MULTISCALE MODEL. SIMUL.
Vol. 6, No. 2, pp. 437–467

MULTISCALE WINDOW TRANSFORM*

X. SAN LIANG† AND DONALD G. M. ANDERSON‡

Abstract. A new analysis apparatus, the multiscale window transform (MWT), is developed to generalize the classical mean-eddy decomposition (MED) in fluid mechanics to include three or more ranges of scales, and to ensure a faithful representation of localized energy processes. The development begins with the introduction of a sequence of finite-dimensional subspaces of $L_2[0,1]$, $\{V_{\varrho,j}\}_{0\le j\le j_2}$, based on a multiresolution analysis of $L_2(\mathbb{R})$. All the $V_{\varrho,j}$ are sampling spaces, i.e., spaces spanned by some translation invariant basis. The upper bound of the index of scale level, j_2, is prescribed in accordance with the signal of concern. Within V_{ϱ,j_2}, the concepts of large-scale, mesoscale, and submesoscale windows are introduced, each being a subspace of V_{ϱ,j_2} and containing exclusively a range of scales. A transform-reconstruction pair is constructed on each window, representing the features for the corresponding range of scales. The resulting MWT has several useful properties. Of particular importance is a property of marginalization, which allows for a simple representation of energy on specified windows and localizations. The MWT has been connected to other existing analysis techniques: The classical MED turns out to be a particular case, with averaging acting as both a transform and a reconstruction in the MWT framework; it has also been compared with wavelet analysis and the Hilbert–Huang transform. As a realization, we show how an MWT is built with an orthonormalized cubic spline system. Computational techniques leading to a fast transform scheme are also presented. The MWT is validated with two examples, and its applications in the study of hydrodynamic stability analysis (fully nonlinear) are discussed and exemplified.

Key words. multiscale window transform, mean-eddy-turbulence interaction, multiresolution analysis, turbulent spot, hydrodynamic stability, perfect energy transfer

AMS subject classifications. 46S99, 42C05, 42C40, 76E99, 76F99, 76M99

DOI. 10.1137/06066895X

1. Introduction.

Mean-eddy decomposition (MED) is a technique used extensively in the studies of multiscale interactive processes in fluid flows such as the mean-eddy interaction as schematized in Figure 1. It is also called Reynolds decomposition in the context of turbulence research, which forms the basis for the renowned Reynolds equations. This technique, however, encounters difficulties when the processes under consideration are inhomogeneous and/or nonstationary, as will be clear in the following. The purpose of this study is to generalize the MED to a new apparatus, the multiscale window transform (MWT), to overcome these difficulties. In a follow-up paper (see section 10), we will see that the complicated mean-eddy-turbulence interactions in fluid flows can be quantitatively expressed in a simple formula in terms of the MWT.

The basic idea of the MED can be easily elucidated with a one-dimensional (1D) signal $x(t)$, $t\in[0,1]$: One takes the mean $\bar{x}=\int_0^1 x(t)\,dt$ and then subtracts $\bar{x}$ from x to obtain the fluctuation $x'=x-\bar{x}$, or eddy field. The corresponding energies/variances are $\bar{x}^2$ and $\overline{x'^2}$, respectively. Notice the overbar or domain average performed on x'^2.

*Received by the editors September 3, 2006; accepted for publication (in revised form) March 12, 2007; published electronically June 1, 2007. This work was supported by the Office of Naval Research under grant N00014-02-1-0989 to Harvard University.

http://www.siam.org/journals/mms/6-2/66895.html

†School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138. Current address: Courant Institute of Mathematical Sciences, 251 Mercer St., New York, NY 10012 (sanliang@courant.nyu.edu).

‡School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138.

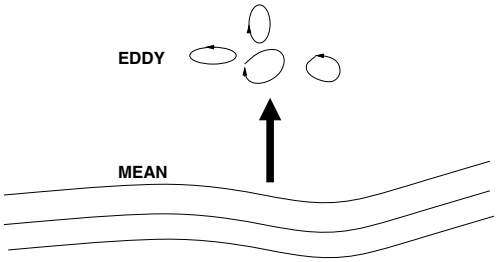

FIG. 1. *A schematic of the multiscale interaction in fluid flows. Central to the problem is the transfer of energy between the designated events.*

This invokes a severe problem for the MED. It eliminates the structure of the eddy field without differentiating one location from another location on $[0, 1]$. While this can be used to characterize eddy processes with homogeneity or stationarity in t, real fluid processes are by nature neither homogeneous nor stationary. The emergence of turbulent spots as illustrated in Figure 2(a) is a good example. A turbulent spot is a highly intermittent event, defined on a localized domain bearing irregularity and variability. Figure 2(b) is a 1D abstraction of the problem. One sees that all the events are limited within a very narrow domain. If an averaging is performed, one obtains an eddy energy evenly distributed everywhere on $[0, 1]$. This is not physically true, as we see clearly that only a burst of eddy energy exists within that limited region, outside which nothing occurs. Another problem for the MED is that the background may be variable (nonstationary background, as it is usually called), while an MED mean is a constant. As shown in Figure 2(c), if the basic profile is a sloped line,[1] even though no perturbation exists, the MED will give an eddy energy positive everywhere except at one point in the middle. Obviously, the classical MED is inadequate for fluid dynamics studies under generic circumstances.

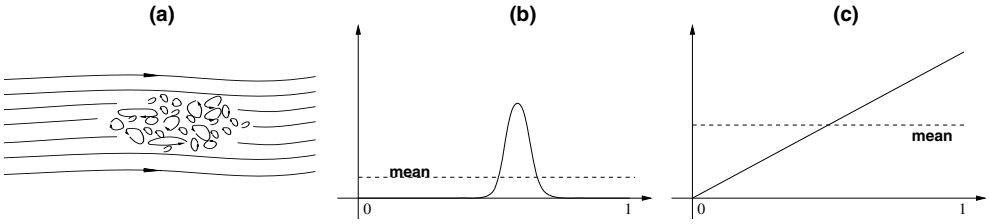

FIG. 2. (a) *schematic of the emergence of a turbulent spot (see, for example, Gad-el-Hak, Blackwelder, and Riley* (1981)*;* (b) *an impulse within a calm fluid;* (c) *a sloping straight line. The abscissa of* (b) *and* (c) *could be time or any dimension of space, while the ordinate is some state variable.*

Different approaches have been sought to remedy this inadequacy. For example, people often use piecewise or subdomain averaging instead of global averaging to gain local information. This in general does not help, as is easily seen with the aid of Figure 2(c): There is always eddy energy through $[0, 1]$, unless the subdomains are infinitely small, which then invokes the issue of inadequate alteration to scale. Besides, localized events tend to appear on varying and irregular geometry, making

[1]For example, if the ordinate represents pressure and abscissa longitude, then the distribution of Figure 2(c) implies a uniform meridional geostrophic flow (see Pedlosky (1979)).

it practically impossible to trace out the right subdomain.

A more sophisticated approach is filtering (e.g., Gaussian filters), a widely employed technique these days in turbulence research. Actually, what we will be building is also related to filters. But, not just any filter will work. Take the schematized decomposition in Figure 1 as an example. Using some filtering (or other) technique, it is always easy to obtain the reconstructions on different scale ranges for the representations of the mean, eddy, and turbulent processes. However, the central issue here is not the reconstructions themselves but the interactions, which are characterized by the energy transferred between these reconstructions. A key question for fluid physics studies is thus, What is the energy of a reconstruction? Or, given a filter, how is energy represented?

Historically, this question has been mostly overlooked, probably because this is not a question in the MED. For a decomposition $x = \bar{x} + x'$, we know that the eddy energy is $\overline{x'^2}$ (constant factors omitted). Look at the overbar (averaging) of $\overline{x'^2}$. That is how eddy energy is represented in the MED. (In section 4, we will see that the MED is very special in that its transform and reconstruction are the same, making this simple representation possible.) If the mean is replaced by a low pass filtered signal, the residue is usually taken as the eddy field. In this case, what is the eddy energy? How should the overbar or averaging be changed to have energy locally represented? This seemingly trivial question is not easy at all. In fact, one had not seen hope for its solution until wavelets and filter banks were recently connected (see Strang and Nguyen (1997)). (Practically, it is not uncommon in engineering literature that the simple square of a reconstruction is taken as the local energy; of course, this is conceptually wrong.) This difficulty may help to explain why the subdomain/subinterval averaging, though its deficiency as illustrated above is well known, is still in use, particularly in oceanographic and atmospheric research, while all kinds of sophisticated filters are readily available.

We proceed to build a physically consistent decomposition technique. Specifically, we need to build a decomposition capable of representing events, characterized by their corresponding energies, localized in space and/or time, out of a variable background, as schematized in Figure 2(b) and (c). In many cases, extension of the two-component MED to include three or more components is needed. An example is the mean-eddy-turbulence decomposition as one may use in studying the energetics of the inertial range of a turbulent flow. Putting all these desiderata together, in a more formal language, we need to generalize the MED so that

(1) the decomposition partitions a Hilbert space into orthogonal subspaces (as many as desired), each containing functions exclusively with a certain range of scales;
(2) each subspace is a sampling space, i.e., a closed subspace in which all the functions can be expanded with respect to a translation invariant orthonormal basis;
(3) energy for each subspace is represented locally.

In item (2), the sampling space requirement ensures the localization as needed. Note the orthogonality between the subspaces and the orthonormality of the translationally invariant bases. As we mentioned above, they are critical. Only with an orthonormal basis can the notion of energy be introduced, and only with an orthogonal decomposition can the total energy be conserved. We insist on this because we are to use the new technique to explore physical laws, not just data analysis. As the resulting subspaces contain exclusively different scale ranges, we call them *scale windows* (formal definition deferred to section 3). Accordingly, the technique we will be developing is

termed the *multiscale window transform* (MWT).

In order to complete the MWT development, classical analysis tools such as the Fourier transform do not work, since they are global in character. Mathematical machinery which might come to aid include the modern Hilbert–Huang transform or HHT (Huang, Shen, and Long (1999)) and wavelet analysis (see, e.g., Strang and Nguyen (1997)). The HHT is a powerful technique whose utility has recently been recognized in the fields of nonlinear wave analysis and fault identification, among many others. But here it does not fit our purpose. The difficulty comes from the translation invariant kernel, $\frac{1}{t-t'}$, which does not form an orthonormal system (see section 6 for more discussion). For orthonormal wavelet transforms, the transform coefficients are defined discretely at different locations in time for different scales, as shown in Figure 3. The associated energies, which are proportional to the square of the transform coefficients, are thus also defined at different locations for different scales. This implies that there is no way to sum across the scales to build scale windows. One cannot apply mathematical manipulations such as interpolation, as the energies are defined discretely. Wavelet analysis as such therefore does not help either. The problem needs to be tackled via a different approach.

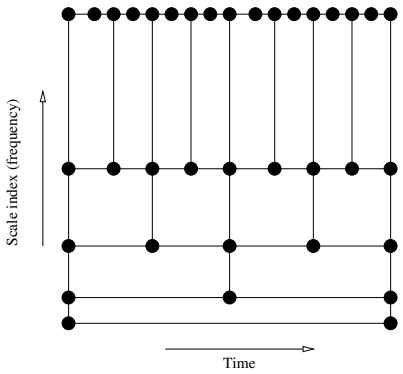

FIG. 3. *Schematic of the time-frequency representation of an orthonormal wavelet transform. The transform coefficients, and hence energies, are defined discretely at different locations for different scale levels.*

The appropriate mathematical framework is introduced in section 2, within which the MWT is defined in section 3. Properties are explored in section 4, and connections to wavelet analysis and the HHT, as well as the MED, are established in section 6. We present a realization in sections 7 and 8 and test it with two examples (section 9); we also briefly present an application to exemplify its utility (section 10). This study is concluded in section 11.

2. Mathematical framework. We consider square integrable functions over some finite domain $\mathcal{D}$. Without loss of generality, let $\mathcal{D}$ be $[0,1]$; if not, a transformation can always make it so. The study is thus conducted in some subspace of $L_2[0,1]$ (or $L_2[0,\varrho]$ for some ϱ restricted on $[0,1]$), which must be a sampling space in order to retain localization. This section introduces the subspace.

2.1. Multiresolution analysis of $L_2(\mathbb{R})$. We rely on the multiresolution analysis (MRA) of $L_2(\mathbb{R})$ by Meyer (1992) to fulfill our objective. By an MRA of $L_2(\mathbb{R})$ we mean a sequence of closed subspaces $\{V_j\}_{j\in\mathbb{Z}}$ such that the following hold (see,

e.g., Hernández and Weiss (1996) and Meyer (1992)):[2]

(1) $\ldots V_{-1} \subset V_0 \subset V_1 \subset V_2 \ldots$ (nestedness).
(2) $\mathrm{Cl} \bigcup_{j\in\mathbb{Z}} V_j = L_2(\mathbb{R})$ (totality).
(3) $\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$ (emptiness).
(4) There exists a translation invariant (affine) orthonormal set, i.e., an orthonormal set formed by shifting the independent variable by integers, $\{\phi(t-n-\frac{1}{2}), n \in \mathbb{Z}\}$, that is total in V_0.
(5) $x(t) \in V_j$ iff $x(2t) \in V_{j+1}$ (refinability).

In item (4), we put a translation $-\frac{1}{2}$ in the argument of ϕ to have the sampling points sit in between adjacent integer locations. We will see the convenience this offers later when boundary extensions are introduced. The closed subspace V_j is called a multiresolution approximation of $L_2(\mathbb{R})$ at scale level j.

