

COMPRESSIVE SAMPLING FOR ENERGY SPECTRUM ESTIMATION OF TURBULENT FLOWS*

GUDMUNDUR F. ADALSTEINSSON[†] AND NICHOLAS K.-R. KEVLAHAN[‡]

Abstract. Recent results from compressive sampling (CS) have demonstrated that accurate reconstruction of sparse signals often requires far fewer samples than suggested by the classical Nyquist–Shannon sampling theorem. Typically, signal reconstruction errors are measured in the ℓ^2 norm and the signal is assumed to be sparse or compressible. Our spectrum estimation by sparse optimization (SpESO) method uses a priori information about isotropic homogeneous turbulent flows with power law energy spectra and applies the methods of CS to one- and two-dimensional turbulence signals to estimate their energy spectra with small logarithmic errors. SpESO is distinct from existing energy spectrum estimation methods which are based on sparse support of the signal in Fourier space. SpESO approximates energy spectra with an order of magnitude fewer samples than needed with Shannon sampling. Our results demonstrate that SpESO performs much better than lumped orthogonal matching pursuit, and as well as or better than wavelet-based best M -term or $M/2$ -term methods, even though these methods require complete sampling of the signal before compression.

Key words. compressive sampling, turbulence, energy spectrum, wavelets, optimization

AMS subject classifications. 76F05, 65F22, 65T60

DOI. 10.1137/140966216

1. Introduction. Sampling and storage of signals becomes challenging for high-wavenumber or high-dimensional signals if the Nyquist–Shannon sampling theorem is followed strictly. The theory of compressive sampling (CS) provides a rigorous framework in which to accurately reconstruct a signal from a few nonadaptive (random) projections, provided it is sufficiently sparse or compressible in some basis [9, 18, 11]. Other signal models considered in the CS literature are not directly relevant to the methods in this paper.

Consider the discrete signal $u \in \mathbb{R}^{n^d}$ of length $N = n^d$ in d dimensions. The traditional fixed-rate sampling, hereafter referred to as *Shannon sampling*, of u is inefficient if the coefficients \hat{u} of u in an orthogonal basis are sufficiently compressible. Shannon sampling is especially wasteful if we are interested only in a particular low-dimensional property of the signal, such as the one-dimensional (1-D) energy spectrum of a two-dimensional (2-D) or three-dimensional (3-D) data set.

This paper focuses on the reconstruction of energy spectra of homogeneous isotropic turbulent flows from a minimal number of samples. A turbulent flow is characterized by a nondimensional number, the Reynolds number Re , which is the ratio of inertial terms to viscous terms in the Navier–Stokes equations governing the flow. Flows become turbulent when Re exceeds a certain threshold (typically $\sim 10^3$), and industrial and natural turbulent flows have very large Reynolds numbers ($\sim 10^5$ – 10^{12}). The minimum length scale of a turbulent flow, the Kolmogorov scale η , decreases with

*Submitted to the journal’s Computational Methods in Science and Engineering section April 23, 2014; accepted for publication (in revised form) April 13, 2015; published electronically June 11, 2015.

<http://www.siam.org/journals/sisc/37-3/96621.html>

[†]School of Computational Science and Engineering, McMaster University, Hamilton, ON L8S 4K1, Canada (adalstgf@mcmaster.ca).

[‡]Department of Mathematics and Statistics, McMaster University, Hamilton, ON L8S 4K1, Canada (kevlahan@mcmaster.ca).

increasing Reynolds number Re like $\eta \propto Re^{-3/4}$ [22], and the number of spatial samples required by the sampling theorem in d dimensions is $N \propto \eta^{-d}$. Therefore, the total number of samples needed to characterize a turbulent flow increases very quickly with Reynolds number: like $Re^{9/4}$ in three dimensions and $Re^{3/2}$ in two dimensions. Thus, straightforward application of Shannon sampling requires huge amounts of regularly sampled data ($\sim 10^{11}$ – 10^{27}) to estimate the complete 1-D energy spectrum of a 3-D turbulent flow.

Although they contain a wide and continuous range of scales, turbulence signals are nevertheless moderately compressible in wavelet bases because they are intermittent in space and time. This compressibility is the basis of several adaptive wavelet techniques for simulating turbulence; see, e.g., [23, 20, 1, 16]. Turbulence researchers are often interested in reconstructing Fourier energy spectra from spatial measurements, but spectrum estimation is not a well-developed area of CS.

Since the range in wavenumber space of the 1-D energy spectrum of u is proportional to $\eta^{-1} \propto Re^{3/4}$, there is definitely room for improved sampling strategies. Even for 1-D signals, such as hot-wire measurements, it should be possible to accurately characterize the energy spectrum using fewer samples than required for the usual Shannon sampling.

In order to accurately estimate the 1-D energy spectra of signals with a very large and continuous range of active length scales, we propose a new method that uses a priori information about the signal, such as the structure and scaling of wavelet coefficients, isotropy, and power law behavior of the energy spectrum. We show that our method is able to approximate energy spectra with an order of magnitude fewer samples than needed with Shannon sampling.

We introduce notation and give a brief introduction to CS in section 2 before we define our problem and introduce two measurement matrix types used in our experiments. In section 3.2 we introduce the relevant wavelet transforms and their application to turbulence, and finally present our spectrum estimation by sparse optimization (SpESO) algorithm for estimating energy spectra. Section 5 verifies the method by applying it to a set of representative test cases: 1-D hot-wire turbulence data, 1-D synthetic power law data, 2-D numerical simulation turbulence data, and 2-D synthetic power law signals.

