, Ro finite. In this regime rotation is less important and inertial forces dominate, leading to vorticity in this regime approximating the barotropic vorticity equation. There is still a scale separation between the fast and slow dynamics, though the slow equation differs from that in strong rotation regimes. The reduced significance of Coriolis forces make this regime more suitable for smaller-scale atmospheric motion than the other regimes. A three dimensional description of

the Low Froude, finite Rossby regime may be found in Embid and Majda [16].

$$Ro\frac{D\hat{u}}{D\hat{t}} - \hat{v} + \frac{\partial\hat{\Phi}'}{\partial\hat{x}} = 0, \tag{3.1.7}$$

$$Ro\frac{D\hat{v}}{D\hat{t}} + \hat{u} + \frac{\partial\hat{\Phi}'}{\partial\hat{y}} = 0, \tag{3.1.8}$$

$$Ro\left(\frac{\partial\hat{u}}{\partial\hat{x}} + \frac{\partial\hat{v}}{\partial\hat{y}}\right) = O(\varepsilon^2). \tag{3.1.9}$$

If we take the x-derivative of (3.1.8) and the y-derivative of (3.1.7) then substituting them into (3.1.9) yields

$$Ro^2\left(\frac{\partial\xi}{\partial t} + u\frac{\partial\xi}{\partial x} + v\frac{\partial\xi}{\partial y}\right) = O(\varepsilon^2). \tag{3.1.10}$$

In the limit $\varepsilon \to 0$ this becomes the barotropic vorticity equation. While not directly enforced, equations such as this can be used in the derivation of balanced models by ensuring that the techniques used respect this scaling as the small Froude limit is approached.

3.1.3 The quasigeostrophic regime

$Ro = Fr = \varepsilon \to 0$. The formulation of the quasigeostrophic equations was one of the most significant advances in numerical weather prediction at the time. The motivation was to produce a set of equations that filtered out the fastest motions of a system - terms that were meterologically insignficant but made integration of the equations almost impossible. This was first achieved by Charney [6] in three dimensions. The quasigeostrophic regime allows inertial forces to have an effect, though they are an order of magnitude smaller than Coriolis forces [25].

$$-\hat{v} + \frac{\partial\hat{\Phi}'}{\partial\hat{x}} = O(\varepsilon), \tag{3.1.11}$$

$$\hat{u} + \frac{\partial\hat{\Phi}'}{\partial\hat{y}} = O(\varepsilon), \tag{3.1.12}$$

$$\frac{\partial\hat{u}}{\partial\hat{x}} + \frac{\partial\hat{v}}{\partial\hat{y}} = O(\varepsilon). \tag{3.1.13}$$

This set of equations cannot describe the evolution of the flow, but we can take the approximation to the next order of ε if we split the momentum into geostrophic and ageostrophic parts and use the fact that the ageostrophic part is of order ε:

$$\varepsilon\frac{D_g\hat{u}_g}{D\hat{t}} - \hat{v}_g - \varepsilon\hat{v}_a + \frac{\partial\hat{\Phi}'}{\partial\hat{x}} = O(\varepsilon^2), \tag{3.1.14}$$

$$\varepsilon\frac{D_g\hat{v}_g}{D\hat{t}} + \hat{u}_g + \varepsilon\hat{u}_a + \frac{\partial\hat{\Phi}'}{\partial\hat{y}} = O(\varepsilon^2), \tag{3.1.15}$$

$$\varepsilon\frac{D_g\hat{\Phi}'}{D\hat{t}} + \frac{\partial\hat{u}_g}{\partial\hat{x}} + \frac{\partial\hat{v}_g}{\partial\hat{y}} + \varepsilon\left(\frac{\partial\hat{u}_a}{\partial\hat{x}} + \frac{\partial\hat{v}_a}{\partial\hat{y}}\right) = O(\varepsilon^2). \tag{3.1.16}$$

Here $D_g/Dt = \partial/\partial t + u_g\partial/\partial x + v_g\partial/\partial y$, the derivative following the geostrophic flow.

By definition, the geostrophic part of the momentum is in exact balance with the geopotential gradiant term thus the order one terms may be eliminated from the momentum equations. Similarly we may use the definition of geostrophic balance to show that $\nabla \cdot \mathbf{u}_g = 0$ to eliminate the order one terms from the mass equation. This gives a set of equations

$$\frac{D_g\hat{u}_g}{D\hat{t}} - \hat{v}_a = O(\varepsilon), \tag{3.1.17}$$

$$\frac{D_g\hat{v}_g}{D\hat{t}} + \hat{u}_a = O(\varepsilon), \tag{3.1.18}$$

$$\frac{D_g\hat{\Phi}'}{D\hat{t}} + \frac{\partial\hat{u}_a}{\partial\hat{x}} + \frac{\partial\hat{v}_a}{\partial\hat{y}} = O(\varepsilon). \tag{3.1.19}$$

In the limit $\varepsilon \to 0$ these become the quasigeostrophic shallow water equations.

In the same limit we can take the $y-$derivative of (3.1.17) and the $x-$derivative of (3.1.18) and substitute them into (3.1.19) to give

$$\frac{D_g\hat{\Phi}'}{D\hat{t}} - \frac{D_g}{D\hat{t}}\left(\frac{\partial\hat{v}_g}{\partial\hat{x}} - \frac{\partial\hat{u}_g}{\partial\hat{y}}\right) = 0. \tag{3.1.20}$$

Re-dimensionalising this yields the quasigeostrophic potential vorticity equation

$$\frac{D_g}{Dt}\left(\xi - \frac{f}{\Phi_0}\Phi'\right) = 0. \tag{3.1.21}$$

Establishing this and similar conservation properties is crucial to the development of accurate balanced models, which should retain such properties.

3.1.4 The semigeostrophic regime

$Ro = Fr^2 = \varepsilon \to 0$. In this regime the momentum is approximated by geostrophic momentum, allowing for the inclusion of ageostrophic advection. The semigeostrophic equations are valid on larger scales than the quasigeostrophic equations and are more accurate when the Burger number is small [9]. They allow for more complicated phenomena, so they could be thought of as being partway between the quasigeostrophic and the primitive equations. They support the representation of several features, such as fronts, that the quasigeostrophic regime struggles with [21].

$$-\hat{v} + \frac{\partial\hat{\Phi}'}{\partial\hat{x}} = O(\varepsilon), \tag{3.1.22}$$

$$\hat{u} + \frac{\partial\hat{\Phi}'}{\partial\hat{y}} = O(\varepsilon), \tag{3.1.23}$$

$$\frac{D\hat{\Phi}'}{D\hat{t}} + \left(1 + \hat{\Phi}'\right)\left(\frac{\partial\hat{u}}{\partial\hat{x}} + \frac{\partial\hat{v}}{\partial\hat{y}}\right) = 0. \tag{3.1.24}$$

Similar to above, splitting momentum into geostrophic and ageostrophic parts yields

$$\varepsilon \frac{D_g \hat{u}_g}{D\hat{t}} - \hat{v}_g - \varepsilon \hat{v}_a + \frac{\partial \hat{\Phi}'}{\partial \hat{x}} = O(\varepsilon^2), \tag{3.1.25}$$

$$\varepsilon \frac{D_g \hat{v}_g}{D\hat{t}} + \hat{u} + \varepsilon \hat{u}_a + \frac{\partial \hat{\Phi}'}{\partial \hat{y}} = O(\varepsilon^2), \tag{3.1.26}$$

$$\frac{D\hat{\Phi}'}{D\hat{t}} + \left(1 + \hat{\Phi}'\right)\left(\frac{\partial \hat{u}}{\partial \hat{x}} + \frac{\partial \hat{v}}{\partial \hat{y}}\right) = 0. \tag{3.1.27}$$

Following the same process as above, we arrive at

$$\frac{D_g \hat{u}_g}{D\hat{t}} - \hat{v}_a = O(\varepsilon), \tag{3.1.28}$$

$$\frac{D_g \hat{v}_g}{D\hat{t}} + \hat{u}_a = O(\varepsilon), \tag{3.1.29}$$

$$\frac{D\hat{\Phi}'}{D\hat{t}} + \left(1 + \hat{\Phi}'\right)\left(\frac{\partial \hat{u}}{\partial \hat{x}} + \frac{\partial \hat{v}}{\partial \hat{y}}\right) = 0. \tag{3.1.30}$$

In the limit $\varepsilon \to 0$ these become the semigeostrophic shallow water equations.

To investigate each of these regimes we identify a small parameter corresponding to either Rossby or Froude number, depending on the regime. These small parameters allow us to define balance equations that show the expected scaling behaviour.

The balance conditions and conservations properties briefly highlighted here, and the discretisation choices discussed in Part I are just some of the decisions that go into defining effective balanced models (for further detail see e.g. Dritschel, Gottwald and Oliver [13]). Next, we will highlight some of the details of the models we have used in Part II.

We run a series of simulations on two shallow water models: a version of the Even Newer Dynamics for General atmospheric modelling of the environment (ENDGame) [62], and a primal-dual finite element model developed as part of the GungHo project [47]. At the time of writing these models could be seen as the present and future of atmospheric modeling: ENDGame is the current dynamical core of the Unified Model employed by the Met Office, and the GungHo project is an ongoing effort to provide its replacement, focussed on running on future massively parallel supercomputers.

From a starting point in parameter space (see figure 3.1) subsequent simulations halve the small parameter each time in order to approach the relevant asymptotic limit.

Regimes 2-4 may use an initial value of *Ro* = *Fr* = 1. Regime 1 will take the lowest value of *Fr* we use in this experiment and keep *Ro* = 1 as the initial value.

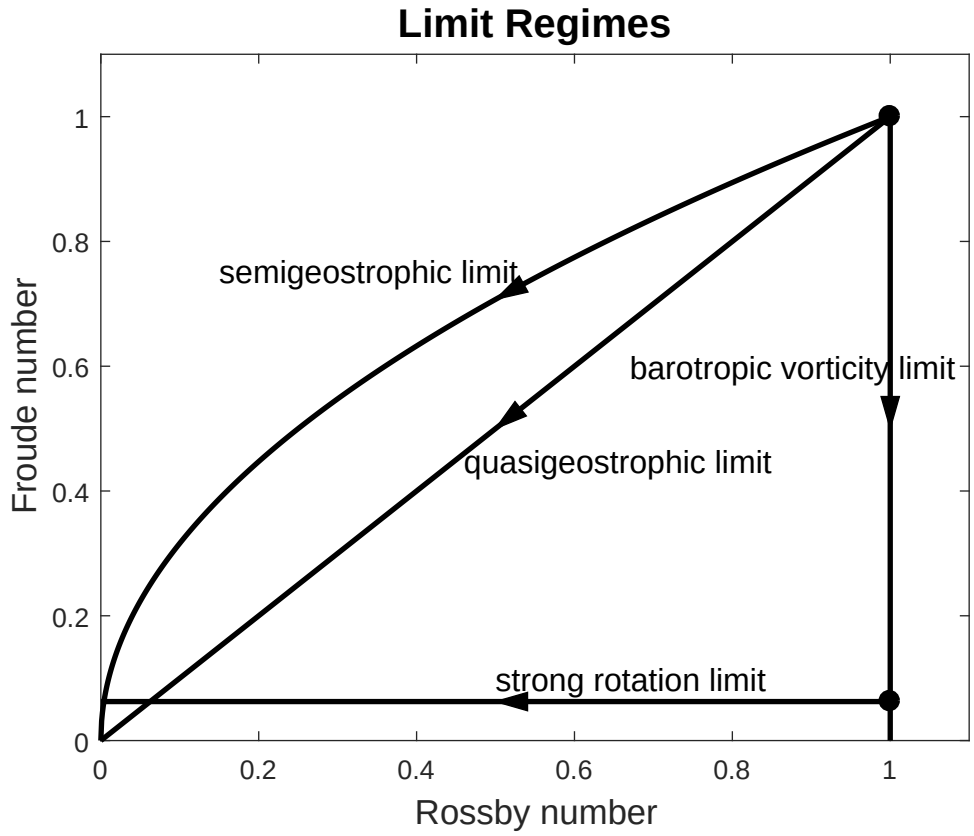

Figure 3.1: Diagram describing the approach to each limit in Rossby-Froude parameter space.

If we allow $Ro < Fr^2$ then it becomes possible for the perturbations in the fluid depth to exceed fluid depth itself! It is therefore necessary to keep *Fr* low in order to allow for a full set of results.

3.2 Experimental setup

It is desirable to keep the experiment as clean as possible, so that we can be confident that the scaling we observe is a result of our varying parameters alone. We desire a system that evolves into a flow of realistic complexity; we wish to avoid the presence of large gravity waves, as we are fundamentally looking at balance and the 'slow equation' form filters such waves; we wish to minimise the sources of errors we looked at in part 1; and we wish to ensure the flow remains as consistent as possible between regimes. We look at each of these issues in more detail.

3.2.1 The initial condition

An ideal test case is one formed of simple, analytic functions but that is able to generate complex flow that is representative of the real atmosphere. For this reason we look to the test case detailed by Galewsky, Scott and Polvani [18] (henceforth refered to as the *Galewsky initial condition*), and modify it as described below. The original initial condition consists of a balanced, barotropically unstable mid-latitude jet with a perturbation added in order to initiate instability. The initial conditions are shown in figure 3.2.

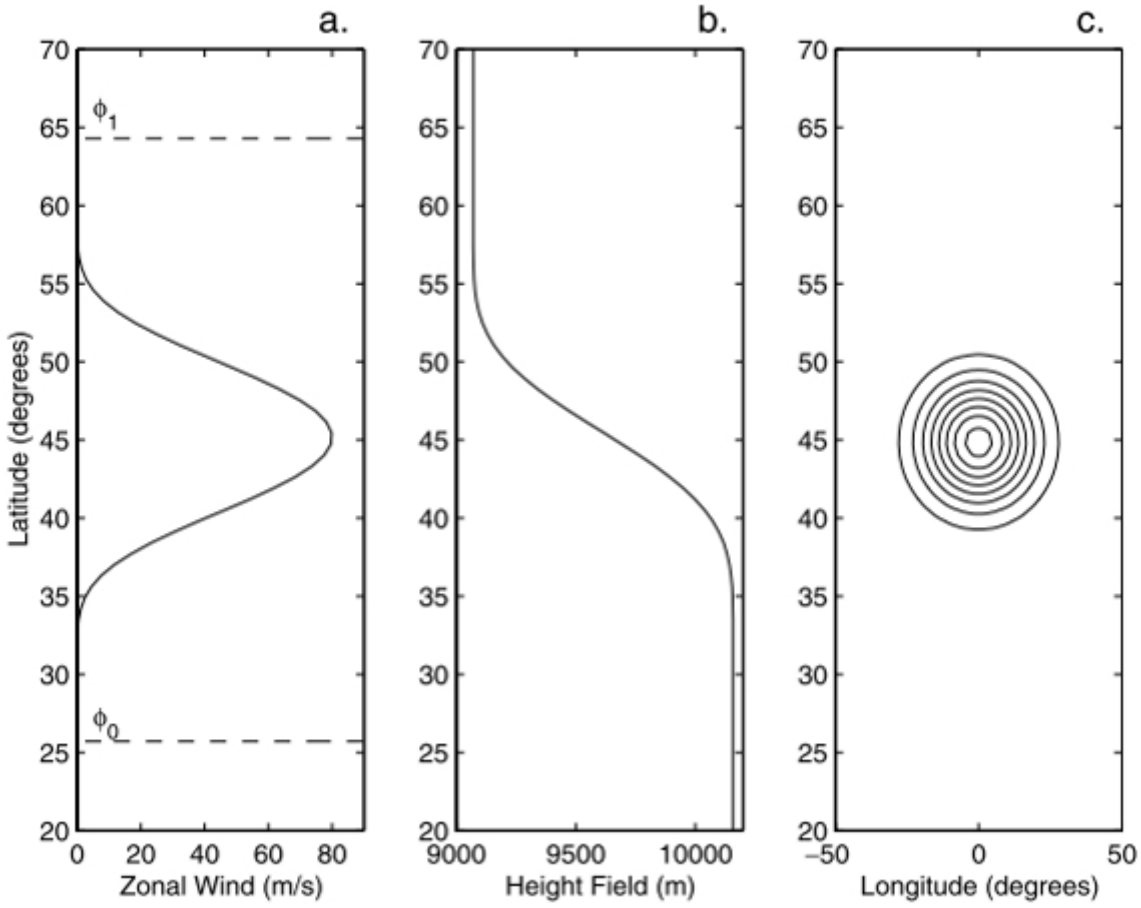

Figure 3.2: The initial conditions detailed in Galewsky et. al.
left: The wind speed near the mid-latitude jet, peaking in the centre.
middle: The fluid depth near the jet.
right: The shape and position of the geopotential perturbation.

We vary the Rossby and Froude numbers in the initial condition by modifying the Coriolis parameter and the mean fluid depth, respectively. Initial wind speed and length scales are constant across all regimes. These parameters are summarised in table 3.1.

Table 3.1: Parameter choices

Variable	Value	Units	Description
u_0	80	m/s	Initial wind speed (mid-jet)
a	6371220	m	Earth's radius
L	$a\pi/9$	m	Typical length scale (jet)
f	u_0/LRo		Coriolis parameter
Φ_0	u_0^2/Fr^2	m	Reference geopotential

Care must be taken in models representative of the real earth, as in such an instance the Coriolis parameter becomes zero at the equator, leading to the Rossby number approaching infinity. To simplify matters, the first alteration to the Galewsky initial condition is that the models used here run on an 'f-sphere' where the Coriolis parameter is constant and independent of latitude.

In slow, balanced atmospheric flow, gravity waves are emitted spontaneously during the flow evolution [52]. We are interested in such gravity waves as a possible diagnostic of imbalance and so efforts have been made to limit the gravity waves created as part of the initialisation of the flow. The Galewsky initial condition effectively introduces the perturbation suddenly at $t = 0$, causing a strong adjustment and waves that we wish to avoid. In order to minimise the impact of this initial perturbation, we introduce a second alteration and slowly force the Φ field of an initially balanced state with a modification of the Galewsky perturbation described above. To the balanced jet we force a function $S(\mathbf{x}, t)$ that is added to the geopotential equation, consisting of the original Galewsky perturbation

$$G(\mathbf{x}) = hg\cos(\phi)e^{-(3\lambda)^2}e^{[15(\phi_1-\phi)]^2}, \tag{3.2.1}$$

where h is the height of the perturbation, g is gravity, ϕ is the latitude, λ is the longitude and $\phi_1 = \frac{\pi}{4}$ is the jet's mid-point, and a smooth forcing function $F(t)$ as follows:

$$S(\mathbf{x}, t) = G(\mathbf{x})F(t). \tag{3.2.2}$$

To keep the forcing relatively smooth we choose the second-order function

$$\gamma\sin^2(\pi t/\tau), \qquad 0 \leq t \leq \tau. \tag{3.2.3}$$

τ is the length of time we wish to introduce the perturbation across, and the coefficient $\gamma = \frac{2}{\tau}$ is chosen such that $\int_0^\tau F(t)dt = 1$.

The hope is that the combination of the gradual introduction of the perturbation and an initial period of time filtering (described below) will be sufficient to damp the gravity waves from the initialisation whilst still allowing for a complex balanced flow to form. In such a flow the fast motion would hopefully be dominated by the spontaneously emitted gravity waves.

3.2.2 The time filter

During the early moments of the model runs, we use time filtering in the form of a backward (or nearly-backward) Euler method for the time integration. The backward time scheme damps the solution, and this damping is strongest on high-frequency waves. Hence this scheme effectively damps gravity waves whilst it is active. Fortunately, the barotropic instability develops over a longer time scale than the gravity wave, and so we have a period in which we can aggressively reduce the gravity wave without affecting the long-term solution too much.

Experiments suggested that instantly switching between the backward Euler and Crank-Nicolson scheme can cause further gravity waves to develop. Thus the off-centering is steadily reduced from a fully backward scheme to a fully centred one.

A generic time integration scheme can be defined as follows:

$$\frac{\mathbf{u}^{n+1} - \mathbf{u}^n}{\Delta t} = \alpha \Psi^{n+1} + \beta \Psi^n, \qquad \alpha + \beta = 1, \tag{3.2.4}$$

where Ψ^n is some function representing the right hand side of the equation of motion evaluated at time n. A similar form will exist for the mass equation.

If $\alpha = 1$ the scheme is Backward Euler, if $\alpha = \beta = 0.5$ the scheme is Crank-Nicolson. We may define α such that for the majority of the time filter's run α remains close to 1, and rapidly but smoothly reduces to 0.5 as the end of the forcing period approaches:

$$\alpha = \begin{cases} \frac{1}{2}\left[1 + \tanh^2\left(\frac{t - t_{off}}{\delta}\right)\right] & \text{if } t \leq t_{off}. \\ \frac{1}{2} & \text{if } t > t_{off}. \end{cases} \tag{3.2.5}$$

t_{off} is the time at which the time integration becomes fully centred, and δ is a parameter that controls the steepness of the adjustment between the two schemes.

In our test case, the initial time filter is playing two roles, or rather, damping two sources of gravity waves.

The first of these sources is the initial jet itself. Although the analytical description of the initial jet is perfectly balanced, truncation errors in the discretisation lead to an imbalance that causes two wave fronts heading north and south of the jet. Although small in amplitude, when these fronts hit the poles they cause an unwanted 'kick' in the divergence field. The effect of this is minimised if the spatial resolution is higher, and it might be wise to run the time filter at least until the southern wave has hit the pole.

The second source is the perturbation itself. No matter how smoothly we force the perturbation, some gravity waves will result, and so the time filtering

must also run for the full duration of the forcing. For all results presented here, the parameters $t_{off} = 144000$ and $\delta = 7200$ were used, corresponding to 40 and 2 hours respectively.

3.2.3 Effects of resolution

In the 1D case, we demonstrated the importance of choosing appropriate temporal and spatial resolutions. A measure of how fast information propagates is the Courant-Friedrichs-Lewy (CFL) condition number (hereafter referred to as the Courant Number):

$$C = \frac{c\Delta t}{\Delta x}, \tag{3.2.6}$$

where c is the velocity of a feature we wish to model, such as the advective wind or the gravity wave speed. For some schemes it is a necessary condition for stability that C remain below a certain value, which varies from scheme to scheme. Typically, an explicit method will require the advective Courant number (C_{adv}) to be 1 or less. As ENDGame is semi-implicit it will allow higher values, but it is still important that C_{adv} be low to keep features well-resolved.

As previously discussed, truly accurate modelling of atmospheric phenomena requires incredibly high resolution so in theory, this suggests that the higher resolution the better. There are some cases where limits in certain parametrisations (of the physics or initial data for example) may introduce errors with excessively high resolution, but in the simulations presented here our grids are unlikely to be so fine as to encounter such issues. A more relevant limitation is that keeping the Courant number sufficiently low requires any increase in spatial resolution to be matched by a respective increase in temporal resolution, which results in a significant increase in computing time. It is no use to have short-range forecasts taking longer to compute than the time they forecast for! In general, we can expect increased resolution to reduce model errors but it is important to balance this against the added cost.

We run with a well-resolved, high resolution run on a 640$\times$320 grid. This resolution is about as high as we can reasonably afford, and will produce sufficiently detailed results.

Where explicit schemes are solved at a grid-point by inspection of nearby grid-points, implicit schemes are solved at all grid-points simultaneously thus increasing the domain of dependence, so the CFL condition is much less strict in implicit schemes. A side-product of this is that waves are slowed down. As this slowing is undesirable, a Semi-Implicit scheme may be used that treats gravity wave terms implicitly to allow for longer timesteps, while evaluating the remaining terms explicitly to retain accuracy. This is one of the benefits of Semi-Implicit

schemes such as ENDGame. As gravity waves will still be slowed for $C_{gw} \gtrapprox 1$ the assumption here is that accurate representation of gravity waves is less important than other factors. What we don't know is whether large C_{gw} will affect the models ability to capture asymptotic limits. To investigate this, we will run a series of tests using the same initial condition, but with C_{gw} varied so as to investigate the sensitivity of the model to changes in the temporal resolution.

The time step for our "base" simulation has been chosen such that the jet remains well resolved. To this end, we require the gravity wave Courant number (C_{gw}) to be around 0.5 at mid-latitudes, and so c is set such that it is the speed of the gravity waves.

In a shallow water system, the wave speed is roughly $\sqrt{\Phi}$. We choose $60\,^\circ$ as the point we calculate the spatial resolution. This is far north enough to contain the entirety of the jet and conveniently simplifies the size of the space step to $\frac{\pi a}{n_x}$, where a is the radius of the earth and n_x is the number of longitudinal grid points, in this case 640.

Keeping the spatial resolution fixed, we adjust c and Δt in relation to each other to maintain $C_{gw} \approx 0.5$. The result of this is that with increasing Froude number, the Fluid depth and thus the gravity wave speed will increase, and the time step must be reduced accordingly.

In order to test whether semi-implicit schemes such as ENDGame will accurately approximate asymptotic limit solutions even with long time steps, the tests are repeated with a time step ten times the size (corresponding to $C_{gw} \approx 5$).

3.2.4 Other considerations

It is worth noting that the above methods for preserving balance are by no means the only options. Of note is the work of David Dritschel, particularly Mohebalhojeh and Dritchel [30] and their hierarchies of balance conditions. Though their approach has not been followed here, some aspects of our initialisations may be coincident (for example our attempts to maintain geostrophic balance below could be considered equivalent to the first "order" of $\delta - \gamma$ balance).

To keep our results as clean as possible, we should make efforts to keep the velocity scale of the response to the initial perturbation the same in each case. In an attempt to do this, we scaled the height of the perturbation so as to maintain geostrophic balance. We assume from the definition of geostrophic balance that we may define a scaling

$$fU \sim \frac{\Phi'_p}{L} \quad \text{or} \quad \Phi'_p \sim Ro^{-1}U^2, \tag{3.2.7}$$

where Φ'_p represents the geopotential of the perturbation. In the regimes we investigate, the scale of U and L does not vary significantly over time. Hence we may best maintain the above by scaling the perturbation in relation to the Coriolis parameter f. Given that our 'baseline' regime has $Ro = O(1)$ it is convenient to define the height of the perturbation in our code as $\Phi'_p = \Phi_{p0}/Ro$, where Φ_{p0} is the height of the original perturbation (120m) as described in Galewsky et al. [18].

Despite this attempt, we find that the system evolves slower as it approaches the small-Fr limit. The likely cause of this was that we did not maintain scale of the vorticity perturbation across all tests - below is how we should have scaled the perturbation with this in mind.

We start with the linearised potential vorticity:

$$\frac{\xi}{\Phi_0} - \frac{f\Phi'}{\Phi_0^2}. \tag{3.2.8}$$

Looking only at the balanced part initially, we can find the scale of Φ' in terms of ξ using the definition of geostrophic balance and taking the gradient of each side:

$$fU \sim \frac{\Phi'}{L} \implies f\xi \sim -\frac{\Phi'}{L^2}. \tag{3.2.9}$$

We can therefore express the PV in terms of the balanced vorticity and the Rossby Radius L_R:

$$\frac{\xi}{\Phi_0} - \frac{f\Phi'}{\Phi_0^2} \sim \frac{\xi}{\Phi_0} + \frac{f^2L^2\xi}{\Phi_0^2} = \frac{\xi}{\Phi_0}\left(1 + \frac{L^2}{{L_R}^2}\right), L_R = \frac{\sqrt{\Phi_0}}{f}. \tag{3.2.10}$$

If we assume that we perturb the fluid swiftly enough that corresponding PV perturbation remains close to the forced location, we can equate the forced and balanced PV:

$$-\frac{f\Phi'_f}{\Phi_0^2} = \frac{\xi_b}{\Phi_0}\left(1 + \frac{L^2}{{L_R}^2}\right). \tag{3.2.11}$$

As we wish the vorticity to remain constant across the range of our tests, we may write

$$\Phi'_f = -f\xi_b\left({L_R}^2 + L^2\right) = -f\xi_b L^2\left(\frac{Ro^2}{Fr^2} + 1\right) = -UL\xi_b\left(\frac{Ro}{Fr^2} + \frac{1}{Ro}\right). \tag{3.2.12}$$

3.2.5 The Rossby radius at asymptotic limits

The Rossby radius will approach different limits depending on the regime we are approaching. The simplest way to evaluate these limits is to look at the scaling of the Rossby radius in each of the four regimes. To this end, we can express the scaling of the Rossby radius in terms of the Rossby and Froude numbers:

$$L_R = \frac{\sqrt{\Phi_0}}{f} \sim \frac{Ro}{Fr}L. \tag{3.2.13}$$

This leads to the the following limits to the Rossby radius as each asymptotic limit is approached:

- Strong rotation (SR) limit: $Ro = \varepsilon \to 0; {L_R}^2 \sim \frac{\varepsilon^2}{\text{constant}} \to 0$

- Barotropic vorticity (BV) limit: $Fr = \varepsilon \to 0; {L_R}^2 \sim \frac{\text{constant}}{\varepsilon^2} \to \infty$

- Quasigeostrophic (QG) limit: $Ro = Fr = \varepsilon \to 0; {L_R}^2 \sim \frac{\varepsilon^2}{\varepsilon^2} \to$ constant

- Semigeostrophic (SG) limit: $Ro = Fr^2 = \varepsilon \to 0; {L_R}^2 \sim \frac{\varepsilon^2}{\varepsilon} = \varepsilon \to 0$

Due to the rapid scaling approaching the barotropic vorticity limit, maintaining the vorticity involves forcing a rather extreme perturbation with the result that it is no longer gradual enough to remain smooth. To counteract this we would need to extend the period of time across which the forcing occurs. Not only would this make the advection of PV non-negligible, but this would defeat the object of attempting to get a consistent response!

Given this fact, it is considered that the slowed evolution is a small price to pay to avoid these complications.

3.3 Diagnostics

The limits that have been mentioned represent idealised forms of atmospheric flow. At these limits, simplified balance (or near-balance) equations may be defined from which we can derive diagnostics that should be matched by the model we are examining. We can parameterise this problem by non-dimensionalising the shallow water equations in order to express them in terms of Rossby and Froude number. Then in each of the limits we described, certain terms will become small. Were these small terms considered to be zero, we would form a series of idealised balance equations. By retaining them as small terms, we can define 'imbalance' equations in which we may predict the scale of these terms. Such imbalance equations are fairly simple at $O(\varepsilon)$, where ε is some small parameter, but it is of particular interest to define imbalance equations at $O(\varepsilon^2)$. If the scaling within the model is not consistent with this imbalance equation then it may provide some important information about the model itself.

Models may employ an Eulerian or Lagrangian framework, and each may require slightly different expressions of imbalance. Both will be presented here.

3.3.1 Eulerian form

The two-dimensional nonlinear shallow water equations may be written in vector invariant form as

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} + f\mathbf{k} \times \mathbf{u} + \nabla \Phi = 0, \tag{3.3.1}$$

$$\frac{\partial \Phi}{\partial t} + \nabla \cdot (\mathbf{u}\Phi) = 0. \tag{3.3.2}$$

We non-dimensionalise as before, using (2.2.4).

This gives the Eulerian non-dimensional form of the shallow water equations:

$$Ro\frac{\partial \hat{\mathbf{u}}}{\partial \hat{t}} + (1 + Ro\hat{\xi})\mathbf{k} \times \hat{\mathbf{u}} + \nabla(\hat{\Phi}' + Ro\frac{1}{2}|\hat{\mathbf{u}}|^2) = 0, \tag{3.3.3}$$

$$Fr^2 \left(\frac{\partial \hat{\Phi}'}{\partial \hat{t}} + \nabla \cdot (\hat{\mathbf{u}}\hat{\Phi}') \right) + Ro\nabla \cdot \hat{\mathbf{u}} = 0, \tag{3.3.4}$$

where $\hat{\xi} = \frac{\partial \hat{v}}{\partial \hat{x}} - \frac{\partial \hat{u}}{\partial \hat{y}}$ and $\mathbf{k}$ is the unit vertical vector.

Low order balance

From (3.3.3) we have

$$\mathbf{k} \times \hat{\mathbf{u}} + \nabla \hat{\Phi}' = O(Ro), \tag{3.3.5}$$

a statement of (approximate) gestrophic balance. In dimensional form this is

$$f\mathbf{k} \times \mathbf{u} + \nabla \Phi' = O(fURo) = O(U^2/L). \tag{3.3.6}$$

The presence of the Coriolis parameter on the right hand side means that when simplified this term only involves U and L, neither of which vary in our simulations. We thus expect this 'geostrophic imbalance' to remain a similar scale across regimes. Taking $\nabla\times$(3.3.5) leads to, in dimensional form,

$$\nabla \cdot \mathbf{u} = O(fRo^2) = O(RoU/L). \tag{3.3.7}$$

This is valid in all small Rossby number regimes, and we therefore expect the divergence to scale like Ro. Similarly, for the barotropic vorticity limit, equation (3.3.4) gives

$$\nabla \cdot \mathbf{u} = O(fRoFr^2) = O(Fr^2U/L). \tag{3.3.8}$$

So we expect the divergence to scale like Fr^2 in the small-Fr regime.

It is simple to derive a scaling prediction for the divergence tendency, δ_t, as well by taking the time derivative of the divergence equations above to give

$$\delta_t = O(f^2Ro^3) = O(RoU^2/L^2) \tag{3.3.9}$$

in the small-Ro regimes (scales like Ro) and

$$\delta_t = O((fRoFr)^2) = O((FrU/L)^2) \tag{3.3.10}$$

in the small-Fr regime (scales like Fr^2).

Higher order balance

For a more stringent test of the model, we find a higher order balance by taking $Ro\frac{\partial}{\partial\hat{t}}$(3.3.3) and substitute in (3.3.3) and (3.3.4) to give

$$\begin{aligned} &Ro^2\frac{\partial^2\hat{\mathbf{u}}}{\partial\hat{t}^2} - (1+Ro\hat{\xi})\mathbf{k}\times((1+Ro\hat{\xi})\mathbf{k}\times\hat{\mathbf{u}} + \nabla(\hat{\Phi}' + Ro\frac{1}{2}|\hat{\mathbf{u}}|^2)) \\ &+Ro^2\frac{\partial\hat{\xi}}{\partial\hat{t}}\mathbf{k}\times\hat{\mathbf{u}} - Ro\nabla\left(\nabla\cdot(\hat{\mathbf{u}}\hat{\Phi}) + \frac{Ro}{Fr^2}\nabla\cdot\hat{\mathbf{u}}\right) + Ro^2\hat{\mathbf{u}}\cdot\frac{\partial\hat{\mathbf{u}}}{\partial\hat{t}} = 0. \end{aligned} \tag{3.3.11}$$

At leading order, the equation of balance is $\hat{\mathbf{u}} - \mathbf{k}\times\nabla\hat{\Phi}' = O(Ro)$, which is equivalent to (3.3.5).

The leading order balance equation for $Ro = \varepsilon$ is $O(\varepsilon)$ accurate. If we include the $O(Ro^2)$ terms, the equation will naturally be $O(\varepsilon^2)$ accurate. Thus the dimensional 2nd-order balance equation (henceforth referred to as the "higher-order" balance) is given by

$$\zeta^2\mathbf{u} - \zeta\mathbf{k}\times\nabla(\Phi + \frac{1}{2}|\mathbf{u}|^2) - \nabla(\nabla\cdot(\mathbf{u}\Phi)) = O(f^2URo^2) = O(U^3/L^2). \tag{3.3.12}$$

We expect the scale of this term to remain constant as we approach a small-Ro asymptotic limit.

There is only one term on the right-hand side of the equation, $\frac{\partial^2\hat{\mathbf{u}}}{\partial\hat{t}^2}$, and so we may intuitively think of this imbalance diagnostic as a measure of the *jerk* of the system.

Higher order balance in terms of Froude number

Personal communication with Mike Cullen after work on Part II was completed led to a suggestion that the higher order imbalance might be expressed in terms of Froude number. Though these diagnostics are not used here, they are included for sake of completeness.

A higher order balance equation for the small-Fr limit is obtained by rearranging (3.3.2) as

$$-\frac{Fr^2}{Ro}\left(\frac{\partial\hat{\Phi}'}{\partial\hat{t}} + \hat{\mathbf{u}}\cdot\nabla\hat{\Phi}' + \hat{\Phi}'\nabla\cdot\hat{\mathbf{u}}\right) = \nabla\cdot\hat{\mathbf{u}}, \tag{3.3.13}$$

and substituting it into itself so

$$\nabla\cdot\hat{\mathbf{u}} + \frac{Fr^2}{Ro}\left(\frac{\partial\hat{\Phi}'}{\partial\hat{t}} + \hat{\mathbf{u}}\cdot\nabla\hat{\Phi}'\right) - \hat{\Phi}'\frac{Fr^4}{Ro^2}\left(\frac{\partial\hat{\Phi}'}{\partial\hat{t}} + \hat{\mathbf{u}}\cdot\nabla\hat{\Phi}' + \hat{\Phi}'\nabla\cdot\hat{\mathbf{u}}\right) = 0. \tag{3.3.14}$$

This gives a higher-order imbalance equation

$$\nabla \cdot \hat{\mathbf{u}} + \frac{Fr^2}{Ro}\left(\frac{\partial \hat{\Phi}'}{\partial \hat{t}} + \hat{\mathbf{u}} \cdot \nabla \hat{\Phi}'\right) = O\left(\frac{Fr^4}{Ro^2}\right). \tag{3.3.15}$$

In dimensional form this is

$$\Phi_0 \nabla \cdot \mathbf{u} + \frac{\partial \Phi'}{\partial t} + \mathbf{u} \cdot \nabla \Phi' = O\left(f\Phi_0 \frac{Fr^4}{Ro}\right) = O\left(\frac{Fr^2}{Ro^2} U^3/L\right). \tag{3.3.16}$$

We can also use (3.3.2) to express this as

$$\Phi' \nabla \cdot \mathbf{u} = O\left(\frac{Fr^2}{Ro^2} U^3/L\right). \tag{3.3.17}$$

In the BV limit this should scale like Fr^2.

In the QG limit the equation becomes

$$\Phi_0 \nabla \cdot \mathbf{u} + \frac{\partial \Phi'}{\partial t} + \mathbf{u} \cdot \nabla \Phi' = O\left(f\Phi_0 RoFr^2\right) = O\left(U^3/L\right). \tag{3.3.18}$$

We expect the scale of this term to remain constant approaching the QG limit.

We might also derive an alternate form for the small-Ro imbalance:

$$(RoFr^2 - \hat{\Phi}' Fr^4)\left(\frac{\partial \hat{\Phi}'}{\partial \hat{t}} + \hat{\mathbf{u}} \cdot \nabla \hat{\Phi}'\right) - \hat{\Phi}'^2 Fr^4 \nabla \cdot \hat{\mathbf{u}} = O(Ro^2) \tag{3.3.19}$$

$$\implies (\Phi_0 - \Phi')\left(\frac{\partial \Phi'}{\partial t} + \mathbf{u} \cdot \nabla \Phi'\right) - \Phi'^2 \nabla \cdot \mathbf{u} = O(f\Phi_0 Ro) = O\left(\frac{1}{Fr^2} U^3/L\right). \tag{3.3.20}$$

The scale should remain constant in the SR limit, however due to the limitations on a small-Ro large-Fr system whereby deviations in the fluid exceed the fluid depth, we are unable to examine this equation.

3.3.2 Lagrangian form

The Lagrangian form of the shallow water equations can be written as

$$\frac{D\mathbf{u}}{Dt} + f\mathbf{k} \times \mathbf{u} + \nabla \Phi = 0, \tag{3.3.21}$$

$$\frac{D\Phi}{Dt} + \Phi \nabla \cdot \mathbf{u} = 0. \tag{3.3.22}$$

Non-dimensionalising as in section 2.2, we achieve the form

$$Ro\frac{D\hat{\mathbf{u}}}{D\hat{t}} + \mathbf{k} \times \hat{\mathbf{u}} + \nabla \hat{\Phi}' = 0, \tag{3.3.23}$$

$$Fr^2\left(\frac{D\hat{\Phi}'}{D\hat{t}} + \hat{\Phi}' \nabla \cdot \hat{\mathbf{u}}\right) + Ro \nabla \cdot \hat{\mathbf{u}} = 0. \tag{3.3.24}$$

Lower-order balance equations obtained from these equations are identical to the Eulerian form:

$$f\mathbf{k}\times\mathbf{u}+\nabla\Phi'=O(fURo), \tag{3.3.25}$$

and

$$\nabla\cdot\mathbf{u}=O(fRo^2), \tag{3.3.26}$$

in the small Rossby number regimes and

$$\nabla\cdot\mathbf{u}=O(fRoFr^2), \tag{3.3.27}$$

in the barotropic vorticity limit.

Higher order balance

As before, we can take $Ro\frac{D}{D\hat{t}}$ (3.3.23) and substitute (3.3.23) back in to give

$$Ro^2\frac{D^2\hat{\mathbf{u}}}{D\hat{t}^2}-\mathbf{k}\times(\mathbf{k}\times\hat{\mathbf{u}}+\nabla\hat{\Phi}')+Ro\frac{D}{D\hat{t}}\nabla\hat{\Phi}'=0. \tag{3.3.28}$$

We have decided to keep the equation in this form (rather than substitute in the mass equation) as it is simpler to work with in the Lagrangian model we are using.

The dimensional balance equation is therefore given by

$$f^2\mathbf{u}-f\mathbf{k}\times\nabla\Phi'+\frac{D}{Dt}\nabla\Phi'=O(f^2URo^2)=O(U^3/L^2). \tag{3.3.29}$$

Higher order balance in terms of Froude number

Similar to above, we can rearrange (3.3.24) and substitute into itself to arrive at

$$\nabla\cdot\hat{\mathbf{u}}+\frac{Fr^2}{Ro}\frac{D\hat{\Phi}'}{D\hat{t}}=O\left(\frac{Fr^4}{Ro^2}\right). \tag{3.3.30}$$

In dimensional form this is

$$\Phi_0\nabla\cdot\mathbf{u}+\frac{D\Phi'}{Dt}=O\left(f\Phi_0\frac{Fr^4}{Ro}\right)=O\left(\frac{Fr^2U^3}{Ro^2L}\right). \tag{3.3.31}$$

The small-Ro would then be:

$$(RoFr^2-\hat{\Phi}'Fr^4)\frac{D\hat{\Phi}'}{D\hat{t}}-\hat{\Phi}'^2Fr^4\nabla\cdot\hat{\mathbf{u}}=O(Ro^2) \tag{3.3.32}$$

$$\implies(\Phi_0-\Phi')\frac{D\Phi'}{Dt}-\Phi'^2\nabla\cdot\mathbf{u}=O(f\Phi_0Ro)=O\left(\frac{U^3}{Fr^2L}\right). \tag{3.3.33}$$

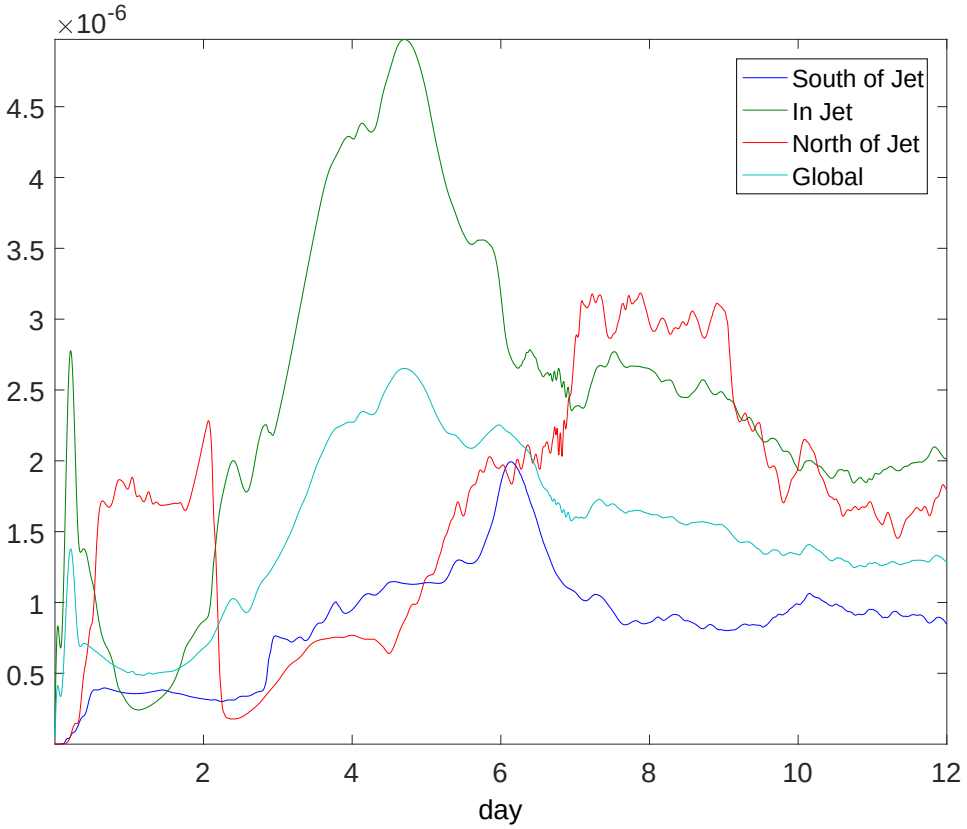

Figure 3.3: RMS higher-order imbalance for $Ro = Fr = 1$. Results have been divided into different regions relative to the initial jet.

3.3.3 Diagnostics

By taking the maximum and root mean square of the divergence, divergence tendency, geostrophic imbalance and higher-order imbalance at each time step, we can develop a picture of how these imbalances scale as we approach each regime. As previously discussed, we expect the geostrophic imbalance and higher-order imbalance to remain the same scale approaching the limit, and we expect the divergence to scale like the Rossby number or the square of the Froude number depending on the regime being approached.

Figure 3.3 shows a typical result of one of these RMS time series across a twelve day period. In addition to the global RMS, we have divided the domain into regions containing the jet and the areas south and north of the jet. The definition of the jet region is based upon the initial condition, and does not change throughout the runs. The south boundary of the jet is defined as $\phi_0 = \pi/7$ and the north boundary of the jet is $\phi_1 = 5\pi/14$. These boundaries are shown in figure 3.2. Due to the differences in behaviour of each regime we cannot expect the destabilisation and wrapping-up of vortices that occurs to be the same each time, so the 'in-jet' region is kept the same and will give us the clearest view of

the specific evolution of the flow, while the other regions can give us information about the general nature of the regime.

3.3.4 An idea for diagnosing gravity wave activity

The scale of gravity waves is difficult to diagnose directly from our system in the manner we have defined our imbalance diagnostics, for example, as there is no variable we have access to that contains only unbalanced motion.

One idea is to look at the divergence and the divergence tendency together. The divergence field is a good indicator of motion within a system, but contains both the slow balanced motion and the faster motion associated with gravity waves. As the two are tricky to unpick, we can also look at the divergence tendency. As the divergence tendency includes a timescale, there will be a separation in the two timescales present when the Rossby or Froude numbers are low. The much shorter timescale of the gravity waves will lead to the divergence tendency being more dominated by such waves. While we will not attempt quantification of gravity wave activity, we will discuss how these results compare to previous work in this area.

We assume that the divergence is the sum of balanced and unbalanced divergence such that

$$\delta = \delta_b + \delta_u, \tag{3.3.34}$$

where δ is the total divergence and δ_b, δ_u are the balanced and unbalanced parts of the divergence respectively. From a scale point of view, we could express the divergence tendency δ_t as the sum of the two divergence components over balanced (slow) and unbalanced (fast) timescales:

$$\delta_t \approx \frac{\delta_b}{\tau_b} + \frac{\delta_u}{\tau_u}. \tag{3.3.35}$$

We are able to obtain values for the divergence and divergence tendency from our analysis of the flow, and we may estimate the timescales as $\tau_b \approx \frac{L}{U}$ and $\tau_u \approx \frac{L}{\sqrt{\Phi_0}}$ or $\frac{1}{f}$, whichever is shorter.

By rearranging (3.3.34) and substituting into (3.3.35) we can obtain an expression only involving δ_u:

$$\frac{\tau_u(\tau_b\delta_t - \delta)}{(\tau_b - \tau_u)} \approx \delta_u. \tag{3.3.36}$$

While we will look briefly at how this value scales within our system, this method is likely not very robust. For example, as we are dealing with RMS values, some of the logic above may not hold. Additionally when $Fr = Ro = 1, \tau_b = \tau_u$ and so (3.3.36) goes to infinity. Nonetheless it may be interesting to see the results and possibly flesh out this idea in future.

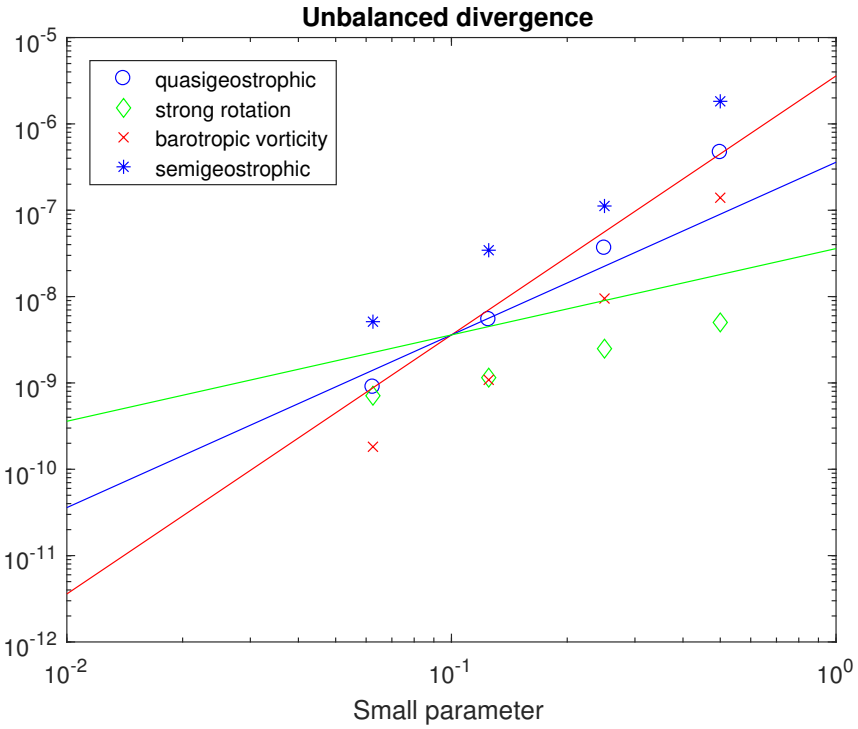

Figure 3.4: Convergence diagram for (3.3.36) in each of our regimes, taken from the ENDGame results in chapter 4. Lines indicate $\varepsilon, \varepsilon^2$ and ε^3 scalings, where ε is the small parameter. Colours indicate the line we expect our results to lie parallel to.

As τ_u depends on either Φ_0 or f, it will thus depend on either Fr or Ro respectively: In the SR, QG and SG limit $\tau_u \approx \frac{1}{f} \sim Ro$, and in the BV limit $\tau_u \approx \frac{L}{\sqrt{\Phi_0}} \sim Fr$. In each of the balance regimes τ_u will become much smaller than τ_b as the limit is approached, meaning that the scale of the denominator of (3.3.36) remains close to τ_b.

Recalling that we expect both δ and δ_t to scale like Ro in the small-Ro regimes and like Fr^2 in the small-Fr regime, this suggests that δ_u will scale like Ro^2 in the small-Ro regimes, and like Fr^3 in the small-Fr regime. It should be noted that due to the fact that $Fr = 0.0625$ in all the Strong Rotation runs, this means that τ_u is actually constant in this regime, and so we expect the SR regime to scale like Ro in our cases.

Figure 3.4 shows the results for each regime. We expect data points to lie parallel with lines of the same colour. This appears to be the case for all but the quasigeostrophic regime, suggesting that although rough, this idea might have some merit. The results from the quasigeostrophic case may be a result of the aforementioned issue with RMS values, where a cancellation could occur if (3.3.36) were calculated in a manner consistent with the diagnostics.

3.4 Summary

This chapter has introduced the basic techniques and methods that will be used in the rest of part II. We have described each of the four regimes we will investigate and defined balance equations relevant in each of them. We have also given an overview of how these diagnostic will be measured and represented in the rest of part II.