There is a localization requirement for the ϕ in item (4) to make $\{\phi(t-n-\frac{1}{2}), n \in \mathbb{Z}\}$ a translation invariant basis. A function $\phi(t)$ is said to have a polynomial localization about the origin (or simply localization) of order γ if

$$|\phi(t)| \le \frac{C}{(1+|t|^2)^{\gamma/2}} \equiv \kappa_\gamma^c(t) \qquad \forall\, t \in \mathbb{R} \tag{2.1}$$

for $\gamma > 0$, where C is some positive constant. There are also other types of localization (cf. Holschneider (1995)), but here we consider polynomial localization only. The parameter γ describes how fast the function decays with increasing $|t|$. The larger the γ, the more localized the function $\phi(t)$. We consider $\gamma > 1$ only. In this case, it is easy to check that, for any $\varepsilon > 0$, there exists an $L_\varepsilon = \left(\frac{2C}{\gamma-1}\right)^{\frac{1}{\gamma-1}} \cdot \varepsilon^{-\frac{1}{\gamma-1}}$, such that $\int_{|t|>L_\varepsilon} |\phi(t)|\, dt \le \varepsilon$. We say $f(t)$ is effectively supported on $[-L_\varepsilon, L_\varepsilon]$ up to ε.

The existence of such an MRA has been established (see, e.g., Hernández and Weiss (1996)), and the nested V_j are all sampling spaces. These subspaces, however, are defined over the entire real line $\mathbb{R}$, while our problems are defined on $[0,1]$. To avail ourselves of the MRA convenience, we need to extend the functions of interest beyond $[0,1]$ to $\mathbb{R}$.

2.2. Boundary extension. There are different ways to carry out the extension. Commonly used schemes include zero-padding (outside $[0,1]$), periodic extension or extension by periodization, and symmetric extension or extension by reflection. The first is the easiest but usually not adequate, and so we consider the other two only. For $x_1(t)$, $t \in [0,1]$, let $x_2(t)$ be an extension of $x_1(t)$. If x_2 is obtained through periodization, then

$$x_2(t+\ell) = x_1(t) \qquad \forall\, t \in [0,1],\ \ell \in \mathbb{Z}; \tag{2.2}$$

if x_2 is obtained by extending x_1 such that $x_1(t)$ is symmetric about the boundary points, then

$$x_2(t+\ell) = \begin{cases} x_1(t), & \ell \text{ even},\ \ell \in \mathbb{Z},\ t \in [0,1], \\ x_1(1-t), & \ell \text{ odd},\ \ell \in \mathbb{Z},\ t \in [0,1]. \end{cases} \tag{2.3}$$

These two extensions can be expressed in a unified way. For this purpose, define a ϱ-periodic extension, $\varrho \in \mathbb{R}$ (or an extension by periodization with period ϱ), of a

[2] In Meyer's original definition, the index of V_j runs in the opposite direction; namely, V_j is associated with basis function $\phi(2^{-j}t)$ rather than $\phi(2^j t)$. This notation is also seen in other literature such as Holschneider (1995). Here we follow the convention of Strang and Nguyen (1997), Louis, Maaß, and Rieder (1997), and Wojtaszczyk (1997).

$[0, \varrho]$-defined function $x_1(t)$ to be another function x_2 on $\mathbb{R}$ such that

$$x_2(t + \varrho\ell) = x_1(t) \qquad \forall\, t \in [0, \varrho],\, \ell \in \mathbb{Z}. \tag{2.4}$$

In this way, the previously introduced periodization becomes a 1-periodic extension; for the symmetric extension, it can be achieved through two steps: an extension from $[0, 1]$ to $[0, 2]$ by reflection, followed by a periodization with period two. These two schemes are essentially the same; we need only study the case of periodization with some generic period ϱ.

2.3. Multiresolution approximation of $L_2[0, \varrho]$. The above boundary extension yields an MRA over $[0, \varrho]$ or a circle $\mathbb{T}$ (a torus in the higher-dimensional case) with period ϱ. Hernández and Weiss (1996) have given a salient treatment for the case $\varrho = 1$. Here we restate their results on $[0, \varrho]$.

Suppose we have an $x_1(t) \in L_2[0, 1]$ and an extension x_2 by ϱ-periodization. The transform of x_2 with respect to the orthonormal scaling basis $\{\phi_n^j(t)\}_n$, where

$$\phi_n^j(t) = 2^{j/2}\phi\left(2^j t - n - \frac{1}{2}\right),$$

of the MRA over $\mathbb{R}$ is thus

$$\begin{aligned}\int_{\mathbb{R}} x_2(t)\phi_n^j(t)\,dt &= \sum_{\ell\in\mathbb{Z}} \int_0^{\varrho} x_1(t)\,\phi_n^j(t + \varrho\ell)\,dt \\ &= \int_0^{\varrho} x_1(t)\left(\sum_{\ell\in\mathbb{Z}} \phi_n^j(t + \varrho\ell)\right) dt. \end{aligned} \tag{2.5}$$

Denote

$$\phi_n^{\varrho,j}(t) = \sum_{\ell\in\mathbb{Z}} \phi_n^j(t + \varrho\ell). \tag{2.6}$$

The transform of $x_2(t)$ with respect to $\{\phi_n^j\}_{n\in\mathbb{Z}}$ over $\mathbb{R}$ is then equal to the transform of $x_1(t)$ with respect to $\{\phi_n^{\varrho,j}\}_{n\in\mathcal{N}_j}$ over $[0, \varrho]$. One naturally conjectures that a finite-dimensional subspace of $L_2[0, \varrho]$ spanned by $\{\phi_n^{\varrho,j}(t)\}_{n\in\mathcal{N}_j}$, $\mathcal{N}_j = \{0, 1, \ldots, 2^j\varrho - 1\}$, may be introduced. In the following, we show this is indeed true, and the resulting subspace possesses some interesting properties.

First, we claim that the infinite sum in (2.6) converges, provided that ϕ satisfies some localization requirement. In fact, we have the following.

PROPOSITION 2.1. *If ϕ has a localization $\gamma > 1$,* (2.6) *converges uniformly on* $[0, \varrho]$ *for* $n = 0, 1, 2, \ldots, 2^j - 1$.

Proof. This follows easily from the definition of polynomial localization (see Hernández and Weiss (1996)). □

The resulting sequence $\{\phi_n^{\varrho,j}(t)\}_{n\in\mathcal{N}_j}$ is finite and square integrable. It thus lies in $L_2[0, \varrho]$. Moreover, we have the following.

PROPOSITION 2.2. $\{\phi_n^{\varrho,j}(t)\}_{n\in\mathcal{N}_j}$ *furnishes an orthonormal system in* $L_2[0, \varrho]$.

Proof. See Appendix A. □

Proposition 2.2 allows us to build a subspace of $L_2[0, \varrho]$

$$V_{\varrho,j} = \operatorname{span}\left\{\phi_n^{\varrho,j}\right\}_{n\in\mathcal{N}_j} \subset L_2[0, \varrho] \tag{2.7}$$

with dimension $2^j\varrho$. This procedure may be repeated again and again for any positive integer j, resulting in a sequence of finite-dimensional (and hence closed) subspaces of $L_2[0,\varrho]$, $\{V_{\varrho,j},\ j=0,1,2,\ldots\}$. The sequence possesses a property of inclusion.

PROPOSITION 2.3. *$V_{\varrho,j}\subset V_{\varrho,j+1}$ for all $j=0,1,2,\ldots$.*

Proof. See Appendix B. □

Moreover, it gives a hierarchy of approximations toward $L_2[0,\varrho]$.

PROPOSITION 2.4. *$\bigcup_{j=0}^{\infty} V_{\varrho,j}$ is dense in $L_2[0,\varrho]$.*

Proof. See Appendix C. □

By Propositions 2.3 and 2.4, $\lim_{j\to\infty} V_j$ is dense in $L_2[0,\varrho]$. This is to say, $V_{\varrho,j}$ is an approximation, called a multiresolution approximation, of $L_2[0,\varrho]$ at level j. Parallel to the MRA of $L_2(\mathbb{R})$, $\{V_{\varrho,j}\}$ is called an MRA of $L_2[0,\varrho]$.

2.4. Mathematical framework. The MRA $\{V_{\varrho,j}\}$ furnishes the mathematical framework for our study. We now show that all our problems can be studied within some multiresolution approximation of $L_2[0,\varrho]$.

In practice, signals for real problems are discrete as well as of finite length. For convenience, suppose the signals of concern are output at $N=2^{j_2}$ equidistant locations for some positive integer j_2. (It will be clear in the next section why we choose j_2 as the notation.) Take, for example, some $x(t)$, t being scaled by the signal duration. This is equivalent to saying that we have a realization $x_n=x(t_n)$, with $t_n=n/N$, and $n=0,1,2,\ldots,N-1$. Since we have no a priori information about the features on scales less than $1/N$, what we need to justify is that the interpolation of these N x_n's with some basis lies in some $V_{\varrho,j}$ (and hence in any multiresolution approximations of $L_2[0,\varrho]$ with level higher than j). Choose the interpolation basis to be $\{\phi_n^{\varrho,j}\}_{n\in\mathcal{N}_j}\subset V_{\varrho,j}$, $\dim=2^j\varrho$. Then we need to show that

$$x_n=\sum_{m\in\mathcal{N}_j}\alpha_m\phi_m^{\varrho,j}(t_n),\qquad t_n=\frac{n}{N},\qquad n=0,1,\ldots,N-1, \tag{2.8}$$

has a unique solution for the coefficients α_m. If we choose $j=j_2$, the set furnishes $N=2^{j_2}$ equations with N unknown α's. Written in a matrix form, this is

$$\underline{\underline{\mathbf{H}}}\,\underline{\alpha}=\underline{\mathbf{x}}, \tag{2.9}$$

with the entries of $\underline{\underline{\mathbf{H}}}$ formed by summing up the scaling basis function:

$$H_{nm}=\phi_m^{\varrho,j_2}(t_n)=\sum_{\ell\in\mathbb{Z}}\phi_{m-\varrho\ell N}^{j_2}(t_n). \tag{2.10}$$

If $\underline{\underline{\mathbf{H}}}$ is invertible, then $\alpha=\underline{\underline{\mathbf{H}}}^{-1}\underline{\mathbf{x}}$, and this x thus interpolated lies in V_{ϱ,j_2}. Therefore, the question whether $x\in V_{\varrho,j_2}$ is justifiable is transformed into a problem regarding the invertibility of $\underline{\underline{\mathbf{H}}}$. The following proposition gives this question an answer for the case $\varrho=1$.

PROPOSITION 2.5. *When $\varrho=1$, the matrix $\underline{\underline{\mathbf{H}}}$ formed above is of full rank if the scaling function $\phi(t)$ is maximized at $t=0$ and*

$$|\phi(n)|\le\frac{\phi(0)}{(1+n^2)^{\gamma/2}}\qquad\forall\ n\in\mathbb{Z} \tag{2.11}$$

for some $\gamma>1$, which satisfies the inequality

$$\frac{2}{2^{\gamma/2}}+\frac{2}{2^{\gamma-1}-1}+\left(1+\frac{4}{1-2^{1-\gamma}}\right)\frac{1}{N^{\gamma-1}}<1, \tag{2.12}$$

where $N = 2^{j_2}$.

Proof. See Appendix E. ∎

Note that we must have $\frac{2}{2^{\gamma/2}} + \frac{2}{2^{\gamma-1}-1} < 1$, or $\gamma > 3.1$, to make the inequality (2.12) valid. Should this be true, (2.12) will hold, provided that N exceeds some threshold value, a requirement which is usually very easy to meet. For example, when $\gamma = 4$, a signal with a length $N \geq 3$ will always have (2.12) satisfied.

When $\varrho = 2$, one may also prove that $\underline{\underline{\mathbf{H}}}$ is of full rank given that (2.11) and (2.12) are satisfied. But note that our purpose is to justify $x \in V_{\varrho,j_2}$. If that is true when $\varrho = 1$, then it is also true for the extension of x with $\varrho = 2$. To see this, we symmetrically extrapolate x to $2N$ points and interpolate to obtain a function $x'(t)$, $t \in [0, 2]$. Rescale the argument of x' to construct a new function

$$y(t) = x'(2t), \qquad t \in [0, 1], \tag{2.13}$$

and a corresponding series

$$y_n = y\left(\frac{n}{2N}\right) = x'_n, \qquad n = 0, 1, \ldots, 2N-1. \tag{2.14}$$

At this point Proposition 2.5 applies. Note that the N in (2.12) now should be replaced with $2N = 2^{j_2+1}$, which further weakens the already weak condition (2.12), given a scaling function. Thereby $y(t)$, $t \in [0, 1]$, lies in V_{ϱ,j_2+1}. That is to say, when (2.12) is satisfied, there exists a unique series α_n such that

$$\begin{aligned} x'(t) = y(t/2) &= \sum_{m=0}^{2N-1} \alpha_m \phi_m^{\varrho,j_2+1}(t/2), \qquad t \in [0, 2] \\ &= \sum_{m=0}^{2N-1} \sqrt{2}\alpha_m \phi_m^{\varrho,j_2}(t). \end{aligned} \tag{2.15}$$

This means that x', a symmetric extension of x, lies in V_{ϱ,j_2} with period 2, just as expected.