In related work, variants of CS have been developed to estimate spectra and other properties of signals, but in different contexts which do not apply in our case. In [15] linear functions of signals were estimated by fast operators. Energy spectra, however, are nonlinear functions of signals. Sparse and locally supported 2-D spectra were estimated in [38], but turbulence is not sparse in Fourier space. Similarly, [25, 4] put some sparsity constraints on their power spectrum estimation. Bands of power spectra are estimated on a linear scale from nonuniform samples in [26]. In [2] the 2-D spectrum itself is sampled and approximated to reduce computational time in spectroscopy. General nonlinear optimization problems for CS are considered in [6]. However, the iterative algorithm proposed is impractical in our case as it requires expensive high-dimensional gradients to be computed at each iteration.

2. Compressive sampling for large signals. In this paper we assume that the turbulent flow is provided as a single component of a turbulent velocity vector field as a discrete sequence $u \in \mathbb{R}^N$. Mathematically, of course, the flow is more accurately described as a vector field of velocity or vorticity defined on a 3-D spatial domain. However, assuming the flow is band-limited in wavenumber, the Nyquist–Shannon sampling theorem allows us to represent it as a sequence of discrete values. The

measurement matrices discussed later are discrete approximations of linear operators in continuous space. We represent 2-D signals of dimension $n \times n$ as vectors of length $N = n^2$.

We first decompose u as a linear combination of vectors in a basis $\Phi \in \mathbb{R}^{N \times N}$,

$$(2.1) \quad u = \Phi \hat{u} = \sum_i \hat{u}_i \phi_i,$$

where \hat{u} are the expansion coefficients and ϕ_i are the basis vectors. A signal u is said to be B -sparse in basis Φ if $|\text{supp}(\hat{u})| = B < N$, where $|\cdot|$ denotes cardinality and $\text{supp}(x) = \{i : x_i \neq 0\}$ is the support.

The signal u is called *compressible* in the basis Φ if it has ordered coefficients $|\hat{u}|_{(1)} \geq \dots \geq |\hat{u}|_{(N)}$ that satisfy the inequality $|\hat{u}|_{(n)} \leq Cn^{-s}$ for $s > 0$ and a constant C [10]. The *best B -term approximation* in an orthonormal basis, u_B , is an approximation with all but the B largest terms of \hat{u} zero. Many signals are highly compressible in a wavelet basis [17] since wavelet basis functions are self-similar and are localized in both position and scale. If the signal is compressible, then the error in the best B -term approximation is $\|u_B - u\| = \mathcal{O}(B^{-s+1/2})$.

The central idea of CS (see, e.g., [9, 18, 11, 12]) is that a few linear nonadaptive (e.g., random) measurements of a signal are sufficient to accurately reconstruct a signal if that signal is compressible in some basis. Note that the measurement scheme (e.g., random samples) and the sparsity system (e.g., a wavelet basis) must be mutually incoherent in the sense of having a sufficiently small maximum inner product between the basis vectors of the measurement scheme and the sparsity system.

Let $A \in \mathbb{R}^{M \times N}$ be a *measurement matrix*, let $g \in \mathbb{R}^M$ be the *compressed samples*, and assume $M < N$. The measurement scheme is defined by the underdetermined system

$$(2.2) \quad g = Au.$$

In a slightly different form, with $\Psi = A\Phi$, which we call the *CS-matrix*, we have

$$(2.3) \quad g = \Psi \hat{u},$$

where \hat{u} is assumed to be sparse or compressible in the basis Φ . Under this framework, the minimization problem [12]

$$(2.4) \quad \hat{u}^* = \arg \min_{\hat{h} \in \mathbb{R}^N} \|\hat{h}\|_{\ell^1} \quad \text{s.t.} \quad \Psi \hat{h} = g$$

is proved to accurately approximate, or exactly reconstruct, the original signal. For \hat{u} B -sparse and Ψ having a restricted isometry constant $\delta_{2B} < 4/\sqrt{41}$ (see (4.1)), the reconstruction is exact [21]. (A star superscript, u^* , denotes approximation.) This method is called *basis pursuit* and can be solved via convex optimization. Unfortunately, although they are moderately compressible in wavelet bases, turbulent signals are not compressible enough at high wavenumbers for basis pursuit to give meaningful results for the entire range of wavenumbers.

For problems that are compressible enough, so-called greedy methods can be faster than basis pursuit methods for (2.4). We have found that basis pursuit methods are very slow and produce poor quality spectrum estimates for turbulence signals, and so we have focused on greedy methods. A popular greedy method is iterative *orthogonal matching pursuit* (OMP) [36]. Our estimation algorithm relies heavily

on a multilevel modification of OMP called QOMOMP; see section 3.2. OMP can be generalized easily to estimate more than one coefficient of the signal at a time [39]. The experiments in section 5 use lumped OMP (LOMP) as a comparison to our SpESO method, where the sparsity B_0 is fixed and L_0 coefficients are estimated in each iteration, requiring a total of B_0/L_0 iterations.

The initial CS literature was largely concerned with full random measurement matrices A , which require $\mathcal{O}(NM)$ operations to apply to a vector. Many CS decoding methods require frequent application of A and its transpose. For very large signals the matrix–vector multiplications are very memory and CPU intensive [8], so a full random matrix is not practical. In our method we consider two matrices with fast matrix-free transforms requiring at most $\mathcal{O}(N \log N)$ operations and $\mathcal{O}(N)$ memory to apply.

The first matrix is intended for measurements of 1-D time-dependent signals—such as hot-wire measurements—without requiring the whole signal for every compressed sample: a random finite impulse response (FIR) *filter* [37]. Let the filter coefficients h be compactly supported with support size K . We can then write

$$(2.5) \quad A = R_\Gamma F^* \Sigma F,$$

where $\Sigma = \text{diag}(Fh)$ is a diagonal matrix where the diagonal elements are the Fourier transform of h , and F is the Fourier transform matrix. Here R_Γ restricts the result to an evenly distributed set Γ of length M . This definition of A assumes periodicity, but our implementation zero pads the signal before the convolution to account for nonperiodic boundary conditions. For a downsampling fraction δ_0 and with $1/\delta_0 \in \mathbb{N}$, the number of samples is

$$(2.6) \quad M = \lceil (N + K - 3)\delta_0 \rceil,$$

and the complexity is $\mathcal{O}(KM)$.

