Available online at www.sciencedirect.com
SCIENCE DIRECT®
Dynamics of Atmospheres and Oceans
38 (2005) 195–230
dynamics of atmospheres and oceans
www.elsevier.com/locate/dynatmoce

Localized multiscale energy and vorticity analysis I. Fundamentals

X. San Liang[a,*], Allan R. Robinson[a,b]

[a] *Harvard University, Division of Engineering and Applied Sciences, 29 Oxford Street, Cambridge, MA 02138, USA*
[b] *Harvard University, Department of Earth and Planetary Sciences, Cambridge, MA, USA*

Received 6 October 2003; received in revised form 10 December 2004; accepted 17 December 2004
Available online 24 March 2005

Abstract

A new methodology, *multiscale energy and vorticity analysis* (MS-EVA), is developed to investigate the inference of fundamental processes from oceanic or atmospheric data for complex dynamics which are nonlinear, time and space intermittent, and involve multiscale interactions. Based on a localized orthogonal complementary subspace decomposition through the multiscale window transform (MWT), MS-EVA is real problem-oriented and objective in nature. The development begins with an introduction of the concepts of scale and scale window and the decomposition of variables on scale windows. We then derive the evolution equations for multiscale kinetic and available potential energies and enstrophy. The phase oscillation reflected on the horizontal maps from Galilean transformation is removed with a 2D large-scale window synthesis. The resulting energetic terms are analyzed and interpreted. These terms, after being carefully classified, provide four types of processes: transport, transfer, conversion, and dissipation/diffusion. The key to this classification is the transfer–transport separation, which is made possible by looking for a special type of transfer, the so-called *perfect transfer*. The intricate energy source information involved in perfect transfers is differentiated through an interaction analysis.

The transfer, transport, and conversion processes form the basis of dynamical interpretation for GFD problems. They redistribute energy in the phase space, physical space, and space of energy types. These processes are all referred to in a context local in space and time, and therefore can be

* Corresponding author. Tel.: +1 617 495 2899; fax: +1 617 495 5192.
E-mail address: liang@deas.harvard.edu (X. San Liang).

0377-0265/$ – see front matter
doi:10.1016/j.dynatmoce.2004.12.004

easily applied to real ocean problems. When the dynamics of interest is on a global or duration scale, MS-EVA is reduced to a classical Reynolds-type energetics formalism.

Keywords: MS-EVA; Multiscale window transform; Perfect transfer; Interaction analysis

1. Introduction

Energy and vorticity analysis is a widely used approach in the diagnosis of geophysical fluid processes. During past decades, much work has been done along this line, examples including Holland and Lin (1975), Harrison and Robinson (1978), Plumb (1983), Pinardi and Robinson (1986), Spall (1989), Cronin and Watts (1996), to name but a few. While these classical analyses have been successful in their respective applications, real ocean processes usually appear in more complex forms, involving interactions among multiple scales and tending to be intermittent in space and time. In order to investigate ocean problems on a generic basis, capabilities of classical energetic analyses need to be expanded to appropriately incorporate and faithfully represent all these processes. This forms the objective of this work.

We develop a new methodology, multiscale energy and vorticity analysis (MS-EVA), to fulfill this objective. MS-EVA is a generic approach for the investigation of multiscale nonlinear interactive oceanic processes which occur locally in space and time. It aims to explore pattern generation and energy and enstrophy budgets, and to unravel the intricate relationships among events on different scales and in different locations. In the sequels to this paper (referred to as LR1), Liang and Robinson (2005, 2004) (LR2 and LR3 hereafter), we will show how MS-EVA can be utilized for instability analysis and how it can be applied to solve real ocean problems which would otherwise be difficult to solve.

In order to be real problem-oriented, MS-EVA should contain full physics. Approximations such as linearization are thus not allowed. It must also have a multiscale representation which retains time and space localization. In other words, the representation should retain time intermittency, and should be able to handle events occurring on limited, irregular and time dependent domains. This makes MS-EVA distinctly different from classical formalism.

MS-EVA should also be *scale windowed*, i.e., the multiscale decomposition must be able to represent events occurring coherently on scale ranges, or *scale windows*. Loosely speaking, a scale window is simply a subspace with a certain range of scales. A rigorous definition is deferred to Section 2. In general, GFD processes tend to occur on scale windows, rather than individual scales. We refer to this phenomenon as scale windowing. Scale windowing requires a special bulk treatment of energy rather than individual scale representations, as transfers between individual scales belonging respectively to different windows could take a direction opposite to the overall transfer between these windows.

Multiscale events could be represented in different forms. One of the most frequently used is wave representation (e.g., Fourier analysis), which transforms events onto many

individual scales; another frequently used form is called eddy representation (Tennekes and Lumley, 1972), in which a process is decomposed into a large-scale part and an eddy part, each part involving a range of scales. Because of its scale window nature, we need an eddy representation for MS-EVA. The resulting energetics will be similar to those of Reynolds formulation, except that the latter is in a statistical context.

To summarize, it is required that MS-EVA handle fairly generic processes in the sense of multiscale windowing, spatial localization, and temporal intermittency; as well as retain full physics. Correspondingly an analysis tool is needed in the MS-EVA formulation such that all these requirements are met. We will tackle this problem in a spirit similar to the wavelet transform, a localized analysis which has been successfully applied to studying energetics for individual scales (e.g., Iima and Toh, 1995; Fournier, 1999). Specifically, we need to generalize the wavelet analysis to handle window or eddy decomposition. The challenge is how to incorporate into a window the transform coefficients (and hence energies) of an orthonormal wavelet transform which are defined discretely at different locations for different scales, while retaining a resolution satisfactory to the problem. (Orthonormality is essential to keep energy conserved.) The next section is intended to deal with this issue. The new analysis tool thus constructed will be termed *multiscale window transform*, or MWT for short. The whole problem is now reduced to first the building of MWT, and then the development of MS-EVA with the MWT. In Sections 3–7, we apply MWT to derive the laws that govern the multiscale energy evolutions. The multiscale decomposition is principally in time, but with a horizontal treatment which preserves spatial localization. Time scale decomposition has been a common practice and meteorologists find it useful for clarifying atmospheric processes. We choose to do so in order to make contacts with the widely used Reynolds averaging formalism, and more importantly, to have the concept *scale* unambiguously defined (cf. Section 2.1), avoiding extra assumptions such as space isotropy or anisotropy. Among these sections, Section 3 is devoted to define energy on scale windows, and Section 4 is for a primary treatment with the nonlinear terms. The multiscale kinetic and potential energy equations are first derived in Sections 5 and 6 based on a time decomposition, and then modified to resolve the spatial issue with a horizontal synthesis (Section 7). In Section 8, we demonstrate how these equations are connected to energetics in the classical formalism. This section is followed by an interaction analysis for the differentiation of transfer sources (Section 9), which allows a description of the energetic scenario with our MS-EVA analysis in both physical and phase spaces (Section 10). As "vorticity" furnishes yet another part of MS-EVA, in Section 11 we briefly present how enstrophy evolves on multiple scale windows. This work is summarized in Section 12, where prospects for application are outlined as well.

2. Multiscale window analysis and marginalization

In this section, we introduce the concept of scale window, multiscale window transform (MWT), and some properties of the MWT, particularly a property referred to as marginalization. A thorough and rigorous treatment is beyond the scope of this paper. For details, the reader is referred to Liang (2002) (L02 hereafter) and Liang and Anderson (2005).

2.1. Scale and scale window

The introduction of MWT relies on how a scale is defined. In this context, our definition of *scale* is based on a modified wavelet analysis (cf., Hernández and Weiss, 1996). For convenience, we limit the initial discussion to 1D functions. The multi-dimensional case is a direct extension and can be found in L02, Section 2.7. For any function $p(t) \in L_2[0, 1]$,[1] it can been analyzed as (L02):

$$p(t) = \sum_{j=0}^{+\infty} \sum_{n=0}^{2^j\varrho-1} \tilde{p}_n^j \psi_n^{\varrho,j}(t), \qquad t \in [0, 1], \tag{1}$$

where

$$\psi_n^{\varrho,j}(t) = \sum_{\ell=-\infty}^{+\infty} 2^{j/2}\psi[2^j(t + \varrho\ell) - n], \qquad n = 0, 1, \ldots, 2^j\varrho - 1 \tag{2}$$

and ψ is some orthonormalized wavelet function.[2] Here we choose it to be the one built from cubic splines, which is shown in Fig. 1a. The "period" ϱ has two choices only: one is $\varrho = 1$, which gives a periodic extension of the signal of interest from [0, 1] to the whole real line $\mathbb{R}$; another is $\varrho = 2$, corresponding to an extension by reflection, which is also an "even periodization" of the finite signal to $\mathbb{R}$ (see L02 for details).

The distribution of $\psi_n^{1,j}(t)$ with $j = 2, 4, 6$ is shown in Fig. 1b. Each j corresponds to a quantity 2^{-j}, which can be used to define a time metric to relate the passage of temporal events since a selected epoch. We call this j a *scale level*, and 2^{-j} the corresponding *scale* over [0, 1].

Given the scale as conceptualized, we proceed to define scale windows. In the analysis (1), we can group together those parts with a certain range of scale levels, say, $(j_1, j_1 + 1, \ldots, j_2)$, to form a subspace of $L_2[0, 1]$. This subspace is called a *scale window* of $L_2[0, 1]$ in L02 with scale levels ranging from j_1 to j_2. In doing this, any function in $L_2[0, 1]$, say $p(t)$, can be decomposed into a sum of several parts, each encompassing exclusively features on a certain window of scales. Specifically for this work, we define three scale windows:

- large-scale window: $0 \le j \le j_0$,
- meso-scale window: $j_0 < j \le j_1$,
- sub-mesoscale window: $j_1 < j \le j_2$.

The scale level bounds j_0, j_1, j_2 are set according to the problem under consideration. Particularly, j_2 corresponds to the finest resolution (sampling interval 2^{-j_2}) permissible by the given finite signals. By projecting $p(t)$ onto these three windows, we obtain its large-scale, meso-scale, and sub-mesoscale features, respectively. This decomposition is orthogonal, so the total energy thus yielded is conserved.

[1] The notation $L_2[0, 1]$ is used to indicate the space of square integrable functions defined on [0, 1].

[2] This is to say, $\{\psi(t - \ell), \ell \in \mathbb{Z}\}$ ($\mathbb{Z}$ the set of integers) forms an orthonormal set.

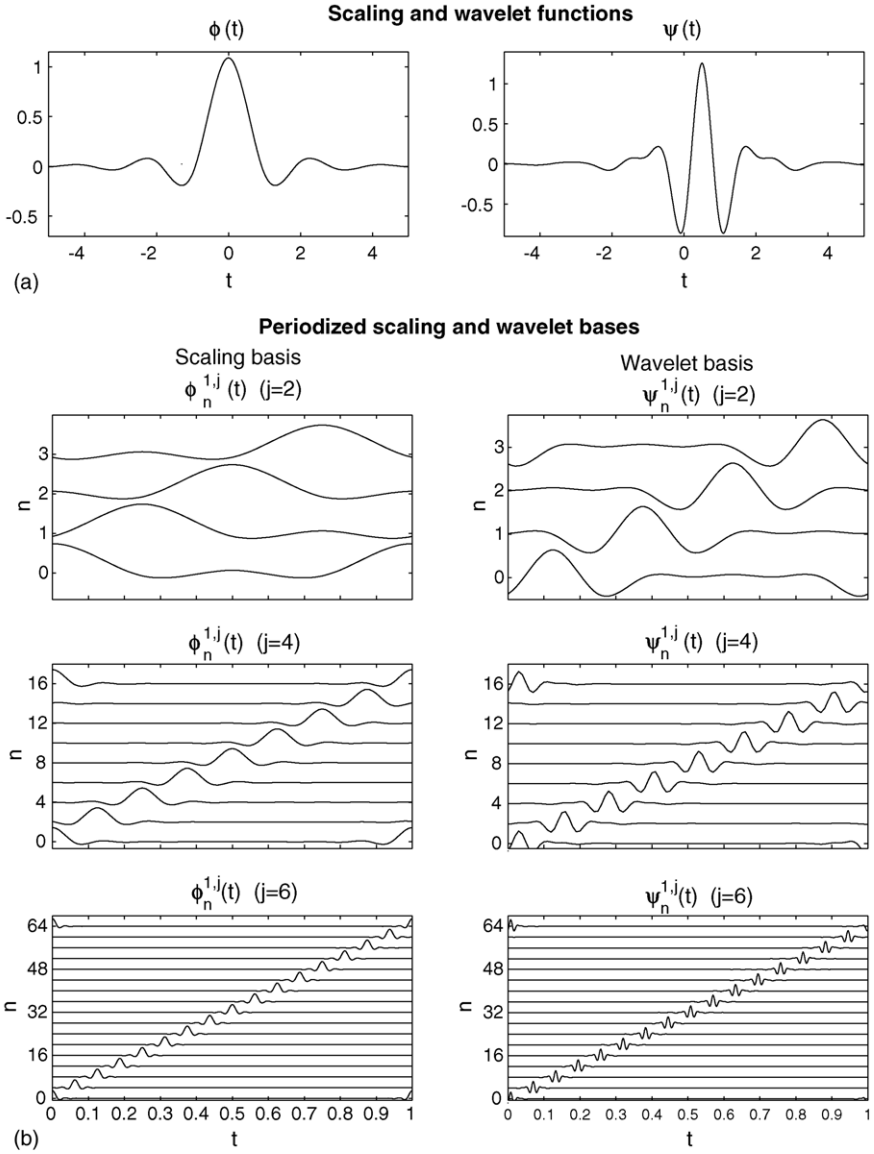

Fig. 1. Scaling and wavelet functions (a) and their corresponding periodized bases ($\varrho = 1$) $\{\phi_n^{\varrho,j}(t)\}_n$ (left panel) and $\{\psi_n^{\varrho,j}(t)\}_n$ (right panel) with scale levels $j = 2$ (top), $j = 4$ (middle), and $j = 6$ (bottom), respectively (b). The scaling and wavelet functions ϕ and ψ are constructed from cubic splines (see Liang, 2002, Section 2.5).

2.2. Multiscale window transform

Scale windows are defined with the aid of wavelet basis, but the definition of multiscale window transform does not follow the same line because of the difficulty we have described

in the introduction, i.e., that orthonormal wavelet transform coefficients are defined discretely on different locations for different scales. To circumvent this problem, we make a direct sum of the subspaces spanned by the wavelet basis $\{\psi_n^{\varrho,m}(t)\}_n$, for all $m \le j$. The shift-invariant basis of the resulting subspace can be shown to be $\phi_n^{\varrho,j}(t)$ (L02), which is the periodization [cf. (2)] of some $\phi(t)$, the orthonormal scaling function in company with the wavelet function $\psi(t)$. Here ϕ is an orthonormalized cubic spline, as shown in Fig. 1a. We utilize the $\phi_n^{\varrho,j}$ thus formed to fulfill our task. In the following only the related formulas and equations are presented. The details are referred to L02.

Let V_{ϱ,j_2} indicate the total (direct sum, to be strict) of the three scale windows. It has been established by L02 that any time signal from a given GFD dataset is justifiably belonging to V_{ϱ,j_2}, with some finite level j_2. Suppose we have $p(t) \in V_{\varrho,j_2}$. Write

$$\hat{p}_n^j = \int_0^\varrho p(t)\phi_n^{\varrho,j}(t)\,\mathrm{d}t, \qquad \text{for all } 0 \le j \le j_2, \quad n = 0, 1, \ldots, 2^j\varrho - 1. \tag{3}$$

Given window bounds j_0, j_1, j_2, and $p \in V_{\varrho,j_2}$, three functions can be accordingly defined:

$$p^{\sim 0}(t) = \sum_{n=0}^{2^{j_0}\varrho-1} \hat{p}_n^{j_0}\phi_n^{\varrho,j_0}(t), \tag{4}$$

$$p^{\sim 1}(t) = \sum_{n=0}^{2^{j_1}\varrho-1} \hat{p}_n^{j_1}\phi_n^{\varrho,j_1}(t) - p^{\sim 0}(t), \tag{5}$$

$$p^{\sim 2}(t) = p(t) - \sum_{n=0}^{2^{j_1}\varrho-1} \hat{p}_n^{j_1}\phi_n^{\varrho,j_1}(t), \tag{6}$$

on the basis of which we will build the MWT later. As a scaling transform coefficient, $\hat{p}_n^j$ contains all the information with scale level lower than or equal to j. The functions $p^{\sim 0}(t)$, $p^{\sim 1}(t)$, $p^{\sim 2}(t)$ thus defined hence include only features of $p(t)$ on ranges $0 - j_0$, $j_0 - j_1$, and $j_1 - j_2$, respectively. For this reason, we term these functions as large-scale, meso-scale, and sub-mesoscale syntheses or reconstructions of $p(t)$, with the notation ~ 0, ~ 1, and ~ 2 in the superscripts signify the corresponding large-scale, meso-scale, and sub-mesoscale windows, respectively.

Using the multiscale window synthesis, we proceed to define a transform

$$\hat{p}_n^{\sim\varpi} = \int_0^\varrho p^{\sim\varpi}(t)\phi_n^{\varrho,j_2}(t)\,\mathrm{d}t \tag{7}$$

for windows $\varpi = 0, 1, 2$, $n = 0, 1, \ldots, 2^{j_2}\varrho - 1$. This is the *multiscale window transform*, or MWT for short, that we want to build. Notice here we use a periodized scaling basis at j_2, the highest level that can be attained for a given time series. As a result, the transform coefficients have a maximal resolution in the sampled t direction.