We have justified that all signals of interest lie in V_{ϱ,j_2} for some j_2 large enough. We have also obtained a sequence of closed spaces $\{V_{\varrho,j}\}$, $j \leq j_2$, which approximate $L_2[0, \varrho]$ at different scale levels. All the $V_{\varrho,j}$, $0 \leq j \leq j_2$, are sampling spaces. They provide a framework for the establishment of the MWT.

3. Multiscale window transform.

3.1. Scale window. We study our problem in V_{ϱ,j_2}, with j_2 prescribed according to the signal of concern. From the closed sequence $\{V_{\varrho,j}\}$, $j < j_2$, built above, the index j represents a level of scale: the larger the j, the finer the scale. So we may construct out of $V_{\varrho,j}$ the scale ranges we need for the decomposition. Suppose V_{ϱ,j_2} is to be split in the domain of scale exclusively into three ranges (more ranges can likely be split), and suppose the ranges are divided at two scale levels, j_0 and j_1, $j_0 < j_1 < j_2$. By the inclusion property (Proposition 2.3), we have $V_{\varrho,j_0} \subset V_{\varrho,j_1} \subset V_{\varrho,j_2}$. This implies that the following orthogonal decompositions may be performed:

$$V_{\varrho,j_2} = V_{\varrho,j_1} \oplus V_{\sim 2}, \tag{3.1}$$

$$V_{\varrho,j_2} = V_{\varrho,j_0} \oplus V_{\sim 1}, \tag{3.2}$$

where $V_{\sim 2}$ and $V_{\sim 1}$ are two subspaces of V_{ϱ,j_2}. They together with $V_{\varrho,j_0} \equiv V_{\sim 0}$ form a mutually orthogonal decomposition of V_{ϱ,j_2}, each containing a certain range

of scales. We will refer to them as *submesoscale window*, *mesoscale window*, and *large-scale window*, respectively. Note that here in the subscripts "$\sim$" is used to stand for "window," and 0, 1, 2 for large-scale, mesoscale, and submesoscale windows. For convenience, they may be alternatively referred to as window 0, window 1, and window 2, respectively. The following is a summary of these notations and definitions:

Scale window	Notation	Definition
Large-scale window	$V_{\sim 0}$	$V_{\varrho,j_0} \quad (j_0 \geq 0)$
Mesoscale window	$V_{\sim 1}$	$V_{\varrho,j_1} \ominus V_{\varrho,j_0} \quad (j_1 > j_0)$
Submesoscale window	$V_{\sim 2}$	$V_{\varrho,j_2} \ominus V_{\varrho,j_1} \quad (j_2 > j_1)$

3.2. Multiscale window transform. The representation of signals on the scale windows motivates the introduction of a transform which we will hereafter refer to as the *multiscale window transform* (MWT). For a function $p \in V_{\varrho,j_2}$, first perform a scaling transform

$$\widehat{p}_n^j = \int_0^\varrho p(t)\, \phi_n^{\varrho,j}(t)\, dt \tag{3.3}$$

with respect to the orthonormal basis $\{\phi_n^{\varrho,j}\}_n$. The reconstructions or syntheses of p on the three windows are

$$p^{\sim 0}(t) = \sum_{n\in\mathcal{N}_{j_2}} \widehat{p}_n^{j_0}\, \phi_n^{\varrho,j_0}(t), \tag{3.4}$$

$$p^{\sim 1}(t) = \sum_{n\in\mathcal{N}_{j_2}} \widehat{p}_n^{j_1}\, \phi_n^{\varrho,j_1}(t) - p^{\sim 0}(t), \tag{3.5}$$

$$p^{\sim 2}(t) = p(t) - \sum_{n\in\mathcal{N}_{j_2}} \widehat{p}_n^{j_1}\, \phi_n^{\varrho,j_1}(t) \tag{3.6}$$

for $t \in [0, \varrho]$. Note that $p^{\sim\varpi} \in V_{\varrho,j_2}$ for all windows $\varpi = 0, 1, 2$; it may be further transformed on the sampling space V_{ϱ,j_2} to have it represented while ensuring its localization with maximal resolution. This defines a transform of p,

$$\widehat{p}_n^{\sim\varpi} = \int_0^\varrho p^{\sim\varpi}(t)\, \phi_n^{\varrho,j_2}(t)\, dt, \tag{3.7}$$

for windows $\varpi = 0, 1, 2$. With (3.7), the three equations, (3.4)–(3.6), may be written in a unified way:

$$p^{\sim\varpi}(t) = \sum_{n\in\mathcal{N}_{j_2}} \widehat{p}_n^{\sim\varpi} \phi_n^{\varrho,j_2}(t), \tag{3.8}$$

where $\varpi = 0, 1, 2$. Equations (3.7) and (3.8) form the transform-reconstruction pair of multiscale window analysis in $L_2[0, 1]$, with n running over $\mathcal{N}_{j_2} = \{0, 1, 2, \ldots, 2^{j_2}\varrho - 1\}$.

3.3. Restriction of representation to [0, 1]. The MWT is developed in V_{ϱ,j_2}, a subspace of $L_2[0, \varrho]$, while our problem is on $[0, 1]$. That is to say, when $\varrho = 2$, what we actually are interested in is the MWT representation on $[0, 1]$. The restriction of representation to $[0, 1]$ is natural, but it raises an issue of energy conservation.

We know for $p, q \in V_{\varrho,j}$ that there is a Parseval relation

$$\int_0^\varrho pq dt = \sum_{n=0}^{2^j\varrho-1} \alpha_n \beta_n, \tag{3.9}$$

where $\alpha_n = \langle p, \phi_n^{\varrho,j} \rangle$, $\beta_n = \langle q, \phi_n^{\varrho,j} \rangle$, because of the orthonormality of $\{\phi_n^{\varrho,j}\}_n$ on $[0, \varrho]$. The Parseval relation is the basis of the power spectrum; only with it can the notion of energy be introduced with respect to an analysis. So when a representation is restricted, we must also have the same relation hold. This forms the following theorem.

THEOREM 3.1. *For all $p, q \in V_{\varrho,j}$,*

$$\int_0^1 pq dt = \sum_{n=0}^{N-1} \alpha_n \beta_n, \tag{3.10}$$

where $N = 2^j$, $\alpha_n = \int_0^\varrho p\phi_n^{\varrho,j}\, dt$, and $\beta_n = \int_0^\varrho q\phi_n^{\varrho,j}\, dt$, if $\phi(t)$ is symmetric about zero.

Proof. See Appendix D. □

Theorem 3.1 is important in that it forms the basis of energy analysis with respect to the MWT. It also sets a requirement for the candidate scaling function.

4. Properties. The MWT and reconstruction as defined by (3.7) and (3.8) possess some interesting properties. Several useful ones are listed in this section.

THEOREM 4.1. *Given $p, q \in V_{\varrho,j_2}$ and two constants $c_1, c_2 \in \mathbb{R}$, we have*

$$\widehat{(c_1 p + c_2 q)}_n^{\sim w} = c_1 \widehat{p}_n^{\sim w} + c_2 \widehat{q}_n^{\sim w}, \tag{4.1}$$

$$(p^{\sim w})^{\sim v} = (p^{\sim v})^{\sim w} = \delta_{vw}\, p^{\sim w}, \tag{4.2}$$

$$\widehat{(p^{\sim v})}_n^{\sim w} = \delta_{vw}\, \widehat{p}_n^{\sim w} \tag{4.3}$$

for $n \in \mathcal{N}_{j_2}$ and $v, w = 0, 1, 2$.

These identities follow directly from definitions (3.7) and (3.8) and the mutual orthogonality between the three windows $V_{\sim 0}$, $V_{\sim 1}$, and $V_{\sim 2}$.

THEOREM 4.2. *Suppose $p(t)$, $t \in [0, 1]$, is a function in V_{ϱ,j_2}, and a periodic extension is used ($\varrho = 1$). Suppose further that the scaling function ϕ has a polynomial localization of order $\gamma > 1$. We have, for $j_0 = 0$,*

$$p^{\sim 0}(t) = \widehat{p}_0^{j_0} = \overline{p(t)} \qquad \forall\ t \in [0, 1], \tag{4.4}$$

where the overline signifies an averaging over the duration, i.e., $\overline{p(t)} = \int_0^1 p(t)\, dt$.

Proof. We appeal to the fact that if the scaling function ϕ is of polynomial localization of order $\gamma > 1$, then

$$\phi_n^{\varrho,0}(t) = \sum_{\ell \in \mathbb{Z}} \phi\left(t + \ell - n - \frac{1}{2}\right) = 1 \tag{4.5}$$

for all $t \in [0, 1]$, a direct consequence of the fact that $\sum_{l \in \mathbb{Z}} \phi(t + l) = 1$ when $\gamma > 1$ (see, e.g., Hernández and Weiss (1996, p. 222); but the condition they impose is weaker). By property (4.3),

$$\widehat{(p^{\sim 0})}_n^{\,0} = \widehat{(p^{\sim 0})}_n^{\sim 0} = \widehat{p}_n^{\sim 0} = \widehat{p}_n^{j_0} = \widehat{p}_n^0.$$

So

$$\begin{aligned} p^{\sim 0}(t) &= \sum_{n=0}^{2^0-1} \widehat{(p^{\sim 0})}_n^{\,0} \cdot \phi_n^{\varrho,0}(t) = \widehat{p}_0^0 \\ &= \int_0^1 p(t)\phi_0^{\varrho,0}(t)\, dt = \int_0^1 p(t)\, dt = \overline{p(t)}. \end{aligned}$$

□

Theorem 4.2 implies that, for a periodically extended signal, when the large-scale window has an upper level bound $j_0 = 0$, its synthesis on this window is simply the mean or average over the duration. If we consider only two windows, the reconstruction on window 1 is the fluctuation from the mean. The classical MED is thus recovered if we set $j_0 = 0$ and $\varrho = 1$ (periodic extension) and consider only two windows. This shows how the classical MED is generalized in our MWT.

An interesting observation is that, in Theorem 4.2, $p^{\sim 0} = 2^{j_2/2}\widehat{p}_n^{\sim 0} = \bar{p}$ for all $n \in \mathcal{N}_{j_2}$. That is to say, in this case, the large-scale transform and large-scale reconstruction are essentially the same, both equal to the mean over the definition domain (except for a constant factor). In other words, when taking a mean in the classical framework, one could be performing a transform, or doing a reconstruction, which are completely different in functional analysis.

Another useful theorem connects quadratic quantities in the physical domain to counterparts in the transform domain.

THEOREM 4.3. *Given $p(t), q(t) \in V_{\varrho,j_2}$, the following equality holds:*

$$\sum_{n=0}^{N-1} \widehat{p}_n^{\sim\varpi}\widehat{q}_n^{\sim\varpi} = \int_0^1 p^{\sim\varpi}(t)q^{\sim\varpi}(t)\,dt \qquad (N = 2^{j_2}) \tag{4.6}$$

for all scale windows $\varpi = 0, 1, 2$.

Proof. For any $p, q \in V_{\varrho,j_2}$, we can find the multiscale synthesis $p^{\sim w}, q^{\sim w} \in V_{j_2}$ for windows $w = 0, 1, 2$. Let $N = 2^{j_2}$, substitute, respectively, $p^{\sim w}$, $q^{\sim w}$ for the p, q in the equality of Theorem 3.1, and use definition (3.7), which may be alternatively stated as

$$(\widehat{p^{\sim w}})_n^{j_2} = \widehat{p}_n^{\sim w},$$

and the result then follows immediately. □

The sum of transform coefficients squared over $n \in \mathcal{N}_{j_2}$ has been called "marginal energy" by Huang, Shen, and Long (1999), in analogy to the nomenclature of probability theory. We will henceforth refer to $\sum_n \widehat{p}_n^{\sim w}\widehat{q}_n^{\sim w}$ as the *marginalization of* $\widehat{p}_n^{\sim w}\widehat{q}_n^{\sim w}$ *over* $\mathcal{N}_{j_2}$ and, accordingly, Theorem 4.3 as the property of marginalization. Using the overline notation for duration averaging, this property may be restated succinctly as

$$\boxed{\sum_{n=0}^{N-1} (\widehat{p}_n^{\sim w}\widehat{q}_n^{\sim w}) \;=\; \overline{p^{\sim w}q^{\sim w}}} \tag{4.7}$$

for the three windows $w = 0, 1, 2$.

5. Energy representation. The property of marginalization allows for a simple representation of the energy on a specified window and location. For $p \in V_{\varrho,j_2}$, $N = 2^{j_2}$, let $E_n^{\varpi *} \equiv (\widehat{p}_n^{\sim\varpi})^2$ $(\varpi = 0, 1, 2)$. By (4.7),

$$\sum_{n=0}^{N-1} E_n^{\varpi *} \;=\; \int_0^1 [p^{\sim w}(t)]^2\,dt, \tag{5.1}$$

which is the energy on window ϖ (up to some constant factor) integrated over $[0, 1]$. $E_n^{\varpi *}$ thus can be viewed as the energy of window ϖ summed over a small interval of length $\Delta t = 2^{-j_2}$ around location $t = 2^{-j_2}(n + \frac{1}{2})$. An energy variable for window ϖ

at $2^{-j_2}(n+\frac{1}{2})$ consistent with the fields at that location is therefore a locally averaged quantity

$$\boxed{E_n^{\varpi} = \frac{1}{\Delta t} E_n^{\varpi *} = 2^{j_2} \cdot (\widehat{\widetilde{p}_n^{\varpi}})^2} \tag{5.2}$$

for $\varpi = 0, 1, 2$.

