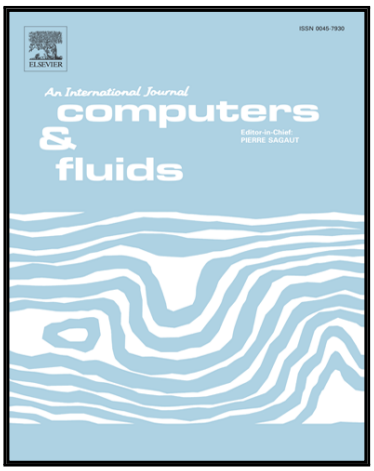
Accepted Manuscript



Survey and Analysis of Multiresolution Methods for Turbulence Data

Jesus Pulido, Daniel Livescu, Jonathan Woodring, James Ahrens, Bernd Hamann

PII: S0045-7930(15)00362-X
DOI: 10.1016/j.compfluid.2015.11.001
Reference: CAF 3034

To appear in: *Computers and Fluids*

Received date: 12 February 2015
Revised date: 12 September 2015
Accepted date: 3 November 2015

Please cite this article as: Jesus Pulido, Daniel Livescu, Jonathan Woodring, James Ahrens, Bernd Hamann, Survey and Analysis of Multiresolution Methods for Turbulence Data, *Computers and Fluids* (2015), doi: 10.1016/j.compfluid.2015.11.001

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Highlights

- We survey various multi-resolution geometric methods for turbulence representation.
- The flow contains novel aspects, not examined before, like variable density.
- The ability to preserve flow properties with reduced set of coefficients is examined.
- An extensive list of recommendations for the best decomposition algorithms is given.
- The cubic B-spline wavelets provide the best tradeoffs in efficiency versus accuracy.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Survey and Analysis of Multiresolution Methods for Turbulence Data

Jesus Pulido[a,b,*], Daniel Livescu[b], Jonathan Woodring[b], James Ahrens[b], Bernd Hamann[a]

[a] *Institute for Data Analysis and Visualization, Department of Computer Science, University of California, Davis, CA 95616, USA*

[b] *Los Alamos National Laboratory, Los Alamos, NM 87544, USA*

Abstract

This paper compares the effectiveness of various multi-resolution geometric representation methods, such as B-spline, Daubechies, Coiflet and Dual-tree wavelets, curvelets and surfacelets, to capture the structure of fully developed turbulence using a truncated set of coefficients. The turbulence dataset is obtained from a Direct Numerical Simulation of buoyancy driven turbulence on a 512^3 mesh size, with an Atwood number, $A = 0.05$, and turbulent Reynolds number, $Re_t = 1800$, and the methods are tested against quantities pertaining to both velocities and active scalar (density) fields and their derivatives, spectra, and the properties of constant density surfaces. The comparisons between the algorithms are given in terms of performance, accuracy, and compression properties. The results should provide useful information for multi-resolution analysis of turbulence, coherent feature extraction, compression for large datasets handling, as well as simulations algorithms based on multi-resolution methods. The final section provides recommendations for best decomposition algorithms based on several metrics related to computational efficiency and preservation of turbulence properties using a reduced set of coefficients.

Keywords: Turbulence, Wavelet, Curvelet, Surfacelet, B-spline Wavelet

*Corresponding author

Email addresses: `jpulido@ucdavis.edu` (Jesus Pulido), `livescu@lanl.gov` (Daniel Livescu), `woodring@lanl.gov` (Jonathan Woodring), `ahrens@lanl.gov` (James Ahrens), `hamann@cs.ucdavis.edu` (Bernd Hamann)

ACCEPTED MANUSCRIPT

1. Introduction

Most datasets encountered in physical applications, similar to most natural images, present lower dimensional structures whose detection, extraction, and characterization are active areas of research. The search for more efficient algorithms to detect and manipulate such structures has led to the development of a multitude of multi-resolution geometric representations, such as curvelet and surfacelet transforms. The curvelet [1] and surfacelet [2] transforms perform spatial partitioning in Fourier space at multiple resolutions by creating bands using discrete frequency tiling that store localized directional coefficients.

One area which has seen significant interest in the application of such methods is fluid turbulence. While turbulence is a strongly multi-scale phenomenon with a large range of dynamically relevant spatio-temporal scales, coherent structures are almost always present, due to initial or boundary conditions, injection mechanisms, or arising from internal dynamics. The characterization of these structures, which interact nonlinearly as they are advected by the background flow and significantly alter the local topology, is one of the fundamental open questions in the study of turbulence. One of the earliest applications of compression algorithms to turbulence is done in Ref. [3]. The focus was on comparing the coherent vortex simulation (CVS) decomposition based on a orthogonal wavelet basis with the Proper Orthogonal Decomposition or Fourier filtering, as applied to a forced homogeneous isotropic turbulence Direct Numerical Simulations (DNS) dataset. It is shown that CVS filtering, which is local in both physical and spectral spaces, can separate the coherent vortex tubes from the incoherent background flow. The latter is structureless, has an equipartition energy spectrum, a Gaussian velocity probability distribution function (PDF), and an exponential vorticity PDF. On the other hand, the Fourier basis does not extract the coherent vortex tubes cleanly and leaves organized structures in the residual high wavenumber modes whose PDFs are stretched exponentials for both the velocity and vorticity.

More recently, curvelets have been briefly evaluated by Ma et al. [4] in comparison to the classical wavelet transform. In their work, multi-scale geometric analysis is systematically applied to turbulent flows in two and three dimensions using curvelets. The analysis is based on the constrained minimization of a total variation functional representing the difference between the

ACCEPTED MANUSCRIPT

data and its representation in the curvelet space. Constrained multi-scale minimization results in a minimum loss of the geometric flow features and the extraction of the coherent structures with their edges and geometry properly preserved, which is significant for turbulence modeling. The results of this work show that curvelets are very effective in edge and geometry preservation in turbulence data when compared to traditional wavelets under a specific series of tests and wavelet coefficient numbers. One goal of the present paper is to expand those tests to the full range of coefficients for reconstruction of turbulence data as well as include tests for quantities dependent on an active scalar field.

These methods have been primarily applied to the quantities directly related to the velocity fields although, recently, the curvelet transform has been used to examine the multi-scale structure of scalar fields, such as mass concentration [5]. These fields are, in general, rougher than the advecting velocity and can present unmixed patches in many practical applications (e.g. non-premixed combustion). Thus, representation methods which are designed to capture surfaces or edges, such as surfacelets or curvelets, appear naturally more suited to capture these fields. For example, Ref. [5] introduces a curvelet-based multi-scale methodology for the study of the non-local geometry of eddy structures in turbulence data. The dataset is from a 512^3 DNS of passive scalar mixing in isotropic turbulence and the curvelet transform is used to extract, characterize and classify structures pertaining to the passive scalar. The classification is based on differential-geometry properties, such as shape index, curvedness, and stretching parameter, which define the geometrical signature of the surfaces of constant scalar value. These properties are discussed with respect to their relation to the dynamical behavior of passive scalar stirring and mixing. Another goal of the present paper is to compare the curvelet transform with other representation methods for their ability to preserve these properties with a reduced set of coefficients.

Accurate simulations of turbulent flows require solving all the dynamically relevant scales of motions. This technique, called DNS (see above), has been successfully applied to a variety of simple flows, however most practical flows would require mesh sizes outside the range of the most powerful supercomputers for the foreseeable future. The resolution requirements can be improved, especially in problems with localized features, by employing an adaptive mesh strategy. However, such approaches often introduce directional bias and use lower order discretization methods, which

ACCEPTED MANUSCRIPT

decreases the accuracy. Adaptive mesh strategies based on wavelet decompositions have been proposed with explicit error control and higher order discretization schemes [6, 7, 8]. While these methods extend the range of DNS applicability, accurately solving all the flow scales still imposes severe limitations on the flows which can be simulated.

One of the avenues being explored for simulating, with feasible meshes, flows with large range of scales, as encountered in most practical applications, explores coherent/incoherent decompositions allowed by multi-resolution geometric representations [3, 4, 9, 10]. This approach relies on the ability of such methods to represent the coherent part of the flow with a significantly reduced set of coefficients (e.g. $1-5\%$ of the coefficients to represent the whole flow) and model the incoherent part using simplifying models (e.g. assume Gaussian statistics). For such applications, the accuracy and computational efficiency of the algorithms are both important. So far, only the curvelet, Dual-tree, and orthogonal wavelet transforms have been used in this context and no comprehensive comparison between these transforms has been made. Classical large eddy simulation (LES) approaches for computing turbulence represent the flow on a coarse mesh in either real or spectral spaces and model the sub-grid contributions. While finding an optimal function set basis to represent turbulence remains an outstanding open question, the representation methods discussed in this paper may offer a better framework for modeling the sub-grid terms in an LES type approach than spectral or physical space based filters, due to their localization in both spectral and real spaces. Here, we rely on this locality to denote the coherent/incoherent decomposition as applied directly to the primary variables, as opposed to CVS-type decompositions. This is along the lines of the SCALES approach [9] and offers easy generalizations to complex flows and direct connection to LES-type approaches.

In addition, there is a significant cost associated with the storage of the data generated by turbulence simulations. Efficient lossy algorithms can take advantage of the coherent/incoherent decompositions of the flow field and significantly reduce the archival requirements. Data retrieval can be optimized by extracting only the coherent structures in the data for faster data visualization and analysis at multiple levels of resolution. By reducing the retrieval and transmission cost, projects such as the Johns Hopkins Turbulence Database (JHTDB) can be improved by reducing the amount of data processed and transmitted to a client [11]. By only sending structures at a

ACCEPTED MANUSCRIPT

resolution relevant for analysis, the reduced cost can allow for real-time remote data visualization and analysis of large datasets.

The focus of this paper is the comparison of new and existing methods used in analysis (feature identification, extraction, and analysis) and simulations (based on coherent/incoherent decompositions) of turbulence. These methods include second-generation wavelets such as Haar, biorthogonal B-spline, Daubechies, Coiflets, Dual-tree, and newer methods such as curvelets and surfacelets. In order to make the comparisons as comprehensive as possible, a flow has been selected in which the turbulence is accompanied by mixing between initially segregated different density materials (see description below) which are subjected to a constant acceleration. The large scales of the flow are anisotropic and the interfaces between the two materials become highly corrugated. The methods considered are compared in their ability to capture the structures of both velocity and density fields. It is hoped that this analysis will help both the simulations of turbulent flows using multi-resolution geometrical representations as well as further the study of turbulence physics using such methods.

1.1. Direct numerical simulation dataset

The dataset used in this paper is from a DNS of homogeneous buoyancy driven turbulence on a 512^3 periodic grid. The simulation used the variable-density version of the petascale CFDNS code [12] to solve the incompressible Navier-Stokes equations for two miscible fluids with different densities, in a triply periodic domain. These equations are obtained from the fully compressible Navier-Stokes equations with two species with different molar masses in the limit $c \to \infty$ (c is the speed of sound) such that the individual densities of the two fluids remain constant [13, 14, 15].

The two fluids are initialized as random blobs, consistent with the homogeneity assumption. The flow starts from rest, with only a small amount of dilatational velocity necessary to satisfy the divergence condition and turbulence is generated as the two fluids start moving in opposite directions due to differential buoyancy forces. However, as the fluids become molecularly mixed, the buoyancy forces decrease and at some point the turbulence starts decaying. For comparison between the different compression algorithms, density and velocity fields are used at the time when the turbulent kinetic energy is maximum. At this time, the turbulent Reynolds number is

ACCEPTED MANUSCRIPT

$Re_t = 1800$ and the turbulence is fully developed. The rest of the non-dimensional parameters characterizing the flow are Atwood number, $A = 0.05$, Schmidt number, $Sc = 1$ and Froude number, $Fr = 1$. A similar dataset [16], on a 1024^3 mesh, covering the whole range of turbulence evolution, from buoyancy driven increase in turbulent kinetic energy to buoyancy mediated turbulence decay, has been recently added to the Johns Hopkins Turbulence databases [11].

The rest of the paper is organized as follows. In section 2, a background is given of the geometric representation methods considered in this paper. Section 3 summarizes the software used, thresholding techniques, and properties for all of the different methods. The testing methodologies are described and the results are quantified in section 4. Finally, section 5 presents the conclusions with the recommendations of the best schemes suited for the representation and preservation of a variety of quantities relevant to fully developed turbulence in the presence of an active scalar field.

2. Background and methods

Wavelets are the generalization of the Fourier transform by using bases that represent both location and spatial frequency [17]. A more fundamental background on wavelets can be found in Appendix A. This section gives a brief overview of each of the representation methods used for comparisons. All methods considered in this paper are the discrete versions of their respective continuous signals. Wavelets used in this paper are second-generation wavelets in implementation [18].

2.1. Orthogonal wavelets and B-spline wavelets

The orthogonal Haar wavelet is the simplest transform and is best known for its top hat, piece wise signal and simplistic representation. These simple wavelets can decompose a discrete signal into two half signals represented by a 0 or 1 per each step as seen in Fig. A.21a in the Appendix. Due to their simplicity, they are the fastest to compute. Farge [19] performed an initial analysis on the use of wavelets, including Haar wavelets, to characterize coherent and incoherent turbulent flow parts.

The orthogonal family of Daubechies wavelets are similar in construction to Haar wavelets but requiring vanishing higher order moments in the mother signal. Their construction and signals are

ACCEPTED MANUSCRIPT

described in [17] and shown for completeness in Fig. A.21 in the Appendix. As the Daubechies family increases in order, more distortions are produced in the original mother signal. For our analysis, the lower order 3rd , 5th, and 7th families of Daubechies wavelets are considered. As one goes to even higher order families of wavelets, the wider signals may cause overlapping in the analysis and introduce a loss in quality as it is shown later. Goldstein et al. [9, 10] used the 5th family of Daubechies wavelets for adaptive eddy simulations.

The orthogonal Coiflet family of wavelets were first proposed by Beylkin et al. [20], then implemented by Daubechies [21] and Tian et al. [22]. As with Daubechies wavelets, Coiflets require vanishing moments in the original reconstruction signal. Unlike Daubechies wavelets, Coiflets focus on vanishing moments in their scaling function in order to further improve the convergence of the original signal. For this analysis, the first three orders are used: Coiflet-6, Coiflet-12 and Coiflet-16. Their respective signals can be seen in Fig. A.21 in the Appendix. Deriaz et al. [23] and Roussel et al. [24] have used Coiflet-12 for coherent vortex extraction in homogeneous turbulence. Farge et al. [3] used Coiflets for CVS decomposition to separate coherent vortex tubes from the incoherent background flow in turbulence.

Biorthogonal B-spline wavelets are the natural extension to the Haar wavelet; similar in usage and implementation but differ in basis functions [25]. B-spline wavelets contain basis functions representing sinusoidal signals of varying magnitudes. The construction of both the decomposition and synthesis filters depend on two control parameters that select the family order and the number of discrete interpolation points to interpret a continuous wave. At the lowest order, constant B-spline wavelets share the same basis function as the Haar wavelets and will therefore refer to this family as Haar wavelets for the remainder of this paper. The first six families of B-spline wavelets are considered for this paper: constant (Haar, 1st degree), linear (2nd degree), quadratic (3rd degree), cubic (4th degree), quartic (5th degree), and quintic (6th degree). Family characteristics are further described in Appendix A.1.

Biorthogonal wavelets in their lowest order have had some limited use in scientific computing. The lack of orthogonality in their higher order basis functions and associated computation cost has hindered their wider adoption so far. Adding to the slightly higher compute cost, B-spline wavelets are also not as easily parallelized compared to orthogonal wavelets. Parallelization is possible by

ACCEPTED MANUSCRIPT

using the second-generation construction of wavelets based on a lifting scheme [26]. Roussel et al. [24] used biorthogonal (constant) Harten wavelets for the extraction of coherent vortices and found that the biorthogonal nature of the wavelets may cause problems in modeling background noise after extraction. Their results showed an improvement by using orthogonal Coiflet-12 over Harten wavelets. These results are expected due to Harten signal being fundamentally similar to the Haar signal as part of the constant family of B-spline wavelets, as seen in Fig. A.22b in the Appendix. Thus, an increase in control points does not generally produce improvements in reconstruction in contrast to increasing the family order. Nevertheless, while one expects a much better performance for higher-order B-spline wavelets, these methods have yet to be tested extensively for turbulent flows.