In terms of $\hat{p}_n^{\sim\varpi}$, Eqs. (4)–(6) can be simplified as

$$p^{\sim\varpi}(t) = \sum_{n=0}^{2^{j_2}\varrho-1} \hat{p}_n^{\sim\varpi}\phi_n^{\varrho,j_2}(t), \tag{8}$$

for $\varpi = 0, 1, 2$. Eqs. (7) and (8) are the transform-reconstruction pair for our MWT. For any $p \in V_{\varrho,j_2}$, it can be now represented as

$$p(t) = \sum_{\varpi=0}^{2}\sum_{n=0}^{2^{j_2}\varrho-1} \hat{p}_n^{\sim\varpi}\phi_n^{\varrho,j_2}(t). \tag{9}$$

A final remark on the choice of extension scheme, or the "period" ϱ in the analysis. In general, we always adopt the extension by reflection $\varrho = 2$, which has proved to be very satisfactory. (Fig. 4 shows such an example.) If the signals given are periodic, then the periodic extension is the exact one, and hence ϱ should be chosen to be 1. In case of linking to the classical energetic formalism, $\varrho = 1$ is also usually used.

2.3. MWT properties and marginalization

Multiscale window transform has many properties. In the following we present two of them which will be used later in the MS-EVA development (for proofs, refer to L02).

Property 1. *For any $p \in V_{\varrho,j_2}$, if $j_0 = 0$, and $\varrho = 1$ (periodic extension adopted), then*

$$\hat{p}_n^{\sim 0} = 2^{-j_2/2}p^{\sim 0}(t) = 2^{-j_2/2}\overline{p} = \textit{constant}, \qquad \textit{for all } n, \textit{ and } t, \tag{10}$$

where the overbar stands for averaging over the duration.

Property 2. *For p and q in V_{ϱ,j_2},*

$$\mathcal{M}_n\hat{p}_n^{\sim\varpi}\hat{q}_n^{\sim\varpi} = \overline{p^{\sim\varpi}(t)q^{\sim\varpi}(t)}, \tag{11}$$

where

$$\mathcal{M}_n(\hat{p}_n^{\sim\varpi}\hat{q}_n^{\sim\varpi}) = \sum_{n=1}^{N-1}\hat{p}_n^{\sim\varpi}\hat{q}_n^{\sim\varpi} + \frac{1}{2}[\hat{p}_0^{\sim\varpi}\hat{q}_0^{\sim\varpi} + \hat{p}_N^{\sim\varpi}\hat{q}_N^{\sim\varpi}]. \qquad (N = 2^{j_2}) \tag{12}$$

Property 1 states that when $j_0 = 0$ and a periodic extension is used, the large-scale window synthesis is simply the duration average. Property 2 involves a special summation over $[0, N]$ (corresponding to $t \in [0, 1]$), which we will call *marginalization* hereafter. The word "marginal" has been used in literature to describe the overall feature of a localized transform (e.g., Huang et al., 1999). We extend this convention to establish an easy reference for the operator $\mathcal{M}_n$. Property 2 can now be restated as: a product of two multiscale window transforms followed by a marginalization is equal to the product of their corresponding syntheses averaged over the duration. For convenience, this property will be referred to as *property of marginalization*.

We close this section by making a comparison between our MWT and wavelet analysis. The commonality is, of course, that both of them are localized on the definition domain. The first and largest difference between them is that the MWT is not a transform in the usual sense. It is an orthogonal complementary subspace decomposition, and as a result, the MWT coefficients contain information for a range of scales, instead of a single scale. For this reason, it is required that three scale bounds be specified a priori in constructing the windows. A useful way to do this is through wavelet spectrum analysis, as is used in LR3. Secondly, the MWT transform is projected on V_{ϱ, j_2}, so transform coefficients obtained for all the windows have the same resolution—the maximal resolution allowed for the signal. This is in contrast to wavelet analysis, whose transform coefficients have different resolution on different scales. We will see soon that, this maximized resolution in MWT transform coefficients puts the embedded phase oscillation under control.

3. Multiscale energies

Beginning this section through Section 7, we will derive the equations that govern the multiscale energy evolutions. The whole formulation is principally based on a time decomposition, but with an appropriate filtering in the horizontal dimensions. It involves a definition of energies on different scale windows, a classification of distinct processes from the nonlinear convective terms, a derivation of time windowed energetic equations, and a horizontal treatment of these equations with a space window reconstruction. In this section, we define the energies for the three time scale windows.

3.1. Primitive equations and kinetic and available potential energies

The governing equations adopted in this study are:

$$\frac{\partial \underline{\mathbf{v}}}{\partial t} = -\nabla\cdot(\underline{\mathbf{v}}\,\underline{\mathbf{v}}) - \frac{\partial(w\underline{\mathbf{v}})}{\partial z} - f\underline{\mathbf{k}}\wedge\underline{\mathbf{v}} - \frac{1}{\rho_0}\nabla P + \underline{\mathbf{F}}_{mz} + \underline{\mathbf{F}}_{mh}, \tag{13}$$

$$0 = \nabla\cdot\underline{\mathbf{v}} + \frac{\partial w}{\partial z}, \tag{14}$$

$$0 = -\frac{\partial P}{\partial z} - \rho g, \tag{15}$$

$$\frac{\partial \rho}{\partial t} = -\nabla\cdot(\underline{\mathbf{v}}\rho) - \frac{\partial(w\rho)}{\partial z} + \frac{N^2\rho_0}{g}w + F_{\rho z} + F_{\rho h}, \tag{16}$$

where $\underline{\mathbf{v}} = (u, v)$ is the horizontal velocity vector, $\nabla = \underline{\mathbf{i}}\frac{\partial}{\partial x} + \underline{\mathbf{j}}\frac{\partial}{\partial y}$ the horizontal gradient operator, $N = (-\frac{g}{\rho_0}\frac{\partial\bar{\rho}}{\partial z})^{1/2}$ the buoyancy frequency ($\bar{\rho} = \bar{\rho}(z)$ is the stationary density profile), ρ the density perturbation with $\bar{\rho}$ excluded, and P the dynamic pressure. All the other notations are conventional. The friction and diffusion terms are just symbolically expressed. The treatment of these subgrid processes in a multiscale setting is not considered in this

paper. From Eqs. (13) and (14), it is easy to obtain the equations that govern the evolution of two quadratic quantities: $K = \frac{1}{2}\underline{\mathbf{v}} \cdot \underline{\mathbf{v}}$, and $A = \frac{1}{2}\frac{g^2}{\rho_0^2 N^2}\rho^2$ (see Spall, 1989). These are the total kinetic energy (KE) and available potential energy (APE), given the location in space and time. The essence of this study is to investigate how KE and APE are distributed simultaneously in the physical and phase spaces.

3.2. Multiscale energies

Multiscale window transforms equipped with the marginalization property (11) allow a simple representation of energy for each scale window $\varpi = 0, 1, 2$. For a scalar field $S(t) \in V_{\varrho, j_2}$, let $E_n^{\varpi *} = (\hat{S}_n^{\sim\varpi})^2$. By (11),

$$\mathcal{M}_n E_n^{\varpi *} = \int_0^1 [S^{\sim\varpi}(t)]^2 \, \mathrm{d}t, \tag{17}$$

which is essentially the energy of S on window ϖ (up to some constant factor) integrated with respect to t over $[0, 1)$. Recall $\mathcal{M}_n$ is a special sum over the 2^{j_2} discrete equi-distance locations $n = 0, 1, \ldots, 2^{j_2} - 1$. $E_n^{\varpi *}$ thus can be viewed as the energy on window ϖ summed over a small interval of length $\Delta t = 2^{-j_2}$ around location $t = 2^{-j_2} n$. An energy variable for window ϖ at time $2^{-j_2} n$ consistent with the fields at that location is therefore a locally averaged quantity

$$E_n^{\varpi} = \frac{1}{\Delta t} E_n^{\varpi *} = 2^{j_2} \cdot (\hat{S}_n^{\sim\varpi})^2, \tag{18}$$

for all $\varpi = 0, 1, 2$. It is easy to establish that

$$\mathcal{M}_n(E_n^0 + E_n^1 + E_n^2)\Delta t = \int_0^1 S^2(t) \, \mathrm{d}t. \tag{19}$$

This is to say, the energy thus defined is conserved.

In the same spirit, the multiscale kinetic and available potential energies now can be defined as follows:

$$K_n^{\varpi} = \frac{1}{2}[2^{j_2}(\hat{u}_n^{\sim\varpi})^2 + 2^{j_2}(\hat{v}_n^{\sim\varpi})^2] = 2^{j_2}\left[\frac{1}{2}\hat{\underline{\mathbf{v}}}_n^{\sim\varpi} \cdot \frac{1}{2}\hat{\underline{\mathbf{v}}}_n^{\sim\varpi}\right] \tag{20}$$

$$A_n^{\varpi} = 2^{j_2}\left[\frac{1}{2}\frac{g^2}{\rho_0^2 N^2}\hat{\rho}_n^{\sim\varpi} \cdot \hat{\rho}_n^{\sim\varpi}\right] = 2^{j_2}\left[\frac{1}{2}c\hat{\rho}_n^{\sim\varpi}\hat{\rho}_n^{\sim\varpi}\right], \tag{21}$$

where the shorthand $c \equiv g^2/(\rho_0^2 N^2)$ is introduced to avoid otherwise cumbersome derivation of the potential energy equation. (Note c is z-dependent.) The purpose of the following sections are to derive the evolution laws for K_n^{ϖ} and A_n^{ϖ}. Note the factor 2^{j_2}, which is a constant once a signal is given, provides no information essential to our dynamics analysis. In the MS-EVA derivation, we will drop it in order to avoid otherwise awkward expressions. Therefore, *all the energetic terms hereafter, unless otherwise indicated, should be multiplied by* 2^{j_2} *before physically interpreted*.

4. Perfect transfer and transfer–transport separation

The MS-EVA is principally developed for time, but with a horizontal treatment for spatial oscillations. Localized energetic study with a time decomposition (and the statistical formulation) raises an issue: the separation of transport from the nonlinear term-related energetics. Here by transport we mean a process which can be represented by some quantity in a form of divergence. It vanishes if integrated over a closed domain. The separation of transport is very important, since it allows the cross-scale energy transfer to come upfront.

Transfer–transport separation is not a problem in a space decomposition-based energetic formulation, e.g., the Fourier formulation. In that case the analysis over the space has already eliminated the transport, and as a result, the summation of the triad interaction terms over all the possible scales vanishes. This problem surfaces in a localized time-based formulation when uniqueness is concerned. In this section, we will show how it is resolved.

We begin by introducing a concept, *perfect transfer process*, for our purpose. The so-called *perfect transfer* is a family of multiscale energetic terms which vanish upon summation over all the scale windows and marginalization over the sampled time locations. A *perfect transfer process*, or simply perfect transfer when no confusion arises in the context, is then a process represented by perfect transfer term(s). Perfect transfers move energy from window to window without destroying or generating energy as a whole. They represent a kind of redistribution process among multiple scale windows. In terms of physical significance, the concept of perfect transfer is a natural choice. We are thence motivated to seek through a larger class of "transfer processes" for perfect transfers, which set a constraint for transport–transfer separation and hence help to solve the above uniqueness problem.

For a detailed derivation of the transport–transfer separation, refer to Liang et al. (2005). Briefly cited here is the result with some modification to the needs in our context. The idea is that, for an incompressible fluid flow, we can have the nonlinear-term related energetics separated into a transport plus a perfect transfer, and the separation is unique. For simplicity, consider a scalar field $S = S(t, x, y)$. Suppose it is simply advected by an incompressible 2D flow $\underline{\mathbf{v}}$, i.e., the evolution is governed by

$$\frac{\partial S}{\partial t} = -\nabla \cdot (\underline{\mathbf{v}} S), \qquad \nabla \cdot \underline{\mathbf{v}} = 0. \tag{22}$$

Let $E_n^{\varpi} = \frac{1}{2}(\hat{S}_n^{\sim\varpi})^2$ be its energy (variance) at time location n on scale window ϖ. The evolution of E_n^{ϖ} can be easily obtained by making a transform of the equation followed by a product with $\hat{S}_n^{\sim\varpi}$. We are tasked to separate the resulting triple product term

$$\mathrm{NL} = -\hat{S}_n^{\sim\varpi} \nabla \cdot (\widehat{\underline{\mathbf{v}} S})_n^{\sim\varpi}$$

as needed. By L02, this is done by performing the separation as

$$\mathrm{NL} = -\nabla \cdot \underline{\mathbf{Q}}_{S_n^{\varpi}} + [-\hat{S}_n^{\sim\varpi} \nabla \cdot (\widehat{\underline{\mathbf{v}} S})_n^{\sim\varpi} + \nabla \cdot \underline{\mathbf{Q}}_{S_n^{\varpi}}] \equiv \Delta_h Q_{S_n^{\varpi}} + T_{S_n^{\varpi}}, \tag{23}$$

where

$$\underline{\mathbf{Q}}_{S_n^{\varpi}} = \lambda_c \hat{S}_n^{\sim\varpi} (\widehat{\underline{\mathbf{v}} S})_n^{\sim\varpi}, \qquad \lambda_c = \tfrac{1}{2}, \tag{24}$$

and

$$\Delta_h Q_{S_n^{\varpi}} \equiv -\nabla \cdot \underline{\mathbf{Q}}_{S_n^{\varpi}} \tag{25}$$

$$T_{S_n^{\varpi}} \equiv -\hat{S}_n^{\sim\varpi} \nabla \cdot \left(\widehat{\underline{\mathbf{v}} S}\right)_n^{\sim\varpi} + \nabla \cdot \underline{\mathbf{Q}}_{S_n^{\varpi}}. \tag{26}$$

It is easy to verify that

$$\sum_{\varpi} \mathcal{M}_n T_{S_n^{\varpi}} = 0, \tag{27}$$

which implies that $T_{S_n^{\varpi}}$ represents a perfect transfer process.

Eq. (23) is the transport–transfer separation for the scalar variance evolution in a 2D flow. For the 3D case, the separation is in the same form. One just needs to change the vectors and the gradient operator in (23) into their corresponding 3D counterparts.

5. Multiscale kinetic energy equation

The formulation of multiscale energetics generally follows from the derivation for the evolutions of K and A. The difference lies in that here we consider our problem in the phase space. Since the basis function $\phi^{\varrho,j}$, for any $0 \le j \le j_2$, is time dependent, and the derivative of $\phi^{\varrho,j}$ does not in general form an orthogonal pair with $\phi^{\varrho,j}$ itself, the local time change terms in the primitive equations need to be pre-treated specially before the energy equations can be formulated. Similar problems also exist in Harrison and Robinson (1978)'s formalism. Appearing on the left hand side of their kinetic energy equation is $\overline{\underline{\mathbf{v}}} \cdot \frac{\partial \overline{\mathbf{v}}}{\partial t}$, not in a form of time change of $\frac{1}{2}\overline{\mathbf{v}} \cdot \overline{\mathbf{v}}$.

To start, first consider $\partial \underline{\mathbf{v}}/\partial t$. Recall that our objective is to develop a diagnostic tool for an existing dataset. Thus every differential term has to be replaced eventually by its difference counterpart. That is to say, we actually do not need to deal with $\partial \underline{\mathbf{v}}/\partial t$ itself. Rather, it is the discretized form (space-dependence suppressed for clarity)

$$\frac{\underline{\mathbf{v}}(t+\Delta t) - \underline{\mathbf{v}}(t-\Delta t)}{2\Delta t} \equiv \delta_t \underline{\mathbf{v}}$$

that we should pay attention to (Δt is the time step size). Viewed as functions of t, $\underline{\mathbf{v}}(t+\Delta t)$ and $\underline{\mathbf{v}}(t-\Delta t)$ make two different series and may be transformed separately. Let

$$\int_0^{\varrho} \underline{\mathbf{v}}^{\sim\varpi}(t+\Delta t)\phi_n^{\varrho,j_2}(t)\,\mathrm{d}t \equiv \hat{\underline{\mathbf{v}}}_{n+}^{\sim\varpi}, \tag{28}$$

$$\int_0^{\varrho} \underline{\mathbf{v}}^{\sim\varpi}(t-\Delta t)\phi_n^{\varrho,j_2}(t)\,\mathrm{d}t \equiv \hat{\underline{\mathbf{v}}}_{n-}^{\sim\varpi}, \tag{29}$$

where ϱ is the periodicity of extension ($\varrho = 1$ and 2 for extensions by periodization and refection, respectively), and define an operator $\hat{\delta}_n$ such that

$$\hat{\delta}_n \hat{\underline{\mathbf{v}}}_n^{\sim\varpi} = \frac{\hat{\underline{\mathbf{v}}}_{n+}^{\sim\varpi} - \hat{\underline{\mathbf{v}}}_{n-}^{\sim\varpi}}{2\Delta t}. \tag{30}$$

$\hat{\delta}_n \underline{\hat{\mathbf{v}}}_n^{\sim\varpi}$ is actually the transform of $\delta_t \underline{\mathbf{v}}$, or the rate of change of $\underline{\hat{\mathbf{v}}}_n^{\sim\varpi}$ on its corresponding scale window. Similarly, define difference operators of the second order as follows:

$$\delta_{t^2}^2 \underline{\mathbf{v}} \equiv \frac{\underline{\mathbf{v}}(t+\Delta t) - 2\underline{\mathbf{v}}(t) + \underline{\mathbf{v}}(t-\Delta t)}{(\Delta t)^2}, \tag{31}$$

$$\hat{\delta}_{n^2}^2 \underline{\hat{\mathbf{v}}}_n^{\sim\varpi} \equiv \int_0^{\varrho} \delta_{t^2}^2 \underline{\mathbf{v}}^{\sim\varpi} \quad \phi_n^{\varrho, j_2}(t)\, \mathrm{d}t. \tag{32}$$

Now take the dot product of $\underline{\hat{\mathbf{v}}}_n^{\sim\varpi}$ with $\hat{\delta}_n \underline{\hat{\mathbf{v}}}_n^{\sim\varpi}$,

$$\begin{aligned}
\underline{\hat{\mathbf{v}}}_n^{\sim\varpi} \cdot \hat{\delta}_n \underline{\hat{\mathbf{v}}}_n^{\sim\varpi} &= \left(-\frac{\underline{\hat{\mathbf{v}}}_{n+}^{\sim\varpi} - 2\underline{\hat{\mathbf{v}}}_n^{\sim\varpi} + \underline{\hat{\mathbf{v}}}_{n-}^{\sim\varpi}}{2} + \frac{\underline{\hat{\mathbf{v}}}_{n+}^{\sim\varpi} + \underline{\hat{\mathbf{v}}}_{n-}^{\sim\varpi}}{2} \right) \cdot \frac{\underline{\hat{\mathbf{v}}}_{n+}^{\sim\varpi} - \underline{\hat{\mathbf{v}}}_{n-}^{\sim\varpi}}{2\Delta t} \\
&= \frac{1}{2\Delta t} \left(\frac{1}{2} \underline{\hat{\mathbf{v}}}_{n+}^{\sim\varpi} \cdot \underline{\hat{\mathbf{v}}}_{n+}^{\sim\varpi} - \frac{1}{2} \underline{\hat{\mathbf{v}}}_{n-}^{\sim\varpi} \cdot \underline{\hat{\mathbf{v}}}_{n-}^{\sim\varpi} \right) - (\Delta t)^2 (\hat{\delta}_{n^2}^2 \underline{\hat{\mathbf{v}}}_n^{\sim\varpi} \cdot \hat{\delta}_n \underline{\hat{\mathbf{v}}}_n^{\sim\varpi}) \\
&= \hat{\delta}_n K_n^{\varpi} - (\Delta t)^2 (\hat{\delta}_{n^2}^2 \underline{\hat{\mathbf{v}}}_n^{\sim\varpi} \cdot \hat{\delta}_n \underline{\hat{\mathbf{v}}}_n^{\sim\varpi}),
\end{aligned} \tag{33}$$

where

$$K_n^{\varpi} = \tfrac{1}{2} \underline{\hat{\mathbf{v}}}_n^{\sim\varpi} \cdot \underline{\hat{\mathbf{v}}}_n^{\sim\varpi} \tag{34}$$

is the kinetic energy at location n (in the phase space) for the window ϖ (the factor 2^{j_2} omitted). Note that K_n^{ϖ} is different from $\hat{K}_n^{\sim\varpi}$. The latter is the multiscale window transform of K, not a concept of "energy". Another quantity that might be confused with K_n^{ϖ} is $K^{\sim\varpi}$, or the field K reconstructed on window ϖ. $K^{\sim\varpi}$ is a property in physical space. It is conceptually different from the phase space-based K_n^{ϖ} for velocity.