6. Connection to other transforms. It is interesting to compare the MWT to some well-known analysis techniques. First, the MWT is a generalization of the classical MED to include localization and more ranges of scales, as established in the preceding section. Particularly, the averaging in the classical MED could mean a transform or a reconstruction in our framework.

The MWT is naturally connected to wavelet analysis in that both of them are based on MRA. The largest difference between them is that the MWT is not a transform in the usual sense; there is no specific basis in the domain of scales. Rather, it is an orthogonal complementary subspace decomposition, and, as a result, the MWT coefficients contain information with a range of scales, instead of a single scale. (Because of this, it is required that window bounds be specified a priori.) Furthermore, the MWT transform is projected onto V_{ϱ,j_2}, and so the transforms obtained for all the windows have the same resolution—the maximal resolution allowed by the signal. This is in contrast to wavelet analysis, whose transform coefficients have a different resolution on different scales (cf. Figure 3).

The HHT is a localized transform which has recently found important applications in nonlinear wave studies, fault identification, etc. Loosely speaking, it separates a signal into intrinsic modes via empirical mode decomposition and then applies the Hilbert transform; details are contained in Huang, Shen, and Long (1999). Notice that the decomposition and transform parallels the two steps with the syntheses (3.4)–(3.6) and the transform (3.7). The difference is that, in the MWT, the second-time transform is with respect to a basis in the same family as the one employed in the a priori decomposition. The largest difference, of course, is that the MWT is built for the derivation of physical laws, and the resulting quantities such as energy are intended for quantitative and objective representation, rather than qualitative description, of physics. As will be elaborated in an upcoming paper, the orthonormality of $\{\phi_n^{\varrho,j_2}\}_n$ allows for simple expression of a law governing the mean-eddy-turbulence interaction in turbulent fluids. On the contrary, the kernel of the Hilbert transform, $\frac{1}{t-t'}$, does not have the orthonormal property, and hence how it may be incorporated into this study is still not clear.

7. The scaling function. A scaling basis $\{\phi(t-k)\}_{k\in\mathbb{Z}}$ is required to fulfill the MWT construction. According to the above sections, the basis function ϕ must meet the following requirements:

1. $\{\phi(t-k)\}_{k\in\mathbb{Z}}$ is orthonormal;

2. $\phi(t)$ is symmetric about and maximized at the origin, and $|\phi(n)| \le \frac{\phi(0)}{(1+n^2)^{\gamma/2}}$, for $n \in \mathbb{Z}$;

3. $\phi(t)$ is polynomially localized up to order $\gamma > 3.1$, i.e., $|\phi(t)| \le \kappa_\gamma^C(t) = \frac{C}{(1+t^2)^{\gamma/2}}$, $\gamma > 3.1$ (see (2.1)).

There is a large family of ϕ meeting these requirements. This section briefly introduces one built out of the cubic spline to realize our MWT.

7.1. B-splines. It is well known that splines, or B-splines as they are usually referred to, provide bases for certain piecewise polynomials (see, e.g., Strang and Nguyen (1997)). They can be constructed to any degree. For the widely used cubic spline, it consists of polynomials of the third degree on unit intervals within $[-2, 2]$ with derivatives continuous up to second order. Splines can be defined with time-domain box functions. $n+1$ box functions convolve to form an nth degree B-spline. Particularly, a cubic B-spline is

$$\phi_3(t) = B(t) * B(t) * B(t) * B(t) \tag{7.1}$$

with the box function $B(t)$ defined as

$$B(t) = \phi_0(t) = \begin{cases} 1, & |t| < \frac{1}{2}, \\ 0 & \text{otherwise.} \end{cases}$$

Figure 4 is a plot of $\phi_3(t)$ versus t.

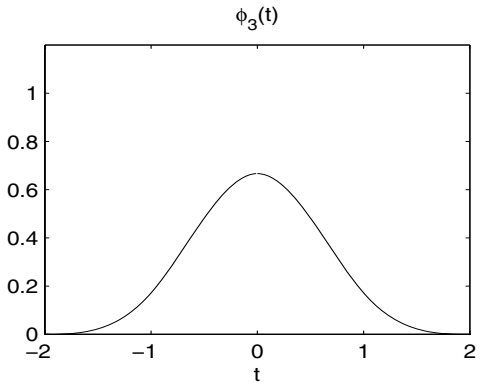

FIG. 4. *B-spline of the third degree.*

The cubic B-spline has a compact support; it is also symmetric if the data points are evenly spaced (we do not consider irregularly spaced series in this study), as can easily be seen from Figure 4. It is, however, not orthonormal on translating (Daubechies (1992) and Strang and Nguyen (1997)). As orthonormality is of paramount importance for this study, we orthonormalize $\phi_3(t)$ in the following.

7.2. Orthonormalization. There are several ways toward the orthonormalization. An easy one, as elucidated in Strang and Nguyen (1997), manipulates the basis in the Fourier domain.

Suppose there is a translation invariant nonorthonormal basis $\{\phi_*(t-k)\}_{k\in\mathbb{Z}}$, which spans $V_0 \subset L_2(\mathbb{R})$. We want to orthonormalize it to $\{\phi(t-k)\}_{k\in\mathbb{Z}}$. In the time domain, the orthonormality of $\{\phi(t-k)\}_{k\in\mathbb{Z}}$ reads that

$$a(\phi, k) \equiv \int_{-\infty}^{\infty} \phi(t)\phi(t-k)\,dt = \delta(k),$$

which, if transformed to the frequency domain, becomes

$$A(\hat{\phi}, \omega) = \sum_{n=-\infty}^{\infty} |\hat{\phi}(\omega + 2\pi n)|^2 = 1.$$

This condition is also sufficient; a detailed proof can be found in Strang and Nguyen (1997). We thus have

$$\{\phi(t-k)\}_{k\in\mathbb{Z}} \text{ orthonormal} \qquad \text{iff} \qquad A(\hat{\phi},\omega)=1. \tag{7.2}$$

If, for some dilation function ϕ_*, $A(\hat{\phi}_*,\omega) > 0$, condition (7.2) implies that we may orthonormalize $B_* = \{\phi_*(t-k)\}_{k\in\mathbb{Z}}$ in the frequency domain by dividing it by $\sqrt{A(\hat{\phi}_*,\omega)}$. The positivity of $A(\hat{\phi}_*,\omega)$ is guaranteed (a property of Riesz basis; see Strang and Nguyen (1997, Theorem 6.13)). It is shown in Liang (2002) that the orthonormalized sequence

$$B_{orth} \equiv \{\phi(t-k)\}_{k\in\mathbb{Z}}, \qquad \phi(t)=\mathcal{F}^{-1}\left(\frac{\hat{\phi}_*(\omega)}{\sqrt{A(\hat{\phi}_*,\omega)}}\right), \tag{7.3}$$

$\mathcal{F}^{-1}$ being an inverse Fourier transform, indeed generates V_0. That is to say, B_* and B_{orth} span the same subspace of $L_2(\mathbb{R})$.

With this result we proceed to orthonormalize the cubic spline. Take the Fourier transform of (7.1),

$$\hat{\phi}_3(\omega)=\operatorname{sinc}^4\left(\frac{\omega}{2}\right)=\left(\frac{\sin\omega/2}{\omega/2}\right)^4, \tag{7.4}$$

where ω is frequency in radians, and then divide by $\sqrt{A(\hat{\phi}_3,\omega)}$ to obtain

$$\hat{\phi}(\omega)=\frac{\hat{\phi}_3(\omega)}{\sqrt{A(\hat{\phi}_3,\omega)}}, \tag{7.5}$$

where

$$A(\hat{\phi}_3,\omega)=\sum_{n=-\infty}^{\infty}|\hat{\phi}_3(\omega+2\pi n)|^2. \tag{7.6}$$

An inverse Fourier transform of the resulting $\hat{\phi}(\omega)$ yields the desired $\phi(t)$. Plotted in Figure 5 are $\phi_3(\omega)$, $\phi(\omega)$, and $\phi(t)$.

The $\phi(t)$ obtained is symmetric about and maximized at $t=0$. It decays exponentially fast as $|t|$ goes to infinity (Holschneider (1995)) and hence should meet any polynomial localization needs. As shown in Figure 6, $|\phi(t)|$ lies below $\kappa_2^c(t)$ with $C=3$ and $\gamma=4>3.1$. It is also easy to check that the condition of Proposition 2.5 is satisfied. In other words, the B_{orth} of (7.3) is what we are seeking.

The resulting bases $\phi_n^{\varrho,j}(t)$ are plotted in Figure 7 for a selection of scale level j. Boundary effects are easily seen, but they are limited within a distance of the order of the size of the main lobe. This implies that if the signal under consideration is long enough, the boundary effects should not be a problem in practice. We will have a chance to see this again with two examples in section 9.

8. Computing the multiscale window transform. Given a $p\in V_{\varrho,j_2}$, the key to its multiscale window analysis is finding

$$\widetilde{p}_n^j=\int_0^{\varrho}p(t)\,\phi_n^{\varrho,j}(t)\,dt \tag{8.1}$$

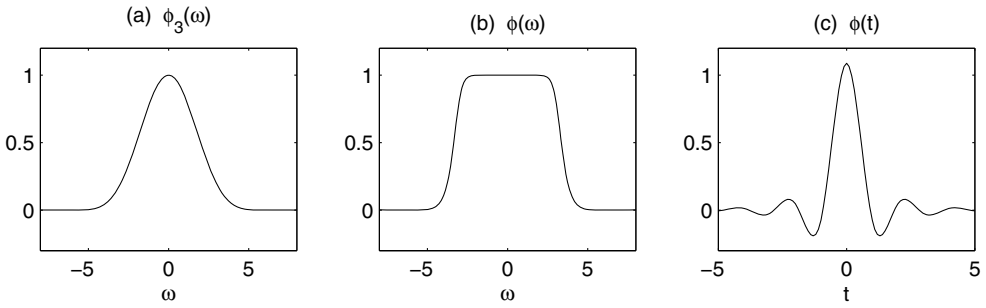

FIG. 5. *The cubic B-spline (in frequency domain ω in radians)* (a) *and its orthonormalization (in both frequency domain* (b) *and time domain* (c)*).*

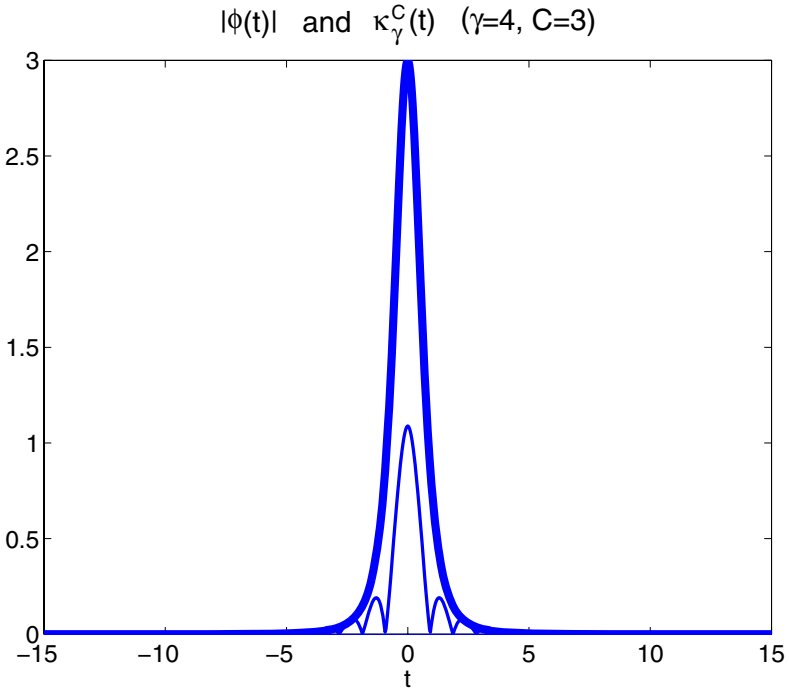

FIG. 6. *The absolute cubic B-spline scaling function $|\phi(t)|$ versus $\kappa_\gamma^C(t) = \frac{C}{(1+t^2)^{\gamma/2}}$ (thick line) with $\gamma = 4 > 3.1$, $C = 3$.*

for the three window bounds $j = j_0, j_1, j_2$. Once this is done, all the window transforms and reconstructions can be directly computed from definitions (3.7) and (3.8).