A *random convolution* and subsampling is a universal sampling strategy [34] (also known as partial random circulant matrix). Consider now a full vector h and a diagonal matrix $\Sigma = \text{diag}(h)$ which randomizes the phase, i.e., $h_k = e^{i\theta_k}$, where θ_k are independent and identically distributed (i.i.d.) uniformly on $(0, 2\pi)$ such that $F^* \Sigma \in \mathbb{R}^N$. We can again write

$$(2.7) \quad A = R_\Gamma F^* \Sigma F,$$

where R restricts the result to a random set Γ . The complexity of this approach is $\mathcal{O}(N \log N)$. Note that the random convolution matrix A has the property that its right pseudoinverse is the transpose, $A^T (AA^T)^{-1} = A^T$ (i.e., $AA^T = I$ or A is right-orthogonal). We use this matrix (or measurement scheme) for analyzing 2-D data.

3. Energy spectrum estimation of turbulence data.

3.1. Problem formulation. Our problem is challenging because we seek to estimate the energy spectrum $E(k)$ from measurements of u , rather than estimating u directly. This problem is challenging because the quantity to be estimated, $E(k)$, is a nonlinear function of the quantity that is sampled, u . In addition, u is not sparse in Fourier space. Dropping the constant normalization factor, let us define $E(k)$ as

$$(3.1) \quad E(k) = \sum_{k \leq |\mathbf{k}'| < k+1} |\hat{u}[\mathbf{k}']|^2,$$

where \hat{u} is the Fourier transform of u , and we use the convention that E_f is the spectrum of signal f . Since we are interested in estimating a turbulence-like power law energy spectrum over all scales, a logarithmic error norm is appropriate and the problem is stated in general terms as

$$(3.2) \quad \min_{u^* \in \mathcal{U}} \|\log(E_u) - \log(E_{u^*})\|_w,$$

where the logarithm is taken componentwise and the solution has properties specified by $\mathcal{U} \subseteq \mathbb{R}^N$, and the ℓ^2 norm has weight w . Obviously, the term $\log(E_u)$ is an unknown in (3.2). Since we have the samples $g = Au$ we can project the signals and recast the problem as

$$(3.3) \quad \min_{u^* \in \mathcal{U}} \|\log(E_{A^T A u}) - \log(E_{A^T A u^*})\|_w,$$

which can be solved using available data. Assuming a sufficiently small coherence [8] (or restricted isometry property (RIP)) of A , the matrix $A^T A$ is a reasonable approximation of the identity matrix I . At the end of section 3.2 we discuss the case when Au^* is close to g and how the deviation of $A^T A$ from I can be dealt with.

Identity (3.3) is a specific case of the general nonlinear minimization $\min_{x \in \mathcal{U}} f(x)$. There exists a gradient method for this problem with an iterative solver [6]

$$(3.4) \quad x^{k+1} = \mathcal{P}_{\mathcal{U}}(x^k - \mu \nabla f(x^k)).$$

However, this method has the drawback that the computation of the gradient of f is very expensive. In section 3.2 we introduce a more efficient method based on OMP to solve this key problem in energy spectrum estimation.

3.2. Estimation algorithm. We now introduce our sparsity system, the orthogonal discrete wavelet transform (DWT) [14, 28]. We choose the DWT because many signals are compressible in a wavelet basis, and the properties of the wavelet transform of turbulence signals are well known [35, 19]. The following wavelet analysis assumes a signal of size 2^J , an integer power of two, with $0 \leq j < J$. A full wavelet decomposition of a signal on J scales is

$$(3.5) \quad u = \hat{s}_0^0 \phi_0^0 + \sum_{j=0}^{J-1} \sum_{i=0}^{2^j-1} \hat{d}_i^j \psi_i^j,$$

where ϕ_i^j and ψ_i^j are, respectively, the scaling and the wavelet functions, and \hat{s}_i^j and \hat{d}_i^j are the expansion coefficients. The level is j , the scale is 2^{-j} , and i is the translation. Without loss of generality, as long as levels $j \geq J_0$ (see Algorithm 1) have detail coefficients, we assume a full transform with a single (coarse) scaling coefficient \hat{s}_0^0 . Note that the basis of scaling functions $\{\phi^j\}$ spans the approximation subspace V^j , while the wavelet basis $\{\psi^j\}$ spans the subspace W^j , which is the orthogonal complement of V^j in V^{j+1} , i.e., $V^{j+1} = V^j \oplus W^j$. Thus, a wavelet coefficient \hat{d}_i^j measures how big the signal variation is at a position i and level j .

The DWT has a fast transform for discrete signals, with complexity $\mathcal{O}(N)$. The decay rate of the wavelet coefficients is determined by the local regularity of the signal [29], and this decay rate can be used to estimate the strength of any (quasi-)singularities in signal. The coefficients \hat{s}_i^j and \hat{d}_i^j are stored in \hat{u} in the standard manner.