Observe that the first term on the right hand side of Eq. (33) is the time change (in difference form) of the kinetic energy on window ϖ at time $2^{-j_2} n$ (scaled by the series length). The second term, which is proportional to $(\Delta t)^2$, is in general very small (of order $O[(\Delta t)^2]$ compared to $\hat{\delta}_n K_n^{\varpi}$). As shown in Appendix A, it could be significant only when processes with scales of grid size are concerned. Besides, it is expressed in a form of discretized Laplacian. We may thereby view it indistinguishably as a kind of subgrid parameterization and merge it into the dissipation terms. The term $\underline{\hat{\mathbf{v}}}_n^{\sim\varpi} \cdot \hat{\delta}_n \underline{\hat{\mathbf{v}}}_n^{\sim\varpi}$, which is akin to Harrison and Robinson's $\overline{\underline{\mathbf{v}}} \cdot \frac{\partial \overline{\underline{\mathbf{v}}}}{\partial t}$, is thus merely the change rate of K_n^{ϖ}, with a small correction of order $(\Delta t)^2$ (t scaled by the series duration).

Terms other than $\partial_t \underline{\mathbf{v}}$ and $\partial_t \rho$ in a 3D primitive equation system do not have time derivatives involved. Multiscale window transforms can be applied directly to every field variable in spite of the spatial gradient operators, if any. To continue the derivation, first take a multiscale window transform of (14),

$$\frac{\partial \hat{w}_n^{\sim\varpi}}{\partial z} + \nabla \cdot \underline{\hat{\mathbf{v}}}_n^{\sim\varpi} = 0. \tag{35}$$

Dot product of the momentum equation reconstructed from (13) on window ϖ with $\underline{\hat{\mathbf{v}}}_n^{\sim\varpi} \phi_n^{\varrho, j_2}(t)$, followed by an integration with respect to t over the domain $[0, \varrho)$, gives the kinetic energy equation for window ϖ. We are now to arrange the right hand side of this equation into a sum of some physically meaningful terms.

Look at the pressure work first. By Eq. (35), it is

$$\begin{aligned}
&\int_0^{\varrho} -\hat{\underline{\mathbf{v}}}_n^{\sim\varpi} \cdot \frac{\nabla P^{\sim\varpi}}{\rho_0} \phi_n^{\varrho,j_2}(t)\,\mathrm{d}t \\
&= -\hat{\underline{\mathbf{v}}}_n^{\sim\varpi} \cdot \frac{\nabla \hat{P}_n^{\sim\varpi}}{\rho_0} = -\frac{1}{\rho_0}\left[\nabla\cdot(\hat{P}_n^{\sim\varpi}\hat{\underline{\mathbf{v}}}_n^{\sim\varpi}) + \frac{\partial}{\partial z}(\hat{P}_n^{\sim\varpi}\hat{w}_n^{\sim\varpi})\right] + \hat{w}_n^{\sim\varpi}\frac{\partial \hat{P}_n^{\sim\varpi}}{\partial z} \\
&= -\frac{1}{\rho_0}\left[\nabla\cdot(\hat{P}_n^{\sim\varpi}\hat{\underline{\mathbf{v}}}_n^{\sim\varpi}) + \frac{\partial}{\partial z}(\hat{P}_n^{\sim\varpi}\hat{w}_n^{\sim\varpi})\right] - \frac{g}{\rho_0}\hat{w}_n^{\sim\varpi}\hat{\rho}_n^{\sim\varpi} \\
&\equiv \Delta_h Q_{P_n^{\varpi}} + \Delta_z Q_{P_n^{\varpi}} - b_n^{\varpi},
\end{aligned} \tag{36}$$

where $\Delta_h Q_{P_n^{\varpi}}$ and $\Delta_z Q_{P_n^{\varpi}}$ (Q_P the pressure flux) are respectively the horizontal and vertical pressure working rates (Q stands for flux, a convention in many fluid mechanics textbooks). The third term, $-b_n^{\varpi} = -\frac{g}{\rho_0}\hat{w}_n^{\sim\varpi}\hat{\rho}_n^{\sim\varpi}$, is the rate of buoyancy conversion between the kinetic and available potential energies on window ϖ.

Next look at the friction terms $\underline{\mathbf{F}}_{mz}$ and $\underline{\mathbf{F}}_{mh}$ in Eq. (13). They stand for the effect of unresolved sub-grid processes. An explicit expression of them is problem-specific, and is beyond of scope of this paper. We will simply write these two terms as $F_{K^{\varpi},z}$ and $F_{K^{\varpi},h}$, which are related to the $\underline{\mathbf{F}}_{mz}$ and $\underline{\mathbf{F}}_{mh}$ in Eq. (13) as follows:

$$F_{K_n^{\varpi},z} = \hat{\underline{\mathbf{v}}}_n^{\sim\varpi} \cdot (\widehat{\underline{\mathbf{F}}_{mz}})_n^{\sim\varpi}, \tag{37}$$

$$F_{K_n^{\varpi},h} = \hat{\underline{\mathbf{v}}}_n^{\sim\varpi} \cdot (\widehat{\underline{\mathbf{F}}_{mh}})_n^{\sim\varpi} + (\Delta t)^2(\hat{\delta}_{n^2}^2\hat{\underline{\mathbf{v}}}_n^{\sim\varpi} \cdot \hat{\delta}_n\hat{\underline{\mathbf{v}}}_n^{\sim\varpi}). \tag{38}$$

In the above, the correction to $\hat{\delta}_n K_n^{\varpi}$ in (33) has been included, as it behaves like a kind of horizontal dissipation.

For the remaining part, the Coriolis force does not contribute to increase K_n^{ϖ}. The nonlinear terms are what we need to pay attention. Specifically, we need to separate

$$\mathrm{NL} = -\hat{\underline{\mathbf{v}}}_n^{\sim\varpi} \cdot \nabla \cdot (\widehat{\underline{\mathbf{v}}\underline{\mathbf{v}}})_n^{\sim\varpi} - \hat{\underline{\mathbf{v}}}_n^{\sim\varpi} \cdot \frac{\partial}{\partial z}(\widehat{w\underline{\mathbf{v}}})_n^{\sim\varpi}$$

into two classes of energetics which represent transport and transfer processes, respectively. This can be achieved by performing a decomposition as we did in Section 4 for the 3D case, with the field variable S in (23) replaced by u and v, respectively. Let

$$\underline{\mathbf{Q}}_h = \lambda_c\hat{\underline{\mathbf{v}}}_n^{\sim\varpi} \cdot (\widehat{\underline{\mathbf{v}}\underline{\mathbf{v}}})_n^{\sim\varpi} = \lambda_c\hat{\underline{\mathbf{v}}}_n^{\sim\varpi} \cdot (\widehat{\underline{\mathbf{v}}\underline{\mathbf{v}}})_n^{\sim\varpi}, \tag{39}$$

$$Q_z = \lambda_c\hat{\underline{\mathbf{v}}}_n^{\sim\varpi} \cdot (\widehat{w\underline{\mathbf{v}}})_n^{\sim\varpi}, \tag{40}$$

where $\lambda_c = \frac{1}{2}$. Further define

$$\Delta_h Q_{K_n^{\varpi}} = -\nabla \cdot \underline{\mathbf{Q}}_h, \tag{41}$$

$$\Delta_z Q_{K_n^{\varpi}} = -\frac{\partial Q_z}{\partial z}, \tag{42}$$

$$T^*_{K_n^\varpi,h} = -\hat{\underline{\mathbf{v}}}_n^{\sim\varpi}\cdot\nabla\cdot(\widehat{\underline{\mathbf{v}}\,\underline{\mathbf{v}}})_n^{\sim\varpi} + \nabla\cdot\underline{\mathbf{Q}}_h, \tag{43}$$

$$T^*_{K_n^\varpi,z} = -\hat{\underline{\mathbf{v}}}_n^{\sim\varpi}\cdot\frac{\partial}{\partial z}(\widehat{w\underline{\mathbf{v}}})_n^{\sim\varpi} + \frac{\partial Q_z}{\partial z}. \tag{44}$$

Then it is easy to show that

$$\mathrm{NL} = (\Delta_h Q_{K_n^\varpi} + \Delta_z Q_{K_n^\varpi}) + (T^*_{K_n^\varpi,h} + T^*_{K_n^\varpi,z}) \tag{45}$$

is the transport–transfer separation for which we are seeking, with

$$T^*_{K_n^\varpi,h} + T^*_{K_n^\varpi,z} = \frac{1}{2}\Big[-\hat{\underline{\mathbf{v}}}_n^{\sim\varpi}\cdot\nabla\cdot(\widehat{\underline{\mathbf{v}}\,\underline{\mathbf{v}}})_n^{\sim\varpi} + \nabla\hat{\underline{\mathbf{v}}}_n^{\sim\varpi} : (\widehat{\underline{\mathbf{v}}\,\underline{\mathbf{v}}})_n^{\sim\varpi} - \frac{\partial}{\partial z}(\widehat{w\underline{\mathbf{v}}})_n^{\sim\varpi}\cdot\hat{\underline{\mathbf{v}}}_n^{\sim\varpi} + \frac{\partial\underline{\mathbf{v}}}{\partial z}\cdot(\widehat{w\underline{\mathbf{v}}})_n^{\sim\varpi}\Big] \tag{46}$$

the perfect transfer.

In (45), although $(T^*_{K_n^\varpi,h} + T^*_{K_n^\varpi,z})$ as a whole is perfect, $T^*_{K_n^\varpi,h}$ or $T^*_{K_n^\varpi,z}$ alone is not. In order to make them so, introduce the following terms:

$$T_{K_n^\varpi,h} = T^*_{K_n^\varpi,h} - \hat{K}_n^{\sim\varpi}\nabla\cdot\hat{\underline{\mathbf{v}}}_n^{\sim\varpi}, \tag{47}$$

$$T_{K_n^\varpi,z} = T^*_{K_n^\varpi,z} - \hat{K}_n^{\sim\varpi}\frac{\partial\hat{w}_n^{\sim\varpi}}{\partial z}, \tag{48}$$

where $\hat{K}_n^{\sim\varpi}$ is the multiscale window transform of $K = \frac{1}{2}\underline{\mathbf{v}}\cdot\underline{\mathbf{v}}$ as a field variable (not K_n^ϖ, the kinetic energy on window ϖ). Clearly $(T^*_{K_n^\varpi,h} + T^*_{K_n^\varpi,z}) = (T_{K_n^\varpi,h} + T_{K_n^\varpi,z})$ by the continuity Eq. (35). It is easy to verify that both $T_{K_n^\varpi,h}$ and $T_{K_n^\varpi,z}$ are perfect transfers using the marginalization property. Decomposition (45) now becomes

$$\mathrm{NL} = (\Delta_h Q_{K_n^\varpi} + \Delta_z Q_{K_n^\varpi}) + (T_{K_n^\varpi,h} + T_{K_n^\varpi,z}). \tag{49}$$

In summary, the kinetic energy evolution on window ϖ is governed by

$$\begin{aligned}\hat{\delta}_n K_n^\varpi = {} & -\nabla\cdot\underline{\mathbf{Q}}_h - \frac{\partial Q_z}{\partial z} + [-\hat{\underline{\mathbf{v}}}_n^{\sim\varpi}\cdot\nabla\cdot(\widehat{\underline{\mathbf{v}}\,\underline{\mathbf{v}}})_n^{\sim\varpi} + \nabla\cdot\underline{\mathbf{Q}}_h - \hat{K}_n^{\sim\varpi}\nabla\cdot\hat{\underline{\mathbf{v}}}_n^{\sim\varpi}] \\ & + \left[-\hat{\underline{\mathbf{v}}}_n^{\sim\varpi}\cdot\frac{\partial}{\partial z}(\widehat{w\underline{\mathbf{v}}})_n^{\sim\varpi} + \frac{\partial Q_z}{\partial z} - \hat{K}_n^{\sim\varpi}\frac{\partial\hat{w}_n^{\sim\varpi}}{\partial z}\right] - \nabla\cdot\left(\hat{\underline{\mathbf{v}}}_n^{\sim\varpi}\frac{\hat{P}_n^{\sim\varpi}}{\rho_0}\right) \\ & - \frac{\partial}{\partial z}\left(\hat{w}_n^{\sim\varpi}\frac{\hat{P}_n^{\sim\varpi}}{\rho_0}\right) - \frac{g}{\rho_0}\hat{w}_n^{\sim\varpi}\hat{\rho}_n^{\sim\varpi} + F_{K_n^\varpi,z} + F_{K_n^\varpi,h},\end{aligned} \tag{50}$$

where $\underline{\mathbf{Q}}_h$ and Q_z are defined in (39) and (40). Symbolically this is,

$$\begin{aligned}\dot{K}_n^\varpi = {} & \Delta_h Q_{K_n^\varpi} + \Delta_z Q_{K_n^\varpi} + T_{K_n^\varpi,h} + T_{K_n^\varpi,z} + \Delta_h Q_{P_n^\varpi} + \Delta_z Q_{P_n^\varpi} \\ & - b_n^\varpi + + F_{K_n^\varpi,z} + F_{K_n^\varpi,h}.\end{aligned} \tag{51}$$

In Appendix D a list of these symbols and their meanings is presented.

6. Multiscale available potential energy equation

To arrive at the multiscale available potential energy equation, take the scale window transform of the time-discretized version of Eq. (16) and multiply it by $c\hat{\rho}_n^{\sim\varpi}$ ($c \equiv g^2/(\rho_0^2 N^2)$). The left hand side becomes, as before,

$$c\hat{\rho}_n^{\sim\varpi}(\widehat{\delta_t \rho})_n^{\sim\varpi} = c\hat{\rho}_n^{\sim\varpi}\hat{\delta}_n\hat{\rho}_n^{\sim\varpi} = \hat{\delta}_n A_n^{\varpi} - (\Delta t)^2 c(\hat{\delta}_{n2}^2\hat{\rho}_n^{\sim\varpi}\cdot\hat{\delta}_n\hat{\rho}_n^{\sim\varpi}),$$

where

$$A_n^{\varpi} = \frac{1}{2}c(\hat{\rho}_n^{\sim\varpi})^2 = \frac{1}{2}\frac{g^2}{\rho_0^2 N^2}(\hat{\rho}_n^{\sim\varpi})^2 \tag{52}$$

(constant multiplier 2^{j_2} omitted) is the available potential energy at location n in the phase space (corresponding to the scaled time $2^{-j_2}n$) for the window ϖ. Compared to $\hat{\delta}_n A_n^{\varpi}$, the correction is of order $(\Delta t)^2$, and could be significant only at small scales, as argued for the kinetic energy case.