Computation of (8.1) is realized via two steps. First, suppose that we already know $\widehat{p}_n^{j_2}$. Notice that (8.1) is equivalent to the scaling transform of the periodic or symmetric extension of p over $\mathbb{R}$ with respect to $\{\phi_n^j\}_{n\in\mathbb{Z}}$. So we can avail ourself of the arsenal of fast scaling transform to conveniently find the $\widehat{p}_n^j$ at all levels below j_2. In a formula unifying the two extension schemes,

$$\widehat{p}_n^j = \sum_{m\in\mathcal{N}_j} \widehat{p}_m^{j+1} \sum_{\ell\in\mathbb{Z}} h_0(m - 2n - 2^{j+1}\varrho\ell), \tag{8.2}$$

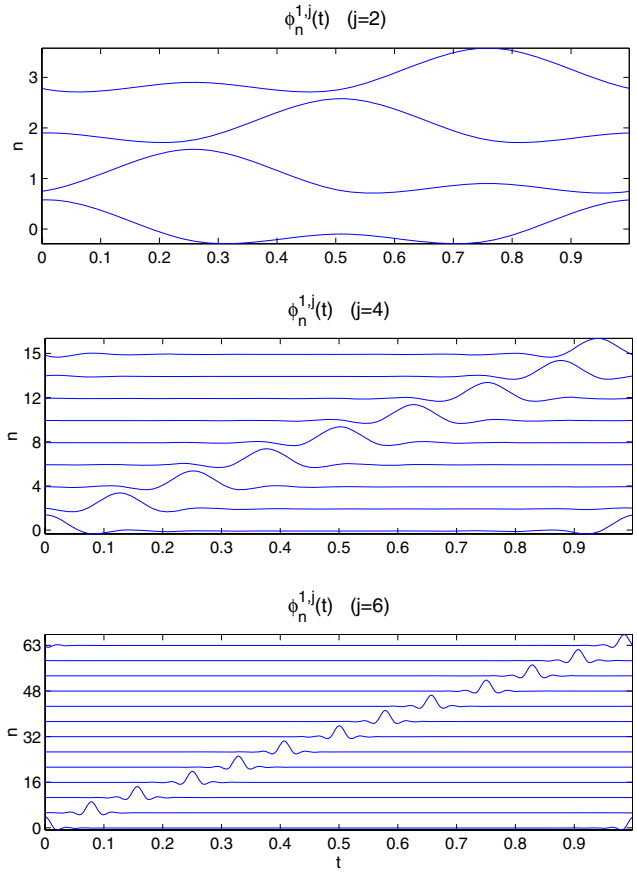

FIG. 7. *The* $\{\phi_n^{\varrho,j}(t)\}_n$ *built out of the cubic spline with* $\varrho = 1$.

where h_0 is a low-pass filter constructed from ϕ:

$$h_0(k) = \frac{1}{\sqrt{2}} \int_{\mathbb{R}} \phi\left(\frac{t}{2}\right) \phi(t-k)\, dt. \tag{8.3}$$

For more about scaling transforms and the filter bank h_0, we refer the reader to Strang and Nguyen (1997).

We still need to find $\widehat{p}_n^{j_2}$. To do this, two cases are distinguished: $\varrho = 1$ and $\varrho = 2$. First, look at the case $\varrho = 1$. Recall that observations of a function $p(t)$ are available only at $2^{j_2} = N$ points: $t_n = \frac{n}{N}$, $n = 0, \ldots, N-1$. It has been justified in section 2 that $p(t) \in V_{\varrho,j_2}$. The values p take at these points then must satisfy

$$p(t_m) = \sum_{n=0}^{N-1} \widehat{p}_n^{j_2} \phi_n^{\varrho,j_2}(t_m), \qquad m = 0, \ldots, N-1. \tag{8.4}$$

In a matrix form, this is $\underline{\underline{\mathbf{H}}}\, \underline{\widehat{\mathbf{p}}}^{j_2} = \underline{\mathbf{p}}$, where $\underline{\mathbf{p}} = \left[p(t_0), p(t_1), \ldots, p(t_{N-1})\right]^T$, $\underline{\widehat{\mathbf{p}}}^{j_2} = \left[\widehat{p}_0^{j_2}, \widehat{p}_1^{j_2}, \ldots, \widehat{p}_{N-1}^{j_2}\right]^T$, and

$$\underline{\underline{\mathbf{H}}} = \begin{bmatrix} H_{0,0} & H_{0,1} & \dots & H_{0,N-1} \\ H_{1,0} & H_{1,1} & \dots & H_{1,N-1} \\ \vdots & \vdots & \vdots & \vdots \\ H_{N-1,0} & H_{N-1,1} & \dots & H_{N-1,N-1} \end{bmatrix}, \tag{8.5}$$

with the entries $H_{mn} = \phi_n^{\varrho,j_2}(t_m)$ for $m, n = 0, 1, \dots, N-1$. Given the ϕ as built in the preceding section, this $\underline{\underline{\mathbf{H}}}$ is invertible. So

$$\underline{\widehat{\mathbf{p}}}^{j_2} = \underline{\underline{\mathbf{H}}}^{-1}\underline{\mathbf{p}}. \tag{8.6}$$

Notice that $\underline{\underline{\mathbf{H}}}$ is determined completely by the structure of the chosen space; it has no dependence on the signal p. In other words, the matrix inverse in (8.6) may be obtained once and for all. Notice further that, when $\varrho = 1$, $H_{m,n+N} = H_{m,n}$, and $H_{m,n} = H_{m+\alpha,n+\alpha}$, for all integers α, m, and n that make the indices meaningful, which is due to the periodic property of $\phi_n^{\varrho,j_2}(t)$. This fact implies that $\underline{\underline{\mathbf{H}}}$ is a circulant matrix, and hence $\underline{\underline{\mathbf{H}}}^{-1}$ is also circulant (see Davis (1979)). This $\underline{\underline{\mathbf{H}}}^{-1}$ therefore behaves like a cyclic filter (except for a constant multiplier), pretreating signals before they enter the fast analysis bank. (Here the "prefilter" is from the rows of the circulant matrix $\underline{\underline{\mathbf{H}}}^{-1}$.) An example of the prefilter for $j_2 = 10$ and $\varrho = 1$ is shown in Figure 8. We see that it is rather weak (side lobes negligible compared to the value at zero) and exerts effects only on grid-size features.

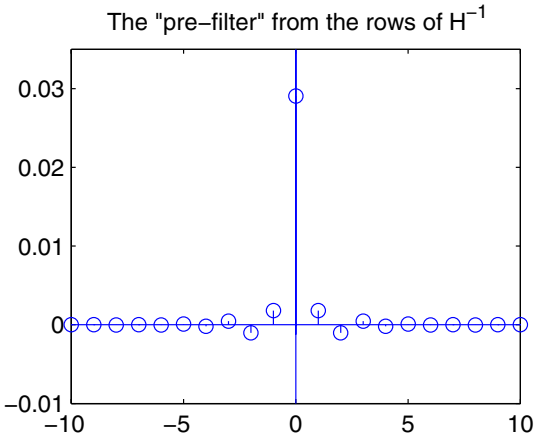

FIG. 8. *The "prefilter" formed from the rows of the* $\underline{\underline{\mathbf{H}}}^{-1}$ *with* $\varrho = 1$ *and* $j_2 = 10$ *for the computation of* $\widehat{p}nj_2$.

For the case $\varrho = 2$, we first extend the discrete signal $p(t_n)$ from $[0, 1]$ to $[0, 2]$ by reflection. The problem is then changed into calculating the transform of $p(t)$ reconstructed from this extended signal. Let $q(t) = p(2t)$. $q(t)$ is hence in V_{ϱ,j_2+1} and is formed by periodic extension from domain $[0, 1]$. By the previous result, we can obtain $\widehat{q}_n^{j_2+1}$ by solving equation

$$\sum_{n=0}^{2N-1} \widehat{q}_n^{j_2+1} \sum_{l\in\mathbb{Z}} \phi_{n-2Nl}^{j_2+1}\left(\frac{m}{2N}\right) = q\left(\frac{m}{2N}\right), \tag{8.7}$$
$$m = 0, 1, \dots, 2N-1.$$

If N satisfies condition (2.12), so does $2N$. This equation must have a unique solution for $\widehat{q}_n^{j_2+1}$. By what we have shown in (2.15), it is easy to get

$$\widehat{p}_n^{j_2} = \sqrt{2}\,\widehat{q}_n^{j_2+1}, \qquad n = 0, \ldots, 2N-1. \tag{8.8}$$

Once $\widehat{p}_n^j$, $j = j_0, j_1, j_2$, are obtained, (3.7) and (3.8) immediately yield the desired MWTs and reconstructions.

9. Validation. We validate the obtained MWT with two highly localized and distinctly scale-windowed signals:

$$f_1(t) = e^{-\alpha(t-\frac{1}{2})^2} + ae^{-\beta(t-\frac{1}{2})^2} \sin\omega\left(t - \frac{1}{2}\right), \tag{9.1}$$

$$f_2(t) = t - \frac{1}{2} + ae^{-\beta(t-\frac{1}{2})^2} \sin\omega\left(t - \frac{1}{2}\right). \tag{9.2}$$

Plotted in Figure 9 are these signals sampled at 2^{j_2} ($j_2 = 10$) locations with $\alpha = 25$, $\beta = 1 \times 10^4$, $\omega = 150\pi$, $a = 0.5$. As the two-scale features are clear, we want to see whether the MWT can faithfully reconstruct them.

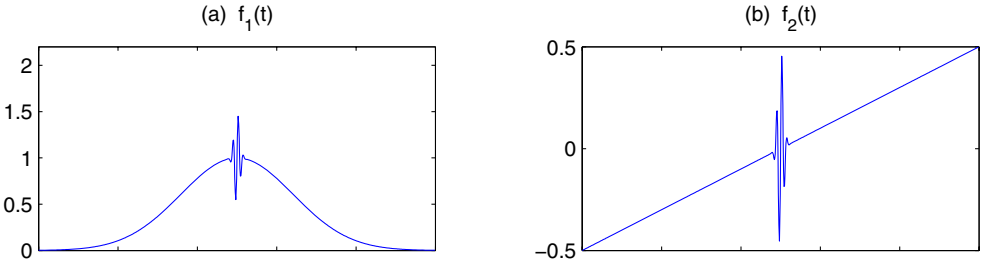

FIG. 9. *The signals sampled from $f_1(t)$ and $f_2(t)$, $t \in [0,1]$, at 2^{10} equidistant locations.*

We perform a two-scale window decomposition (windows 0 and 1), the large-scale window bound set to be $j_0 = 5$. The analysis results with $\varrho = 1$ are plotted in Figure 10, column (a) for f_1 and column (b) for f_2. In either case, the large-scale background and the highly localized oscillation in the interior region have been well reconstructed. The transform coefficients, and hence the energy, also recover the highly localized features, just as one expects. (Note that the energy has a distribution as reconstructions have, in contrast to the one-value energy in the MED framework.) However, a severe problem arises at the endpoints of case f_2: Large spurious oscillations occur due to the mismatch between the boundary values f_2 takes. Clearly, the periodic extension is not appropriate for signals such as f_2.

A different scenario in the analysis result is seen when the extension by reflection is used, i.e., when $\varrho = 2$ is chosen. As shown in Figure 11, all the features are just as expected, and the boundary oscillations for the f_2 case are almost indiscernible. In reality, symmetric extension ($\varrho = 2$) can usually meet the needs of different problems because of the continuity at the boundaries. It should always be used unless the function is precisely periodic, i.e., unless $\varrho = 1$ is a precise extension.

10. Discussion and illustration of an application. In this section we briefly show how the MWT can be applied to the study of the complex multiscale interaction as schematized in Figure 1. This kind of interaction is very important in fluid mechan-

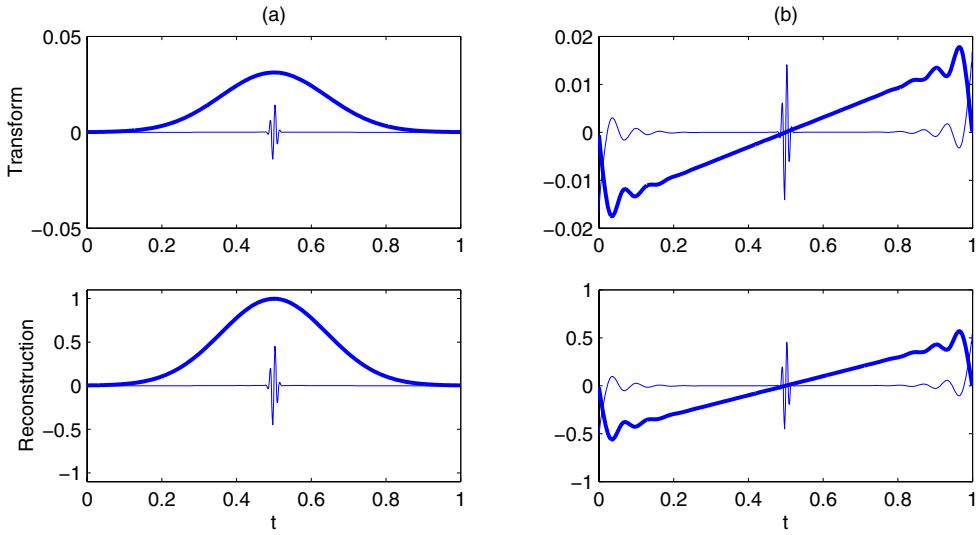

FIG. 10. *The transforms (upper row) and reconstructions (lower row) of* $f_1(t)$ *(column* (a)*) and* $f_2(t)$ *(column* (b)*), which are shown in Figure* 9*. The MWT parameters used include* $j_0 = 5$ *and* $\varrho = 1$*. In each subplot, the thick line represents the large-scale window. Note that the transform coefficients are discrete in character; they look continuous because there are* $2^{j_2} = 1024$ *sampling points.*

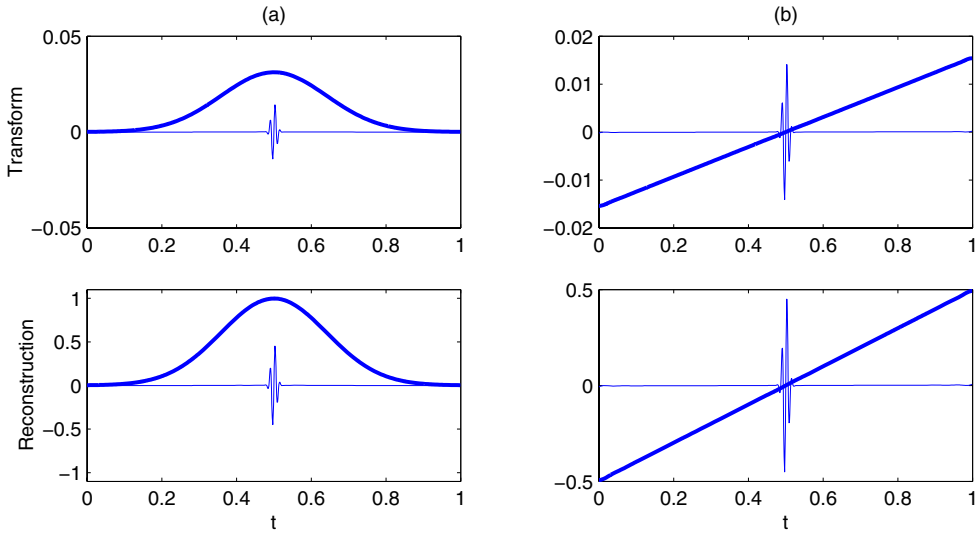

FIG. 11. *Same as Figure* 10*, but with* $\varrho = 2$*.*

ics, as it is related to hydrodynamic stability, turbulence production, laminarization, atmospheric cyclogenesis, and ocean eddy shedding, to name a few. Our intention is just to show the utility of the MWT. The reader should focus on the overall pattern in the discussion; details are deferred to other papers cited hereafter.