ALGORITHM 1

Quasi-oracle multilevel orthogonal matching pursuit (QOMOMP). Approximates a sparse solution to $g = \Psi \hat{u}$, where $g \in \mathbb{R}^M$ and $\hat{u} \in \mathbb{R}^N$. It approximates all wavelet coefficients at level $j < J_0$ and L_j coefficients at level $j \geq J_0$. Inputs: g, Ψ, J_0, L . Output: \hat{u}^ .*

```

 $\Omega \leftarrow \cup_{j < J_0} \Gamma_j$   $\triangleright$  initial coefficient index set by oracle,  $\Gamma_j$  is the index set for level  $j$ 
 $\hat{u}^* \leftarrow 0$   $\triangleright$  the decoded signal initial guess
 $\hat{u}_\Omega^* \leftarrow \arg \min_x \|\Psi_\Omega x - g\|^2$   $\triangleright$  least squares,  $\Psi_\Omega$  are columns of  $\Psi$ 
for  $j = J_0$  to  $J - 1$  do  $\triangleright$  for each level  $j \geq J_0$ 
     $r \leftarrow g - \Psi \hat{u}^*$   $\triangleright$  update residual
     $a \leftarrow \Psi^T r$   $\triangleright$  project residual
     $i \leftarrow \text{supp}(|\text{tree}(a)|_{L_j}^>)$   $\triangleright$  the largest  $L_j$  coefficients of  $\text{tree}(a)$  in  $\Gamma_j$ ; see Algorithm 2
     $\Omega \leftarrow \Omega \cup i$   $\triangleright$  update current index set
     $x_0 \leftarrow \hat{u}_\Omega^*$   $\triangleright$  initial guess for iterative methods
     $\hat{u}_\Omega^* \leftarrow \arg \min_{x_0} \|\Psi_\Omega x_0 - g\|^2$   $\triangleright$  least squares solved using tfqmr (1-D) or symmlq (2-D)
end for

```

In two dimensions a separable multiresolution analysis (MRA) includes three components of scale variation [29], decomposing a $2^J \times 2^J$ signal similarly into

$$(3.6) \quad u = \hat{s}_{0,0} \phi_{0,0} + \sum_{k=1}^3 \sum_{j=0}^{J-1} \sum_{i_1, i_2=0}^{2^j-1} \hat{d}_{j,i}^k \psi_{j,i}^k$$

with $i = (i_1, i_2)$ and where $k = 1, 2, 3$ includes contributions from wavelets measuring variation in the horizontal, vertical, and diagonal directions.

Turbulent flows have been analyzed and computed adaptively using wavelet methods for almost two decades [19]. Although homogeneous turbulence is not highly compressible, wavelet coefficients approximate local structures much better than Fourier modes due to the intermittent multiscale structure of turbulence. This multiscale structure is characterized by a continuous range of active length scales which grows like $\text{Re}^{3/4}$ and has a power law energy spectrum like $E(k) \propto k^{-5/3}$ in three dimensions.

The multiscale structure of turbulence and the DWT leads us to propose a multi-level version of OMP that uses our knowledge about the multiscale turbulent flows to predict the typical space and scale structure of the wavelet coefficients. For example, a priori we know that wavelet coefficients are relatively large above a certain scale and, on average, the magnitude of wavelet coefficients decreases monotonically with decreasing scale. We call this method *quasi-oracle multilevel orthogonal matching pursuit* (QOMOMP); see Algorithm 1. QOMOMP will be used to efficiently solve the minimization problem (2.4), which is the key computational step of our energy spectrum estimation method.

Most of the steps in QOMOMP are equivalent to the standard OMP, except for the following differences. QOMOMP estimates all coefficients at levels less than a predefined coarsest level $j < J_0$. The “initial coefficient index set by oracle” defined by J_0 is chosen such that almost all wavelet coefficients up to level J_0 are large, approximately large enough to be included in the best $M/2$ -term approximation. At each finer scale $j \geq J_0$ a predefined number of coefficients, L_j , is estimated. We will see later that the choice of the sequence $L = \{L_j\}$ is a key factor determining the performance of the method.

Discrete wavelet coefficients have a tree-like structure, where (in one dimension) the two child coefficients at a fine scale j are more likely to be large if their parent coefficient at the coarse scale $j - 1$ is large. To enforce this tree-like structure of the

ALGORITHM 2

Description of the function $\text{tree}(a)$. Returns adjusted elements of $a = \Psi^T r$ at level j to enforce tree-like structure of the estimated coefficients \hat{u}_Ω^* . The function depends on the parameter $\beta \geq 1$ and a threshold defined by $\Lambda : \mathbb{R}^{|\omega|} \rightarrow \mathbb{R}_+$.

$\omega \leftarrow \Omega \cap \Gamma_{j-1}$	$\triangleright \Omega$ is current coefficient index set, Γ_{j-1} is index set for level $j-1$
$\omega^* \leftarrow \{i \in \omega : \hat{u}_i^* > \Lambda(\hat{u}_\omega^*)\}$	\triangleright locate large coefficients in Γ_{j-1}
$\Omega^* \leftarrow \text{children}(\omega^*)$	\triangleright corresponding coefficients in Γ_j
$a_{\Gamma_j^c} \leftarrow 0$	\triangleright coefficients outside Γ_j will not be selected
$a_{\Omega^*} \leftarrow \beta a_{\Omega^*}$	\triangleright adjust elements of a with a large parent coefficient (in \hat{u}^*)

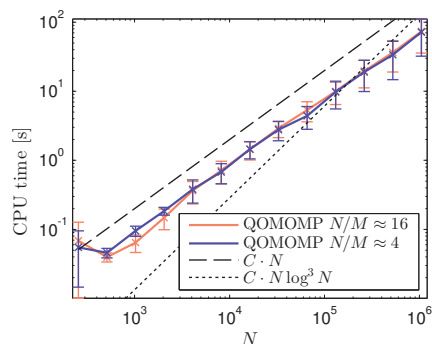


FIG. 1. Computational cost of QOMOMP (measured by CPU time) versus signal length N , showing mean curves and standard deviation bars of 16 random simulations. The number of samples is a fixed ratio of N , either $N/M = 16$ or $N/M = 4$, and the measurement matrix is a filter of length $K = 284$. The number of coefficients L is a fixed ratio of N such that the sparsity is $B/M \approx 0.79$.

nonzero wavelet coefficients \hat{u}^* we apply the function $\text{tree}(a)$ (see Algorithm 2) to modify the raw wavelet coefficients of the residual in the QOMOMP Algorithm 1. This is similar to the method used in [24], but enforces the tree structure less strictly.