For simple applications, orthogonal wavelets have overshadowed B-spline wavelets; however, higher order transforms such as B-splines may perform better in applications requiring the preservation of structures. For example, Bertram et al. [27] have shown that B-splines do very well in representing smooth surfaces and curved structures with a minimal set of coefficients. For the purpose of this paper, the lack of orthogonality is less important than the better preservation of various turbulence quantities.

2.2. Dual-tree Wavelet Transform

The Dual-tree complex wavelet transform (DTWT) is an enhancement of the discrete wavelet transform [28]. The DTWT implements shift invariance and directionality in multiple dimensions. The DTWT performs two discrete wavelet transform operations to construct a tree-like structure. Two paths are branched off, each representing the real and imaginary components of the transform. The real component is utilized to represent the majority of the energy in a dataset where the imaginary component captures small-scale details. The DTWT transform addresses issues present with real wavelets, those being oscillations, shift variance, aliasing, and lack of directionality. The redundancy of having two branches provides additional information for analysis, approximate shift-invariance, and a perfect signal reconstruction unlike the discrete wavelet transform. The DTWT contains filters designed to represent many properties, such as approximate half-sample delays, orthogonal or bi-orthogonal signals, finite support, vanishing moments/good

ACCEPTED MANUSCRIPT

stop-band, and linear-phase filters [29]. The filters considered in this paper include the Farras filter for the initial scale, and Kingsman's Q-shift filter for the subsequent scales. The DTWT has been extensively used by Nguyen et al. [30, 31] for signal filtering on Galerkin methods.

2.3. Curvelets

The curvelet transform is a filter-based wave-signal decomposition algorithm that is designed to represent signals by their edges [1]. Curvelets are a non-adaptive method for multi-scale signal representation. Curvelets are derived from ridgelets, which are linear features that represent different lengths with respect to scale in different positions and as many orientations as can be defined for a dataset. Fig. 1 demonstrates the frequency tiling for the 2D and 3D discrete transform, the latter being used for this paper. The frequency tiling is defined in Fourier space, where certain ranges of Fourier coefficients are defined as the directional components of the curvelet transform. The curvelet transform extends the bi-directionality of wavelets to multiple directions where basis functions are localized to each direction. Within each direction, the degree of localization varies per scale allowing multi-resolution extraction of signals.

Due to the large number of orientations, positions and scales available, the computation of this transform is expensive compared to any of the simpler wavelets. The result of a decomposition produces a large number of complex coefficients where their amplitudes refer to the underlying structure of the original dataset.

So far, curvelets have not been proven to have an orthogonal basis function. The discrete implementation aims to decompose a dataset based on the described band scheme and sub-components. For 3D datasets, this scheme introduces modest levels of data redundancy in order to properly encode directional/angular data. One of the major incentives to consider curvelets is their ability to express orientation versus its predecessors. Although curvelets have proven to be superior in the field of 2D image processing, similar areas have not been fully explored in 3D turbulence. Directional structures embedded in the velocity gradient tensor or related to constant active scalar surfaces present in 3D turbulent flows may lend themselves in structure and orientation to the curvelet transform. Opposite to wavelets, curvelets take a bottom-up approach in extracting coefficients. The largest, coarsest coefficients are first extracted in the first bands of the decomposition

ACCEPTED MANUSCRIPT

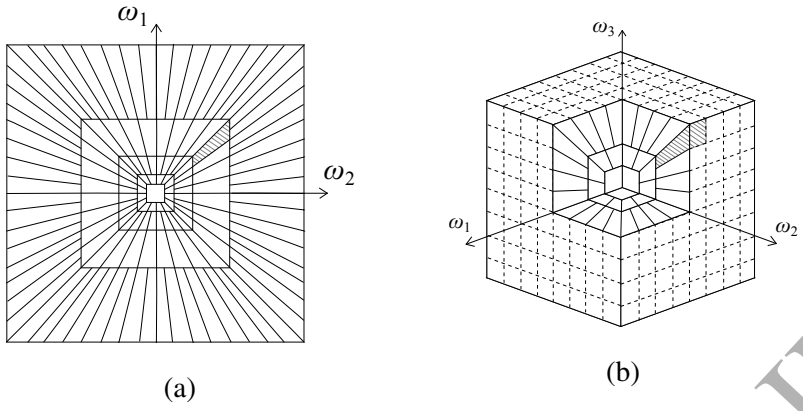


Figure 1: **Discrete Curvelet**. The continuous curvelet transform is discretized in Fourier space for each dimension ω. Directional wedges are darkened in both figures where (a) 2-D frequency tiling is performed along with (b) 3-D frequency tiling. Figure is reproduced from [1].

and the smallest, finest coefficients are in the highest bands. Curvelets have been used for geometric analysis of turbulent data by Ma et al. [4] and Bermejo-Moreno and Pullin [5].

2.4. Surfacelets

The surfacelet transform uses a combination of directional filter banks and contourlets in order to represent signal singularities that lie on smooth surfaces. In order to analyze multi-dimensional signals for multiple scales and directions, the directional filter bank (DFB) was proposed by Bamberger and Smith for 2-D signals [32]. The DFB is a tree-structured decomposition that creates 2^k sub-bands with directional partitioning as in Fig. 2(a). The value k is defined with respect to the number of directions to be extracted. Do and Vetterli later constructed contourlets by combining the DFB with a Laplacian pyramid structure [33].

Previously restricted to 2D, these directional filter banks were extended into higher dimensions creating the N-dimensional directional filter bank (NDFB) by Lu and Do [2]. The surfacelet transform combines the extension of contourlets to 3-D space and the NDFB in order to represent surface-like singularities in multidimensional data. The NDFB uses a frequency partitioning scheme that resembles rectangular-based pyramids radiating out from the origin in different directions and multiple tiles as seen in Fig. 2(b). Similar to curvelets, the surfacelet partitioning is defined in Fourier space. Each direction is represented by a number as seen in Fig. 2(a) within

ACCEPTED MANUSCRIPT

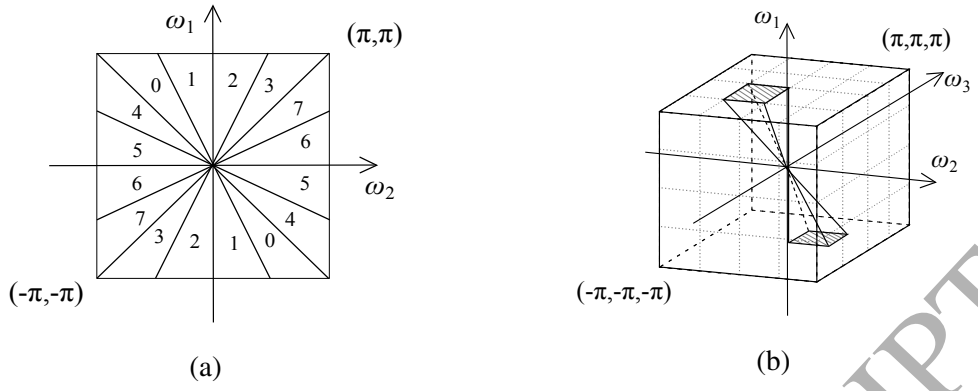


Figure 2: **Discrete Surfacelet**. Surfacelet frequency tiling of the N-dimensional filter bank (NDFB) in Fourier space bounded by π on each dimension ω. (a) Discrete 2D tiling for each numbered angular direction. (b) Discrete 3D tiling highlighting a single angular direction. Figure is reproduced from [2].

Fourier space. The inclusion of higher dimension filter banks allows surfacelets to have the following distinct properties: directional decomposition and construction, angular resolution, perfect reconstruction and small redundancy [2].

One of the features of surfacelets is the frequency partitioned directional filter bank with multiple levels in a tree-like structure. As opposed to curvelets, a single directional signal is spread out across its domain requiring the extraction of a lesser number of coefficients. The introduction of this filter bank allows aliasing to occur for each band in the frequency domain as opposed to being restricted to alias-free bands in curvelets. Due to the alias-free restriction, curvelets may require much more bands and coefficients to capture certain frequencies. By utilizing the tree-structured NDFB, aliasing can be removed in surfacelets by combining multiple levels and overall producing a less redundant extraction of coefficients compared to curvelets. The surfacelet transform has not been previously tested in the context of turbulence. The ability of identifying and preserving rough surfaces could make this transform well suited for representing material fields in turbulent flow datasets.

2.4.1. Orthogonality

Orthogonality is an attractive feature in mathematical transforms since it infers energy preservation. The wavelet families that are fully orthogonal in this analysis are the Haar, Daubechies and Coiflet families. The B-spline wavelets lose orthogonality due to their signal and become

ACCEPTED MANUSCRIPT

biorthogonal. Dual-tree wavelets contain two biorthogonal filter sets with slightly different frequency responses. When combined, the two biorthogonal filters can become near orthogonal through the use of long filters leading to a higher computation penalty. Curvelets contain an orthogonal wavelet component during their directional, finest-scale decomposition but overall are not orthogonal. Candes et al. [1] claim that curvelets behave similar to an orthogonal decomposition but are unable to prove it. Surfacelets, like curvelets, are only orthogonal at their lowest directional component but not in their entirety. The issue of needing orthogonal wavelets is not a priority in our analysis since the preservation of turbulence structures is prioritized over full preservation of energy. In Ref. [24], it is stated that a drawback of using non-orthogonal wavelet methods is a loss in energy. It is shown below that non-orthogonal methods still preserve a large portion of energy and do better in representing smooth structures compared to orthogonal methods that lack directionality support.

2.4.2. Complexity and redundancy

Complexities are unknown for higher-order methods such as curvelets and surfacelets but they are much higher compared to the other wavelets. Since they are based on discretizations of continuous signals, Haar, B-spline, Daubechies and Coiflet wavelets are of linear complexity $O(N)$ where N is the total number of points in the data. The efficiency of any method can be asserted by observing its redundancy.

Redundancy represents a combination of computational effort and space consumption. Redundancy can also be seen in the number of coefficients extracted and comparable to the ratio between the total number of elements in a dataset and the number of coefficients extracted. The number of coefficients extracted and ratios for a sample 512^3 dataset are shown in Table 1.

While Haar, B-spline, Daubechies and Coiflet wavelets have a redundancy near 1, surfacelets and Dual-tree wavelets have a redundancy of about 3 to 4 in the 3-D scale. Curvelets also have a redundancy between 4 to 5 if they use wavelets for their finest scale, or 40 if they use curvelets again. Curvelets and surfacelets are more redundant due to their native multi-resolution storage components and Dual-tree wavelets due to the capture of directional features. Although this is the case, the curvelet, surfacelet, and Dual-tree wavelet implementations with the lowest redundancy

ACCEPTED MANUSCRIPT

Method	# of coefficients extracted	Ratio	Method	# of coefficients extracted	Ratio
Source Grid	134,217,728	1.00	Daubechies-5	142,978,473	1.07
Haar	134,217,728	1.00	Daubechies-7	147,717,100	1.10
Linear B-spline	138,553,029	1.03	Coiflet-6	138,553,029	1.03
Quadratic B-spline	136,538,751	1.02	Coiflet-12	145,512,802	1.08
Cubic B-spline	142,978,473	1.07	Coiflet-18	152,517,105	1.14
Quartic B-spline	145,512,802	1.08	Dual-Tree	536,870,720	4.00
Quintic B-spline	152,517,105	1.14	Surfacelet	459,538,432	3.42
Daubechies-3	138,553,029	1.03	Curvelet	553,095,117	4.12

Table 1: **Coefficients**. Number of extracted coefficients and ratio compared to the total grid points of a 512^3 dataset.

are used for the course of this paper. To formulate a fair comparison, a number of four directional bands for curvelets is used and matched by surfacelets. For further tests which measure the amount of the energy acquired per each band, the number of extracted bands is increased to better measure the energy distribution per band for both methods. Our redundancy evaluations are similar to those presented in Ref. [2].

2.4.3. Coefficients and their storage components

Haar, B-spline, Daubechies and Coiflet wavelets have no specific multi-band coefficient extraction scheme and therefore typically decompose N^3 coefficients, where N is the grid size in each direction. The implementations used in this paper support wavelet signal extrapolation on boundary data through symmetrization. This results in numbers of coefficients slightly larger than a ratio of 1.0 compared to the original dataset but also in much more accurate representations of the boundaries. Curvelets and surfacelets require multi-band storage schemes and, therefore, they extract many more coefficients, as suggested by their redundancy amounts. Similarly, Dual-tree wavelets are also very redundant since they also extract four sets of directional coefficients per scale. Curvelet coefficients are stored with respect to their extraction based on their frequency partitioning per scale and direction. Surfacelet coefficients are similarly stored but differ in partitioning schemes in scale and direction compared to Curvelets. Table 1 shows the number of extracted coefficients in the system for our sample dataset.

ACCEPTED MANUSCRIPT

3. Implementation and thresholding techniques

For this paper, the discrete variants of all of the methods are utilized. The original software provided by the respective authors of each method is used, except for B-spline, Daubechies and Coiflet wavelets. Although some methods such as curvelets provide parallel and out-of-core methods to increase the computation speed which may be beneficial in high performance computing environments, the single threaded implementations of all methods is used for fair compute benchmarks. Haar Wavelets, B-spline, Daubechies and Coiflet wavelets are tested through the GNU Scientific Library (GSL). The GSL was modified to add support for 3D wavelet decomposition, as well as expand the variety of wavelets supported. All methods use their C++ variants and are compiled in Windows 7 Professional 64-bit using Microsoft Visual Studio 2008.

The term thresholding is used to summarize the act of selecting distinct coefficients, after a decomposition operation, using a certain criterium. In general, the magnitude of these coefficients for all transforms represents the energy for that specific component. Typical thresholding is done by considering a set of coefficients where all values are sorted by magnitude. Once sorted, a subset of coefficients may be chosen by percentage or by a specific number. Hard thresholding is defined by the selection of only specific coefficients and removing the rest. This approach is utilized for the extent of this paper when comparing the different representation methods. Hard thresholding is used either a priori or, after domain-specific knowledge of the data is gained, the coefficients are thresholded for specific magnitudes above a certain value.

For this paper, the decomposition of the original turbulent fields into coherent/incoherent parts is targeted in addition to the ability of the multi-resolution methods to reproduce turbulence structures with a minimal set of coefficients. The ansatz that the coherent, non-universal part of the flow can be captured with a reduced set of coefficients and, thus, does not require the representation and preservation of the highest frequency, smallest coefficients is tested. In representations and analysis of developed turbulence, the hope is that these coefficients correspond to some universal properties (often understood and describable as white noise).

Specialized optimal values for thresholding are only available for orthogonal methods and Gaussian noise [3, 24]. For example, the maximum quadratic error, $E\|\hat{f} - f\|_{2,N}^2/N$, between the

ACCEPTED MANUSCRIPT

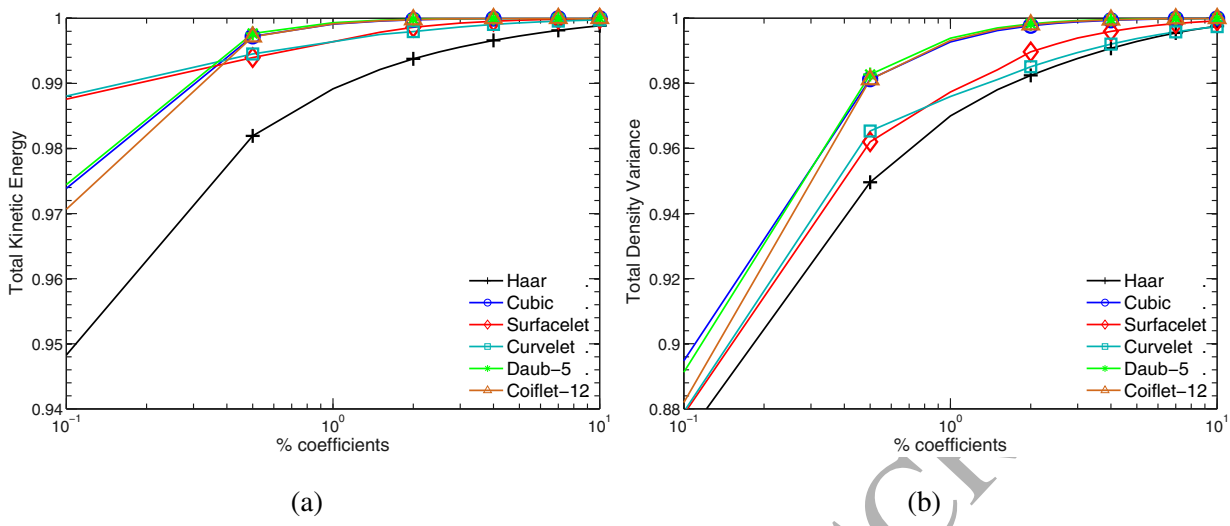
denoised signal, $\hat{f}$, and the original signal, f, can be minimized using the minimax threshold [46]. Here, $||v||^2_{2,N}$ denotes the usual l^2_N norm and N is the resolution. For large data samples, this threshold approaches its upper bound defined by $\epsilon = (2\sigma lnN)^{1/2}$, where σ is the variance of the noise. For turbulence data, however, the noise is likely not identically distributed Gaussian (both spatial correlation and non-Gaussian PDF are likely present) and the noise level is not known a priori.