For the advection-related terms, the transform followed by a multiplication with $c\hat{\rho}_n^{\sim\varpi}$ yields

$$\begin{aligned}(\mathrm{AD}) &= c\hat{\rho}_n^0\int_0^{\varrho}\left(-\nabla\cdot(\underline{\mathbf{v}}\rho)^{\sim\varpi} - \frac{\partial(w\rho)^{\sim\varpi}}{\partial z}\right)\phi_n^{\varrho,j_2}(t)\,\mathrm{d}t\\ &= -c\hat{\rho}_n^{\sim\varpi}\nabla\cdot(\widehat{\underline{\mathbf{v}}\rho})_n^{\sim\varpi} - c\hat{\rho}_n^{\sim\varpi}\frac{\partial}{\partial z}(\widehat{w\rho})_n^{\sim\varpi}.\end{aligned}$$

As has been explained in Section 4, we need to collect flux-like terms. In the phase space, these terms are:

$$\Delta_h Q_{A_n^{\varpi}} \equiv -\nabla\cdot[\lambda_c c\hat{\rho}_n^{\sim\varpi}(\widehat{\underline{\mathbf{v}}\rho})_n^{\sim\varpi}], \tag{53}$$

$$\Delta_z Q_{A_n^{\varpi}} \equiv -\frac{\partial}{\partial z}[\lambda_c c\hat{\rho}_n^{\sim\varpi}(\widehat{w\rho})_n^{\sim\varpi}], \tag{54}$$

where $\lambda_c = \frac{1}{2}$. With this flux representation, (AD) is decomposed as

$$\begin{aligned}(\mathrm{AD}) = {}& \Delta_h Q_{A_n^{\varpi}} + \Delta_z Q_{A_n^{\varpi}} - [c\hat{\rho}_n^{\sim\varpi}\nabla\cdot(\widehat{\underline{\mathbf{v}}\rho})_n^{\sim\varpi} + \Delta_h Q_{A_n^{\varpi}}]\\ &- \left[c\hat{\rho}_n^{\sim\varpi}\frac{\partial}{\partial z}(\widehat{w\rho})_n^{\sim\varpi} + \Delta_z Q_{A_n^{\varpi}}\right].\end{aligned}$$

The two brackets as a whole represent a perfect transfer process. However, neither of them alone does so. For physical clarity, we need to make some manipulation.

Making use of Eq. (35), and denoting

$$TS_{A_n^{\varpi}} \equiv \lambda_c\hat{\rho}_n^{\sim\varpi}(\widehat{w\rho})_n^{\sim\varpi}\frac{\partial c}{\partial z}, \tag{55}$$

the above decomposition can be written as

$$\begin{aligned}(\mathrm{AD}) &= \Delta_h Q_{A_n^\varpi} + \Delta_z Q_{A_n^\varpi} - [c\hat{\rho}_n^{\sim\varpi}\nabla\cdot(\widehat{\underline{\mathbf{v}}\rho})_n^{\sim\varpi} + \Delta_h Q_{A_n^\varpi} - \lambda_c c((\widehat{\rho^2})_n^{\sim\varpi}\nabla\cdot\hat{\underline{\mathbf{v}}}_n^{\sim\varpi})] \\ &\quad - \left[c\hat{\rho}_n^{\sim\varpi}\frac{\partial}{\partial z}(\widehat{w\rho})_n^{\sim\varpi} + \Delta_z Q_{A_n^\varpi} + TS_{A_n^\varpi} - \lambda_c c\left((\widehat{\rho^2})_n^{\sim\varpi}\frac{\partial\hat{w}_n^{\sim\varpi}}{\partial z}\right)\right] + TS_{A_n^\varpi} \\ &\equiv \Delta_h Q_{A_n^\varpi} + \Delta_z Q_{A_n^\varpi} + T_{A_n^\varpi,\partial_h\rho} + T_{A_n^\varpi,\partial_z\rho} + TS_{A_n^\varpi}, \end{aligned} \tag{56}$$

where $\Delta_h Q_{A_n^\varpi}$ and $\Delta_z Q_{A_n^\varpi}$ are, as we already know, the horizontal and vertical transports. The other pair,

$$T_{A_n^\varpi,\partial_h\rho} \equiv -c\hat{\rho}_n^{\sim\varpi}\nabla\cdot(\widehat{\underline{\mathbf{v}}\rho})_n^{\sim\varpi} - \Delta_h Q_{A_n^\varpi} + \lambda_c c((\widehat{\rho^2})_n^{\sim\varpi}\nabla\cdot\hat{\underline{\mathbf{v}}}_n^{\sim\varpi}) \tag{57}$$

$$T_{A_n^\varpi,\partial_z\rho} \equiv -c\hat{\rho}_n^{\sim\varpi}\frac{\partial}{\partial z}(\widehat{w\rho})_n^{\sim\varpi} - \Delta_z Q_{A_n^\varpi} - TS_{A_n^\varpi} + \lambda_c c\left((\widehat{\rho^2})_n^{\sim\varpi}\frac{\partial\hat{w}_n^{\sim\varpi}}{\partial z}\right) \tag{58}$$

represent two perfect transfer processes, as can be easily verified with the definition in Section 4.

If necessary, $\Delta_h Q_{A_n^\varpi}$ and $T_{A_n^\varpi,\partial_h\rho}$ can be further decomposed as

$$\Delta_h Q_{A_n^\varpi} = \Delta_x Q_{A_n^\varpi} + \Delta_y Q_{A_n^\varpi}, \tag{59}$$

$$T_{A_n^\varpi,\partial_h\rho} = T_{A_n^\varpi,\partial_x\rho} + T_{A_n^\varpi,\partial_y\rho}, \tag{60}$$

where $\Delta_x Q_{A_n^\varpi}$ ($T_{A_n^\varpi,\partial_x\rho}$) and $\Delta_y Q_{A_n^\varpi}$ ($T_{A_n^\varpi,\partial_y\rho}$) are given by the equation for $\Delta_h Q_{A_n^\varpi}$ ($T_{A_n^\varpi,\partial_h\rho}$) with the gradient operator ∇ replaced by $\partial/\partial x$ and $\partial/\partial y$, respectively.

Besides the above fluxes and transfers, there exists an extra term

$$TS_{A_n^\varpi} \equiv \lambda_c\hat{\rho}_n^{\sim\varpi}(\widehat{w\rho})_n^{\sim\varpi}\frac{\partial c}{\partial z} = -\lambda_c c\hat{\rho}_n^{\sim\varpi}(\widehat{w\rho})_n^{\sim\varpi}\frac{\partial(\log N^2)}{\partial z} \tag{61}$$

in the (AD) decomposition (recall $c = g^2/\rho_0^2 N^2$). This term represents an apparent source/sink due to the stationary vertical shear of density, as well as an energy transfer.

Next consider the term $w\frac{N^2\rho_0}{g}$. Recall that N^2 is a function of z only. It is thus immune to the transform. So

$$c\hat{\rho}_n^{\sim\varpi}\frac{\rho_0}{g}\cdot(\widehat{wN^2})_n^{\sim\varpi} = c\frac{N^2\rho_0}{g}\hat{\rho}_n^{\sim\varpi}\hat{w}_n^{\sim\varpi} = \frac{g}{\rho_0}\hat{w}_n^{\sim\varpi}\hat{\rho}_n^{\sim\varpi} = b_n^\varpi, \tag{62}$$

which is exactly the buoyancy conversion between available potential and kinetic energies on window ϖ.

The diffusion terms are treated the same way as before, they are merely denoted as

$$F_{A_n^\varpi,z} = c\hat{\rho}_n^{\sim\varpi}(\widehat{F_{\rho,z}})_n^{\sim\varpi}, \tag{63}$$

$$F_{A_n^\varpi,h} = c\hat{\rho}_n^{\sim\varpi}(\widehat{F_{\rho,h}})_n^{\sim\varpi} + (\Delta t)^2 c(\hat{\delta}_{n^2}^2\hat{\rho}_n^{\sim\varpi}\cdot\hat{\delta}_n\hat{\rho}_n^{\sim\varpi}). \tag{64}$$

Put all the above equations together (with the aid of notations (53), (54) and (61)),

$$\begin{aligned}\hat{\delta}_n A_n^\varpi &= \Delta_h Q_{A_n^\varpi} + \Delta_z Q_{A_n^\varpi} \\ &\quad + [-c\hat{\rho}_n^{\sim\varpi}\nabla\cdot(\widehat{\underline{\mathbf{v}}\rho})_n^{\sim\varpi} - \Delta_h Q_{A_n^\varpi} + \lambda_c c((\widehat{\rho^2})_n^{\sim\varpi}\nabla\cdot\hat{\underline{\mathbf{v}}}_n^{\sim\varpi})]\end{aligned}$$

$$+\left[-c\hat{\rho}_n^{\sim\varpi}\frac{\partial}{\partial z}(\widehat{w\rho})_n^{\sim\varpi}-\Delta_z Q_{A_n^\varpi}-TS_{A_n^\varpi}+\lambda_c c\left((\widehat{\rho^2})_n^{\sim\varpi}\frac{\partial\hat{w}_n^{\sim\varpi}}{\partial z}\right)\right]$$
$$+TS_{A_n^\varpi}+\frac{g}{\rho_0}\hat{w}_n^{\sim\varpi}\hat{\rho}_n^{\sim\varpi}+F_{A_n^\varpi,z}+F_{A_n^\varpi,h}, \tag{65}$$

or, in a symbolic form,

$$\dot{A}_n^\varpi=\Delta_h Q_{A_n^\varpi}+\Delta_z Q_{A_n^\varpi}+T_{A_n^\varpi,\partial_h\rho}+T_{A_n^\varpi,\partial_z\rho}+TS_{A_n^\varpi}+b_n^\varpi+F_{A_n^\varpi,z}+F_{A_n^\varpi,h}. \tag{66}$$

For a list of the meanings of these symbols, refer to Appendix D.

7. Horizontal treatment

As in Fourier analysis, the transform coefficients of MWT contain phase information; unlike Fourier analysis, the energies defined in Section 3.2, which are essentially the transform coefficients squared, still contain phase information. This is fundamentally the same as what happens with the real-valued wavelet analysis, which has been well studied in the context of fluid dynamics (e.g., Farge, 1992; Iima and Toh, 1995).

In the presence of advection, the phase information problem leads to superimposed oscillations with high wavenumbers on the spatial distribution of obtained energetics. This may be understood easily, following an argument in the wavelet energetic analysis of shock waves by Iima and Toh (1995). While in the sampling space[3] the phase oscillation might not be obvious or even ignored because of the discrete nature in time, in the spatial directions it surfaces through a Galilean transformation. Look at the transform (7). The characteristic frequency is $f_c\sim 2^{j_2}$ cycles over the time duration. (Recall the signals are equally sampled on 2^{j_2} points in time.) Now suppose there is a flow with constant speed u_0. The oscillation in time with f_c is then transformed to the horizontal plane with a wavelength on the order of u_0/f_c. Suppose the sampling interval is Δt, the time step size for the dataset. Suppose further the spatial grid size is Δx. In a numerical scheme explicit in advection (which is true for most numerical models), it must be smaller than or equal to $\Delta x/u_0$ to satisfy the CFL condition. So the oscillation has a wavenumber $k_c\sim O(\frac{1}{\Delta x})$ or larger, as $f_c\sim\frac{1}{\Delta t}$. Fig. 2a shows a typical example of the energetic term for the Iceland-Faeroe Frontal variability (cf. Robinson et al., 1996a,b; LR3). Notice how the substantial energetic information (Fig. 2b) is buried in the oscillations with short wavelengths. (The time sampling interval is $10\Delta t$ here.)

The phase oscillation as in Fig. 2a is a technique problem deeply rooted in the nature of localized transforms. It must be eliminated to keep the energetic terms from being blurred. In our case, this is easy to be done. As the characteristic frequency is always 2^{j_2}, the highest for the signal under concern, the oscillation energy peaks at very high wavenumbers, far away from the substantial energy on the spectrum. Except for energetics on the sub-mesoscale

[3] Given a scale window, the MWT transform coefficients form a complete function space. We here refer to it as a sampling space.

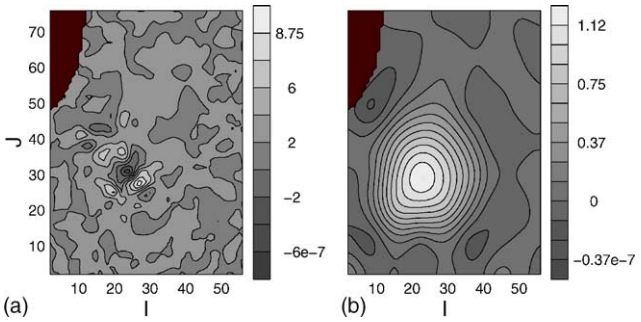

Fig. 2. (a) The total transfer of APE from the large-scale window to the meso-scale window for the Iceland-Faeroe Frontal variability at depth 300 m on August 21, 1993 (cf. LR3, and Robinson et al., 1996a,b). (b) The horizontally filtered map (units: $m^2 s^{-3}$).

window, a horizontal scaling synthesis with a proper upper scale level (lower enough to avoid the phase problem but higher enough to encompass all the substantial information) will give us all what we want. As a scaling synthesis is in fact a low-pass filtering which may also be loosely understood as a "local averaging", we are taking a measure essentially similar to the time averaging approach of Iima and Toh (1995), except that we are here dealing with the horizontal rather than temporal direction. *From now on, all the energetics should be understood to be "locally averaged" with appropriate spatial window bounds, though for notational laconism, we will keep writing them in their original forms.*

One thing that should be pointed out regarding the MWT is that the phase information to be removed is always located around the highest wavenumbers on the energy spectrum. The reason is that in Eq. (7) a scaling basis at the highest scale level j_2 is used for transforms on all windows. This is in contrast to wavelet analyses, in which the larger the scale for the transform, the larger the scale for the phase oscillation (see Iima and Toh, 1995). The special structure of the MWT transform spectrum is very beneficial to the phase removal. Generally no aliasing will happen in separating the substantial processes from the phase oscillation.

8. Connection to the classical formalism

The MS-EVA can be easily connected to a classical energetics formalism, with the aid of the MWT properties presented in Section 2.3, particularly the property of marginalization. For kinetic energy, Appendix C shows that, when

(1) $j_0 = 0$, $j_1 = j_2$ (i.e., only *two-scale windows* are considered), and
(2) a *periodic extension* ($\varrho = 1$) is employed,

Eq. (50) for $\varpi = 0$ and $\varpi = 1$ are reduced respectively to the mean and eddy kinetic energy equations in Harrison and Robinson (1978)'s Reynolds-type energetics adapted for open ocean problems [see Eqs. (A.28) and (A.33)]. For available potential energy, the classical

formulation (2D only) in a statistical context gives the following mean and eddy equations (e.g., Tennekes and Lumley, 1972)

$$\frac{\partial A_{\text{mean}}}{\partial t} + \nabla\cdot(\underline{\bar{\mathbf{v}}}A_{\text{mean}}) = -c\bar{\rho}\nabla\cdot\overline{\underline{\mathbf{v}}'\rho'}, \tag{67}$$

$$\frac{\partial A_{\text{eddy}}}{\partial t} + \nabla\cdot\overline{\left(\underline{\mathbf{v}}\frac{1}{2}c\rho'^2\right)} = -c\overline{\rho'\underline{\mathbf{v}}'}\cdot\nabla\bar{\rho}, \tag{68}$$

where $A_{\text{mean}} = \frac{1}{2}c\bar{\rho}^2$, $A_{\text{eddy}} = \frac{1}{2}c\overline{(\rho)'^2}$. Eqs. (67) and (68) can be adapted for open ocean problems by modifying the time rates of change using the approach by Harrison and Robinson (1978). Following the same way as that for KE, these modified equations can be derived directly from the MS-EVA APE Eq. (65) under the above two assumptions.

It is of interest to notice that the multiscale energy Eqs. (50) and (65) appear in the same form for different windows. This is in contrast to the classical Reynolds-type formalism, where the eddy energetics are usually quite different in form from their mean counterparts. This difference disappears if the averaging and deviating operators in (67), (68), (A.28), and (A.33), are rewritten in terms of multiscale window transform. One might have been using the averaging-deviating approach for years without realizing that they actually belong to a kind of transform and synthesis.

Consequently, the classical energetic formalism is equivalent to our MS-EVA under a two-window decomposition with $j_0 = 0$ and $\varrho = 1$. The latter can be viewed as a generalization of the former for GFD processes occurring on arbitrary scale windows. The MS-EVA capabilities, however, are not limited to this. In (67) and (68), the rhs terms, or transfers as usually interpreted, sum to $-c\nabla\cdot(\overline{\rho\rho'\underline{\mathbf{v}}'})$, which is generally not zero. That is to say, these "transfers" are not "perfect". They still contain some information of transport processes. Our MS-EVA, in contrast, produces transfers on a different basis. The concept of perfect transfer defined through transfer–transport separation allows us to make physically consistent inference of the energy redistribution through scale windows. In this sense, the MS-EVA has an aspect which is distinctly different from the classical formalism.

9. Interaction analysis

Different from the classical energetics, a localized energy transfer involves not only interactions between scales, but also interactions between locations in the sampling space. We have already seen this in the definition of perfect transfer processes. A schematic is shown in Fig. 3. The addition of sampling space interaction compounds greatly the transfer problem, as it mingles the inter-scale interactions with transfers within the same scale window, and as a result, useful information tends to be disguised, especially for those processes such as instabilities. We must single out this part in order to have the substantial dynamics up front.

In the MS-EVA, transfer terms are expressed in the form of triple products. They are all like

$$T(\varpi, n) = \hat{\mathcal{R}}_n^{\sim\varpi}(\widehat{pq})_n^{\sim\varpi}, \qquad \text{for } \mathcal{R},\ p, q \in V_{\varrho, j_2}, \tag{69}$$

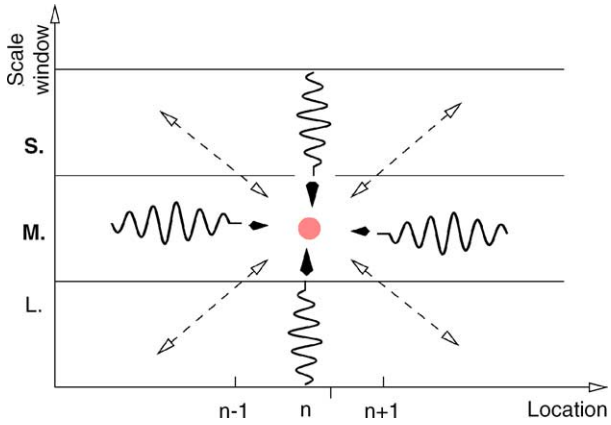

Fig. 3. A schematic of the energy transfers toward a meso-scale process at location n. Depicted are the transfers from different time scales at the same location (vertical arrows), transfers from surrounding locations at the same scale level (horizontal arrows), and transfers from different scales at different locations (dashed arrows).

a form which we call *basic transfer function* for reference convenience. Using the representation (9), it may be expanded as

$$T(\varpi, n) = \sum_{\varpi_1,\varpi_2} \sum_{n_1,n_2} Tr(n, \varpi | n_1, \varpi_1; n_2, \varpi_2), \tag{70}$$

where

$$Tr(n, \varpi | n_1, \varpi_1; n_2, \varpi_2) = \hat{\mathcal{R}}_n^{\sim\varpi} \cdot [\hat{p}_{n_1}^{\sim\varpi_1} \hat{q}_{n_2}^{\sim\varpi_2} (\widehat{\phi_{n_1}^{\varrho,j_2} \phi_{n_2}^{\varrho,j_2}})_n^{\sim\varpi}], \tag{71}$$

and the sums are over all the possible windows and locations. $Tr(n, \varpi | n_1, \varpi_1; n_2, \varpi_2)$ is a *unit expression* of the interaction amongst the triad $(n, w; n_1, w_1; n_2, w_2)$. It stands for the rate of energy transferred to (n, ϖ) from the interaction of (n_1, ϖ_1) and (n_2, ϖ_2). We will refer to the pairs (n_1, w_1) and (n_2, w_2) as the *giving modes*, and (n, w) the *receiving mode*, a naming convention after Iima and Toh (1995).