The tree algorithm, Algorithm 2, works as follows. Let Γ_j be the index set for level j , and let Ω be the current support of wavelet coefficients \hat{u}^* at iteration j in QOMOMP. The index set ω^* identifies those coefficients at the coarse level $j-1$ above a threshold defined by Λ . Then $\Omega^* = \text{children}(\omega^*)$ are the child coefficients at level j of the significant parent coefficients ω^* at level $j-1$. Finally, the tree function scales the residuals a in Ω^* by a constant, $a_{\Omega^*} \leftarrow \beta a_{\Omega^*}$. If $\beta > 1$, this makes the residuals corresponding to children at scale j of significant wavelet coefficients at scale $j-1$ more likely to be selected as the L_j largest coefficients. If $\beta = 1$, $\text{tree}(a)$ returns a unchanged, while in the limit $\beta \rightarrow \infty$ it exactly enforces a tree structure; i.e., every nonzero coefficient in the wavelet tree has a nonzero parent coefficient.

Isotropy of the signal is not of concern in one dimension. In two dimensions, however, the diagonal wavelet coefficients, denoted by $k = 3$ in (3.6), of a best B -term approximation of an isotropic signal become a smaller proportion of the total for a particular level as the scale decreases. To account for this we let the operator $|\cdot|_{L_j}^>$ in QOMOMP in two dimensions choose the coefficients such that the diagonal ones are a ratio q_j of the total for level j .

The least squares problem in Algorithm 1 is solved using an iterative method for the normal equation. The relative tolerances are fixed, except for the last level, where we decrease the tolerance for higher accuracy. Numerical verification of the computational cost of QOMOMP—Figure 1—confirms that it scales linearly with the signal size N for typical parameters. Intermediate and final tolerances are set to

ALGORITHM 3

Spectrum estimation by sparse optimization (SpESO). The signal $u^(L)$ is an estimate provided by QOMOMP with input L . Inputs: $L^0, \mathcal{J}, A, g(Au), h, w, n_{it}$. Output: energy spectrum estimate E_u^* .*

```

 $L \leftarrow L^0$  ▷ initial guess for number of coefficients at each level
for  $i = 1$  to  $n_{it}$  do ▷ iterate  $n_{it}$  times
  for each  $j \in \mathcal{J}$  in increasing order do ▷ at each level  $j$ ,  $L_j$  in the sequence  $L$ 
     $\omega \leftarrow \begin{cases} \sum_t w_t & \text{if } i = n_{it} \\ w_j & \text{else} \end{cases}$  is adjusted to find an approximate minimization
    (3.7)  $\text{approx } \min_{L_j} \|\log(E_{A^T g}) - \log(E_{A^T A u^*(L)})\|_\omega, \quad \text{s.t. } h_j \geq 0$ 
  end for
end for
 $E_u^* \leftarrow E_{u^*(L)}$  ▷ estimate the energy spectrum of  $u$ 

```

$\epsilon_i = 2 \times 10^{-2}$ and $\epsilon_f = 3.3 \times 10^{-6}$, respectively.

Now, let us return to the energy spectrum estimation problem stated in (3.3). Algorithm 3 defines a low-dimensional optimization algorithm that approximates a solution to (3.3), which we call *spectrum estimation by sparse optimization*¹ (SpESO). The sequence L^0 is the initial sequence of the number of nonzero coefficients at each level for QOMOMP, and the index set \mathcal{J} specifies those levels for which we want to optimize the sequence L . The weight $w_j = 1$ for indices in $2^{j-1} < k \leq 2^j$ and is zero elsewhere (essentially an index set). We put the constraints $h_j = L_{j-1} - L_j$ in one dimension and $h_j = 2L_{j-1} - L_j$ in two dimensions. This is a sort of bootstrapping from large scales (small j , where the energy spectrum is likely most accurate) to smaller scales (where the energy spectrum is initially less accurate). Since the computation of u^* is expensive and the optimization function is nonsmooth, we do not solve (3.7) exactly. Instead, we search among 5 to 7 values uniformly distributed on a log scale and narrow the search after each iteration of (3.7) over all levels \mathcal{J} . The minimization (3.7) over \mathcal{J} is iterated a few times (n_{it}) until L stabilizes or converges to a (probably) local minimum.² The overall error over all levels \mathcal{J} is measured in the final iteration. We estimate the overall computational complexity of SpESO to be $\mathcal{O}(N|\mathcal{J}|n_{it})$.

Our experiments show that decoupling the matrix used in SpESO from the one used in QOMOMP improves the convergence properties. By that we mean that the measurement matrix is split horizontally into two parts, \bar{A} and \tilde{A} , giving a set of measurements $\bar{g} = \bar{A}u$ and $\tilde{g} = \tilde{A}u$. For QOMOMP we use $\bar{\Psi} = \bar{A}\Phi$ and \bar{g} , and for SpESO we use $\tilde{A}^T\tilde{A}$, and vice versa. The two estimated spectra are then combined proportionally to their relative errors. A simplistic argument for the decoupling is that since QOMOMP minimizes the error $Au^* - g$ to a small or zero value regardless of L , then the difference between $A^T Au$ and $A^T Au^*$ will be small and (3.7) will not converge to any meaningful minimum. By using two separate matrices this problem disappears and results in a better correlation between a good choice of L and a low energy spectrum error. The downside is that QOMOMP only uses half of the measurements for each estimation.

¹The code for SpESO with QOMOMP is available at github.com as SpESO.

²One might think that an optimal solution to the minimization would be the maximum value of L_j , but this not the case since the estimates provided by QOMOMP start to deteriorate for too large values of L . No restriction is placed on $\sum_j L_j$.