In order to compare the methods considered, the hard thresholding of the coefficients based on the percentage of coefficients in the system is used. The minimax threshold was derived for selected orthogonal methods and approximated to be below 0.5% coefficients. Due to the large differences in the ratio of captured kinetic energy between methods, a larger percentage amount was selected for a fairer comparison. As in Ref. [3], comparisons are made after the selection of 3% coefficients. However, unlike CVS type decompositions, where the vorticity field is decomposed, thresholded and then reconstructed, while the coherent and incoherent velocity parts are obtained from the corresponding vorticity parts via the Bio-Savart law, here, the thresholding is performed on the decompositions of the primary variables, velocity and density. The motivation, described in the Introduction, follows from the applications considered in the comparison of the methods: LES-type simulations of more general flows than incompressible single-fluid turbulence and turbulence data compression for storage, analysis and visualization.

Fig. 3 compares the amounts of kinetic energy and density variance captured by several of the methods, as a function of the percentage of the coefficients. The ratios to the kinetic energy and density variance in original fields show that, even at low percentage of coefficients, both quantities can be well captured. Large gains in accuracy occur up to 1% coefficients, when all methods capture more than 98% of the kinetic energy and 96% of the density variance; the further increase in the number of coefficients above 1% results in smaller accuracy gains. Both surfacelets and curvelets are able to more accurately capture the kinetic energy below extreme percentages of 0.4% coefficients but are then overtaken by other methods. The 3% coefficient mark ensures that the kinetic energy captured exceeds 99% and density variance exceeds 98% for all methods. The remaining 97% coefficients are also computed to estimate the residual coherence or departure from a Guassian, white noise.

Hard thresholding at equivalent percentage of coefficients is a good indicator of the potential

ACCEPTED MANUSCRIPT



(a) (b)

Figure 3: **Total kinetic energy and density variance**. The ratios of total kinetic energy (a) and density variance from the original data compared to the compressed fields by percentage of coefficients. Large gains in accuracy occur up to 1% coefficients, when all methods capture more than 98% of the kinetic energy and 96% of the density variance.

of the different transforms considered in this paper. For Haar, B-spline, Daubechies and Coiflet wavelets, which do not have explicitly defined multi-bands, the coefficients can be ordered and thresholded without ambiguity. However, in general, transforms with multi-band support decompose into a series of coefficients for multiple bands. These bands are further subdivided into a larger series of angular directions. These transforms also extract real and complex coefficients. These features add additional layers of complexity and expand the capabilities of thresholding coefficients. The same number of bands and number of directions are extracted when comparing both surfacelets and curvelets.

More advanced thresholding schemes were tested for these multi-band methods, including the localization of features by scale and direction. Due to the difference in directional representation of each method, to perform a fair comparison, coefficients are thresholded by their largest magnitudes as a whole per direction and bands. Although thresholding coefficients per individual band or per angle may be more intuitive in finding specific features, reconstruction behavior is more consistently obtained when considering all coefficients as a whole. Each method has various parameters that change their functionality such as number of bands to decompose and variables related to the behavior of their filters. More modern, higher-order methods such as curvelets and

ACCEPTED MANUSCRIPT

surfacelets contain band systems with multiple levels of redundancy in order to capture angularity based on their filter type.

Due to each method having different filters, basis functions and lack of complete orthogonality, the relation between coefficients from different methods with respect to a given frequency is not trivial. Such a relation can be estimated by performing a power spectrum analysis in Fourier space, which is also covered in the results section.

4. Results and discussion

4.1. Test types and definitions

A series of tests are performed that compare the compressed datasets to the originals for quantities related to the velocity and density fields. For the velocity field, the comparison metrics used are the PDF of the velocity field itself and of enstrophy ($\Omega \equiv |\vec{\omega}|^2$), as well as enstrophy peak signal-to-noise ratio (PSNR) and mean square error (MSE) and the PDF of the unbiased measure of the state of the deviatoric strain rate tensor, s^* [41]. For the density field, the comparison metrics used are the density PSNR and MSE, density power spectrum, coherent/incoherent representations of the density field itself and iso-density surface corresponding to the pure light fluid. The pure light fluid iso-density surface is also characterized using several differential geometry tools: surface area, mean and Gaussian curvatures as well as the surface signature.

Previous comparisons related to the coherent/incoherent decompositions of turbulent fields have been restricted to orthogonal and biorthogonal (Harten) wavelets as applied to velocity and enstrophy PDFs, kinetic energy spectrum and the PDF of helicity and Lamb vector (e.g. [24]). While CVS type decompositions have been mainly used to investigate the coherent/incoherent parts of the flow, here the decomposition is performed at the level of the primary variables, velocity and density. This approach is not entirely new, for example it is the base of the SCALES method [9] and has also been used for the analysis of a passive scalar [5]. Here, the results presented address a more complex flow and encompass more representation methods, with more quantities related to the velocity field (e.g. the state of the deviatoric strain rate tensor) and, for the first time, quantities related to an active scalar field and its constant property surfaces.

ACCEPTED MANUSCRIPT

For all tests, the strongest 3% coefficients are referred to as defining the coherent part and the weakest 97% coefficients as defining the incoherent (residual) part of the respective turbulent quantity. A subset of methods are compared looking at the progression of reconstruction for the number and percentage of coefficients. It is noted that the relationship between percentage and number of coefficients is not the same for all methods as some methods such as curvelets and surfacelets may extract many more coefficients compared to the B-spline wavelets. The subset of methods include Haar, cubic B-spline, Daubechies-5, Coiflet-12, Dual-tree wavelets, surfacelets, and curvelets. For clarity, Dual-tree wavelets are only added in a subset of the figures, especially where they excel beyond the other methods; nevertheless, they are added in all the tables. These methods were selected as representative for each class of methods considered.

Finally, the last set of tests evaluate the performance and memory requirements of each of the methods. The performance is evaluated based on the runtime of each transform during its decomposition phase for a single scalar field. Only the single-threaded variant of each of the methods is considered. The hardware used to perform these tests contains an Intel Core i7 930 processor running at 3.80GHz and 48GB of DDR3 RAM.

4.1.1. Definitions and discussion on the quantities used for comparisons

The definitions and short explanations for the various test metrics and turbulence quantities used in this paper are given below.

PSNR in Eq.(1) is the primary measurement to evaluate point-wise accuracy. The signal-to-noise-ratio (SNR) was used in Ref. [39] in a similar analysis to compare the effects of compression on data. PSNR is used in this paper rather than SNR since it better reproduces the behavior of the representation methods over all percentages of coefficients thresholded. MSE is superior in showing the behavior at lower percentages with the eventual drop-off to zero due to numerical precision.

$$PSNR = 10 \cdot \log_{10}\left(\frac{MAX^2}{MSE}\right), \tag{1}$$

where MAX is the maximum possible value in the range of the data and MSE is the mean square error.

ACCEPTED MANUSCRIPT

Velocity ($\vec{u}$) and vorticity ($\vec{\omega} \equiv \nabla \times \vec{u}$) PDFs have been been used to ascertain the ability of orthogonal and B-spline wavelets to decompose turbulence fields into coherent and incoherent parts in Ref. [24]. This was accomplished by comparing the PDFs corresponding to the fields reconstructed using the largest 3% of the coefficients with the PDFs of the original signal and the PDFs corresponding to the fields reconstructed using the remaining 97% of the coefficients with a Gaussian. While these quantities, together with the helicity ($\vec{u} \cdot \vec{\omega}$) and Lamb vector ($\vec{u} \times \vec{\omega}$), are useful in assessing the representation of turbulence with a limited number of coefficients, they do not fully address the local structure of turbulence and, thus, the incoherence of the turbulence data. Therefore, here, we also consider the quantity s^*, which uniquely defines the state of the deviatoric strain rate tensor for incompressible flows [41]:

$$s^* = \left(\frac{-3\sqrt{6}\alpha\beta\gamma}{(\alpha^2 + \beta^2 + \gamma^2)^{3/2}} \right), \tag{2}$$

where α, β and γ are the eigenvalues of the deviatoric strain rate tensor, $\underline{\underline{S}} \equiv 1/2(\nabla\vec{u} + (\nabla\vec{u})^T)$, or by components, $S_{ij} = 1/2(\partial u_i/\partial x_j + \partial u_j/\partial x_i)$. For the flow considered here, the divergence of the velocity is not zero, but it is small at the time when the flow is being analyzed. In addition, the enstrophy PSNR and MSE are also shown.

The pure light, pure heavy and mixed fluid regions are defined as all the points with density values satisfying the relations [14]:

$$\rho \leq \rho_1 + 0.05\,(\rho_2 - \rho_1)\,, \tag{3}$$

$$\rho \geq \rho_2 - 0.05\,(\rho_2 - \rho_1)\,, \tag{4}$$

$$0.45\,(\rho_2 - \rho_1) \leq \rho \leq 0.55\,(\rho_2 - \rho_1)\,, \tag{5}$$

where ρ_1 and ρ_2 are densities of the two pure fluids. For comparing the different representation methods, in this paper we focus on the surface bounding the pure light fluid regions, which is defined by the formula $\rho(x, y, z) = \rho_l \equiv \rho_1 + 0.05\,(\rho_2 - \rho_1)$. This surface is visualized using the coherent and incoherent parts of the density field. In addition, the ability of the representation

ACCEPTED MANUSCRIPT

methods to capture the surface area, mean and Gaussian curvatures and surface signature with a reduced set of coefficients is also discussed. The area of the surface defined by $\rho(x, y, z) = \rho_l$ can be calculated using Theorem 1.2.4. from Ref. [42], by replacing the Borel measurable nonnegative function Φ with the delta function, $\delta(\rho(x, y, z) - \rho_l)$, to obtain:

$$SA_{\rho_l} = \int_{\Omega} |\nabla\rho|\delta(\rho - \rho_l)\, \mathrm{d}v. \tag{6}$$

If one further assumes homogeneity of the flow, so that averages can be calculated as volume averages, then $SA_{\rho_l} =< |\nabla\rho||_{\rho=\rho_l} > f(\rho_l)V$, which is the product of the conditional average of the density gradient magnitude and density PDF at $f(\rho(x, y, z)) = f(\rho_l)$, and the total volume of the flow. For a uniform discrete representation of the data, as obtained from DNS, the surface area formula becomes

$$SA_{\rho_l} = \left(\frac{\sum ||\nabla\rho||}{n\,(\rho_{max} - \rho_{min})}\right)(2\pi)^3\,, \tag{7}$$

where n is the ratio between the total number of points, N, and the number of bins, n_b, used to calculate the density PDF. The discretization errors becomes small for large values of n, n_b. Following Ref. [5], the surface signature is defined as the joint PDF of the absolute value of the shape index, S, and curvedness, C:

$$S = \left|\left(\frac{-2}{\pi}\right) tan^{-1}\left(\frac{H}{\sqrt{H^2 - K}}\right)\right|, \tag{8}$$

$$C = \sqrt{2H^2 - K}, \tag{9}$$

where H and K are the mean and Gaussian curvatures. Since the same surface is considered for all the representation methods and this surface may be multiply connected, the curvedness is not normalized using the surface area and volume as done in Ref. [5].

4.2. Quantities related to the velocity field

In this section, quantities related to velocity field are compared for the representation methods discussed, with respect to their ability to separate the flow into coherent and incoherent parts and

ACCEPTED MANUSCRIPT

represent the turbulence structure with a reduced set of coefficients. Some of the quantities considered here (velocity and enstrophy PDFs) have been investigated in previous studies; however, only the orthogonal (Coiflet) and bi-orthogonal (Harten) wavelets have been explored. Here, we calculate these quantities as a consistency check with the previous studies and also for comparison with the additional representation methods considered. In addition, several new metrics are examined to better assess the capturing of the local turbulence structure.

4.2.1. Velocity probability density function

The velocity PDF itself is Gaussian in most fully developed turbulent flows and this is reflected in the PDF of the fields obtained from the largest 3% of the coefficients obtained for all methods considered here (not shown). Figure 4 compares the PDFs of the residual vertical velocity PDF for a subset of the methods, which includes one method for each of the classes considered. All methods yield symmetrical residual velocity PDFs; however, Haar and curvelet transforms lead to exponential shapes, while the others are closer to Gaussian shapes. The Kurtosis values are: 6.13 (Haar), 6.44 (cubic B-spline), 4.62 (surfacelet), 7.10 (curvelet), 5.79 (Daubechies-5), and 6.10 (Coiflet-12). In addition, the field's extrema are the smallest for the cubic B-spline, with relatively close values for the other high order wavelet methods. The surfacelet transform has slightly larger extrema, but still in the range of the higher order wavelets, while the curvelets and Haar wavelets have much larger extrema. Thus, while there is no bias towards positive or negative values, the residual still retains large velocity values.

4.2.2. Vorticity peak signal-to-noise ratio and mean square error

Velocity PDF, while useful as a first metric to assess the residual part of the velocity field, does not contain any spatial information. The next quantity considered for assessing the representation methods is the vorticity. Fig. 5 shows the PSNR for enstrophy (vorticity square) as a function of the percentage of coefficients retained. In this case, the signal is obtained from three lossy scalar components, as each velocity direction is separately represented, then each vorticity component is calculated from the reconstructed velocities. The cubic B-spline wavelets are observed to have the highest overall reconstruction accuracy spanning percentages above 10% coefficients.

ACCEPTED MANUSCRIPT

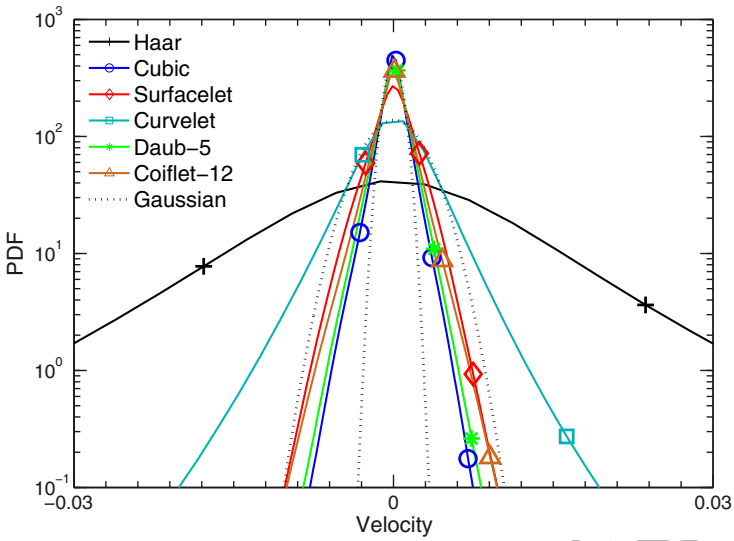


Figure 4: **Residual velocity PDF.** The PDF is obtained from the remaining 97% coefficients in the vertical velocity transform. Two Gaussian signals matching the variances of the curvelet and B-spline wavelet results are plotted.