The initial condition has been outlined along with a number of methods we applied in order to limit the presence of gravity waves not generated by the flow itself. Considerations that have helped us define the parameters of the experiments have been discussed. We also briefly discussed the role of the Rossby radius, which provides further information about the nature of the flow in each regime.

Chapter 4

Finite difference Model

4.1 ENDGame

The diagnostics derived in chapter 3 will be tested on a Semi-implicit Semi-Lagrangian model known as ENDGame.

ENDGame (Even Newer Dynamics for General atmospheric modelling of the environment) is the current dynamical core of the Met Office's Unified Model (UM). ENDGame was introduced as the dynamics core in July 2014 as a replacement for New Dynamics [1]. It offered improvements both in terms of accuracy but also scalability with more powerful computers, allowing the model to utilise more processors as the capacity of supercomputers increase.

The model uses a latitude-longitude grid with a C-grid placement of variables such that v-points are located at the poles. The version we are using is a shallow water model using the same methods as ENDGame, though the full version is for deep atmospheres. As the shallow water equations on the sphere involve a zero Coriolis parameter at the equator, this would cause problems for our diagnostics as we would have an infinite Rossby Number. In order to avoid this we have modified the model to use a constant Coriolis Parameter typical of mid-latitudes.

The semi-Lagrangian method involves 'tracing' a parcel of air back to its previous location and evaluating its properties there via interpolation. For a fully Lagrangian scheme, this would involve restructuring the grid with each timestep and eventually the parcels deform into long filaments, leading to a non-uniform grid [38]. The semi-Lagrangian scheme avoids this by interpolating the deformed grid back onto the structured grid at each timestep. In general this method is non-conservative, but the shallow water version ENDGame includes an option to use an inherently conserving algorithm [61]. Though it is not used operationally, we employ it here. This technique inherently satisfies the condition for stability that the numerical domain of dependence includes the analytical domain of de-

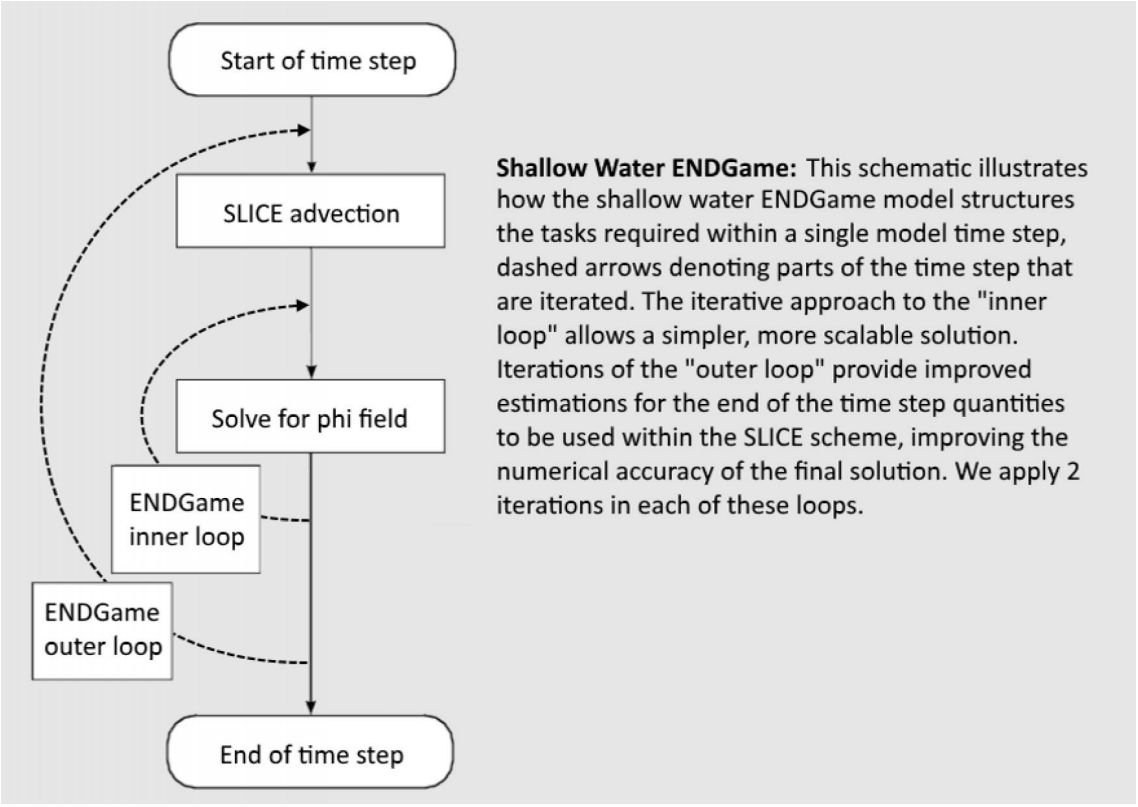

Figure 4.1: An outline of the solution algorithm for each time step, modified from ENDGame overview report, July 2014 [1].

pendence.

The semi-implicit scheme ensures stability by averaging terms over each time step. This increases the largest possible timestep for stability, and the fastest motions (waves) are slowed down such that they satisfy the CFL condition. This method assumes that accurate representation of these waves is unimportant to the evolution of the flow. Both these methods contribute to the model being able to run at large Courant numbers. For more detail see Zerroukat et al. [62].

Work has already begun on the replacement for ENDGame, known as GungHo. A shallow water candidate model from this project is used as our Eulerian model in chapter 5.

Figure 4.1 shows a representation of the model we use here, based on a similar figure used for the full ENDGame model. Before the outer loop, the current-time-level diagnostics are calculated. During the advection routine velocity terms are interpolated to their departure points, and rotated to the arrival point local coordinate system. Within the inner loop, next-time-level Coriolis terms are included and the system is iteratively solved for the geopotential. Next-time-level geopotential terms are then added in. Once the iteration is complete, the diagnostics are averaged to phi points (cell centres).

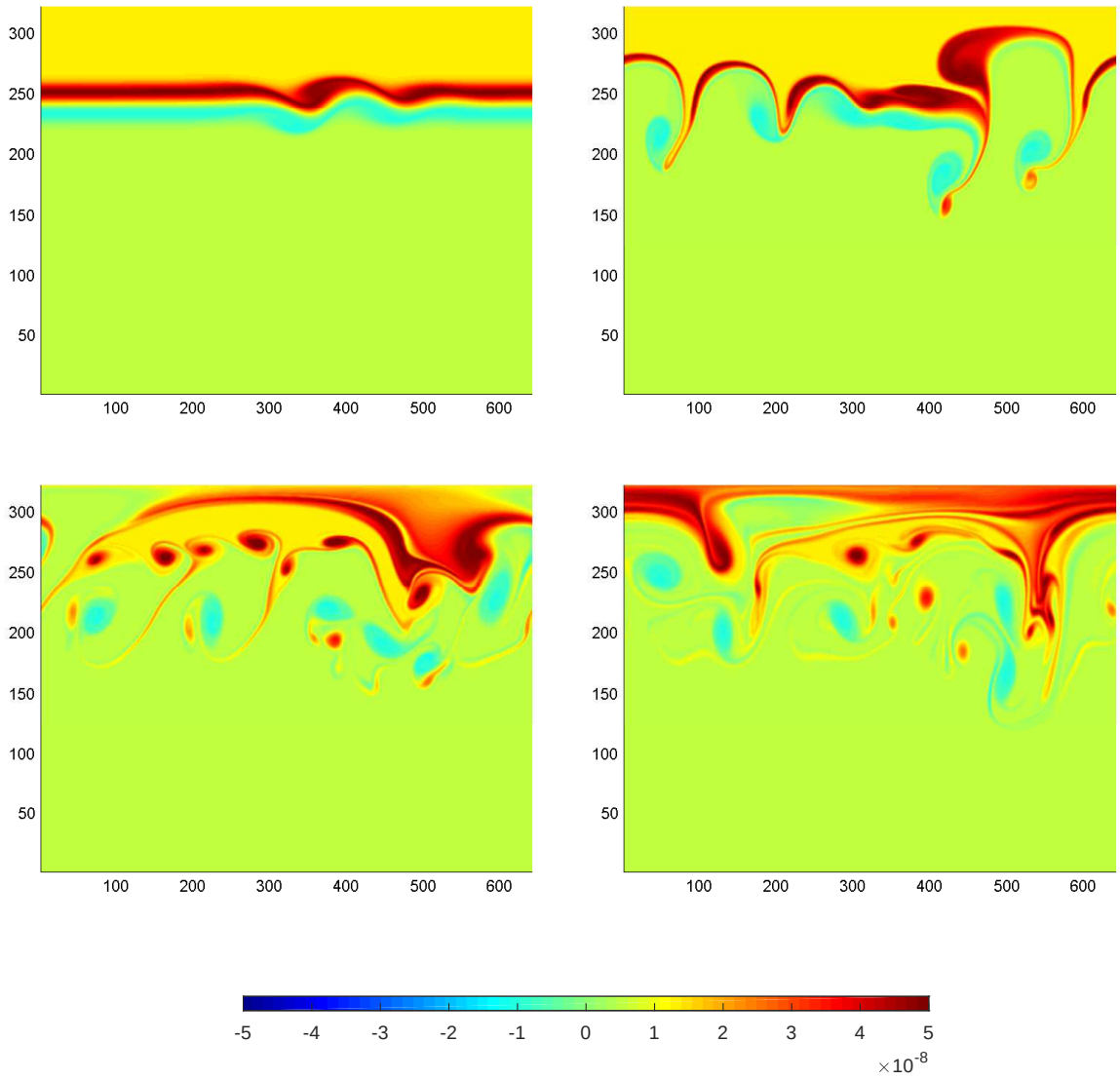

Figure 4.2: Potential vorticity maps showing the evolution of the flow at days 3 (top left), 6 (top right), 9 (bottom left) and 12 (bottom right). We can see the beginnings of instability and the wrapping of up vortices within the jet. The grid is 640x320 and parameters for this run are $Ro = 1$, $Fr = 1$.

4.2 Results

4.2.1 Field maps

Figure 4.2 shows the evolution of potential vorticity on a 640x320 grid in the $Ro = Fr = 1$ case matching figure 3.3. The beginnings of the instability can be seen at day 3. The peak in the in-jet diagnostic around day 5 in figure 3.3 coincides with first wrapping-up of vortices in the jet. These vortices can be seen in figure 4.2 at day 6. At days 9 and 12 the flow has developed a number of small vortices within the jet region, with cyclone-anticyclone pairs predominantly south of the jet. There is strong positive vorticity north of the jet, while the southernmost regions

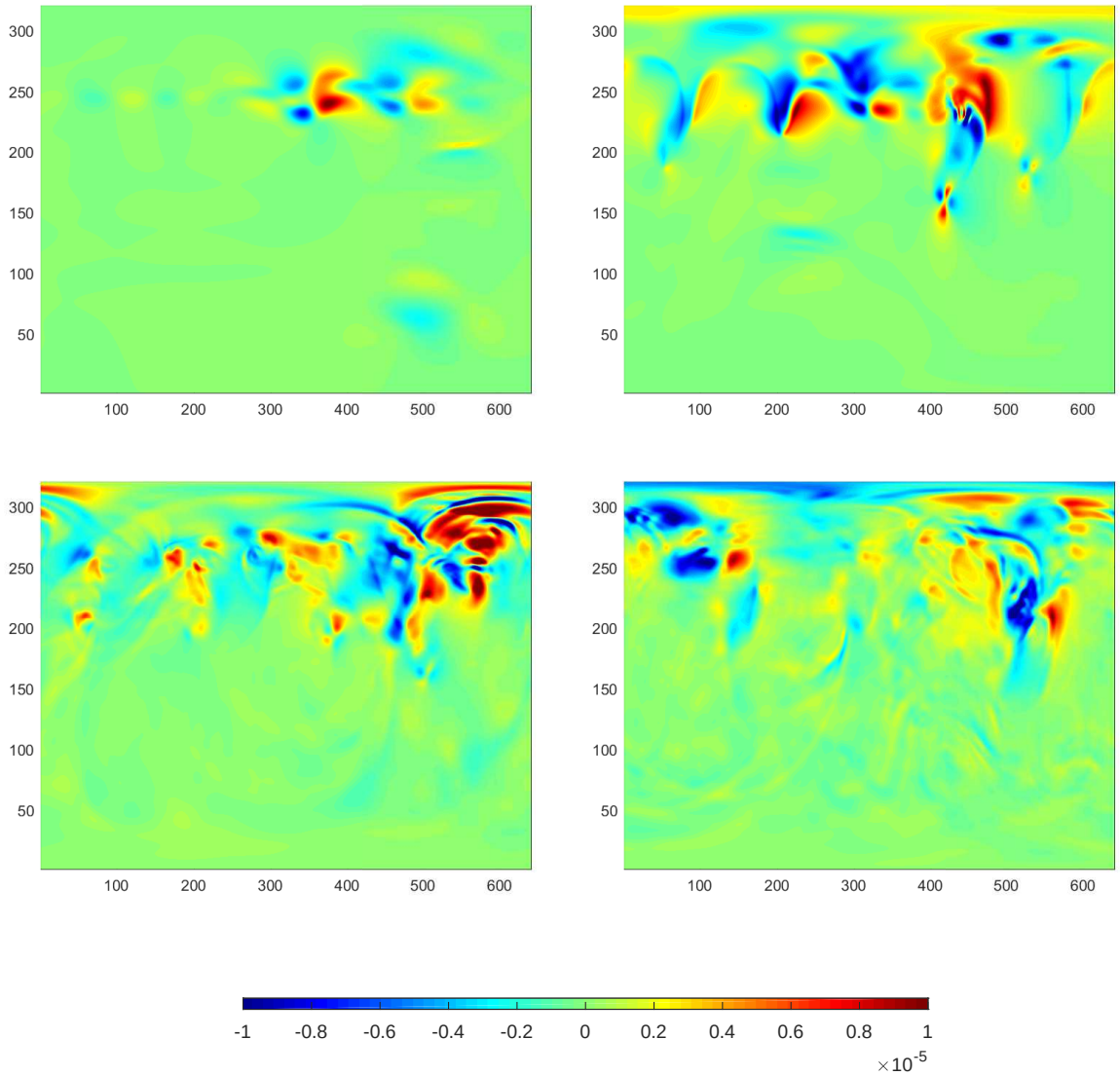

Figure 4.3: Divergence maps showing the evolution of the flow at days 3 (top left), 6 (top right), 9 (bottom left) and 12 (bottom right). Activity far from the jet could be weak gravity waves. Parameters for this run are $Ro = 1$, $Fr = 1$.

remain free of vorticity.

Figure 4.3 shows the evolution of divergence in the same case. Gravity wave activity at day 3 appears relatively weak, with areas of higher divergence concentrated around the beginnings of the instability. At day 6 there is higher divergence across the jet region, particularly around vortices (compare with figure 4.2). Day 9 in particular shows what look like gravity waves in the upper right quarter of the field, in the region of several vortices. This could suggest that they have been spontaneously emitted by the flow. Unlike vorticity, we can see variations in divergence throughout the entire map. It is likely that these are the very weak gravity waves that did result from the discretisation of the initial condition. The weakness of this divergence compared to that found in the jet suggests that we have been largely successful in our attempts to curb gravity wave activity deriving from the initial conditions.

4.2.2 Time series

Figure 4.4 shows an example of the time series for each of the 4 diagnostic quantities as they approach the barotropic vorticity limit regime. These diagnostics are calculated by dividing the map into 4 regions, and calculating the root mean square (RMS) values across each of these regions at each time step. The top figures have a small parameter $\varepsilon = O(1)$ and in each subsequent figure the small parameter is halved, so the bottom figure has a small parameter $\varepsilon = O(0.0625)$. As in figure 3.3 blue represents the region south of the jet, green is the 'in-jet' region, red is the region north of the jet and cyan is the global RMS. Full results from every regime may be found in Appendix A.

There is spurious scaling in the higher-order imbalance diagnostic north of the jet region (which we will discuss later). In cases where this is far larger than the other regions' diagnostics the axes have been scaled so as to keep the in-jet diagnostic visible. This may have the effect of 'cutting off' the north-of-jet and global diagnostics from the figure.

4.2.3 Larger time step

Figure 4.5 shows the results of a run with all parameters identical to the barotropic vorticity regime run (figure 4.4) except for the time step, which was 10 times the size. This corresponds with $C_{gw} = 5$. Note the similarity between the small- and large-timestep runs with identical Rossby and Froude numbers, especially in the important jet region.

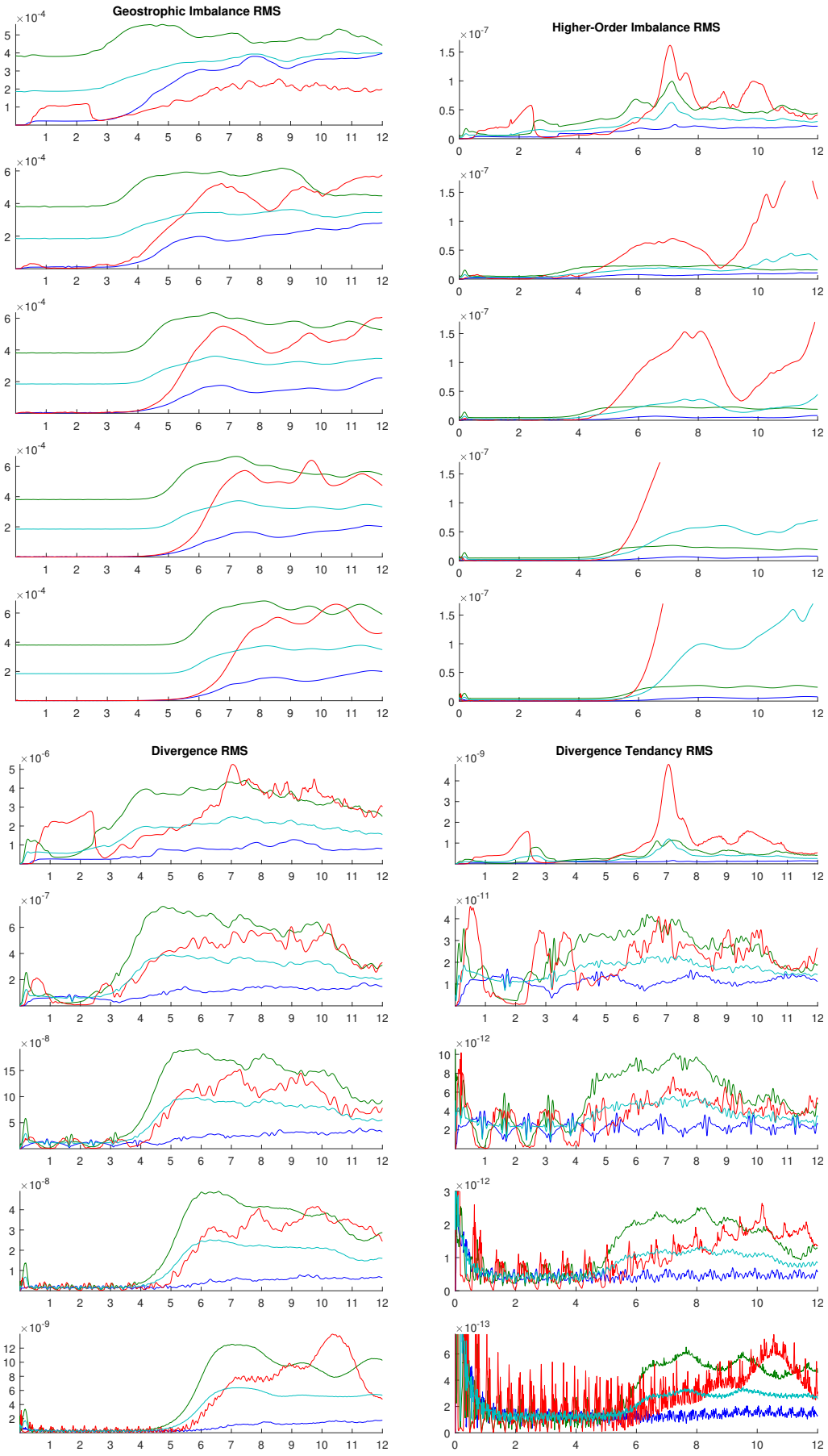

Figure 4.4: Time series of diagnostic quantities approaching the barotropic vorticity limit in the ENDGame model, $C_{gw} \approx 0.5$. For each subfigure the small parameter ε is, from top to bottom, = 1, 0.5, 0.25, 0.125, 0.0625.

The scaling of the activity north of jet seems unusual in the higher order imbalance. Investigation shows that this scaling is roughly inversely proportional to the size of the time step (so ten times smaller with the ten times time step). Removing each of the terms in turn from the higher-order imbalance diagnostic suggests that this is due to incorrect interpolation of the $\frac{D}{Dt}\nabla\Phi'$ term over the pole, though we have not been able to find the exact reason for this. It is logical that the lack of this scaling as we approach the semigeostrophic regime is due to the reducing Rossby radius, meaning the higher-order imbalance remains minimal beyond the jet region. This effect is also likely to be present in the south-of-jet region, but it is diluted by the region being larger and $\nabla\Phi'$ being smaller. The spurious values exist only at the northernmost gridpoints, and do not affect adjacent gridpoints as the higher-order imbalance equation is purely diagnostic. Therefore while this particular behaviour is undesirable, the in-jet region is the region of most interest, so this will not impact significantly upon the results we investigate here.

It is clear that despite small differences between the results, the most crucial aspect (being the amplitude of the measures of imbalance) remains almost identical in each case. This supports the view that the semi-implicit model enables simulations at large time steps without significantly impacting upon the accuracy of the results, specifically the model's ability to capture balance.

4.2.4 Convergence

In order to satisfy ourselves that the system is converging towards one of the asymptotic limits it is necessary to represent an entire time series with a single number. As previously discussed, we must accept that our time series are going to look different in each regime and demonstrate slightly different behaviour.

For example, if we view our 'base' simulation as the $Ro = Fr = 1$ case, then we can see from the time series that the general shape of the evolution is consistent approaching the QG limit (see figure 4.6 for an example). The initial condition is designed to be divergence-free, but the jet is not geostrophically balanced. Thus where the divergence-based diagnostics start near zero, we see this is not the case for the geostrophic imbalance diagnostic. The higher-order imbalance diagnostic to leading order is equivalent to the geostrophic imbalance, so the fact that this diagnostic is also near zero at $t = 0$ (as shown in figure 4.4) demonstrates the effects of cancellation from the higher order terms. Each diagnostic remains the same scale through the initial period of the runs, and increases only when the vortices begin to wrap up. This occurs first in the jet region and then shortly after in the north-of-jet region and then the south-of-jet as the vorti-

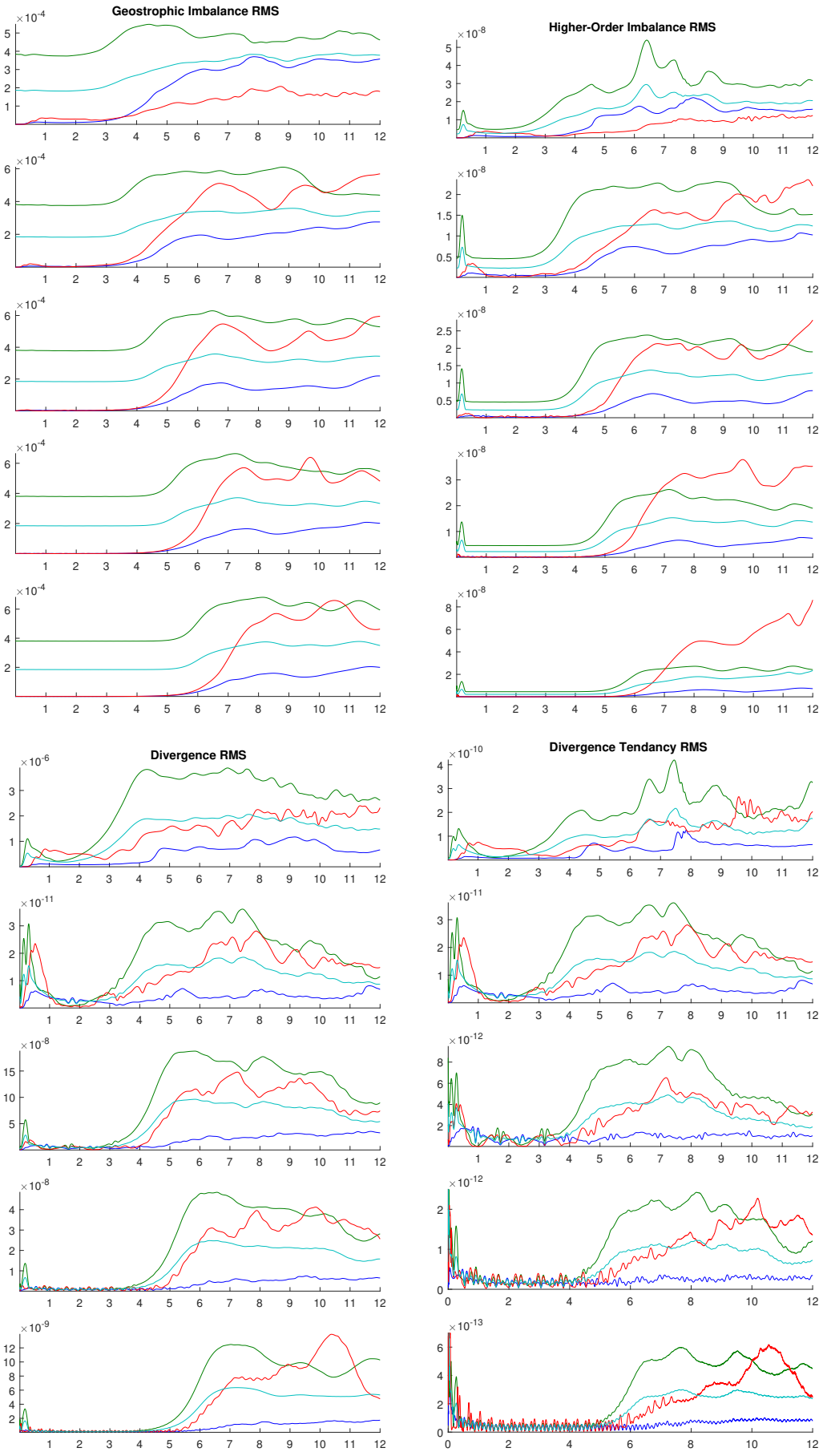

Figure 4.5: Time series of diagnostic quantities approaching the barotropic vorticity limit in the ENDGame model, $C_{gw} \approx 5$. For each subfigure the small parameter ε is, from top to bottom, = 1, 0.5, 0.25, 0.125, 0.0625.

ces spread out from the jet. The difference in the mean values of the regions is due to the south-of-jet region being larger and the vortices do fill it. The last two panels (closest to the asymptotic limit) in each figure of each diagnostic look very similar, appearing to demonstrate a strong convergence to a solution.

The time series in figure 4.7 appears to become 'smoother' approaching the SG limit, peaking later and the wave-like oscillations shown predominantly in the north-of-jet series becoming less visible. The scales of the diagnostics in each regions' time series also separate, indicating that the behaviour within each region is becoming more distinct. This fits with the Rossby radius shrinking and so 'compacting' the vortices as we approach the SG limit.

The BV limit retains its shape but takes longer to destabilise. This fact makes convergence less obvious just from looking at the time series, though consistent traits do appear as we approach the limit.

The SR limit as we have defined it effectively transitions from the run closest to the BV limit ($Fr = \varepsilon = 0.0625$) to the run closest to the QG Limit ($Ro = Fr = \varepsilon = 0.0625$). Compare figuresAs such we see the flow evolving sooner as we approach the limit. Similar to the figure of the BV limit, the convergence is not certain, though there are consistent features.

If we are looking for evidence of gravity waves in the system, the natural place to look is in the divergence tendency diagnostic, which we touched on briefly in chapter 3. Previously we have discussed a number of factors that limit both the amplitude and the representation of gravity waves within the regimes we are looking at, which can make it hard to spot gravity waves from the diagnostics we have available. A convenient signal of gravity waves that we might best use is when gravity waves crash into each other, as this might result in a brief 'kick' in the diagnostic. Because our model operates on a sphere, this will happen reliably as north-moving gravity waves move over the pole and meet other north-moving gravity waves. To see this we can look at the diagnostics from the barotropic vorticity and strong rotation regimes (figures 4.4 (bottom right) and 4.8) where there is a very clear series of such kicks in the north-of-jet region. In particular, looking at the barotropic vorticity regime we see that the frequency of these kicks is increasing in direct proportion to the Froude number decreasing - which is exactly what we would expect to see as the gravity wave speed increases with increased fluid depth.

The most consistent trait in all the tests is the major initial peak that corresponds with the wrapping-up of large vortices as mentioned in the "Figures" section above. Comparative field maps of the PV are presented here that show

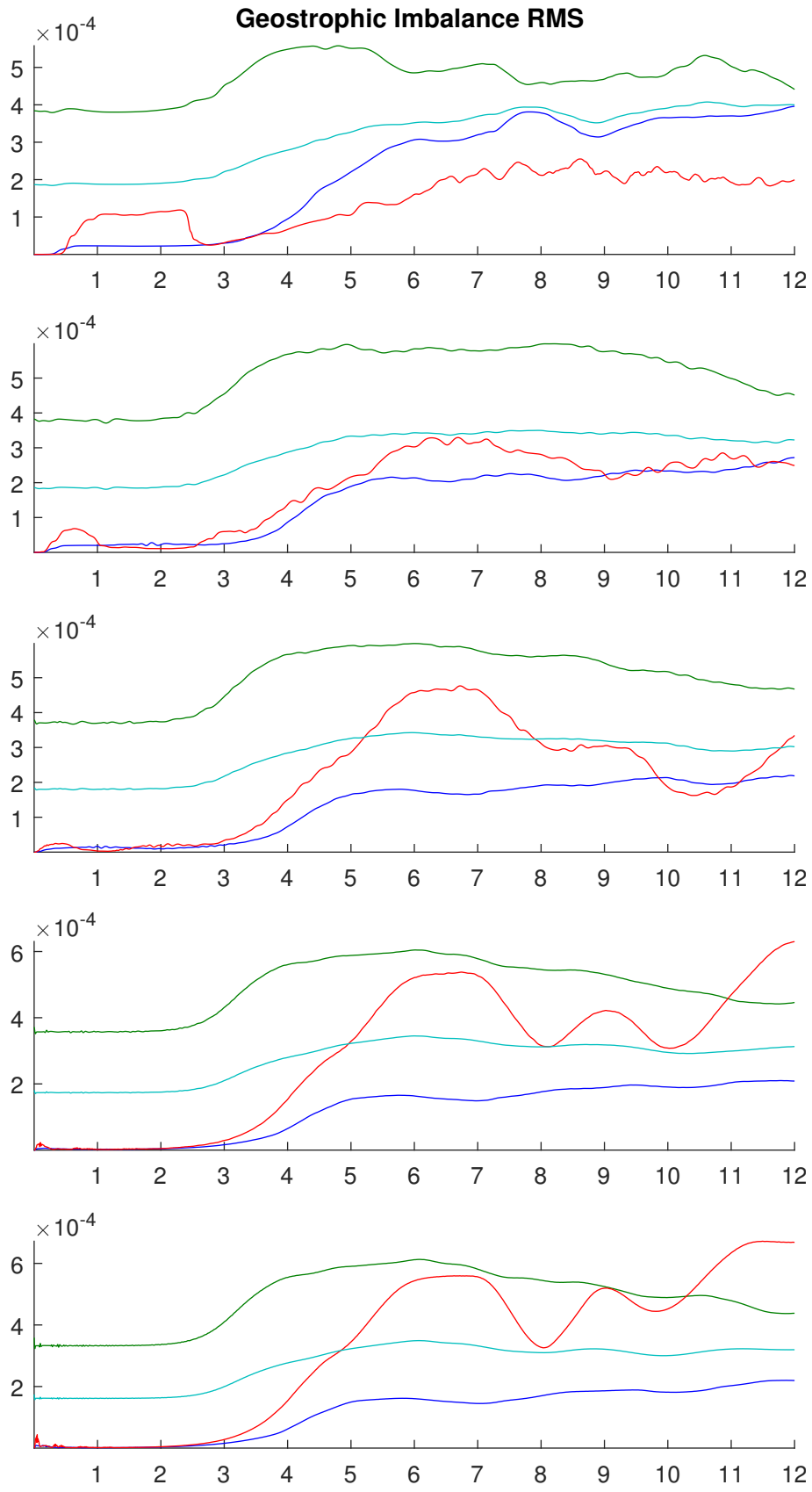

Figure 4.6: Time series of geostrophic imbalance approaching the quasigeostrophic limit in the ENDGame model, $C_{gw} \approx 0.5$. For each subfigure the small parameter ε is, from top to bottom, = 1, 0.5, 0.25, 0.125, 0.0625. The similarity of the bottom two panels suggests convergence to a solution as we approach the asymptotic limit.

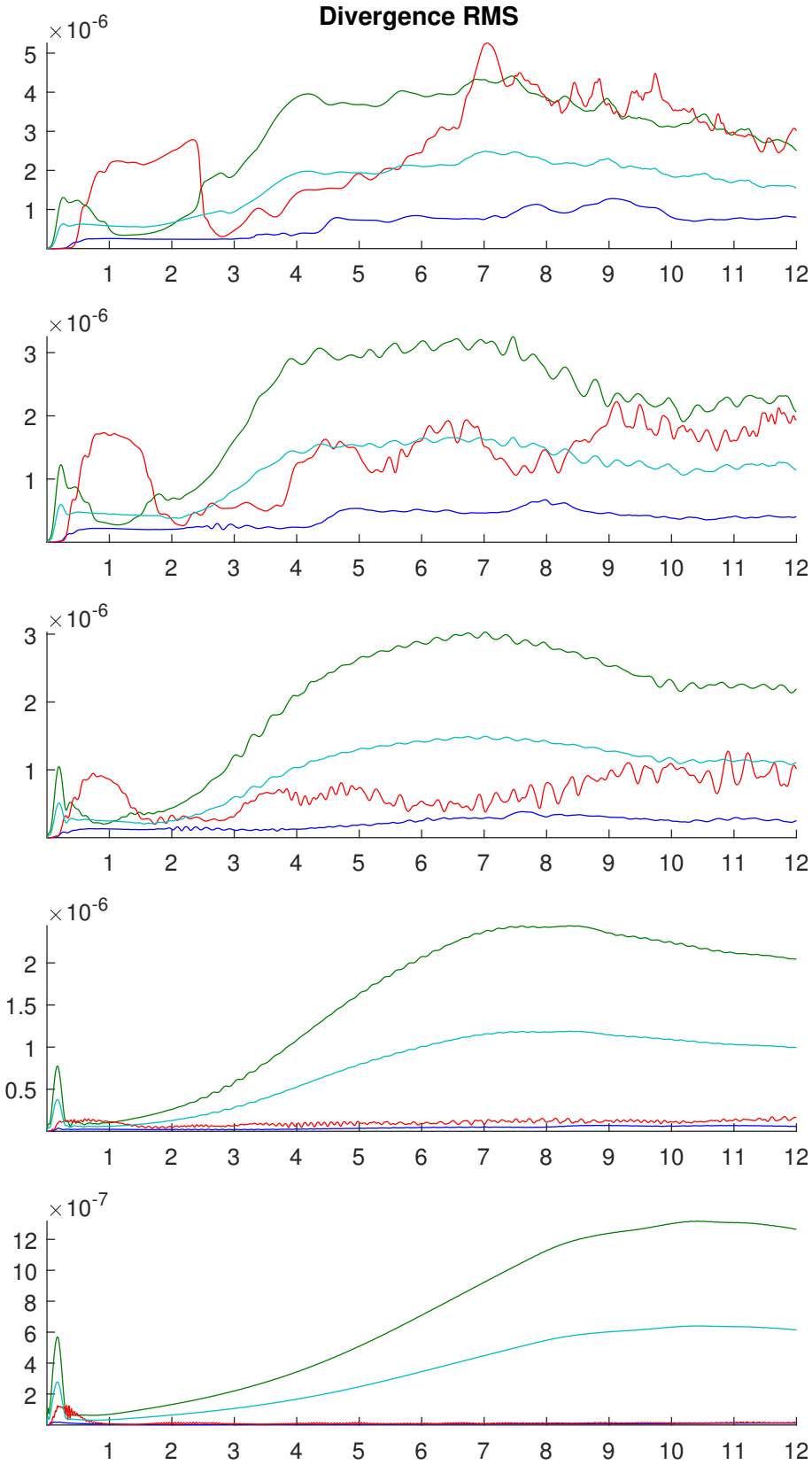

Figure 4.7: Time series of divergence approaching the semigeostrophic limit in the ENDGame model, $C_{gw} \approx 0.5$. For each subfigure the small parameter ε is, from top to bottom, = 1, 0.5, 0.25, 0.125, 0.0625. The flow appears to be evolving slower, and divergence becomes predominantly contained within the jet as we would expect as the Rossby radius shrinks.

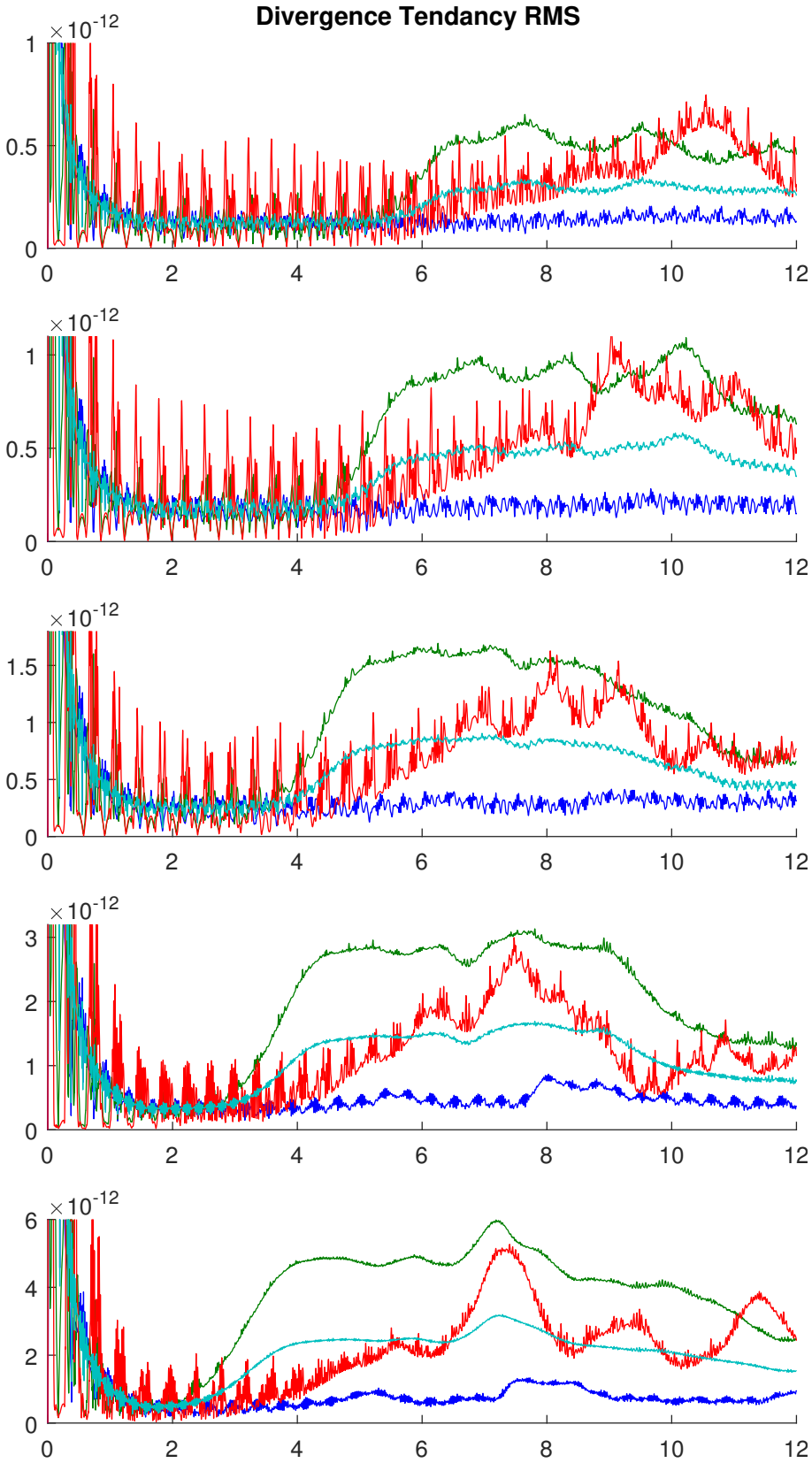

Figure 4.8: Time series of divergence tendency approaching the strong rotation limit in the ENDGame model, $C_{gw} \approx 0.5$. For each subfigure the small parameter ε is, from top to bottom, = 1, 0.5, 0.25, 0.125, 0.0625. There is a clear pattern of "bursts" in divergence tendency north of the jet, believed to be evidence of gravity waves.

the similarity of the field at this point.

Figure 4.9 shows PV maps approaching the quasigeostrophic limit at days 4 and 7. There is a clear similarity between runs with smaller small parameters strongly suggesting convergence to a solution.

Figure 4.10 shows the same fields, this time approaching the semigeostrophic limit. We can clearly see the effects of the shrinking Rossby radius, which make it harder to be confident that the time chosen is equivalent to runs with different small parameters. However, we can see that some features (such as number of vortices) persist despite the 'compacting', allowing us to be relatively confident that we are looking at equivalent points in time in the evolution of the flow.

Figures 4.11-12 show the PV field approaching the barotropic vorticity limit. Figure 4.11 show the same times as the previous two figures. It is clear that the flow is not evolving on the same timescale as we approach the limit. As we discussed in chapter 3, the most likely cause for this is that we failed to appropriately scale the perturbation in the initial condition and as we approached the barotropic vorticity limit (the scaling we used has no Fr-dependence). For this reason we looked at offsetting the time by 16 hours for each halving of the small parameter. The results can be seen in figure 4.12. Here the leftmost figures look much more like those we saw for the QG limit. Although the later time means the pattern of vortices has been advected further eastward by the jet we can still be confident that we are looking at a roughly equivalent point in the evolution.

The peak in the higher order imbalance diagnostic at around day 7 was therefore chosen as the most reliable point at which to evaluate our diagnostics. The results are presented below for each of the diagnostics, with all relevant regimes included. Points represent the RMS of the diagnostic in the in-jet region taken at the time corresponding with this peak, while lines indicate the scaling expected from the theory. If the numerical model respects the scaling then the points should lie parallel to these lines. The results are shown in figure 4.16.

The scaling of divergence is close to the expectation in the quasigeostrophic and barotropic vorticity regimes. The semigeostrophic regime scales slower than expected initially, though it appears as though the scaling is beginning to converge as the small parameter gets smaller.

A similar picture is evident in the geostrophic imbalance - again, the QG and SR limit behave consistently with predictions while convergence in the SG limit only seems to appear as the regime becomes more enforced.

The three regimes tested for convergence of the higher-order imbalance appear to scale as expected initially, however as the small parameter gets smaller we see the imbalance in the system increasing. A possible explanation for this is

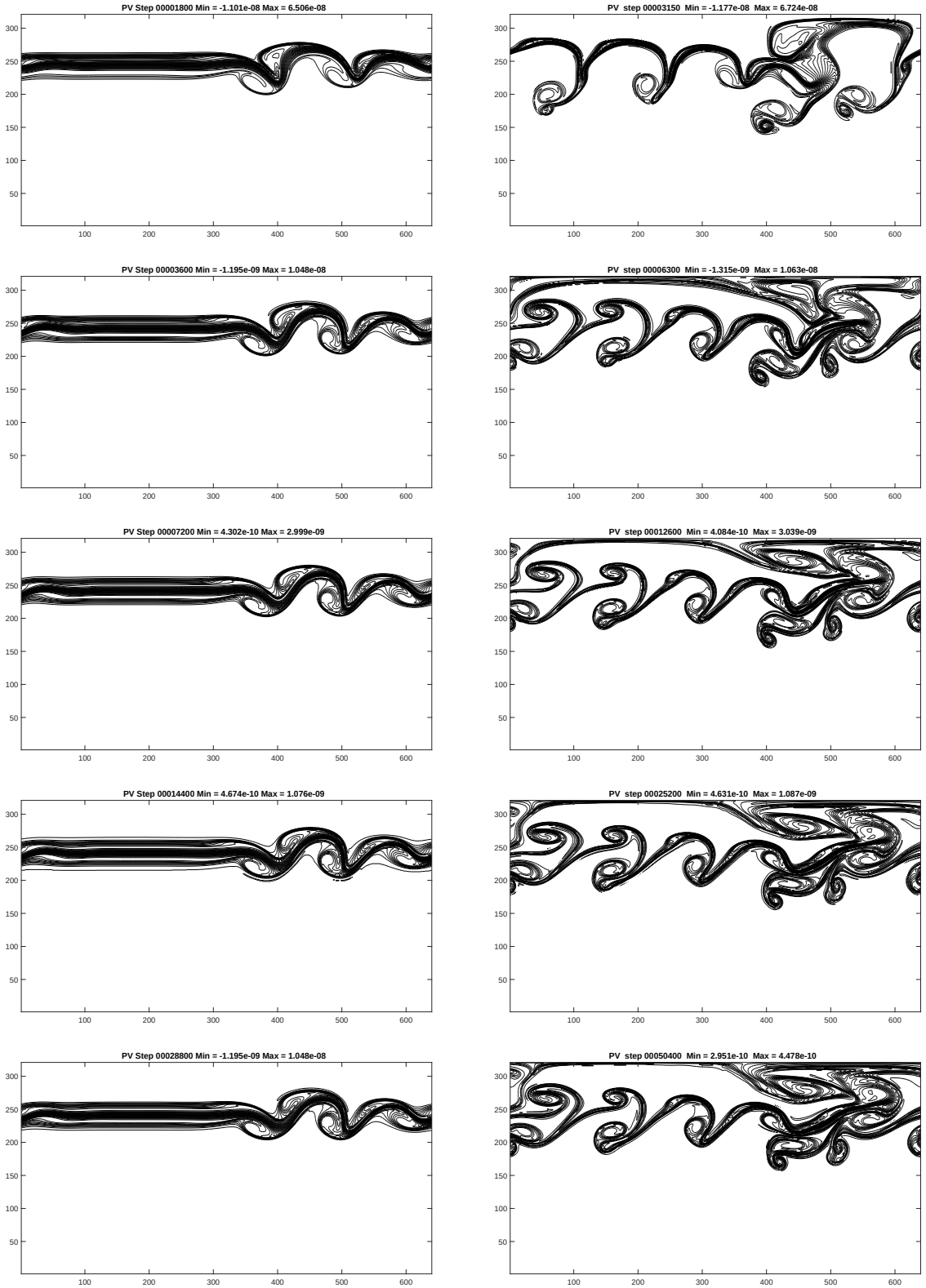

Figure 4.9: PV maps at day 4 (left) and day 7 (right) approaching the quasigeostrophic limit. The small parameters from top to bottom are ε =1, 0.5, 0.25, 0.125 and 0.0625. $C_{gw} = 0.5$. As seen previously the bottom two panels suggest convergence to a solution as the asymptotic limit is approached.

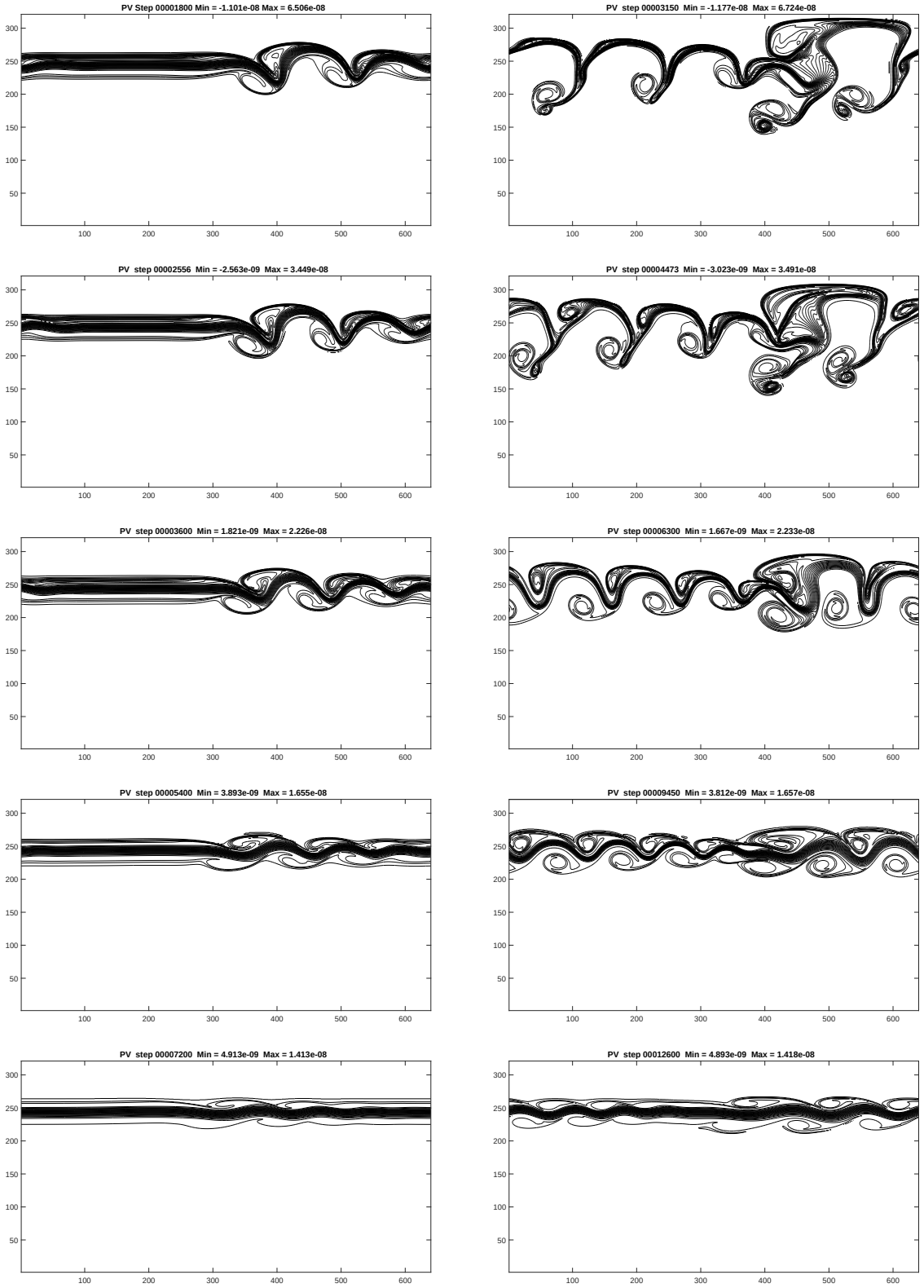

Figure 4.10: PV maps at day 4 (left) and day 7 (right) approaching the semigeostrophic limit. The small parameters from top to bottom are ε =1, 0.5, 0.25, 0.125 and 0.0625. $C_{gw} = 0.5$. As the Rossby radius shrinks, we see features of the flow compacted within the jet.