Theoretically, expansion of a basic transfer function in terms of unit expression allows one to trace back to all the sources that contributes to the transfer. Practically, however, it is not an efficient way because of the huge number of mode combinations and hence the huge number of triads. In our problem, such a detailed analysis is not at all necessary. If (70) is modified such that some terms are combined, the computational redundancy would be greatly reduced whereas the physical interpretation could be even clearer. We now present the modification.

Look at the meso-scale window ($\varpi = 1$) first. It is of particular importance because it mediates between the large scales and sub-mesoscales on a spectrum. For a field p, make the decomposition

$$p = \hat{p}_n^{\sim 1} \phi_n^{\varrho,j_2}(t) + p_{*^1} = p^{\sim 0} + \hat{p}_n^{\sim 1} \phi_n^{\varrho,j_2}(t) + p_{*^1}^{\sim 1} + p^{\sim 2}, \tag{72}$$

where

$$p_{*^1} = p - \hat{p}_n^{\sim 1} \phi_n^{\varrho,j_2}(t) \tag{73}$$

and $p^{\sim 1}_{*^1}$ is the meso-scale part of p_{*^1},

$$p^{\sim 1}_{*^1} = p^{\sim 1} - \hat{p}^{\sim 1}_n \phi^{\varrho, j_2}_n = \sum_{i \in \mathcal{N}^{j_2}_{\varrho}, i \neq n} \hat{p}^{\sim 1}_i \phi^{\varrho, j_2}_i. \tag{74}$$

The new interaction analysis concerns the relationship between scales and locations, instead of between triads. The advantage of this is that we do not have to resort to those triad modes, which may not have physical correspondence in the large-scale window, to make interpretation. Note not any $\hat{p}^{\sim 1}_n \phi^{\varrho, j_2}_n$ can convincingly characterize $p^{\sim 1}(t)$ at location n. But in this context, as the basis function $\phi^{\varrho, j_2}_n(t)$ we choose is a very localized one (localization order delimited, see L02), we expect the removal of $\hat{p}^{\sim 1}_n \phi^{\varrho, j_2}_n$ will effectively (though not totally) eliminate from $p^{\sim 1}$ the contribution from location n. This has been evidenced in the example of of a meridional velocity series v (Fig. 4), where at $n = 384$, $v^{\sim 1}_{*^1}$ is only about 6% ($|\frac{-0.0106}{0.17}|$) of the $v^{\sim 1}$ in magnitude, while at other locations v and $v^{\sim 1}_{*^1}$ are almost the same (fluctuations negligible around n). Therefore, one may practically, albeit not perfectly, take $\hat{p}^{\sim 1}_n \phi^{\varrho, j_2}_n$ as the meso-scale part of p with contribution from location n only (corresponding to $t = 2^{-j_2} n$), and $p^{\sim 1}_{*^1}$ the part from all locations other than n. Note $p^{\sim 1}_{*^1}$ has an n-dependence. For notational clarity, it is suppressed henceforth.

Likewise, for field $q \in V_{\varrho, j_2}$, it can also be decomposed as

$$q = q^{\sim 0} + q^{\sim 1} + q^{\sim 2} \tag{75}$$

$$q = q^{\sim 0} + \hat{q}^{\sim 1}_n \phi^{\varrho, j_2}_n + q^{\sim 1}_{*^1} + q^{\sim 2}, \tag{76}$$

with interpretation analogous to that of $p^{\sim 1}_{*^1}$ for the starred term. The decompositions for p and q yield an analysis of the basic transfer function $T(1, n) = \hat{\mathcal{R}}^{\sim 1}_n \cdot (\widehat{pq})^{\sim 1}_n$ into an interaction matrix, which is shown in Table 1. In this matrix, L stands for large-scale window and S for sub-mesoscale window (all locations). $\mathrm{M_n}$ is used to denote the meso-scale contribution from location n, while M_* signifies the meso-scale contributions *other than* that location. Among these interactions, $\mathrm{M_n}$–M_* and M_*–M_* contribute to $T(1, n)$ from the same scale window (meso-scale, without inter-scale transfers being involved. We may sub-total all the resulting 16 terms into 5 more meaningful terms:

$$\begin{aligned} T^{0\to 1}_n &= \hat{\mathcal{R}}^{\sim 1}_n \cdot [(\widehat{p^{\sim 0} q^{\sim 0}})^{\sim 1}_n + \hat{q}^{\sim 1}_n (\widehat{p^{\sim 0} \phi^{\varrho, j_2}_n})^{\sim 1}_n + (\widehat{p^{\sim 0} q^{\sim 1}_{*^1}})^{\sim 1}_n \\ &\quad + \hat{p}^{\sim 1}_n (\widehat{\phi^{\varrho, j_2}_n q^{\sim 0}})^{\sim 1}_n + (\widehat{p^{\sim 1} q^{\sim 0}})^{\sim 1}_n] \\ &= \hat{\mathcal{R}}^{\sim 1}_n \cdot [(\widehat{p^{\sim 0} q^{\sim 0}})^{\sim 1}_n + (\widehat{p^{\sim 1} q^{\sim 0}})^{\sim 1}_n + (\widehat{p^{\sim 0} q^{\sim 1}})^{\sim 1}_n] \end{aligned} \tag{77}$$

$$\begin{aligned} T^{2\to 1}_n &= \hat{\mathcal{R}}^{\sim 1}_n \cdot [\hat{p}^{\sim 1}_n (\widehat{\phi^{\varrho, j_2}_n q^{\sim 2}})^{\sim 1}_n + (\widehat{p^{\sim 1}_{*^1} q^{\sim 2}})^{\sim 1}_n + \hat{q}^{\sim 1}_n (\widehat{p^{\sim 2} \phi^{\varrho, j_2}_n})^{\sim 1}_n \\ &\quad + (\widehat{p^{\sim 2} q^{\sim 1}_{*^1}})^{\sim 1}_n + (\widehat{p^{\sim 2} q^{\sim 2}})^{\sim 1}_n] \\ &= \hat{\mathcal{R}}^{\sim 1}_n \cdot [(\widehat{p^{\sim 1} q^{\sim 2}})^{\sim 1}_n + (\widehat{p^{\sim 2} q^{\sim 2}})^{\sim 1}_n + (\widehat{p^{\sim 2} q^{\sim 1}})^{\sim 1}_n] \end{aligned} \tag{78}$$

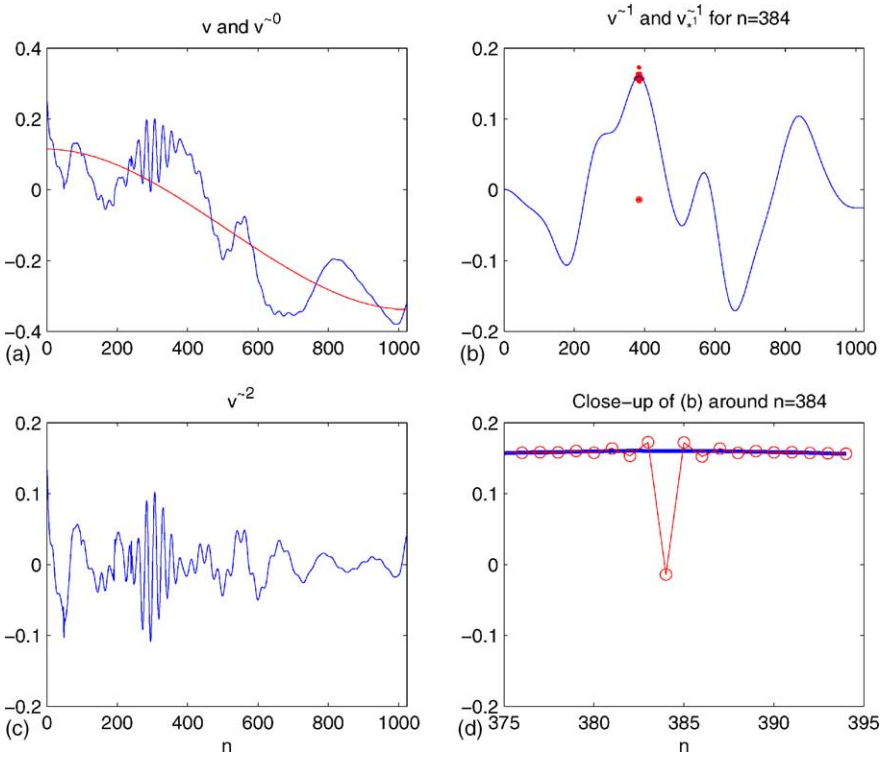

Fig. 4. A typical time series of v (in cm/s) from the Iceland-Faeroe Frontal variability simulation (point (35, 43, 2). Refer to Fig. 2 for the location) and its derived series (cf. LR3). There are $2^{j_2} = 1024$ data points, and scale windows are chosen such that $j_0 = 0$ and $j_1 = 4$. The original series v and its large-scale reconstruction $v^{\sim 0}$ are shown in (a), and the meso-scale and sub-mesoscale are plotted in (b) and (c) respectively. Also plotted in (b) is the "starred" series (dotted) $v_*^{\sim 1}$ for location $n = 384$. (d) is the close-up of (b) around $n = 384$. Apparently, $v_*^{\sim 1}$ is at least one order smaller than $v^{\sim 1}$ in size at that point, while these two are practically the same at other points. Location n corresponds to a scaled time $t = 2^{-j_2} n$ (here forecast day 8).

$$T_n^{0\oplus 2\to 1} = \hat{\mathcal{R}}_n^{\sim 1} \cdot [\widehat{(p^{\sim 2} q^{\sim 0})}_n^{\sim 1} + \widehat{(p^{\sim 0} q^{\sim 2})}_n^{\sim 1}] \tag{79}$$

$$T_{n\to n}^{1\to 1} = \hat{\mathcal{R}}_n^{\sim 1} \cdot \left[\hat{p}_n^{\sim 1} \hat{q}_n^{\sim 1} \widehat{(\phi_n^{\varrho, j_2})}_n^{2^{\sim 1}}\right] \tag{80}$$

$$T_{\text{other}\to n}^{1\to 1} = \hat{\mathcal{R}}_n^{\sim 1} \cdot [\widehat{(p^{\sim 1} q_*^{\sim 2})}_n^{\sim 1} + \hat{q}_n^{\sim 1} \widehat{(p_*^{\sim 2} \phi_n^{\varrho, j_2})}_n^{\sim 1}]. \tag{81}$$

Table 1
Interaction matrix for basic transfer function $T(1, n) = \hat{\mathcal{R}}_n^{\sim 1} \cdot (\widehat{pq})_n^{\sim 1}$

	$p^{\sim 0}$	$\hat{p}_n^{\sim 1} \phi_n^{\varrho, j_2}$	$p_*^{\sim 1}$	$p^{\sim 2}$
$q^{\sim 0}$	L–L	L–M_n	L–M_*	L–S
$\hat{q}_n^{\sim 1} \phi_n^{\varrho, j_2}$	M_n–L	M_n–M_n	M_n–M_*	M_n–S
$q_*^{\sim 1}$	M_*–L	M_*–M_n	M_*–M_*	M_*–S
$q^{\sim 2}$	S–L	S–M_n	S–M_*	S–S

If necessary, $T^{1\to1}_{n\to n}$ and $T^{1\to1}_{\text{other}\to n}$ may also be combined to one term. The result is denoted as $T^{1\to1}_n$.

The physical interpretations of above five terms are embedded in the naming convention of the superscripts, which reveals how energy is transferred to mode $(1, n)$ from other scales. Specifically, $T^{0\to1}_n$ and $T^{2\to1}_n$ are transfer rates from windows 0 and 1, respectively, and $T^{0\oplus2\to1}_n$ is the contribution from the window 0–window 2 interaction over the meso-scale range. The last two terms, $T^{1\to1}_{n\to n}$ and $T^{1\to1}_{\text{other}\to n}$, sum up to $T^{1\to1}_n$, which represents the part of transfer from the same window.

Above are the interaction analysis for $T(1, n)$. Using the same technique, one can obtain a similar analysis for $T(0, n)$ and $T(2, n)$. The results are supplied in Appendix B.

What merits mentioning is that different analyses may be obtained by making different sub-grouping for Eq. (70). The rule of thumb here is to try to avoid those starred terms as in Eq. (81), which makes the major overhead in computation (in terms of either memory or CPU usage). In the above analyses, say the meso-scale analysis, if a whole perfect transfer is calculated, the sum of those terms in the form of $T^{1\to1}_{n\to n}$ will vanish by the definition of perfect transfer processes. This also implies that the sum of those transfer functions in the form of $T^{1\to1}_{\text{other}\to n}$ will be equal to the sum of terms in the same form but with all the stars dropped. Hence in performing interaction analysis for a perfect transfer process, we may simply ignore the stars for the corresponding terms. But if it is an arbitrary transfer term which does not necessarily represent a perfect transfer process (e.g, $TS_{A^1_n}$), the starred-term-caused heavy computational overhead will still be a problem.

In practice, this overhead may be avoided under certain circumstances. Recall that we have built a highly localized scaling basis function ϕ. For any $p \in V_{\varrho,j_2}$, it yields a function $p(t)\phi^{\varrho,j_2}_n(t)$ with an effective support of the order of the grid size. The large- or meso-scale transform of this function is thence negligible, should j_1 be smaller than j_2 by some considerable number (3 is enough). Only when it is in the sub-mesoscale window need we really compute the starred term. An example with a typical time series of ρ and u is plotted in Fig. 5. Apparently, for the large-scale and meso-scale cases, $\hat{\rho}^{\sim0}_n\widehat{(u\phi^{\varrho,j_2}_n)}^{\sim0}_n$ and $\hat{\rho}^{\sim1}_n\widehat{(u\phi^{\varrho,j_2}_n)}^{\sim1}_n$ (red circles) are very small and hence $\widehat{(\rho^{\sim0}{}_{*^0}u)}^{\sim0}_n$ and $\widehat{(\rho^{\sim1}_{*^1}u)}^{\sim1}_n$ can be approximated by $\widehat{(\rho^{\sim0}u)}^{\sim0}_n$ and $\widehat{(\rho^{\sim1}u)}^{\sim1}_n$, respectively. This approximation fails only in the sub-mesoscale case, where the corresponding two parts are of the same order.

It is of interest to give an estimation of the relative importance of all these interaction terms obtained thus far. For the mesoscale transfer function $T(1, n)$, $T^{0\oplus2\to1}_n$ is generally not significant (compared to other terms). This is because, on a spectrum, if two processes are far away from each other (as is the case for large scale and sub-mesoscale), they are usually separable and the interaction are accordingly very weak. Even if there exists some interaction, the spawned new processes generally stay in their original windows, seldom going into between. Apart from $T^{0\oplus2\to1}_n$, all the others are of comparable sizes, though more often than not $T^{0\to1}_n$ dominates the rest (e.g., Fig. 6b).

For the large-scale window, things are a little different. This time it is term $T^{2\to0}_n$ that is not significant, with the same reason as above. But term $T^{1\oplus2\to0}_n$ is in general not negligible. In this window, the dominant energy transfer is usually not from other scales, but from other locations at the same scale level. Mathematically this is to say, $T^{0\to0}_{\text{other}\to n}$ usually dominates

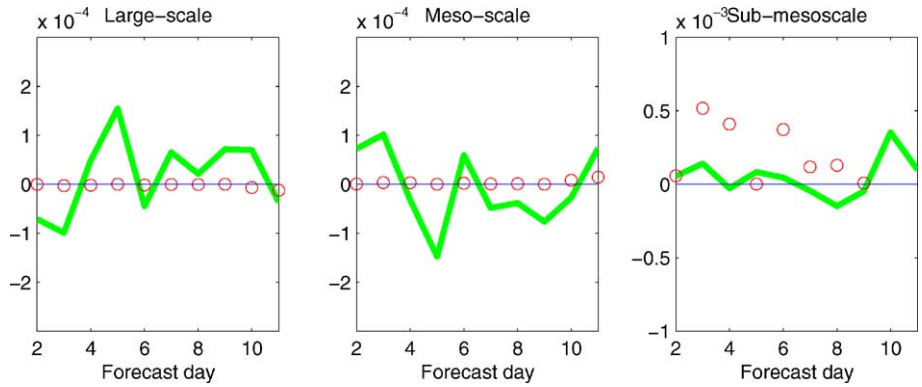

Fig. 5. An example showing relative importance of the decomposed terms from $T_{A_n^{\varpi},\partial_h\rho}$. Data source: same as that in Fig. 4 (zonal velocity only). Units: kg/m^{2}s. Left: $(\widehat{\rho_{*^0}^{\sim 0}u})_n^{\sim 0}$ (heavy solid line) and $\hat{\rho}_n^{\sim 0}(\widehat{u\phi_n^{\varrho,j_2}})_n^{\sim 0}$ (circle); middle: $(\widehat{\rho_{*^1}^{\sim 1}u})_n^{\sim 1}$ (heavy solid line) and $\hat{\rho}_n^{\sim 1}(\widehat{u\phi_n^{\varrho,j_2}})_n^{\sim 1}$ (circle); right: $(\widehat{\rho_{*^2}^{\sim 1}u})_n^{\sim 2}$ (heavy solid line) and $\hat{\rho}_n^{\sim 2}(\widehat{u\phi_n^{\varrho,j_2}})_n^{\sim 2}$ (circle). Obviously, the $(\widehat{\rho_{*^w}^{\sim w}u})_n^{\sim w}$ in the decomposition $(\widehat{\rho^{\sim w}u})_n^{\sim w} = (\widehat{\rho_{*^w}^{\sim w}u})_n^{\sim w} + \hat{\rho}_n^{\sim w}(\widehat{u\phi_n^{\varrho,j_2}})_n^{\sim w}$ can be well approximated by $(\widehat{\rho^{\sim w}u})_n^{\sim w}$ for windows $w = 0, 1$.

the other terms. This is understandable since a large-scale feature results from interactions with modes covering a large range of location on the time series. If each location contributes even a little bit, the grand total could be huge. This fact is seen in the example in Fig. 6a.