Table 2 shows several numerical results for 3% of the coefficients retained (coherent part). All the B-splines methods above linear show consistently high values for PSNR and low values for MSE and maximum error (L_∞ norm), with the best results obtained for cubic B-splines. Overall, all wavelets methods give close results (except the Haar wavelets, which is expected), with slightly worse values obtained for Daubechies and Coiflets (in this order). The surfacelets yield results within the range of Coiflets (but worse than higher order Coiflets) and curvelets show the largest point-wise departure from the original signal (in all three metrics considered in Table 2) for 3% of coefficients retained. Interestingly, for each family there is an optimal order which yields the best point-wise representation, so that further increasing the order of the method starts to deteriorate the results. Thus, the discrete signal in lower families tends to not have enough representation to capture the original field while the higher order representations within each family introduce oscillations which reduce the overall quality.

The main focus for the remaining tests will be on the subset of representation methods, consisting of Haar, cubic B-splines, Daubechies-5, Coiflet-12, Dual-tree wavelets, surfacelets and curvelets. The 3% thresholding for the velocity representation captures 94.4%, 99.4%, 99.6%, 99.6%, 98.1%, 97.7%, and 95.4% (respectively) of the total enstrophy for this subset. Within each

ACCEPTED MANUSCRIPT

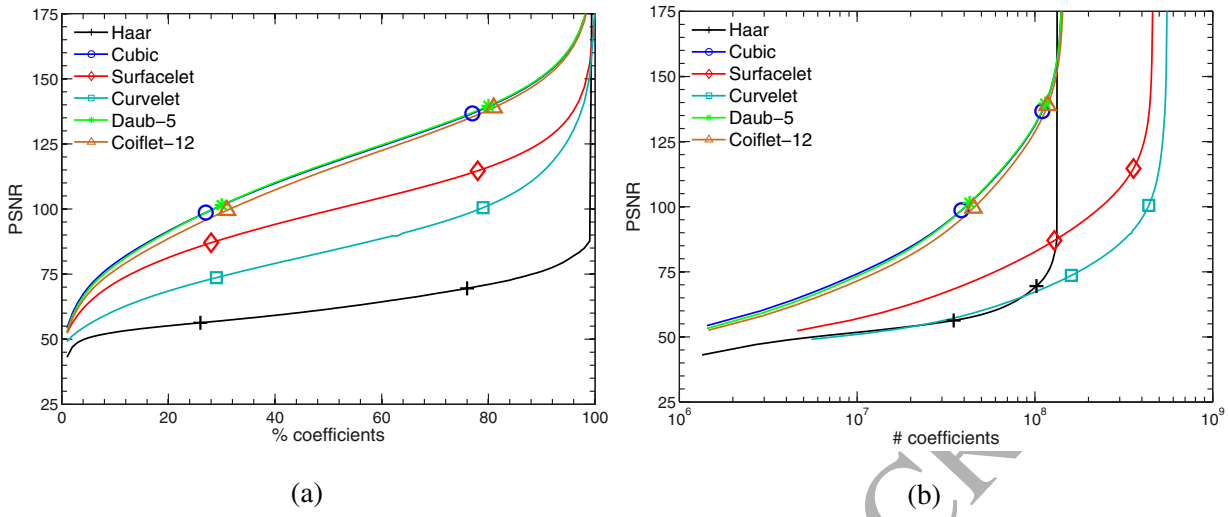


Figure 5: **Enstrophy PSNR** as a function of the percentage of coefficients (a) and number of coefficients (b) retained for reconstruction. Higher PSNR values represent a more accurate point-wise reconstruction. Cubic B-splines have the highest quality through all percentages and number of coefficients.

class of the methods, the total vorticity captured using this thresholding generally increases as the order of the method increases, but decreases using the highest methods.

4.2.3. Vorticity probability density function

Vorticity PDF, computed for the remaining 97% coefficients, is shown in Fig. 7. Similar to velocity, all methods yield symmetrical residual PDFs. Haar and curvelet transforms lead to wide tails shapes, while the others have sharper shapes, but still different than Gaussian. The Kurtosis values are: 7.415 (Haar), 9.27 (Cubic B-spline), 8.72 (Surfacelet), 23.9 (Curvelet), 8.26 (Daubechies-5), and 8.70 (Coiflet-12). The field's extrema are the smallest for the cubic B-spline, with relatively close values for the other high order wavelet methods. The moderate increase in family order for B-splines achieved the smallest extrema compared to constant B-splines (Harten-3) used in Ref. [24]. The surfacelet transform has slightly larger extrema, but still in the range of the higher order wavelets, while the curvelets and Haar wavelets have much larger extrema. Similar to velocity, while there is no bias towards positive or negative values, the residual still retains large vorticity events.

ACCEPTED MANUSCRIPT

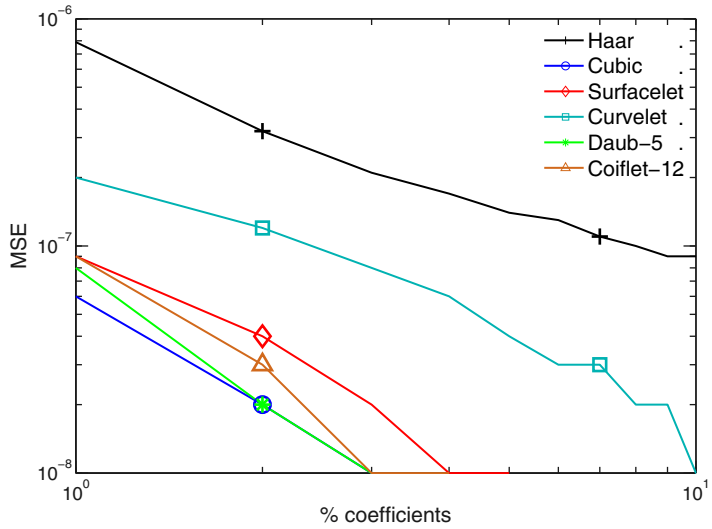


Figure 6: **Enstrophy MSE** variation with the percentage of coefficients retained for reconstruction. Lower MSE values represent a more accurate point-wise reconstruction. Cubic B-splines have the lowest MSE at very low percentages of coefficients, but are matched by Daubechies-5 and Coiflet-12 wavelets at higher percentages.

4.2.4. The state of the deviatoric strain rate tensor

Previous studies, while examining several wavelet transforms for their ability to represent quantities related to the velocity field, have not investigated in detail the local structure of turbulence. The state of the strain rate tensor can identify the tendency of the flow to organize in certain types of structures (e.g. stable-focus/stretching or unstable-node/saddle/saddle) characteristic of fully developed turbulence [44, 43] and may be a better indication of the residual coherence than velocity or vorticity PDFs. One quantity which can uniquely define the state of the deviatoric strain rate tensor, s^*, Eq. (2), was first derived in Ref. [41]. For most fully developed turbulent flows, the PDF of s^* peaks at 1, while it becomes flat for a Gaussian signal. Therefore, a structureless incoherent part of the flow should exhibit a flat s^* PDF; while the departure from a flat PDF indicates that the residual is not structureless. Fig. 8 shows the PDF of s^* for the velocity field reconstructed from the remainder 97% of the coefficients and the skewness and kurtosis of the PDFs are tabulated in 3. A flat PDF has skewness zero and kurtosis equal to 1.8. We observe that for most methods, the remaining signal is mostly Gaussian noise in the remaining 97% coefficients. The method closest to a flat s^* PDF is the quadratic B-splines, while surfacelets and

ACCEPTED MANUSCRIPT

Method	PSNR	MSE	Max Error	Method	PSNR	MSE	Max Error
Haar	48.786	$1.9921x10^{-8}$	$5.1897x10^{-2}$	Daubechies-7	61.784	$1.0695x10^{-8}$	$1.2053x10^{-2}$
Linear B-spline	60.691	$1.3756x10^{-8}$	$0.9685x10^{-2}$	Coiflet-6	57.632	$2.7824x10^{-8}$	$1.2457x10^{-2}$
Quadratic B-spline	63.692	$0.6893x10^{-8}$	$0.8786x10^{-2}$	Coiflet-12	61.865	$1.0499x10^{-8}$	$0.9154x10^{-2}$
Cubic B-spline	**64.030**	$\mathbf{0.6377}x10^{-8}$	$\mathbf{0.7181}x10^{-2}$	Coiflet-18	61.504	$1.1407x10^{-8}$	$1.0614x10^{-2}$
Quartic B-spline	63.645	$0.6969x10^{-8}$	$0.7374x10^{-2}$	Dual-Tree	60.132	$1.5648x10^{-8}$	$1.3191x10^{-2}$
Quintic B-spline	62.099	$0.9947x10^{-8}$	$1.1021x10^{-2}$	Surfacelet	59.319	$1.8867x10^{-8}$	$1.1105x10^{-2}$
Daubechies-3	61.437	$1.1586x10^{-8}$	$0.7198x10^{-2}$	Curvelet	52.929	$8.2165x10^{-8}$	$4.8001x10^{-2}$
Daubechies-5	63.010	$0.8065x10^{-8}$	$0.8216x10^{-2}$				

Table 2: **Enstrophy comparison at 3% coefficients**. The methods with the most accurate properties are represented in bold. Cubic splines have the best point-wise characteristics for enstrophy representation compared to any other method using these metrics.

curvelets present the largest departures from a flat PDF.

4.3. Quantities related to the density field

Previous comparisons between representation methods have been focused on quantities related to the velocity field only. Since many flows occur in the presence of passive or active scalars, here we are also comparing the representation methods with respect to their potential to separate the scalar field into coherent and incoherent parts and represent the structure of the field with a reduced set of coefficients.

In this flow, the scalar (density) is an active scalar, since the differential buoyancy force which drives the turbulence is generated by the variations of density. Since density comparisons have not been performed before, visualizations of the density field itself and constant density surfaces are also shown, together with the tools considered above, namely PSNR, MSE and PDFs. The ability of the representation methods to reproduce the density power spectrum with a limited number of coefficients is also discussed. For the surfacelet and curvelet transforms, the power spectra are calculated for each band separately, which shows the wavenumbers captured by each band, as well as the overlap between bands. In order to characterize the representation of constant density surfaces, surface area, mean and Gaussian curvatures, as well as the surface signature are also calculated.

ACCEPTED MANUSCRIPT

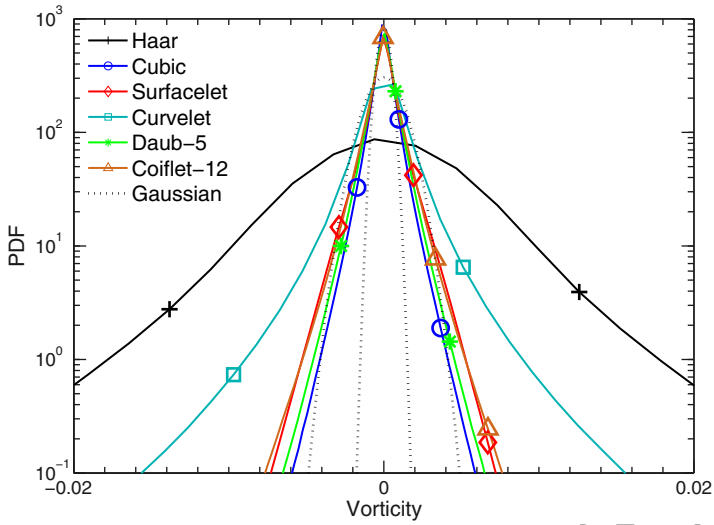


Figure 7: **Residual vorticity PDF.** The PDF is obtained from the remaining 97% coefficients in the vertical vorticity transform. Two Gaussian signals matching the variances of the curvelet and B-spline wavelet results are plotted.

4.3.1. Density visualization

The subset of the methods described above are visually compared by using the center 2D slice of the dataset (Fig. 9) and a log-difference map comparing this slice to the original (Fig. 10) for the density field reconstructed from the 3% largest coefficients. Except the Haar wavelets, which approximate curved variations by a staircase-like variation, the rest of the wavelet methods reconstruct the density field reasonably well. The closest results with the original dataset is obtained using cubic B-splines see also Fig. 10), followed by Daubechies-5, Dual-tree, surfacelets and curvelets (in this order). Both curvelet and surfacelet transforms show relatively large differences between the original field and the field reconstructed for the 3% percent largest coefficients.

The residual parts based on the remaining 97% coefficients are also compared and visualized in Fig. 11. Similar to Fig. 10, Cubic B-spline wavelets perform the best, as they exhibit the least amount of residual density in their remaining coefficients. The Haar wavelet reconstruction contains the most remaining structures. Daubechies-5 and Dual-tree wavelets contain slightly more residual density than cubic B-splines. Relative to cubic B-splines, curvelet and surfacelet coefficients contain larger amounts of remaining structures. Between the two multi-band methods, the curvelets reconstruction leads to the least structureless residual density field.

ACCEPTED MANUSCRIPT

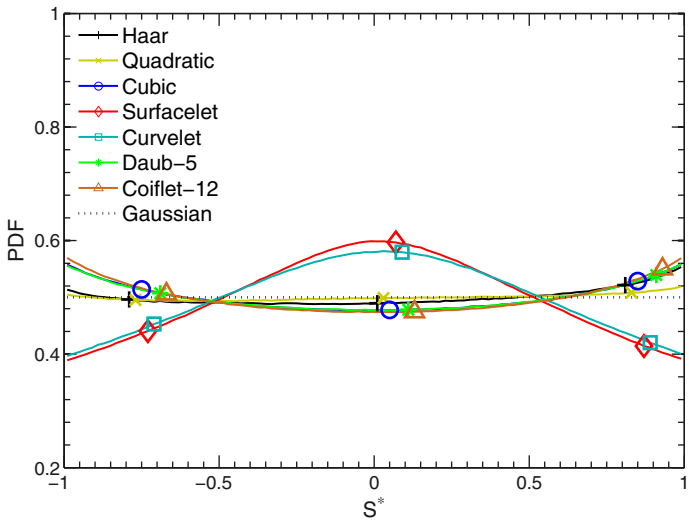


Figure 8: **PDF of** s^* corresponding to the remaining 97% coefficients. Although all methods yield a near-horizontal PDF, indicating a quasi-Gaussian white noise residual, quadratic B-spline wavelets are observed to have residual signal closest to a Gaussian.

4.3.2. Density peak signal-to-noise ratio and mean square error

Fig. 12 shows the PSNR variation as a function of the percentage and number of coefficients, while Fig. 13 shows the MSE variation for the subset of the methods defined above. As the number or percentage of coefficients are increased, the amount of increase in accuracy is similar across all methods tested. Numerical results for all methods at 3% coefficients are presented in Table 4. Similar to vorticity PSNR and MSE, at low percentages of coefficients used in reconstruction, cubic B-spline wavelets are superior by having the highest PSNR and lowest MSE. The higher-order B-spline wavelets (quadratic, cubic, and quartic), in general, out-perform the rest of the methods tested. When comparing methods within each individual family at 3% coefficients, there is an observed parabolic behavior in reconstruction quality where the best reconstruction is achieved by using the middle members. These results are consistent to those obtained for vorticity, since the lower families tend to not have enough representation to capture the original density field, while the higher families introduce oscillations which reduce the overall quality of the representation.

Surfacelets have the benefit of a PSNR that is close to B-splines and a low point-wise maximum error while achieving multi-band functionality. As expected, curvelets perform better than

ACCEPTED MANUSCRIPT

Method	Skewness	Kurtosis	Method	Skewness	Kurtosis
Haar	-0.018671	1.7837	Daubechies-7	-0.000893	1.7773
Linear B-spline	0.000594	1.7502	Coiflets-6	0.000221	1.7485
Quadratic B-spline	-0.010178	**1.7970**	Coiflets-12	0.000523	1.7566
Cubic B-spline	-0.000261	1.7622	Coiflets-18	-0.000581	1.7663
Quartic B-spline	**-0.000031**	1.7774	Dual-Tree	0.005532	1.8299
Quintic B-spline	-0.000402	1.7675	Surfacelet	-0.007940	1.9495
Daubechies-3	-0.002019	1.7605	Curvelet	-0.010240	1.9269
Daubechies-5	-0.000803	1.7627			

Table 3: **Statistical properties of residual** s^*. The methods with the best properties are represented in bold. Skewness and kurtosis are computed for the remaining 97% coefficients for all methods. A flat PDF (corresponding to a Gaussian noise residual) has a kurtosis of 1.8. All methods have skewness close to zero, while quadratic B-spline wavelets have the kurtosis closest to 1.8.