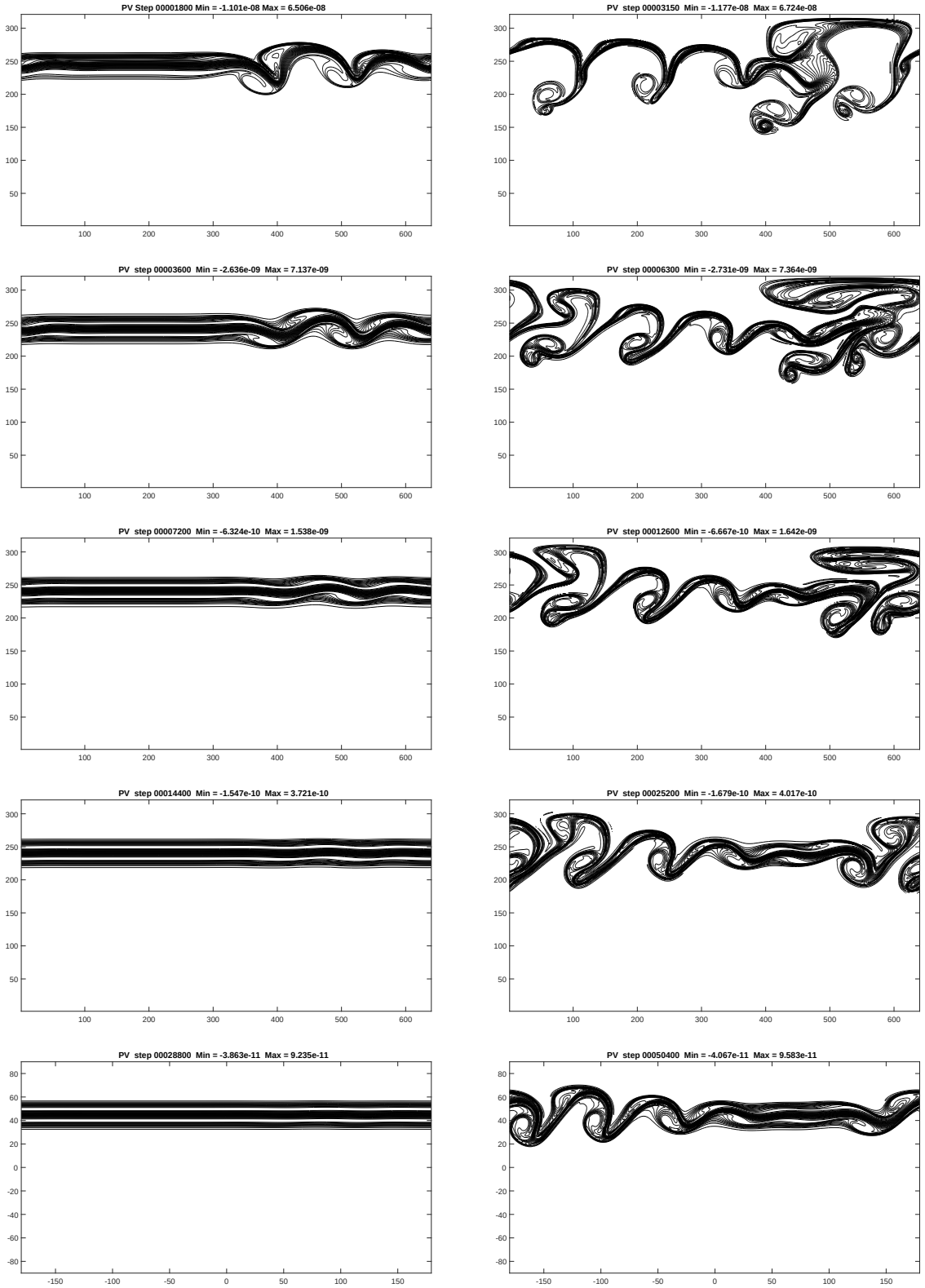

Figure 4.11: PV maps at day 4 (left) and day 7 (right) approaching the barotropic vorticity limit. The small parameters from top to bottom are ε =1, 0.5, 0.25, 0.125 and 0.0625. $C_{gw} = 0.5$. The flow appears to develop more slowly in this regime.

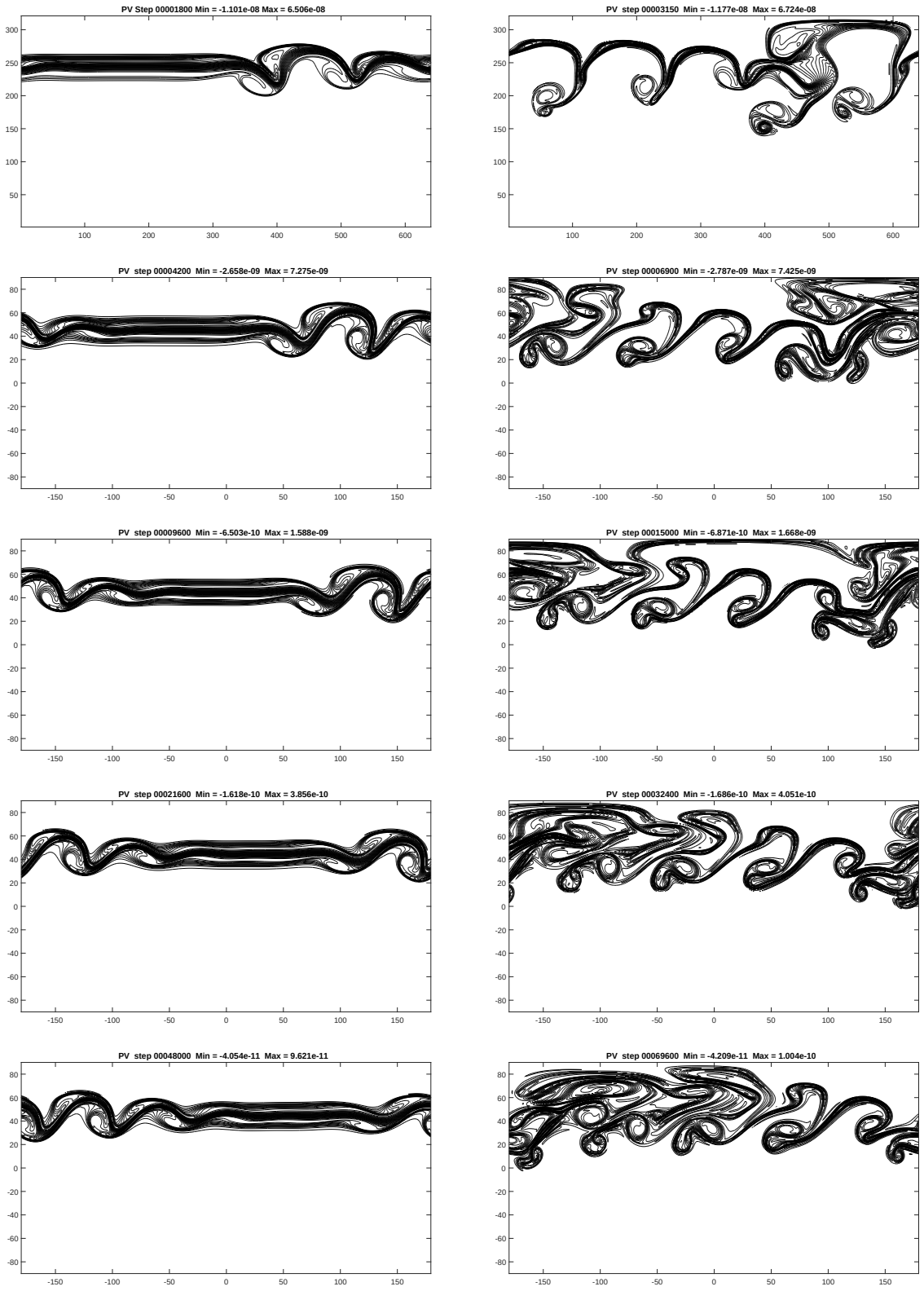

Figure 4.12: PV maps approaching the barotropic vorticity limit. From top to bottom the time is:

Left - 4 days, 4 days 16 hours, 5 days 8 hours, 6 days, 6 days 16 hours.

Right - 7 days, 7 days 16 hours, 8 days 8 hours, 9 days, 9 days 16 hours.

The small parameters from top to bottom are ε =1, 0.5, 0.25, 0.125 and 0.0625. $C_{gw} = 0.5$. Although the flow is evolving more slowly, its behaviour is very similar.

given below.

In both the time series and the convergence figures, there is minimal difference between the results from the two different time steps. One exception to this is when both the Rossby and Froude number are approximately unity. There is spurious scaling in the north-of-jet higher-order imbalance RMS (see, for example, Figure 4.5). This seems to be due to the ENDGame's handling of the time derivative term when interpolating it across the pole. Inspection shows that the spurious values of the imbalance occur only at the northernmost 'ring' of gridpoints, and that these values scale with the size of the timestep (like $\frac{1}{\Delta t}$ as can be seen in the diagnostics). However this ring exists purely in the higher-order diagnostic and it doesn't seem to affect the behaviour of the rest of the system.

Unexpected scaling

We have shown that the low-order diagnostics of geostrophic imbalance and divergence respect the scalings we expected. Although the higher-order imbalance initially appears to scale in a manner consistent with our predictions we see a sudden increase in the RMS values in the limit $\varepsilon < 0.25$, scaling like $\frac{1}{\varepsilon}$ (so doubling with each halving of the small parameter).

We believe that the source of the scaling of the higher-order imbalance is a result of an inconsistency between the analytic and numerical implementation of our imbalance equations. It has been assumed that terms that cancel in the calculation of the analytic form of the higher-order imbalance would be so close to zero in their numerical form that we could treat them as such. To test this, we must reconstruct the higher order imbalance equation numerically.

From the shallow water momentum equation in Lagrangian form:

$$\frac{D\mathbf{u}}{Dt} + f\mathbf{k} \times \mathbf{u} + \nabla\Phi = 0 \tag{4.2.1}$$

Define an interpolation operator $\mathcal{I}^n$ for the semi-Lagrangian advection at step n, and write the discretised momentum equation:

$$\frac{\mathbf{u}^{n+1} - \mathcal{I}^n\mathbf{u}^n}{\Delta t} + f\left(\alpha\mathbf{k} \times \mathbf{u}^{n+1} + \beta\mathcal{I}^n(\mathbf{k} \times \mathbf{u}^n)\right) + \alpha\nabla\Phi^{n+1} + \beta\mathcal{I}^n\nabla\Phi^n = 0 \tag{4.2.2}$$

where ∇ is now the discrete operator, and $f\mathbf{k} \times \mathbf{u}$ involves spatial averaging. The same equation may be written one time step earlier:

$$\frac{\mathbf{u}^{n} - \mathcal{I}^{n-1}\mathbf{u}^{n-1}}{\Delta t} + f\left(\alpha\mathbf{k} \times \mathbf{u}^{n} + \beta\mathcal{I}^{n-1}(\mathbf{k} \times \mathbf{u}^{n-1})\right) + \alpha\nabla\Phi^{n} + \beta\mathcal{I}^{n-1}\nabla\Phi^{n-1} = 0 \tag{4.2.3}$$

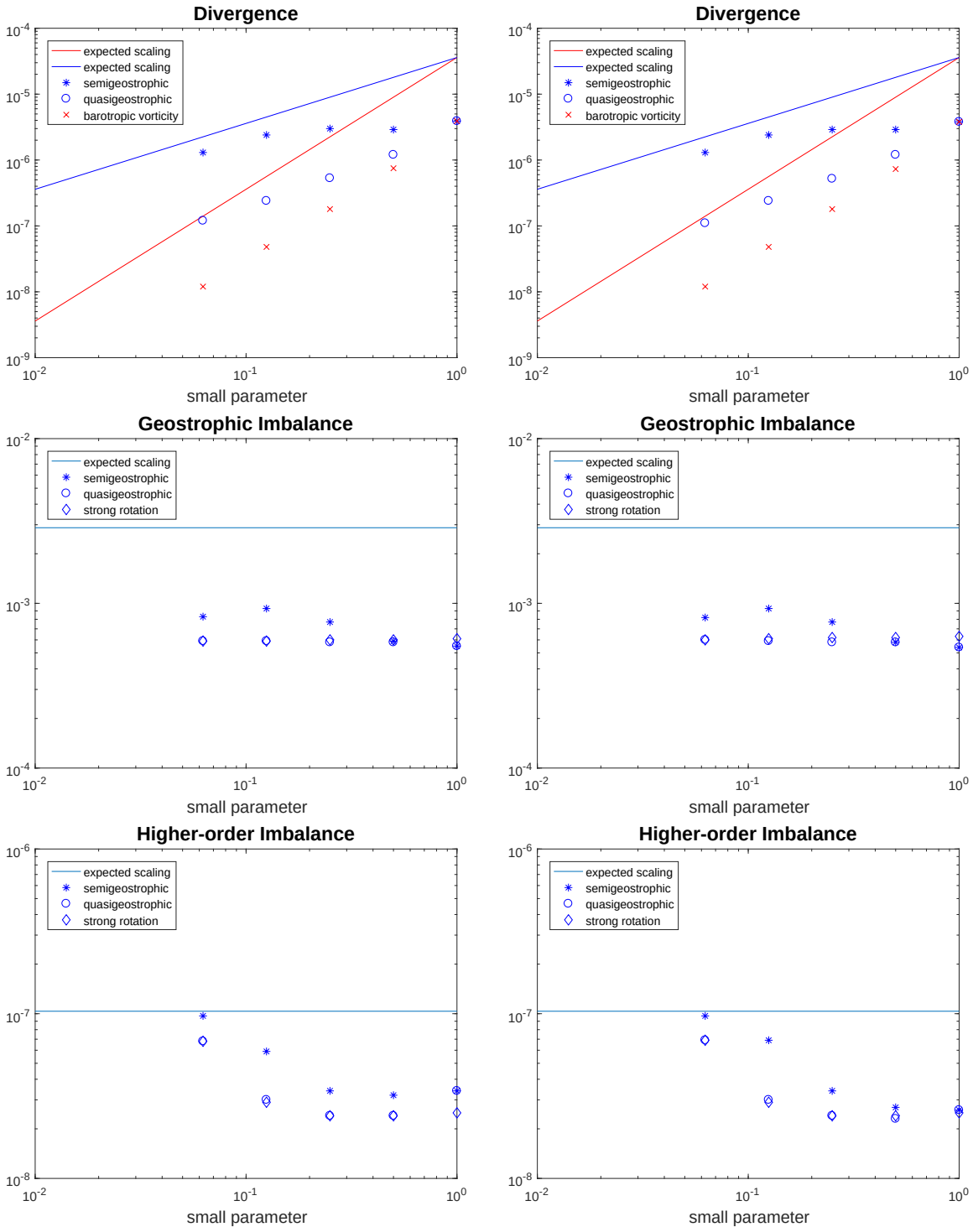

Figure 4.13: Convergence diagrams for each in-jet diagnostic. Left hand side is small time step ($C_{gw} \approx 0.5$), right hand side is $10\times$ time step ($C_{gw} \approx 5$). Red points are expected to lie parallel to the red lines, and blue points are expected to lie parallel to the blue lines. Results are encouraging for divergence and geostrophic imbalance diagnostics, but there is clear spurious scaling in the higher-order imbalance.

With these two equations we can make the first step towards the higher-order imbalance equation by taking the time derivative of the momentum equation $\frac{1}{\Delta t}$((4.2.2) - $\mathcal{I}^n$(4.2.3)):

$$\begin{aligned}
&\frac{\mathbf{u}^{n+1} - 2\mathcal{I}^n\mathbf{u}^n + \mathcal{I}^n(\mathcal{I}^{n-1}\mathbf{u}^{n-1})}{\Delta t^2} \\
&+\frac{f}{\Delta t}\left(\alpha(\mathbf{k}\times\mathbf{u}^{n+1} - \mathcal{I}^n(\mathbf{k}\times\mathbf{u}^n)) + \beta(\mathcal{I}^n(\mathbf{k}\times\mathbf{u}^n) - \mathcal{I}^n(\mathcal{I}^{n-1}(\mathbf{k}\times\mathbf{u}^{n-1})))\right) \\
&+\frac{1}{\Delta t}\left(\alpha(\nabla\Phi^{n+1} - \mathcal{I}^n\nabla\Phi^n) + \beta(\mathcal{I}^n\nabla\Phi^n - \mathcal{I}^n(\mathcal{I}^{n-1}\nabla\Phi^{n-1}))\right) = 0
\end{aligned} \tag{4.2.4}$$

Now we can substitute terms from (4.2.2) and (4.2.3) to give the numerical equivalent of the full imbalance equation. The second term of (4.2.4) can therefore be written:

$$\begin{aligned}
&\frac{f}{\Delta t}\Bigg[\alpha\Big(\mathbf{k}\times\left(\mathcal{I}^n\mathbf{u}^n - \Delta t\left(f(\alpha\mathbf{k}\times\mathbf{u}^{n+1} + \beta\mathcal{I}^n(\mathbf{k}\times\mathbf{u}^n)) - \alpha\nabla\Phi^{n+1} - \beta\mathcal{I}^n\nabla\Phi^n\right)\right) - \mathcal{I}^n(\mathbf{k}\times\mathbf{u}^n)\Big) \\
&\quad+\beta\Big(\mathcal{I}^n(\mathbf{k}\times(\mathcal{I}^{n-1}\mathbf{u}^{n-1} - \Delta t\left(f(\alpha\mathbf{k}\times\mathbf{u}^n + \beta\mathcal{I}^{n-1}(\mathbf{k}\times\mathbf{u}^{n-1})) - \alpha\nabla\Phi^n - \beta\mathcal{I}^{n-1}\nabla\Phi^{n-1}\right)\Big) \\
&\quad-\mathcal{I}^n(\mathcal{I}^{n-1}(\mathbf{k}\times\mathbf{u}^{n-1}))\Big)\Bigg]
\end{aligned} \tag{4.2.5}$$

or

$$\begin{aligned}
&\frac{f}{\Delta t}\left[\alpha\Big(\mathbf{k}\times\mathcal{I}^n\mathbf{u}^n - \mathcal{I}^n(\mathbf{k}\times\mathbf{u}^n)\Big) + \beta\mathcal{I}^n\Big(\mathbf{k}\times\mathcal{I}^{n-1}\mathbf{u}^{n-1} - \mathcal{I}^{n-1}(\mathbf{k}\times\mathbf{u}^{n-1})\Big)\right] \\
&-f^2\left[\alpha^2\Big(\mathbf{k}\times\mathbf{k}\times\mathbf{u}^{n+1}\Big) + \alpha\beta\Big(\mathbf{k}\times\mathcal{I}^n(\mathbf{k}\times\mathbf{u}^n) + \mathcal{I}^n(\mathbf{k}\times\mathbf{k}\times\mathbf{u}^n)\Big) + \beta^2\Big(\mathcal{I}^n(\mathbf{k}\times\mathcal{I}^{n-1}(\mathbf{k}\times\mathbf{u}^{n-1})\Big)\right] \\
&-f\left[\alpha^2\Big(\mathbf{k}\times\nabla\Phi^{n+1}\Big) + \alpha\beta\Big(\mathbf{k}\times\mathcal{I}^n\nabla\Phi^n + \mathcal{I}^n(\mathbf{k}\times\nabla\Phi^n)\Big) + \beta^2\Big(\mathcal{I}^n(\mathbf{k}\times\mathcal{I}^{n-1}\nabla\Phi^{n-1}\Big)\right]
\end{aligned} \tag{4.2.6}$$

To simplify our calculations, some assumptions are made about the accuracy of approximations within the code. Of interest here are the following:

- $\mathbf{k}\times\mathbf{k}\times\mathbf{u}$ is a good approximation to $-\mathbf{u}$
- $\mathbf{k}\times\mathcal{I}^n$ is close to $\mathcal{I}^n(\mathbf{k}\times)$

$\mathbf{k}\times\mathbf{u}$ terms involve an average, and so $\mathbf{k}\times\mathbf{k}\times\mathbf{u}$ effectively involves reconstructing $\mathbf{u}$ from an average of an average.

If either assumption proves false then the code may not be able to successfully approximate the higher-order imbalance equation. Additionally, while lines 2 and 3 of equation (4.2.6) correspond to numerical approximations of terms in our imbalance equation, line 1 is assumed to disappear. If our second assumption is unjustified, then this will not be the case.

The terms on the first line of (4.2.6) feature an $\frac{f}{\Delta t}$ factor, yet our results have shown that the choice of Δt does not significantly affect the diagnostics. It is

plausible that the choice of interpolation implies that $\mathbf{k} \times \mathcal{I}^n \mathbf{u}^n - \mathcal{I}^n(\mathbf{k} \times \mathbf{u}^n)$ scales like Δt, which would result in the first line scaling like f. This would match the spurious scaling we observed. Investigation of the scale of these terms in our time series shows that they are of similar initial scale to the level of imbalance and so if they did scale with f then they would be expected to dominate the imbalance equation as we approach the limit. Evaluation of the right-hand side of the higher order imbalance equation ($\frac{D^2\mathbf{u}}{Dt^2}$) does not demonstrate this spurious scaling, further suggesting that the error is in the definition of the diagnostic rather than in the model itself.

4.3 Summary

We have introduced the ENDGame model with a brief overview of its key features and integration algorithm. Where appropriate we have also highlighted the differences between the model used in the thesis and the full ENDGame. Potential vorticity and divergence field maps were presented to demonstrate the evolution of the flow as the jet destabilises and vortices begin to form. The divergence field in particular showed what look like gravity waves in the southern hemisphere, far from the jet. The diagnostic time series were given alongside those of the equivalent runs with the $\times 10$ time step. These are seen to be very similar in terms of scale and general evolution. We went into more detail, describing how the 'shape' of the time series indicates the emergence of phenomena like vortex wrapping, and the variations between different regimes caused by how the Rossby radius scales. A signal in the divergence tendency diagnostic was suggested as a good indicator of gravity wave activity due to the frequency of the signal appearing to scale as gravity wave speed would scale, and that it is most clearly seen near the north pole. Some of the time series show that the timescale of the jet destabilising and vortices wrapping up is not consistent. We discuss this and suggest that this is due to an inadequacy in the way we formulated the perturbation that initiates the instability in the jet - leading to a smaller response and longer timescale for the destabilisation in barotropic vorticity regime. Convergence diagrams for each of the diagnostics have also been shown. Where the diagnostics have failed to converge to the expected scaling as predicted in chapter 3 an explanation has been attempted, though it would take further study to properly convince ourselves that the given explanation is the correct one.

The next chapter will follow a similar structure in the context of the Eulerian model.

Chapter 5

Finite element model

5.1 GungHo

The gridpoints on a latitude-longitude grid are very close together at the poles, and models using such a grid are expected to scale poorly on massively parallel computers as a result of this [44]. Models that are structured so as to apply to quasi-uniform spherical grids are expected to fare better, though key conservation, balance and potential vorticity properties are no longer straightforward to achieve. These properties of the continuous governing equations are mimicked by careful design of the finite element function spaces of the numerical scheme. The finite element model we look at here is designed for effective scalability on massively parallel supercomputers, though the model itself is not parallel. In addition to the scalability requirements, the aim is to achieve comparable accuracy to a latitude-longitude model without excessive computational cost. We will briefly look at some of the properties desirable for such an endeavour.

5.1.1 Key features

To ensure accurate representation of geostrophic adjustment [2] this model uses C-grid primal finite element spaces and dual spaces that are mapped to the primal spaces through a set of Hodge star operators. This results in a rather beautiful group of relations that link the function spaces together (see figure 5.1). The Hodge star operations are indicated by the vertical arrows, and the solid horizontal lines indicate strong (derivative) operators that hold at every point in the domain exception of discontinuities. As the discrete divergence is such an operator, this means that mass conservation holds in this model.

The discretisation that allows mass conservation is known as a *mimetic* property, as it mimics key properties of the continuous governing equations. Another example is the conservation of energy, which is achieved by ensuring certain

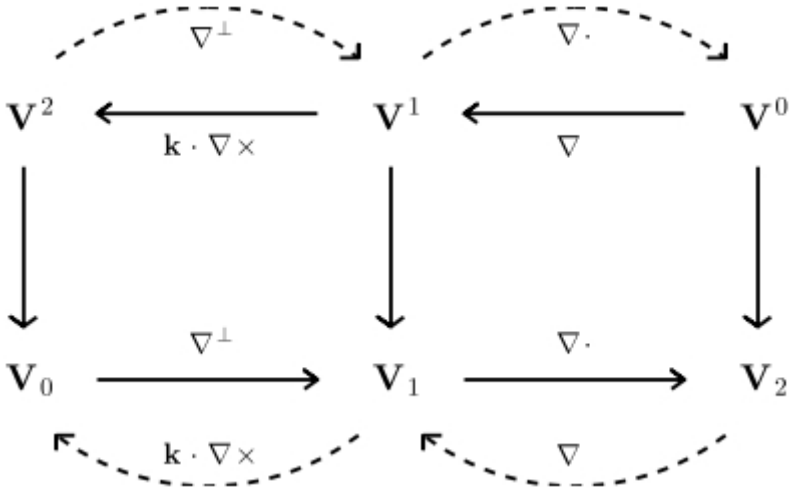

Figure 5.1: The relationship between function spaces. Primal spaces are on the bottom row, and dual spaces are on the top row. Sub- and super-scripts indicate the degrees of freedom of each space.

mass matrices or operators are symmetric or antisymmetric.

Geostrophic balance is a fundamental principle of some of the asymptotic regimes that we look at. For a model to capture geostrophic balance it must be capable of supporting steady geostrophic modes. This is achieved by ensuring that the result when taking the divergence of the averaged velocities used in construction of the Coriolis terms (W, see below) is the same as the result when calculating the velocity divergence and averaging. This implies that non-divergent velocity fields are balanced by a geopotential field and there is no contribution to vorticity from the Coriolis term. Thus geostrophic balance is supported. For a fuller explanation see [46]. A consequence of the ability to support steady geostrophic modes is that the linear potential vorticity is also steady.

This model has been in development for some time. Testing so far has shown the scheme to be stable for large C_{gw} and C_{adv} up to 0.75 for centred time integration, with only minor off-centering required to achieve stability for C_{adv} up to 1 [45]. At high resolution, the model produces comparable results to ENDGame, though there is evidence of grid imprinting at coarser resolutions, with some grid choices more susceptible than others.

As with all models, the ability to respect balance is important and so asymptotic limit tests will be an important measure of this models accuracy. Previous works (e.g. Thuburn, Cotter and Dubos [45]) have looked at the barotropic vorticity case and used divergence as a measure of imbalance. This work found that imbalance due to the initial perturbation persisted for the duration of the run, and so the integration used a fully off-centred scheme for an initial period before swit-

ching to a centred one. Using this method it was found that imbalance remains extremely small in both the finite volume and finite element versions of the model, giving similar results to ENDGame.

The diagnostics we use are a stricter measure of imbalance, so comparison to the results from the finite difference model in Chapter 4 will be interesting in light of the above results. As this model is a prototype for possible future Met Office models the results could help to inform future improvements.

5.1.2 Discrete shallow water equations

As the resulting balance equations are perhaps in a less intuitive form compared to the finite difference model, the discrete versions are now given. For full details see the paper by Thuburn and Cotter [47].

The nonlinear shallow water equations may be written as

$$\dot{\Phi} + D_2 F = 0, \tag{5.1.1}$$

$$M\dot{U} + MQ^{\perp} + \bar{D}_1 L(\Phi_T + K) = 0. \tag{5.1.2}$$

$\bar{D}_1$ and D_2 are matrices that perform the ∇ operation on dual spaces and the $\nabla\cdot$ operation on primal spaces, respectively; F, Q, K, Φ and U are vectors of coefficients representing the mass flux, PV flux, kinetic energy per unit mass, geopotential and velocity; L and M are mass matrices; Φ_T is the total geopotential (depth + topography); and the $\dot{X}$ and $X^{\perp}$ are the standard operators for time derivative and orthogonal complement.

5.1.3 Discrete diagnostics

We can now define the divergence as the cell integral, divided by the cell area A to give a value for the divergence at cell centres:

$$\nabla \cdot \mathbf{u} \overset{\text{da}}{\equiv} D_2 U/A. \tag{5.1.3}$$

$X \overset{\text{da}}{\equiv} Y$ is used here to indicate to indicate that Y is the discrete approximation to X.

Similarly, the absolute vorticity as a dual cell integral is

$$\zeta \overset{\text{da}}{\equiv} \bar{D}_2 U + \bar{f}, \tag{5.1.4}$$

where $\bar{D}_2$ is a matrix that performs the $\mathbf{k}\cdot\nabla\times$ operation on dual cells and $\bar{f}$ is used here to indicate the coriolis parameter integrated across the dual cells.

The geostrophic imbalance diagnostic requires use of the W operator, which constructs an edge integral on a dual cell from the edge integral of the primal cell perpendicular to it:

$$f\mathbf{k}\times\mathbf{u}+\nabla\phi \overset{\text{da}}{\equiv} -fWU+\bar{D}_1L\Phi. \tag{5.1.5}$$

Finally, the higher order imbalance may be expressed as:

$$\zeta^2\mathbf{u}-\zeta\mathbf{k}\times\nabla(\phi+\frac{1}{2}|\mathbf{u}|^2)-\nabla(\nabla\cdot(\mathbf{u}\phi)) \overset{\text{da}}{\equiv} M\Xi^2U+M\Xi M^{-1}WM^{-1}\bar{D}_1L(\Phi+K)-\bar{D}_1LD_2F, \tag{5.1.6}$$

where Ξ is the absolute vorticity taken from cell area integrals to a cell edge integral (note: we are using lowercase ϕ here to differentiate it from the coefficient vector form Φ).

To calculate the higher order balance we needed to make use of some derivative functions that must be defined in a weak sense due to the discontinuity of fields. These functions are indicated in figure 5.1 as those lying on dotted lines. The gradient on primal spaces is approximated by $\nabla\psi \overset{\text{da}}{\equiv} M^{-1}\bar{D}_1L\Psi$ and the $\perp$ operator on dual spaces is approximated by $U^{\perp}=-M^{-1}WU$. The higher order balance diagnostic we used (5.1.6) was then calculated by discretising the left-hand side term-by-term. It should be pointed out that the right-hand side is of different dimension to the left, being integrals over cell edges. To aid comparison, the diagnostics we used in our tests have been divided by cell edge length in order to make them dimensionally consistent.

5.1.4 Grid

The finite element model is designed to function on arbitrary grids, so we have decided to use two different choices of grid to compare. During the model's development, a hexagonal-icosahedral and a cubed-sphere grid have been used as standard test cases and so we have used them here too. Figure 5.2 shows two low resolution examples of these grids to give an idea of how the primal and dual grids align with each other.

To better compare the results from the two models, we attempt to match the degrees of freedom (cells plus edges) between the grids. On the cubed sphere grid this corresponds to a grid with 221184 faces, and on the Hexagonal-Icosohedral it corresponds to a grid with 163842 faces. These values are also chosen so as to closely match the degrees of freedom of the 640x320 grid we used with the ENDGame model.

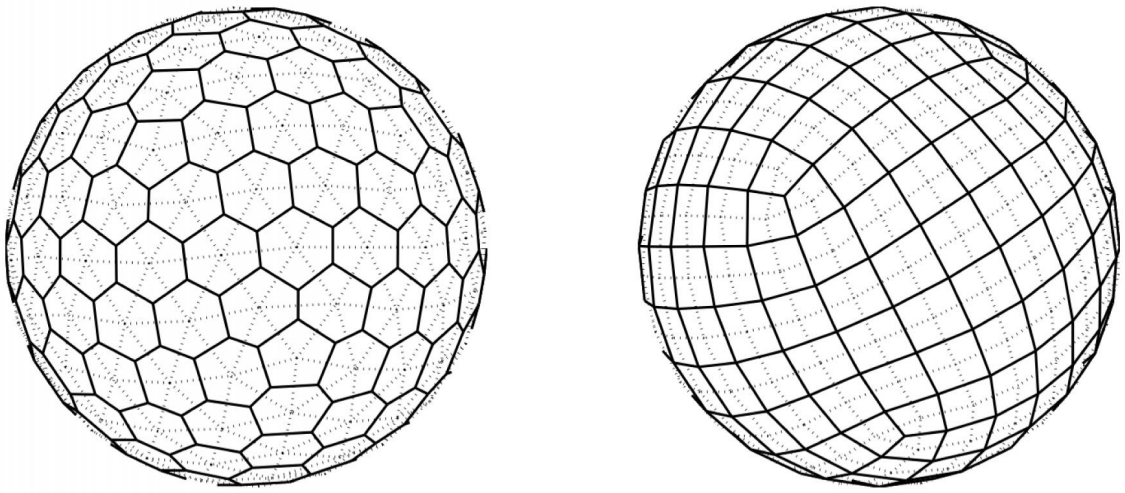

Figure 5.2: Left: a hexagonal-icosahedral mesh with 162 cells and 642 degrees of freedom. Right: a cubed-sphere mesh with 216 cells and 648 degrees of freedom. Continuous lines are primal grid edges, dotted lines are dual grid edges.

Courant number(s)

Ideally we would have scaled our time step to match the typical wave Courant number $\frac{\sqrt{\Phi}\Delta t}{\Delta x} \approx 0.5$ used in the ENDGame runs. Unfortunately, the finite element model is unstable if the advective number violates the CFL condition $\frac{\mathbf{u}\Delta t}{\Delta x} \leq 1$. As C_{adv} tends to be maximised on the dual grid (due to smaller cells), this leads to $C_{adv} < 1$ on the primal grid. Indeed in the cubed sphere grid when $\varepsilon = 1$ the CFL condition seemed to be violated when time steps longer than $182s$ were taken. This may be the result of the strongest region of the jet coinciding with a 'corner' of the cubed sphere grid, thus maximising u while minimising Δx. This resulted in the wave courant number for the cubed sphere grid having a maximum value of around $C_{gw} \approx 0.4$. In order to more directly compare the two grids, we chose to use the same size of time step in the hexagonal-icosahedral grid, which results in a maximum value of around $C_{gw} \approx 0.2$.

5.2 Results

As in the ENDGame model, the semigeostrophic time series (e.g. figure 5.3) show regional behaviour becoming more distinct as we approach the limit - almost all the divergence and geostrophic and higher-order imbalance is localised

in the jet region. The general shape of the series is also similar to that of ENDGame, with the exception that a peak occuring in the higher-order imbalance and divergence tendency diagnostics occurs later here, around day 10 rather than day 7. This is only noticeable in the $\varepsilon = 1$ case though, so the simulations are not particularly close to the semigeostrophic limit.

The same difference in peak time is true of the quasigeostrophic limit. Here we also note that the FEM does not demonstrate the same difficulties in interpolating a time derivative over the pole, and so we do not see the spurious scaling in the diagnostic at the north pole. It should be noted that the shape of the geostrophic imbalance, higher-order imbalance and divergence diagnostics show what look like two peaks where the imbalance peaks then deacreases and then peaks again. On both grids these peaks seem to occur around 2 days apart, which can be most clearly seen in the higher-order imbalance (figure 5.4). However, the scale seems to vary such that the first peak is larger in the cubed sphere grid and the second is larger in the hexagonal-icosahedral grid.

The barotropic vorticity and strong rotation limits show an intriguing difference to the ENDGame model. While the general shape of the time series are similar as in the other limits, we do not see the flow developing later as we approach the BV limit. It is possible that instability in this regime is being triggered by grid imprinting, which would make the scaling of the forcing irrelevant. However, investigation of the PV field shows no indication of grid imprinting, and the PV perturbation in the BV limit looks very similar to the PV perturbation in the Quasigeostrophic case with the same Froude number, so for the time this behaviour remains unexplained. Both the BV and SR limit show the flow reaching the "developed" state we use for our convergence diagnostics at day 5. This is compared to the same state being reached at around day 7 in the ENDGame model. As the development of the strong rotation limit (e.g. figure 5.5) effectively shows the flow moving from the BV limit to the QG limit, it is probably that this difference in behaviour is a consequence of the BV limit alone. As this difference seems to be a factor of the model itself rather than the diagnostics we are using, examination of this behaviour falls out of the scope of this thesis. It might warrant further investigation at a later date.

In the divergence tendency diagnostic of each figure we see 'bumps' of imbalance at a fairly regular intervals. As they are most evident in the north-of-jet region, it seems probable that this is caused by gravity waves moving over the pole and hitting other gravity waves. This is supported by the fact that these bumps seem to occur more frequently as the gravity wave speed increases. The barotropic limit figure 5.6 shows this most clearly. The strong rotation figures seem to have the same frequency of bumps in each run, which is logical as the

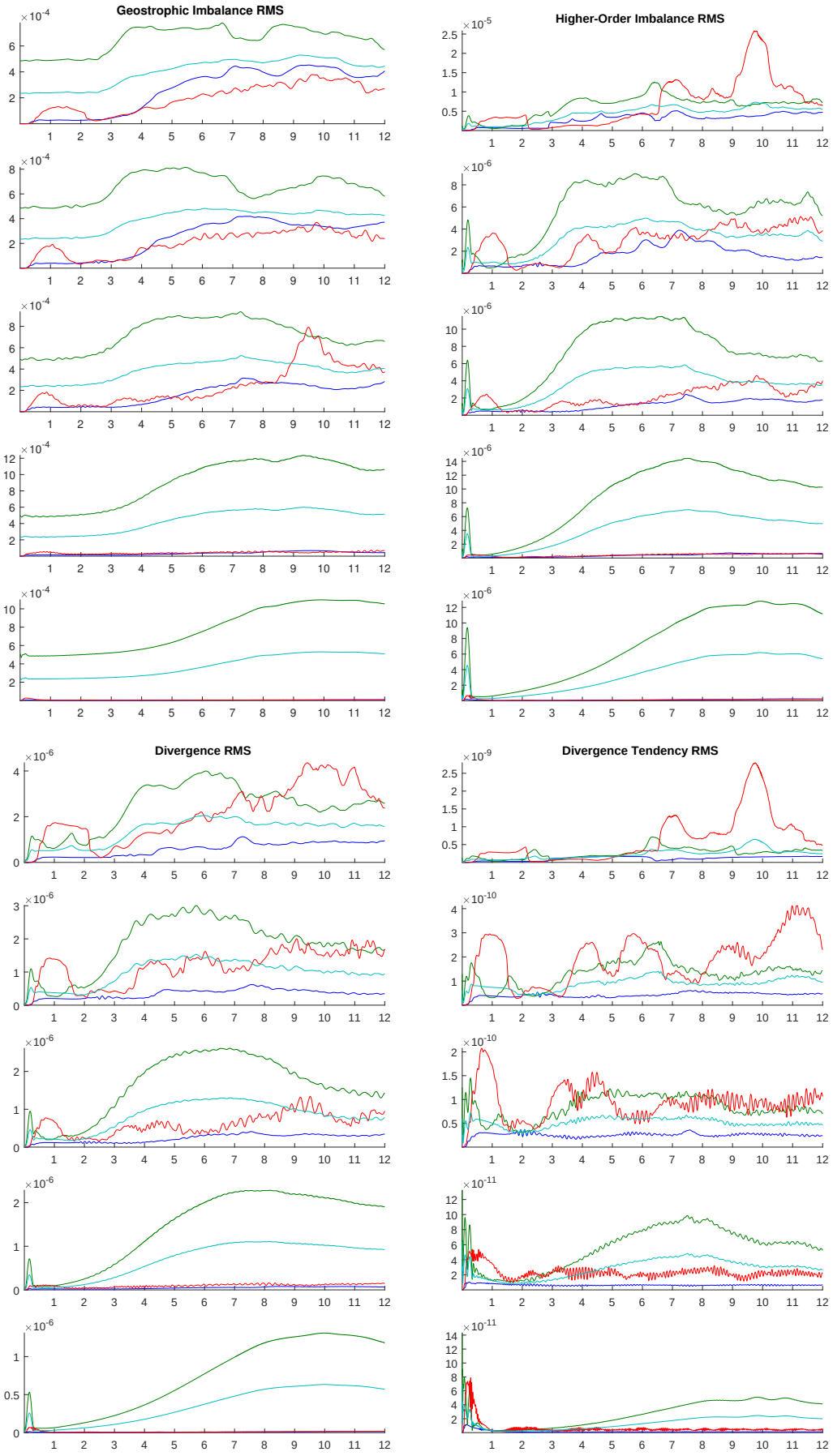

Figure 5.3: Time series of diagnostic quantities approaching the semigeostrophic limit in the Hexagonal-Icosahedral FEM. For each subfigure the small parameter ε is, from top to bottom, = 1, 0.5, 0.25, 0.125, 0.0625.

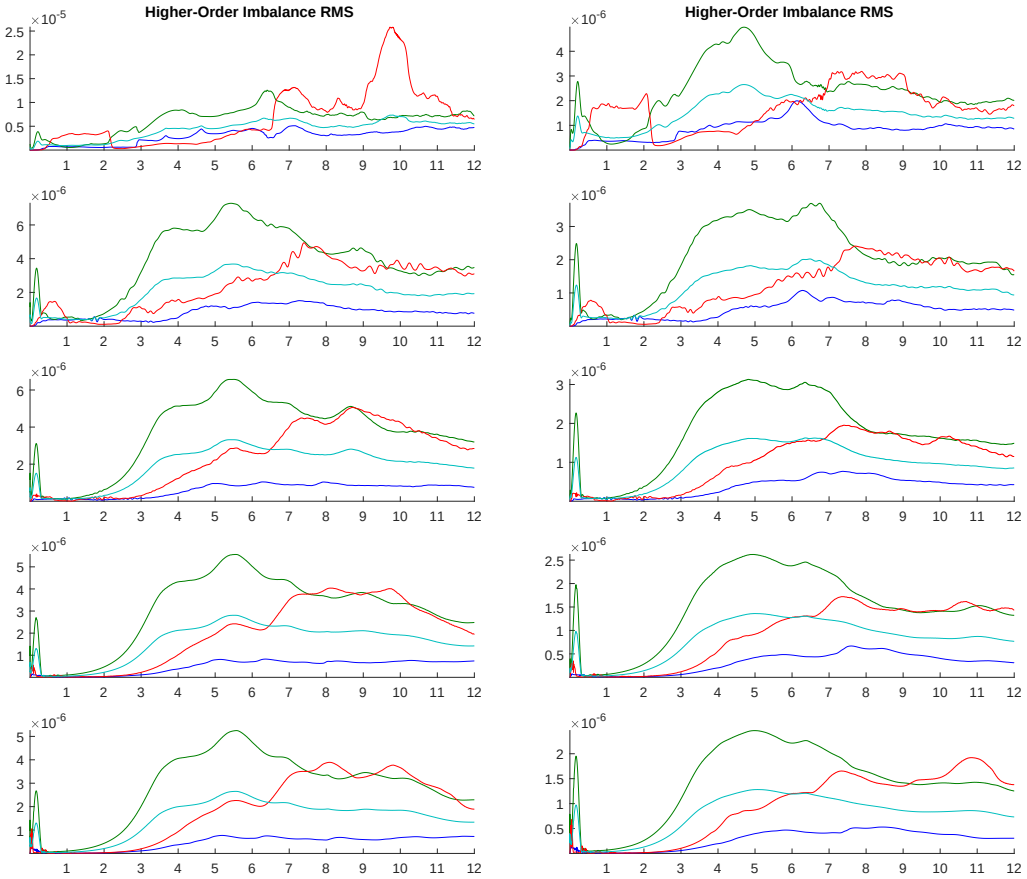

Figure 5.4: Time series of diagnostic quantities approaching the quasigeostrophic limit in the Hexagonal-Icosahedral (left) and Cubed Sphere (right) FEM. For each subfigure the small parameter ε is, from top to bottom, = 1, 0.5, 0.25, 0.125, 0.0625. Both grids show two peaks in the in-jet diagnostic at around days 4 and 6, though the relative strength of this peak varies in each grid.

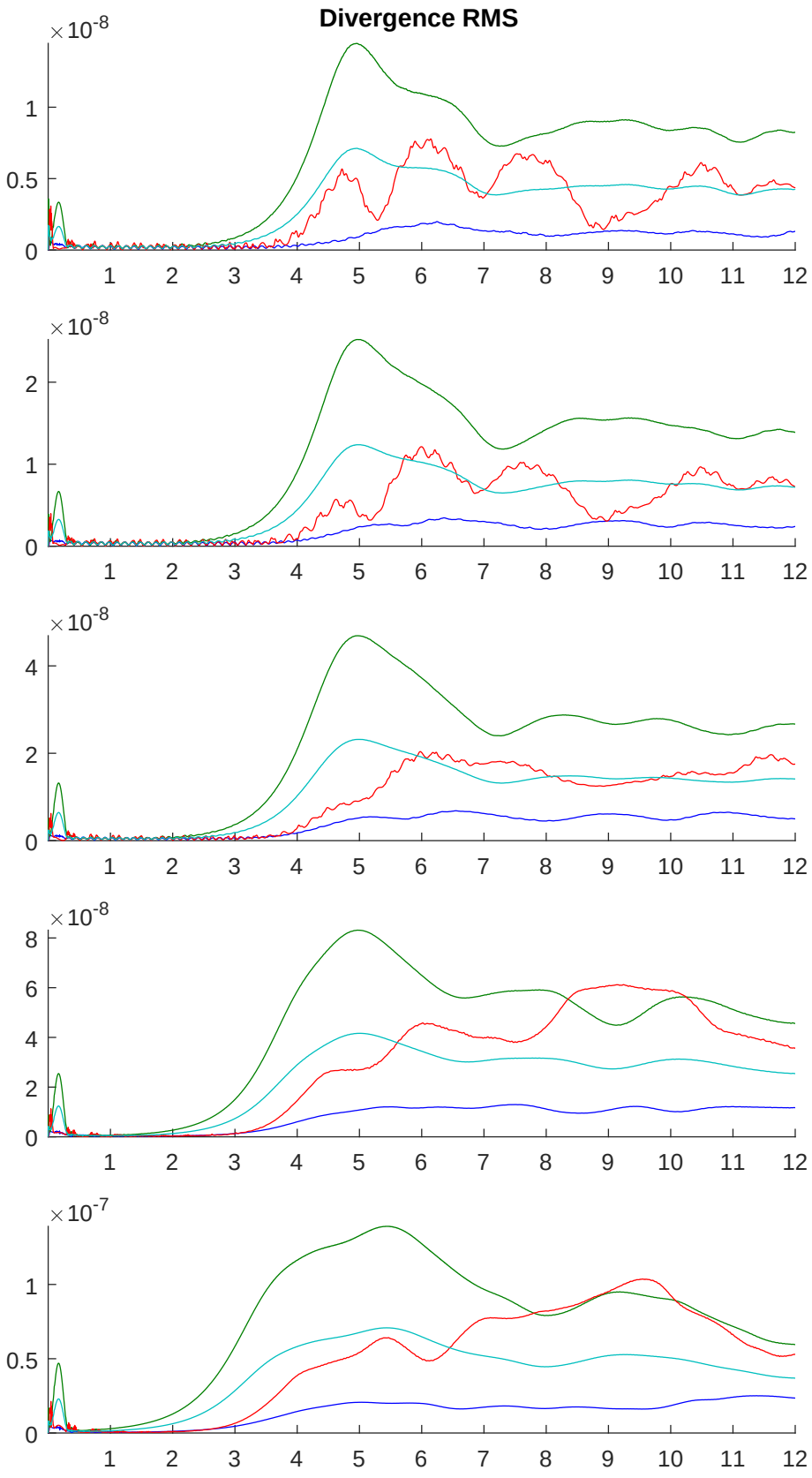

Figure 5.5: Time series of divergence approaching the strong rotation limit in the Hexagonal-Icosahedral FEM. For each subfigure the small parameter ε is, from top to bottom, = 1, 0.5, 0.25, 0.125, 0.0625. The peak we use as the mark for the flow reaching its developed state is clearly visible at day 5. Note that unlike ENDGame the time this peak occurs does not vary.

gravity wave speed does not change in these runs.

Inspection shows that the two FE grids used here produce almost identical results to the extent that the time series are almost indistinguishable in their key features. The strong rotation and barotropic vorticity tests all show a peak of imbalance at day 5, and the semigeostophic tests show their peak developing later as the limit is approached (around day 5 at $\varepsilon = 0.5$ to around day 10 at $\varepsilon = 0.0625$). The runs also show very similar behaviour in north-of-jet activity that we suspect to be caused by gravity waves, with the frequency of the bumps indicating that gravity wave speeds in both grids are very close. This is encouraging as it shows that the choice of grid has limited impact upon the results, at least in these examples.

Full results may be found in Appendix B.

5.2.1 Convergence diagnostics

Looking at the convergence diagrams for divergence and geostrophic imbalance in figure 5.11, both of the grids using the Eulerian model produce very similar results to ENDGame, with some minor differences in scale. Also similar to the ENDGame model, it appears that only the two runs closest to the limit are matching the correct scaling in the SG regime. This suggests that this behaviour is a robust signal and could merit further investigation.

The higher-order imbalance figure looks rather different in the Eulerian model. Here the pattern of the SG diagnostics look very similar to those of the SG geostrophic imbalance, and we do not seem the same kind of spurious scaling as each limit was approached. Both QG and SR regimes scale at a consistent rate, though in the QG case the imbalance grows smaller and in the SR case it is increasing, rather than remaining the same scale as we would expect. It is also notable that in the ENDGame model the QG and SR higher-order imbalance was almost identical, which is not the case here. This implies that the scale of the diagnostic has a dependence on the Froude number that was not present in the ENDGame tests.

This cannot be exactly the same scaling that was suggested in the semi-Lagrangian model, as we believe that relies on errors in the time step interpolation process, but perhaps there is a similar root cause - the derivation of the analytic diagnostic might not be consistent with approximations used when defining the discrete version of the diagnostic.

It is worth noting that the higher order imbalance diagnostic is of a very different scale in the ENDGame and FE models. At $\varepsilon = 0.0625$ the typical scale in the ENDGame model is $\sim 10^{-7}$, whereas it is $\sim 10^{-5}$ in the FE models. This

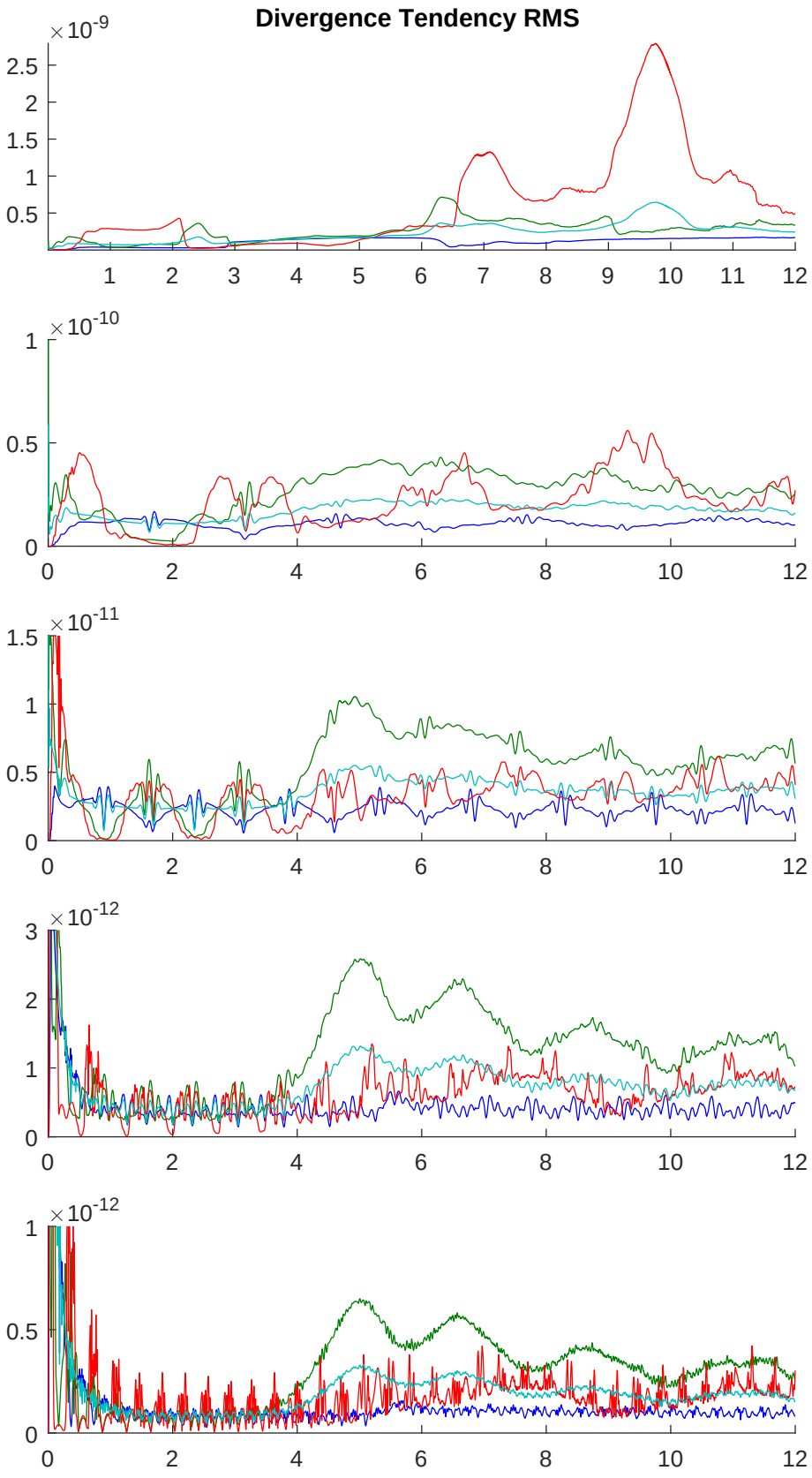

Figure 5.6: Time series of divergence tendency approaching the barotropic vorticity limit in the Hexagonal-Icosahedral FEM. For each subfigure the small parameter ε is, from top to bottom, = 1, 0.5, 0.25, 0.125, 0.0625. The frequency of bursts of imbalance north of the jet shows a correlation with the Froude number and thus fluid depth/gravity wave speed.