4. Analysis of the performance of SpESO for ideal signals. We now analyze mathematically the convergence and accuracy of SpESO. Let us consider the restricted isometry property (RIP) of the CS matrices that determines the accuracy of reconstructions. The restricted isometry constant of a matrix Ψ is the smallest number δ_B such that [13, 7]

$$(4.1) \quad (1 - \delta_B)\|x\|^2 \leq \|\Psi x\|^2 \leq (1 + \delta_B)\|x\|^2$$

holds for all x at most B -sparse. If the OMP algorithm is applied with a matrix Ψ satisfying $\delta_{13B} < \frac{1}{6}$ [21] or $\delta_{B+1} < 1/(\sqrt{B} + 1)$ [40], then it recovers a B -sparse signal exactly. The proof in [40] is mainly concerned with showing that at each iteration the index chosen is in the true support T . Given the true support at the final iteration, the reconstruction is trivial.

Assume T is the true support of the best B -term approximation u_B , where again $u = \Phi \hat{u}$. In the case of a perfect oracle where $\Omega = T$ in QOMOMP, the solution to the final least squares problem is (Ψ_T are columns of Ψ index by T)

$$(4.2) \quad \hat{u}_T^* = \Psi_T^+ \Psi \hat{u} = \Psi_T^+ (\Psi_T \hat{u}_T + \Psi_{T^c} \hat{u}_{T^c}),$$

where Ψ_T^+ is the Moore–Penrose pseudoinverse. With $\Psi_T^* \Psi_T$ nonsingular ($\delta_B < 1$) we get $\hat{u}_T^* = \hat{u}_T + \Psi_T^+ \Psi_{T^c} \hat{u}_{T^c}$, where Ψ_T^* is Ψ_T transposed. Therefore, the error is

$$(4.3) \quad \|\hat{u}_T - \hat{u}_T^*\| = \|\Psi_T^+ \Psi_{T^c} \hat{u}_{T^c}\|.$$

Now considering the case where Φ is orthonormal, we get

$$(4.4) \quad \|u - u^*\|^2 = \|\hat{u}_{T^c}\|^2 + \|\Psi_T^+ \Psi_{T^c} \hat{u}_{T^c}\|^2 \leq \|\hat{u} - \hat{u}_B\|^2 + \frac{1}{1 - \delta_B} \|\Psi(\hat{u} - \hat{u}_B)\|^2$$

(the inequality follows from RIP [30]). For a compressible signal u , the error depends on the best B -term approximation error $\|\hat{u}_{T^c}\| = \|u - u_B\|$ and the least squares error term, which depends on the RIP of the matrix Ψ . Given u is B -sparse ($u = u_B$), the error vanishes.

Now consider our QOMOMP method in a very simple 1-D setting to obtain some quantitative performance estimates. Let QOMOMP be applied to a signal u with a power law energy spectrum $E(k) \lesssim k^{-\alpha}$, where the number of vanishing moments n of the wavelet used in the sparsity system is limited by $0 < \alpha \leq 2n + 1$. The variance of the wavelet coefficients at each level then scales like $\text{Var}(\hat{d}_i^j) \sim 2^{-j\alpha}$ [33]. Assuming u is a Fourier synthetic signal like those considered in section 5, then \hat{d}_i^j for each level is well approximated as i.i.d. with a Gaussian distribution and zero mean. If $\Omega_{J_0} = \cup_{j < J_0} \Gamma_j$ is the initial QOMOMP setup, the probability that the true support T_B of the best B -term approximation contains Ω_{J_0} is

$$(4.5) \quad P_{J_0} = \Pr(\Omega_{J_0} \subset T_B) = \prod_{j < J_0} \Pr(|\hat{d}_i^j| \geq \epsilon_B)^{2^j},$$

where the threshold ϵ_B is the best B -term threshold such that

$$(4.6) \quad B/N = \frac{1}{N} \sum_{j < J} \sum_i \Pr(|\hat{d}_i^j| \geq \epsilon_B) = \frac{1}{N} \sum_{j < J} 2^j \left[2 - 2F(\epsilon_B 2^{j\alpha/2}) \right],$$

where F is the standard cumulative distribution.

Assume now, as in the 1-D experiments in section 5, that $N = 2^{15}$ and $B = N/8$. If $J_0 = 5$, then $P_5 = 82\%$ for energy spectrum slope $\alpha = 5/3$ and $P_5 = 99.5\%$ for energy spectrum slope $\alpha = 3$. There is a reasonable probability that *every* coefficient in Ω_{J_0} is in the true support. We can also consider how many coefficients L_j should be estimated at finer levels $j \geq J_0$. At the single level $j = 7$, 94.9% of the coefficients are included in u_B for $\alpha = 5/3$, and 99.4% are included for $\alpha = 3$. Thus, the exact value for the number of coefficients to estimate at this level is $L_7 = 0.95 \times 2^7$ for $\alpha = 5/3$. Recall that in practice the initial sequence guess $L^0 = \{L_j\}$ must be estimated a priori, but this sort of analysis gives us a good ansatz for determining it. Note that adding some excess coefficients that are not in the true support is not a serious problem; the algorithm in [39] defines an iteration as successful if at least one coefficient is correctly chosen.

A full theoretical analysis of SpESO is beyond the scope of the present paper, i.e., a proof of concept of an experimental algorithm. In addition, since CS algorithm estimates are usually conservative, they are not a good indicator of the actual performance of the method. Therefore, in the next section we rely instead on a wide range of representative computational experiments to assess the actual performance of SpESO.

5. Numerical tests of SpESO.

5.1. Turbulence test signals and computational parameters. To test the effectiveness of SpESO we need signals with energy spectra and arbitrary power law scaling. For this purpose, synthetic turbulence signals with power law energy spectra are particularly appropriate, in addition to experimental signals and data from numerical simulations of turbulence.