Haar wavelets but not as well as the remaining methods. When increasing the number of coefficients, curvelets do not surpass the PSNR quality of the alternative methods mainly due to the high levels of redundancy needed to support energy separation in bands and directional coefficients. By decreasing the number of bands, curvelets can be made to extract a smaller number of coefficients causing a left-ward shift of the PSNR plot. Even though this will increase the point-wise accuracy of the curvelet transform for a given percentage of coefficients, it will still be lower than that obtained using cubic B-spline wavelets. Daubechies wavelets generally perform better than Coiflets, but both are still below B-spline wavelets. When considering the accuracy for percentages of coefficients above 20%, the highest member of B-spline wavelets family tested, quintic, is the most accurate.

When instead comparing total number of coefficients, which is more relevant for compression and storage purposes, cubic B-spline wavelets once again provide the most accurate reconstruction using the least number of coefficients. The relative ranking in reconstruction quality from previous tests is observed again for quadratic, cubic, and quartic B-splines. These methods once again show slight improvement in accuracy when the number of coefficients is increased. Finally, due to the number of bands extracted causing high redundancy levels, the PSNR results for curvelets and surfacelets are generally worse than the rest of the methods. After a certain point, surfacelets

ACCEPTED MANUSCRIPT

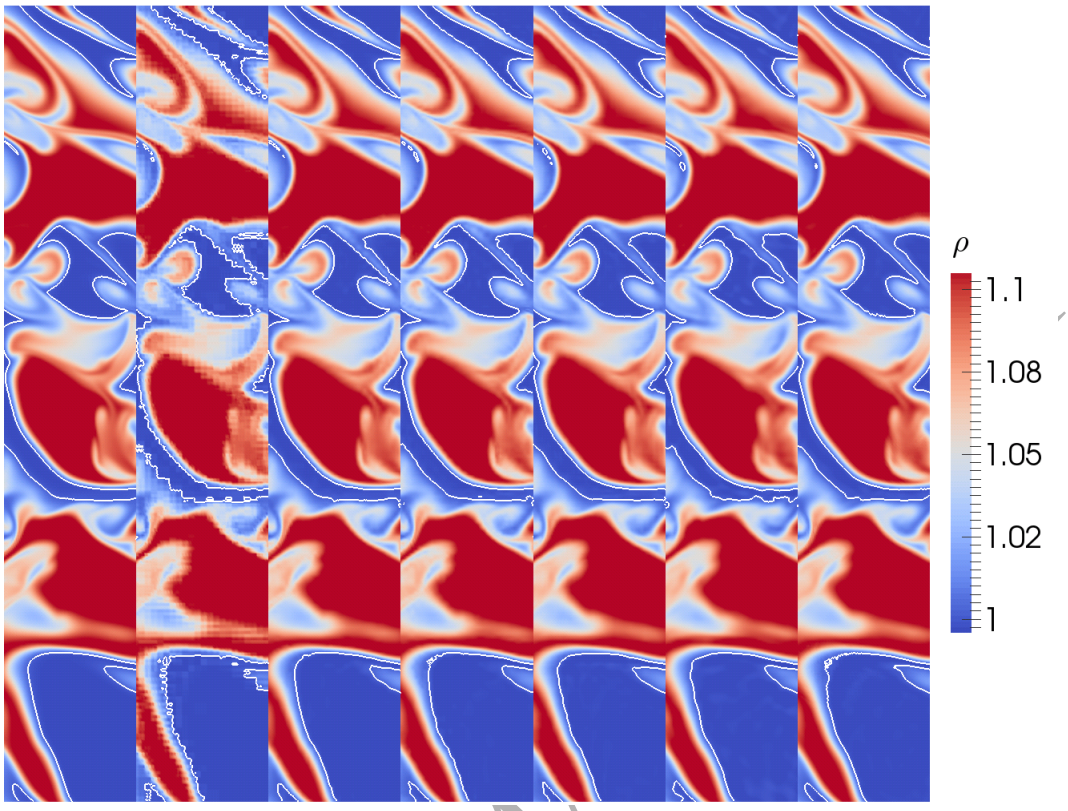


Figure 9: **Density visualizations**. 2D center slices for a subset of the methods are compared at 3% coefficients. From left to right: original, Haar, cubic B-spline, Daubechies-5, Dual-tree wavelets, curvelets, and surfacelets. The white iso-line represents the boundary of the pure-light fluid (blue). Haar wavelets perform noticeably worse and have strong artifacts. The rest of the methods are able to better reconstruct the original field.

manage to overtake the accuracy results of Haar wavelets despite extracting significantly more coefficients. When comparing both band-based methods, surfacelets perform better than curvelets.

4.3.3. Residual density probability density function

Density PDF, computed for the remaining 97% coefficients, is shown in Fig. 14. Similar to velocity, all methods yield symmetrical residual PDFs. Haar and curvelet transforms lead to widest tails, while the other methods have narrower shapes, but still wider than a Gaussian. The Kurtosis values are: 8.00 (Haar), 6.63 (Cubic B-spline), 4.38 (Surfacelet), 5.53 (Curvelet), 5.79 (Daubechies-5), and 6.20 (Coiflet-12). The field's extremas are the smallest for the cubic B-spline, with relatively close values for the other high order wavelet methods. The surfacelet transform has

ACCEPTED MANUSCRIPT

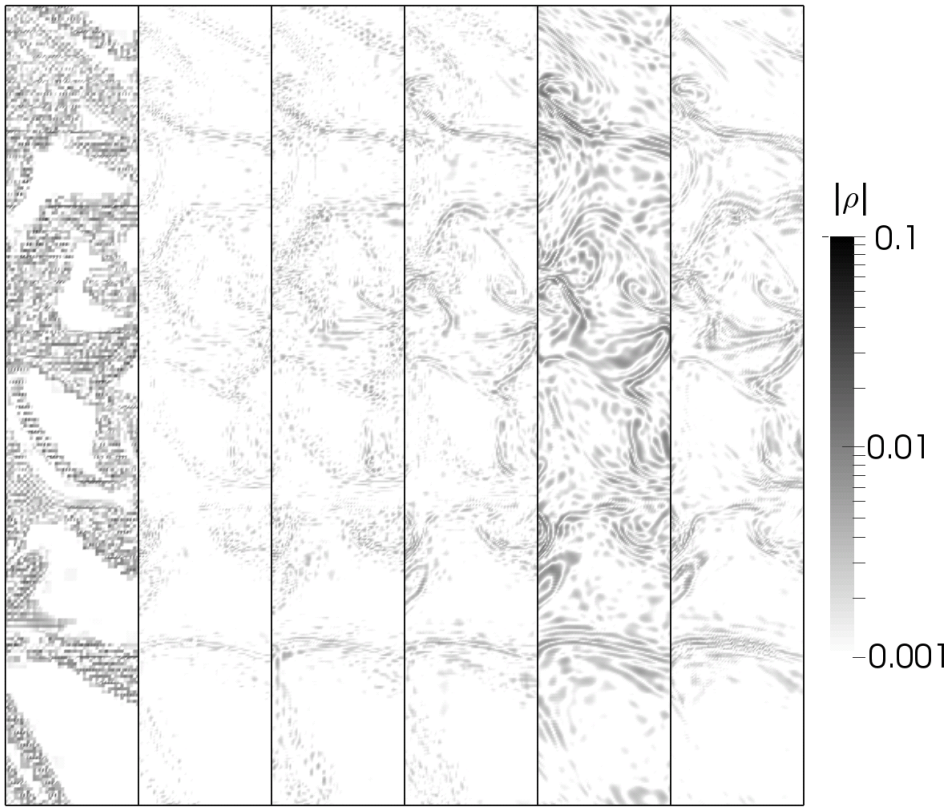


Figure 10: **Density logarithmic difference** between the original 2D center slice for methods at 3% coefficients and plotted in a logarithmic scale. From left to right: Haar, cubic B-spline, Daubechies-5, Dual-tree wavelets, curvelets and surfacelets. Haar wavelets have the most visible error, and cubic B-spline wavelets the least.

slightly larger extremas, but still in the range of the higher order wavelets, while the curvelets and Haar wavelets have much larger extremas.

4.3.4. Density power spectrum and multi-band analysis

In order to examine how the coherent and incoherent part of the density are distributed across scales, Figs. 15 and 16 compare the power spectrum at coherent 3% coefficients, and the remaining 97% coefficients. All methods capture the large and part of the inertial range scales, and the results depart from the original spectrum at small scales. The wavelet methods introduce small scale noise, which translates into increased values at higher wave-numbers. Within each class, higher order methods capture better the high frequency features of the density field. The Dual-tree wavelet transform is able to provide the closest reconstruction, using 3% of the coefficients, to the original

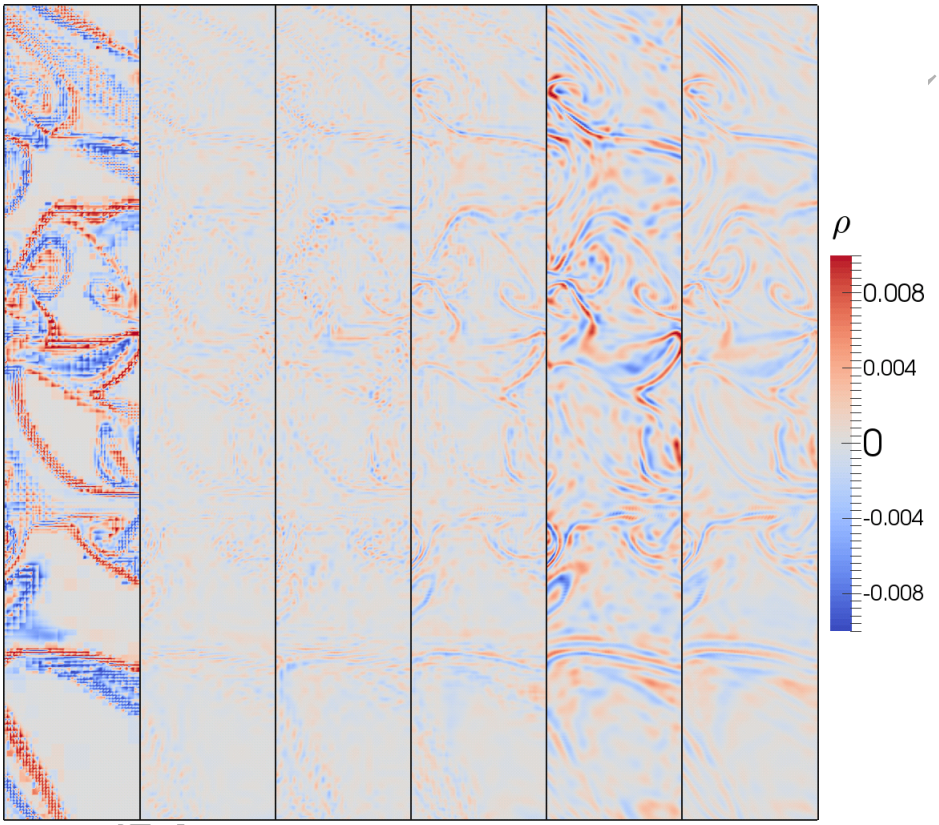


Figure 11: **Residual density visualization**. The remaining density residing in the weakest 97% coefficients is shown for Haar, Cubic B-spline, Daubechies-5 and Dual-tree wavelets, curvelets and surfacelets (left to right). Cubic B-spline wavelets show the weakest structures in the remaining density. Curvelets and surfacelets extract many more redundant coefficients, and their remaining structures are clearly visible.

ACCEPTED

ACCEPTED MANUSCRIPT

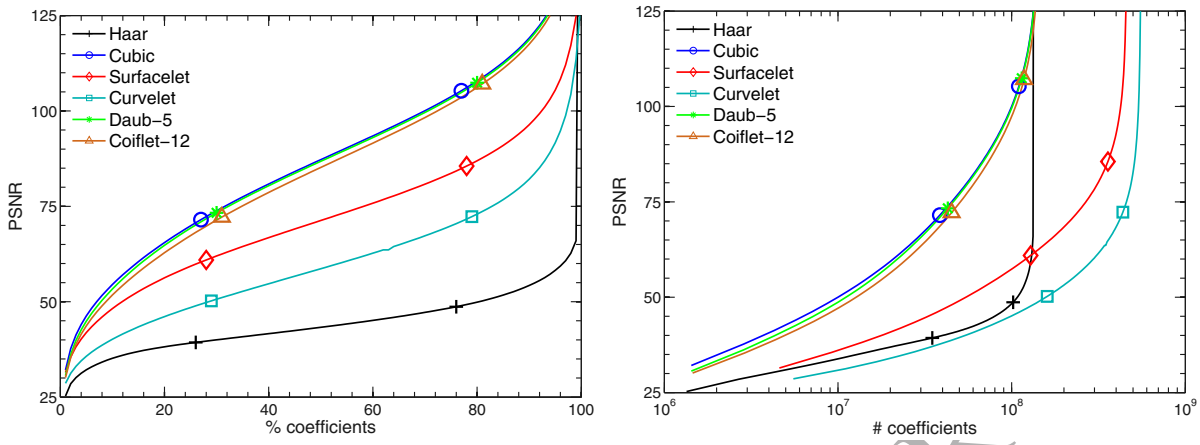


Figure 12: **Density field PSNR** as a functions of the percentage of coefficients (left) and number of coefficients (right). Higher PSNR values represents a more accurate point-wise reconstruction. The cubic B-splines provide the best accuracy at the lowest percentages and number of coefficients compared to other methods. Although not shown, for larger percentages of coefficients retained (i.e. $> 20\%$), within each class, higher order methods such as quintic B-splines perform slightly better.

spectrum. The same order of B-spline family is able to achieve the next closest reconstruction. On the other hand, the surfacelet and curvelet transforms tend to smooth out the high frequency components and the resulting spectrum is lower than the original spectrum at high wave-numbers. Both methods also capture a large amount of energy in the remaining 97% coefficients due to their high redundancy.

To compare the surfacelet and curvelet band systems with respect to the wavenumber coverage, a number of 6 bands are chosen, where band 0 represents the largest features and band 5 represents the smallest. As previously explained, a larger number of bands are used in this set of the calculations to better distribute the energy and better measure the strengths of each method with respect to the coherent/incoherent decomposition and capture of the flow structure within each band. Except band 0, the curvelet bands are narrower than the surfacelet bands, providing more localization in wavenumber space (Fig. 17). This can be seen, quantitatively, from table 5 which measures the contribution of each band to the total density variance. As proposed by Pullin et al. [5], the integral velocities for each band despite the different length of scales sum up to virtually 1 although regions overlap. This shows both methods, although overly redundant, are able to preserve energy

ACCEPTED MANUSCRIPT

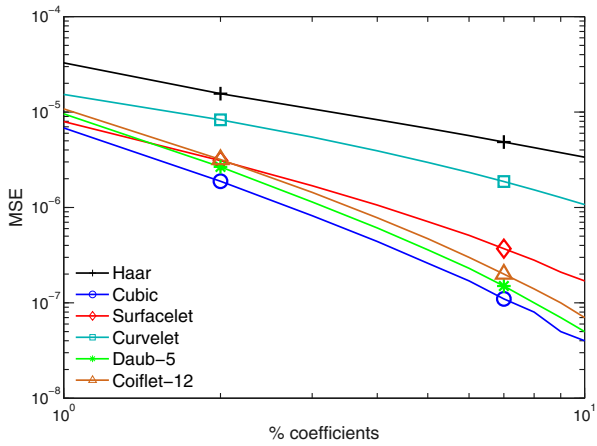


Figure 13: **Density field MSE** variation with the percentage of coefficients. Lower MSE values represent a more accurate point-wise reconstruction. The error becomes indistinguishable for more than 10% coefficients used in the reconstruction. The results show that cubic B-splines have the lowest MSE at very low percentages of coefficients.

well when factoring in all bands. From the point of view of flow analysis, the narrower curvelet bands may allow better representation of specific flow structures associated with certain frequencies. However, the wider bands surfacelets provide lead to better overall representation of the data in terms of percentage of coefficients as seen in Fig. 12 above.