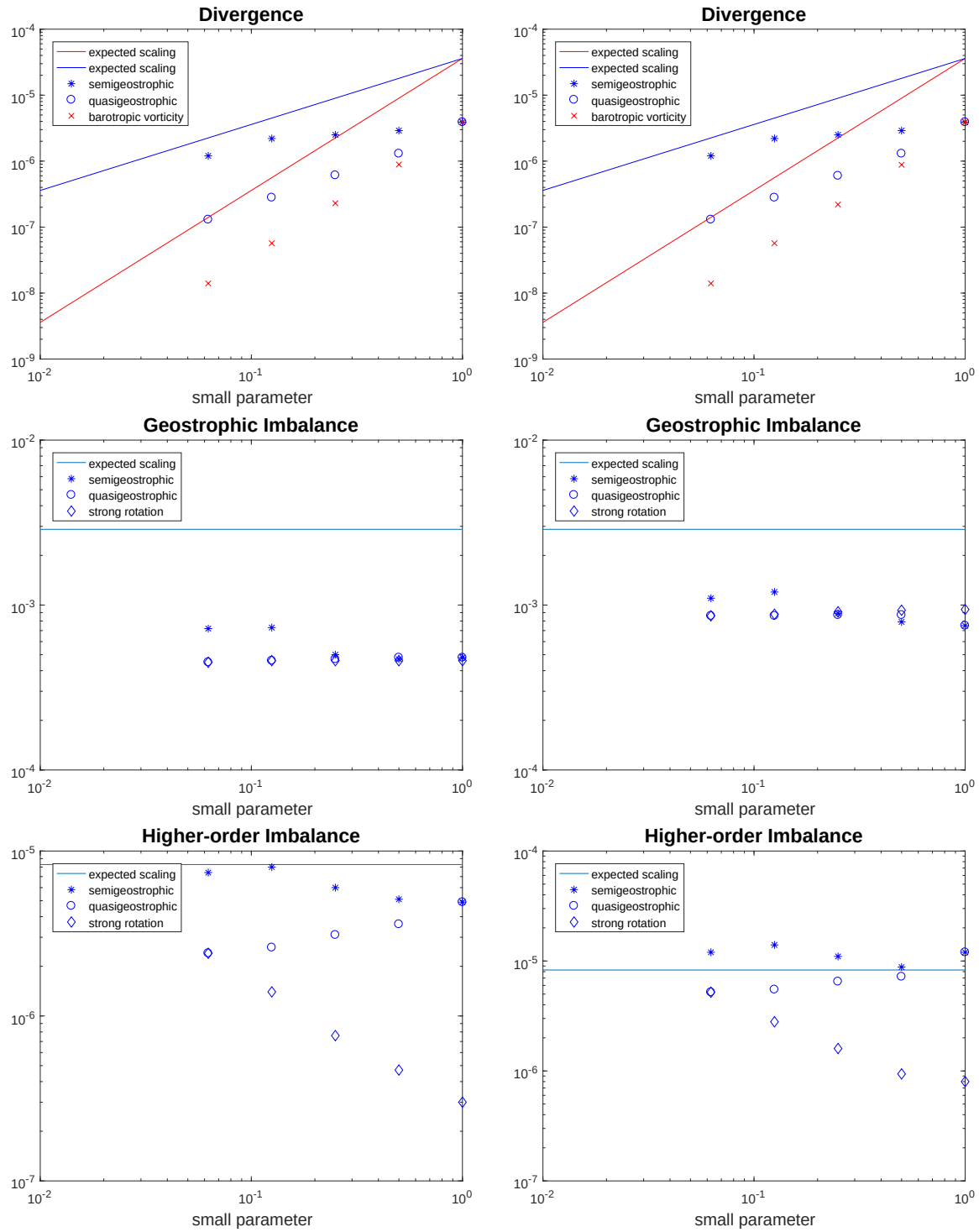

Figure 5.7: Convergence diagrams for each in-jet diagnostic. Left hand side is the cubed sphere grid, right hand side is the hexagonal-icosahedral grid. Red points are expected to lie parallel to the red lines, and blue points are expected to lie parallel to the blue lines. As in the ENDGame model, results for the lower-order diagnostics look good while there is strange scaling in the higher-order diagnostic.

discrepancy comes from the $\bar{D}_1 L D_2 F$ term in the Finite Element formulation. It is notably smaller than the equivalent term in ENDGame whereas the other terms are very close in scale. Additionally there is also less cancellation of terms, leaving the higher order imbalance equation significantly larger in the FE models.

5.2.2 Ideas

It is possible that the unexpected scaling in the finite element model is related to something other than errors amplifying as the asymptotic limits are approached (as we suspect in ENDGame). Looking at the convergence diagram for the higher order balance it is especially curious that the QG and SR do not lie on top of each other, as they did in the other diagnostic quantities and also in the higher-order diagnostic in ENDGame. This implies a dependence on the Froude number (or related quantity) that should not exist. It seems noteworthy that the high-order imbalance diagnostic retains its shape in the SG case, and we might reasonably expect the other two regimes to follow suit. The form of the SG result also suggests that is not an error in the definition of the high-order imbalance.

The suggestion is that while the source of the errors may differ from those in ENDGame, we might be seeing another inconsistency between the diagnostic and the discrete implementation of the diagnostic - a strong candidate would be the weak definitions of certain functions that was mentioned briefly in section 5.1.1.

5.3 Summary

As in the previous chapter, we have given an overview of the model we use. More detail has been gone into here, as it is arguably less intuitive than the finite difference model.

Diagnostic time series and convergence diagrams are presented and discussed. We discover that the flow does not show the same late development in the barotropic vorticity limit that we saw in the ENDGame diagnostics. We have found no compelling evidence that this is a result of grid imprinting destabilising the flow.

The time series of both FE grids demonstrate a number of features consistent with the ENDGame results, and the lower-order diagnostics are also very similar. While we would not expect to see the same spurious scaling in the higher-order imbalance diagnostics that we did in ENDGame, we nonetheless find some curious behaviour. This is most notable in the strong rotation regime, where the

results imply a dependence on Froude Number that did not appear to be present in the ENDGame runs. In addition, the third term of the FE diagnostic does not appear to cancel other terms as strongly as in the ENDGame model, meaning that the higher order imbalance diagnostics are around 2 orders larger in the FE model.

As before, some ideas have been given regarding failure for diagnostics to converge to the expected scaling. It is suspected that the weak definitions used in the model might mean that we do not achieve second-order convergence to the asymptotic limit solutions.

Chapter 6

Part II summary

We have demonstrated an application of the theory of using asymptotic limit solutions to assess the effectiveness of two shallow water models. We derived both lower- and higher-order diagnostics that we expected to converge to a given scaling as we approached the strong rotation, barotropic vorticity, quasigeostrophic and semigeostrophic limits of the shallow water equations.

The lower-order diagnostics (divergence and geostrophic imbalance) perform well in both models, converging to the expected scaling. This compares well to previous works that looked at similar diagnostic. However, there were a number of problems with the higher-order diagnostics in both models. ENDGame appeared to incorrectly interpolate $\frac{D}{Dt}\nabla\Phi'$ over the pole, leading to large errors in the north-of-jet diagnostic. Fortunately these errors did not affect the rest of the results. This apparent shortcoming of ENDGame is surprising - checking the code did not find any error in the way this term was calculated, which might warrant further investigation.

Both models' higher order diagnostics failed to converge to the expected scaling, though the FE model's results perhaps look slightly better. Though the exact mechanics of these errors are different, we believe that the fundamental problem was a failure of the discrete diagnostics we used to respect approximations used when deriving them. These errors are small enough to not matter to the lower-order diagnostics but they become significant when we look for higher-order versions. This view is supported by examination of the RHS of the higher order imbalance equations, which do not demonstrate the same spurious behaviour.

It is unfortunate that such errors were present in both models, and that there was not time during this thesis to properly examine and correct the shortcomings of the method we employed. However, the results we have obtained should be instructive to those looking to employ similar methods themselves.

On a brighter note, the results from the lower-order diagnostic have shown that the finite element model is capable of producing very similar similar results

to those of ENDGame, regardless of the grid used. We also showed how similar results using a large time step in ENDGame were to those using short time steps.

Part III

Subgrid interactions

Chapter 7

Turbulent cascade theory

In the previous parts of this thesis we discussed a general method of developing balance equations whose scaling should match that of a model as appropriate asymptotic limits are approached. Due to the separation in the timescales, these equations tended to represent the balanced part of the flow. However, we did discuss the topic of spontaneous gravity wave generation. While the exact mechanisms are a matter of some debate, this process can fairly be described as unbalanced motion created from balanced motion.

This process is one example of the ways in which balanced and unbalanced parts of a flow can interact and exchange energy. It is distinct from the notion of cascades in that it is different forms of motion interacting, possibly across different time scales, rather than a 'chain' of increasingly small spatial scales. Energy should be conserved, so it will be useful to see where the energy goes within a cascade. This could inform the development of future subgrid models.

Whatever the mechanism, energy transfers between scales are a valuable feature to accurately model. This chapter will give a brief introduction to the theory behind energy cascades, expanding on the introduction section and also highlighting some specific consequences of the theory regarding timescales. The rest of part III will involve the development and application of a method of measuring energy transfers from scale to scale.

7.1 Introduction to Kolmogorov theory

Turbulence can be described as chaotic fluid motion with high nonlinearity and eddies at a range of scales. This nonlinearity implies interactions and a transfer of energy between different scales of motion. If there is a systematic transfer of energy between smaller and smaller eddies (or vice versa), this is known as a

cascade. Kolmogorov [24] developed a theory for the form of the energy spectrum in a turbulent flow. The following is a brief introduction to the arguments regarding energy and enstrophy cascades.

If we assume a three dimensional flow is statistically steady and isotropic, we can consider a large-scale stirring of the fluid and its corresponding energy input. If this energy is cascaded to small scales and dissipated by viscosity, there should be some intermediate range of scales known as the *inertial range* where neither the forcing nor the dissipation are important. If we are in a statistically steady state then the rate of energy input, ε, must be equal to the dissipation rate and thus also to the transfer rate between the input and dissipation scales.

We consider the spectral energy density

$$\frac{1}{V}\int_V \hat{E}dV = \int \mathcal{E}dk,\ k = |\mathbf{k}|, \tag{7.1.1}$$

where $\hat{E}$ is the energy density per unit mass, V is the volume of the fluid and $\mathcal{E}$ is the energy spectrum. We know the dimension of $\hat{E}$ is U^2 and the dimension of k is $1/L$, and we can define U in terms of length and time scales as $U = L/T$. Hence the dimension of $\mathcal{E}$ is $[\mathcal{E}(k)] = \hat{E}L = L^3/T^2$ and the dimension of ε is $[\varepsilon] = \hat{E}/T = U^2/T^2 = L^2/T^3$. If we wish to construct $\mathcal{E}$ only as a function of ε and k then we require a factor of $\varepsilon^{2/3}$ as k has no time dependence, and a factor of $k^{-5/3}$ is needed for balance. Thus

$$\mathcal{E}(k) = \mathcal{K}_\varepsilon \varepsilon^{2/3} k^{-5/3}, \tag{7.1.2}$$

where $\mathcal{K}_\varepsilon$ is a dimensionless constant of order one.

7.1.1 Two-dimensional turbulence

In two dimensions another conserved quantity known as the enstrophy is cascaded downscale, and energy is cascaded to large scales instead. By a similar argument, we still expect to see the $k^{-5/3}$ spectrum, only now at scales larger than the forcing scale. In addition we expect to see an enstrophy transfer, η, whose dimensions are

$$[\eta] = U^2/L^2T = 1/T^3. \tag{7.1.3}$$

Following the derivation of (7.1.2) we can obtain the downscale energy spectrum

$$\mathcal{E}(k) = \mathcal{K}_\eta \eta^{2/3} k^{-3}, \tag{7.1.4}$$

where $\mathcal{K}_\eta$ is another dimensionless constant of order one.

Polvani, McWilliams, Spall and Ford [33] have shown that in two-dimensional shallow water flows the upscale energy cascade is limited by the Rossby radius of

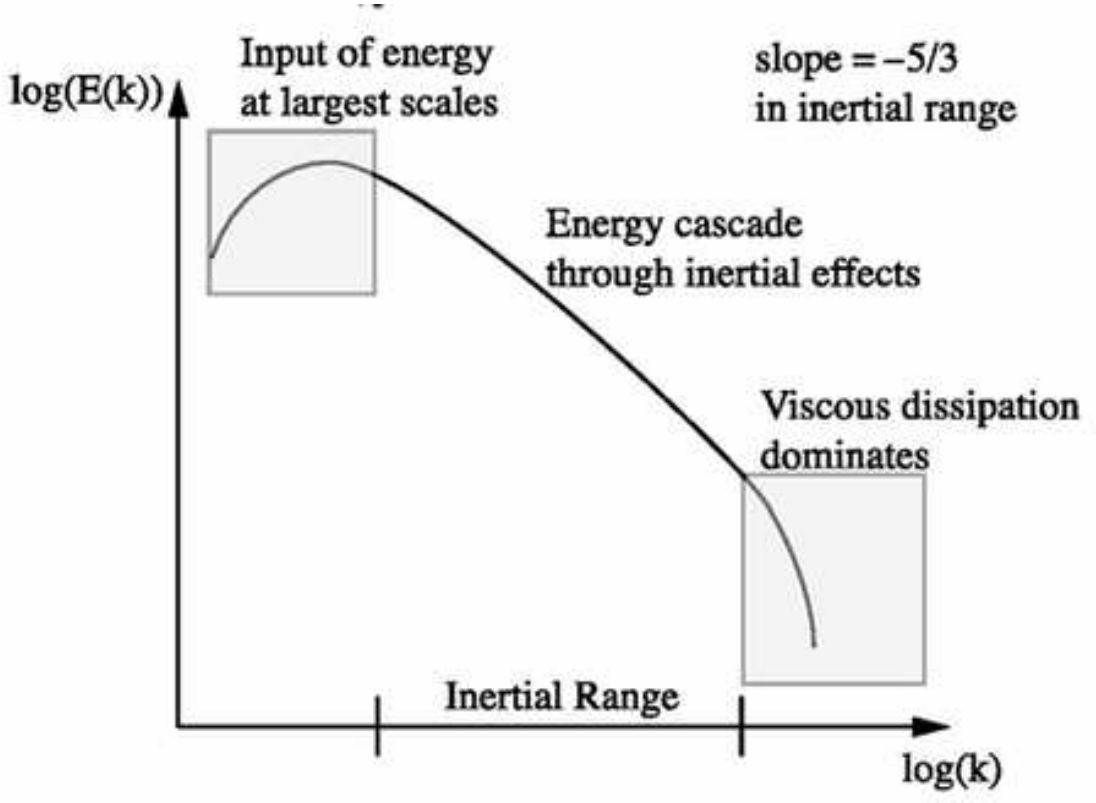

Figure 7.1: A depiction of the observed energy cascade, taken from Berselli, Iliescu and Layton [4].

deformation: As vortices approach L_R they interact more weakly, thus slowing the rate of energy exchange and effectively restricting the inertial range. The upscale energy cascade may also be halted by large-scale dissipation.

7.2 The cascade timescale

When choosing the integration time of our simulations, it would be helpful if we had some idea of the timescales in the flow. An estimate of the time it takes for energy to move from the forcing scale to the dissipation scale will give us a good idea of how long it will take the energy added through forcing to populate the entire spectrum, and thus suggest a minimum time to run the model for.

We can calculate a speed at which energy cascades between wavenumbers by solving dk/dt, equal to the input rate (ε or η) divided by the 'capacity' or spectrum ($\mathcal{E}$) [34]. So for the upscale energy cascade we have

$$\frac{dk}{dt}_{\varepsilon} = \frac{\varepsilon}{\mathcal{E}_{\varepsilon}} = \frac{\varepsilon^{1/3}k^{5/3}}{\mathcal{K}_{\varepsilon}}, \tag{7.2.1}$$

and for the downscale enstrophy cascade

$$\frac{dk}{dt}_{\eta} = \frac{\eta}{\mathcal{E}_{\eta}} = \frac{\eta^{1/3}k^{3}}{\mathcal{K}_{\eta}}. \tag{7.2.2}$$

From this, we can determine the time it takes for a quantity to move between a lower wavenumber, k_a, and a higher one, k_b.

This gives us the following timescales:

$$\tau_\varepsilon = \frac{3}{2}\mathcal{K}_\varepsilon \varepsilon^{-1/3}(k_b^{-2/3} - k_a^{-2/3}), \tag{7.2.3}$$

$$\tau_\eta = \frac{1}{2}\mathcal{K}_\eta \eta^{-1/3}(k_a^{-2} - k_b^{-2}). \tag{7.2.4}$$

We can also use this method to optimise the computation time by spinning up the simulation at a lower resolution and steadily increasing the resolution until the energy has filled the spectrum. Details are given in chapter 8.

7.3 Shock formation

Test cases involving forcing of the unbalanced flow resulted in the formation of shocks. As we only wish to observe the effects of shocks in one of our regimes, it is beneficial to classify the conditions for shock formation in our system so that we may avoid them in the gravity-wave-dominated case.

Using the method of characteristics, we may find constants along these characteristics known as Riemann invariants. Shock formation happens where these characteristics cross. The method for obtaining these crossing points is as follows:

It is sufficient to view the shallow water equations as 2D in the (x, t) plane and hence write them as

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(\Phi + \frac{1}{2}|\mathbf{u}|^2) = 0, \tag{7.3.1}$$

$$\frac{\partial \Phi}{\partial t} + \frac{\partial u\Phi}{\partial x} = 0. \tag{7.3.2}$$

We can write these equations in terms of the wave speed $c = \sqrt{\Phi}$:

$$\frac{\partial u}{\partial t} + 2c\frac{\partial c}{\partial x} + u\frac{\partial u}{\partial x} = 0, \tag{7.3.3}$$

$$2\frac{\partial c}{\partial t} + c\frac{\partial u}{\partial x} + 2u\frac{\partial c}{\partial x} = 0. \tag{7.3.4}$$

Then, we can take 7.3.3 $\pm$ 7.3.4 to give

$$\frac{\partial}{\partial t}(u \pm 2c) + (u \pm c)\frac{\partial}{\partial x}(u \pm 2c) = 0. \tag{7.3.5}$$

This gives characteristics $\frac{\mathrm{d}x_\pm}{\mathrm{d}t} = u \pm c$, with corresponding Riemann invariants $r_\pm = u \pm 2c$ if Φ is constant.

If variations in Φ are small compared to the mean geopotential we may take it to be constant here. Hence as $u \pm 2c$ is invariant along each $\pm$ characteristic, then

both u and c must be constant along the characteristics, so the characteristics themselves are straight lines. As we are interested in the time of first shock formation, the characteristic that matters is the one that minimises $|u \pm c|$. As c is positive by definition, and u can be either positive or negative, this suggests that a regime where $Fr \approx 1$ is the most likely to produce shocks. Further, these 'shallow' characteristics will intersect the soonest with 'steep' characteristics in close proximity thus the maximum divergence will also be a good predictor of shock formation (this should be intuitive).

It is logical that the first intersection will be between two adjacent points in the domain (x_1 and x_2, say). Thus the time of first shock formation, t_{shock} may be calculated as follows, considering right-characteristics:

$$\begin{aligned}
(u_1 + c_1)t_{shock} + x_1 &= (u_2 + c_2)t_{shock} + x_2 \\
\implies x_1 - x_2 &= (u_2 - u_1 + c_2 - c_1)t_{shock} \\
\implies t_{shock} &= \frac{x_1 - x_2}{u_2 - u_1 + c_2 - c_1} \\
\implies t_{shock} &\approx \frac{1}{\frac{\partial}{\partial x}(u + c)}
\end{aligned}$$

To arrive at a regime where we avoid shocks, we should maximise t_{shock}, or minimise

$$\frac{\partial}{\partial x}(u + c) = \frac{\partial u}{\partial x} + \frac{\partial c}{\partial x} = \delta + \frac{1}{2\sqrt{\Phi}}\frac{\partial \Phi}{\partial x}.$$

From part 2 we know that δ scales like Ro or Fr^2 depending on the regime, and $\frac{1}{2\sqrt{\Phi}}\frac{\partial \Phi}{\partial x} \sim O(\frac{Fr}{Ro})$. This suggests that for t_{shock} to be large, we require that Fr is small, and smaller than Ro.

Chapter 8

Wave-turbulence interactions in spectral space

Accurate numerical simulations must replicate transfers between scales, including those scales too small to be directly resolved. Reproduction of energy and enstrophy cascades have been investigated in the case of two-dimensional turbulence by Thuburn, Kent and Wood [48]. The aim is to perform a similar investigation in the case featuring both turbulence and waves.

Where the previous work looked at energy and enstrophy, we will decompose the energy into different components in order to investigate them individually. Rather than assessing different subgrid models, we will instead use this method to look more closely at the dynamics below the subgrid scale, providing information on how to construct these subgrid models.

We therefore seek to express the energy in such a way that it can be conveniently separated into different components: vortical, divergent, potential, balanced and unbalanced.

8.1 Alternate forms of the shallow water equations

Different forms of the shallow water equations make it easier to extract different information from the resulting energy equations. Similar to previous parts of this thesis, we use an f-plane approximation. The domain is also taken to be doubly periodic.

We may express the shallow water equations as:

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{k} \times \mathbf{Q} + \nabla(\Phi + \mathbf{K}) = 0, \tag{8.1.1}$$

$$\frac{\partial \Phi}{\partial t} + \nabla \cdot \mathbf{F} = 0, \tag{8.1.2}$$

where the vorticity flux $\mathbf{Q} = \mathbf{u}(f + \xi)$, the kinetic energy per unit mass $\mathbf{K} = \frac{1}{2}|\mathbf{u}|^2$ and the mass flux $\mathbf{F} = \mathbf{u}\Phi$.

The momentum equation can be split into vortical (ξ) and divergent (δ) parts :

$$\frac{\partial \xi}{\partial t} + \nabla \cdot \mathbf{Q} = 0, \tag{8.1.3}$$

$$\frac{\partial \delta}{\partial t} - \mathbf{k} \cdot \nabla \times \mathbf{Q} + \nabla^2(\Phi + \mathbf{K}) = 0. \tag{8.1.4}$$

The velocity field may also be decomposed into an irrotational and a solenoidal part by expressing it as the sum of a velocity potential, χ, and a stream function, ψ, such that $\mathbf{u} = \nabla\chi + \nabla^{\perp}\psi$:

$$\frac{\partial}{\partial t}\nabla^2\psi + \nabla \cdot \mathbf{Q} = 0, \tag{8.1.5}$$

$$\frac{\partial}{\partial t}\nabla^2\chi - \mathbf{k} \cdot \nabla \times \mathbf{Q} + \nabla^2(\Phi + \mathbf{K}) = 0. \tag{8.1.6}$$

These different forms of the shallow water equations are useful for highlighting different components of the energy. As we will shortly see, the vorticity-divergence form is helpful for separating the energy into balanced and unbalanced components, while the irrotational-solenoidal (Helmholtz) form is helpful for separating the energy into divergent, rotational and potential components.

8.2 Energy in shallow water systems

The kinetic energy and potential energy densities are (with density terms omitted):

$$KE = \Phi\frac{|\mathbf{u}|^2}{2}, \tag{8.2.1}$$

$$PE = \frac{\Phi^2}{2}. \tag{8.2.2}$$

Taking the time derivative of these terms and substituting in the mass and momentum equations where appropriate, we derive the following equations for the evolution of the kinetic energy density and potential energy density.

$$\frac{\partial}{\partial t}\left(\Phi\frac{|\mathbf{u}|^2}{2}\right) + \nabla \cdot \left(\Phi\frac{\mathbf{u}^3}{2}\right) + \mathbf{u} \cdot \nabla\frac{\Phi^2}{2} = 0, \tag{8.2.3}$$

$$\frac{\partial}{\partial t}\left(\frac{\Phi^2}{2}\right) + \Phi\nabla \cdot (\mathbf{u}\Phi) = 0. \tag{8.2.4}$$

Therefore we can obtain an equation for the evolution of total energy:

$$\frac{\partial TE}{\partial t} = \frac{\partial}{\partial t}(KE + PE) = \frac{\partial}{\partial t}\frac{1}{2}(\Phi\mathbf{u}^2 + \Phi^2) = -\nabla \cdot \left[\Phi\mathbf{u}(\Phi + \frac{\mathbf{u}^2}{2})\right]. \tag{8.2.5}$$

Integrating over the area and using Gauss's theorem, we see that the energy is conserved, and we can write

$$E = \frac{1}{2}\int (\Phi|\mathbf{u}|^2 + \Phi^2) dA, \tag{8.2.6}$$

where A is the area. The potential energy of a system has available and unavailable components. The unavailable potential energy, UPE, is the potential energy that cannot be converted to kinetic energy. Intuitively, an arbitrary fluid will have points that are deeper or shallower than the mean. Forces such as gravity will convert this potential into kinetic energy until the fluid is at rest. This fluid still has potential energy, but there is no way to convert it into kinetic energy - such a fluid would have zero depth. This unavailable potential energy will thus be dependent on the mean geopotential $\bar{\Phi}$ and the available potential energy will be dependent on the geopotential perturbation Φ'.

The available energy, AE, of a system is the total energy that can be converted into work. Kinetic energy and available potential energy are components of the available energy.

It is convenient to write the potential energy as the sum of the available energy, and the unavailable potential energy such that

$$E = \frac{1}{2}\int (\Phi|\mathbf{u}|^2 + \bar{\Phi}^2 + \Phi'^2) dA, \tag{8.2.7}$$

where $AE = \int \frac{\Phi|\mathbf{u}|^2}{2} + \frac{\Phi'^2}{2} dA$ and $UPE = \int \frac{\bar{\Phi}^2}{2} dA$

8.2.1 Quadratic approximation

Calculating the integral of the nonlinear kinetic energy term will be difficult, and so it is desirable to find a good approximation to the energy equation that is more simple. For an energy spectrum it makes the most sense to think about energy at some wavenumber if the quantity is quadratic.

If we non-dimensionalise the terms inside the integral (8.2.7) using the same scalings as previously, we can write:

$$(\bar{\Phi}U^2 + fU^3L\hat{\Phi})|\hat{\mathbf{u}}|^2 + \bar{\Phi}^2 + f^2U^2L^2\hat{\Phi}'^2. \tag{8.2.8}$$

Rearranging these terms, we obtain

$$\bar{\Phi}U^2\left[\left(1 + \frac{Fr^2}{Ro}\hat{\Phi}\right)|\hat{\mathbf{u}}|^2 + \frac{1}{Fr^2} + \frac{Fr^2}{Ro^2}\hat{\Phi}'^2\right] \tag{8.2.9}$$

The aim is to be able to neglect the third order term, for which we require either (a) $\frac{Fr^2}{Ro} \ll 1$ or (b) $\frac{Fr^2}{Ro} \ll \frac{Fr^2}{Ro^2} \implies Ro \ll 1$. The $\frac{1}{Fr^2}$ term represents the UPE and is constant, so we can ignore it.

If this condition is met we may neglect the third order term and approximate the energy as

$$E \approx \frac{1}{2}\int (\bar{\Phi}|\mathbf{u}|^2 + \bar{\Phi}^2 + \Phi'^2) dA. \tag{8.2.10}$$

8.3 Potential enstrophy in shallow water systems

The potential enstrophy is given by:

$$Z = \frac{1}{2}\int \frac{(\xi+f)^2}{\Phi} dA = \frac{1}{2}\int \frac{\zeta^2}{\Phi} dA = \frac{1}{2}\int \Phi q^2 dA, \tag{8.3.1}$$

where q is the potential vorticity. We can seek a quadratic approximation as follows:

$$\begin{aligned} Z = \frac{1}{2}\int \frac{(\xi+f)^2}{(\bar{\Phi}+\Phi')} dA = \frac{1}{2}\int \frac{(\xi+f)^2}{\bar{\Phi}(1+\frac{\Phi'}{\bar{\Phi}})} dA = \frac{1}{2}\int \frac{(\xi+f)^2}{\bar{\Phi}} \left(1 - \frac{\Phi'}{\bar{\Phi}} + \frac{\Phi'^2}{\bar{\Phi}^2} - ...\right) dA \\ \approx \frac{1}{2\bar{\Phi}}\int f^2 + 2f\xi - f^2\frac{\Phi'}{\bar{\Phi}} + \xi^2 - 2f\xi\frac{\Phi'}{\bar{\Phi}} + f^2\frac{\Phi'^2}{\bar{\Phi}^2} dA. \end{aligned} \tag{8.3.2}$$

In order to drop the third order terms we require $\frac{Fr^2}{Ro}$ to be small by a similar argument to that in the previous section. All the order one terms integrate to zero, leaving a constant and the order two terms which may be written as:

$$Z \approx \frac{f^2 A}{2\bar{\Phi}} + \frac{1}{2\bar{\Phi}}\int \left(\xi - f\frac{\Phi'}{\bar{\Phi}}\right)^2 dA. \tag{8.3.3}$$

8.4 Energy via the Helmholtz decomposition

We wish to obtain the energy in the form of its divergent, rotational and potential components. Recalling equations (8.1.5-6), we may use the Helmholtz decomposition to express the energy in terms of χ and ψ. Neglecting terms that integrate to zero, this allows us to express (8.2.10) as

$$E \approx \frac{1}{2}\int (\bar{\Phi}[|\nabla\chi|^2 + |\nabla\psi|^2] + \bar{\Phi}^2 + \Phi'^2) dA \tag{8.4.1}$$

or

$$E \approx -\frac{\bar{\Phi}}{2}\int \left(\chi\delta + \psi\xi + \nabla\cdot(\chi\nabla\chi) + \nabla\cdot(\psi\nabla\psi) - \bar{\Phi} - \frac{\Phi'^2}{\bar{\Phi}}\right) dA. \tag{8.4.2}$$

With suitable boundary conditions, we may neglect the divergence terms such that

$$E \approx -\frac{\bar{\Phi}}{2}\int \left(\chi\delta + \psi\xi - \bar{\Phi} - \frac{\Phi'^2}{\bar{\Phi}}\right) dA. \tag{8.4.3}$$

with

$$AE \approx -\frac{\bar{\Phi}}{2}\int\left(\chi\delta + \psi\xi - \frac{\Phi'^2}{\bar{\Phi}}\right)dA. \tag{8.4.4}$$

The first two terms may be thought of as the kinetic energy relating to divergence and rotation, respectively.

8.5 Energy in terms of Fourier coefficients

To perform any detailed scale analysis of the energy, we must calculate its spectrum. We do this by the use of Fourier transforms to measure the energy by wavenumber. If we separate (8.4.4) into rotational, divergent and potential parts:

$$AE \approx \frac{1}{2}\int\bar{\Phi}|\nabla\chi|^2 dA + \frac{1}{2}\int\bar{\Phi}|\nabla\psi|^2 dA + \frac{1}{2}\int\Phi'^2 dA, \tag{8.5.1}$$

In a square domain of dimension N we may relate a function f to its fourier transform $\hat{f}$ by

$$f = \frac{1}{N^2}\sum_{\mathbf{k}}\hat{f}e^{i\mathbf{k}\cdot\mathbf{x}}, \tag{8.5.2}$$

where $\mathbf{k} = (k, l)$. Using this and Parseval's theorem, we write (8.5.1) as

$$E \approx \frac{1}{2}\int\bar{\Phi}|\widehat{\nabla\chi}|^2 d\mathbf{k} + \int\bar{\Phi}|\widehat{\nabla\psi}|^2 d\mathbf{k} + \int\bar{\Phi}^2 + |\hat{\Phi}|^2 d\mathbf{k}, \tag{8.5.3}$$

or

$$E \approx -\frac{1}{2}\int\bar{\Phi}|\mathbf{k}|^2|\widehat{\chi}|^2 d\mathbf{k} + \int\bar{\Phi}|\mathbf{k}|^2|\widehat{\psi}|^2 d\mathbf{k} - \int\bar{\Phi}^2 + |\hat{\Phi}|^2 d\mathbf{k}. \tag{8.5.4}$$

We can then express the divergence and vorticity as:

$$\hat{\delta} = -K^2\hat{\chi} \text{ and } \hat{\xi} = -K^2\hat{\psi} \tag{8.5.5}$$

where $K^2 = k^2 + l^2$. This leads to

$$\hat{u} = ik\hat{\chi} - il\hat{\psi} \text{ and } \hat{v} = il\hat{\chi} + ik\hat{\psi}. \tag{8.5.6}$$

And so we can express the energy associated with divergence as

$$E_{div} = -\frac{\bar{\Phi}}{2}\int\chi\delta\, dA = \frac{\bar{\Phi}}{2N^4}\int\left(\sum_{\mathbf{k}}\hat{\delta}e^{i(\mathbf{k}\cdot\mathbf{x})}\right)\left(\sum_{\mathbf{k}'}\frac{\hat{\delta}'}{K'^2}e^{i(\mathbf{k}'\cdot\mathbf{x})}\right)dA. \tag{8.5.7}$$

All terms in this form integrate to zero except where $\mathbf{k}' = -\mathbf{k}$. Hence we can write

$$E_{div} = \frac{\bar{\Phi}}{2N^4}\int\sum_{\mathbf{k}}\frac{|\hat{\delta}|^2}{K^2}\, dA = \frac{\bar{\Phi}D^2}{2N^4\Delta k^2}\sum_{\mathbf{k}}\frac{|\hat{\delta}|^2}{K^2} = \bar{\Phi}\mathcal{N}\sum_{\mathbf{k}}\frac{|\hat{\delta}|^2}{K^2}, \tag{8.5.8}$$

where Δk is the wave number interval and D is the domain size, $\mathcal{N} = \frac{D^2}{2N^4\Delta k^2}$ is a normalisation constant used for the Fourier transform. Similarly, we can express the kinetic energy associated with rotation as

$$E_{rot} = \bar{\Phi}\mathcal{N}\sum_{\mathbf{k}} \frac{|\hat{\xi}|^2}{K^2}. \tag{8.5.9}$$

Thus we can express the kinetic energy as

$$KE = E_{div} + E_{rot} = \bar{\Phi}\mathcal{N}\sum_{\mathbf{k}} \frac{|\hat{\delta}|^2 + |\hat{\xi}|^2}{K^2}. \tag{8.5.10}$$

The available potential energy is given by

$$APE = \mathcal{N}\sum_{\mathbf{k}} |\hat{\Phi}'|^2, \tag{8.5.11}$$

and the unavailable potential energy is a constant

$$UPE = \frac{D\bar{\Phi}^2}{2}. \tag{8.5.12}$$

The total energy then should be $AE = KE + APE$, which in spectral space is

$$AE = \mathcal{N}\sum_{\mathbf{k}} \left(\frac{\bar{\Phi}(|\hat{\delta}|^2 + |\hat{\xi}|^2)}{K^2} + |\hat{\Phi}'|^2 \right). \tag{8.5.13}$$

8.6 Balanced-unbalanced decomposition

Given (8.5.5) we may also decompose the rotational and potential energy contributions into balanced and unbalanced parts.

We start by linearising the vorticity-divergence form of the shallow water equations to give:

$$\frac{\partial \xi}{\partial t} + f\delta = 0, \tag{8.6.1}$$

$$\frac{\partial \delta}{\partial t} - f\xi + \nabla^2\Phi' = 0. \tag{8.6.2}$$

We also require the linearised potential vorticity $q' = \xi - \frac{f\Phi'}{\bar{\Phi}}$.

8.6.1 Balanced modes

For balanced modes we require $\delta = 0$, and so from (8.6.2) we see that $f\xi = \nabla^2\Phi'$.

We can write

$$q' = \frac{\nabla^2\Phi'}{f} - \frac{f\Phi'}{\bar{\Phi}} = \left(\frac{\nabla^2}{f} - \frac{f}{\bar{\Phi}} \right) \Phi', \tag{8.6.3}$$

and so for a single Fourier mode

$$q' = \left(\frac{K^2\bar{\Phi} - f^2}{f\bar{\Phi}}\right)\Phi', \quad \xi = \frac{\nabla^2\Phi'}{f} = K^2\left(\frac{\bar{\Phi}q'}{K^2\bar{\Phi} - f^2}\right) = \frac{K^2\bar{\Phi}q'}{\bar{\Phi}K^2 + f^2}. \tag{8.6.4}$$

Similarly,

$$\psi = -\frac{\xi}{K^2} = -\frac{\bar{\Phi}q'}{\bar{\Phi}K^2 + f^2}. \tag{8.6.5}$$

The balanced energy can therefore be expressed as

$$E_{bal} = \frac{\bar{\Phi}}{2}\int -\psi\xi + \frac{\Phi'^2}{\bar{\Phi}}dA = \frac{1}{2}\int \frac{\bar{\Phi}K^2\bar{\Phi}^2|\hat{q'}|^2}{(\bar{\Phi}K^2 + f^2)^2} + \frac{\bar{\Phi}^2 f^2|\hat{q'}|^2}{(\bar{\Phi}K^2 + f^2)^2}dA = \frac{1}{2}\int \frac{\bar{\Phi}^2|\hat{q'}|^2}{\bar{\Phi}K^2 + f^2}dA \tag{8.6.6}$$

8.6.2 Unbalanced modes

For unbalanced modes, we define a quantity related to the divergence tendency as the imbalance $s = \xi - \frac{\nabla^2\Phi'}{f}$ and say $q' = 0$. Hence $\xi = \frac{f\Phi'}{\bar{\Phi}}$ and

$$s = \frac{f\Phi'}{\bar{\Phi}} - \frac{\nabla^2\Phi'}{f} = \left(\frac{f}{\bar{\Phi}} - \frac{\nabla^2}{f}\right)\Phi' = \frac{1}{f\bar{\Phi}}\left(f^2 - \bar{\Phi}\nabla^2\right)\Phi'. \tag{8.6.7}$$

Thus for a single Fourier mode

$$\xi = \frac{f^2 s}{f^2 - \bar{\Phi}\nabla^2} = \frac{f^2 s}{\bar{\Phi}K^2 + f^2} \tag{8.6.8}$$

and

$$\psi = \frac{f^2 s}{K^2(\bar{\Phi}K^2 + f^2)}. \tag{8.6.9}$$

So we can express the energy associated with imbalance as

$$E_{imb} = \frac{\bar{\Phi}}{2}\int -\psi\xi + \frac{\Phi'^2}{\bar{\Phi}}dA = \frac{1}{2}\int \frac{\bar{\Phi}f^4|\hat{s}|^2}{K^2(\bar{\Phi}K^2 + f^2)^2} + \frac{\bar{\Phi}^2 f^2|\hat{s}|^2}{(\bar{\Phi}K^2 + f^2)^2}dA = \frac{1}{2}\int \frac{\bar{\Phi}f^2|\hat{s}|^2}{K^2(\bar{\Phi}K^2 + f^2)}dA \tag{8.6.10}$$

8.6.3 Total energy

With these contributions and the contribution from divergence:

$$E_{div} = \frac{1}{2}\int \frac{\bar{\Phi}|\hat{\delta}|^2}{K^2}dA \tag{8.6.11}$$

we can write the energy in terms of these modes $E_{bal} + E_{imb} + E_{div} = E_{bal} + E_{unb}$:

$$E = \frac{1}{2}\int \frac{\bar{\Phi}^2|\hat{q'}|^2}{\bar{\Phi}K^2 + f^2} + \frac{\bar{\Phi}f^2|\hat{s}|^2}{K^2(\bar{\Phi}K^2 + f^2)} + \frac{\bar{\Phi}|\hat{\delta}|^2}{K^2}dA \tag{8.6.12}$$

where E_{unb} is the energy associated with unbalanced energy. In spectral space this is

$$AE = \bar{\Phi}\mathcal{N}\sum_{\mathbf{k}}\left(\frac{\bar{\Phi}|\hat{q'}|^2}{\bar{\Phi}K^2 + f^2} + \frac{f^2|\hat{s}|^2}{K^2(\bar{\Phi}K^2 + f^2)} + \frac{|\hat{\delta}|^2}{K^2}\right). \tag{8.6.13}$$

This value should be equal to (8.5.13).

8.7 Energy and Enstrophy tendencies

The cascade of energy up- or downscale can be most easily seen in the tendencies of energy and enstrophy in spectral space. The formulations of the energy we derived in this chapter also allow us to separate the energy into divergent, rotational and potential parts, and balanced and unbalanced parts. This will enable us to not only see the movement of energy from scale to scale, but also the transfers between different components of the energy. As such:

$$\dot{E}_{div}(\mathbf{k}) = Re\left\{\frac{2\bar{\Phi}\mathcal{N}\hat{\delta}^*\dot{\hat{\delta}}}{K^2}\right\} = -Re\left\{\frac{2\bar{\Phi}\mathcal{N}\hat{\delta}^*\hat{D}}{K^2}\right\} = Re\left\{2\bar{\Phi}\mathcal{N}\hat{\chi}^*\hat{D}\right\}, \tag{8.7.1}$$

$$\dot{E}_{rot}(\mathbf{k}) = Re\left\{\frac{2\bar{\Phi}\mathcal{N}\hat{\xi}^*\dot{\hat{\xi}}}{K^2}\right\} = -Re\left\{\frac{2\bar{\Phi}\mathcal{N}\hat{\xi}^*\hat{Z}}{K^2\Delta k^2 N^4}\right\} = Re\left\{2\bar{\Phi}\mathcal{N}\hat{\psi}^*\hat{Z}\right\}, \tag{8.7.2}$$

$$\dot{E}_{pot}(\mathbf{k}) = Re\left\{2\mathcal{N}\hat{\Phi}'^*\dot{\hat{\Phi}}'\right\} = -Re\left\{2\mathcal{N}\hat{\Phi}'^*\hat{M}\right\}, \tag{8.7.3}$$

where $\hat{D} = \dot{\hat{\delta}}$, $\hat{Z} = \dot{\hat{\xi}}$ and $\hat{M} = \dot{\hat{\Phi}}$ will be useful when defining our spectral scheme. Also

$$\begin{aligned}\dot{E}_{bal}(\mathbf{k}) = Re\left\{\frac{2\bar{\Phi}^2\mathcal{N}\hat{q}^*\dot{\hat{q}}}{\bar{\Phi}K^2 + f^2}\right\} = \\ Re\left\{\frac{2\bar{\Phi}^2\mathcal{N}\hat{q}^*(\dot{\hat{\xi}} - \frac{f}{\bar{\Phi}}\dot{\hat{\Phi}}')}{\bar{\Phi}K^2 + f^2}\right\} = -Re\left\{\frac{2\bar{\Phi}\mathcal{N}\hat{q}^*(\bar{\Phi}\hat{Z} - f\hat{M})}{\bar{\Phi}K^2 + f^2}\right\},\end{aligned} \tag{8.7.4}$$

$$\begin{aligned}\dot{E}_{imb}(\mathbf{k}) = Re\left\{\frac{\bar{\Phi}f^2\hat{s}^*\dot{\hat{s}}}{\Delta k^2 N^4 K^2(\bar{\Phi}K^2 + f^2)}\right\} = \\ Re\left\{\frac{\bar{\Phi}f^2\hat{s}^*(\dot{\hat{\xi}} + \frac{K^2}{f}\dot{\hat{\Phi}}')}{\Delta k^2 N^4 K^2(\bar{\Phi}K^2 + f^2)}\right\} = -Re\left\{\frac{\bar{\Phi}f^2\hat{s}^*(\hat{Z} + \frac{K^2}{f}\hat{M})}{\Delta k^2 N^4 K^2(\bar{\Phi}K^2 + f^2)}\right\},\end{aligned} \tag{8.7.5}$$

where $\cdot^*$ denotes the complex conjugate.

Chapter 9

Experimental design

Designing a model that can mimic the effects of motion below its resolvable limit requires information on the impact of the subgrid scales on the resolvable scales. An intuitive means to accomplish this is to run a test case at high resolution and at low resolution and compare the two solutions. The difference between the solutions will represent the effects of the subgrid scale motion.

In order to assess these effects at a single timestep one may run the high resolution model for one step, truncate the data at the desired resolution, and re-run the step. As before the difference between the two solutions represents the effect of the subgrid scales. This method can be helpful as it avoids the cumulative effects of such scales that can distort the picture - for example a buildup of grid-scale noise or integration of processes that cannot be resolved at the lower resolution.

A spectral model allows us to very easily truncate the resolution by simply setting the energy at wavenumbers below a certain limit to zero. We can then re-run the step with the truncated data and difference the high and low resolution solutions. Averaging these data across several steps will show the 'trend' of the subgrid-scale impact on the resolvable scales. This was the method applied by Thuburn et. al. [48] on a spectral model of the barotropic vorticity equation. It would also be a simple matter to remove the energy at a single wavenumber and thus obtain the effects of that specific wavenumber on the flow, though we will not do this here.

To expand upon the work of Thuburn et al. [48], we look at a shallow water model and using the tendencies derived in chapter 8.

Where this previous work looked at the energy cascade as a whole, by expressing the energy contribution from each component we open up new options. The truncation method described above can be applied to each individual component, and by looking at the energy spectra from the other component (and the total energy spectrum) we can see what the effect of the subgrid energy of that

component is. For instance, we can look at the effect that truncating the rotational energy has on the divergent energy spectrum. Not only will this give an indication of the contribution from each component to the total energy, but it will let us see exchanges of energy between these components.

9.1 Spectral scheme

As our diagnostics involve spectra and are working in simple geometry, it is logical to use a spectral shallow water model for our experiments. We can easily calculate the (fast) Fourier transforms of the geopotential, vorticity and divergence at each time step. The proposed solution method for the shallow water equations is as follows:

1. Given the transforms of the vorticity and divergence, calculate the transforms of the stream function and velocity potential:
$$\hat{\psi} = -\frac{\hat{\xi}}{K^2},\ \hat{\chi} = -\frac{\hat{\delta}}{K^2}$$

2. Calculate the transforms of the horizontal velocity:
$$\hat{u} = ik\hat{\chi} - il\hat{\psi},\ \hat{v} = il\hat{\chi} + ik\hat{\psi}$$

3. Use an inverse Fourier transform to obtain the geopotential, vorticity and horizontal velocity in grid space.

4. Calculate mass and vorticity fluxes and the kinetic energy per unit mass.

5. Transform these terms back into spectral space.

6. Calculate the divergence of the spectral mass flux and vorticity flux:
$$\hat{M} = \nabla \cdot \hat{\mathbf{F}} = ik\hat{F}_x + il\hat{F}_y,\ \hat{Z} = \nabla \cdot \hat{\mathbf{Q}} = ik\hat{Q}_x + il\hat{Q}_y$$

7. We now have all the components needed to calculate the imbalance in spectral space, and hence the divergence tendency:
$$\hat{S}_x = -\hat{Q}_y + ik(\hat{\Phi} + \hat{K}),\ \hat{S}_y = \hat{Q}_x + il(\hat{\Phi} + \hat{K}),\ \hat{D} = ik\hat{S}_x + il\hat{S}_y$$

8. We can then form a spectral time stepping scheme using the mass, vorticity and divergence shallow water equations:
$$\hat{\phi}_t + \hat{M} = 0,\ \hat{\xi}_t + \hat{Z} = 0,\ \hat{\delta}_t + \hat{D} = 0$$

9. A simple leapfrog scheme is used for the time stepping:
$$\hat{\phi}_t^{n+1} = \hat{\phi}_t^{n-1} - 2\Delta t\hat{M},\ \hat{\xi}_t^{n+1} = \hat{\xi}_t^{n-1} - 2\Delta t\hat{Z},\ \hat{\delta}_t^{n+1} = \hat{\delta}_t^{n-1} - 2\Delta t\hat{D}$$

In addition to the method of solving the shallow water equations above, additional steps may be taken which are detailed below:

10. Apply scale-selective dissipation.

11. Apply scale-independent dissipation.

12. Apply forcing.

13. Apply Robert-Asselin time filter.

9.2 Forcing

Returning briefly to figure 7.1, we see the depicted energy cascading from large-scale energy input to small-scale energy dissipation. We add this energy at large scales through forcing. We can choose the exact details of the forcing to stimulate different behaviour in the system as we will see.

We define a forcing factor

$$F = 2\mathcal{F}N^2\Delta t\Re\{e^{i\tau_F t}\} \tag{9.2.1}$$

where $\mathcal{F}$ is the forcing coefficient, N^2 is the resolution of the $N \times N$ square domain and τ_F is a forcing timescale. If we define a forcing wavenumber $K_F = \sqrt{k_F^2 + l_F^2}$ we can force a 'ring' of random wavenumbers in $k - l$ space between K_F and $2K_F$. Note that there will be a different $\mathcal{F}$, and by extension a different F, for each forced wavenumber.

9.2.1 Balanced forcing

To investigate energy transfers in a regime where balance dominates, we wish to force only the balanced part of the spectrum. We can do this directly via a forcing in the PV field.

While we want to force PV directly, the prognostic variables our models works with are vorticity, divergence and geopotential. We therefore need to translate the PV forcing into geopotential and vortical forcing values that are added to the relevant prognostic variable at each time step.

We start with the linearised potential vorticity $q' = \xi - \frac{f\Phi'}{\bar{\Phi}}$, and to keep imbalance zero we require $f\xi = \nabla^2\Phi'$. Hence a PV forcing F becomes a geopotential forcing $\hat{\Phi}'_F$

$$F = -\left(\frac{K^2}{f} + \frac{f}{\bar{\Phi}}\right)\hat{\Phi}'_F \implies \hat{\Phi}'_F = -F\left(\frac{K^2}{f} + \frac{f}{\bar{\Phi}}\right)^{-1}, \tag{9.2.2}$$

and a vorticity forcing $\hat{\xi}_F$

$$F = \left(1 + \frac{f^2}{K^2\bar{\Phi}}\right)\hat{\xi}_F \implies \hat{\xi}_F = F\left(1 + \frac{f^2}{K^2\bar{\Phi}}\right)^{-1}. \tag{9.2.3}$$

Velocity response

As we are forcing the vorticity, it is a simple matter to calculate the corresponding velocity response.

$$\hat{\psi}_F = -\frac{\hat{\xi}_F}{K^2} = -\frac{F}{\left(K^2 + \frac{f^2}{\bar{\Phi}}\right)} \tag{9.2.4}$$

and so

$$\hat{u}_F = \frac{ilF}{\left(K^2 + \frac{f^2}{\bar{\Phi}}\right)},\ \hat{v}_F = -\frac{ikF}{\left(K^2 + \frac{f^2}{\bar{\Phi}}\right)} \tag{9.2.5}$$

9.2.2 Unbalanced forcing

We hope to be able to directly force the gravity wave part of the spectrum via forcing of the unbalanced component. We define the imbalance as

$$s = \xi - \frac{\nabla^2 \Phi'}{f}, \tag{9.2.6}$$

and given $q' = \xi - \frac{\Phi' f}{\hat{\Phi}} = 0$ we can define as before

$$F = \left(\frac{f}{\bar{\Phi}} + \frac{K^2}{f}\right)\hat{\Phi}'_F \implies \hat{\Phi}'_F = F\left(\frac{f}{\bar{\Phi}} + \frac{K^2}{f}\right)^{-1}, \tag{9.2.7}$$

or

$$F = \left(1 + \frac{K^2\bar{\Phi}}{f^2}\right)\hat{\xi}_F \implies \hat{\xi}_F = F\left(1 + \frac{K^2\bar{\Phi}}{f^2}\right)^{-1}. \tag{9.2.8}$$

We could also force the unbalanced spectrum by directly forcing the divergence.