By the same argument as above, within the sub-mesoscale window, the dominant term is $T_n^{1\to 2}$. But $T_n^{0\oplus 1\to 2}$ could be of some importance also. In comparison to these two, $T_n^{0\to 2}$ and $T_n^{2\to 2} = T_{\text{other}\to n}^{2\to 2} + T_{n\to n}^{2\to 2}$ are not significant.

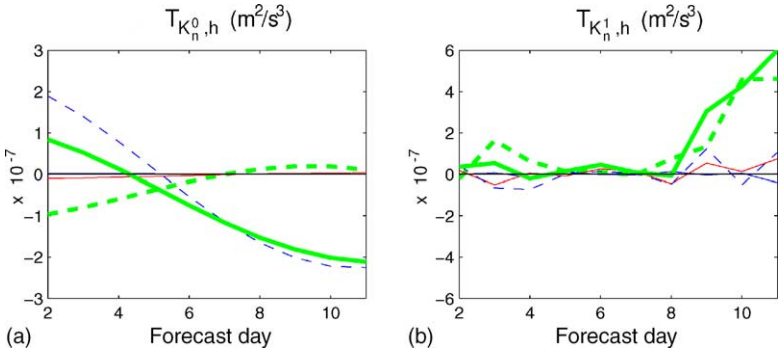

Fig. 6. An example showing the relative importance of analytical terms of $T_{K_n^{\varpi},h}$ at 10 (time) locations. The data source and parameter choice are the same as that of Fig. 4. Here the constant factor 2^{j_2} has been multiplied. (a) Analysis of $T_{K_n^0,h}$ (thick solid): $T_{K_n^0,h}^{1\to 0}$ (thick dashed), $T_{K_n^0,h}^{2\to 0}$ (solid), and $T_{K_n^0,h}^{0\to 0}$ (dashed). $T_{K_n^0,h}^{1\oplus 2\to 0}$ is also shown but unnoticeable. (b) Analysis of $T_{K_n^1,h}$ (thick solid): $T_{K_n^1,h}^{0\to 1}$ (thick dashed), $T_{K_n^1,h}^{2\to 1}$ (solid), and $T_{K_n^1,h}^{1\to 1}$ (dashed). $T_{K_n^1,h}^{0\oplus 2\to 1}$ is also shown but unnoticeable.

We finish up this section with two observations of Fig. 6. (1) During the forecast days, $T_{K_n^0,h}$ and $T_{K_n^0,h}^{1\to 0}$ are almost opposite in sign. That is to say, the transfer term without interaction analysis could be misleading in inter-scale energy transfer study. (2) The transfer rates change with time continuously. Analyses in a global time framework apparently do not work here, as application of a global analysis basically eliminates the time structure. This from one aspect demonstrates the advantage of MS-EVA in diagnosing real problems.

10. Process classification and energetic scenario

From the above analysis, energetic processes for a geophysical fluid system can be generally classified into the following four categories: transport, perfect transfer, buoyancy conversion, and dissipation/diffusion. (The apparent source/sink in the multiscale APE equation is usually orders smaller than other terms and hence is negligible.) Dissipation/diffusion is beyond the scope of this paper. All the remaining categories belong to some "conservative" processes. Transport vanishes if integrated over a closed domain; perfect transfer summarizes to zero over scale windows followed by a marginalization in the sampling space; buoyancy conversion serves as a protocol between the two types of energy.

The energetic scenario is now clear. If a system is viewed as defined in a space which includes physical space, phase space, and the space of energy type, then transport, transfer and buoyancy conversion are three mechanisms that redistribute energy through this super space. In a two-window decomposition, communication between the windows are achieved via $T_K^{0\leftrightarrow 1}$ and $T_A^{0\leftrightarrow 1}$. (Here T stands for total transfer, and the superscript $0 \leftrightarrow 1$ for either $0 \to 1$ or $1 \to 0$.) the two types of energy are converted on each window; while transport brings every point to connection in the physical space. The whole scenario is like an energetic cycle, which is pictorially presented in the left part of Fig. 7 (with all the sub-mesoscale window-related arrows dropped), where arrows are utilized to indicate energy flows, and box and discs for the KE and APE, respectively.

When the number of windows increase from 2 to 3, the scenario of energetic processes becomes much more complex. Besides the addition of a sub-mesoscale window, and the corresponding transports, conversions, and the window 1–2 and 0–2 transfers, another process appears. Schematized in Fig. 7 by dashed arrows, it is a transfer to a window from the interaction between another two windows. In traditional jargon, it is a "non-local" transfer, i.e., a transfer between two windows which are not adjacent in the phase space. We do not adopted this language as by "local" in this paper we refer to a physical space context. If the number of windows increases, these "nonlocal" transfers will compound the problem very much, and as a result, the complexity of the energetic scenario will increase exponentially. In a sense, this is one of the reasons why an eddy decomposition is preferred to a wave decomposition for multiscale energy study.

11. Multiscale enstrophy equation

Vorticity dynamics is an integral part of the MS-EVA. In this section we develop the laws for multiscale enstrophy evolution, which are derived from the vorticity equation.

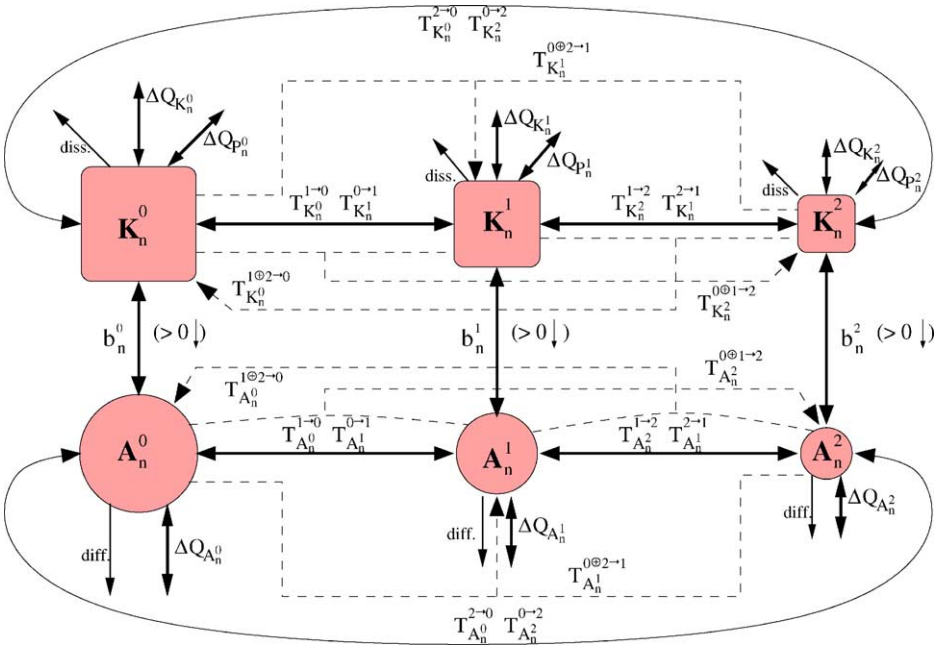

Fig. 7. A schematic of the multiscale energetics for location n. Arrows are used to indicate the energy flow, both in the physical space and phase space, and labeled over these arrows are the processes associated with the flow. The symbols adopted are the same as those listed in Table A.2, except that transport and transfer are the total processes. Interaction analyses are indicated in the superscripts of the T-terms, whose interpretation is referred to Section 9. For clarity, transfers from the same window are not shown. From this diagram, we see that transports ($\Delta Q_{K_n^\varpi}$, $\Delta Q_{P_n^\varpi}$, $\Delta Q_{A_n^\varpi}$, for windows $\varpi = 0, 1, 2$) occur between different locations in physical space, while transfers (the T-terms) mediate between scale windows in phase space. The connection between the two types of energy is established through buoyancy conversion (positive if in the direction as indicated in the parenthesis), which invokes neither scale–scale interactions nor location–location energy exchange.

The equation for vorticity $\zeta = \underline{\mathbf{k}} \cdot \nabla \wedge \underline{\mathbf{v}}$ is obtained by crossing the momentum Eq. (13) followed by a dot product with $\underline{\mathbf{k}}$,

$$\frac{\partial \zeta}{\partial t} = \underline{\mathbf{k}} \cdot \nabla \wedge w \frac{\partial \underline{\mathbf{v}}}{\partial z} - \underline{\mathbf{k}} \cdot \nabla \wedge [(f+\zeta)\underline{\mathbf{k}} \wedge \underline{\mathbf{v}}] + F_{\zeta,z} + F_{\zeta,h}, \tag{82}$$

where $F_{\zeta,z}$ and $F_{\zeta,h}$ denote respectively the vertical and horizontal diffusion. Making use of the continuity Eq. (14), we get,

$$\frac{\partial \zeta}{\partial t} = \underbrace{-\nabla \cdot (\underline{\mathbf{v}}\zeta) - \frac{\partial}{\partial z}(w\zeta)}_{\text{(I)}} \underbrace{-\beta v}_{\text{(II)}} \underbrace{+(f+\zeta)\frac{\partial w}{\partial z}}_{\text{(III)}} \underbrace{+\underline{\mathbf{k}} \cdot \frac{\partial \underline{\mathbf{v}}}{\partial z} \wedge \nabla w}_{\text{(IV)}} \underbrace{+F_{\zeta,z} + F_{\zeta,h}}_{\text{(V)}}. \tag{83}$$

Here $\beta = \partial f/\partial y$ is a constant if a β-plane is approximation is assumed. But in general, it does not need to be so. In Eq. (83), there are five mechanisms that contribute to the change of relative vorticity ζ (e.g., Spall, 1989). Apparently, term (I) is the advection of ζ by the flow, and term (V) the diffusion. β-Effect comes into play through term (II). It is the advection

of planetary vorticity f by meridional velocity v. Vortex tubes may stretch or shrink. The vorticity gain or loss due to stretching or shrinking is represented in term (III). Vortex tube may also tilt. Term (IV) results from such a mechanism.

Enstrophy is the *"energy" of vorticity*, a positive measure of rotation. It is the square of vorticity: $\mathcal{Z} = \frac{1}{2}\zeta^2$. Following the same practice for multiscale energies, the enstrophy on scale window ϖ at time location n is defined as (factor 2^{j_2} omitted for brevity)

$$\mathcal{Z}_n^{\varpi} = \frac{1}{2}(\hat{\zeta}_n^{\sim\varpi})^2. \tag{84}$$

The evolution of $\mathcal{Z}_n^{\varpi}$ is derived from Eq. (83).

As before, first discretize the only time derivative term in Eq. (83), $\partial\zeta/\partial t$, to $\delta_t\zeta$. Take a multiscale transform of the resulting equation and then multiply it by $\hat{\zeta}_n^{\sim\varpi}$. The left hand side results in the evolution $\hat{\delta}_n\mathcal{Z}_n^{\varpi}$ plus a correction term which is of the order Δt^2, Δt being the time spacing of the series. Merging the correction term into the horizontal diffusion, we get an equation

$$\begin{aligned}\dot{\mathcal{Z}}_n^{\varpi} = & \underbrace{-\hat{\zeta}_n^{\sim\varpi}\left[\nabla\cdot(\widehat{\underline{\mathbf{v}}\zeta})_n^{\sim\varpi} + \frac{\partial(\widehat{w\zeta})_n^{\sim\varpi}}{\partial z}\right]}_{\text{(AD)}} - \beta\hat{\zeta}_n^{\sim\varpi}\hat{v}_n^{\sim\varpi} + f\hat{\zeta}_n^{\sim\varpi}\left(\widehat{\frac{\partial w}{\partial z}}\right)_n^{\sim\varpi} \\ & + \hat{\zeta}_n^{\sim\varpi}\left(\widehat{\zeta\frac{\partial w}{\partial z}}\right)_n^{\sim\varpi} + \hat{\zeta}_n^{\sim\varpi}\underline{\mathbf{k}}\cdot\left(\widehat{\frac{\partial\underline{\mathbf{v}}}{\partial z}\wedge\nabla w}\right)_n^{\sim\varpi} + F_{\mathcal{Z}_n^{\varpi},z} + F_{\mathcal{Z}_n^{\varpi},h}.\end{aligned}$$

Again, $F_{\mathcal{Z}_n^{\varpi},z}$ and $F_{\mathcal{Z}_n^{\varpi},h}$ here are just symbolic representations of the vertical and horizontal diffusions. Following the practice in deriving the APE equation, the process represented by the advection-related terms (AD) can be decomposed into a sum of transport processes and transfer processes. Denote

$$\Delta_h Q_{\mathcal{Z}_n^{\varpi}} = -\nabla\cdot[\lambda_c\hat{\zeta}_n^{\sim\varpi}(\widehat{\underline{\mathbf{v}}\zeta})_n^{\sim\varpi}], \tag{85}$$

$$\Delta_z Q_{\mathcal{Z}_n^{\varpi}} = -\frac{\partial}{\partial z}[\lambda_c\hat{\zeta}_n^{\sim\varpi}(\widehat{w\zeta})_n^{\sim\varpi}] \tag{86}$$

then it is

$$\begin{aligned}\mathrm{AD} = & \Delta_h Q_{\mathcal{Z}_n^{\varpi}} + \Delta_z Q_{\mathcal{Z}_n^{\varpi}} + [-\Delta_h Q_{\mathcal{Z}_n^{\varpi}} - \hat{\zeta}_n^{\sim\varpi}\nabla\cdot(\widehat{\underline{\mathbf{v}}\zeta})_n^{\sim\varpi} + \lambda_c(\widehat{\zeta^2})_n^{\sim\varpi}\nabla\cdot\hat{\underline{\mathbf{v}}}_n^{\sim\varpi}] \\ & + \left[-\Delta_z Q_{\mathcal{Z}_n^{\varpi}} - \hat{\zeta}_n^{\sim\varpi}\frac{\partial(\widehat{w\zeta})_n^{\sim\varpi}}{\partial z} + \lambda_c(\widehat{\zeta^2})_n^{\sim\varpi}\frac{\partial\hat{w}_n^{\sim\varpi}}{\partial z}\right] \\ \equiv & \Delta_h Q_{\mathcal{Z}_n^{\varpi}} + \Delta_z Q_{\mathcal{Z}_n^{\varpi}} + T_{\mathcal{Z}_n^{\varpi},\partial_h\zeta} + T_{\mathcal{Z}_n^{\varpi},\partial_z\zeta},\end{aligned}$$

where $\Delta_h Q_{\mathcal{Z}_n^{\varpi}}$ and $\Delta_z Q_{\mathcal{Z}_n^{\varpi}}$ represent the horizontal and vertical transports, and $T_{\mathcal{Z}_n^{\varpi},\partial_h\zeta}$, $T_{\mathcal{Z}_n^{\varpi},\partial_z\zeta}$ the transfer rates for two distinct processes. It is easy to prove that both of these processes are perfect transfers. Note the multiscale continuity Eq. (35) has been used in obtaining the above form of decomposition. If necessary, $\Delta_h Q_{\mathcal{Z}_n^{\varpi}}$ and $T_{\mathcal{Z}_n^{\varpi},\partial_h\zeta}$ may be further decomposed into contributions from x and y directions, respectively.

The enstrophy equation now becomes, after some algebraic manipulation,

$$\begin{aligned}\dot{\mathcal{Z}}_n^{\varpi} &= \Delta_h Q_{\mathcal{Z}_n^{\varpi}} + \Delta_z Q_{\mathcal{Z}_n^{\varpi}} + [-\Delta_h Q_{\mathcal{Z}_n^{\varpi}} - \hat{\zeta}_n^{\sim\varpi}\nabla\cdot(\widehat{\underline{\mathbf{v}}\zeta})_n^{\sim\varpi} + \lambda_c(\widehat{\zeta^2})_n^{\sim\varpi}\nabla\cdot\hat{\underline{\mathbf{v}}}_n^{\sim\varpi}] \\ &+ \left[-\Delta_z Q_{\mathcal{Z}_n^{\varpi}} - \hat{\zeta}_n^{\sim\varpi}\frac{\partial(\widehat{w\zeta})_n^{\sim\varpi}}{\partial z} + \lambda_c(\widehat{\zeta^2})_n^{\sim\varpi}\frac{\partial\hat{w}_n^{\sim\varpi}}{\partial z}\right] \\ &- \beta\hat{\zeta}_n^{\sim\varpi}\hat{v}_n^{\sim\varpi} + f\hat{\zeta}_n^{\sim\varpi}\frac{\partial\hat{w}_n^{\sim\varpi}}{\partial z} + \hat{\zeta}_n^{\sim\varpi}\left(\widehat{\zeta\frac{\partial w}{\partial z}}\right)_n^{\sim\varpi} \\ &+ \hat{\zeta}_n^{\sim\varpi}\underline{\mathbf{k}}\cdot\left(\widehat{\frac{\partial\underline{\mathbf{v}}}{\partial z}\wedge\nabla w}\right)_n^{\sim\varpi} + F_{\mathcal{Z}_n^{\varpi},z} + F_{\mathcal{Z}_n^{\varpi},h}. \end{aligned} \tag{87}$$

Or, symbolically,

$$\begin{aligned}\dot{\mathcal{Z}}_n^{\varpi} &= \Delta_h Q_{\mathcal{Z}_n^{\varpi}} + \Delta_z Q_{\mathcal{Z}_n^{\varpi}} + T_{\mathcal{Z}_n^{\varpi},\partial_h\zeta} + T_{\mathcal{Z}_n^{\varpi},\partial_z\zeta} + S_{\mathcal{Z}_n^{\varpi},\beta} + S_{\mathcal{Z}_n^{\varpi},f\nabla\cdot\underline{\mathbf{v}}} \\ &+ TS_{\mathcal{Z}_n^{\varpi},\zeta\nabla\cdot\underline{\mathbf{v}}} + TS_{\mathcal{Z}_n^{\varpi},\text{tilt}} + F_{\mathcal{Z}_n^{\varpi},z} + F_{\mathcal{Z}_n^{\varpi},h}. \end{aligned} \tag{88}$$

The meanings of these symbols are tabulated in Appendix D.