4.3.5. Pure light fluid isosurface visualization

So far, the performance of the representation methods has been investigated with respect to describing various turbulent fields and their separation into coherent and incoherent parts. However, in turbulence problems involving scalar mixing, e.g. pollutant dispersion or non-premixed combustion, accurately capturing constant property surfaces is also very important. Fig. 18 visualizes isosurfaces of pure light fluid obtained using the reconstruction based on 3% coefficients. Although the wavelet methods may have higher point-wise PSNR, the curvelet and surfacelet isosurfaces are much smoother and very near to the original dataset. Both curvelets and surfacelets sacrifice point-wise accuracy in exchange for superior curved structure representation.

Among the other methods, the Haar wavelets have the coarsest representation of isosurfaces, as expected. The representation becomes very blocky and it has the highest visible difference compared to the original isosurfaces. The coarsening effect is amplified when derivatives are cal-

ACCEPTED MANUSCRIPT

Method	PSNR	MSE	Max Error	Method	PSNR	MSE	Max Error
Haar	30.114	$1.0791x10^{-5}$	$5.0793x10^{-2}$	Daubechies-7	39.142	$1.3502x10^{-6}$	$1.6532x10^{-2}$
Linear B-spline	37.940	$1.7807x10^{-6}$	$1.9407x10^{-2}$	Coiflet-6	34.760	$3.7028x10^{-6}$	$2.2010x10^{-2}$
Quadratic B-spline	41.095	$8.6113x10^{-7}$	$1.4169x10^{-2}$	Coiflet-12	38.860	$1.4406x10^{-6}$	$1.4867x10^{-2}$
Cubic B-spline	**41.282**	$\mathbf{8.2476}x10^{-7}$	$1.1944x10^{-2}$	Coiflet-18	38.830	$1.4505x10^{-6}$	$1.5929x10^{-2}$
Quartic B-spline	40.553	$9.7552x10^{-7}$	$\mathbf{1.1424}x10^{-2}$	Dual-Tree	39.020	$1.3884x10^{-6}$	$1.4690x10^{-2}$
Quintic B-spline	39.753	$1.1729x10^{-6}$	$1.4073x10^{-2}$	Surfacelet	38.169	$1.6892x10^{-6}$	$1.3875x10^{-2}$
Daubechies-3	38.029	$1.7442x10^{-6}$	$1.5653x10^{-2}$	Curvelet	33.075	$5.4580x10^{-6}$	$2.8111x10^{-2}$
Daubechies-5	39.858	$1.1449x10^{-6}$	$1.5414x10^{-2}$				

Table 4: **Density comparison at 3% coefficients**. The methods with the most accurate properties are represented in bold. The cubic B-splines have the Cubic splines have the best point-wise characteristics for enstrophy representation compared to any other method using these metrics. Although not as accurate from the point of view of PSNR and MSE, quartic B-splines exhibit a more consistent maximum point-wise error throughout the entire dataset.

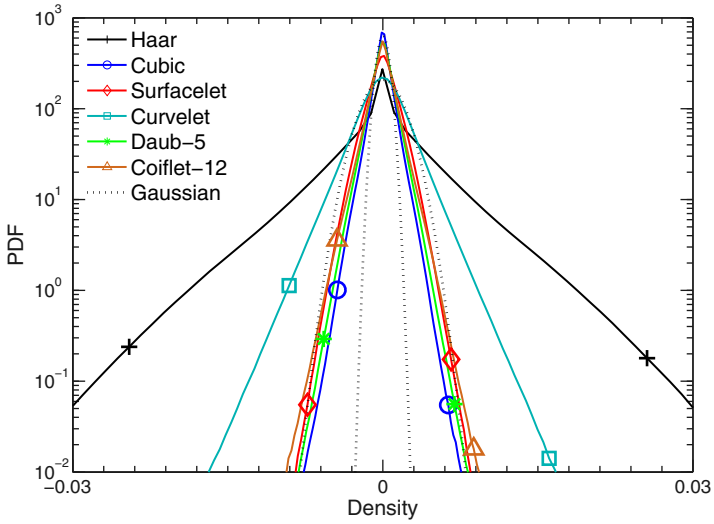


Figure 14: **Residual density PDF**. PDF for the remaining 97% coefficients. Two Gaussian signals matching the variances of the curvelet and B-spline wavelet results are plotted.

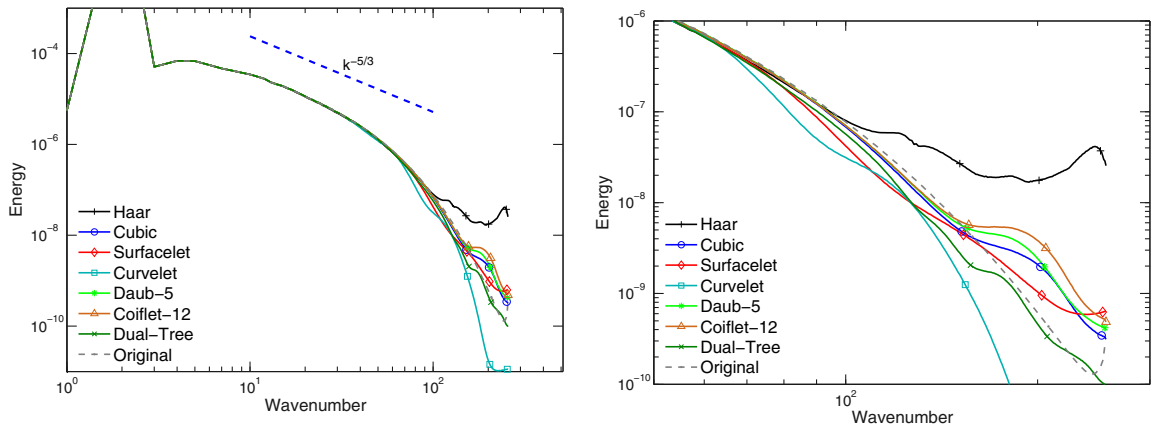


Figure 15: **Density power spectrum** for the coherent density field. The $k^{-5/3}$ variation is also plotted for comparison. A close-up for higher wavenumbers is provided in b).

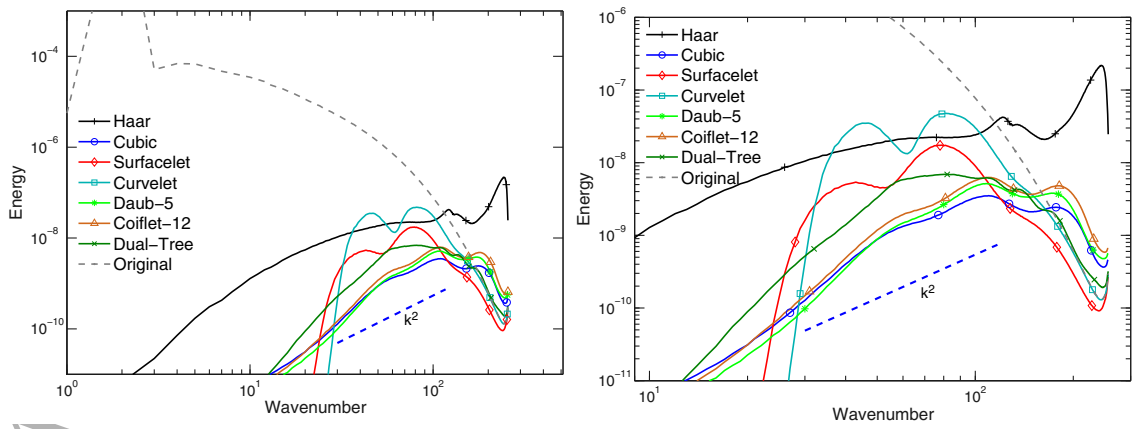
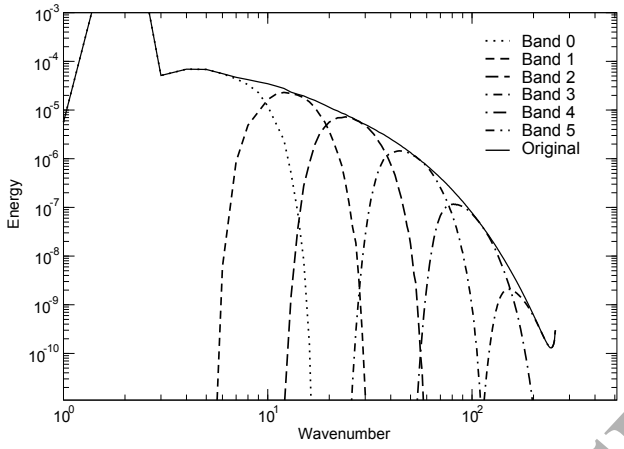
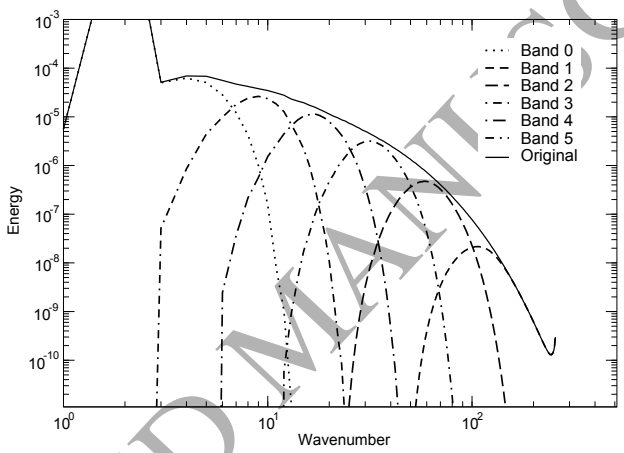


Figure 16: **Density power spectrum** for the residual density. The k^2 variation corresponding to a Gaussian signal is also plotted. A close-up for higher wavenumbers is provided in b).

ACCEPTED MANUSCRIPT



(a)



(b)

Figure 17: **Density power spectrum** band-based decomposition using six bands for (a) curvelets and (b) surfacelets. Higher level surfacelet bands are wider compared to those of curvelets.

Band	Surfacelet $< \rho'^2 >_i / < \rho'^2 >$	Curvelet $< \rho'^2 >_i / < \rho'^2 >$
Original	1.0000000	1.0000000
0	0.6785391	0.8000469
1	0.2080657	0.1447320
2	0.0861972	0.0465525
3	0.0237471	0.0080909
4	0.0032868	0.0005679
5	0.0001637	0.0000095

Table 5: **Band contribution to density variance.**

ACCEPTED MANUSCRIPT

culated making the Haar wavelets unsuitable for the preservation of constant property surfaces. Although cubic B-splines, Daubechies, and Coiflets are able to reproduce smoother surfaces than Haar wavelets, there are visible ripples in the isosurface visualization related to the sinusoidal-like reconstruction signal that these methods are representing (seen in Fig. A.21 in the Appendix). In comparison, curvelets, surfacelets and Dual-tree wavelets, which are designed to preserve smooth structures, perform much better. The exchange of point-wise accuracy for smooth structure preservation by these methods can be seen by comparing Fig. 10, showing the log-difference of the 2D slice, and Fig. 18, showing the pure light fluid isosurface. Surfacelets are able to provide the smoothest reconstruction at 3% coefficients with the lowest difference to the original visualization.

4.3.6. Pure light fluid isosurface area, mean and Gaussian curvatures

In addition to qualitative results by visually comparing the smoothness of isosurfaces, quantitative assessments can also be made using differential geometry tools. In this section, the surface area and average mean and Gaussian curvatures for the light and fully mixed isosurfaces obtained using the largest 3% of the coefficients are compared with the original field, while in the next section the surface signature is discussed. The averages are calculated as sums over all points on the surface divided by the number of points. The definitions and formulas used are given in section 4.1.1. Table 6 shows the ratios of this quantities to the original dataset for all methods considered. Values closer to 1 indicate better quality of the reconstruction. In general, the surface area is relatively well captured by all methods (with larger departures from the original for the light fluid isosurface), while the average mean and Gaussian curvatures can be significantly under- or over-predicted without an overall winner. Even though surfacelets and curvelets show smoothest surfaces, they can still miss both the surface area and average curvatures.

At the time during the flow evolution when the methods are compared, there is relatively little pure light fluid left, while the fully mixed fluid occupies a larger volume. Consequently, the surface area of the mixed fluid isosurface is overall better captured by all methods. The large blockiness in the Haar wavelets representation of constant property surfaces can be clearly seen in the large departures from 1 in all entries of table 6. For the pure light fluid isosurface, again, within each

ACCEPTED MANUSCRIPT

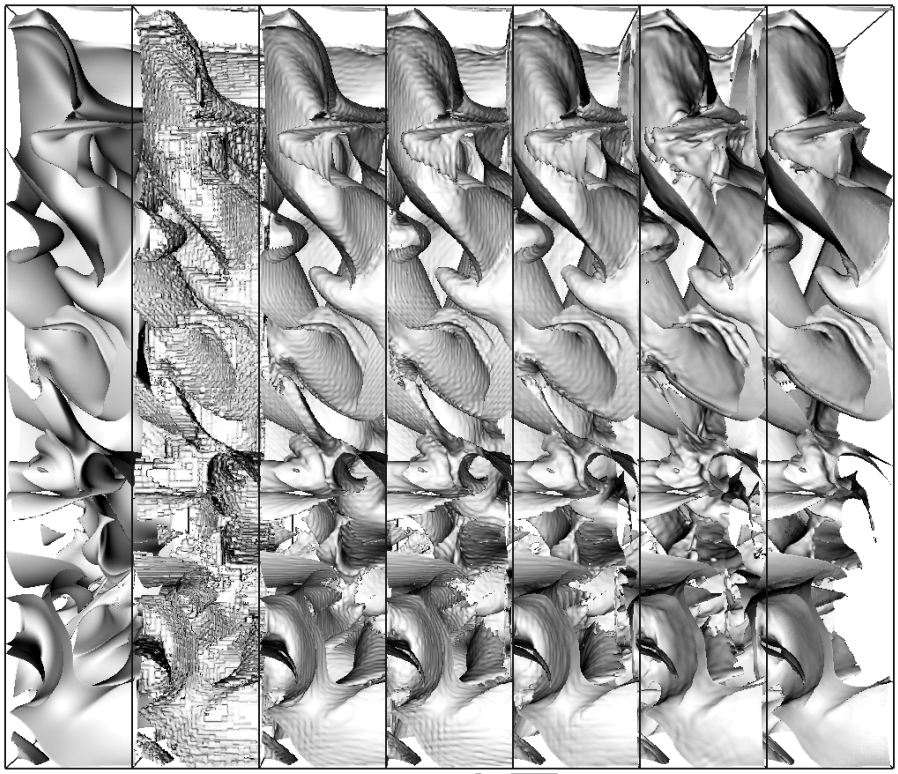


Figure 18: **Density Isosurface comparison**. Isosurfaces representing the pure light fluid are constructed using the largest 3% coefficients. From left to right: original, Haar, cubic B-spline, Daubechies-5, Dual-tree wavelets, curvelet, and surfacelet transforms. Haar wavelet isosurfaces are extremely blocky. Surfacelets and curvelets provide the smoothest representation, with cubic B-splines, Daubechies, and Dual-tree having noticeable ripples in their isosurfaces.

class, the middle members provide the best quality reconstruction. Thus, lower families tend to not have enough representation to capture all the folds in the original surfaces, while the higher families introduce oscillations which affect the curvature of the surface. Such a trend is not seen for the fully mixed fluid isosurface, which has more folds and wrinkles due to the larger volume encompassed. In this case, the oscillations introduced by the higher order methods may help capture more of the details of the surface when they overlap to the surface wrinkles; however, such a match cannot not lead to a consistent trend.