9.3 Robert-Asselin time filter

One problem with the leapfrog time scheme is the amplification over time of a spurious computational mode (see e.g. [59]). This has the effect of time splitting, whereby the odd and even time steps have different dependences.

The solution we employ here is a time filter developed by Robert [39] and was shown by Asselin [3] to suppress the computational mode with little damping of the physical mode:

$$\overline{\Psi} = \Psi^n + \rho(\Psi^{n+1} - 2\Psi^n + \overline{\Psi}^{n-1}), \tag{9.3.1}$$

where Ψ^n is some field (vorticity, geopotential or divergence here) at time level n, and $\overline{\Psi}$ is the filtered form of the field. ρ is a coefficient determining the strength of the filtering. Ideally this is kept as low as possible to suppress the computational mode without impacting the numerics too much. For our experiments we have used $\rho = 0.1$.

9.4 Aliasing

In order to correctly represent waves, it is necessary to sample each wavelength at least twice. Thus on a grid of N points the maximum wavenumber that can be accurately represented is $k_{max} = \frac{N}{2}$. Any wavenumber above this limit will be aliased back onto wavenumber $|N - k|$. We therefore may wish to truncate our Fourier series such that the highest wavenumber is k_{max}, and aliasing will not occur. However, nonlinear (quadratic) terms in a Fourier series truncated at wavenumber k_T will produce wavenumbers up to $2k_T$ (see, for example [15]) and hence we need to ensure that $|N - 2k_T| > k_T$ so that the wavenumbers that are being aliased onto are above the truncation limit. Thus we require $N > 3k_T$.

9.5 Test cases

There are a number of things we wish to focus our investigation on: We wish to examine the interactions between different component, and to examine the differences in these interactions between different flow regimes.

The first case we are interested in is a balanced regime. This has been defined to be as close as possible to Thuburn, Kent and Wood (TKW) [48], in order that we might compare results to that work too.

This thesis is particularly interested in the interactions between waves and turbulence and as such we would like a regime in which the balanced energy dominates at large scales, and unbalanced energy dominates at small scales. As mentioned in chapter 1, this is believed to be representative of the real atmosphere, and hence as model resolution increases we will find ourselves resolving more and more of the unbalanced part of the spectrum. In practise, we found that this could only be achieved in our model with significant forcing of the unbalanced spectrum, resulting in an unbalanced regime. Nonetheless this regime is of particular interest in the design of subgrid models as it is far less studied than balanced regimes.

Experiments with the unbalanced regime demonstrated the apparent formation of shocks, which we discussed in chapter 7. We therefore also sought out a second unbalanced regime which minimised the presence of these shocks.

To keep the different regimes as comparable as possible, we keep the energy added from forcing constant and equal to that in TKW. The dissipation coefficients and timescales are also kept constant and equal to TKW across all regimes. The initial condition provides our initial vorticity and we set the flow to be divergence free at $t = 0$. These define our initial velocities.

Hence the balanced regime was differentiated from the first unbalanced regime only by which component of the flow was forced. This regime was where we discovered shocks (the shock regime), and hence we also defined a second regime (the unbalanced regime) that was a modified version of the shock regime with the aim of minimising these shocks.

All regimes had their spectra truncated at wavenumber $k_T = 85$ as in the majority of cases in TKW. Motions below this wavenumber represent the unresolved scales whose effects would be represented by a subgrid model. Comparison of the full-resolution and truncated energy spectra averaged over several steps will show what the average effect of these subgrid scales was on the entire spectrum, hopefully providing a clearer picture of energy transfers between different components and scales within the system.

9.6 Energy Spectra

The energy spectra at $t = 211$ are shown for the following regimes:

9.6.1 Balanced regime

The balanced regime is characterised by the dominance of the balanced spectrum. Forcing operates by adding energy directly to the balanced spectrum at a single wavenumber. The example here was designed to resemble that of Thuburn et al. [48], though in the shallow water rather than barotropic vorticity equation(s). This allows us a point of reference for comparison of results. The flow has $Ro \approx 0.5$ and $Fr \approx 0.1$

In the figures we can clearly see how balance is dominant at every scale: the total and balanced energy are nearly identical. Looking at figure 9.1 one can easily surmise that the balanced energy is predominantly rotational. Similarly figure 9.2 suggests the unbalanced energy is largely made up of imbalanced energy.

9.6.2 Unbalanced regime

Forcing of the unbalanced spectrum in the unbalanced regime was spread between wavenumbers 16-32, creating the 'peak' we can see in the spectra of figures 9.3-4. This was sufficient to keep the unbalanced spectrum above the balanced spectrum below the forcing scale(s). Other aspects of the flow (and so Ro and Fr) were kept the same as the balanced regime. Note the pattern of reliable peaks at harmonic wavenumbers down the scale. This pattern is a consequence of how

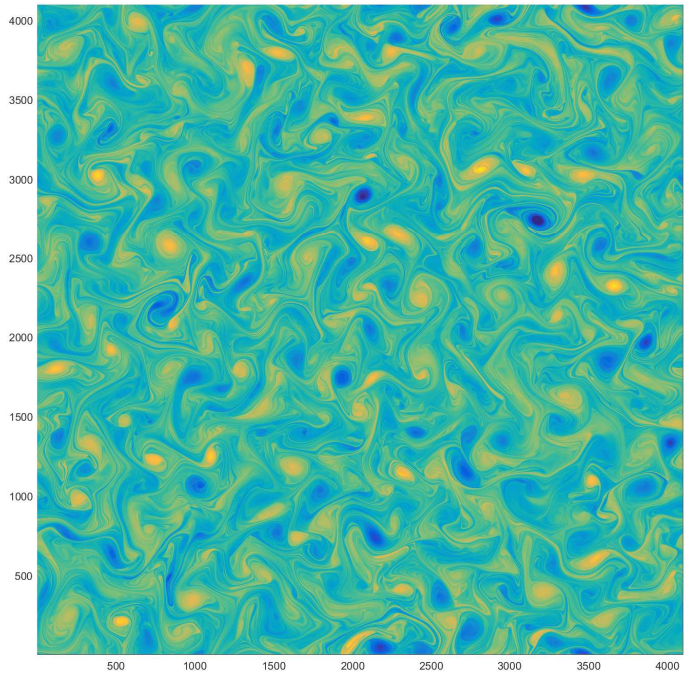

Figure 9.1: Field map of the potential vorticity in the balanced regime.

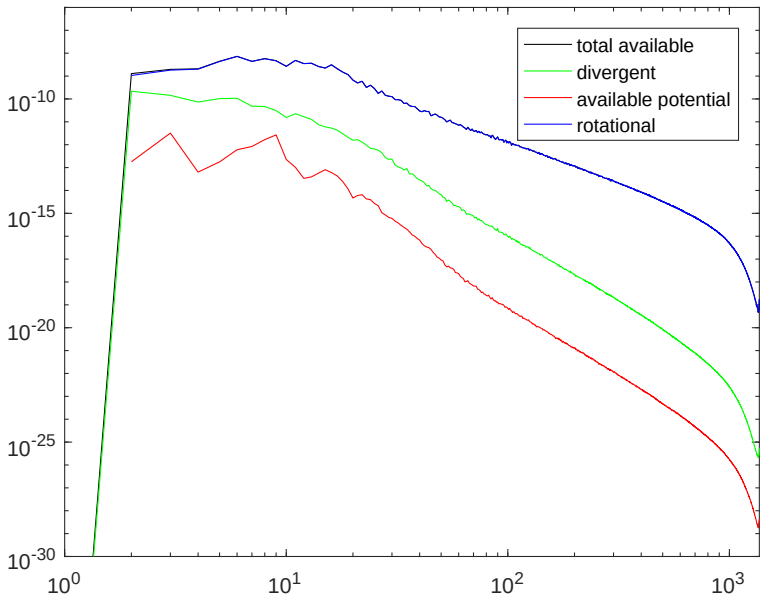

Figure 9.2: Energy spectra of the balanced regime. The total available energy is shown with the rotational, divergent and potential components. The rotational energy clearly dominates the flow.

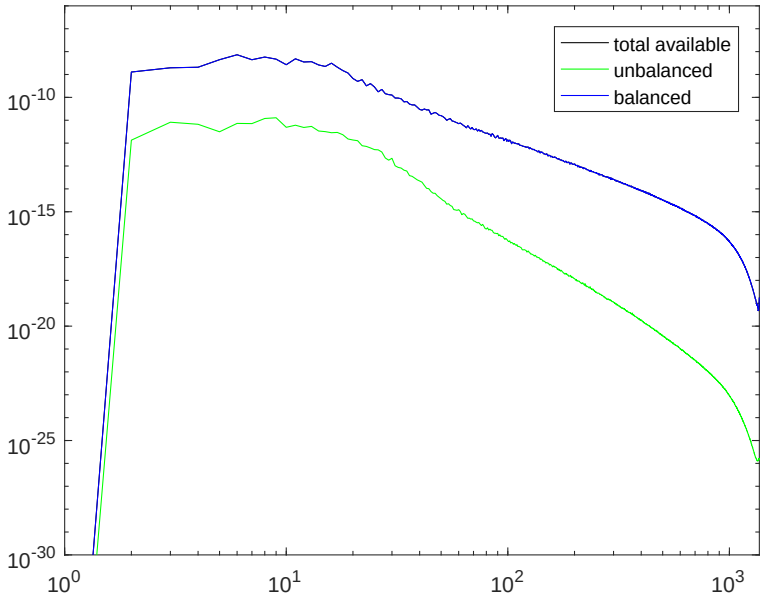

Figure 9.3: Energy spectra of the balanced regime. The total available energy is shown with the balanced and unbalanced components. The balanced energy clearly dominates the flow.

little the balanced and unbalanced energy interact - if the unbalanced energy occurs predominantly at a single wavenumber k, and unbalanced spectrum primariy interacts with itself, then wavenumber k interacts with wavenumber k resulting in a wavenumber $2k$ response. Energy at this wavenumber interacts with the wavenumber k energy for a wavenumber $3k$ response and so on, creating a pattern of energy concentrated at harmonic wavenumbers of k. Although we distributed the forcing randomly across wavenumbers 16-32, the pattern we happened to achieve concentrated the majority of energy close to wavenumber 25. The unbalanced regime shows a 'bump' at very high wavenumbers. At first glance this might suggest insufficient dissipation or evidence that we have not run the simulation at high resolution for long enough to have energy propagate the higher wavenumbers. Investigation however reveals that it is unaffected by modifying the dissipation and running the simulation for the entire run at the highest resolution. Later results also suggest that this is a robust aspect of the dynamics.

9.6.3 Shock regime

This regime was obtained by modifying the unbalanced regime with a smaller mean depth resulting in a greater Froude number ($Fr \approx 0.25$), as indicated in chapter 7. Figures 9.5-6 demonstrate 'sharper' peaks than in the unbalanced regime, indicating that energy is more localised at specific wavenumbers. Similar

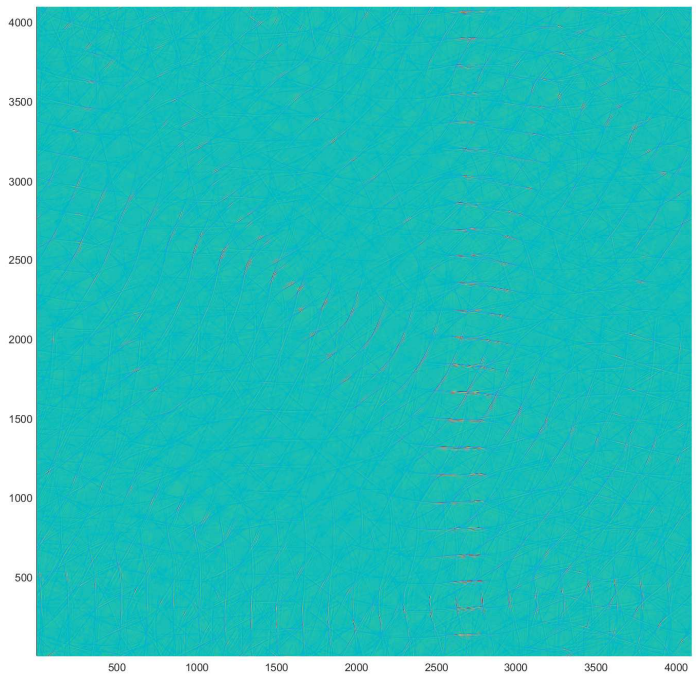

Figure 9.4: Field map of the divergence in the unbalanced regime. There is evidence of what appear to be trains of gravity waves.

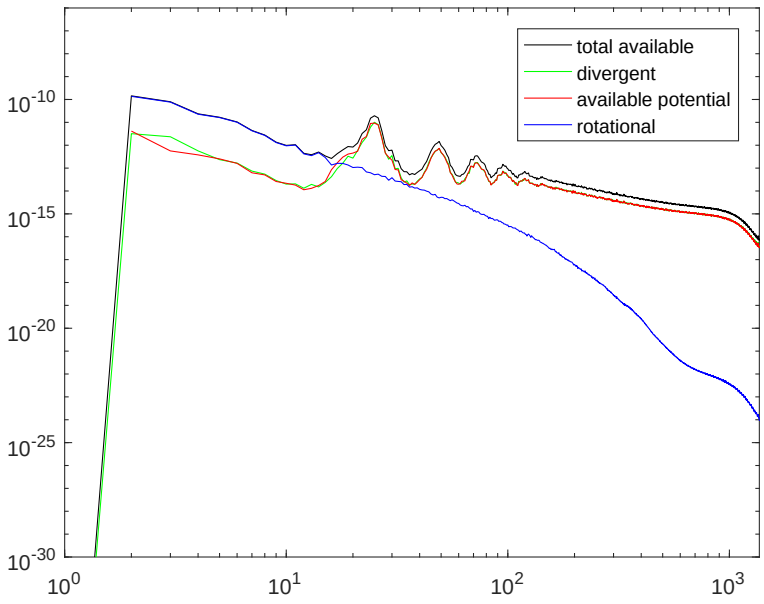

Figure 9.5: Energy spectra of the unbalanced regime. The total available energy is shown with the rotational, divergent and potential components. The divergent and potential energy clearly dominates the flow below the forcing scale.

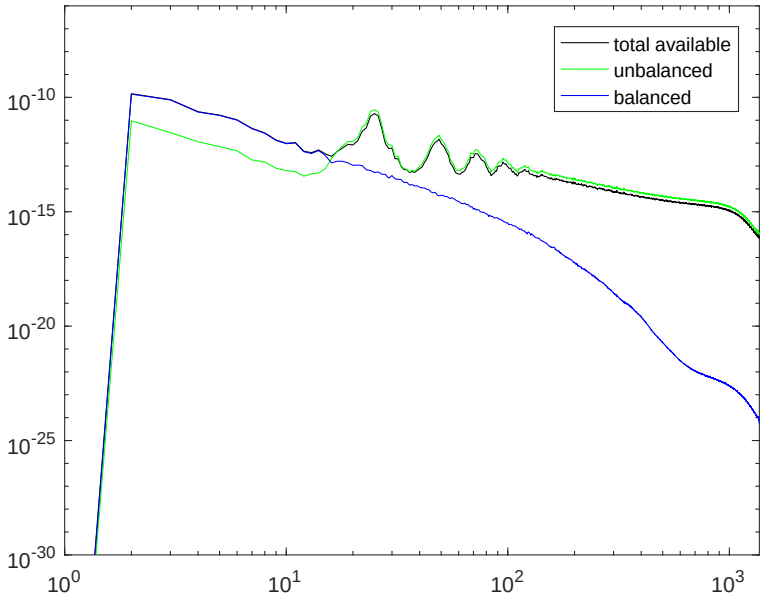

Figure 9.6: Energy spectra of the unbalanced regime. The total available energy is shown with the balanced and unbalanced components. The unbalanced energy clearly dominates the flow below the forcing scale.

to the unbalanced regime, we see a 'bump' in the spectrum at high wavenumbers.

9.7 Model setup

A shallow water spectral model has been initialised on a 4096^2 grid in each of the test cases described. For efficiency's sake, the model is not run for the full length at this resolution. Following on from the cascade timescale theory in chapter 7 and as the initial condition contains only large wavenumbers, we initialise the model at 512^2 before increasing to 1024^2 then to 2048^2 and finally to 4096^2. The cascade timescale suggests that we can safely run the model for 180 time units at 512^2, followed by 20 time units at 1024^2, 10 time units at 2048^2 and finally 2 time units at 4096^2.

To ensure that this approach was not significantly affecting th long-term dynamics of the system, this method was compared against running the model for 210 time units at 2048^2 and no notable differences were found in the results. TKW indicated that the energy tendency spectra was consistent after only a few steps, and that was also the case here. For this reason we were able to average the results for the truncated spectra over only a single time unit following an initial time unit run to allow the spectra to fill out. Running the truncation routine for longer would be prohibitively expensive as even a single time unit took around a week of

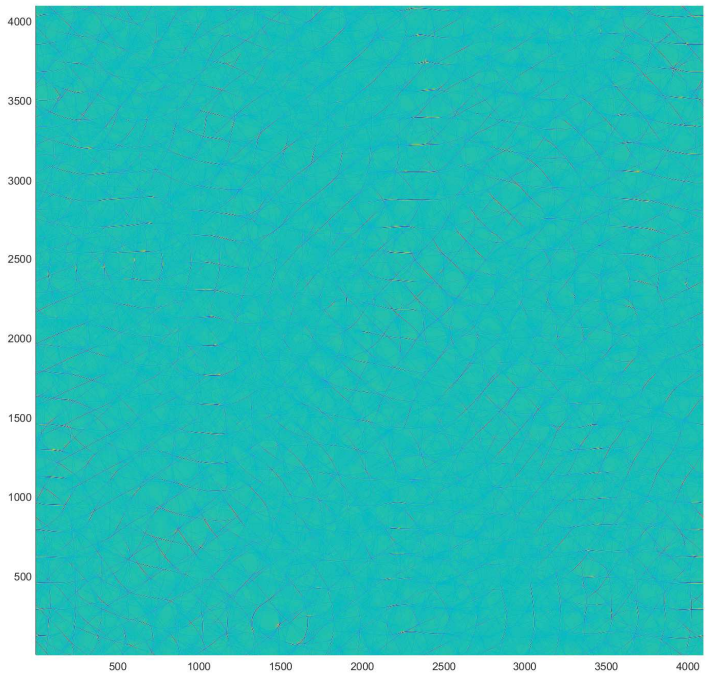

Figure 9.7: Field map of the divergence in the shock regime. What we believe to be shocks are clearly found throughout the domain.

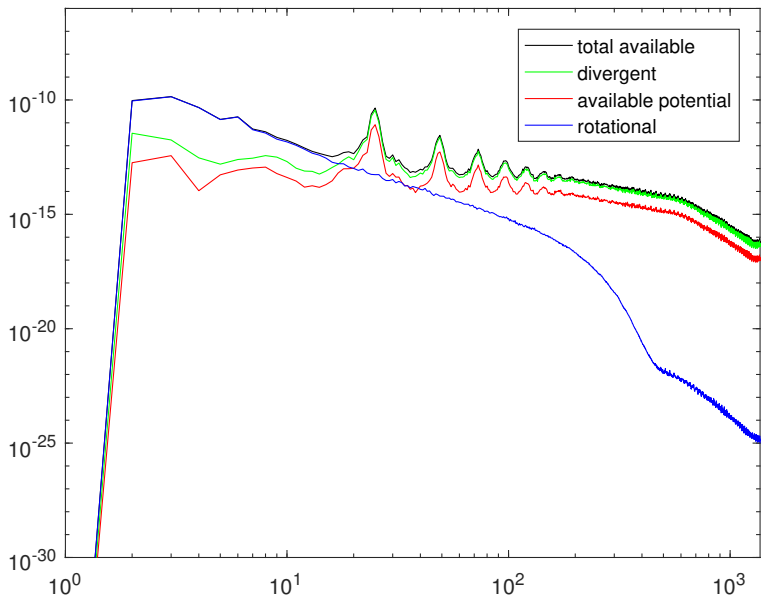

Figure 9.8: Energy spectra of the shock regime. The total available energy is shown with the rotational, divergent and potential components. The divergent and potential energy clearly dominates the flow below the forcing scale.

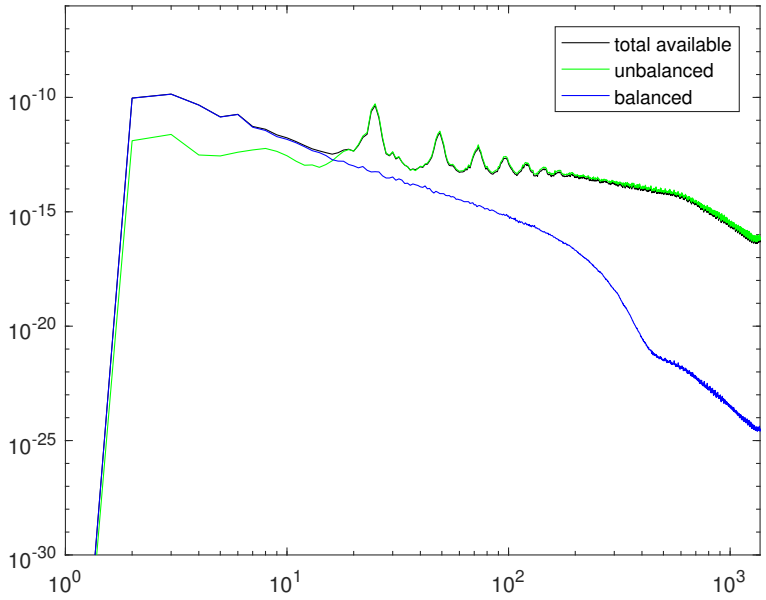

Figure 9.9: Energy spectra of the shock regime. The total available energy is shown with the balanced and unbalanced components. The unbalanced energy clearly dominates the flow below the forcing scale.

computation time on the supercomputer that was available.

Chapter 10

Results

10.1 Energy tendency spectra

The energy tendency spectra are a good indicator of where energy is 'going' in a flow. As total energy should remain constant in the absence of forcing and dissipation, the energy tendency shows where energy is being added and taken out of the spectrum, which should reveal transfers between scales and different components of the energy. By truncating the scales of different components of the energy, the difference between the truncated and untruncated energy spectra will demonstrate what the effect of those truncated scales was. Full results may be found in appendix C.

10.1.1 Truncating all components

Truncating all the components of the energy will show the overall impact of the subgrid scales - those scales that fall below the truncation wavenumber represent those scales that we cannot directly resolve in a model, and hence their effect might need to be replicated by a subgrid model.

This set of results should most closely resemble those of Thuburn et al. Similar to those results, figure 10.1 shows the subgrid scales of the total energy predominantly add energy to wavenumbers above the forcing scale (dashed line) and predominantly remove it at the dissipation scale (solid line). In addition to this result we see that the scales at which energy is added most lie between the forcing scale and the Rossby deformation scale (dotted line).

As the energy is dominated by the rotational energy, this structure is mirrored in the energy tendencies, with the total energy tendency being almost identical to the rotational energy tendency. We also see that a very small amount of divergent energy is being added at the large scales, though rather than being removed at the dissipation scale, we see it is removed just below the forcing scale. The

potential energy tendency is very small scale compared to the other components and shows no clear trend, suggesting that what we see is effectively noise.

Both the shock and unbalanced regimes (e.g. figure 10.2) show that the subgrid scales are taking energy directly from the most energetic wavenumbers - those that have been directly (or harmonically) forced. Figure 10.3 shows this effect when truncating the geopotential energy. Some energy is also being taken from the balanced energy at the largest scales, though it is comparatively small.

10.1.2 Truncating the geopotential energy

Truncating the geopotential energy will show the impact of the unresolved available potential energy on resolved scales.

Truncating the geopotential shows that the subgrid-scale geopotential energy follows roughly the same behaviour as the total energy by feeding energy to the large scales while removing them at the small. This mostly affects the balanced spectrum, though there is an impact on the unbalanced spectrum too. The rotational spectrum is completely independent of the geopotential, as is shown by the rotational energy tendency being zero everywhere in as in figure 10.3. The divergent spectrum shows zero energy tendency above the truncation wavenumber and mostly negative values below the truncation wavenumber, indicating that on average geopotential wavenumbers remove divergent energy from smaller wavenumbers. A small amount of energy is added to the unbalanced spectrum around the forcing wavenumber, and also just below the truncation wavenumber.

In the unbalanced and shock regimes, we see that the geopotential energy is a major cause of the energy removal from the forced wavenumbers. There also seems to be a small amount of energy being added to the largest scales of the balanced energy. Given that the rotational field is independent of the effects of the geopotential field, this energy must be being added to the potential energy.

10.1.3 Truncating the vortical energy

As the balanced energy is dominated by vorticity, this picture looks very similar to that when the total energy is truncated. The primary difference here is the effect on the potential energy, where figure 10.4 shows energy is primarily being removed from large scales, and added around the truncation scale. This indicates that there is a small upscale energy cascade from the rotational to the potential energy. It is also of note that the difference in the potential energy tendency is several times larger than that present in when the total energy is truncated, suggesting that there is a balancing effect from the process that adds energy to

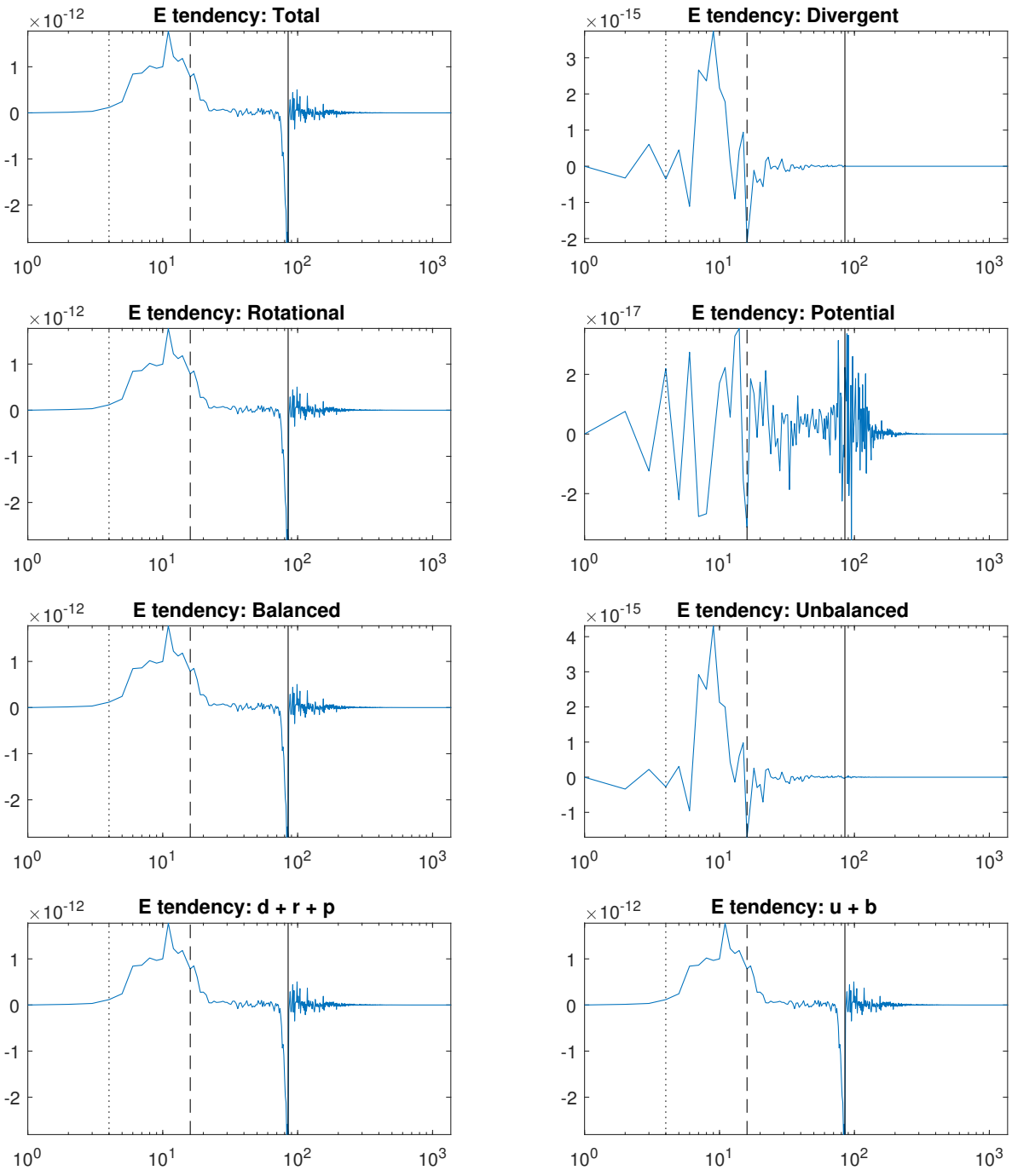

Figure 10.1: Difference in the energy tendency spectra resulting from truncation of all fields in the balanced regime. The dotted line indicates the Rossby deformation scale, the dashed line indicates the forcing scale and the solid line indicates the dissipation scale. Energy is predominantly being added between the forcing and deformation scale and removed at the dissipation scale.

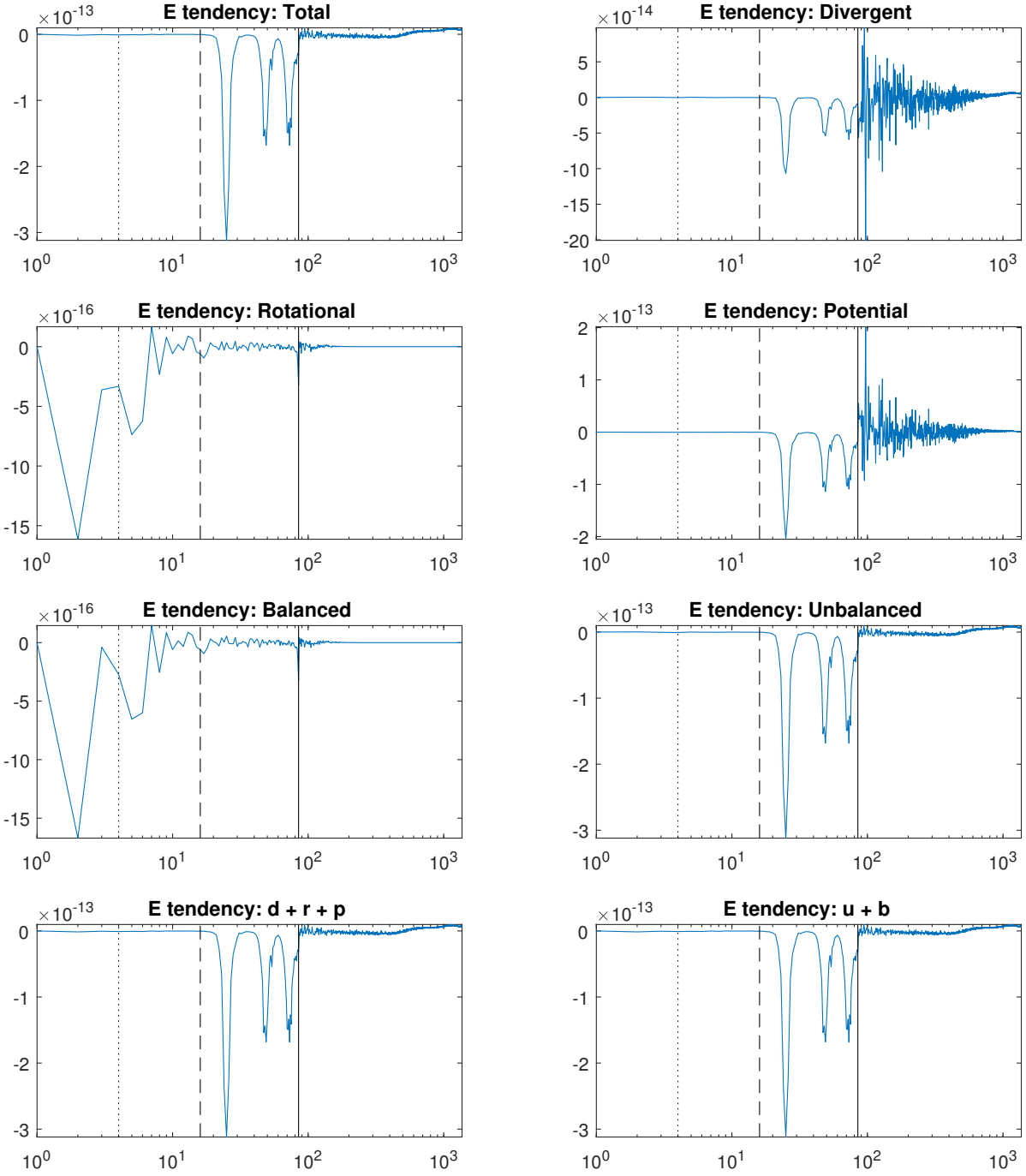

Figure 10.2: Difference in the energy tendency spectra resulting from truncation of all fields in the unbalanced regime. There is a clear pattern of energy being removed from the most energetic scales - those that are directly or harmonically forced.

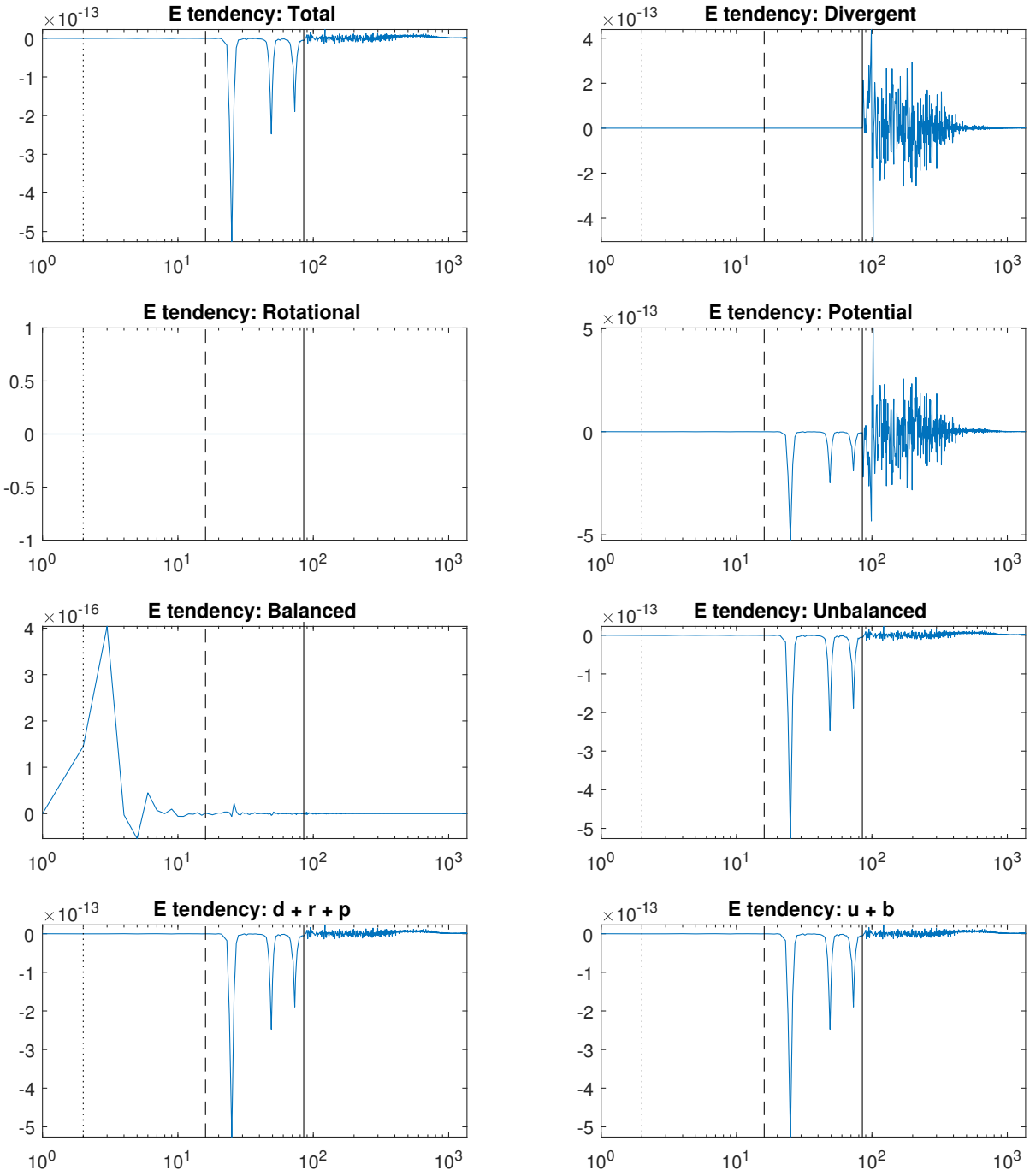

Figure 10.3: Difference in the energy tendency spectra resulting from truncation the geopotential in the shock regime. In addition to the pattern of energy removal we saw before, note the rotational energy subfigure, which demonstrates that the rotational and geopotential spectra are independent.

the geopotential field at large scales and removes it at small scales.

Vorticity plays a small role in the shock and unbalanced regimes and so in both these regimes, the energy tendency differences look like noise with the exception of a consistent removal of rotational energy from the largest scales (those above the Rossby deformation scale).

10.1.4 Truncating the divergent energy

In the balanced regime, figure 10.5 shows that the subgrid scales of divergence are adding energy to the scales between the forcing and Rossby deformation scales, much as the geopotential subgrid scales did. Though it is not clear from the figure due to the log axis used, rotational energy is being removed around wavenumbers 20-80. We also see that divergent wavenumbers add a small amount of potential energy at smaller wavenumbers.

Much like the the geopotential field, the divergent energy has the effect of removing energy from the forced wavenumbers, though the divergent energy is able to interact with both the divergent and potential energy at all wavenumbers. We also see that like vorticity, there is a minor removal of energy from wavenumbers above the Rossby deformation scale. The divergence appears to be responsible for the 'bump' in the spectra we observed earlier, as we can see that energy is being added at the same wavenumbers that this bump occurs.

10.1.5 Truncating the kinetic energy

Truncating the kinetic energy is achieved by truncating the vortical and divergent energy together.

In the balanced regime, the rotational energy dominates the kinetic energy and hence the results for truncating the kinetic energy look very similar to those for truncating the rotational energy.

Similarly, in the shock and unbalanced regimes, the kinetic energy is dominated by the divergence, and so the results look very similar to those for truncating the divergent energy.

10.1.6 Truncating the balanced energy

Truncating the balanced energy is achieved by truncating the potential vorticity.

The balanced energy is dominated by the rotational energy in all regimes, so there is again little difference in the energy tendency spectrum here and those of the rotational energy truncation figures.

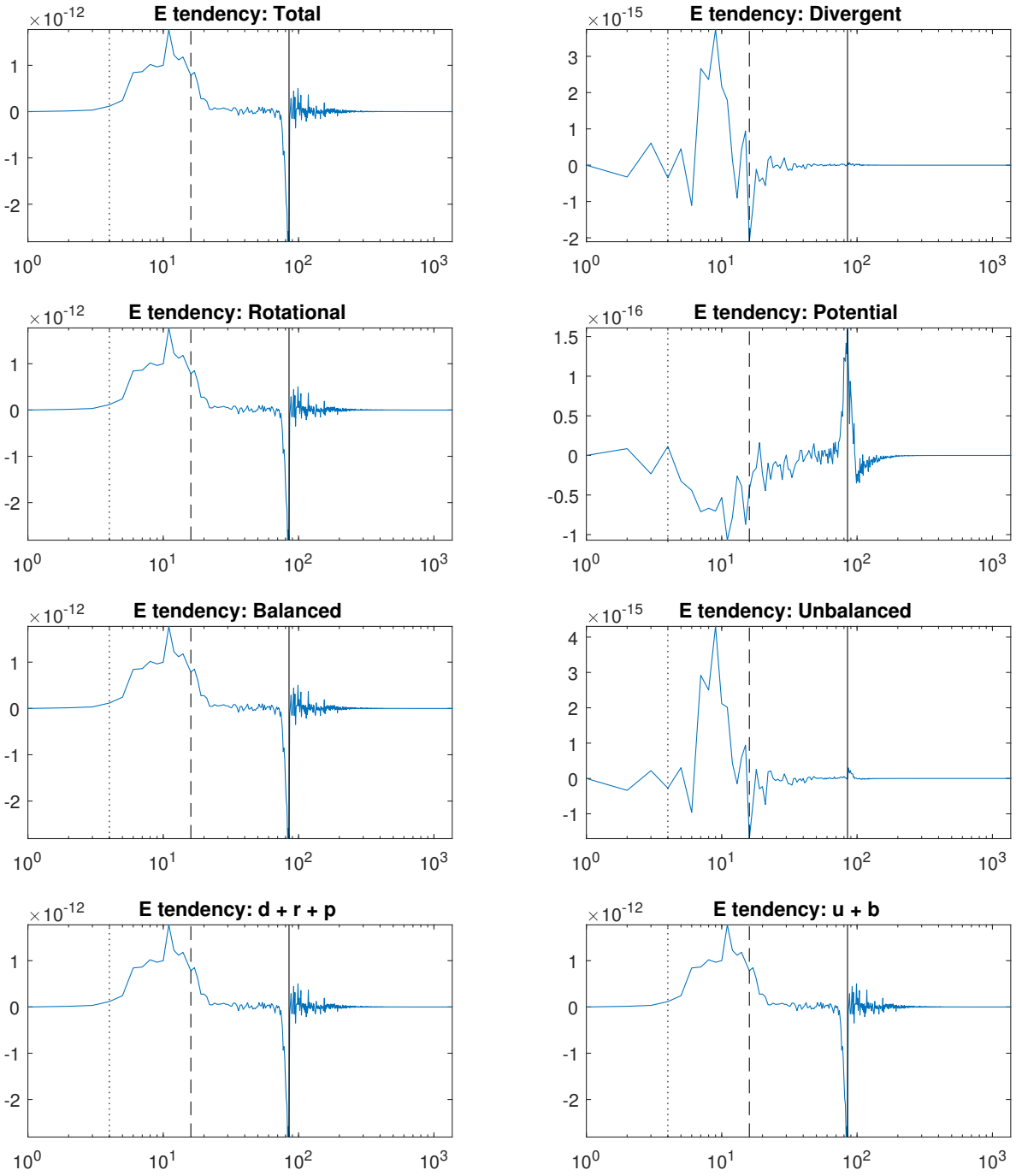

Figure 10.4: Difference in the energy tendency spectra resulting from truncation of the vorticity in the balanced regime. The potential energy subfigure shows that subgrid vorticity has the effect of removing large-scale potential energy and adding it around the truncation scale - an upscale energy cascade.

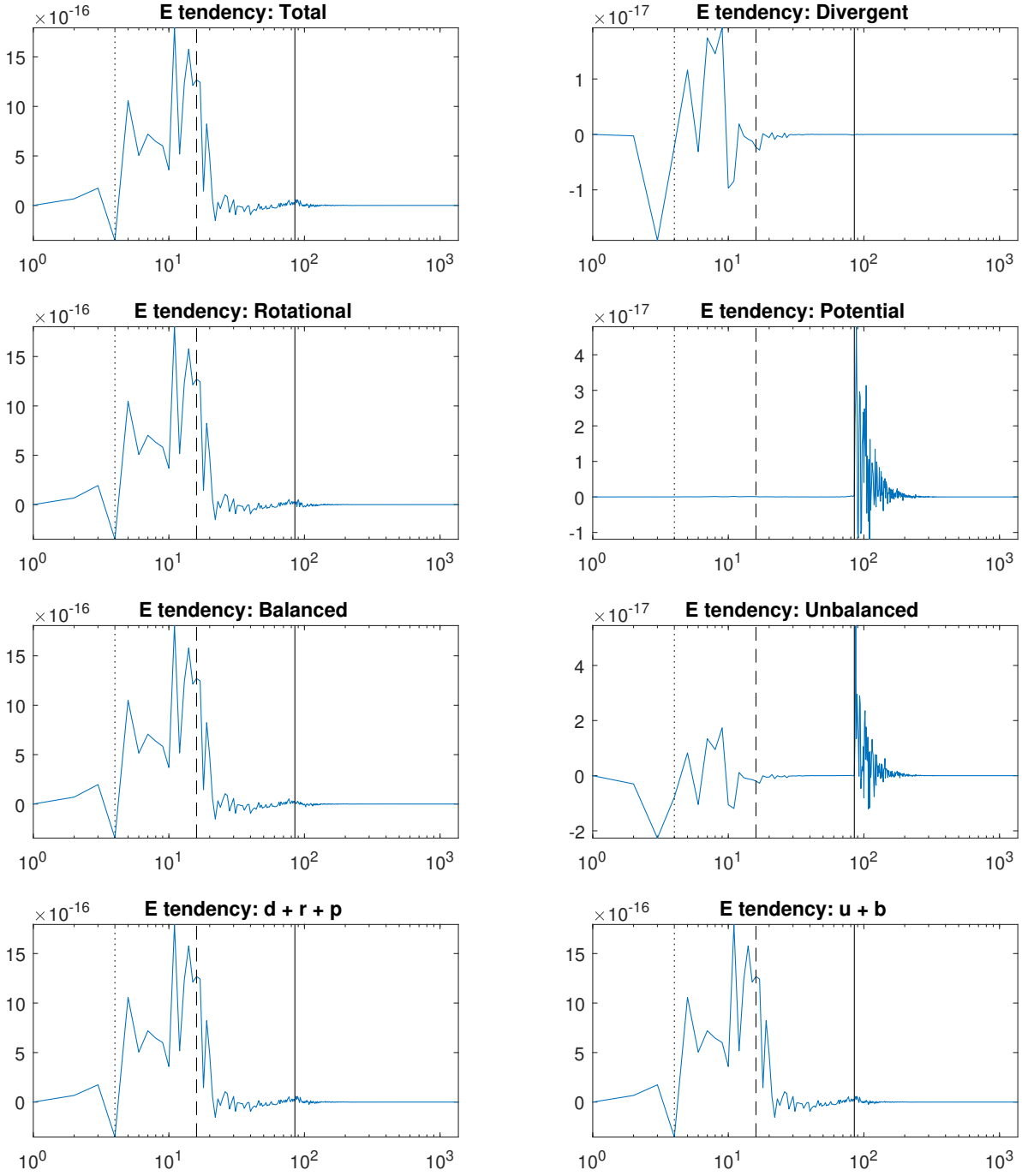

Figure 10.5: Difference in the energy tendency spectra resulting from truncation of the divergence in the balanced regime. Subgrid divergence has the effect of adding energy between the forcing and deformation scales, similar to the subgrid geopotential.

10.1.7 Truncating the unbalanced energy

In the balanced regime, imbalance is dominated by the divergence and the regime itself is dominated by the rotation. Hence the effect on the total energy looks like the effect of truncating the divergence on the rotational energy tendency spectrum. The effect on the unbalanced spectrum of the subgrid scales is to add energy around the forcing wavenumber and remove it from near the truncation wavenumber. The same is true of the effect on the potential spectrum.

Similar to the balanced regime, the energy tendency spectra for the shock and unbalanced regimes look nearly identical to those of the results from truncation of the divergence field.

Chapter 11

Summary of Part III

We have sought to expand upon the work of Thuburn et al. [48] by applying their technique for analysing spectral energy transfers to a shallow water model, and decomposing the energy into distinct parts. The hope was to gain an insight into the finer details of how energy transfers below the subgrid scale affect the resolvable scales and so inform the development of more accurate subgrid models.

We chose to split the energy into rotational, divergent and potential parts, as well as into balanced and unbalanced parts. We then expressed these components as quadratic approximations and calculated the energy tendencies for these components. The resulting diagnostics enabled us to look at the effect of each component's truncated scales on the others.

In addition to this we investigated these diagnostics in three different regimes: a balanced regime similar to that of Thuburn et al., and two unbalanced regimes - one exhibiting shocks and the other not.

In the balanced case, we see that the total energy tendency is dominated by the rotational energy and that on average the subgrid scales remove energy from just above the truncation scale and feed it to the large scales, as we saw in Thuburn et al. Truncating the geopotential produces similar results.

When looking at truncating the rotational energy alone, we see that although the results are similar to those from truncating the total energy, the subgrid scales associated with rotation appear to remove large-scale potential energy and add potential energy around the truncation scale. This process seems to be balanced by the effect of the subgrid potential scales on the potential energy. The subgrid scales associated with the kinetic and balanced energy behave very similarly to those associated with the rotational energy.

The divergent subgrid scales appear to primarily add divergent and rotational energy at large scales as well as adding potential energy at small scales in the balanced regime. The behaviour is similar when truncating the unbalanced

energy.

Despite the efforts we made to differentiate the two unbalanced regimes, they appear very similar in the diagnostics. This may suggest that the presence of shocks has a minimal effect on the energy spectra of the system, or that we were unsuccessful in removing shocks from the unbalanced regime (so both regimes remain essentially shock-dominated).

The most notable feature of our unbalanced regimes is the consistent removal of energy from the forced wavenumbers. This pattern is seen when truncating the total, geopotential, divergent, kinetic or unbalanced energy. The rotational and balanced subgrid scales seem no consistent pattern of energy input or removal except at the largest scales where energy is removed. It is possible that this removal of energy is the result of wavebreaking, which again may suggest that we have been unsuccessful in the removal of shocks from the gravity wave regime.

The distinct differences between the balanced and unbalanced regimes might necessitate that a different subgrid model be required for each regime.

Chapter 12

Conclusion

In this thesis we have sought two methods that will aid in the development of models, particularly in cases where waves and turbulence interact.

The first of these methods was the use of asymptotic limit solutions to define imbalance equations whose scaling should be known in appropriate asymptotic regimes.

We begun with demonstration of some examples in 1D, introducing the “slow equation” and highlighting some of the consequences of grid choice and time step size.

We then moved on to part two of the thesis in which we applied the method to two shallow water models, one using an Eulerian framework and the other using a semi-Lagrangian. Four different regimes were introduced - strong rotation (SR), barotropic vorticity (BV), quasigeostrophic (QG) and semigeostrophic (SG).

Low- and higher-order imbalance equations were defined, and their scalings given in each of the regimes. The low-order represent approximate geostrophic balance, divergence and divergence tendency. The higher-order imbalance equation may be thought of as the jerk of the system. All imbalance equations were defined for both the Eulerian and semi-Lagrangian models. The final diagnostics given were an outline of an idea for measuring gravity wave activity. The results for the SR, BV and SG limits suggest that the method may have some merit, but the QG results show that more effort would need to be spent on such an approach before it could be considered robust enough to be reliable.

By measuring these imbalance equations against our expected scaling we hoped to be able to assess the success of the models ability to capture balance. While the low-order equations scale as we expect, we found issues with the higher-order. In each case it appears that there was a crucial distinction between the continuous and discrete forms of the equations. The result was that the discrete forms of the imbalance equations we used did not appear to properly

respect the models' structure. This manifested as spurious scaling in almost all the higher-order test cases. Nevertheless, we believe that this method, properly applied, will provide a good check on a models ability to respect balance.

The final part of this thesis looked at a spectral shallow water model. We split the energy into different components, and by truncating these components in turn we were able to build up a picture of the truncated scales affected the energy spectra of each component. This method aims to provide information on the behaviour of a system below the subgrid scale, and so aid in the development of subgrid models.

This technique is applied to three different regimes - a balanced regime and two unbalanced regimes, one dominated by shocks and one in which we hoped shocks were minimised. The results from the balanced regime look similar to those from previous work, but the unbalanced regimes are not well studied. We discover a consistent pattern where subgrid scales remove energy from the most energetic wavenumbers (representing gravity/shock waves). In addition we see that energy seems to be added at very low wavenumbers in the unbalanced regimes, creating a 'bump' in the spectra. We do not know the mechanics behind these behaviours, but they suggest that subgrid scales behave significantly differently depending on the regime we are in, and perhaps a different subgrid model is required in each regime. It would be interesting to investigate these aspects further in order to find answers about the processes behind these results.

In some sense it may feel that we have raised more questions than we have answered. However, in part II we provided sufficient cause to see the merits of the asymptotic limit approach to assessing models. With relatively small changes to the technique we used it should be possible to avoid the pitfalls that we demonstrated here. Part III showed the amount of information that we were able to obtain by separating the energy into different components. There is evidence that as affordable resolutions increase and we find ourselves more in a regime where unbalanced energy dominates, then we may need to develop a new class of subgrid model that is capable of replicating the interactions we have seen in this thesis. Additionally, some of the patterns of energy transfers in the unbalanced regimes merit further study.