In our results, synthetic signal type (x, y) denotes a signal with two energy spectrum power laws $-x$ and $-y$, i.e., $E(k) \sim k^{-x}$ and $E(k) \sim k^{-y}$, split at $k = N/32$ in one dimension (unless specified otherwise) and $k = N/8$ in two dimensions. Signal (x) denotes a signal with a single power law. $F(x, y)$ denotes a synthetic Fourier signal, and $W(x, y)$ denotes a synthetic wavelet signal. The Fourier and wavelet synthetic signals are described below. Note that signals with a change in slope are particularly challenging for energy spectrum estimation when this change occurs at wavenumbers larger than the Nyquist wavenumber corresponding to the Shannon sampling rate since the second slope would not be resolvable using a standing Shannon sampling technique.

The first method constructs a synthetic signal in Fourier space. The Fourier coefficients of the signal u are determined by the desired energy spectrum but with random complex phases,

$$(5.1) \quad \hat{u}_k = \sqrt{E(k)}e^{i\theta_k},$$

where θ_k are i.i.d. uniformly on $[0, 2\pi)$ and $E(k)$ is the specified spectrum. In higher dimensions, the coefficients $\hat{u}[\mathbf{k}]$ for which $|\mathbf{k}| = k$ have variance proportional to $E(k)$. The resulting signal is homogeneous Gaussian statistics. Shown in Figure 2 are typical realizations of the Fourier-based synthetic 1-D test signal (top left) and 2-D data (bottom left).

In addition to the random phase Fourier synthetic signal described above, we also consider a synthetic multiscale signal, generated by a random process in wavelet space. This allows us to generate a synthetic signal that is closer to a true turbulent signal. Kolmogorov's original statistical theory of turbulence [22] predicts a structure function

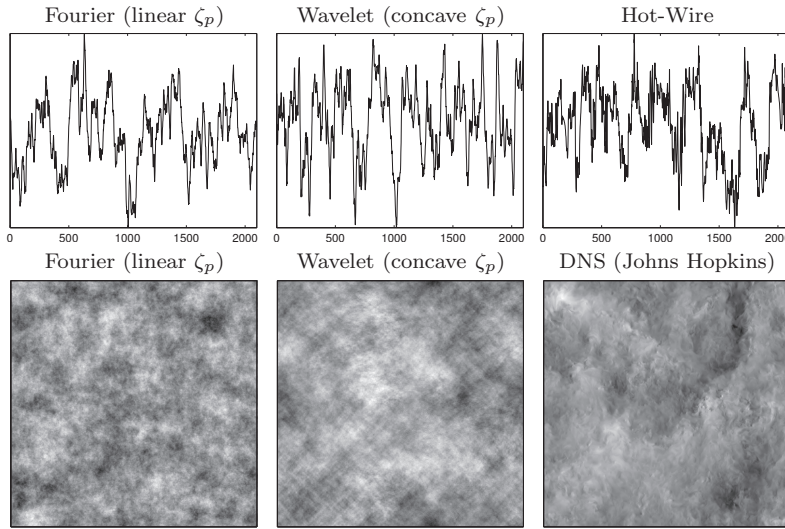


FIG. 2. 1-D and 2-D turbulence test signals. The Fourier and wavelet synthetic signals both have $(5/3, 3)$ energy spectra typical of a 2-D flow, while the experimental hot-wire signal and the direct numerical simulation (DNS) data have a signal power law scaling $(5/3)$ typical of 3-D flow. Note that the Fourier signal is nonintermittent (with linear ζ_p), while the wavelet signal has been designed to have more realistic intermittent statistics (with concave ζ_p).

scaling $S_p(r) = \frac{1}{L} \int_0^L |u(x) - u(x+r)|^p dx \sim r^{\zeta_p}$ with $\zeta_p = p/3$, in the limit of infinite inertial subrange, and this is approximately the scaling produced by the Fourier-based synthetic signal process described above. However, actual experimental measurements show that as a result of intermittency ζ_p increases more slowly than linearly with p ; i.e., it is concave. The lack of intermittency in the Fourier synthetic signal means that the resulting data sequence is more homogeneous locally than real turbulence. In order to assess the ability of SpESO to cope with intermittency, we have also used the wavelet-based method of [5] to synthesize a signal with a more realistic concave function ζ_p . The wavelet-based signal has a realistic concave, intermittent scaling of structure function exponents ζ_p , while the Fourier signal has a nonphysical slightly convex scaling, in contradiction to experimental measurements [3]. The scaling of the energy spectrum is then defined implicitly by the second-order structure function. Typical realizations of the wavelet synthetic 1-D test signal $W(5/3, 3)$ (top middle) and 2-D data (bottom middle) are shown in Figure 2. Note that we generate this wavelet-based synthetic data using symlet 12 wavelets with six vanishing moments, rather than the coiflet wavelets used for the energy spectrum estimation algorithm, to ensure that the data are independent of the sparsity system used in the compressive sampling.

Finally, we consider two realistic turbulence data sets: a 1-D times series measurement of a single velocity component of an axisymmetric jet [31], and a 2-D slice of a 3-D direct numerical simulation (DNS) of homogeneous isotropic turbulence [27, 32]. The 1-D data are from hot-wire measurements at 20kHz and $\text{Re} = 4 \times 10^4$, and a typical section is shown in Figure 2 (top right). Note that the energy spectrum of this data has a signal power law scaling $k^{-5/3}$. The second data set is from a high Reynolds number 3-D pseudospectral turbulence simulation stored in the Johns Hopkins University (JHU) turbulence database cluster [27, 32]. The Taylor scale Reynolds number of this flow is $R_\lambda \sim 433$ (corresponding to $\text{Re} \approx 2 \times 10^5$). This simulation has a resolution of 1024^3 , and therefore the 2-D slice has a resolution of 1024^2 . To

simplify the analysis, we consider a single velocity component. A typical example of this data is shown Figure 2 (bottom right).