Consider a scalar field T advected by a flow $\underline{\mathbf{v}}$. Coherent and smaller structures may be generated if both T and $\underline{\mathbf{v}}$ vary. Represent the basic profile and these struc-

tures on the large-scale, mesoscale, and submesoscale windows, respectively. (Here we consider scale windows with respect to time.) The structure genesis problem is then boiled down to finding how energy (quadratic quantities related to T) is transferred between these windows. Denote by E_n^{ϖ} the energy $\frac{1}{2}\widehat{T}_n^{\sim\varpi}$ for scale window ϖ at time step n. It has been proved that this transfer is[3]

$$\Gamma_n^{\varpi} = E_n^{\varpi}\nabla\cdot\underline{\mathbf{v}}_T, \qquad \text{where } \underline{\mathbf{v}}_T = \frac{\widehat{(\underline{\mathbf{v}}T)}_n^{\sim\varpi}}{\widehat{T}_n^{\sim\varpi}}, \tag{10.1}$$

where $\underline{\mathbf{v}}_T$ has the dimensions of velocity, and has been referred to as T-coupled velocity. It may be viewed as a kind of averaged velocity with weights derived from the MWT transforms of T. So with the MWT, the multiscale interaction, or, more concretely, the mean-eddy-turbulence interaction, can be expressed succinctly as the divergence of the T-coupled velocity.

The above transfer Γ_n^{ϖ} is Eulerian and localized. It possesses a very interesting property, i.e.,

$$\sum_{\varpi}\sum_{n}\Gamma_n^{\varpi} = 0, \tag{10.2}$$

as proved in Liang and Robinson (2005), with the aid of Theorem 4.3 (property of marginalization). That is to say, the transfer is a mere redistribution of energy among the scale windows. It does not generate or destroy energy as a whole. For this reason, it has been called *perfect transfer*, to distinguish it from those transfers one might have encountered in the literature.

The concept of perfect transfer as defined above is the key to the understanding of the multiscale interactions in fluid flows. One may obtain the perfect transfer of kinetic energy among windows by replacing the T in (10.1) by the velocity components and then summing the resulting Γ's; one may also obtain the perfect transfer of potential energy by replacing the T by density anomaly (up to some constant factor). The resulting perfect transfers correspond to the two important geophysical fluid dynamics (GFD) instabilities, the barotropic instability and baroclinic instability (Liang and Robinson (2005)). This way the GFD instabilities can be studied locally, without the difficulties embedded in the classical formalisms (see, e.g., Pedlosky (1979)).

The localized GFD instability theory in terms of MWT has been applied to the investigation of a variety of complex oceanic processes, which otherwise would be very difficult, if not impossible, to investigate. Hereafter we briefly summarize some results from an application by Liang and Robinson (2004) on the variability of the Iceland–Faeroe front (IFF).

The IFF is a highly variable oceanic front between Iceland and the Faeroe Islands (Figure 12) which separates the waters from two major oceans, the North Atlantic and the Arctic. Understanding the dynamical processes governing its variability is very important, not only for oceanography and climate research but also for military operations and fishery industry. Extensive research effort has been invested in this region ever since the 1958 International Geophysical Year surveys (see Robinson et al. (1996)). However, because of the high nonlinearity and complexity, the dynamical processes had not been clear until the MWT and the MWT-based theory of perfect transfer recently became available (Liang and Robinson (2004)).

[3]A brief treatment can be seen in Liang and Robinson (2005); detailed derivation is deferred to a forthcoming paper. The key thing here is a unique separation of transport processes from transfer processes, which is made precise within the MWT framework.

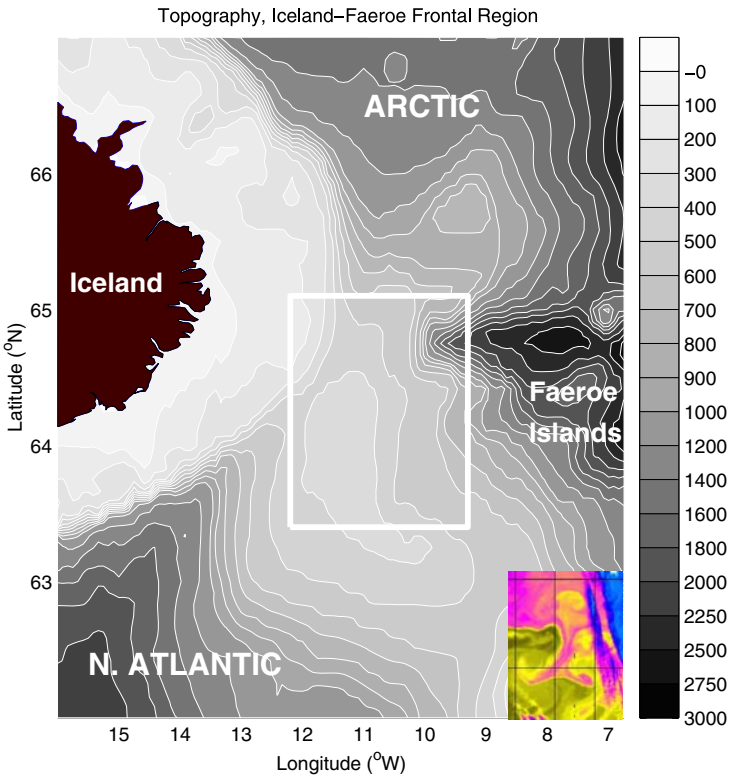

FIG. 12. *Topography of the IFF region. Inserted into the middle is the survey domain. A satellite image showing the sea surface temperature on August* 22, 1993 *is also inserted into the bottom-right corner.*

By observation, the IFF variability is basically a mesoscale process highly intermittent in space and time; it is not a process on some individual scale but on a window with a range of scales. The central issue about the mesoscale dynamics is how it derives energy from the background. So the aforementioned theory is expected to apply here.

Liang and Robinson (2004) examined a dataset acquired in August of 1993 which captured the processes toward the formation of a highly localized meandering intrusion on August 22, as shown in Figure 12 in the inserted satellite picture of sea surface temperature. This dataset has been well studied, with different results available for comparison. They first determined the window bounds. In the wavelet spectrum of the temperature series in the meandering region, a time scale window between 0.75–2.6 days is shown to fully capture the mesoscale meandering event. Plotted in Figure 13 are the large-scale and mesoscale reconstructions or syntheses for a typical time series. With the window bounds the computation of the perfect transfers is straightforward. As an example, the potential energy transfer for the depth of 300 m is computed (kinetic energy transfer is insignificant).[4] The resulting transfer sequence is drawn in Figure 14. Remarkably, appearing on the map is a clear solitary positive regime around the center during the intrusion event, while in other subregions it

[4]The depth and time steps are chosen the same as those in Figure 9 of Liang and Robinson (2004). The result here is similar but slightly different, as the program codes have since been optimized.

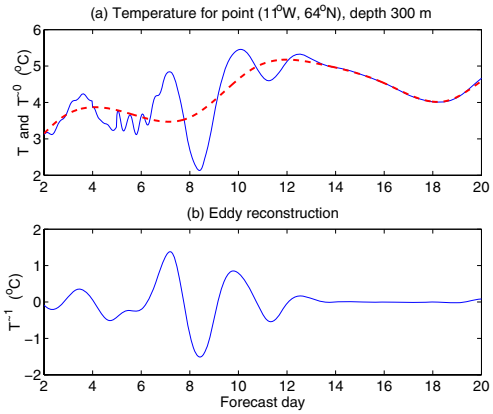

FIG. 13. *A time series of the temperature (top, solid line), and its large-scale window synthesis (top, dashed line) and mesoscale or eddy window synthesis (bottom) for point (11° W, 64° N), at depth 300 m. The decomposition is such that processes with periods shorter than and equal to 2.6 days (characteristic of the meandering intrusion) are included in the eddy window. In the bottom panel, signals with periods shorter than 0.75 day have been filtered out. The forecast starts on August 14 (day 0), 1993.*

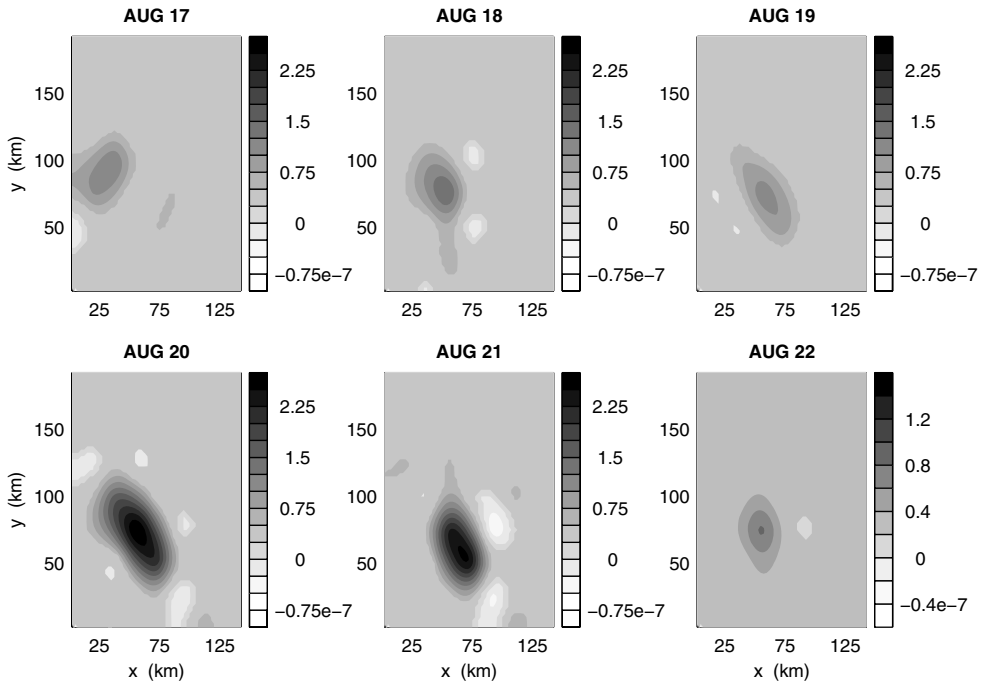

FIG. 14. *The evolution of the perfect potential energy transfer (constant factor multiplied to ensure physical interpretation) for the IFF region from August 17–22, 1993 at depth 300 m. The units are in m^2/s^3.*

is virtually zero. This is in contrast to the complicated maps of other diagnostic properties (Liang and Robinson (2004)) and the previous dynamical analyses (cf. the references in Robinson et al. (1996) and Liang and Robinson (2004)). Clearly, a

baroclinic instability is occurring here.

Moreover, the baroclinic instability actually has an interesting spatial-temporal structure. Originally, the hotspot sits near the western boundary; it then moves into the interior, where it halts and amplifies, and diminishes to zero by August 23, just after the meandering matures. That is to say, the instability originally appears in the west. Disturbances are introduced eastward into the domain along the front in a form of spatial growing mode. This is a convective instability. After August 19, the disturbances become strong enough to counteract the propagation. Correspondingly, the process is switched into a time growing mode, i.e., an absolute instability, leading to the meandering intrusion.

In a brief summary, underlying the complex IFF variability, at least in August 1993, turns out to be a baroclinic instability, first in the form of a convective instability and then switching to an absolute instability. This clear-cut dynamical scenario results from a direct application of the MWT.

11. Summary. A new apparatus, the *multiscale window transform* (MWT), has been developed for the analysis of highly localized multiscale processes. Based on Meyer's multiresolution analysis, the MWT generalizes the classical mean-fluctuation or mean-eddy decomposition (MED) to include three or more ranges of scales and to ensure a faithful representation of localized events in terms of energy.