Appendix A

Finite difference model figures

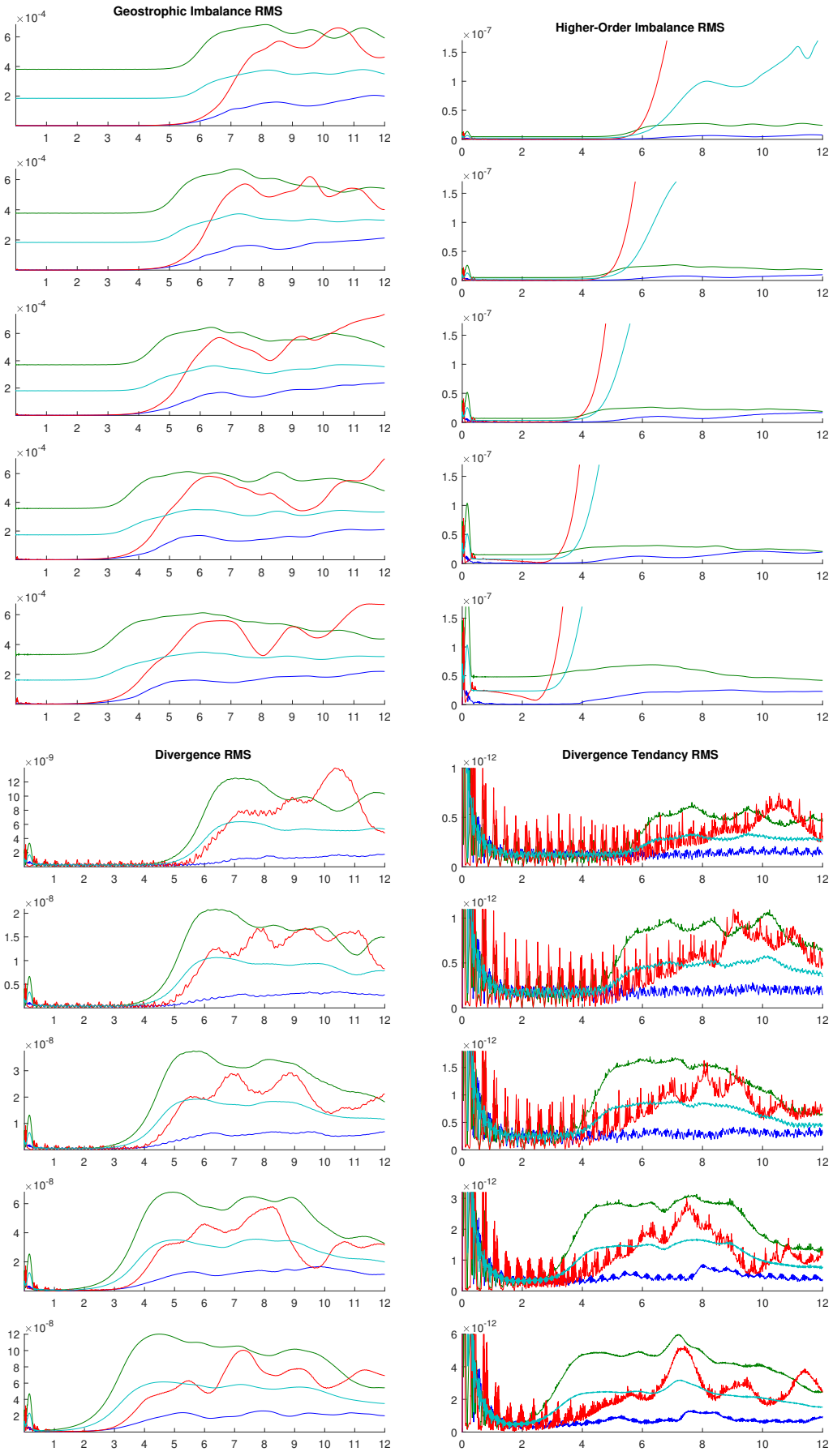

Figure A.1: Time series of diagnostic quantities approaching the strong rotation limit in the ENDGame model, $C_{gw} \approx 0.5$. For each subfigure the small parameter ε is, from top to bottom, = 1, 0.5, 0.25, 0.125, 0.0625.

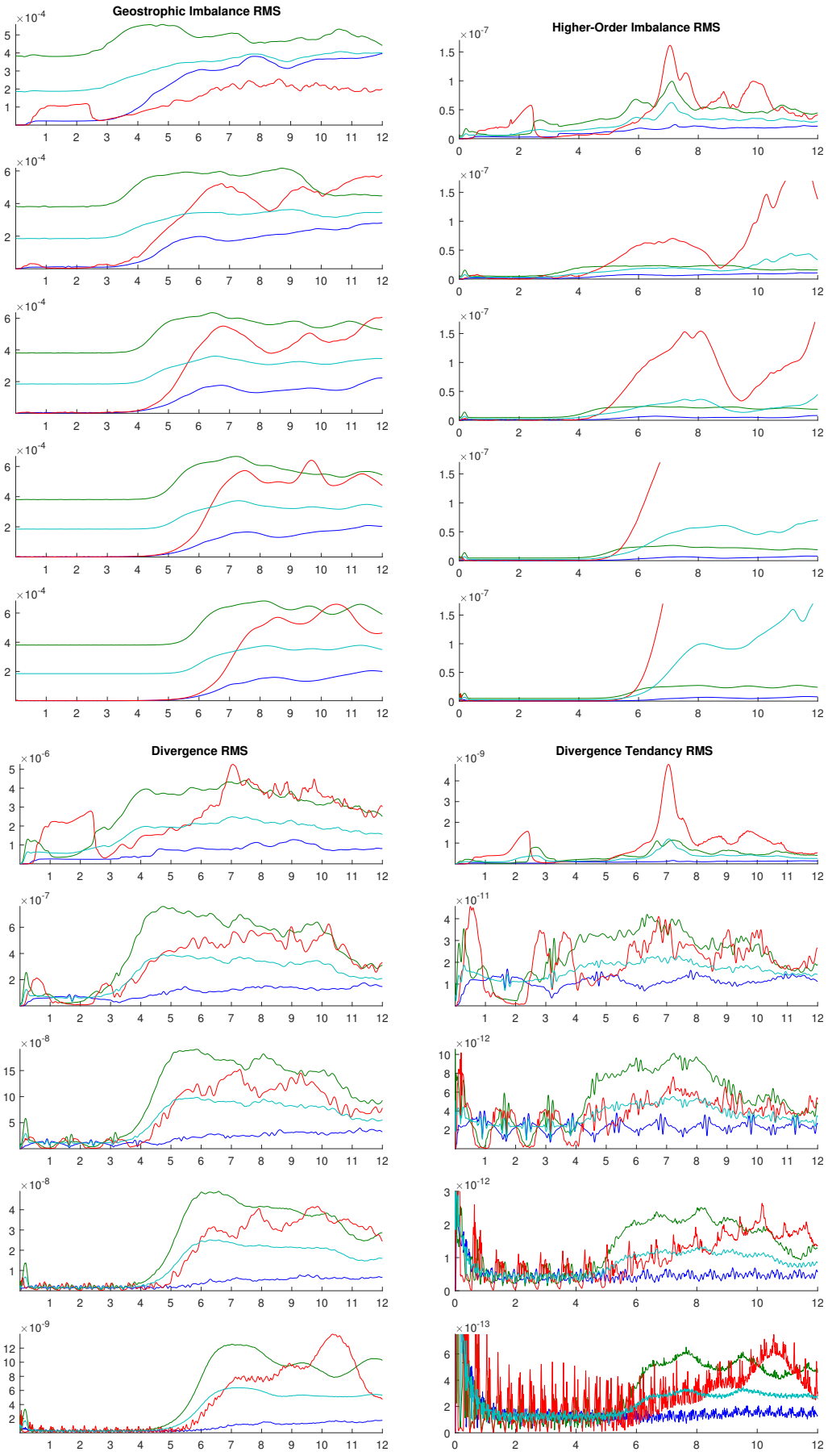

Figure A.2: Time series of diagnostic quantities approaching the barotropic vorticity limit in the ENDGame model, $C_{gw} \approx 0.5$. For each subfigure the small parameter ε is, from top to bottom, = 1, 0.5, 0.25, 0.125, 0.0625.

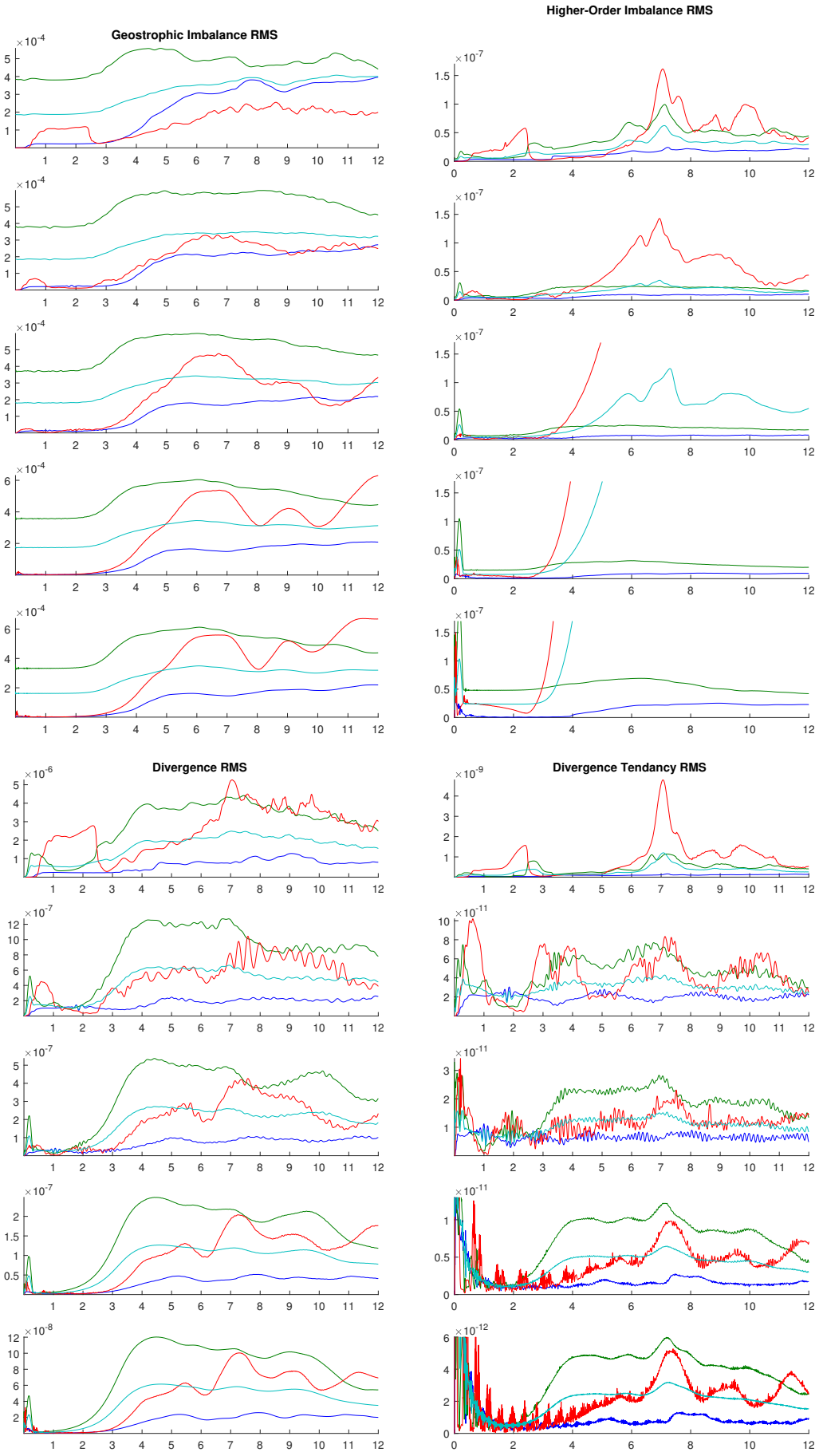

Figure A.3: Time series of diagnostic quantities approaching the quasigeostrophic limit in the ENDGame model, $C_{gw} \approx 0.5$. For each subfigure the small parameter ε is, from top to bottom, = 1, 0.5, 0.25, 0.125, 0.0625.

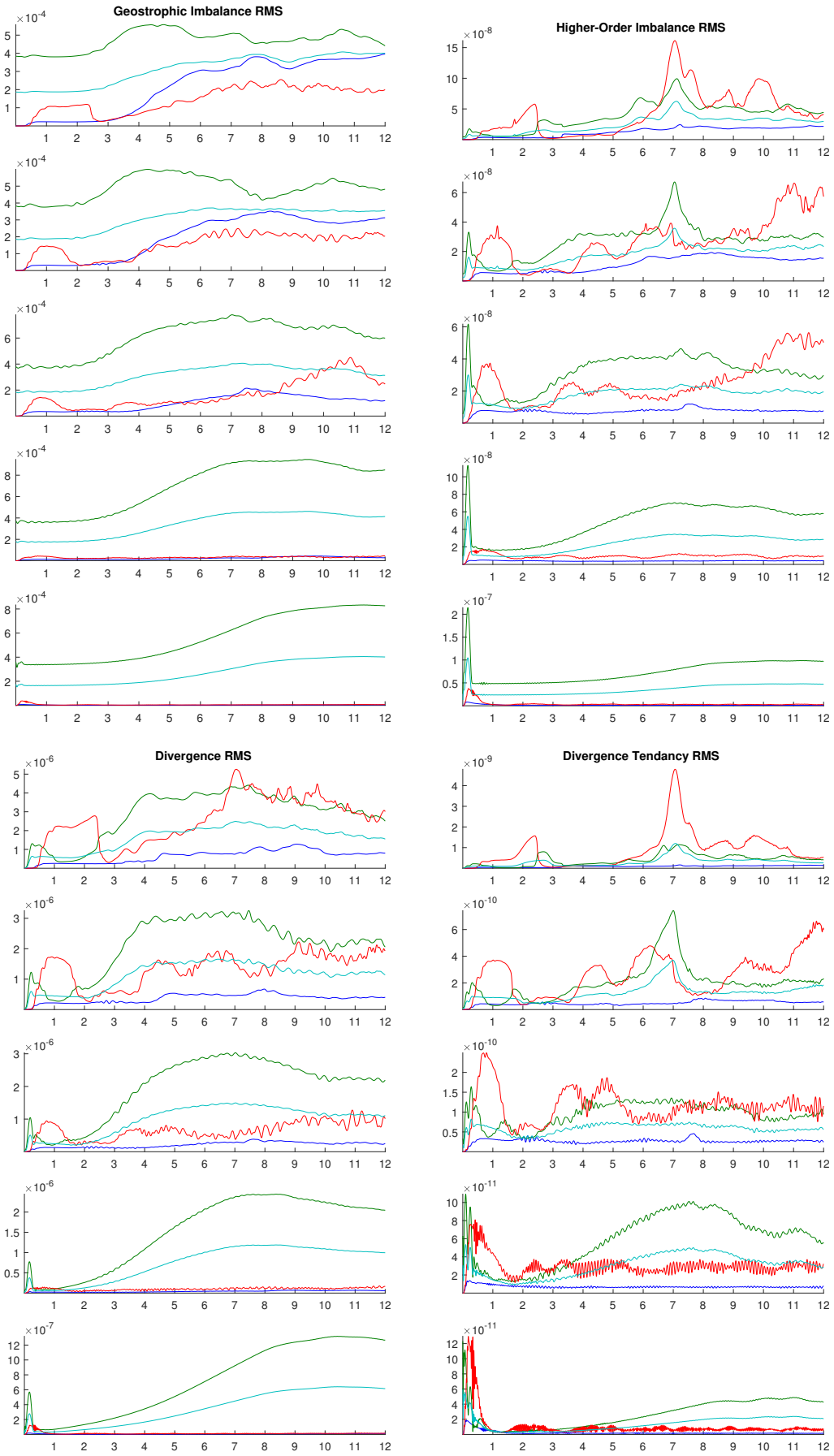

Figure A.4: Time series of diagnostic quantities approaching the semigeostrophic limit in the ENDGame model, $C_{gw} \approx 0.5$. For each subfigure the small parameter ε is, from top to bottom, = 1, 0.5, 0.25, 0.125, 0.0625.

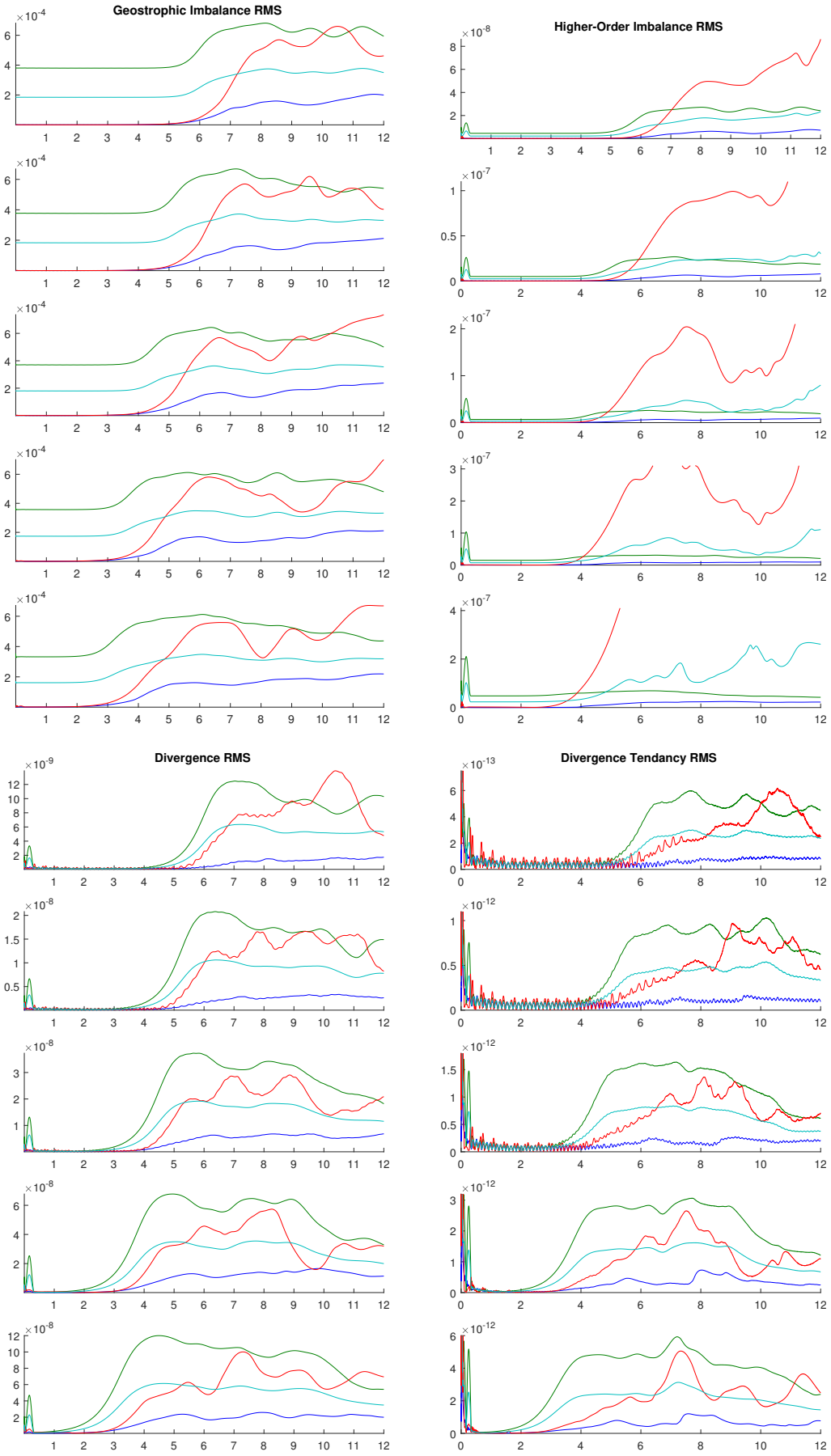

Figure A.5: Time series of diagnostic quantities approaching the strong rotation limit in the ENDGame model, $C_{gw} \approx 5$. For each subfigure the small parameter ε is, from top to bottom, = 1, 0.5, 0.25, 0.125, 0.0625.

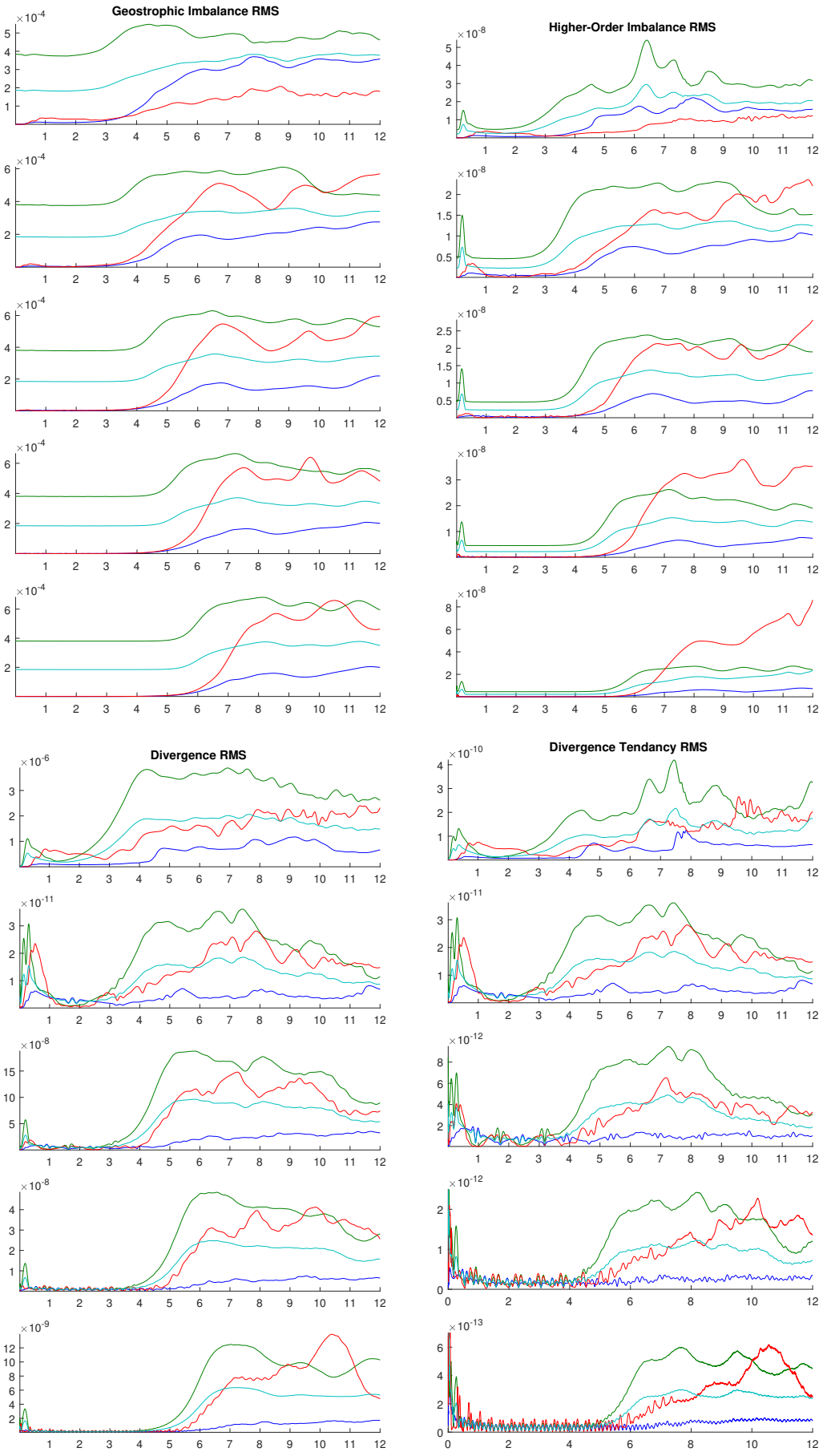

Figure A.6: Time series of diagnostic quantities approaching the barotropic vorticity limit in the ENDGame model, $C_{gw} \approx 5$. For each subfigure the small parameter ε is, from top to bottom, = 1, 0.5, 0.25, 0.125, 0.0625.

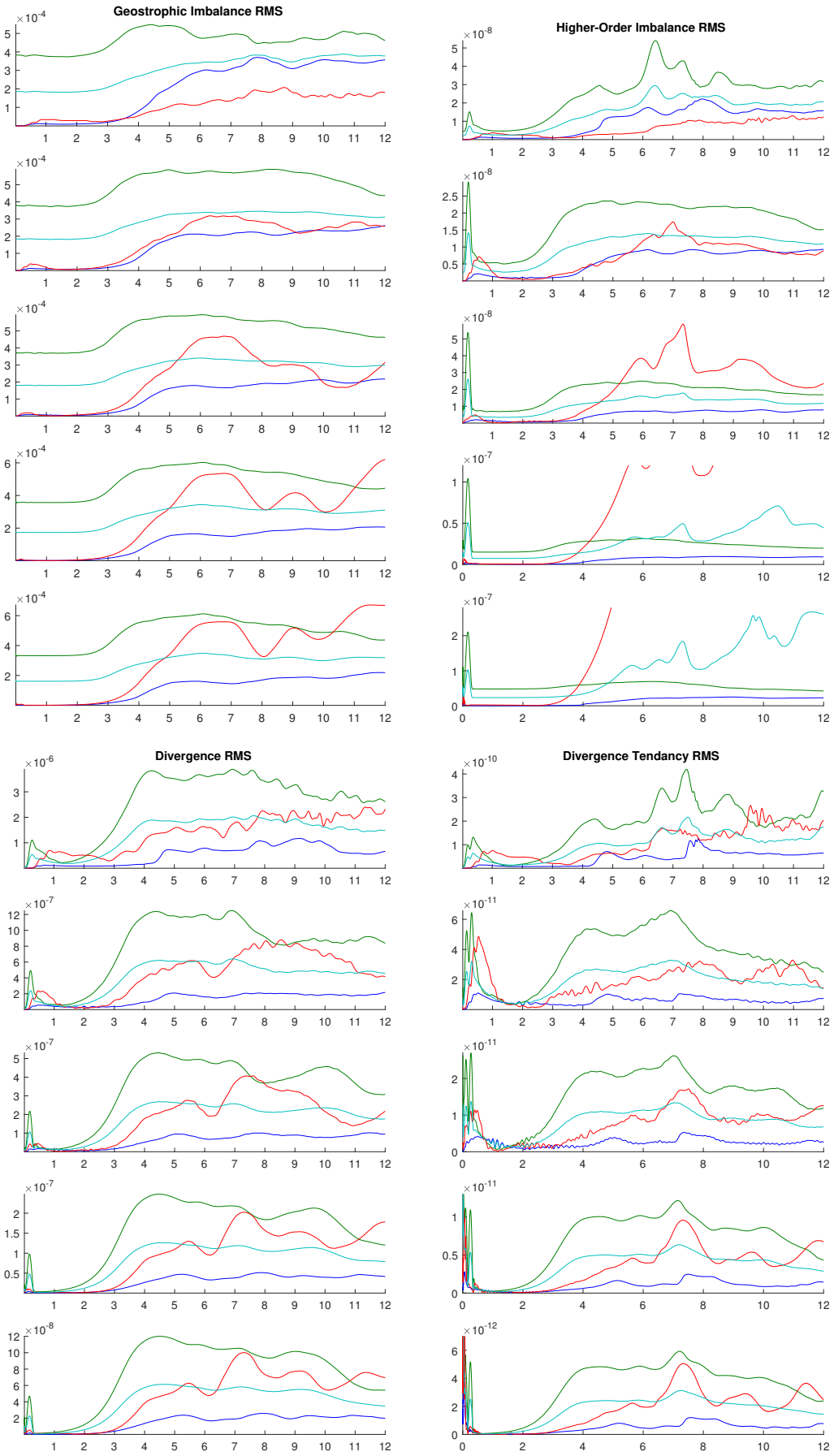

Figure A.7: Time series of diagnostic quantities approaching the quasigeostrophic limit in the ENDGame model, $C_{gw} \approx 5$. For each subfigure the small parameter ε is, from top to bottom, = 1, 0.5, 0.25, 0.125, 0.0625.

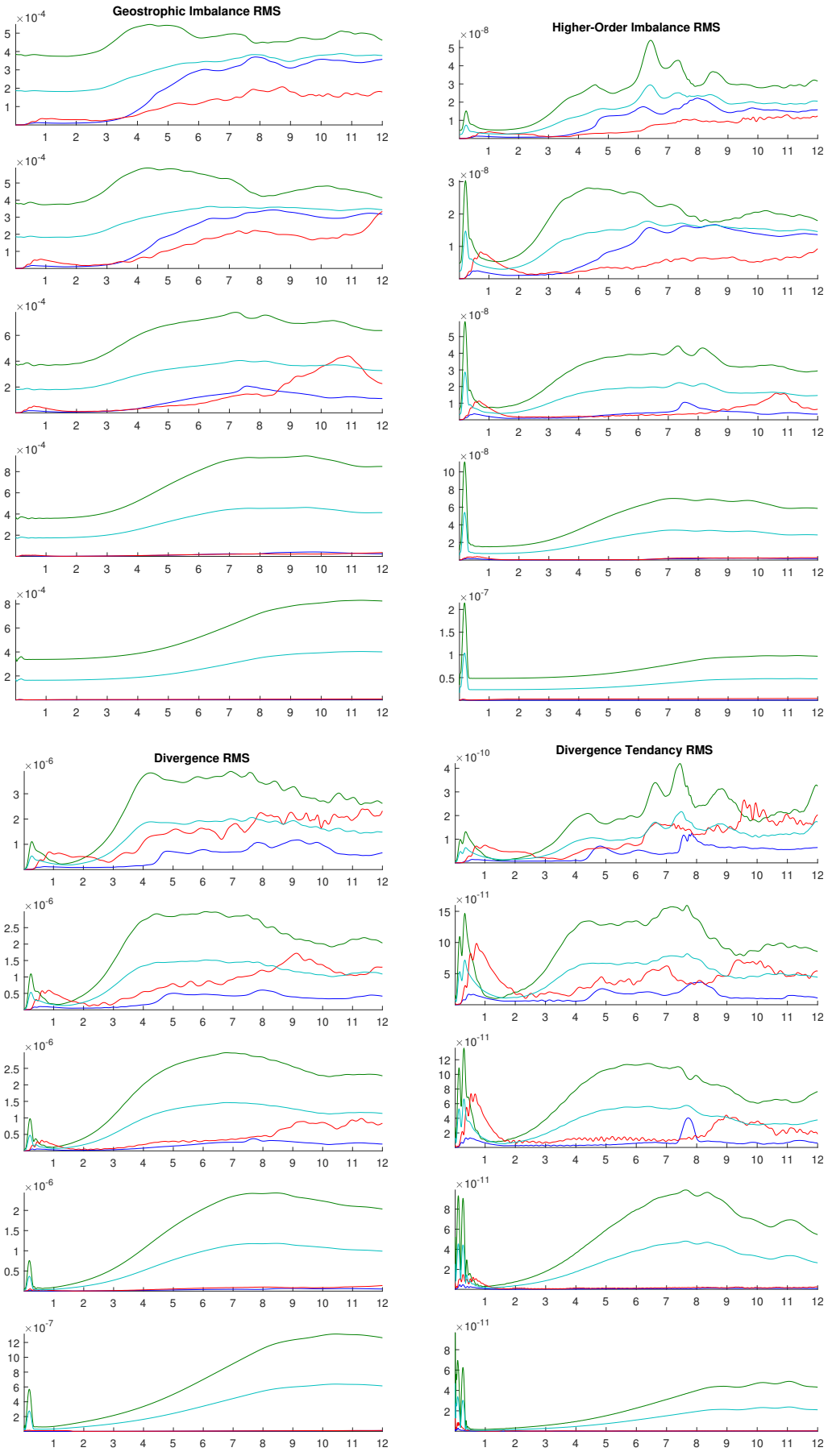

Figure A.8: Time series of diagnostic quantities approaching the semigeostrophic limit in the ENDGame model, $C_{gw} \approx 5$. For each subfigure the small parameter ε is, from top to bottom, = 1, 0.5, 0.25, 0.125, 0.0625.

Appendix B

Finite element model figures

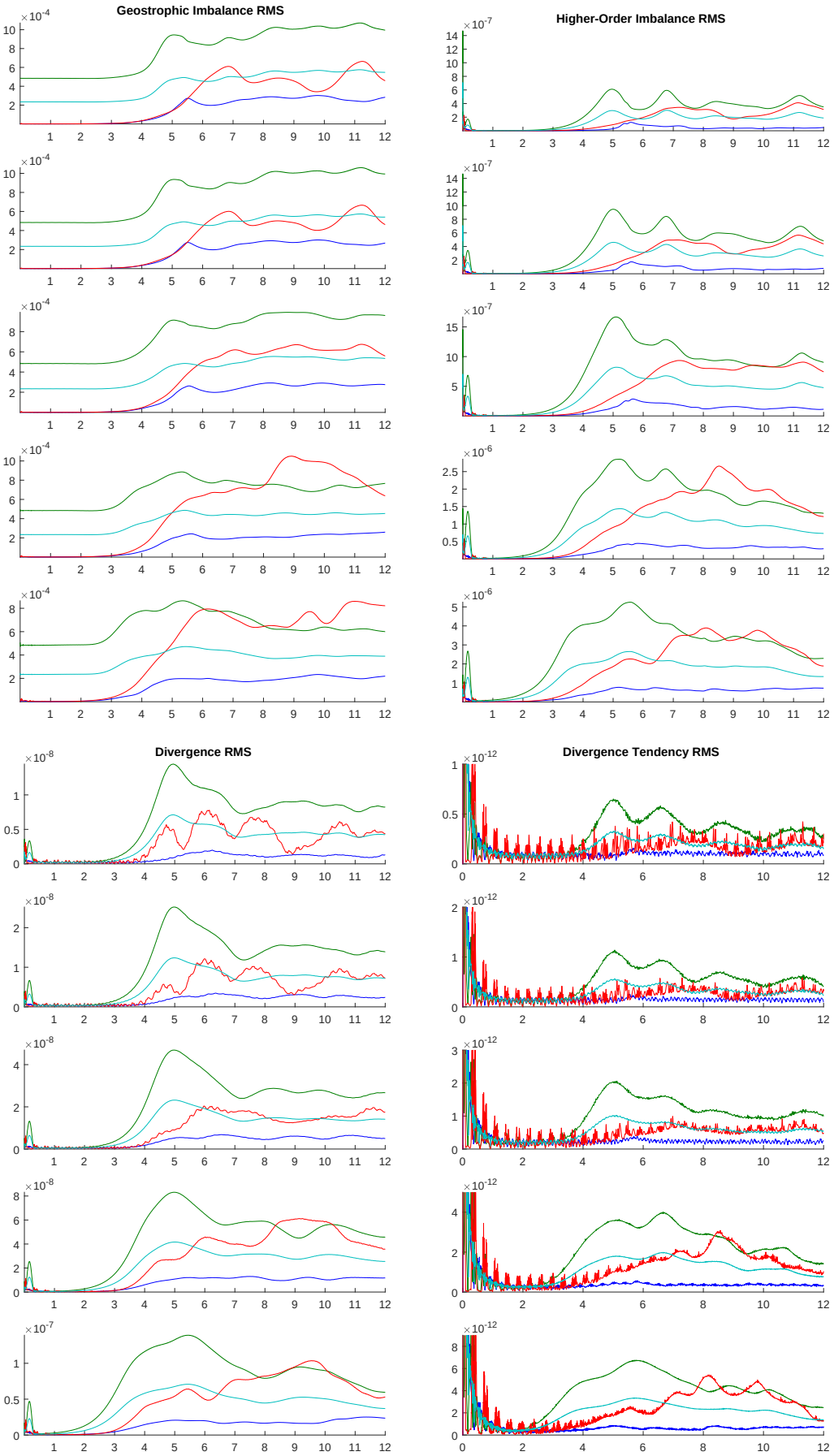

Figure B.1: Time series of diagnostic quantities approaching the strong rotation limit in the Hexagonal-Icosahedral FEM. For each subfigure the small parameter ε is, from top to bottom, = 1, 0.5, 0.25, 0.125, 0.0625.

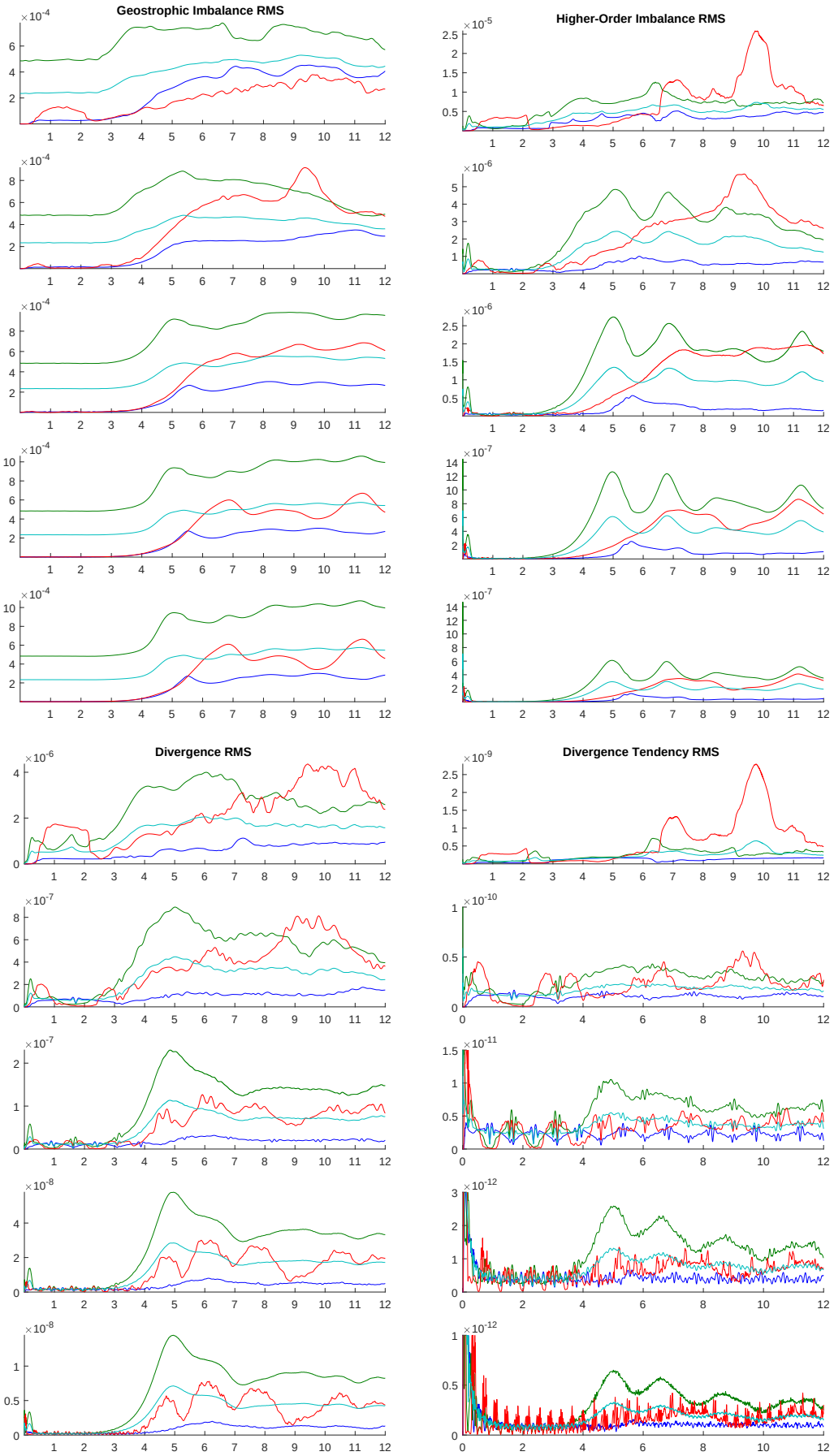

Figure B.2: Time series of diagnostic quantities approaching the barotropic vorticity limit in the Hexagonal-Icosahedral FEM. For each subfigure the small parameter ε is, from top to bottom, = 1, 0.5, 0.25, 0.125, 0.0625.

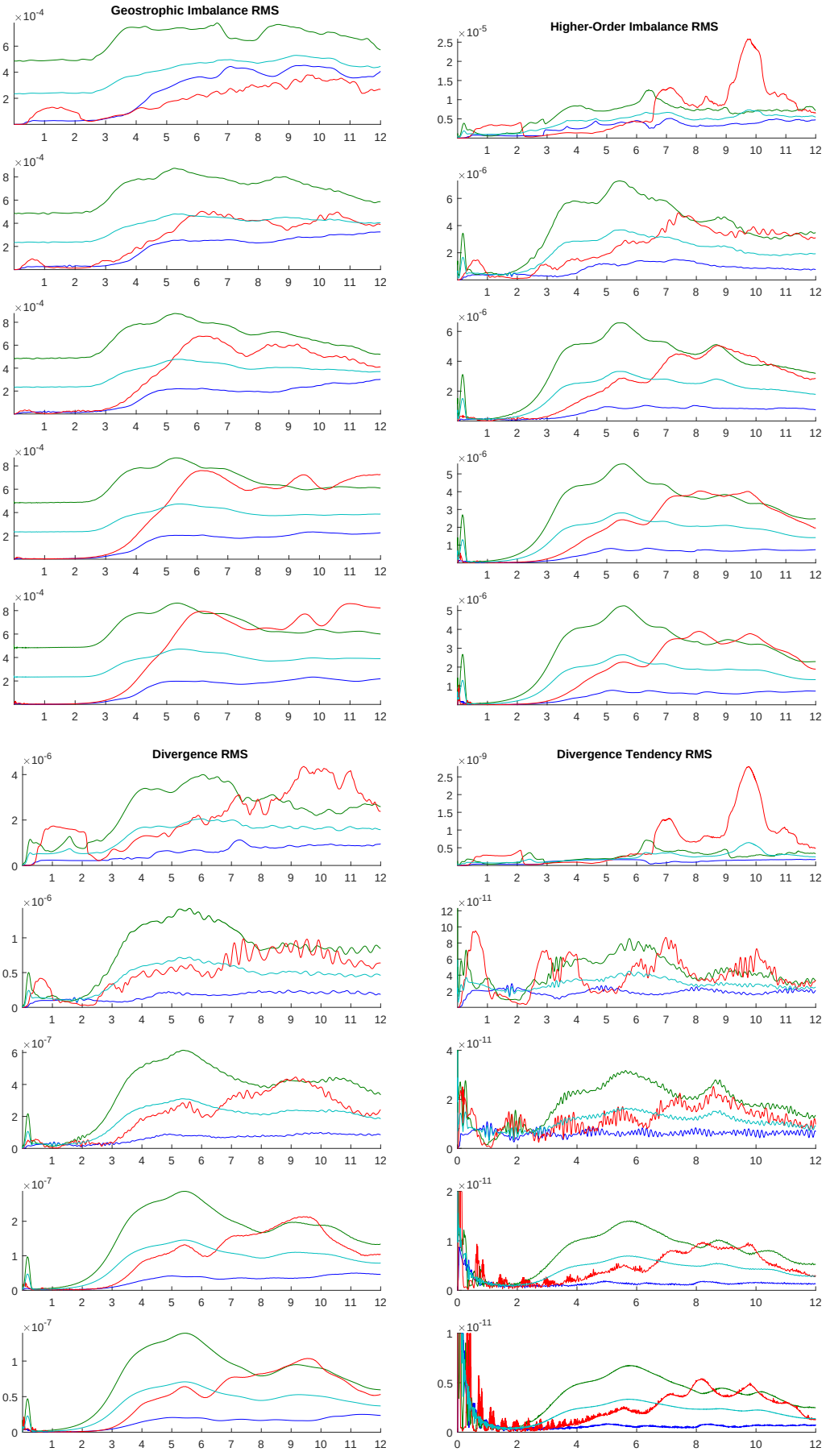

Figure B.3: Time series of diagnostic quantities approaching the quasigeostrophic limit in the Hexagonal-Icosahedral FEM. For each subfigure the small parameter ε is, from top to bottom, = 1, 0.5, 0.25, 0.125, 0.0625.

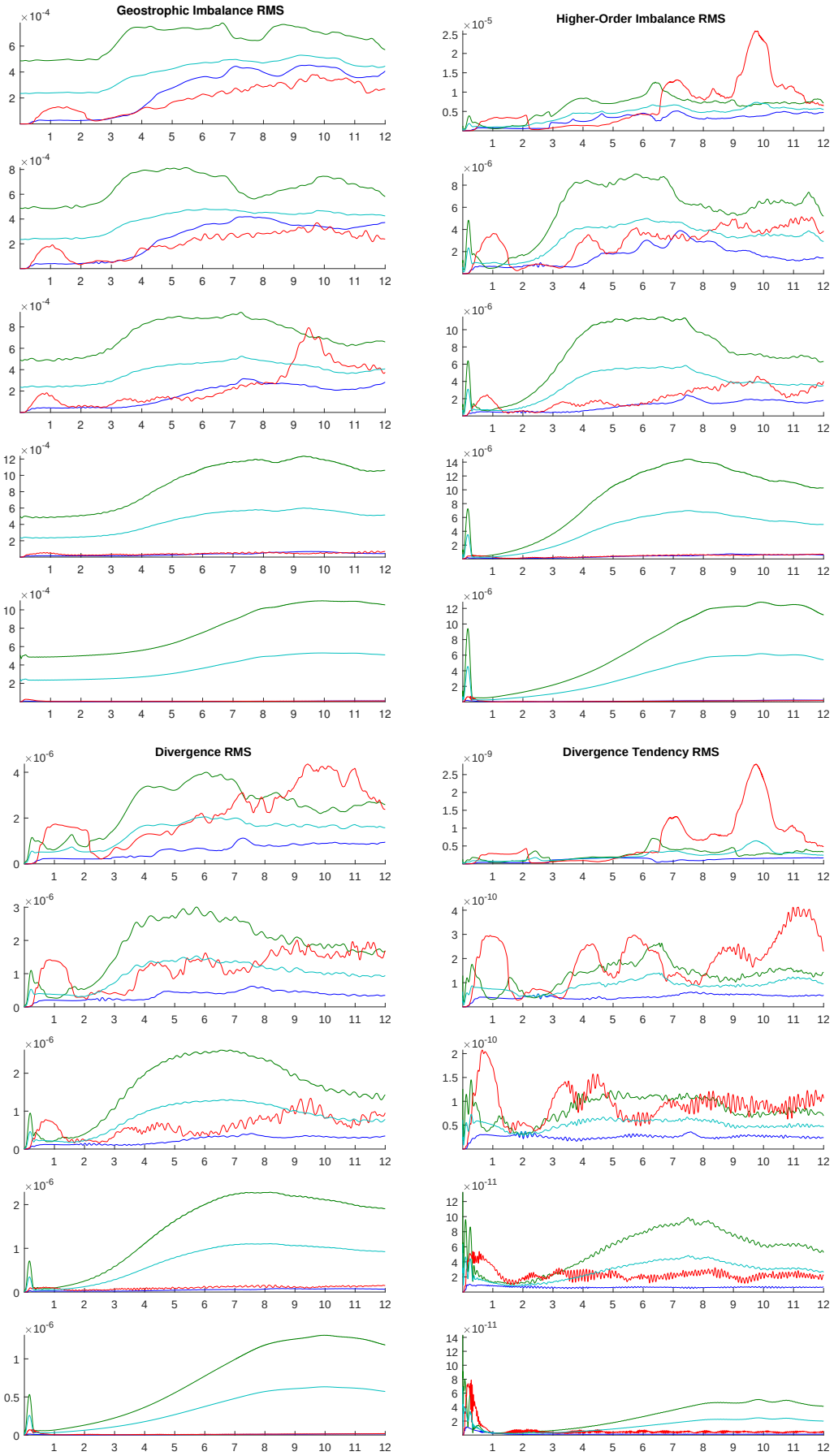

Figure B.4: Time series of diagnostic quantities approaching the semigeostrophic limit in the Hexagonal-Icosahedral FEM. For each subfigure the small parameter ε is, from top to bottom, = 1, 0.5, 0.25, 0.125, 0.0625.

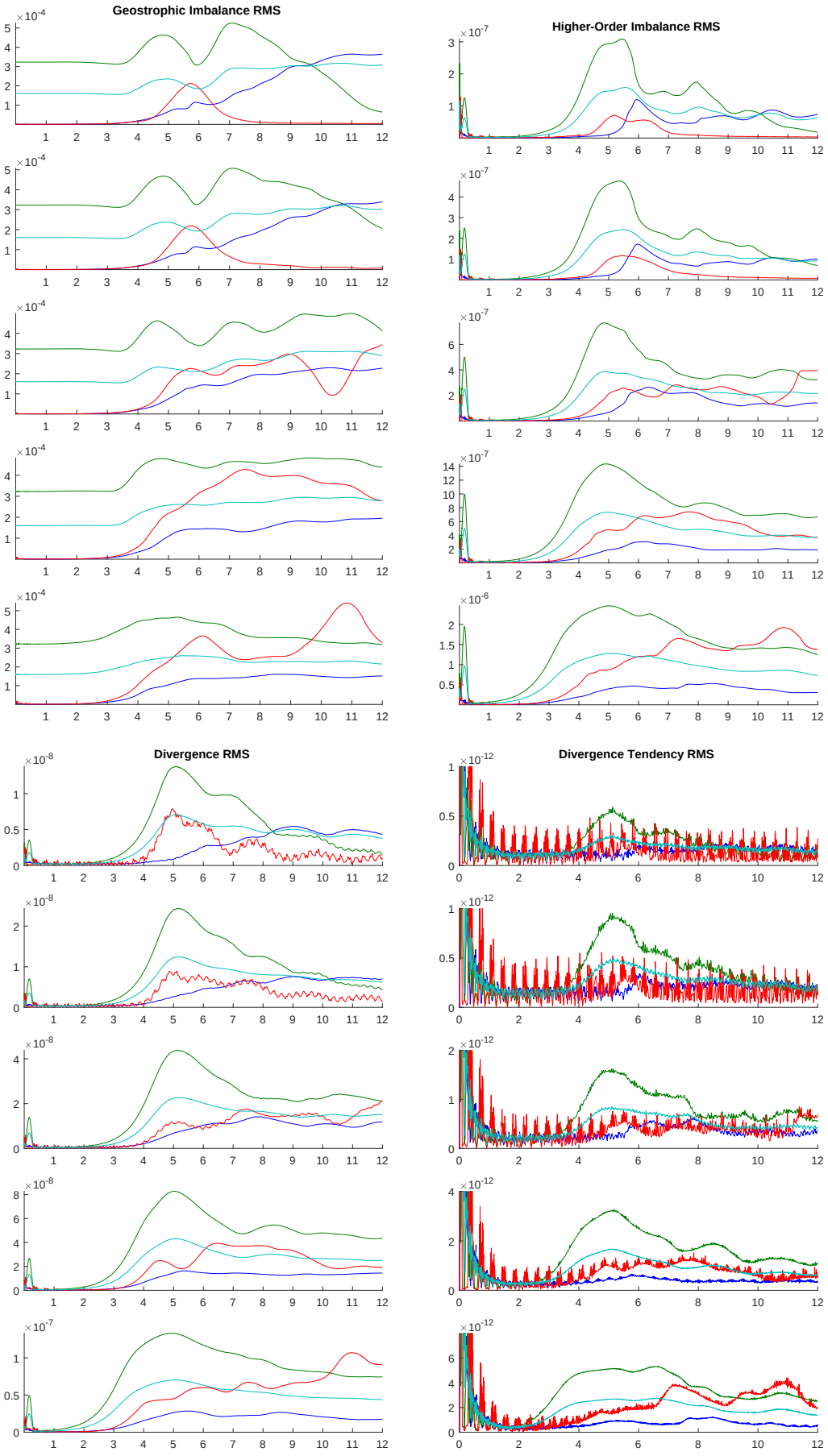

Figure B.5: Time series of diagnostic quantities approaching the strong rotation limit in the Cubed-Sphere FEM. For each subfigure the small parameter ε is, from top to bottom, = 1, 0.5, 0.25, 0.125, 0.0625.