Each term of Eq. (88) has a corresponding physical interpretation. We have known that $\Delta_h Q_{\mathcal{Z}_n^{\varpi}}$ and $\Delta_z Q_{\mathcal{Z}_n^{\varpi}}$ are horizontal and vertical transports of $\mathcal{Z}_n^{\varpi}$, respectively, and $T_{\mathcal{Z}_n^{\varpi},\partial_h\zeta}$ and $T_{\mathcal{Z}_n^{\varpi},\partial_z\zeta}$ transfer rates for two perfect transfer processes. If ζ is horizontally and vertically a constant, then $T_{\mathcal{Z}_n^{\varpi},\partial_z\zeta}$ and $T_{\mathcal{Z}_n^{\varpi},\partial_h\zeta}$ sum up to zero. We have also explained $F_{\mathcal{Z}_n^{\varpi},z} + F_{\mathcal{Z}_n^{\varpi},h}$ represents the diffusion process. Among the rest terms, $S_{\mathcal{Z}_n^{\varpi},\beta}$ and $S_{\mathcal{Z}_n^{\varpi},f\nabla\cdot\underline{\mathbf{v}}}$ stand for two sources/sinks of $\mathcal{Z}$ due to β-effect and vortex stretching, and $TS_{\mathcal{Z}_n^{\varpi},\zeta\nabla\cdot\underline{\mathbf{v}}}$ and $TS_{\mathcal{Z}_n^{\varpi},\text{tilt}}$ transfer as well as generate/destroy enstrophy. Processes cannot be well separated for them. In a 2D system, both $TS_{\mathcal{Z}_n^{\varpi},\zeta\nabla\cdot\underline{\mathbf{v}}}$ and $TS_{\mathcal{Z}_n^{\varpi},\text{tilt}}$ vanish. As a result, the multiscale enstrophy equation is expected to be more useful for a plane flow than for a 3D flow.

12. Summary and discussion

A new methodology, *multiscale energy and vorticity analysis*, has been developed to investigate the inference of fundamental processes from real oceanic or atmospheric data for complex dynamics which are nonlinear, time and space intermittent, and involve multiscale interactions. Multiscale energy and enstrophy equations have been derived, interpreted, and compared to the energetics in classical formalism.

The MS-EVA is based on a localized orthogonal complementary subspace decomposition. It is formulated with the multiscale window transform, which is constructed to cope with the problem between localization and multiscale representation.[4] The concept of scale and scale window is introduced, and energy and enstrophy evolutions are then formulated for the large-scale, meso-scale, and sub-mesoscale windows. The formulation is principally in time and hence time scale window, but with a treatment in the horizontal dimension. We emphasize that, before physically interpreted, *all the final energetics should be multiplied by a*

[4] In the classical framework, multiscale energy does not have location identity of the dimension (time or space) to which the multiscale decomposition is performed.

constant factor 2^{j_2}*, and horizontally filtered with a 2D large-scale window synthesis*. When the large-scale window bound $j_0 = 0$, and a periodic extension scheme ($\varrho = 1$) is adopted, the multiscale energy Eqs. [(50) and (65)] in a two-window decomposition are reduced to the mean and eddy energy equations in a classical framework. In other words, our MS-EVA is a generalization of the classical energetics formalism to scale windows for generic purposes.

We have paid particular attention to the separation of transfers from the energetics resulting from nonlinearity. The separation is made possible by looking for a special type of process, the so-called perfect transfer. A perfect transfer process carries energy through scale windows, but does not generate nor destroy energy as a whole in the system.

Perfect transfer terms can be further decomposed to unravel the complicated window-window interactions. This is the so-called interaction analysis. Given a transfer function T, an interaction analysis results in many interaction terms, which can be cast into the following four groups:

$$T^{\varpi_1\to\varpi}, \qquad T^{\varpi_2\to\varpi}, \qquad T^{\varpi_1\oplus\varpi_2\to\varpi}, \qquad T^{\varpi\to\varpi},$$

each characteristic of an interaction process. Here superscripts $\varpi = 0, 1, 2$ stand for large-, meso-, and sub-meso-scale windows, respectively, and $\varpi_1 = (\varpi + 1) \bmod 3$, $\varpi_2 = (\varpi + 2) \bmod 3$. Explicit expressions for these functions are given in Eqs. (77)–(80).

By collecting the MS-EVA terms, energetic processes have been classified into four categories: transport, perfect transfer, buoyancy conversion, and dissipation/diffusion processes. Transport vanishes if integrated over a closed physical space; buoyancy conversion mediates between KE and APE on each individual window; while perfect transfer acts merely to redistribute energy between scale windows. The whole scenario is like a complex cycle, as shown in Fig. 7. These "conservative mechanisms" can essentially make energy reach anywhere in the super space formed with physical space, phase space, and space of energy type. It is not unreasonable to conjecture that, many patterns generated in geophysical fluid flows, complex as they might appear to be, could be a consequence of these energy redistributions.

Our MS-EVA therefore contains energetic information which is fundamental to GFD dynamics. It is expected to provide a useful platform for understanding the complexity of the fluids in which all life on Earth occurs. Direct applications may be set up for investigating the processes of turbulence, wave-current and wave-wave interaction, and the stability for infinite dimensional systems. In the sequels to this paper, we will show how this MS-EVA can be adapted to study a more concrete GFD problem. An avenue to application will be established for localized stability analysis (LR2), and two benchmark stability models will be utilized for validation. In another study (LR3), this methodology will be applied to a real problem to demonstrate how process inference is made easy with otherwise a very intricate dynamical system.

Acknowledgements

We would like to thank Prof. Donald G.M. Anderson, Dr. Kenneth Brink, and Dr. Arthur J. Miller for important and interesting scientific discussions. X. San Liang also thanks

Dr. Joseph Pedlosky for first raising the issue of transport–transfer separation, and thanks Prof. Brian Farrell, Prof. Yaneer Bar-Yam, Mr. Wayne Leslie, Dr. Patrick Haley, Dr. Pierre Lermusiaux, Dr. Carlos Lozano, Ms. Gioia Sweetland and Dr. James Wang for their generous help. This work was supported by the Office of Naval Research under Contracts N00014-95-1-0371, N00014-02-1-0989 and N00014-97-1-0239 to Harvard University.

Appendix A. Correction to the time derivative term

We have shown in Section 5 that there exists a correction term in the formulas with time derivatives. For a kinetic equation, this formula is

$$\underbrace{\hat{\delta}_n K_n}_{\text{(K)}} - \underbrace{(\Delta t)^2(\hat{\delta}^2_{n^2}\hat{\underline{\mathbf{v}}}_n \cdot \hat{\delta}_n\hat{\underline{\mathbf{v}}}_n)}_{\text{(C)}}, \tag{A.1}$$

where (C) is the correction term. Scale superscripts are omitted here since we do not want to limit the discussion to any particular scale window. Let's first do some nondimensional analysis so that a comparison is possible. Scale $\hat{\underline{\mathbf{v}}}_n$ with U, t with T, then

$$\text{Term (K)} \sim \frac{U^2}{T}, \qquad \text{Term (C)} \sim (\Delta t)^2 \frac{U}{T^2}\cdot\frac{U}{T} = (\Delta t)^2\frac{U^2}{T^3}.$$

This enables us to evaluate the weight of (C) relative to (K):

$$\frac{\text{Term (C)}}{\text{Term (K)}} \sim \frac{(\Delta t)^2U^2/T^3}{U^2/T} = \left(\frac{\Delta t}{T}\right)^2.$$

Apparently, this ratio will become significant only when $T \sim \Delta t$, i.e., when the time scale is of the time step size. In our MS-EVA formulation, the correction term (C) is hence not

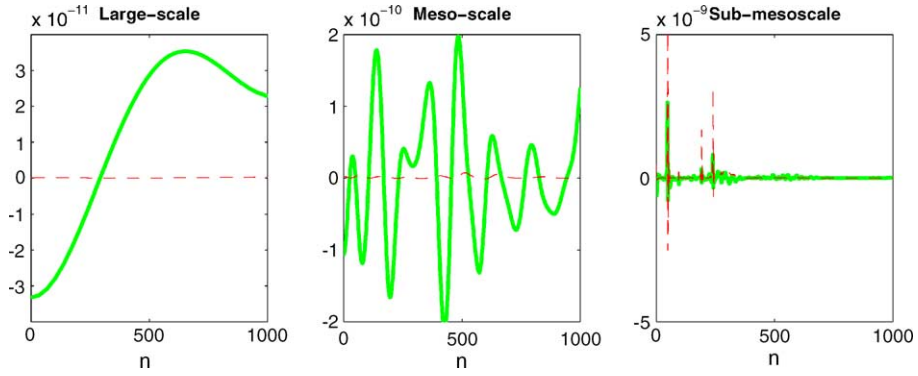

Fig. A.1. $\hat{\delta}_n K_n$ (thick solid) and its correction term (dashed) for the large-scale (left), meso-scale (middle), and sub-mesoscale (right) kinetic energy equations. Data source and parameter choice are the same as those of Fig. 4 (units in m^2/s^3; factor 2^{j_2} not multiplied).

significant for both large-scale and meso-scale equations. Fig. A.1 confirms this conclusion. The correction (dashed line) is so small in either the left or middle plots that it is totally negligible. Only in the sub-mesoscale window can its effect be seen, which, as argued before, might be parameterized into the dissipation/diffusion.

Appendix B. Interaction analysis for $T(0, n)$ and $T(2, n)$

Using the technique same as that for $T(1, n)$ in Section 9, we obtain a similar analysis for $T(0, n)$:

$$T(0,n) = \hat{\mathcal{R}}_n^{\sim 0} \cdot (\widehat{pq})_n^{\sim 0} = T_n^{1\to 0} + T_n^{2\to 0} + T_n^{1\oplus 2\to 0} + T_{n\to n}^{0\to 0} + T_{\text{other}\to n}^{0\to 0}, \tag{A.2}$$

where

$$T_n^{1\to 0} = \hat{\mathcal{R}}_n^{\sim 0} \cdot [(\widehat{p^{\sim 1}q^{\sim 1}})_n^{\sim 0} + (\widehat{p^{\sim 1}q^{\sim 0}})_n^{\sim 0} + (\widehat{p^{\sim 0}q^{\sim 1}})_n^{\sim 0}] \tag{A.3}$$

$$T_n^{2\to 0} = \hat{\mathcal{R}}_n^{\sim 0} \cdot [(\widehat{p^{\sim 0}q^{\sim 2}})_n^{\sim 0} + (\widehat{p^{\sim 2}q^{\sim 2}})_n^{\sim 0} + (\widehat{p^{\sim 2}q^{\sim 0}})_n^{\sim 0}] \tag{A.4}$$

$$T_n^{1\oplus 2\to 0} = \hat{\mathcal{R}}_n^{\sim 0} \cdot [(\widehat{p^{\sim 2}q^{\sim 1}})_n^{\sim 0} + (\widehat{p^{\sim 1}q^{\sim 2}})_n^{\sim 0}] \tag{A.5}$$

$$T_{n\to n}^{0\to 0} = \hat{\mathcal{R}}_n^{\sim 0} \cdot [\hat{p}_n^{\sim 0}\hat{q}_n^{\sim 0}(\widehat{\phi_n^{\varrho, j_2}})_n^{2^{\sim 0}}] \tag{A.6}$$

$$T_{\text{other}\to n}^{0\to 0} = \hat{\mathcal{R}}_n^{\sim 0} \cdot [(\widehat{p^{\sim 0}q^{\sim 0}}_{*^0})_n^{\sim 0} + \hat{q}_n^{\sim 0}(\widehat{p^{\sim 0}}_{*^0}\widehat{\phi_n^{\varrho, j_2}})_n^{\sim 0}], \tag{A.7}$$

and $T(2, n)$:

$$T(2,n) = \hat{\mathcal{R}}_n^{\sim 2} \cdot (\widehat{pq})_n^{\sim 2} = T_n^{0\to 2} + T_n^{1\to 2} + T_n^{0\oplus 1\to 2} + T_{n\to n}^{2\to 2} + T_{\text{other}\to n}^{2\to 2}, \tag{A.8}$$

where

$$T_n^{0\to 2} = \hat{\mathcal{R}}_n^{\sim 2} \cdot [(\widehat{p^{\sim 0}q^{\sim 0}})_n^{\sim 2} + (\widehat{p^{\sim 2}q^{\sim 0}})_n^{\sim 2} + (\widehat{p^{\sim 0}q^{\sim 2}})_n^{\sim 2}] \tag{A.9}$$

$$T_n^{1\to 2} = \hat{\mathcal{R}}_n^{\sim 2} \cdot [(\widehat{p^{\sim 1}q^{\sim 2}})_n^{\sim 2} + (\widehat{p^{\sim 1}q^{\sim 1}})_n^{\sim 2} + (\widehat{p^{\sim 2}q^{\sim 1}})_n^{\sim 2}] \tag{A.10}$$

$$T_n^{0\oplus 1\to 2} = \hat{\mathcal{R}}_n^{\sim 2} \cdot [(\widehat{p^{\sim 0}q^{\sim 1}})_n^{\sim 2} + (\widehat{p^{\sim 1}q^{\sim 0}})_n^{\sim 2}] \tag{A.11}$$

$$T_{n\to n}^{2\to 2} = \hat{\mathcal{R}}_n^{\sim 2} \cdot [\hat{p}_n^{\sim 2}\hat{q}_n^{\sim 2}(\widehat{\phi_n^{\varrho, j_2}})_n^{2^{\sim 2}}] \tag{A.12}$$

$$T_{\text{other}\to n}^{2\to 2} = \hat{\mathcal{R}}_n^{\sim 2} \cdot [(\widehat{p^{\sim 2}q^{\sim 2}}_{*^2})_n^{\sim 2} + \hat{q}_n^{\sim 2}(\widehat{p^{\sim 2}_{*^2}\phi_n^{\varrho, j_2}})_n^{\sim 2}]. \tag{A.13}$$

In these analyses, p_{*^0} and p_{*^2} are defined as

$$p_{*^0} = p - \hat{p}_n^{\sim 0}\phi_n^{\varrho, j_2}(t), \tag{A.14}$$

$$p_{*^2} = p - \hat{p}_n^{\sim 2}\phi_n^{\varrho, j_2}(t). \tag{A.15}$$

The physical meaning of the interaction terms is embedded in these mnemonic notations. In the superscripts, arrows signify the directions of energy transfer and the numbers 0–2 represent the large-scale, meso-scale, and sub-mesoscale windows, respectively.

Appendix C. Connection between the MS-EVA KE equations and the mean and eddy KE equations in a classical Reynolds formalism

To connect our MS-EVA to the classical energetic formulation, rewrite Eq. (50) (dissipation omitted) as

$$\hat{\underline{\mathbf{v}}}_n^{\sim\varpi}\cdot\left(\widehat{\frac{\partial\underline{\mathbf{v}}}{\partial t}}\right)_n^{\sim\varpi}=\underbrace{\hat{\underline{\mathbf{v}}}_n^{\sim\varpi}\cdot\left[-\nabla\cdot(\widehat{\underline{\mathbf{v}}\,\underline{\mathbf{v}}})_n^{\sim\varpi}-\frac{\partial}{\partial z}(\widehat{w\underline{\mathbf{v}}})_n^{\sim\varpi}\right]}_{\textbf{(I)}}+\underbrace{\Delta_h Q_{P_n^\varpi}+\Delta_z Q_{P_n^\varpi}-b_n^\varpi}_{\textbf{(II)}},\tag{A.16}$$

We want to see what this equation reduces to if $j_1=j_2$ (that is to say, only *two-scale windows* are considered), $j_0=0$, and a *periodic extension* is employed.