ACCEPTED MANUSCRIPT

	Pure light fluid			Fully mixed fluid		
Method	Surface area	Mean curvature	Gaussian curvature	Surface area	Mean curvature	Gaussian curvature
Haar	1.346	0.523	219.030	0.991	0.627	11.483
Linear B-spline	1.173	0.298	0.577	0.983	0.914	0.655
Quadratic B-spline	1.143	0.453	0.829	0.971	0.880	0.835
Cubic B-spline	**1.014**	**1.192**	1.078	0.995	0.981	0.625
Quartic B-spline	0.954	2.867	1.346	0.995	0.947	0.682
Quintic B-spline	0.988	2.607	1.259	0.994	0.968	0.661
Daubechies-3	1.080	1.745	0.094	0.991	1.020	0.590
Daubechies-5	1.024	2.710	1.356	0.994	0.264	0.605
Daubechies-7	1.017	3.985	1.505	0.991	1.060	0.609
Coiflet-6	1.199	1.355	1.062	**1.001**	0.960	0.531
Coiflet-12	1.053	2.164	**1.051**	0.998	0.970	0.613
Coiflet-18	1.015	3.073	1.250	**0.999**	0.903	0.574
Dual-Tree	0.987	2.643	1.085	0.965	0.900	0.653
Surfacelet	0.921	2.862	1.994	0.957	**0.995**	**0.931**
Curvelet	1.183	2.005	1.126	0.920	0.891	0.574

Table 6: **Isosurface area and average curvatures**. Methods closest to 1.0 are in bold. Ratios of surface area and average mean and Gaussian curvatures for the pure light and fully mixed fluids isosurfaces reconstructed using 3% coefficients to the original dataset.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

4.3.7. Surface signature of the pure light, pure heavy and mixed fluids isosurfaces

The surface signature can provide details on the structure of the surface [5]. Following the definition given in section 4.1.1, Fig. 19 compares the surface signature properties of the original dataset and a subset of the methods.

There is a large disparity between the original dataset, the Haar wavelets and the remaining schemes. To start, the original surface signatures characterize all coarse and fine features in the dataset. The dark region near the center of the 2D histogram of the original surface signature corresponds to the fine features of the isosurfaces. None of the methods is able to reproduce the fine portion of the original dataset. In general, all higher order methods over-predict the amount of small scale features of the pure fluids isosurfaces and under-predict this amount for the mixed fluid isosurface. This is consistent with the representation of the average mean and Gaussian curvatures discussed in the previous section. Again, at the time in the flow evolution when the comparison are made, the amount of mixed fluid is much larger than either pure fluids and, consequently, the mixed fluid isosurface presents more folds and wrinkles. The oscillations introduced by the representation methods discussed will over-predict these features for the pure fluids isosurfaces and under-predict for the mixed fluid isosurface. In addition, the departure from the original dataset can not be attributed to Gaussian white noise. These results underline the specific problem of representing turbulence fields with non-physics based, universal basis functions. While in many problems the details of the small scales are not important, there are instances, e.g. reaction fronts in non-premixed combustion, when the lack of accurate representation of the scalar fields can lead to serious errors. For these cases, physics based sub-grid modeling can remedy the situation; however, finding optimal such models is still an outstanding open question.

4.4. Performance and memory usage

Performance is evaluated by measuring the runtime of each transform during its decomposition stage for a single scalar field from the dataset. Only the time taken during decomposition is presented as the synthesis operation generally takes the same amount of time. As before, the single-threaded variant of all methods is evaluated. Memory usage is defined as the storage space consumed for the decomposed coefficients and does not include the space used to hold the input

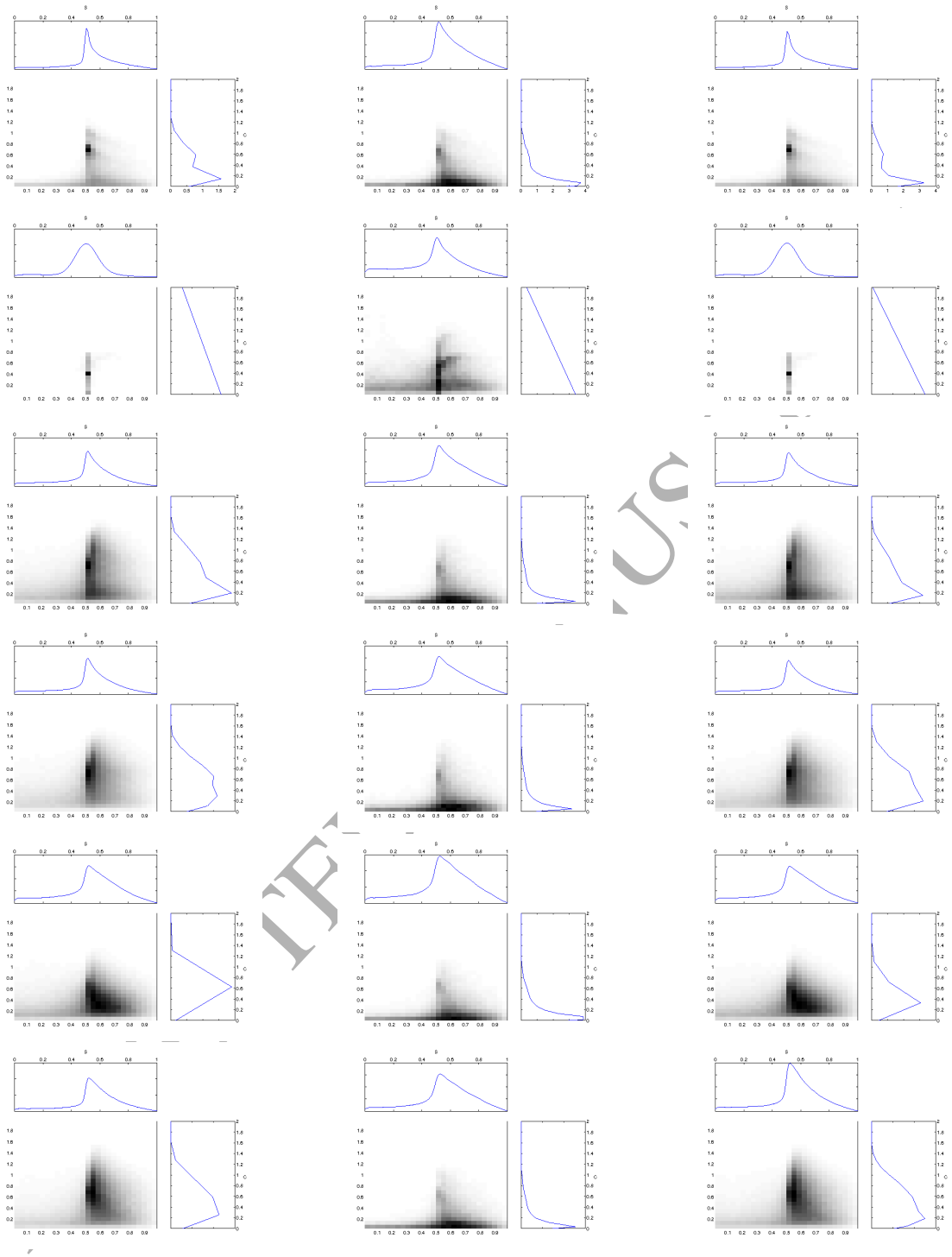


Figure 19: **Surface signatures**. Surface signatures described as a joint probability density function of the shape index, S, and curvedness, C, together with their marginal PDFs. From top to bottom: original, Haar, cubic B-spline, Daubechies-5 wavelets, curvelets and surfacelets. From left to right: pure light, fully mixed and pure heavy fluid isosurfaces.

ACCEPTED MANUSCRIPT

Method	Compute (s)	Memory (MB)	Efficiency	Method	Compute (s)	Memory (MB)	Efficiency
Haar	16.390	1065	1.2211	Daubechies-7	25.342	1125	1.7156
Linear B-spline	17.282	1072	1.2474	Coiflet-6	20.130	1059	1.4528
Quadratic B-spline	17.601	1080	1.2891	Coiflet-12	25.221	1113	1.7333
Cubic B-spline	22.320	1092	1.5611	Coiflet-18	30.045	1152	1.9699
Quartic B-spline	24.304	1106	1.6703	Dual-Tree	79.667	4096	1.4839
Quintic B-spline	30.373	1179	1.9914	Surfacelet	31.666	4928	**0.6891**
Daubechies-3	18.866	1057	1.3616	Curvelet	77.660	8214	1.4041
Daubechies-5	21.863	1090	1.5291				

Table 7: **Performance**. Compute time in seconds, memory usage in megabytes (MB) and Efficiency as compute time per coefficient in seconds * 10^{-7} . The simplest methods are faster to compute compared to the higher order methods. Surfacelets have the best efficiency by extracting the most coefficients for their measured compute time.

of the original dataset.

The results in Table 7 represent the compute time required only for the forward transform of a targeted 3% coefficients, averaged over the series of five runs. The amount of effort required in extracting the strongest 3% coefficients is evaluated. During this evaluation, it is considered that the multi-band transforms can be performed without the need to extract all coefficients. The term 'Efficiency' is defined as the amount of effort needed to compute one coefficient. Effort is a function of time taken over the total number of coefficients extracted. Due to their simplistic nature, the lowest-order wavelets perform the fastest and require the lowest amount of storage. The increase in wavelet complexity and family order coincides with the increase of compute time as well as number of coefficients extracted and memory usage. Surfacelets provide the best of compromises in terms of compute time per coefficient, overall compute time, and the ability to extract multiple bands from the data at a comparable cost to other methods. The curvelet transform takes the longest amount of time to compute due to the extraction of many coefficients. When requesting the same number of bands as surfacelets, curvelets take much longer to compute and extract more coefficients.

In terms of efficiency, the surfacelet transform is the clear winner. Although the curvelet transform may take long to compute, it does performs remarkably well for the amount of coefficients extracted compared to the simpler B-splines and Daubechies wavelets. The curvelet transform

ACCEPTED MANUSCRIPT

Test Type	Best Method	Runner-up
Velocity PDF	Cubic B-spline	Daubechies-5
Vorticity PSNR and MSE	Cubic B-spline	Quadratic B-spline
Vorticity PDF	Cubic B-spline	Quadratic B-spline
Strain rate tensor	Quadratic B-spline	Cubic B-spline
Density visualization	Surfacelets	Cubic B-spline
Density PSNR and MSE	Cubic B-spline	Quadratic B-spline
Density PDF	Cubic B-spline	Daubechies-5
Density power spectrum	Dual-Tree	Quintic B-Spline
Isosurface Visualization	Surfacelets	Curvelets
Curvature quantities	Cubic B-spline	Surfacelets
Surface signatures	All methods except Haar	—-
Performance	Surfacelets	Haar

Table 8: **Summary**. Compilation of the best respective methods for each test performed.

implementation provided by the respective authors contains a message passing interface (MPI), out-of-core C++ implementation where an almost linear speedup is observed in relation to the number of additional processors used when increased by powers of two. Although costly in performance, the MPI implementation makes the curvelet transform readily usable for large scientific computing applications where distributed compute nodes can be used to offset their high compute cost and memory requirements.

5. Conclusions

We have compared the effectiveness of various multi-resolution representation methods, including B-spline, Daubechies, Coiflet and Dual-tree wavelets, curvelets and surfacelets, to evaluate these methods' ability to capture the structure of fully developed turbulence using a truncated set of coefficients. The methods were tested by considering quantities pertaining to both velocity and active scalar (density) fields and their derivatives, spectra, and the properties of constant density surfaces. Previous comparisons related to such decompositions of turbulent fields were restricted to orthogonal and biorthogonal (Harten) wavelets as applied to velocity and enstrophy PDFs, the kinetic energy spectrum and the PDF of helicity and Lamb vector [24] and the curvelet transform as a multi-scale analysis tool of turbulence [4, 5]. While CVS-type decompositions have

ACCEPTED MANUSCRIPT

been mainly used to investigate the coherent/incoherent decomposition of flow, here the decomposition is performed at the level of the primary variables, velocity and density. This approach is not entirely new, for example, it is the base of the SCALES method [9] and has also been used for the analysis of a passive scalar [5]. The results presented address a more complex flow and encompass more representation methods, with more quantities related to the velocity field (e.g., the state of the deviatoric strain rate tensor, which may be a better measure of the local flow coherence/incoherence than the velocity and vorticity PDFs) and, for the first time, quantities related to an active scalar field and its constant property surfaces.

In addition, comparisons between the algorithms are given in terms of performance, accuracy, and compression properties. The results provide useful information for multi-resolution analysis of turbulence, coherent feature extraction, compression for large datasets handling, as well as simulations algorithms based on multi-resolution methods.

A list of recommended methods for each test is provided in Table 8. In general, any method is superior to Haar wavelets as well as the constant members of the wavelet classes considering our series of tests. On the other hand, the results show that increasing the family order above a certain value is not always the ideal solution towards higher accuracy. At 3% coefficients, large structures within the flow are well preserved and at that percentage it is not necessary to go to very high order methods. Since the results are very similar across the higher orders of B-splines, it is not recommended to go above cubic B-splines, sacrificing more compute time and questionable accuracy gains. Daubechies wavelets are generally overshadowed by B-splines so their use is not recommended unless orthogonal properties must be preserved. Although Coiflets have a few advantages to Daubechies in derived surface quantities and preservation of kinetic energy, their representations of curved surfaces and scalar quantities are not ideal. Dual-tree wavelets contain properties that make them useful for curved surface preservation and energy separation, especially above 3% coefficients, but at the cost of high redundancy, point-wise accuracy, and a large compute cost. Based on their overall performance, the cubic B-spline wavelet is recommended as the general method for the turbulence data applications considered here. Surfacelets and curvelets have specific applications and advantages where they are able to identify specific features at different scales in turbulence, taking full advantage of the multi-scale interface of these methods

ACCEPTED MANUSCRIPT

but further analysis is required. Both surfacelets and curvelets provide superior representation of smooth surfaces compared to any other method for general applications in turbulence. Surfacelets are recommended over curvelets since they are much more efficient in computation, reconstruct more accurately at the same number of coefficients, and capture curved surfaces closest to the original data even compared to all the methods tested.

The selection process described in this paper should be useful in several areas, including multi-resolution analysis of turbulence, coherent feature extraction, compression for large datasets handling, as well as simulations algorithms based on multi-resolution methods. While some of the algorithms discussed have already been used for simulations algorithms based on multi-resolution methods and others for multi-resolution analysis of turbulence, this survey offers a comprehensive view of most of the methods which are candidates for multi-resolution analysis and computation of turbulence. In addition, with the continuous increase in the computational platform speed and size, we would like to stress the emerging importance of compression algorithms for large dataset handling. For example, multi-resolution visualization for large datasets [38] and simulation applications for in-situ analysis and visualization can be facilitated through the use of these algorithms. Projects such as the JHTDB [11] can be improved to support multi-resolution analysis and visualization to both reduce storage and transmission costs as well as create faster data visualizations of interesting structures. Opportunities for remote visualization are also available through the application of these methods at the data and image level.

Acknowledgements

We would like to thank our colleagues John Patchett and Curt Canada in the Computer, Computational, and Statistical Sciences division for their feedback and support, as well as the Institute for Next Generation Visualization and Analysis (INGVA). Los Alamos National Laboratory is operated by Los Alamos National Security, LLC for the US Department of Energy National Nuclear Security Administration under Contract No. DE-AC52-06NA25396. Computational resources were provided by the Los Alamos National Laboratory Institutional Computing Program.

ACCEPTED MANUSCRIPT

Appendix A. Wavelet construction and their signals

Wavelet signals in general have several vanishing moments. This property allows for a sparse but accurate representation of an input dataset with only a small number of coefficients. A signal is decomposed through multiple steps involving ”folds” at the largest scale until it reaches the smallest data scale. For the 3-D Wavelet Transform, an input data of size X (length) by Y (width) by Z (height) is processed by utilizing the 1-D signal decomposition step in the multiple dimensions. After a single iteration of the 1-D decomposition for each dimension, the result consists of a series of seven high-pass coefficients and one set of pure low-pass coefficients shown in Fig. A.20. Any region that has full or partial high-pass coefficients such as W_{HLL} and W_{HHH} can be used to describe the original data at that resolution. The low-pass coefficients W_{LLL} are recursively computed by halving the size of each dimension per iteration until the data can no longer be halved further. This method can be described as a top-down approach, where the small-scale features in the data are captured initially and for each iteration, larger features are described. The finest coefficients, smallest in magnitude are extracted first and in the end, the coarsest, largest in magnitude coefficients are extracted.