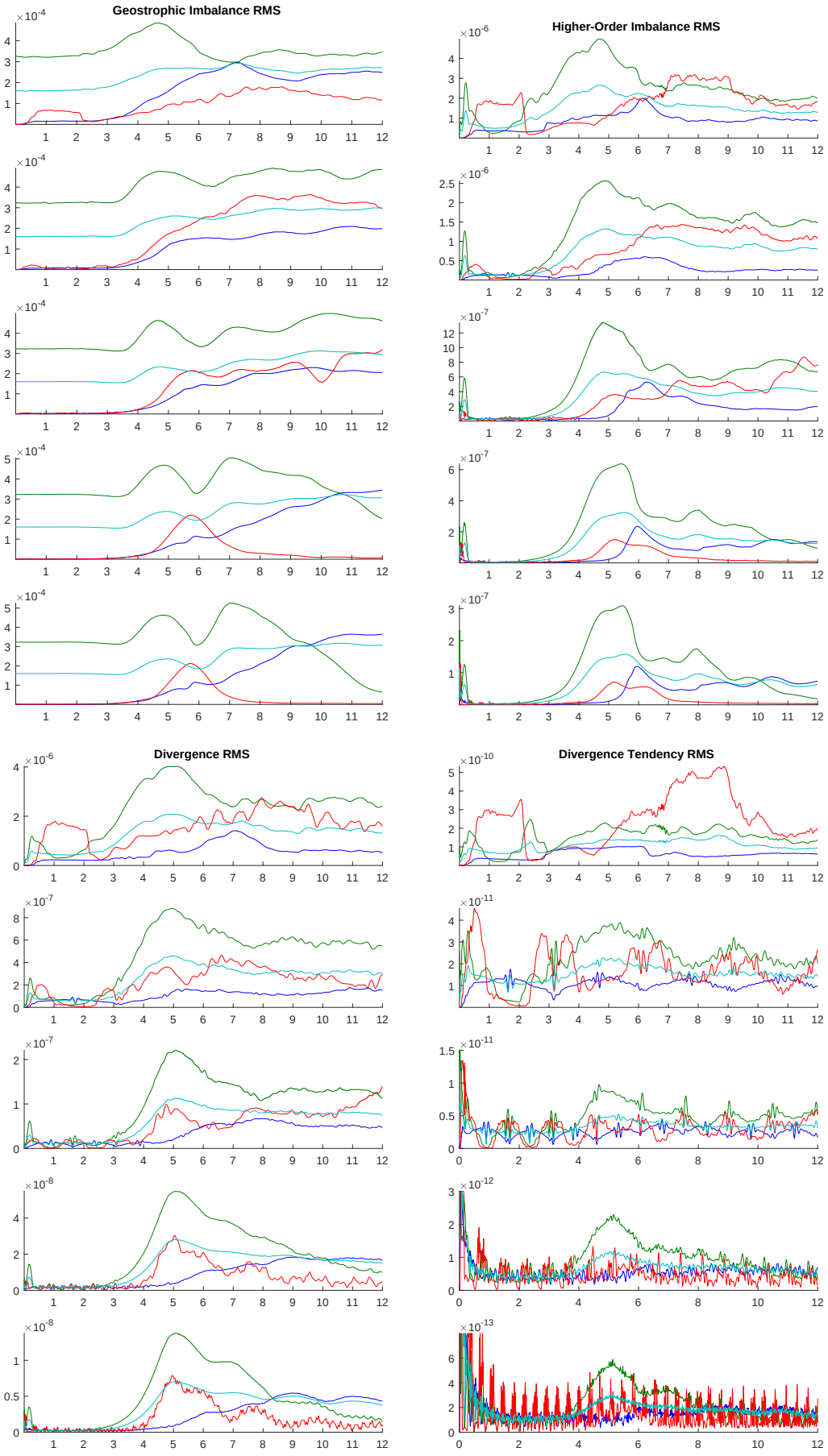

Figure B.6: Time series of diagnostic quantities approaching the barotropic limit limit in the Cubed-Sphere FEM. For each subfigure the small parameter ε is, from top to bottom, = 1, 0.5, 0.25, 0.125, 0.0625.

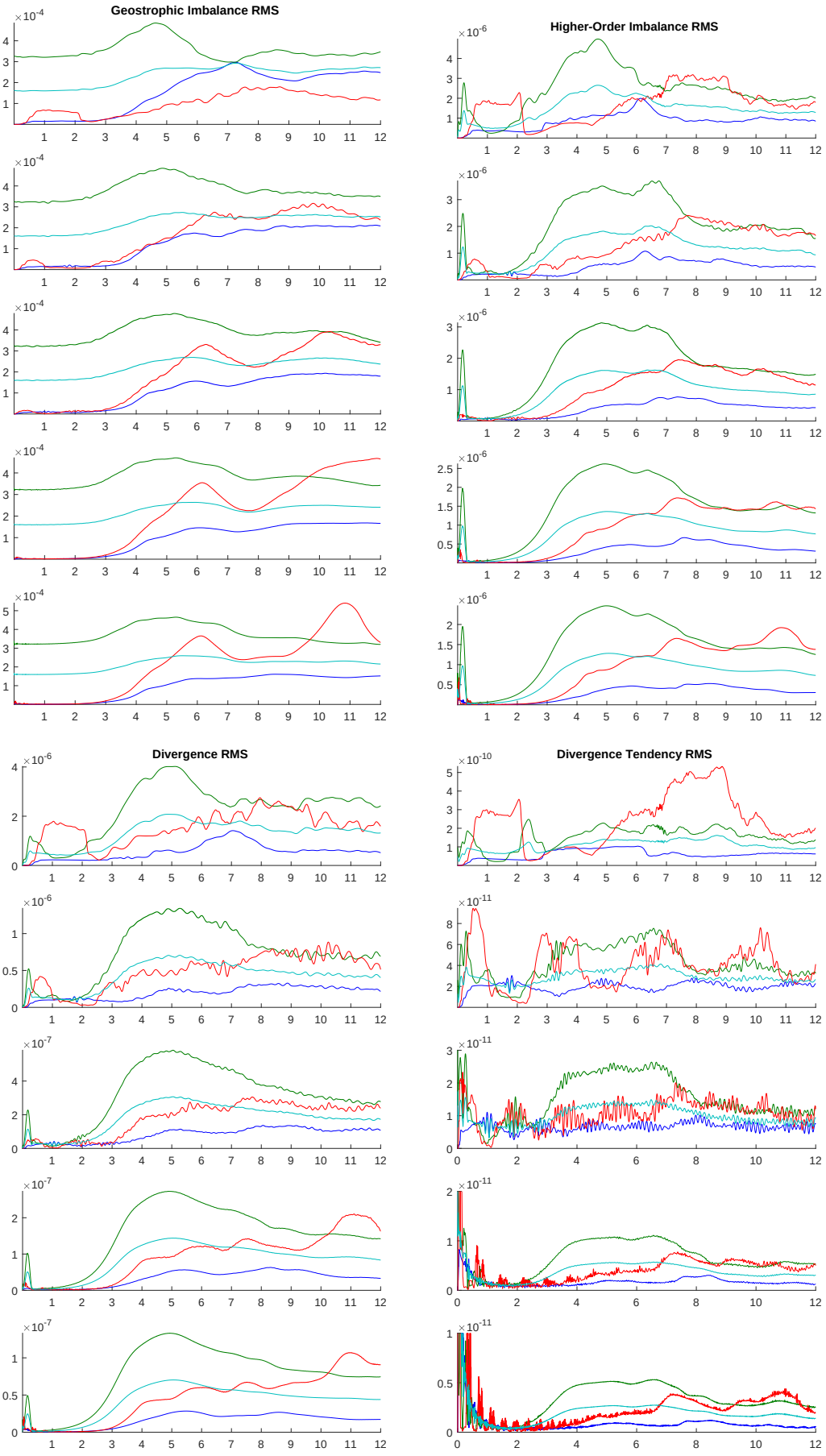

Figure B.7: Time series of diagnostic quantities approaching the quasigeostrophic limit in the Cubed-Sphere FEM. For each subfigure the small parameter ε is, from top to bottom, = 1, 0.5, 0.25, 0.125, 0.0625.

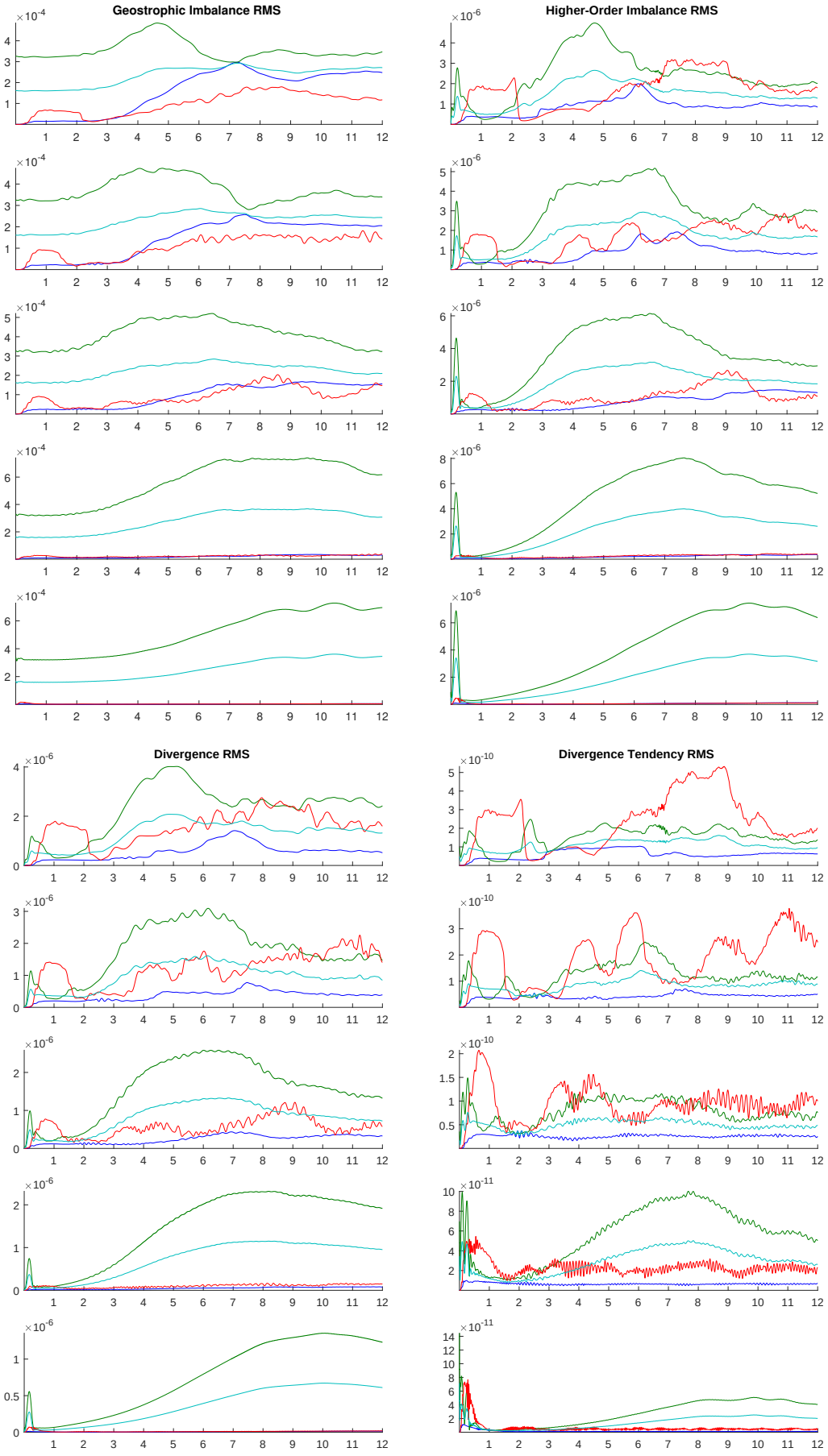

Figure B.8: Time series of diagnostic quantities approaching the semigeostrophic limit in the Cubed-Sphere FEM. For each subfigure the small parameter ε is, from top to bottom, = 1, 0.5, 0.25, 0.125, 0.0625.

Appendix C

Subgrid interaction figures

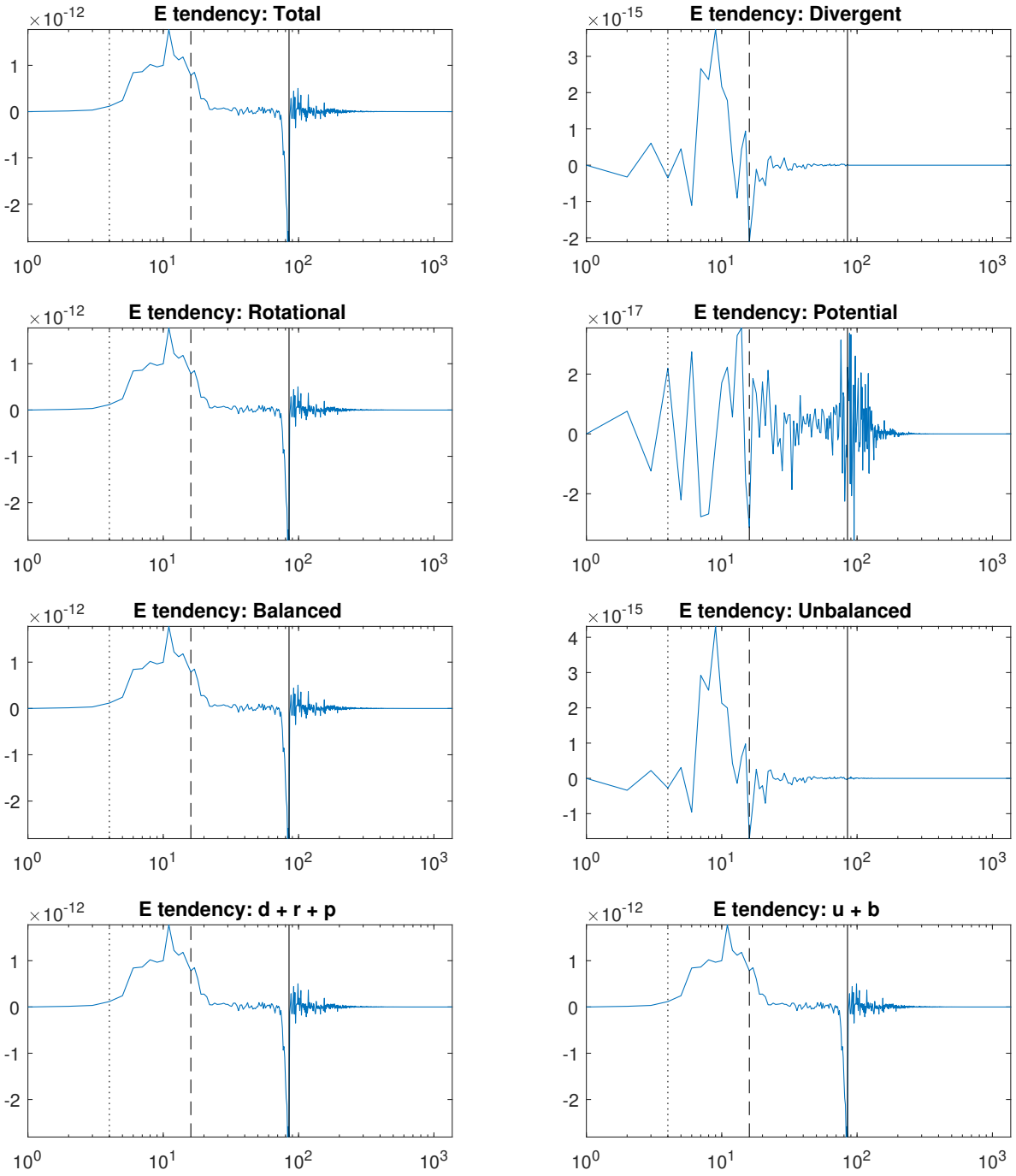

Figure C.1: Difference in the energy tendency spectra resulting from truncation of all fields in the balanced regime. The dotted line indicates the Rossby deformation scale, the dashed line indicates the forcing scale and the solid line indicates the dissipation scale.

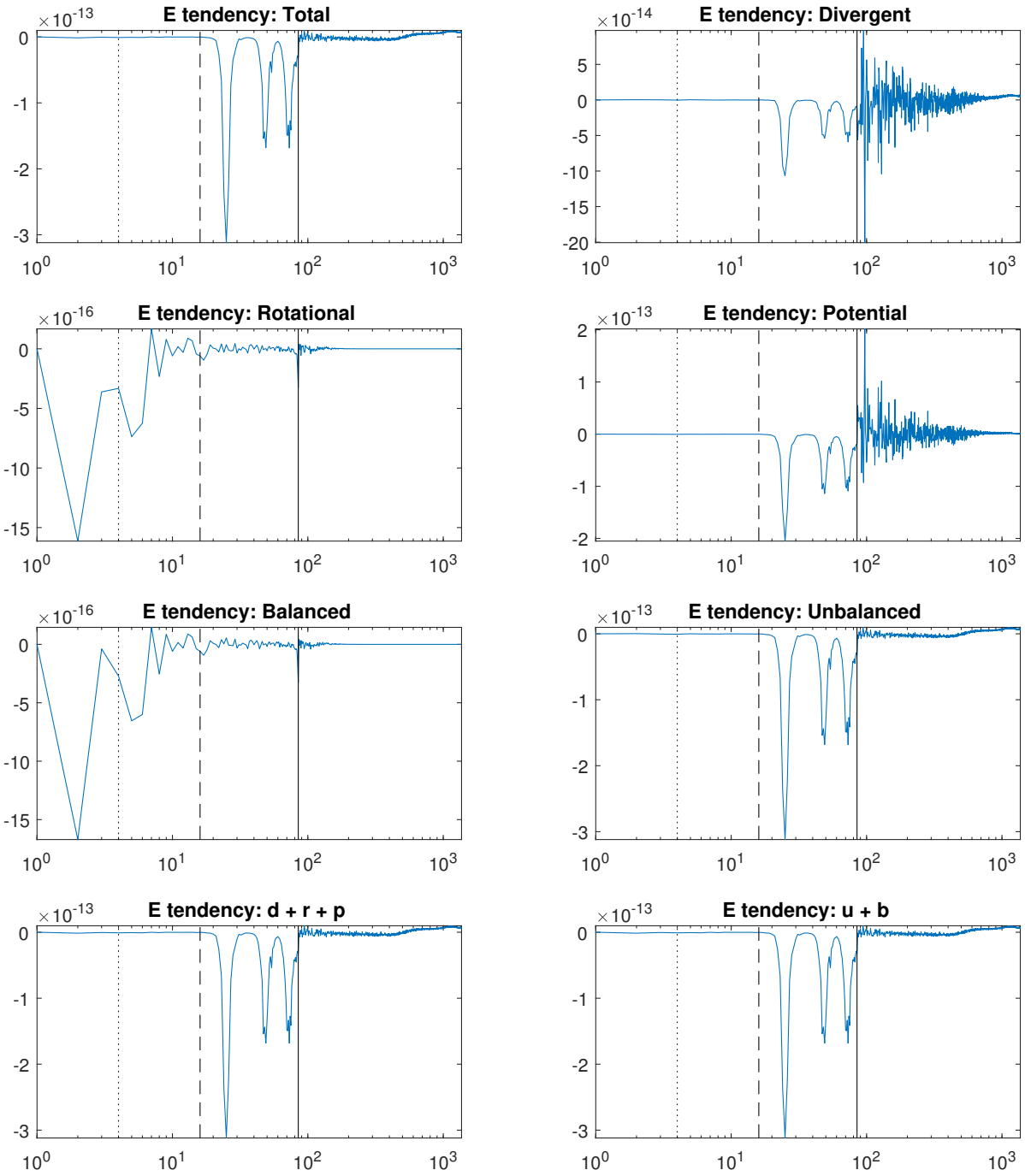

Figure C.2: Difference in the energy tendency spectra resulting from truncation of all fields in the unbalanced regime

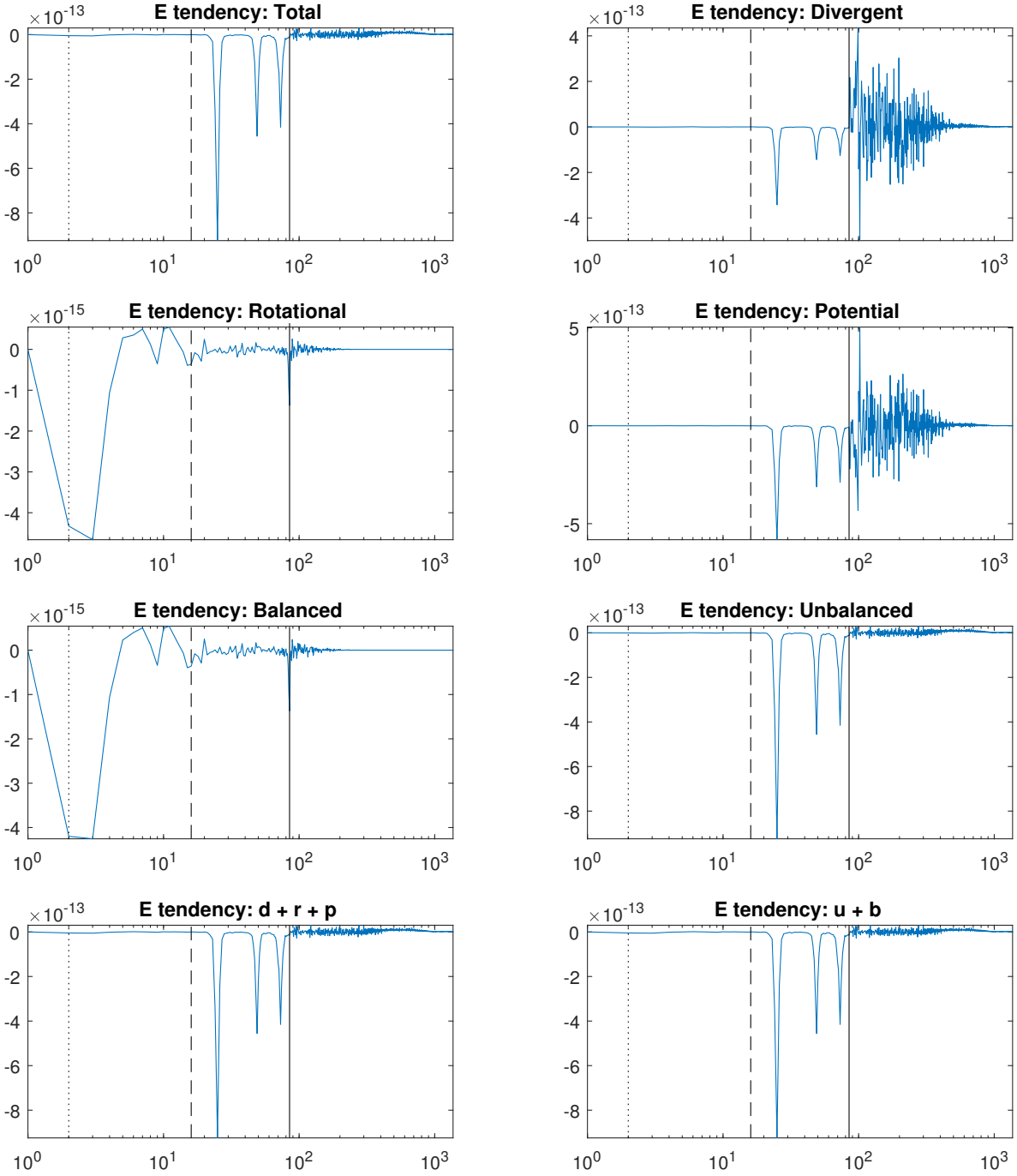

Figure C.3: Difference in the energy tendency spectra resulting from truncation of all fields in the shock regime

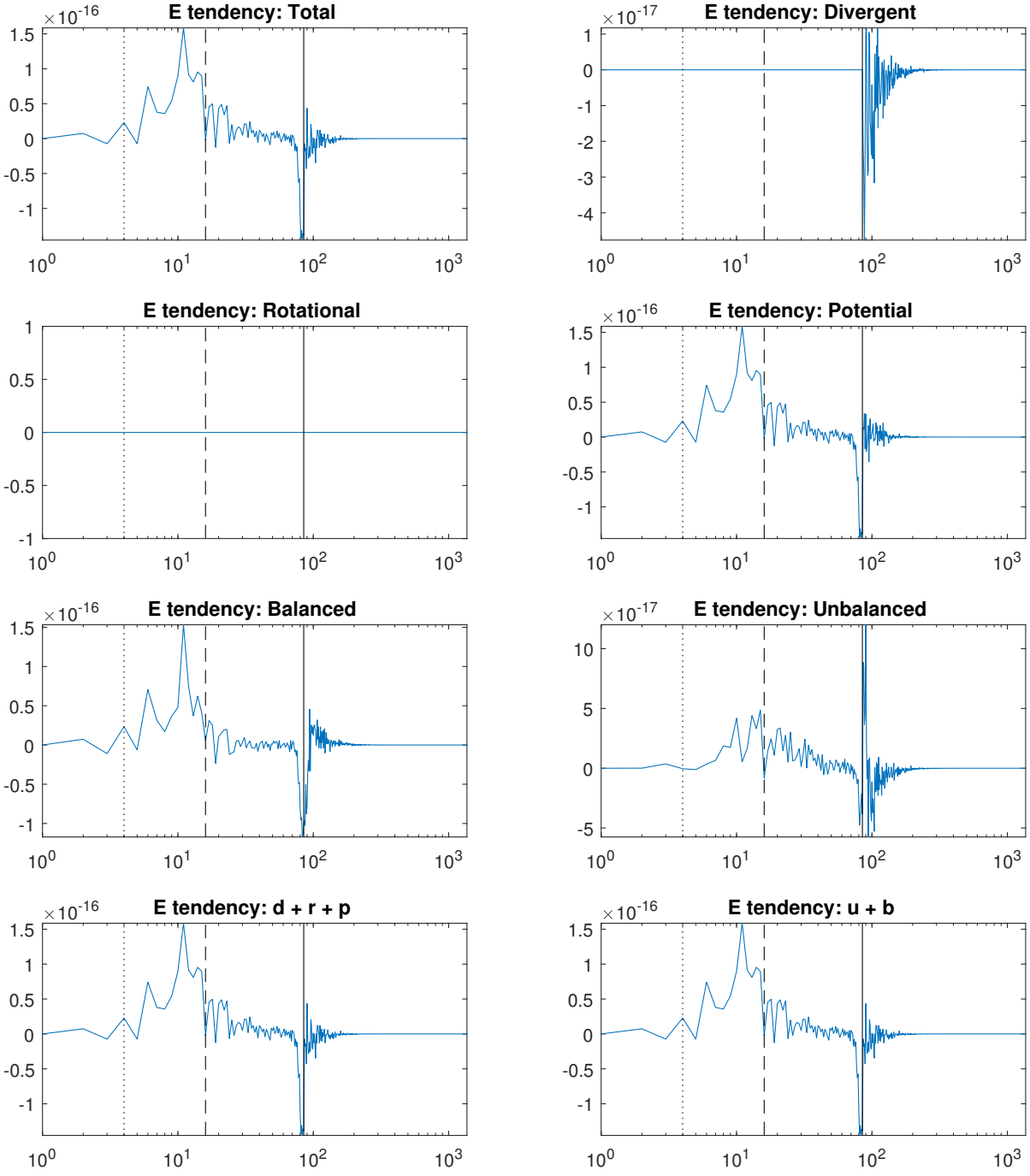

Figure C.4: Difference in the energy tendency spectra resulting from truncation the geopotential in the balanced regime

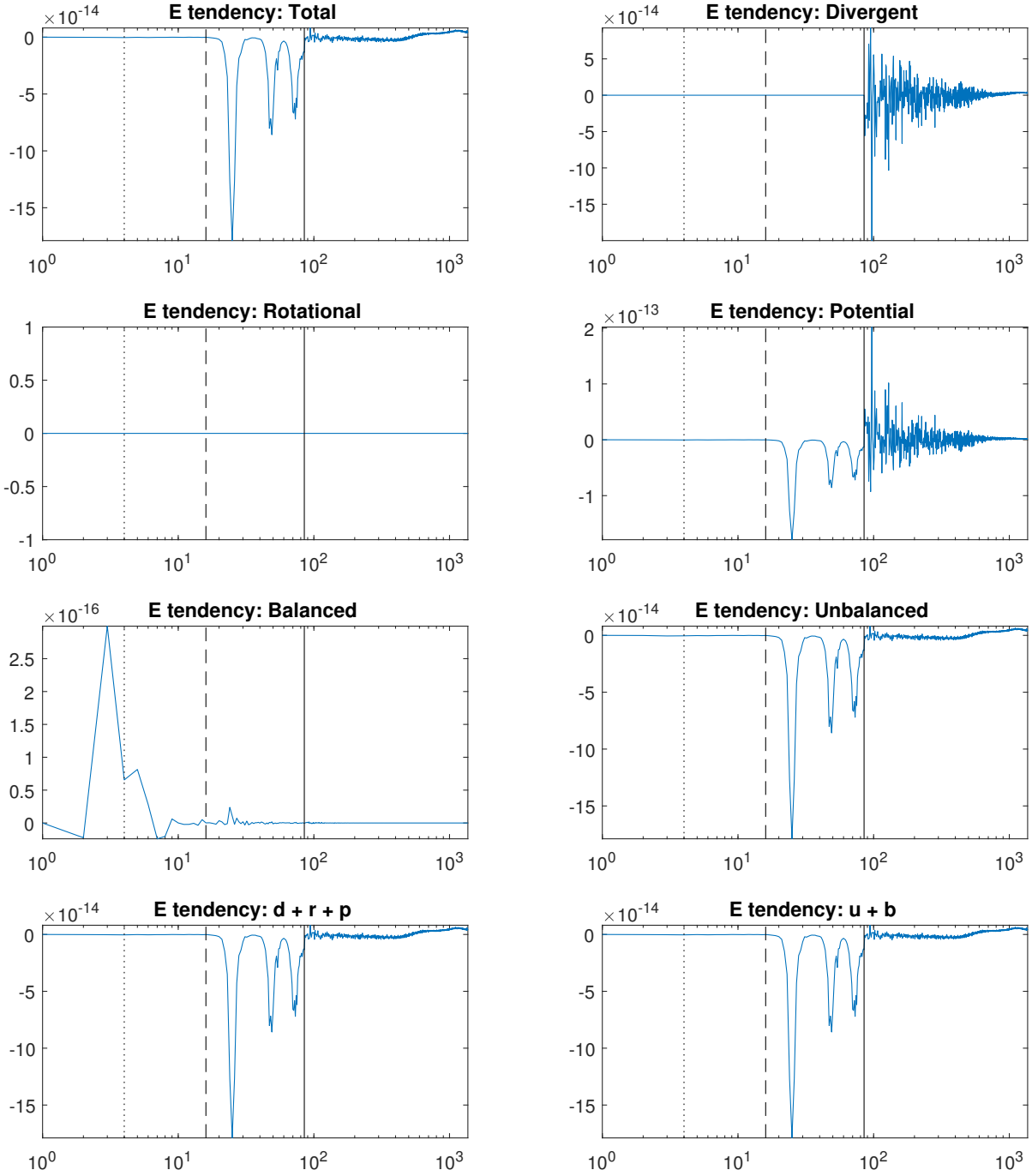

Figure C.5: Difference in the energy tendency spectra resulting from truncation the geopotential in the unbalanced regime

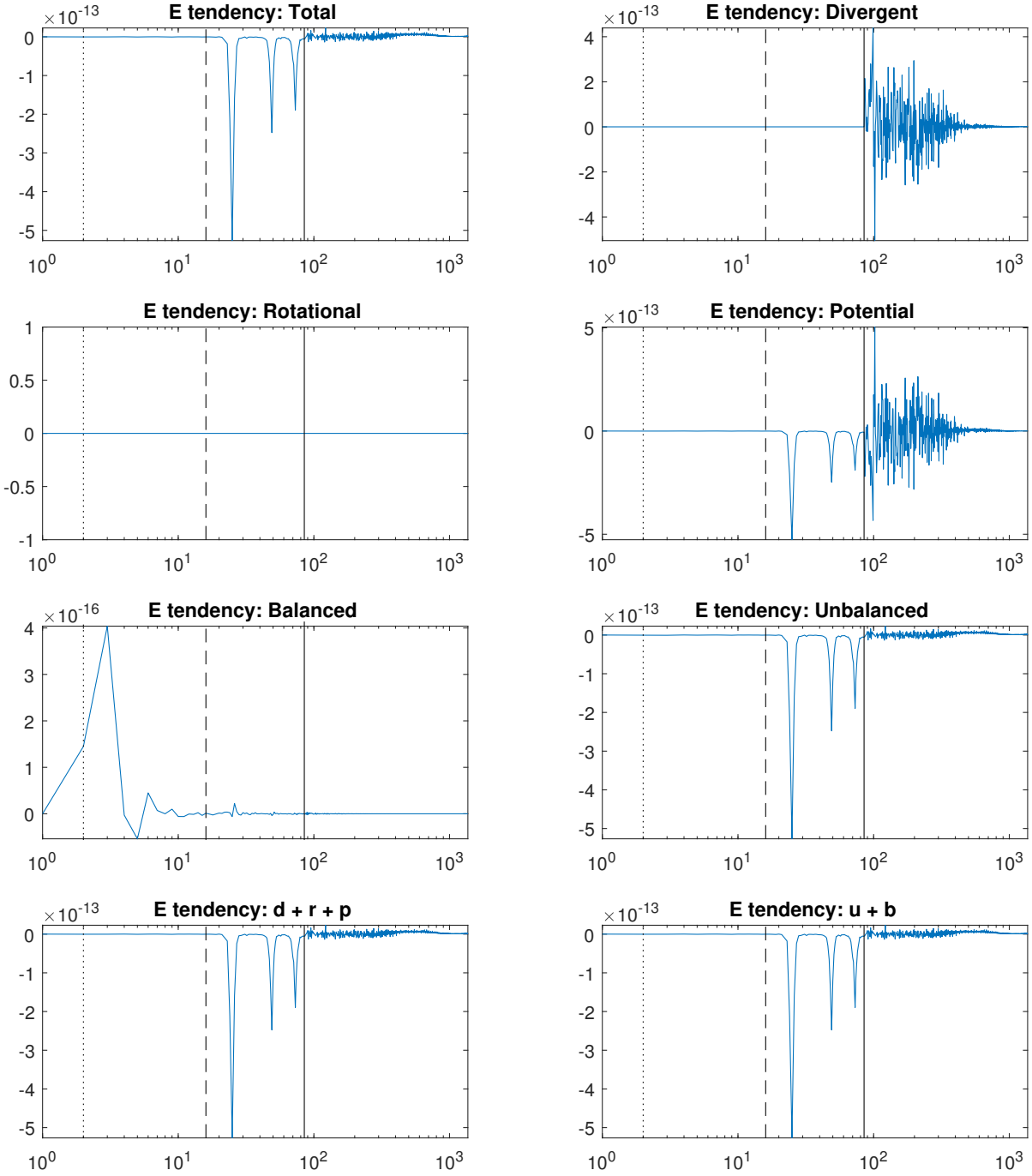

Figure C.6: Difference in the energy tendency spectra resulting from truncation the geopotential in the shock regime

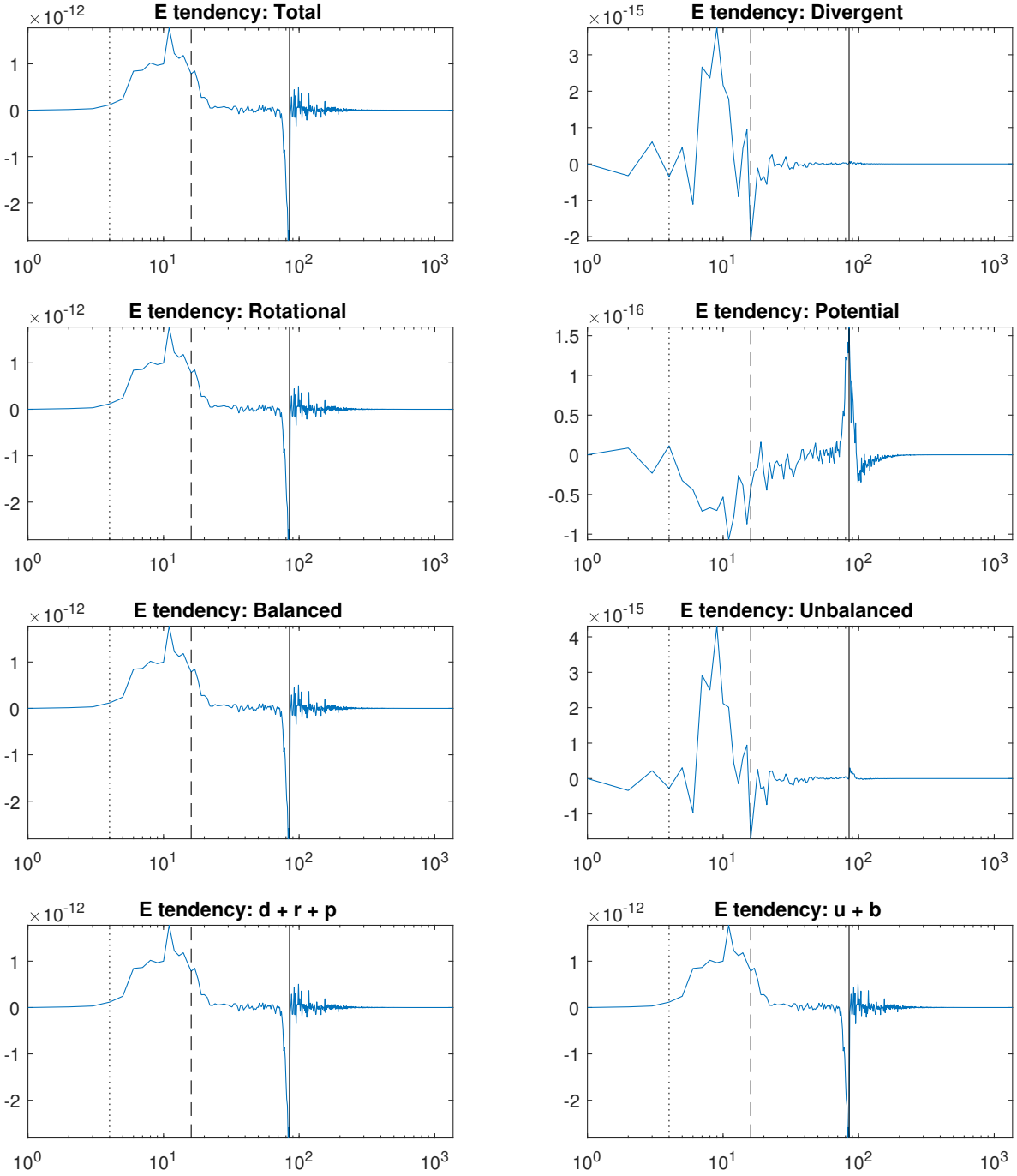

Figure C.7: Difference in the energy tendency spectra resulting from truncation of the vorticity in the balanced regime

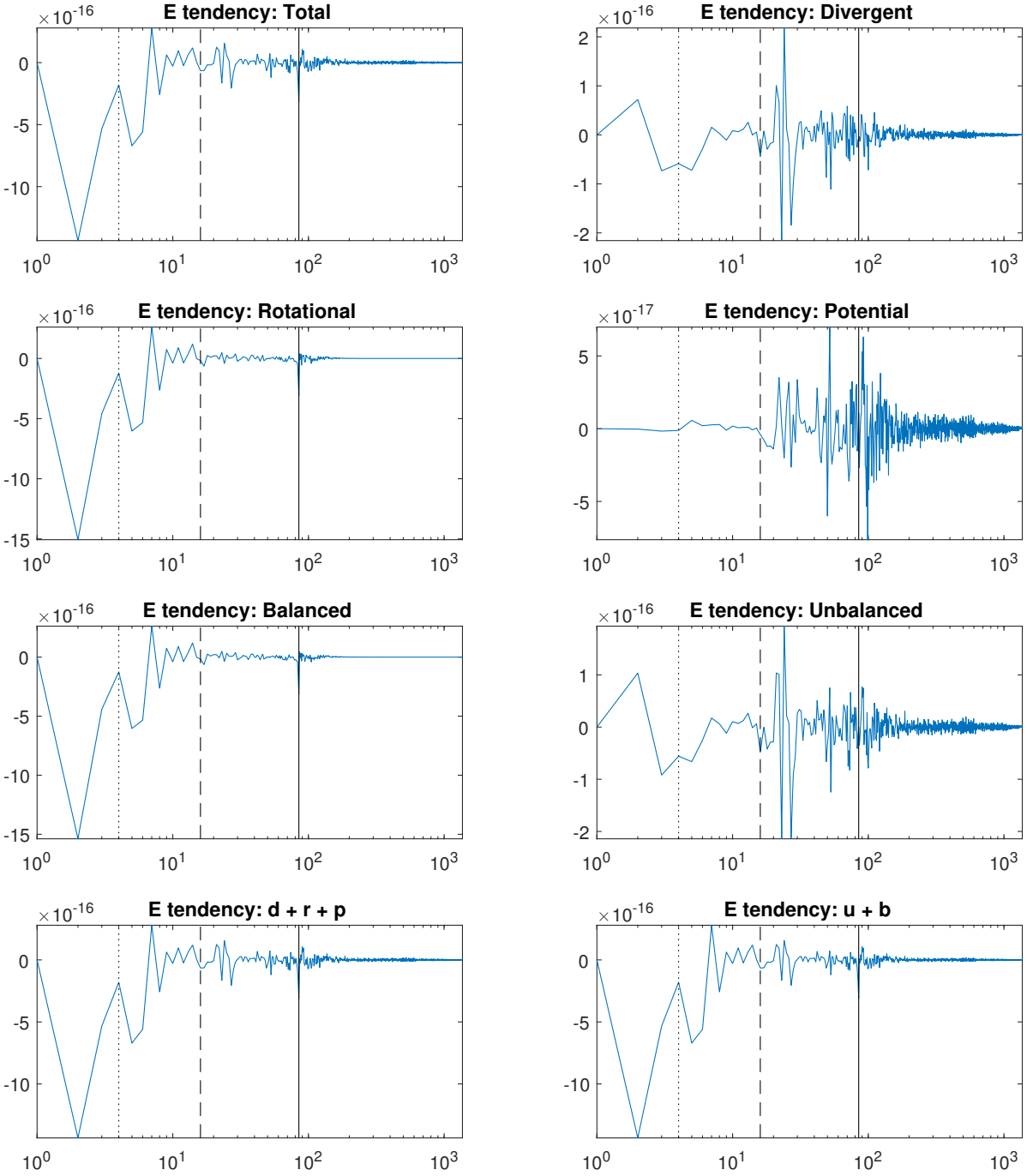

Figure C.8: Difference in the energy tendency spectra resulting from truncation of the vorticity in the unbalanced regime

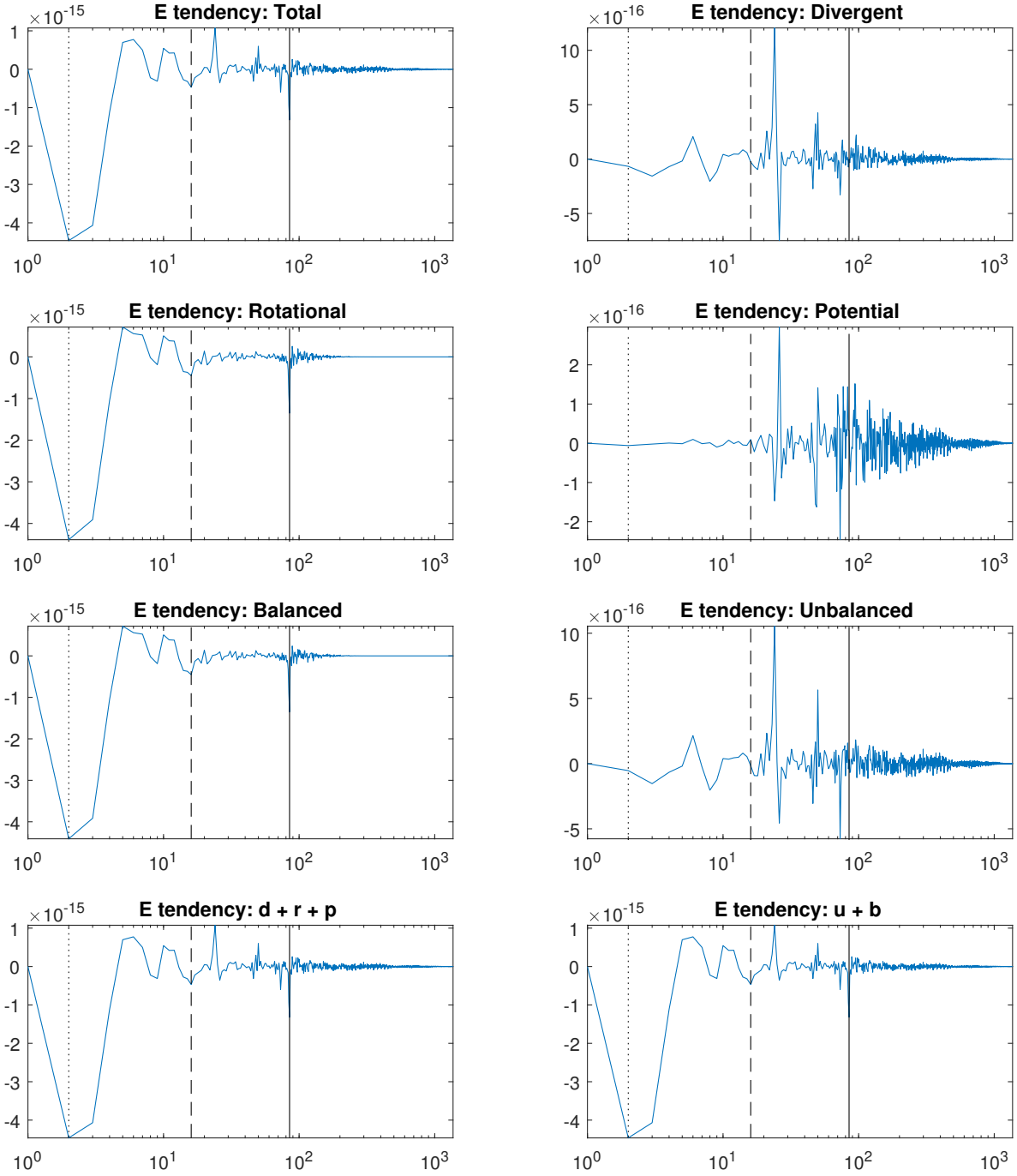

Figure C.9: Difference in the energy tendency spectra resulting from truncation of the vorticity in the shock regime

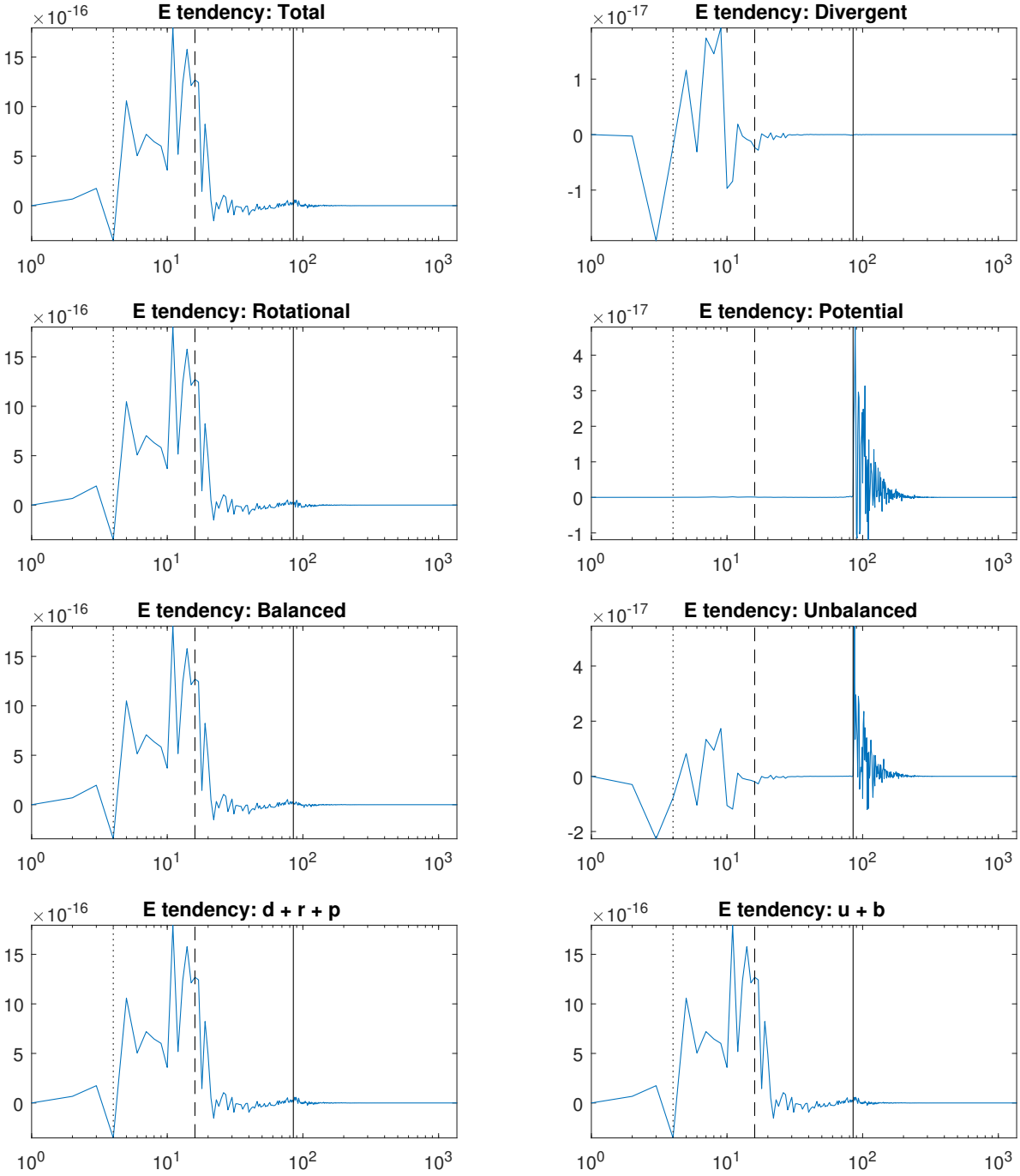

Figure C.10: Difference in the energy tendency spectra resulting from truncation of the divergence in the balanced regime