First consider the large scale window $\varpi=0$. Let q be any field variable (u, v, w, or P). A two-scale window decomposition means

$$q=q^{\sim 0}+q^{\sim 1}.\tag{A.17}$$

With the choice of zero j_0 and periodic extension, we know from the MWT properties (see Section 2.3) that $q^{\sim 0}$ is constant in time and is equal to $\bar{q}$ or $2^{j_2/2}\hat{q}_n^{\sim 0}$ in magnitude, that is,

$$q^{\sim 0}=\bar{q}=2^{j_2/2}\,\hat{q}_n^{\sim 0},\qquad q^{\sim 1}=q-\bar{q}=q'.\tag{A.18}$$

Hence

$$(\hat{\bar{q}})_n^{\sim 0}=(\widehat{q^{\sim 0}})_n^{\sim 0}=q^{\sim 0}=2^{-j_2/2}\bar{q},\tag{A.19}$$

$$(\widehat{q'})_n^{\sim 0}=(\widehat{q^{\sim 1}})_n^{\sim 0}=0.\tag{A.20}$$

Substituting $\underline{\mathbf{v}}$ and w for the q in (A.17), the velocity field is decomposed as $\underline{\mathbf{v}}=\bar{\underline{\mathbf{v}}}+\underline{\mathbf{v}}'$, and $w=\bar{w}+w'$. Let $K_{\text{mean}}=\frac{1}{2}\bar{\underline{\mathbf{v}}}\cdot\bar{\underline{\mathbf{v}}}$. The equivalence between the large-scale transform and duration average allows an expression of the large-scale kinetic energy K_n^0 in terms of K_{mean}. In fact,

$$K_n^0=2^{j_2}\left(\frac{1}{2}\hat{\underline{\mathbf{v}}}_n^{\sim 0}\cdot\hat{\underline{\mathbf{v}}}_n^{\sim 0}\right)=\frac{1}{2}\bar{\underline{\mathbf{v}}}\cdot\bar{\underline{\mathbf{v}}}=K_{\text{mean}}.\tag{A.21}$$

Note here we have taken into account the multiplier 2^{j_2}. These facts are now used to simplify the term (**I**) of Eq. (A.16). With the two-scale decomposition, the dyad $(\underline{\mathbf{v}}\,\underline{\mathbf{v}})$ after transform is expanded as

$$(\widehat{\underline{\mathbf{v}}\,\underline{\mathbf{v}}})_n^{\sim 0}=(\widehat{\bar{\underline{\mathbf{v}}}\,\bar{\underline{\mathbf{v}}}})_n^{\sim 0}+(\widehat{\bar{\underline{\mathbf{v}}}\,\underline{\mathbf{v}}'})_n^{\sim 0}+(\widehat{\underline{\mathbf{v}}'\,\bar{\underline{\mathbf{v}}}})_n^{\sim 0}+(\widehat{\underline{\mathbf{v}}'\,\underline{\mathbf{v}}'})_n^{\sim 0}\tag{A.22}$$

$$(\widehat{\underline{\mathbf{v}}\,\underline{\mathbf{v}}})_n^{\sim 0}=\bar{\underline{\mathbf{v}}}\hat{\underline{\mathbf{v}}}_n^{\sim 0}+(\widehat{\underline{\mathbf{v}}'\,\underline{\mathbf{v}}'})_n^{\sim 0}.\tag{A.23}$$

Likewise,

$$(\widehat{w\underline{\mathbf{v}}})_n^{\sim 0} = \overline{w}\hat{\underline{\mathbf{v}}}_n^{\sim 0} + (\widehat{w'\underline{\mathbf{v}}'})_n^{\sim 0}. \tag{A.24}$$

These allow term (**I**) to be written as

$$\begin{aligned}(\mathbf{I}) &= \hat{\underline{\mathbf{v}}}_n^{\sim 0} \cdot \left[-\nabla \cdot (\overline{\underline{\mathbf{v}}}\,\hat{\underline{\mathbf{v}}}_n^{\sim 0}) - \frac{\partial}{\partial z}(\overline{w}\hat{\underline{\mathbf{v}}}_n^{\sim 0}) \right] + \hat{\underline{\mathbf{v}}}_n^{\sim 0} \cdot \left[-\nabla \cdot (\widehat{\underline{\mathbf{v}}'\underline{\mathbf{v}}'})_n^{\sim 0} - \frac{\partial}{\partial z}(\widehat{w'\underline{\mathbf{v}}'})_n^{\sim 0} \right] \\ &= 2^{-j_2} \left\{ -\nabla \cdot (\overline{\underline{\mathbf{v}}} K_{\text{mean}}) - \frac{\partial}{\partial z}(\overline{w} K_{\text{mean}}) + \overline{\underline{\mathbf{v}}} \cdot \left[-\nabla \cdot \overline{(\underline{\mathbf{v}}'\underline{\mathbf{v}}')} - \frac{\partial}{\partial z}\overline{(w'\underline{\mathbf{v}}')} \right] \right\} \\ &= 2^{-j_2} \left\{ -\nabla \cdot (\overline{\underline{\mathbf{v}}} K_{\text{mean}}) - \frac{\partial}{\partial z}(\overline{w} K_{\text{mean}}) + \overline{\underline{\mathbf{v}}} \cdot \nabla_3 \cdot \underline{\underline{\mathbf{T}}} \right\}, \end{aligned} \tag{A.25}$$

where

$$\nabla_3 = \underline{\mathbf{i}}\frac{\partial}{\partial x} + \underline{\mathbf{j}}\frac{\partial}{\partial y} + \underline{\mathbf{k}}\frac{\partial}{\partial z},$$

and

$$\underline{\underline{\mathbf{T}}} = \begin{bmatrix} -\overline{(u'u')} & -\overline{(u'v')} & -\overline{(u'w')} \\ -\overline{(v'u')} & -\overline{(v'v')} & -\overline{(v'w')} \\ -\overline{(w'u')} & -\overline{(w'v')} & -\overline{(w'w')} \end{bmatrix}. \tag{A.26}$$

For term (**II**), it is equal to, in the present setting,

$$(\mathbf{II}) = 2^{-j_0} \left\{ -\frac{1}{\rho_0}\nabla \cdot (\overline{P}\overline{\underline{\mathbf{v}}}) - \frac{1}{\rho_0}\frac{\partial}{\partial z}(\overline{P}\overline{w}) - \frac{g}{\rho_0}\overline{w}\overline{\rho}. \right\} \tag{A.27}$$

Substitute (**I**) and (**II**) back to Eq. (A.16). Considering that the left hand side is now $2^{-j_0}\overline{\underline{\mathbf{v}}} \cdot \overline{\left(\frac{\partial \underline{\mathbf{v}}}{\partial t}\right)}$, we have, with the common factor 2^{-j_0} cancelled out,

$$\begin{aligned}\overline{\underline{\mathbf{v}}} \cdot \overline{\left(\frac{\partial \underline{\mathbf{v}}}{\partial t}\right)} &= -\nabla \cdot (\overline{\underline{\mathbf{v}}} K^L) - \frac{\partial}{\partial z}(\overline{w} K^L) - \frac{1}{\rho_0}\nabla \cdot (\overline{P}\overline{\underline{\mathbf{v}}}) - \frac{1}{\rho_0}\frac{\partial}{\partial z}(\overline{P}\overline{w}) \\ &\quad - \frac{g}{\rho_0}\overline{w}\overline{\rho}. + \overline{\underline{\mathbf{v}}} \cdot \nabla_3 \cdot \underline{\underline{\mathbf{T}}}. \end{aligned} \tag{A.28}$$

This is exactly what Harrison and Robinson (1978) have obtained for the mean kinetic energy, with $\underline{\underline{\mathbf{T}}}$ the Reynolds stress tensor in their formulation.

Above is about the large-scale energetics. For the meso-scale window ($\varpi = 1$), things are more complicated. In order to make Eq. (A.16) comparable to the classical eddy KE equation, just $j_0 = 0$ and periodic extension are not enough, as now there no longer exists for variable p a linear relation between $\hat{p}_n^{\sim 1}$ and p'. We have to marginalize (A.16) to the physical space to fulfill this mission. In this particular case, the marginalization equality

(11) in Section 2.3 is simply

$$\mathcal{M}_n \hat{p}_n^{\sim 1} \hat{q}_n^{\sim 1} = \overline{p'q'}, \tag{A.29}$$

since here the deviation operation (prime) and the meso-scale synthesis operator are identical. Marginalization of (A.16) with $\varpi = 1$ yields

$$\overline{\underline{\mathbf{v}}' \cdot \left(\frac{\partial \underline{\mathbf{v}}}{\partial t}\right)'} = -\underbrace{\overline{\underline{\mathbf{v}}' \cdot \nabla \cdot (\underline{\mathbf{v}}\underline{\mathbf{v}})'}}_{(\mathbf{I}')} - \underbrace{\overline{\underline{\mathbf{v}}' \cdot \frac{\partial}{\partial z}(w\underline{\mathbf{v}})'}}_{(\mathbf{II}')} - \underbrace{\overline{\underline{\mathbf{v}}' \cdot \nabla \left(\frac{P'}{\rho_0}\right)}}_{(\mathbf{III}')}. \tag{A.30}$$

It is easy to show, as we did before,

$$(\mathbf{III}') = \nabla \cdot \overline{\left(\underline{\mathbf{v}}' \frac{P'}{\rho_0}\right)} + \frac{\partial}{\partial z}\overline{\left(w' \frac{P'}{\rho_0}\right)} + \frac{g}{\rho_0}\overline{w'\rho'}. \tag{A.31}$$

The other two terms sum up to

$$(\mathbf{I}') + (\mathbf{II}') = \nabla \cdot \overline{\left(\underline{\mathbf{v}} \frac{\underline{\mathbf{v}}' \cdot \underline{\mathbf{v}}'}{2}\right)} + \frac{\partial}{\partial z}\overline{\left(w \frac{\underline{\mathbf{v}}' \cdot \underline{\mathbf{v}}'}{2}\right)} + \overline{\underline{\mathbf{v}}'\underline{\mathbf{v}}'} : \nabla \underline{\overline{\mathbf{v}}} + \overline{\underline{\mathbf{v}}' w'} \cdot \frac{\partial \underline{\overline{\mathbf{v}}}}{\partial z}. \tag{A.32}$$

Therefore,

$$\begin{aligned}\overline{\underline{\mathbf{v}}' \cdot \left(\frac{\partial \underline{\mathbf{v}}}{\partial t}\right)'} = &-\nabla \cdot \overline{\left(\underline{\mathbf{v}} \frac{\underline{\mathbf{v}}' \cdot \underline{\mathbf{v}}'}{2}\right)} - \frac{\partial}{\partial z}\overline{\left(w \frac{\underline{\mathbf{v}}' \cdot \underline{\mathbf{v}}'}{2}\right)} - \nabla \cdot \overline{\left(\underline{\mathbf{v}}' \frac{P'}{\rho_0}\right)} \\ &- \frac{\partial}{\partial z}\overline{\left(w' \frac{P'}{\rho_0}\right)} - \frac{g}{\rho_0}\overline{w'\rho'} - \overline{\underline{\mathbf{v}}'\underline{\mathbf{v}}'} : \nabla \underline{\overline{\mathbf{v}}} - \overline{\underline{\mathbf{v}}' w'} \cdot \frac{\partial \underline{\overline{\mathbf{v}}}}{\partial z}.\end{aligned} \tag{A.33}$$

Again, this is exactly the eddy KE equation obtained by Harrison and Robinson (1978).

Appendix D. Glossary

Tables A.1–A.3.

Table A.1
General symbols

A_n^{ϖ}	Available potential energy on window ϖ at time $2^{-j_2}n$
j_0, j_1, j_2	Upper bounds of scale level for the three scale windows
K_n^{ϖ}	Kinetic energy on window ϖ at time $2^{-j_2}n$
V_{ϱ, j_2}	Direct sum of the three scale windows.
ϖ	Window index ($\varpi = 0, 1, 2$ for large-scale, meso-scale, and sub-mesoscale windows, respectively)
$\mathcal{Z}_n^{\varpi}$	Enstrophy on window ϖ at time $2^{-j_2}n$
$\hat{z}_n^{\sim \varpi}$	Multiscale window transform of variable z
$z^{\sim \varpi}$	Multiscale window synthesis of variable z
$\overline{z}$	Duration average of variable z
$\phi_n^{\varrho, j}$	Periodized scaling basis function at level j
$\psi_n^{\varrho, j}$	Periodized wavelet basis function at level j

Table A.2
Symbols for the multiscale energy equations (time $2^{-j_2}n$, window ϖ)

Kinetic energy (KE)		Available potential energy (APE)	
$\dot{K}_n^{\varpi}$	Time rate of change of KE	$\dot{A}_n^{\varpi}$	Time rate of change of APE
$\Delta_h Q_{K_n^{\varpi}}$	Horizontal KE advective working rate	$\Delta_h Q_{A_n^{\varpi}}$	Horizontal APE advective working rate
$\Delta_z Q_{K_n^{\varpi}}$	Vertical KE advective working rate	$\Delta_z Q_{A_n^{\varpi}}$	Vertical APE advective working rate
$T_{K_n^{\varpi},h}$	Rate of KE transfer due to the horizontal flow	$T_{A_n^{\varpi},\partial_h\rho}$	Rate of APE transfer due to the horizontal gradient density
$T_{K_n^{\varpi},z}$	Rate of KE transfer due to the vertical flow	$T_{A_n^{\varpi},\partial_z\rho}$	Rate of APE transfer due to the vertical gradient density
$-b_n^{\varpi}$	Rate of buoyancy conversion	b_n^{ϖ}	Rate of inverse buoyancy conversion
$\Delta_h Q_{P_n^{\varpi}}$	Horizontal pressure working rate	$TS_{A_n^{\varpi}}$	Rate of an imperfect APE transfer due to the stationary shear of the density profile
$\Delta_z Q_{P_n^{\varpi}}$	Vertical pressure working rate	$F_{A_n^{\varpi},h}$	Horizontal diffusion
$F_{K_n^{\varpi},z}$	Vertical dissipation	$F_{A_n^{\varpi},z}$	Vertical diffusion
$F_{K_n^{\varpi},h}$	Horizontal dissipation		

Table A.3
Symbols for the multiscale enstrophy equation (time $2^{-j_2}n$, window ϖ)

$\dot{\mathcal{Z}}_n^{\varpi}$	Time rate of change of $\mathcal{Z}$ on window ϖ at time $2^{-j_2}n$
$\Delta_h Q_{\mathcal{Z}_n^{\varpi}}$	Horizontal transport rate
$\Delta_z Q_{\mathcal{Z}_n^{\varpi}}$	Vertical transport rate
$T_{\mathcal{Z}_n^{\varpi},\partial_h\zeta}$	Rate of enstrophy transfer due to the horizontal variation of ζ
$T_{\mathcal{Z}_n^{\varpi},\partial_z\zeta}$	Rate of enstrophy transfer due to the vertical variation of ζ
$S_{\mathcal{Z}_n^{\varpi},\beta}$	β-Effect-caused source/sink
$S_{\mathcal{Z}_n^{\varpi},f\nabla\cdot\underline{\mathbf{v}}}$	Source/sink of enstrophy due to horizontal divergence
$TS_{\mathcal{Z}_n^{\varpi},\zeta\nabla\cdot\underline{\mathbf{v}}}$	Rate of $\mathcal{Z}$ transfer and generation due to rotation-divergence correlation
$TS_{\mathcal{Z}_n^{\varpi},\text{tilt}}$	Rate of $\mathcal{Z}$ transfer and generation due to the vortex tube tilting
$F_{\mathcal{Z}_n^{\varpi},h}$	Horizontal diffusion rate
$F_{\mathcal{Z}_n^{\varpi},z}$	Vertical diffusion rate

References

Cronin, M., Watts, R., 1996. Eddy-mean flow interaction in the Gulf Stream at 68°W. Part I: Eddy energetics. J. Phys. Oceanogr. 26, 2107–2131.

Farge, M., 1992. Wavelet transforms and their applications to turbulence. Annu. Rev. Fluid Mech. 24, 395–457.

Fournier, A., 1999. Atmospheric energetics in the wavelet domain I: Governing equations and interpretation for idealized flows. Ph.D. Thesis, University of Maryland.

Harrison, D.E., Robinson, A.R., 1978. Energy analysis of open regions of turbulent flows-mean eddy energetics of a numerical ocean circulation experiment. Dyn. Atmos. Oceans 2, 185–211.

Hernández, E., Weiss, G., 1996. A First Course on Wavelets. CRC Press, 489 pp.

Holland, W.R., Lin, L.B., 1975. On the generation of mesoscale eddies and their contribution to the ocean general circulation. I. A preliminary numerical experiment. J. Phys. Oceangr. 5, 642–657.

Huang, N.E., Shen, Z., Long, S.R., 1999. A new view of nonlinear water waves: The Hilbert spectrum. Annu. Rev. Fluid Mech. 31, 417–457.

Iima, M., Toh, S., 1995. Wavelet analysis of the energy transfer caused by convective terms: Application to the Burgers shock. Phys. Rev. E 52 (6), 6189–6201.

Liang, X.S., 2002. Wavelet-based multiscale window transform and energy and vorticity analysis. Harvard Reports in Physical/Interdisciplinary Ocean Science, Rep. No. 66, Harvard University, Cambridge, MA, 411 pp.

Liang, X.S., Anderson, D.G.M., 2005. Multiscale window analysis, in preparation.

Liang, X.S., Robinson, A.R., 2005. Localized multiscale energy and vorticity analysis. II. Instability theory and validation, Dyn. Atmos. Oceans (this issue).

Liang, X.S., Robinson, A.R., 2004. A study of the Iceland-Faeroe frontal variability using the multiscale energy and vorticity analysis, J. Phys. Oceanogr. 34, 2571–2591.

Liang, X.S., Anderson, D.G.M., Wang, M., 2005. A unified localized hydrodynamic instability analysis for incompressible fluid flows, J. Fluid Mech., submitted.

Pinardi, N., Robinson, A.R., 1986. Quasigeostrophic energetics of open ocean regions. Dyn. Atmos. Oceans 10, 185–219.

Plumb, R.A., 1983. A new look at the energy cycle. J. Atmos. Sci. 40, 1669–1688.

Robinson, A.R., Arango, H.G., Miller, A.J., Warn-Varnas, A., Poulain, P.-M., Leslie, W.G., 1996a. Real-time operational forecasting on shipboard of the Iceland-Faeroe Frontal variability. Bull. Am. Meteorol. Soc., 243–259.

Robinson, A.R., Arango, H.G., Warn-Varnas, A., Leslie, W., Miller, A.J., Haley, P., Lozano, C., 1996b. Real-time regional forecasting. In: Malanotte-Rizzoli, P. (Ed.), Modern Approaches to Data Assimilation in Ocean Modeling. Elsevier Oceanogr. Ser., vol. 61., pp. 377–410.

Spall, M., 1989. Regional primitive equation modeling and analysis of the POLYMODE data set. Dyn. Atmos. Oceans 14, 125–174.

Tennekes, H., Lumley, J.L., 1972. A First Course in Turbulence. MIT Press, 300 pp.