The MWT is developed within the framework of $V_{\varrho,j_2} \subset L_2[0,1]$, a finite-dimensional sampling space spanned by a translational invariant basis built out of some scaling function ϕ. A sequence of subspaces of V_{ϱ,j_2}, $\{V_{\varrho,j}\}_{0\le j\le j_2}$, is then constructed, which allows for the introduction of the concept of scale window, i.e., a subspace of V_{ϱ,j_2} containing a range of scales. We have defined three mutually orthogonal windows: the large-scale window, mesoscale window, and submesoscale window. Given a signal p, reconstructions on these three windows followed by a scaling transform with respect to the orthonormal basis of V_{ϱ,j_2} yield the large-scale, mesoscale, and submesoscale transforms of p, respectively. These transforms and the reconstructed components form the transform-reconstruction pairs of the MWT. Properties have been explored. Of particular importance is the property of marginalization, which allows for an easy representation of energy for specified windows and locations. We have presented a realization with a scaling function ϕ orthonormalized from the cubic spline. Also presented is a fast transform scheme. Two examples have been supplied as a validation of this machinery.

The MWT generalizes the classical MED in that it retains localized information, has faithful local energy representation, and may include three or more ranges in the domain of scale. It is interesting to note that the simple operator of averaging or mean in the classical sense may be both a transform and a reconstruction, two distinctly different concepts in functional analysis. The MWT has also been compared to wavelet analysis and the HHT. Its utility was exemplified in an oceanographic application.

In this paper the transform is over one dimension. More dimensions may be added, but the development is much more difficult. For example, in the two-dimensional (2D) case, the sequence of subspaces of V_{ϱ,j_2}, $\{V_{\varrho,j}\}_j$, are not as simple as their 1D counterparts. A framework has been built in Liang (2002) to deal with these issues, within which we have studied a variety of decompositions in forming the sequence (particularly the tensor product decomposition and quincunx decomposition) (ibid.). The so-formed 2D MWT has been applied in several real problems, e.g., Liang and Robinson (2004). We will elaborate it in another paper.

Appendix A. Proof of Proposition 2.2.

Proof. We need only to show that

$$\int_0^{\varrho} \phi_n^{\varrho,j}(t)\,\phi_{n'}^{\varrho,j}(t)\,dt = \delta(n-n'), \qquad n,n' = 0,1,2,\ldots,2^j\varrho - 1.$$

By the definition of periodized bases,

$$\begin{aligned}
&\int_0^{\varrho} \phi_n^{\varrho,j}(t)\,\phi_{n'}^{\varrho,j}(t)\,dt \\
&\quad= \sum_{\ell\in\mathbb{Z}}\sum_{\ell'\in\mathbb{Z}} \int_0^{\varrho} \phi_n^j(t+\varrho\ell)\,\phi_{n'}^j(t+\varrho\ell')\,dt \\
&\quad= \sum_{\Delta\ell\in\mathbb{Z}}\sum_{\ell'\in\mathbb{Z}} \int_{\varrho\ell'}^{\varrho(\ell'+1)} \phi_n^j(t'+\varrho\Delta\ell)\,\phi_{n'}^j(t')\,dt' \\
&\qquad\qquad\qquad\qquad (t' = t+\varrho\ell',\ \Delta\ell = \ell-\ell') \\
&\quad= \sum_{\Delta\ell\in\mathbb{Z}} \int_{\mathbb{R}} \phi_{n-2^j\varrho\Delta\ell}^j(t')\,\phi_{n'}^j(t')\,dt' \\
&\quad= \int_{\mathbb{R}} \phi_n^j(t')\,\phi_{n'}^j(t')\,dt' + \sum_{\Delta\ell\neq 0} \int_{\mathbb{R}} \phi_{n-2^j\varrho\Delta\ell}^j(t')\,\phi_{n'}^j(t')\,dt'.
\end{aligned}$$

In the derivation we have used term-by-term integration, which is made legitimate by the localization assumption for ϕ (polynomially localized up to an order $\gamma > 1$) we made in the beginning. Since $n, n' \in \mathcal{N}_j$, an integer $\Delta\ell$ which is different from zero will not give any chance for $n - 2^j\varrho\Delta\ell$ to be equal to n'. By the orthonormality of $\{\phi_n^j\}_{n\in\mathbb{Z}}$, the second term of the above equation is hence zero, and the whole equation is simply $\delta(n-n')$. This finishes the proof of Proposition 2.2. ∎

Appendix B. Proof of Proposition 2.3.

Proof. We have assumed that ϱ is a power of 2. Let it be 2^λ, with λ a positive integer. This yields the following two useful identities:

$$\phi_n^{1,j+\lambda}\left(\frac{t}{\varrho}\right) = \sqrt{\varrho}\,\phi_n^{\varrho,j}(t), \tag{B.1}$$

$$\mathcal{N}_j = \mathcal{N}_1^{j+\lambda}. \tag{B.2}$$

Any $p \in V_{\varrho,j}$ has a representation

$$p(t) = \sum_{n\in\mathcal{N}_j} \alpha_n \phi_n^{\varrho,j}(t) \tag{B.3}$$

with some expansion coefficients α_n. Make a transformation

$$t' = t/\varrho, \qquad t = \varrho t'.$$

By (B.1) and (B.2),

$$\begin{aligned}
p(\varrho t') &= \sum_{n\in\mathcal{N}_j} \alpha_n \phi_n^{\varrho,j}(\varrho t') \\
&= \sum_{n\in\mathcal{N}_1^{j+\lambda}} \frac{\alpha_n}{\sqrt{\varrho}} \phi_n^{1,j+\lambda}(t').
\end{aligned}$$

This is to say,

$$p_\varrho(t') \equiv p(\varrho t') \in V_{1,j+\lambda}.$$

As proved in Hernández and Weiss (1996), we have

$$V_{1,j+\lambda} \subset V_{1,j+1+\lambda}.$$

So we see that

$$\begin{aligned} p(\varrho t') = p_\varrho(t') &= \sum_{n\in\mathcal{N}_1^{j+\lambda+1}} \beta_n \phi_n^{1,j+\lambda+1}(t') \\ &= \sum_{n\in\mathcal{N}_\varrho^{j+1}} \sqrt{\varrho}\beta_n \phi_n^{\varrho,j+1}(\varrho t'), \end{aligned}$$

where β_n are some expansion coefficients. Transformed back to t, this is

$$p(t) = \sum_{n\in\mathcal{N}_\varrho^{j+1}} \beta_n \phi_n^{\varrho,j+1}(t),$$

which means $p \in V_{\varrho,j+1}$. Since p is chosen arbitrarily from $V_{\varrho,j}$, we have

$$V_{\varrho,j} \subseteq V_{\varrho,j+1}.$$

But $\dim V_{\varrho,j+1} > \dim V_{\varrho,j}$, and so

$$V_{\varrho,j} \subset V_{\varrho,j+1}. \qquad \square$$

Appendix C. Proof of Proposition 2.4.

Proof. By the inclusion property proved above, we need only show that $\lim_{j\to\infty} V_{\varrho,j}$ is dense in $L_2[0,\varrho]$. The proof of case $\varrho = 1$ has been provided by Hernández and Weiss (1996). Moreover, they have shown that, for any function $g \in L_2[0,1]$,

$$\lim_{j\to\infty} \|g - P_1^j g\|_{C_\infty} = 0, \tag{C.1}$$

where P_1^j denotes the projection operator from $L_1[0,1]$ onto $V_{1,j}$ ($C_\infty[a,b]$ is the normed space of continuous functions that have finite extrema over $[a,b]$). Now consider function $f \in L_2[0,\varrho]$. We want to examine the performance of

$$\|f - P_\varrho^j f\|_{C_\infty}$$

when j is very large. Here $P_\varrho^j : L_2[0,\varrho] \longrightarrow V_{\varrho,j}$ projects functions in $L_2[0,\varrho]$ onto $V_{\varrho,j}$, (notice that $V_{\varrho,j} \subset C_\infty(\mathbb{R})$ for $j \geq 0$), and ϱ is by assumption a power of 2. Make a transformation of variable:

$$t' = t/\varrho, \qquad t = \varrho t',$$

and let

$$g(t') = f(\varrho t');$$

then $g \in L_2[0,1]$. By (C.1), for any $\varepsilon > 0$, there always exists a $J = J(\varepsilon) > 0$, such that

$$\|g - P_1^{j+\lambda} g\|_{C_\infty} < \varepsilon \tag{C.2}$$

if $j > J$. Here $\lambda = \log_2 \varrho$ is a positive integer, and

$$P_1^{j+\lambda} g(t') = \sum_{n \in \mathcal{N}_1^{j+\lambda}} \alpha_n \phi_n^{1,j+\lambda}(t')$$

with the expansion coefficients being

$$\begin{aligned}
\alpha_n &= \left\langle g, \phi_n^{1,j+\lambda} \right\rangle \\
&= \int_0^1 g(t')\, \phi_n^{1,j+\lambda}(t')\, dt' \\
&= \int_0^1 f(\varrho t') \cdot \sqrt{\varrho} \phi_n^{\varrho,j}(\varrho t')\, dt' \qquad \text{((B.1) applied)} \\
&= \frac{1}{\sqrt{\varrho}} \int_0^\varrho f(t)\, \phi_n^{\varrho,j}(t)\, dt \\
&= \frac{1}{\sqrt{\varrho}} \left\langle f, \phi_n^{\varrho,j} \right\rangle \equiv \frac{1}{\sqrt{\varrho}} \beta_n,
\end{aligned}$$

where $\{\beta_n\}$ are the expansion coefficients of f with respect to $\{\phi_n^{\varrho,j}\}_n$. So

$$\begin{aligned}
P_1^{j+\lambda} g(t') &= \sum_{n \in \mathcal{N}_1^{j+\lambda}} \frac{1}{\sqrt{\varrho}} \beta_n \cdot \sqrt{\varrho} \phi_n^{\varrho,j}(\varrho t') \qquad \text{((B.1) applied)} \\
&= \sum_{n \in \mathcal{N}_j} \beta_n \phi_n^{\varrho,j}(t) \qquad \text{((B.2) applied)} \\
&= P_\varrho^j f(t).
\end{aligned}$$

But we also know $g(t') = f(\varrho t') = f(t)$; what (C.2) actually states is thus

$$\|f - P_\varrho^j f\|_{C_\infty} < \varepsilon$$

as $j > J$. That is to say, given any $\varepsilon > 0$, when $j > J(\varepsilon)$, $V_{\varrho,j}$ is an ε-net of $L_2[0,\varrho]$, and hence $\lim_{j\to\infty} V_{\varrho,j}$ is dense in $L_2[0,\varrho]$. □

Appendix D. Proof of Theorem 3.1.

Proof. When $\varrho = 1$, this is precisely (3.9). When $\varrho = 2$, the Parseval relation reads

$$\int_0^2 p(t)\, q(t)\, dt = \sum_{n=0}^{2N-1} \alpha_n \beta_n. \tag{D.1}$$

Since both p and q are obtained by symmetric extension, we have

$$p(2-t) = p(t), \qquad q(2-t) = q(t). \tag{D.2}$$

So the left-hand side of (D.1) is

$$\begin{aligned}\int_0^2 p(t)\,q(t)\,dt &= \int_0^1 p(t)\,q(t)\,dt + \int_1^2 p(t)\,q(t)\,dt \\ &= \int_0^1 p(t)\,q(t)\,dt - \int_1^0 p(2-t')\,q(2-t')\,dt' \qquad (t'=2-t) \\ &= \int_0^1 p(t)\,q(t)\,dt + \int_0^1 p(t')\,q(t')\,dt' \\ &= 2\int_0^1 p(t)\,q(t)\,dt. \qquad \text{(D.3)}\end{aligned}$$

To perform a similar decomposition for the right-hand side of (D.1), notice that

$$\begin{aligned}\alpha_n &= \int_0^2 p(t)\,\phi_n^{2,j}(t)\,dt \\ &= \int_0^2 p(t)\sum_{\ell\in\mathbb{Z}} \phi^j_{n-2N\ell}(t)\,dt \\ &= \int_0^2 p(t)\sum_{\ell\in\mathbb{Z}} \sqrt{N}\phi(Nt-n+2N\ell)\,dt \\ &= -\int_2^0 p(2-t')\sum_{\ell\in\mathbb{Z}} \sqrt{N}\phi(N(2-t')-n+2N\ell)\,dt' \qquad (t'=2-t) \\ &= \int_0^2 p(t')\sum_{\ell\in\mathbb{Z}} \sqrt{N}\phi(Nt'-2N+n-2N\ell)\,dt' \qquad (p\ \&\ \phi \text{ symmetric}) \\ &= \int_0^2 p(t')\,\phi_{2N-n}^{2,j}(t')\,dt' \\ &= \alpha_{2N-n},\end{aligned}$$

and this is true for all $n = 0, 1, \ldots, 2N$. Likewise,

$$\beta_n = \beta_{2N-n} \qquad \forall\ n = 0, 1, 2, \ldots, 2N-1.$$

So

$$\sum_{n=0}^{2N-1} \alpha_n\beta_n = 2\sum_{n=1}^{N-1} \alpha_n\beta_n. \qquad \text{(D.4)}$$

Equations (D.3) and (D.4) substituted back into (D.1) yield

$$\int_0^1 p(t)\,q(t)\,dt = \sum_{n=0}^{N-1} \alpha_n\beta_n. \qquad \square$$

Appendix E. Proof of Proposition 2.5.