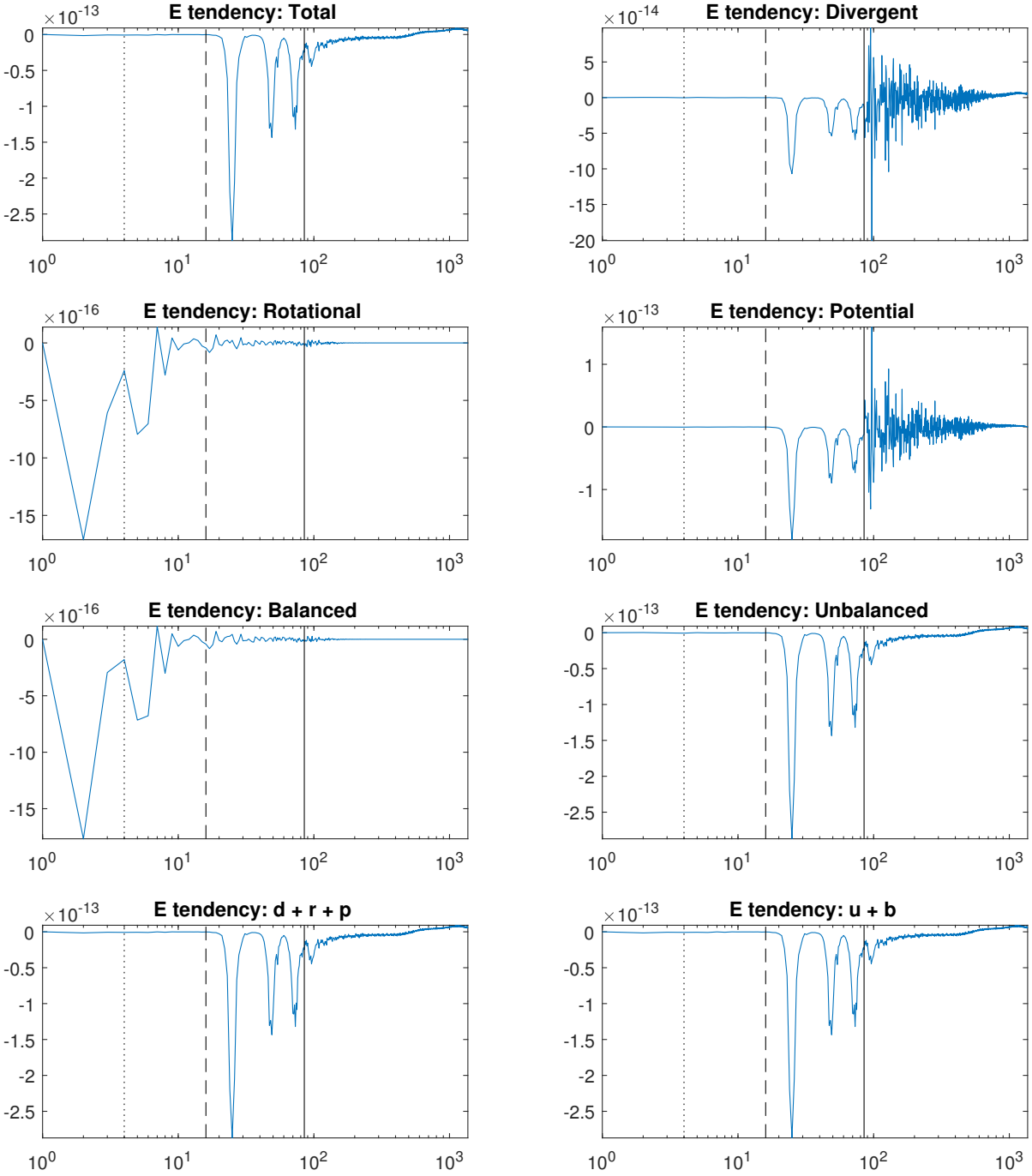

Figure C.11: Difference in the energy tendency spectra resulting from truncation of the divergence in the unbalanced regime

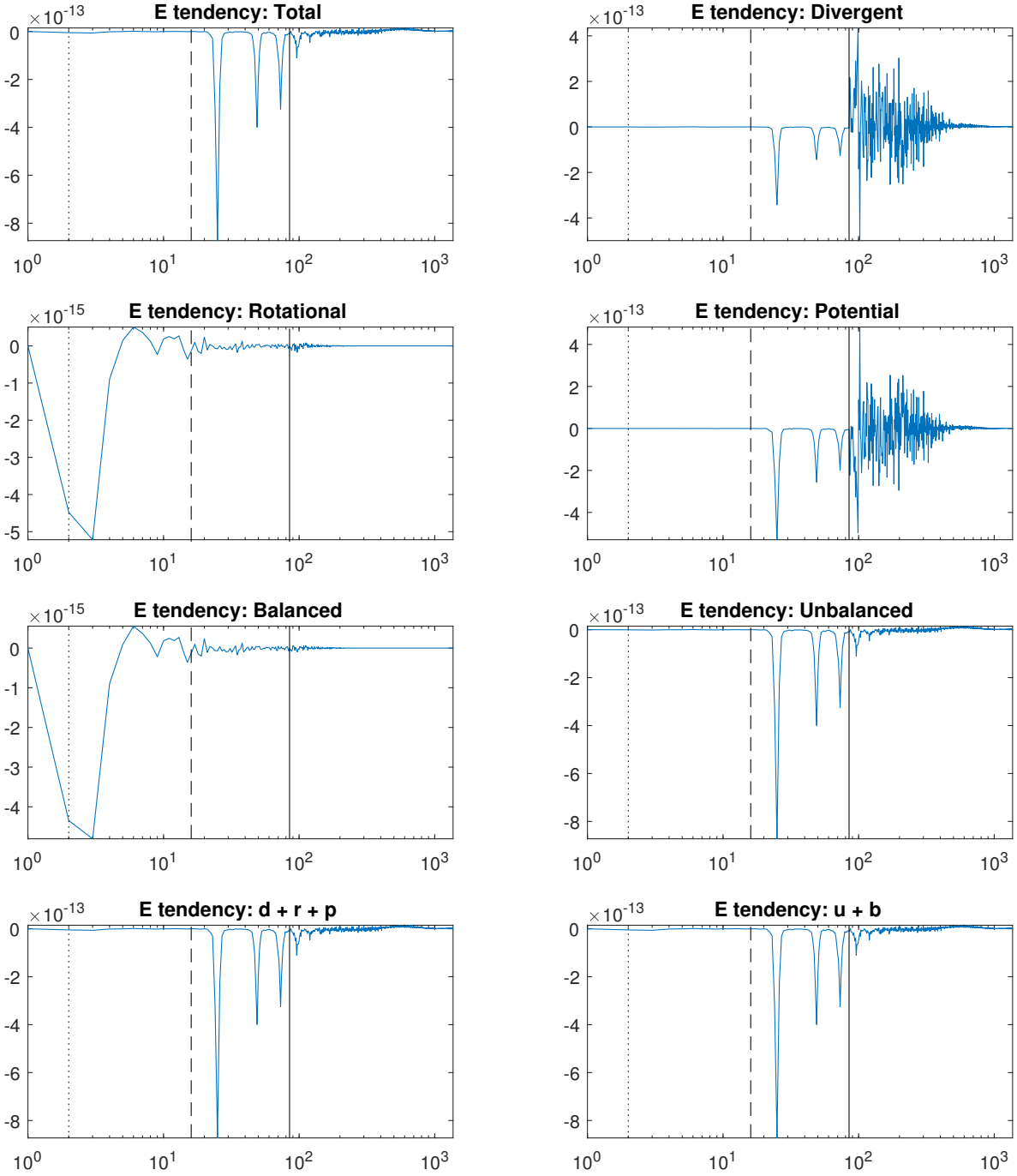

Figure C.12: Difference in the energy tendency spectra resulting from truncation of the divergence in the shock regime

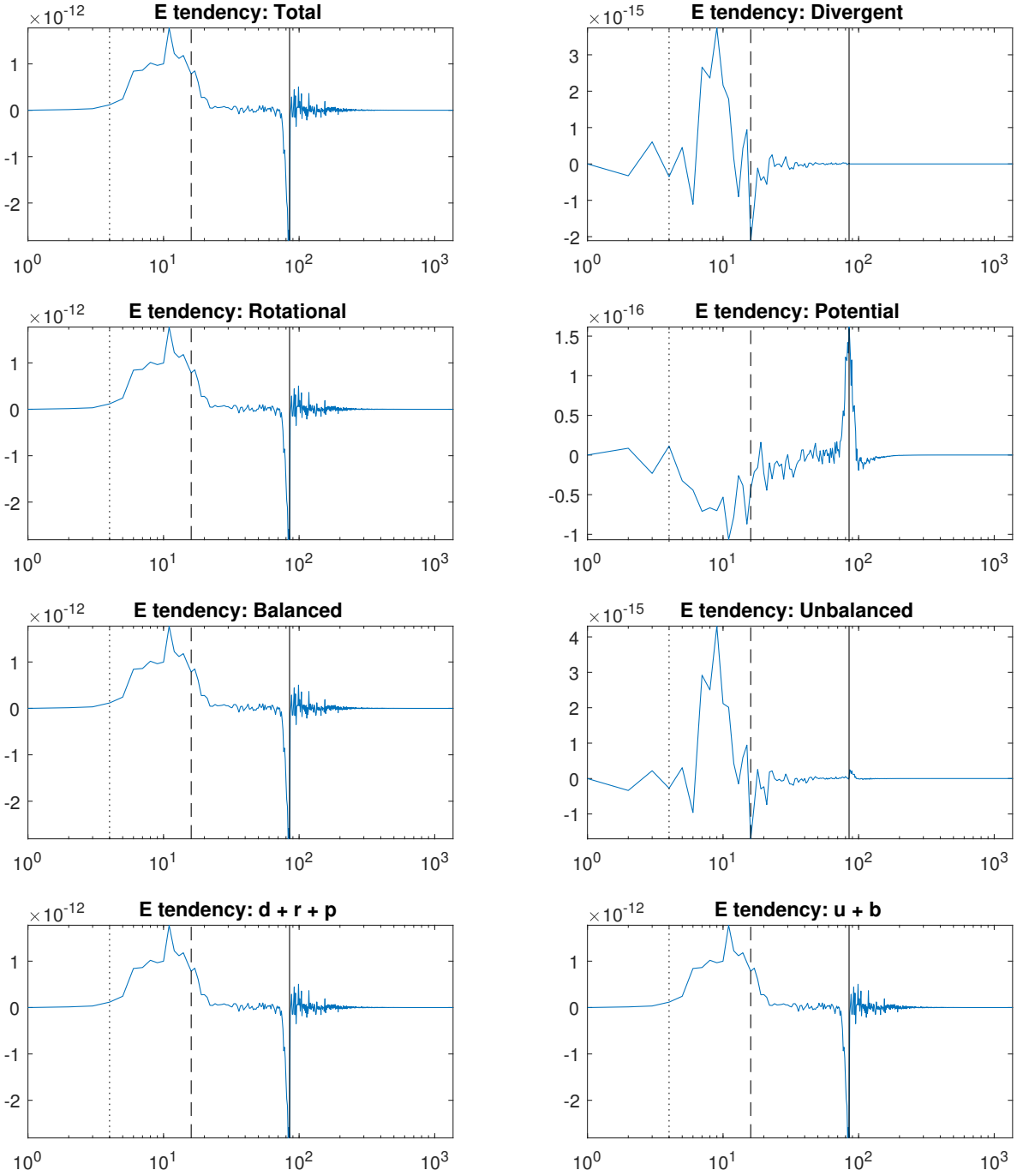

Figure C.13: Difference in the energy tendency spectra resulting from truncation of the kinetic energy in the balanced regime

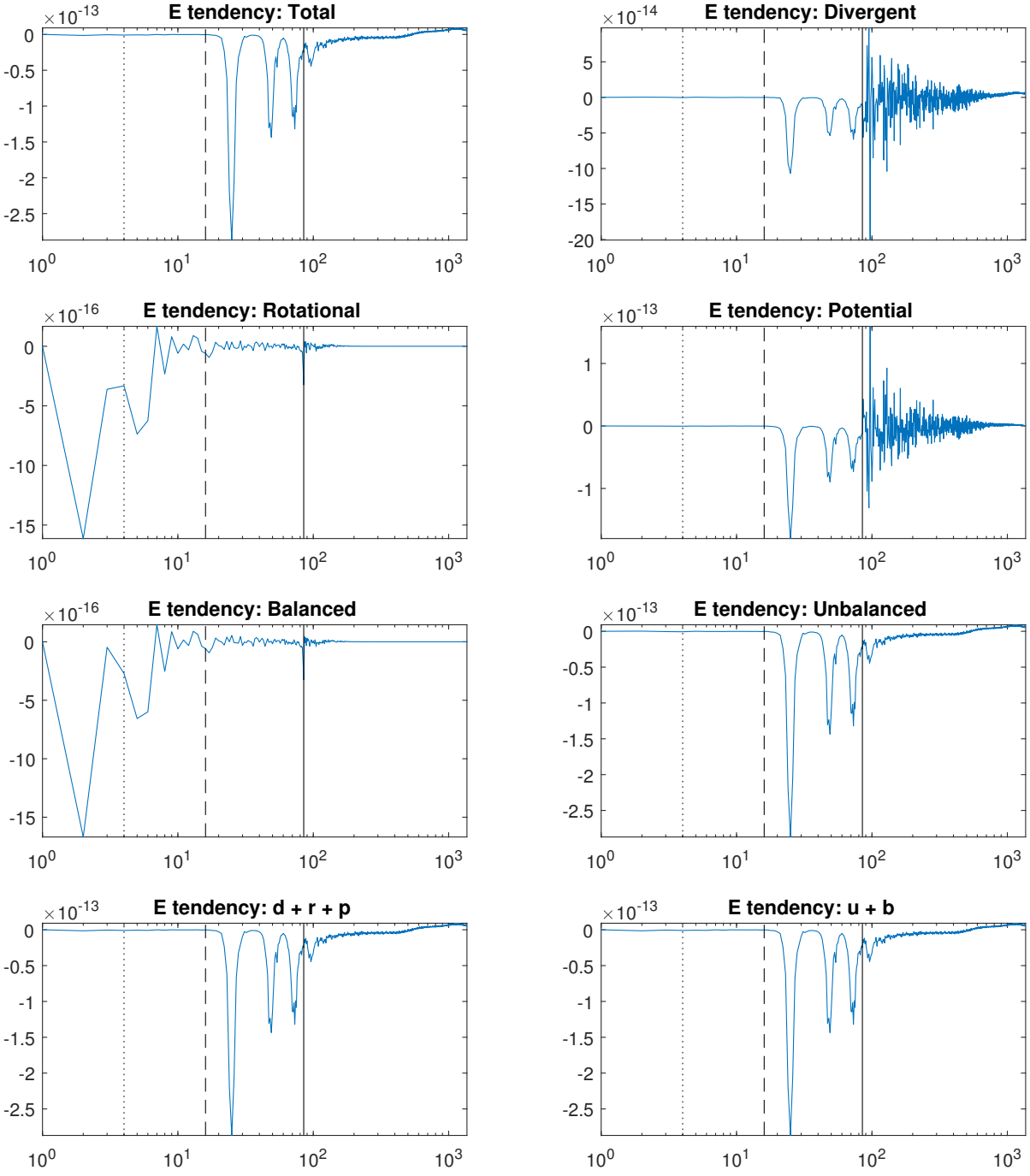

Figure C.14: Difference in the energy tendency spectra resulting from truncation of the kinetic energy in the unbalanced regime

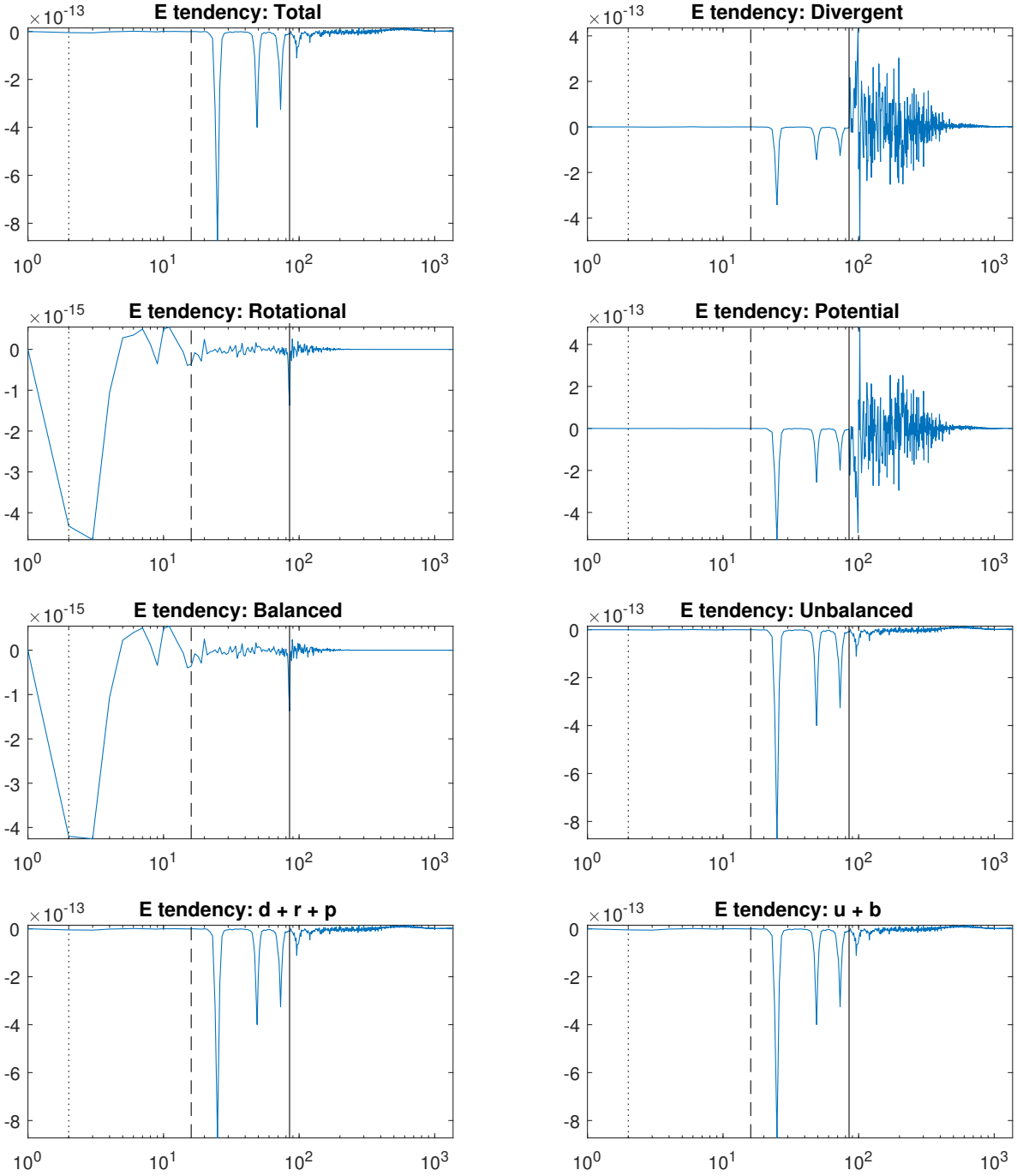

Figure C.15: Difference in the energy tendency spectra resulting from truncation of the kinetic energy in the shock regime

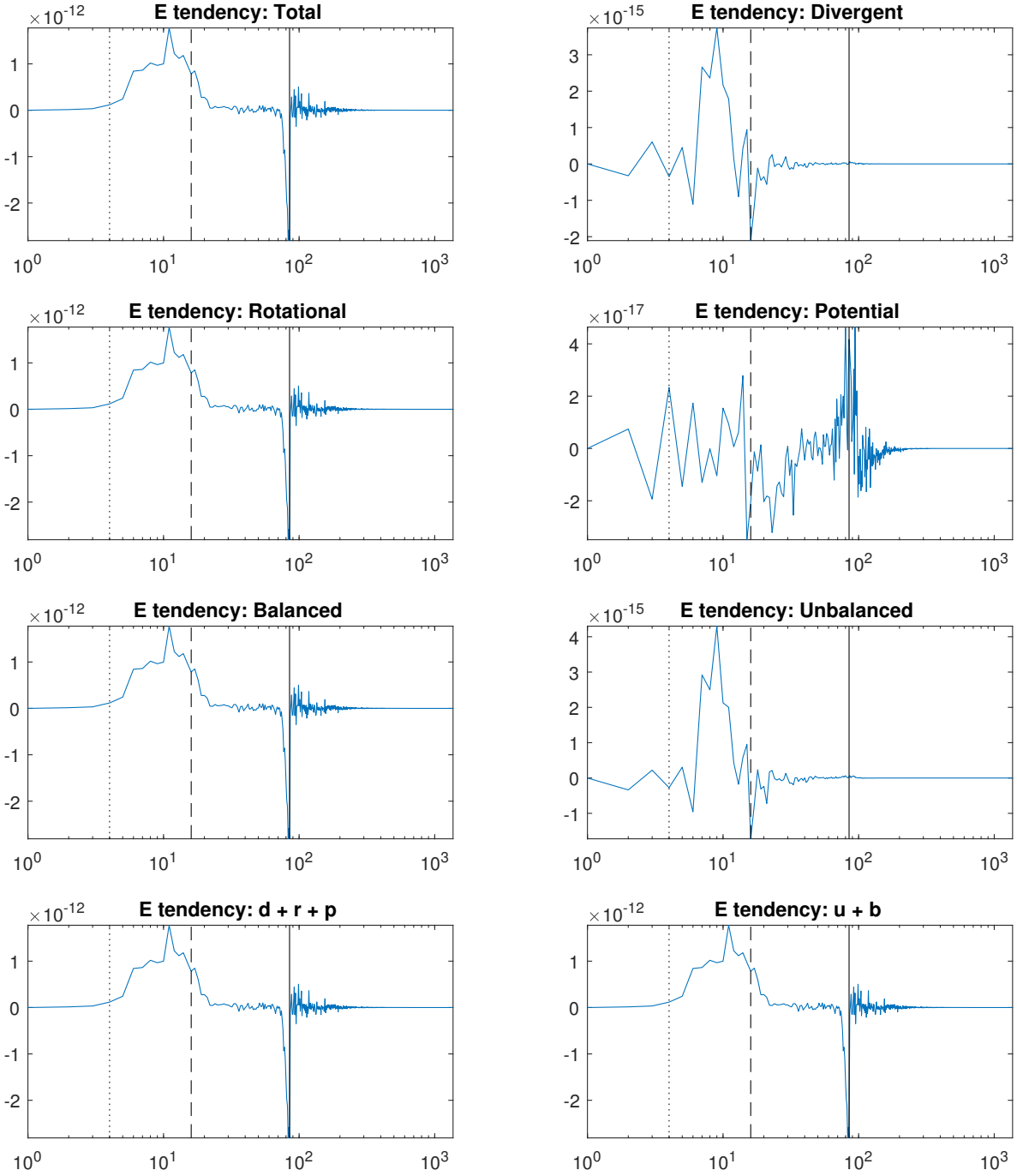

Figure C.16: Difference in the energy tendency spectra resulting from truncation of the potential vorticity in the balanced regime

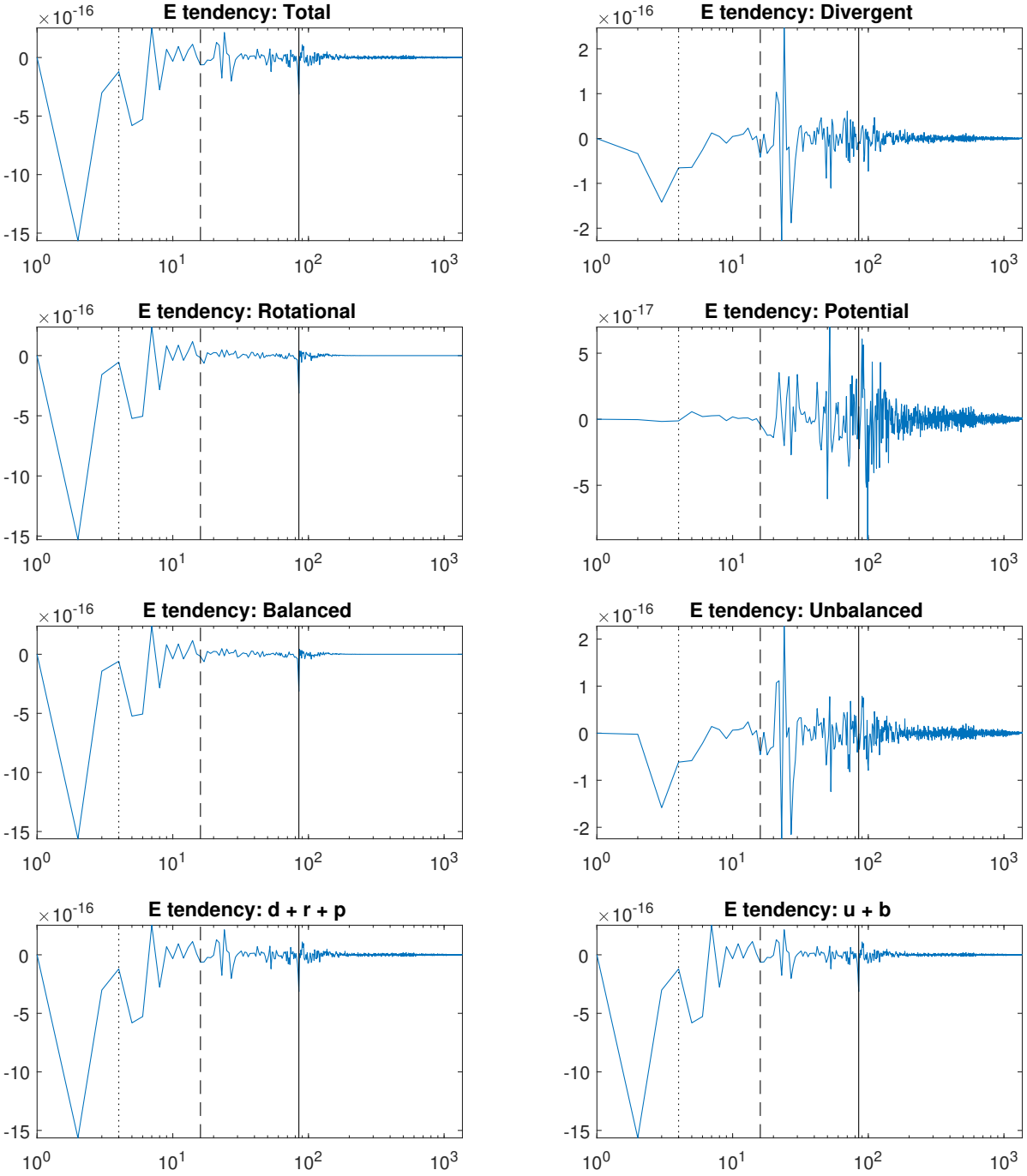

Figure C.17: Difference in the energy tendency spectra resulting from truncation of the potential vorticity in the unbalanced regime

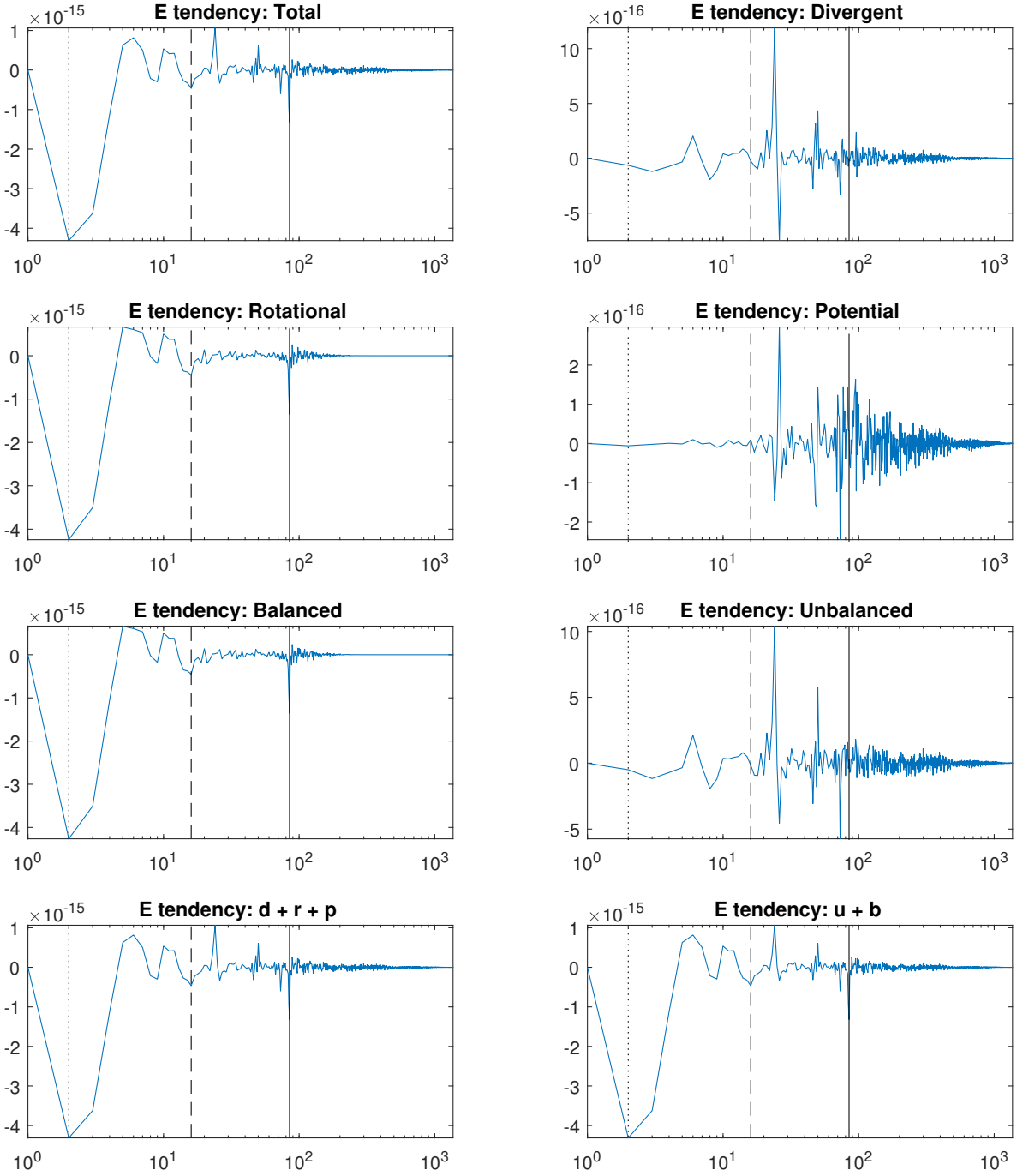

Figure C.18: Difference in the energy tendency spectra resulting from truncation of the potential vorticity in the shock regime

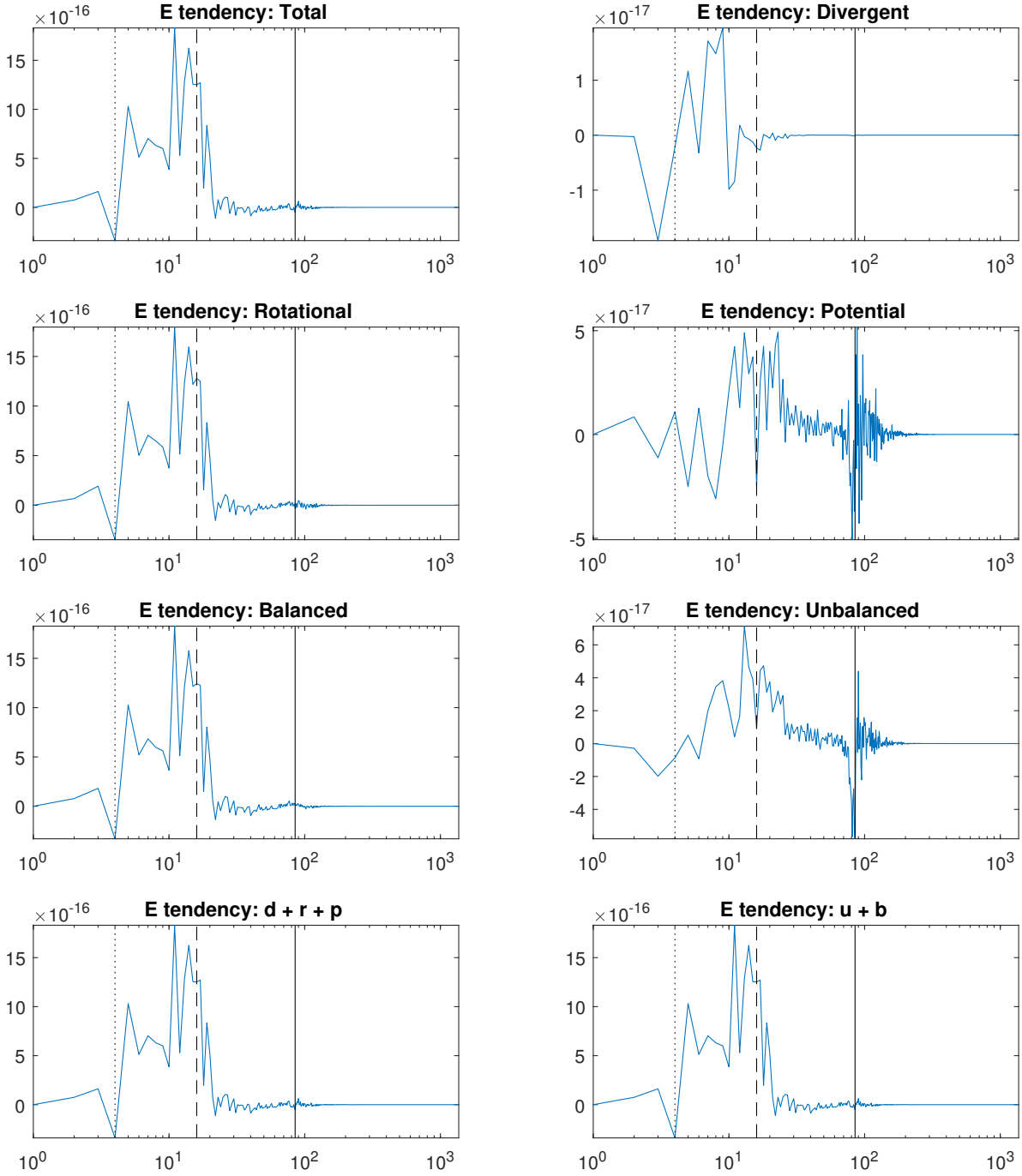

Figure C.19: Difference in the energy tendency spectra resulting from truncation of the unbalanced energy in the balanced regime

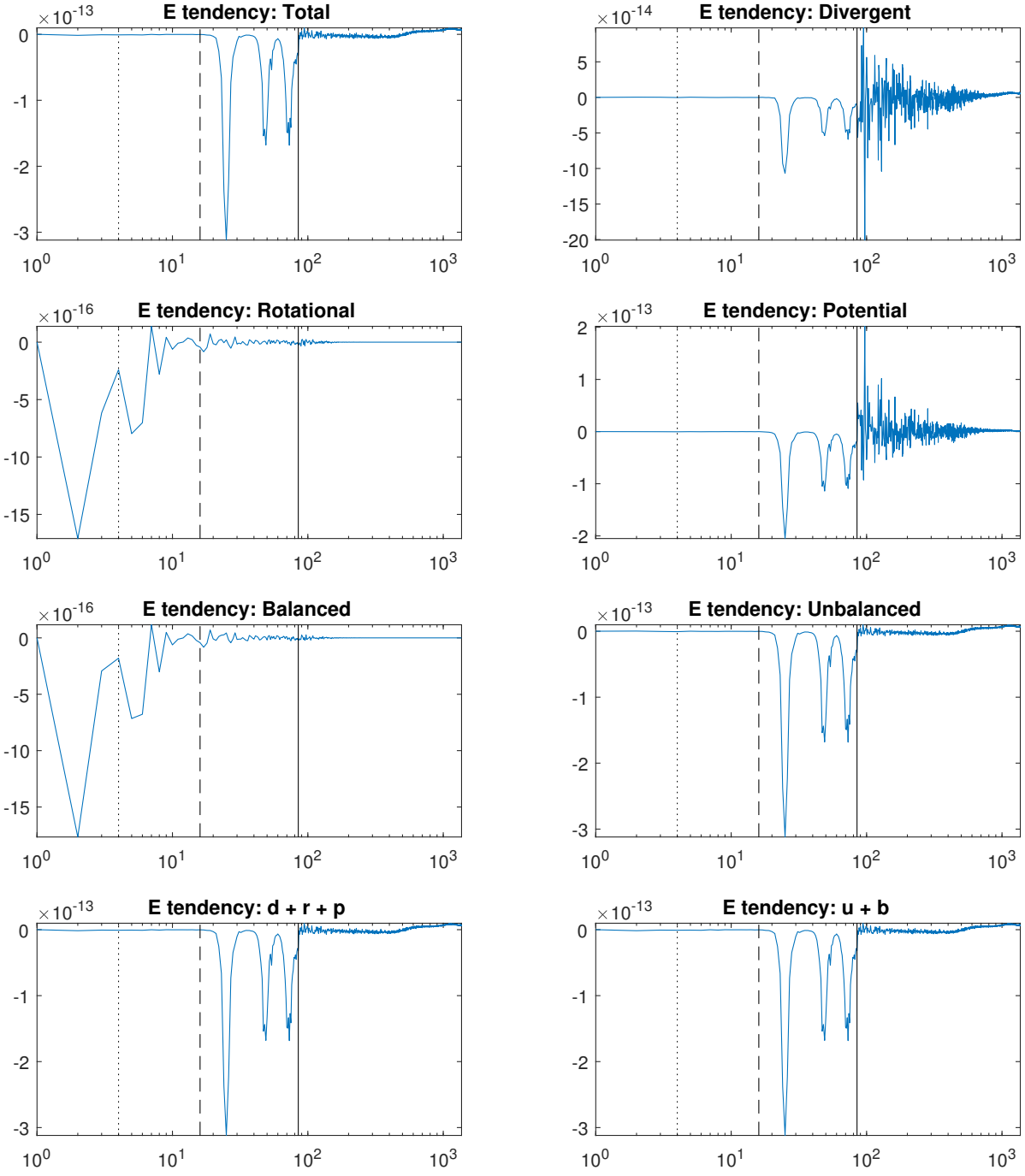

Figure C.20: Difference in the energy tendency spectra resulting from truncation of the unbalanced energy in the unbalanced regime

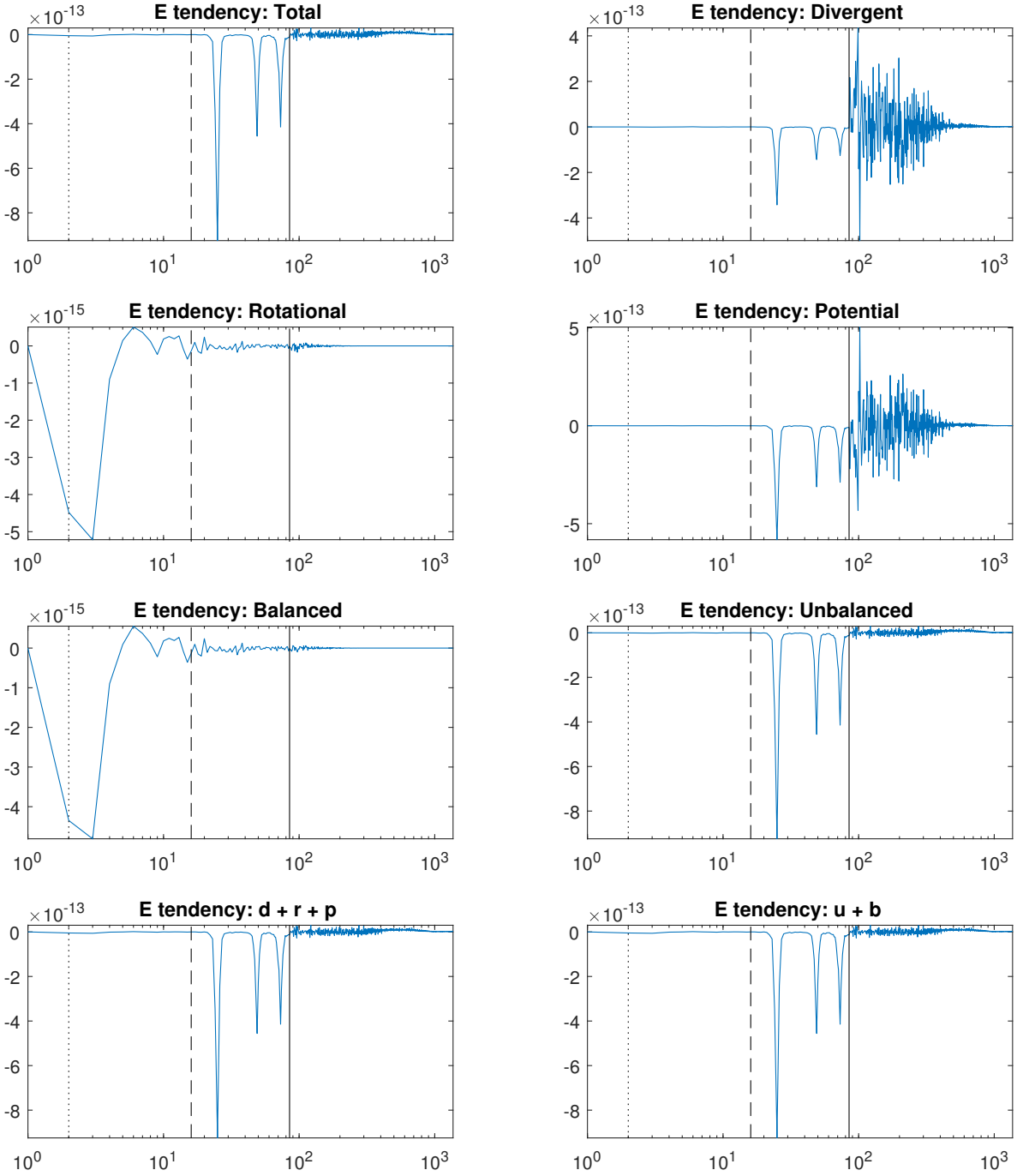

Figure C.21: Difference in the energy tendency spectra resulting from truncation of the unbalanced energy in the shock regime

Bibliography

[1] ENDGame: A new dynamical core for seamless atmospheric prediction. http://www.metoffice.gov.uk/research/news/2014/endgame-a-new-dynamical-core. Accessed: 2017/05/26.

[2] A. Arakawa and V. R. Lamb. Computational design of the basic dynamical processes of the UCLA general circulation model. Methods in computational physics, 17:173–265, 1977.

[3] R. Asselin. Frequency filter for time inegrations. Monthly Weather Review, 100:487–490, 1972.

[4] L. Berselli, T. Iliescu, and W. J. Layton. Mathematics of Large Eddy Simulation of Turbulent Flows. Springer, 2006.

[5] G. F. Carnevale, M. Briscolini, and P. Orlandi. Buoyancy- to inertial-range transition in forced stratified turbulence. Journal of Fluid Mechanics, 427:205–239, January 2001.

[6] J. G. Charney. On the scale of atmospheric motions. Geofysiske publikasjoner, 17:1–17, 1948.

[7] J. G. Charney. Geostrophic turbulence. Journal of the Atmospheric Sciences, 28:1087–1095, 1971.

[8] M. Cullen. Modelling atmospheric flows. Acta Numerica, pages 1–87, 2007.

[9] M. J. P. Cullen. On the accuracy of the semi-geostrophic approximation. Quarterly Journal of the Royal Meteorological Society, 126:1099–1115, 2000.

[10] M. J. P. Cullen. Semi-geostrophic solutions for flow over a ridge. Quarterly Journal of the Royal Meteorological Society, 133:491–501, 2007.

[11] M. J. P. Cullen. A comparison of numerical solutions to the Eady frontogenesis problem. Quarterly Journal of the Royal Meteorological Society, 134:2143–2155, 2008.

[12] Davies, T. and Cullen, M. J. P. and Malcom, A. J. and Mawson, M. H. and Staniforth, A. and White, A. A. and Wood, N. A new dynamical core for the Met Office's global and regional modelling of the atmosphere. Quarterly Journal of the Royal Meteorological Society, 131:1759–1782, 2005.

[13] D. G. Dritschel, G. A. Gottwald, and M. Oliver. Comparison of variational balance models for the rotating shallow water equations. Journal of Fluid Mechanics, 822:689–716, 2017.

[14] P. Edwards. Before 1955: Numerical models and the prehistory of AGCMs. `http://pne.people.si.umich.edu/sloan/mainpage.html`. Accessed: 2018/08/03.

[15] M. Ehrendorfer. Spectral Numerical Weather Prediction Models. Society for Industrial and Applied Mathematics, 2012.

[16] P. F. Embid and A. J. Madja. Low Froude number limiting dynamics for stably stratified flow with small or finite Rossby numbers. Geophysical and Astrophysical Fluid Dynamics, 87:1–50, 1988.

[17] R. Ford, M. E. McIntyre, and W. A. Norton. Balance and the slow quasimanifold: Some explicit results. Journal of the Atmospheric Sciences, 57:1236–1254, May 2000.

[18] J. Galewsky, R. K. Scott, and L. M. Polvani. An initial-value problem for testing numerical models of the global shallow-water equations. Tellus, 56:429–440, October 2004.

[19] A. Gassmann. A global hexagonal C-grid non-hydrostatic dynamical core (ICON-IAP) designed for energetic consistency. Quarterly Journal of the Royal Meteorological Society, 139:152–175, January 2013.

[20] Z. Günther, D. Reinert, P. Rípodas, and M. Baldauf. The ICON (ICOsahedral non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core. Quarterly Journal of the Royal Meteorological Society, 141:563–579, January 2015.

[21] B. J. Hoskins. The geostrophic momentum approximation and the semigeostrophic equations. Journal of the Atmospheric Sciences, 32:233–242, 1975.

[22] J. Kent, J. Thuburn, and N. Wood. Assessing implicit large eddy simulation for two-dimensional flow. Quarterly Journal of the Royal Meteorological Society, 138:365–376, January 2012.

[23] A. Kolmogorov. The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. Comptes rendus de l'Académie des sciences de l'URSS, 30:301–305, 1941.

[24] A. N. Kolmogorov. Dissipation of energy in locally isotropic turbulence. Doklady Akademii Nauk SSSR, 32:16–18, 1941.

[25] P. K. Kundu and I. M. Cohen. Fluid Mechanics, 4th edition. Elsevier, 2008.

[26] C. E. Leith. Nonlinear normal mode initialization and quasi-geostrophic theory. Journal of the Atmospheric Sciences, 37:958–968, 1980.

[27] D. K. Lilly. Stratified turbulence and the mesoscale variability of the atmosphere. Journal of the Atmospheric Sciences, 40:749–761, 1983.

[28] E. N. Lorenz. Attractor sets and quasi-geostrophic equilibrium. Journal of the Atmospheric Sciences, 37:1685–1699, 1980.

[29] J. L. Lumley. The spectrum of nearly inertial turbulence in a stably stratified fluid. Journal of the Atmospheric Sciences, 21:99–102, 1964.

[30] A. R. Mohebalhojeh and D. G. Dritschel. Hierarchies of balance conditions for the f-plane shallow-water equations. Journal of the Atmospheric Sciences, 58:2411–2426, 2001.

[31] G. D. Nastrom and K. S. Gage. A climatology of atmospheric wavenumber spectra of wind and temperature observed by commercial aircraft. Journal of the Atmospheric Sciences, 42:950–960, 1985.

[32] A. Obukhov. On the energy distribution in the spectrum of a turbulent flow. Comptes rendus de l'Académie des sciences de l'URSS, 32:19–21, July 1941.

[33] L. M. Polvani, J. C. McWilliams, M. A. Spall, and R. Ford. The coherent structures of shallow-water turbulence: Deformation-radius effects, cyclone/anticyclone asymmetry and gravity-wave generation. Chaos: An Interdisciplinary Journal of Nonlinear Science, 4:177–186, 1994.

[34] S. B. Pope. Turbulent Flows. Cambridge University Press, 2000.

[35] P. B. Rhines. Waves and turbulence on a beta-plane. Journal of Fluid Mechanics, 69:417–443, June 1975.

[36] L. F. Richardson. The supply of energy from and to atmospheric eddies. Proceedings of the Royal Society of London, 97:354–373, July 1920.

[37] J. J. Riley and S. M. deBruynKops. Dynamics of turbulence strongly influenced by bouyancy. Physics of Fluids, 15:2047–2059, July 2003.

[38] T. D. Ringler. Momentum, vorticity and transport: Considerations in the design of a finite-volume dynamical core. In P. H. Lauritzen, C. Jablonowski, M. A. Taylor, and R. D. Nair, editors, Numerical Techniques for Global Atmospheric Models. Springer.

[39] A. J. Robert. The integration of a low order spectral form of the primitive meteorological equations. Journal of the Meteorological Society of Japan, 44:237–245, 1966.

[40] W. C. Skamarock. Evaluating mesoscale NWP models using kinetic energy spectra. Monthly Weather Review, 132:3019–3032, 2004.

[41] W. C. Skamarock, S.-H. Park, J. B. Klemp, and C. Snyder. Atmospheric kinetic energy spectra from global high-resolution nonhydrostatic simulations. Journal of the Atmospheric Sciences, 71:4369–4381, 2014.

[42] Skamarock, W. C. and Kelmp, J. B. and Dudhia, J. and Gill, D. O. and Barker, D. M. and Duda, M. G. and Huang, XY. and Yang, W. and Powers, J. G. A description of the advanced research WRF version 3. NCAR Technical Note, 2008.

[43] J. Thomas. Resonant fast-slow interactions and breakdown of quasi-geostrophy in rotating shallow water. Journal of Fluid Mechanics, 788:492–520, 2016.

[44] J. Thuburn and C. J. Cotter. A framework for mimetic discretization of the rotating shallow-water equations on arbitrary polygonal grids. Society for Industrial and Applied Mathematics, 34:B203–B225, 2012.

[45] J. Thuburn, C. J. Cotter, and T. Dubos. A mimetic, semi-implicit, forward-in-time, finite volume shallow water model: comparison of hexagonal-icosahedral and cubed sphere grids. Geoscientific Model Development, 7:909–929, 2014.

[46] J. Thuburn, T. D. Ringler, W. C. Skamarock, and J. B. Klemp. Numerical representation of geostrophic modes on arbitrarily structured C-grids. Journal of Computational Physics, 228:8321–8335, 2009.

[47] Thuburn, J. and Cotter, C. J. A primal-dual mimetic finite element scheme for the rotating shallow water equations on polygonal spherical meshes. Journal of Computational Physics, 290:274–297, June 2015.

[48] Thuburn, J. and Kent, J. and Wood, N. Cascades, backscatter and conservation in numerical models of two-dimensional turbulence. Quarterly Journal of the Royal Meteorological Society, 140:626–638, 2013.

[49] A. Vallgren, E. Deusebio, and E. Lindborg. Possible explanation of the atmospheric kinetic and potential energy spectra. Physical Review Letters, 107:268501, 2011.

[50] G. K. Vallis. Atmospheric and Oceanic Fluid Dynamics. Cambridge University Press, 2006.

[51] J. Vanneste. Exponential smallness of inertia-gravity wave generation at small Rossby number. Journal of the Atmospheric Sciences, 65:1622–1637, May 2008.

[52] J. Vanneste. Balance and spontaneous wave generation in geophysical flows. Annual Review of Fluid Mechanics, 45:147–172, 2013.

[53] A. J. Visram, C. J. Cotter, and M. J. P. Cullen. A framework for evaluating model error using asymptotic convergence in the Eady model. Quarterly Journal of the Royal Meteorological Society, 140:1629–1639, 2014.

[54] T. von Larcher and P. D. Williams. Modeling Atmospheric and Oceanic Flows: Insights from Laboratory Experiments and Numerical Simulations. Wiley, 2015.

[55] M. L. Waite and P. Bartello. Stratified turbulence dominated by vortical motion. Journal of Fluid Mechanics, 517:281–308, October 2004.

[56] M. L. Waite and P. Bartello. Stratified turbulence generated by internal gravity waves. Journal of Fluid Mechanics, 546:313–339, January 2006.

[57] M. L. Waite and P. Bartello. The transition from geostrophic to stratified turbulence. Journal of Fluid Mechanics, 568:89–108, January 2006.

[58] G. B. Whitham. Linear and Nonlinear Waves. Wiley, 1974.

[59] P. D. Williams. A proposed modification to the Robert-Asselin time filter. Monthly Weather Review, 137:2538–2546, 2009.

[60] B. A. Wingate, P. Embid, M. Holmes-Cerfon, and M. Taylor. Low Rossby limiting dynamics for stably stratified flow with finite Froude number. Journal of Fluid Mechanics, 676:546–571, June 2011.

[61] M. Zerroukat, N. Wood, and A. Staniforth. An improved version of SLICE for conservative monotonic remapping on a C-grid. Quarterly Journal of the Royal Meteorological Society, 135:541–546, 2009.

[62] Zerroukat, M. and Wood, N. and Staniforth, A. and White, A. A. and Thuburn, J. An inherently mass-conserving semi-implicit semi-Lagrangian discretisation of the shallow-water equations on the sphere. Quarterly Journal of the Royal Meteorological Society, 135:1104–1116, 2009.