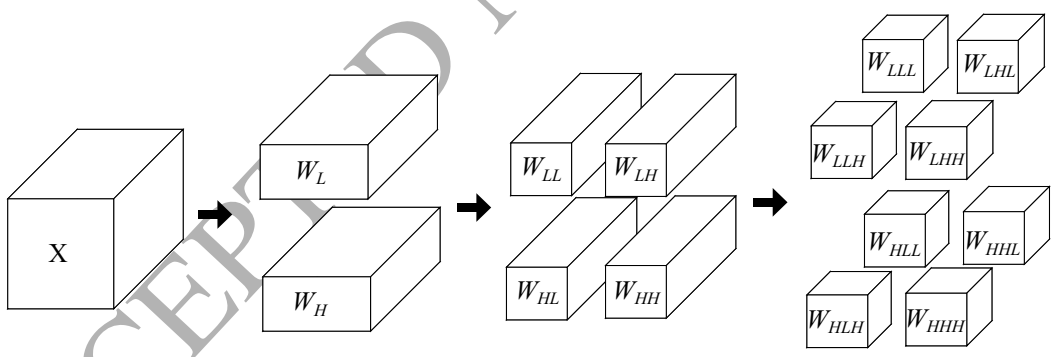


Figure A.20: **Signal Decomposition**. Decomposition of a wavelet for dataset X using a single step in three dimensions with the remaining coarse (W_L) and fine (W_H) coefficients. For each subsequent step, the low-pass coefficients W_{LLL} are iterated upon.

The continuous representations of the wave signals we use here can be found in Fig. A.21.

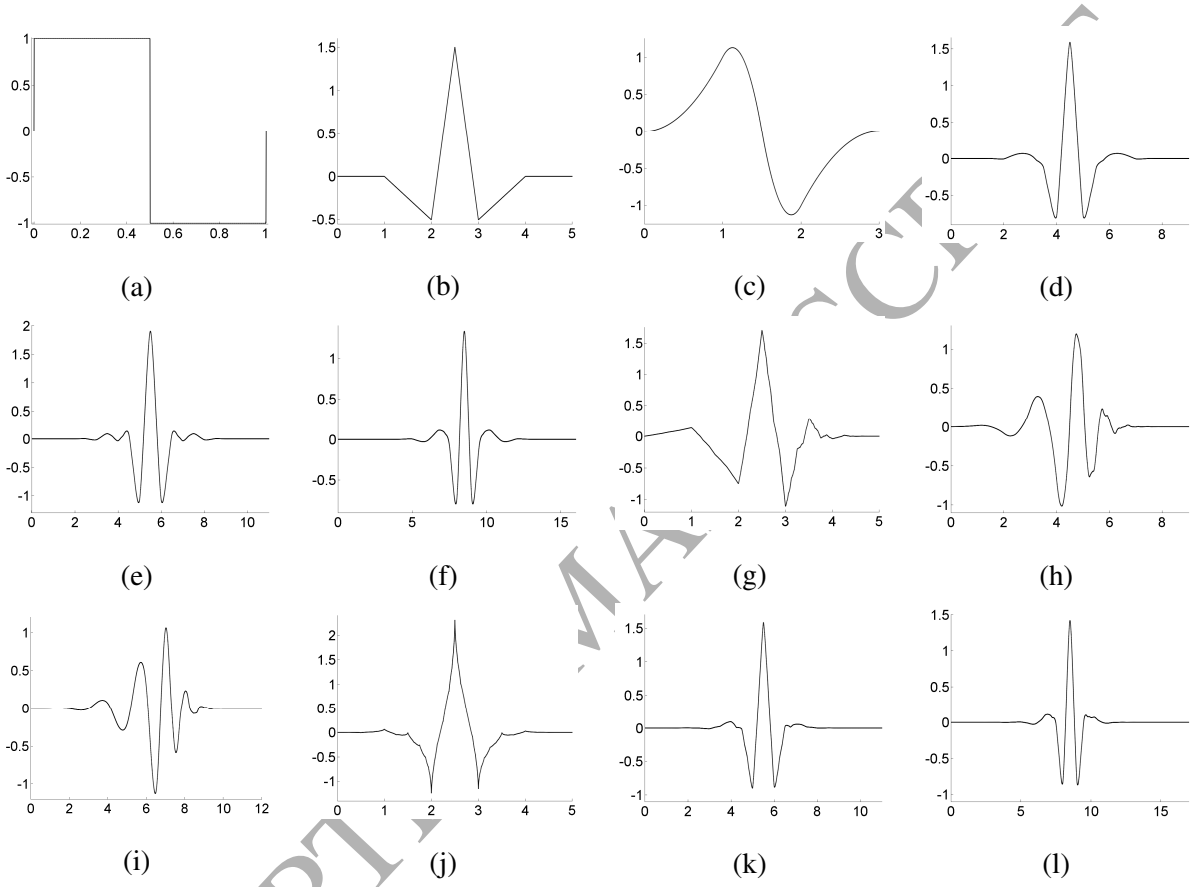
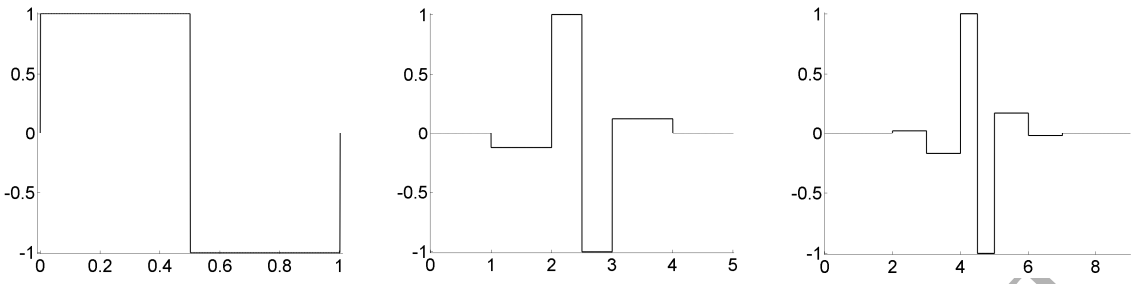


Figure A.21: **Wavelet mother reconstruction signals.** As the family order increases, signals in the frequency domain become more complex creating more vanishing moments. The following signals represent biorthogonal wavelets: (a) Haar, constant (b) linear B-spline (c) quadratic B-spline (d) cubic B-spline (e) quartic B-spline and (f) quintic B-spline. Orthogonal signals are also shown in comparison as (g) Daubechies-3 (h) Daubechies-5 (i) Daubechies-7 (j) Coiflet-6 (k) Coiflet-12 and (l) Coiflet-18.

ACCEPTED MANUSCRIPT



(a) (b) (c)

Figure A.22: **Constant biorthogonal B-splines and control points.** As the number of control parameters increases, signals become more complex. The reconstruction wavelet functions presented are subsets of the constant biorthogonal wavelet family using (a) 1 control point, Haar (b) three control points, Harten and (c) five control points

Appendix A.1. B-spline wavelet families

When increasing the family order for B-spline wavelets, an improvement is expected in the representation of non-rigid surfaces. The change in signal behavior by increasing the family order can be seen in Fig. A.21. An increase in the number of control points within each family generally introduces a large number of oscillations in the data and reduces the overall quality of a reconstruction. The effect of increased control points on the constant family of signals can be seen in Fig. A.22. As a result, the B-spline wavelets chosen for this paper were tested with the least number of control parameters per family. Rather than increasing the number of control point parameters, an increase in family order enhances the complexity of the signal and improves the quality of a reconstruction. The discrete versions of these signals were used in form of filters as described by [36]. By considering a sinusoidal signal instead of the top-hat, a wavelet transform can capture superior directionality of a dataset.

References

[1] E. Candes, L. Demanet, D. Donoho, and L. Ying. Fast discrete curvelet transforms, 2005.

[2] Y. M. Lu and M. N. Do. Multidimensional directional filter banks and surfacelets. *IEEE Trans. Image Process*, pages 918–931, 2007.

[3] M. Farge, K. Schneider, G. Pellegrino, A. Wray, and R. Rogallo. Cvs decomposition of 3d homogeneous turbulence using orthogonal wavelets. *Proceedings of the Summer Program*, page 305, 2000.

ACCEPTED MANUSCRIPT

[4] J. Ma, M. Yousuff Hussaini, O. V. Vasilyev, and F. Le Dimet. Multiscale geometric analysis of turbulence by curvelets. *Physics of Fluids*, 21(7), 2009.

[5] I. Bermejo-Moreno and D. I. Pullin. On the non-local geometry of turbulence. *Journal of Fluid Mechanics*, 603:101–135, 2008.

[6] O. V. Vasilyev and S. Paolucci. A fast adaptive collocation algorithm for multidimensional PDEs. *J. Comput. Phys.*, 138:16–56, 1997.

[7] O. V. Vasilyev and C. Bowman. Second generation wavelet collocation method for the solution of partial differential equations. *J. Comput. Phys.*, 165:660–693, 2000.

[8] S. Paolucci, Z. J. Zikoski, and D. Wirasaet. WAMR: An adapative method for the simulation of compressible reacting flow. Part I. accuracy and efficiency of the algorithm. *J. Comput. Phys.*, 272:814–841, 2014.

[9] D. Goldstein and O. V. Vasilyev. Stochastic adaptive large eddy simulation method. *Phys. Fluids*, 16:2497, 2004.

[10] D. E. Goldstein, O. V. Vasilyev, and N. K. Kevlahan. Cvs and scales simulations of 3d isotropic turbulence. *J. Turbul*, 2005.

[11] E. Perlman, R. Burns, Y. Li, and C. Meneveau. Data exploration of turbulence simulations using a database cluster. In *Proc. of the ACM/IEEE Conference on Supercomputing*, 2007.

[12] D. Livescu, J. Mohd-Yusof, M. R. Petersen, and J. W Grove. CFDNS: A computer code for direct numerical simulation of turbulent flows. Technical report, Los Alamos National Laboratory, 2009. LA-CC-09-100.

[13] D. Livescu and J. R. Ristorcelli. Buoyancy-driven variable-density turbulence. *J. Fluid Mech.*, 591:43–71, 2007.

[14] D. Livescu and J. R. Ristorcelli. Variable-density mixing in buoyancy-driven turbulence. *J. Fluid Mech.*, 605:145–180, 2008.

[15] D. Livescu. Numerical simulations of two-fluid turbulent mixing at large density ratios and applications to the rayleightaylor instability. *Phil. Trans. R. Soc. A*, 371:20120185, 2013.

[16] D. Livescu, C. Canada, K. Kalin, R. Burns, IDIES staff, and J. Pulido. Homogeneous buoyancy driven turbulence data set. Technical report, Los Alamos National Laboratory, 2014. LA-UR-14-20669, available at http://turbulence.pha.jhu.edu/docs/README-HBDT.pdf.

[17] I. Daubechies. Ten lectures on wavelets. *PA: SIAM*, 1992.

[18] W. Sweldens. The lifting scheme: A construction of second generation wavelets, 1997.

[19] M. Farge. Wavelet transforms and their applications to turbulence. *Annual Review of Fluid Mechanics*, 24(1):395–458, 1992.

[20] G. Beylkin, R. R. Coifman, and V. Rokhlin. Fast wavelet transforms and numerical algorithm. *Pure Applied Mathematics*, 44:141–183, 1991.

[21] I. Daubechies. Orthonormal bases of compactly supported wavelets ii. variations on a theme. *SIAM Journal on Mathematical Analysis*, 24(2), 1993.

ACCEPTED MANUSCRIPT

[22] J. Tian and R. O. Wells Jr. Vanishing moments and wavelet approximation. Technical report, Computational Math. Lab., Rice Univ, 1995.

[23] E. Deriaz, M. O. Domingues, V. Perrier, K. Schneider, and M. Farge. Divergence-free wavelets for coherent vortex extraction in 3d homogeneous isotropic turbulence. *ESAIM: Proc.*, 16:146–163, 2007.

[24] O. Roussel, K. Schneider, and M. Farge. Coherent vortex extraction in 3d homogeneous turbulence: comparison between orthogonal and biorthogonal wavelet decompositions. *Journal of Turbulence*, page N11, 2005.

[25] A. Cohen, I. Daubechies, and J. C. Feauveau. Biorthogonal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, 45(5):485–500, 1992.

[26] P. González, J. C. Cabaleiro, and T. F. Pena. Parallel computation of wavelet transforms using the lifting scheme. *J. Supercomput.*, 18(2):141–152, February 2001.

[27] M. Bertram, M. A. Duchaineau, B. Hamann, and K. I. Joy. Generalized b-spline subdivision-surface wavelets for geometry compression. *IEEE Transactions on Visualization and Computer Graphics*, 10(3):326–338, May 2004.

[28] N.G. Kingsbury. Complex wavelets for shift invariant analysis and filtering of signals, May 2001.

[29] I.W. Salesnick, R.G. Baraniuk, and N.G. Kingsbury. The dual-tree complex wavelet transform, 2005.

[30] R. Nguyen van yen, M. Farge, and K. Schneider. Wavelet regularization of a fourier-galerkin method for solving the 2d incompressible euler equations, 2009.

[31] R. M. Pereira, R. Nguyen van yen, M. Farge, and K. Schneider. Wavelet methods to eliminate resonances in the galerkin-truncated burgers and euler equations, 2013.

[32] R. H. Bamberger and M. J. T. Smith. A filter bank for the directional decomposition of images: Theory and design. *IEEE Trans. Signal Processing*, 40(4):882–893, 1992.

[33] M. N. Do and M. Vetterl. The contourlet transform: An efficient directional multiresolution image representation. *IEEE Trans. Image Processing*, 14(12):2091–2106, 2005.

[34] Y. Yang, D. I. Pullin, and I. Bermejo-Moreno. Multiscale geometric analysis of lagrangian structures in isotropic turbulence. *Journal of Fluid Mechanics*, 654:233–270, 2010.

[35] E. J. Stollnitz, T. D. DeRose, and D. H. Salesin. *Wavelets for computer graphics: theory and applications*. Morgan Kaufmann Publishers Inc, San Francisco, CA, USA, 1996.

[36] M. J. Shensa. The discrete wavelet transform: wedding the a trous and mallat algorithms. *IEEE Trans. on Signal Processing*, 40(10):2464–2482, 1992.

[37] M. Bertram. *Multiresolution Modeling for Scientific Visualization*. PhD thesis, Center for Image Processing and Integrated Computing, University of California, Davis, aug 2000.

[38] D. F. Wiley, H. R. Childs, B. Hamann, K. I. Joy, and N. L. Max. Best quadratic spline approximation for hierarchical visualization. In *Data Visualisation 2002, Eurographics Association, Aire-la-Ville, Switzerland*, pages 133–140, 2002.

ACCEPTED MANUSCRIPT

[39] J. Woodring, S. M. Mniszewski, C. M. Brislawn, D. E. DeMarle, and J. P. Ahrens. Revisiting wavelet compression for large-scale climate data using jpeg 2000 and ensuring data precision. In David Rogers and Cludio T. Silva, editors, *LDAV*, pages 31–38. IEEE, 2011.

[40] D. L. Donoho. Unconditional bases are optimal bases for data compression and for statistical estimation. *Applied and Computational Harmonic Analysis*, pages 100–115, 1993.

[41] T. S. Lund and M. M. Rogers. An improved measure of strain state probability in turbulent flows. *Physics of Fluids*, 6(5):1838–1847, 1994.

[42] V. Maz'ja. *Sobolev spaces*. Springer-Verlag, Berlin, Germany, 1985.

[43] J. Ryu and D. Livescu. Turbulence structure behind the shock in canonical shock-vortical turbulence interaction. *J. Fluid Mech.*, 756:R1, 2014.

[44] A. E. Perry and M.S. Chong. A description of eddying motions and flow patterns using critical-point concepts. *Annu. Rev. Fluid Mech.*, 19:125, 1987.

[45] S. J. Reckinger, D. Livescu, and O. V. Vasilyev. Comprehensive numerical methodology for Direct Numerical Simulations of compressible Rayleigh-Taylor instability. *under review J. Comput. Phys.*

[46] D. L. Donoho and I. M. Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81:425–455, 1994.

ACCEPTED MANUSCRIPT