Proof. We first show that

$$\sum_{m\ge m_0>0} \frac{1}{m^\gamma} \le \frac{m_0^{1-\gamma}}{1-2^{1-\gamma}} \qquad \forall\ \gamma>1,\ m\in\mathbb{N}. \qquad \text{(E.1)}$$

Let p_0 and p_1 be positive integers such that

$$2^{p_0-1} \le m_0 < 2^{p_0}, \qquad 2^{p_1-1} \le N_m < 2^{p_1}; \tag{E.2}$$

then

$$\begin{aligned}
\sum_{m=m_0}^{N_m} \frac{1}{m^\gamma} &\le \sum_{m=2^{p_0-1}}^{2^{p_1}-1} \frac{1}{m^\gamma} \\
&= \left[\frac{1}{(2^{p_0-1})^\gamma} + \cdots + \frac{1}{(2^{p_0}-1)^\gamma}\right] + \cdots + \left[\frac{1}{(2^{p_1-1})^\gamma} + \cdots + \frac{1}{(2^{p_1}-1)^\gamma}\right] \\
&\le 2^{p_0-1} \cdot \frac{1}{2^{(p_0-1)\gamma}} + \cdots + 2^{p_1-1} \cdot \frac{1}{2^{(p_1-1)\gamma}} \\
&= \frac{1}{2^{(p_0-1)(\gamma-1)}} + \cdots + \frac{1}{2^{(p_1-1)(\gamma-1)}} + \cdots \\
&= \frac{1}{1-2^{1-\gamma}} \cdot \left[2^{-(p_0-1)(\gamma-1)} \left(1 - 2^{-(p_1-p_0)(\gamma-1)}\right)\right].
\end{aligned}$$

This summation converges when $\gamma > 1$. Taking the limit as $N_m \to \infty$, we get

$$\sum_{m \ge m_0 > 0} \frac{1}{m^\gamma} \le \frac{2^{-(p_0-1)(\gamma-1)}}{1-2^{1-\gamma}}. \tag{E.3}$$

As $2^{p_0-1} \le m_0$, this is exactly what we want.

We are now ready to prove Proposition 2.5. By assumption, the localized function $\phi(t)$ attains its maximum at $t = 0$. Let this maximum be C; then

$$\phi(0) = C.$$

We also know, for $n \in \mathbb{Z}$, that

$$|\phi(n)| \le \frac{C}{(1+n^2)^{\gamma/2}}, \qquad \gamma > 1. \tag{E.4}$$

This "localization" can be used to estimate bounds for the entries of matrix $\underline{\underline{\mathbf{H}}}$:

$$\begin{aligned}
H_{nm} &= \sum_{l\in\mathbb{Z}} \phi^{j_2}_{m+lN}(t_n) \\
&= \sum_{l\in\mathbb{Z}} \sqrt{N}\phi\left(N\frac{n}{N} - m - lN\right) \\
&= \sum_{l\in\mathbb{Z}} \sqrt{N}\phi(n-m-lN), \qquad n,m = 0,1,\ldots,N-1.
\end{aligned} \tag{E.5}$$

Obviously, $\underline{\underline{\mathbf{H}}} = \{H_{nm}\}_{N\times N}$ forms a circulant matrix (see, e.g., Davis (1979)). On the diagonal,

$$H_{nn} = \sqrt{N}\sum_{l\in\mathbb{Z}} \phi(-lN) = C\sqrt{N} + r_{nn}, \tag{E.6}$$

where

$$
\begin{aligned}
|r_{nn}| &= \sqrt{N} \sum_{l\in\mathbb{Z},\, |l|\neq 0} |\phi(lN)| \\
&\leq \sqrt{N} \cdot \sum_{|l|\neq 0} \frac{C}{(1+|lN|^2)^{\gamma/2}} \\
&< C\sqrt{N} \cdot \sum_{|l|>0} \frac{1}{|lN|^{\gamma}} \\
&= 2C\sqrt{N} \frac{1}{N^{\gamma}} \sum_{l=1}^{\infty} \frac{1}{l^{\gamma}} \\
&\leq \frac{2C}{1-2^{1-\gamma}} \sqrt{N} \frac{1}{N^{\gamma}} \qquad \text{by (E.1).}
\end{aligned} \tag{E.7}
$$

The off-diagonal elements, on the other hand, are equal to

$$
H_{nm} = \sum_{l\in\mathbb{Z}} \sqrt{N}\phi(n-m-lN), \qquad \text{where } 0<|n-m|<N. \tag{E.8}
$$

We now examine $\sum_{m\neq n} |H_{nm}|$. As $\underline{\underline{\mathbf{H}}}$ is circulant, this sum is the same for all $0 \leq n < N$. It suffices to consider the case $n=0$: $\sum_{m=1}^{N-1} |H_{0m}|$. For convenience, split the summation into two parts:

$$
\begin{aligned}
|H_{0m}| &\leq \sqrt{N} \sum_{l\in\mathbb{Z}} |\phi(m+lN)| \\
&\leq \sqrt{N} \sum_{l\in\mathbb{Z}} \frac{C}{[1+(m+lN)^2]^{\gamma/2}} \\
&= \underbrace{C\sqrt{N} \left\{ \frac{1}{[1+m^2]^{\gamma/2}} + \frac{1}{[1+(m-N)^2]^{\gamma/2}} \right\}}_{(\mathrm{PRIN})_{\mathrm{m}}} \\
&\quad + \ (\mathrm{OTHER})_{\mathrm{m}},
\end{aligned} \tag{E.9}
$$

where

$$
(\mathrm{OTHER})_{\mathrm{m}} = C\sqrt{N} \underbrace{\frac{1}{[1+(m+N)^2]^{\gamma/2}}}_{\mathrm{(I)}} + C\sqrt{N} \underbrace{\sum_{|l|\geq 2} \frac{1}{[1+(m+lN)^2]^{\gamma/2}}}_{\mathrm{(II)}}. \tag{E.10}
$$

As $m>0$, we have

$$
\mathrm{(I)} < \frac{1}{N^{\gamma}}
$$

and

$$
\begin{aligned}
\mathrm{(II)} &= \sum_{l\geq 2} \frac{1}{[1+(m+lN)^2]^{\gamma/2}} + \sum_{l\geq 2} \frac{1}{[1+(m+lN)^2]^{\gamma/2}} \\
&< \sum_{l\geq 2} \frac{1}{(N-1-lN)^{\gamma}} + \sum_{l\geq 2} \frac{1}{[-(N-1)+lN]^{\gamma}}
\end{aligned} \tag{E.11}
$$

$$= 2\sum_{l\geq 2}\frac{1}{[(l-1)N+1]^\gamma} < 2\sum_{l'\geq 1}\frac{1}{(l'N)^\gamma}$$
$$= \frac{2}{N^\gamma}\sum_{l'\geq 1}\frac{1}{l'^\gamma} \leq \frac{2}{1-2^{1-\gamma}}\cdot\frac{1}{N^\gamma} \qquad \text{for } \gamma > 1$$

by what we showed in the beginning. Hence, in (E.9),

$$(\text{OTHER})_\text{m} < C\sqrt{N}\left[1+\frac{2}{1-2^{1-\gamma}}\right]\frac{1}{N^\gamma}. \tag{E.12}$$

So

$$\sum_{m=1}^{N-1}(\text{OTHER})_\text{m} < C\sqrt{N}\left[1+\frac{2}{1-2^{1-\gamma}}\right]\frac{1}{N^\gamma}\cdot(N-1)$$
$$< C\sqrt{N}\left[1+\frac{2}{1-2^{1-\gamma}}\right]\frac{1}{N^{\gamma-1}}. \tag{E.13}$$

The other part, $(\text{PRIN})_\text{m}$, sums to

$$\sum_{m=1}^{N-1}(\text{PRIN})_\text{m} = \sum_{m=1}^{N-1} C\sqrt{N}\left\{\frac{1}{(1+m^2)^{\gamma/2}}+\frac{1}{[1+(m-N)^2]^{\gamma/2}}\right\}$$
$$= 2C\sqrt{N}\sum_{m=1}^{N-1}\frac{1}{(1+m^2)^{\gamma/2}}$$
$$= 2C\sqrt{N}\left[\frac{1}{2^{\gamma/2}}+\sum_{m=2}^{N-1}\frac{1}{m^\gamma}\right]$$
$$< 2C\sqrt{N}\left[\frac{1}{2^{\gamma/2}}+\frac{2^{1-\gamma}}{1-2^{1-\gamma}}\right] \tag{E.14}$$

by (E.1). Put $(\text{PRIN})_\text{m}$ and $(\text{OTHER})_\text{m}$ together,

$$\sum_{m=1}^{N-1}|H_{0m}| < C\sqrt{N}\left[2\left(\frac{1}{2^{\gamma/2}}+\frac{1}{2^{\gamma-1}(1-2^{1-\gamma})}\right)+\frac{1}{N^{\gamma-1}}\left(1+\frac{2}{1-2^{1-\gamma}}\right)\right], \tag{E.15}$$

and compare $\sum_{m=1}^{N-1}|H_{0m}|$ to $|H_{00}|$:

$$|H_{00}| - \sum_{m=1}^{N-1}|H_{0m}| = \left|C\sqrt{N}+r_{00}\right| - \sum_{m=1}^{N-1}|H_{0m}|$$
$$> C\sqrt{N} - |r_{00}| - \sum_{m=1}^{N-1}|H_{0m}|.$$

Substitution of (E.7) and (E.15) for $|r_{00}|$ and $\sum_{m=1}^{N-1}|H_{0m}|$ gives

(E.16)

$$
\begin{aligned}
&|H_{00}| - \sum_{m=1}^{N-1} |H_{0m}| \\
&> C\sqrt{N}\left\{1 - \frac{2}{1-2^{1-\gamma}}\frac{1}{N^{\gamma}} - \left(\frac{2}{2^{\gamma/2}} + \frac{2}{2^{\gamma-1}(1-2^{1-\gamma})}\right) - \frac{1}{N^{\gamma-1}}\left(1 + \frac{2}{1-2^{1-\gamma}}\right)\right\} \\
&> C\sqrt{N}\left\{1 - \frac{2}{2^{\gamma/2}} - \frac{2}{2^{\gamma-1}(1-2^{1-\gamma})} - \left(1 + \frac{4}{1-2^{1-\gamma}}\right)\frac{1}{N^{\gamma-1}}\right\},
\end{aligned}
$$

which is greater than zero given that (2.12) is satisfied. This is to say, $\underline{\underline{\mathbf{H}}}$ is strictly diagonally dominant under condition (2.12). Nonsingularity follows immediately from this diagonal dominance (see Ortega (1987)). □

Acknowledgments. This work was motivated by the imperative need of dynamical diagnostics as oceanographic data keeps building up. XSL thanks Allan R. Robinson and the Harvard Ocean Group, including Carlos Lozano, Wayne Leslie, Pierre Lermusiaux, Patrick Haley, Patricia Moreno, Oleg Logoutov, Gioia Sweetland, Peggy Zaldivar, and Marsha Glass for their generous support and the important scientific interactions with them. Particularly, it was Carlos who first suggested the terminology *scale window*, and Pierre's insights in wavelets were very helpful. During the study, XSL also gained generous help from Gilbert Strang on wavelet analysis, and Kenneth Brink and Arthur Miller on geophysical fluid dynamics.

REFERENCES

I. Daubechies (1992), *Ten Lectures on Wavelets*, CBMS-NSF Regional Conf. Ser. in Appl. Math. 61, SIAM, Philadelphia.

P. J. Davis (1979), *Circulant Matrices*, John Wiley and Sons, New York.

M. Gad-el-Hak, R. F. Blackwelder, and J. J. Riley (1981), *On the growth of turbulent regions in laminar boundary layers*, J. Fluid Mech., 110, pp. 73–95.

E. Hernández and G. Weiss (1996), *A First Course on Wavelets*, CRC Press, Boca Raton, FL.

M. Holschneider (1995), *Wavelets: An Analysis Tool*, Clarendon Press, Oxford, UK.

N. E. Huang, Z. Shen, and S. R. Long (1999), *A new view of nonlinear water waves: The Hilbert spectrum*, in Annual Review of Fluid Mechanics, Annu. Rev. Fluid Mech. 31, Annual Reviews, Palo Alto, CA, pp. 417–457.

X. S. Liang (2002), *Wavelet-Based Multiscale Window Transform and Energy and Vorticity Analysis*, Ph.D. thesis, Harvard University, Cambridge, MA.

X. S. Liang and A. R. Robinson (2004), *A study of the Iceland-Faeroe frontal variability using the multiscale energy and vorticity analysis*, J. Phys. Oceanogr., 24, pp. 2571–2591.

X. S. Liang and A. R. Robinson (2005), *Localized multiscale energy and vorticity analysis. I. Fundamentals*, Dyn. Atmos. Oceans, 38, pp. 195–230.

A. K. Louis, P. Maaß, and A. Rieder (1997), *Wavelets: Theory and Applications*, John Wiley and Sons, Chichester, UK.

Y. Meyer (1992), *Wavelets and Operators*, Cambridge University Press, Cambridge, UK.

J. M. Ortega (1987), *Matrix Theory*, Plenum Press, New York.

J. Pedlosky (1979), *Geophysical Fluid Dynamics*, Springer-Verlag, New York.

A. R. Robinson, H. G. Arango, A. J. Miller, A. Warn-Varnas, P.-M. Poulain, and W. G. Leslie (1996), *Real-time operational forecasting on shipboard of the Iceland-Faeroe frontal variability*, Bull. Amer. Meteorol. Soc., 77, pp. 243–259.

W. H. Press, S. A. Teukolsky, and W. T. Vetterling (1993), *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed., Cambridge University Press, Cambridge, UK.

G. Strang and T. Nguyen (1997), *Wavelets and Filter Banks*, Wellesley-Cambridge Press, Wellesley, MA.

P. Wojtaszczyk (1997), *A Mathematical Introduction to Wavelets*, Cambridge University Press, Cambridge, UK.