The results are computed using the coiflet 18 wavelet basis with six vanishing moments for 1-D signals and coiflet 12 wavelet basis with four vanishing moments for 2-D signals. The random filter is i.i.d. uniform in $\{\pm 1\}$, and the length for all 1-D cases is $K = 284$. For the tree function, the threshold function is $\Lambda : a \rightarrow \frac{1}{2}\sqrt{\text{Var}(a)}$, and $\beta = 3$ ($\beta = 2$ in two dimensions). The set of levels to optimize over is approximately $\mathcal{J} = \{j : j > \log_2(M/2)\}$. The a priori-based initial guess of coefficients L^0 is set approximately to the number of coefficients of each level in a typical best M -term approximation, obtained by analyzing turbulence signals of similar parameters, and the level $J_0 = 5$ in one dimension and $J_0 = 4$ in two dimensions. Logarithmic scale averages of spectra are essentially geometric means of spectra.

Dashed lines in the figures show the theoretical slope of the power law parts of the spectrum. In each case, we compare SpESO with the original signal, best M -term wavelet approximation, and the usual fixed rate Shannon sampling. In many cases we also compare results with the best $M/2$ -term approximation and the LOMP method.

The differences between each of the spectrum estimation methods are as follows. The best M -term approximation first takes the wavelet transform of the entire signal and then selects the largest M wavelet coefficients for the estimation. It is therefore not a sampling scheme, but rather an ideal benchmark to which the CS methods are compared. We expect that the M -best term approximation to be the best possible estimate using M samples. The Shannon scheme subsamples u at a uniform rate (without low-pass filtering), followed by a Fourier interpolation. LOMP is an iterative CS method like the well-known OMP that estimates a few ($L_0 \geq 1$; we use a moderate value of 20 from computational evidence) coefficients at a time, without using any a priori information or the tree structure of the wavelet coefficients. Finally, SpESO is a CS optimization method that uses a priori information to minimize the logarithmic scale error of the energy spectra. It also enforces a realistic tree structure for the estimated wavelet coefficients.

5.2. Results for 1-D Signals. The performance of SpESO compared to other methods is tested numerically for a range of signal length to measurement length ratios N/M (i.e., sampling ratios). The other methods are fixed-rate Shannon sampling, best $M/2$ and M -term wavelet approximations in the coiflet 12 basis, and the CS reconstruction method LOMP. For the tested N/M ratios 4, 8, 16, 32, SpESO and LOMP have the nearest ratio satisfying (2.6), namely, 3.97, 7.93, 15.86, 31.72, respectively. The results for the Fourier (F) and wavelet (W) synthetic 1-D signals are shown in Figures 3, 4, 5, 6, 7 and Table 1. The results for hot-wire signals are shown in Figure 8.

The energy spectrum errors for each level shown in Figure 3 behave as expected for Shannon sampling: they increase dramatically at the Nyquist wavenumber. The best-term approximations are the most accurate method at large scales, but have a steeply rising error at smaller scales. The performance of SpESO is almost independent of level j , except for the highest level $j = 14$. The figures show that SpESO has lower errors than the best $M/2$ -term approximation at levels 11 to 14. Compared to the other methods, LOMP is clearly not competitive at large scales, but at small scales it is similar. It is important to remember that the good relative performance of SpESO is especially significant since the best $M/2$ -term approximation requires full sampling of the signal (the nonlinear wavelet filtering is based on the full set of wavelet coefficients).

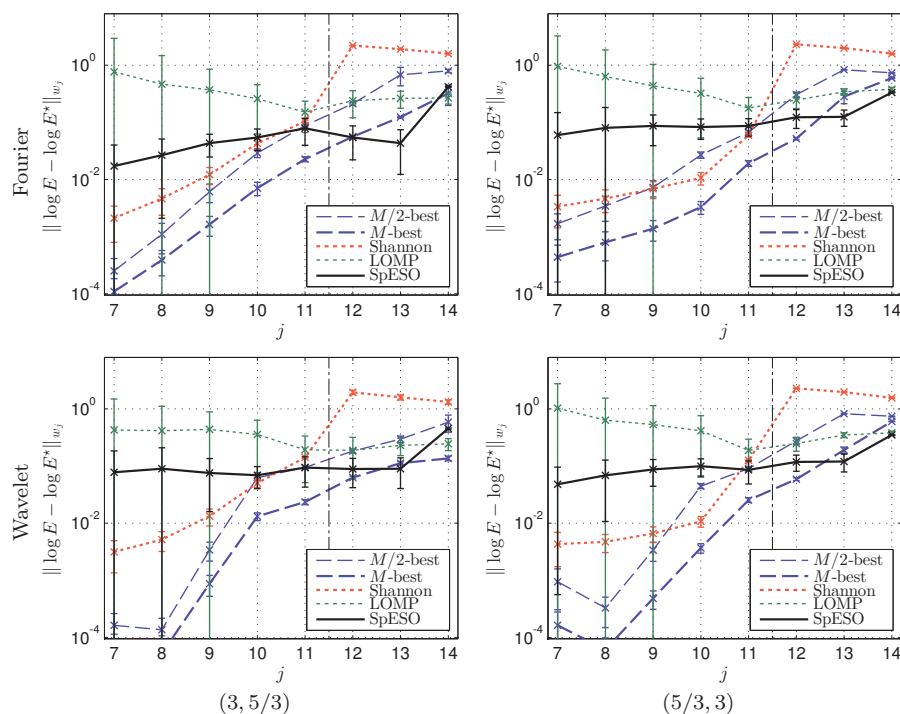


FIG. 3. Logarithmic scale errors of spectrum estimations versus level j , on a logarithmic scale, with $N = 2^{15}$ and $N/M = 8$ (Nyquist wavenumber at dash-dot vertical line). The support of the norm weight w_j is compact and centered at a wave number corresponding to level j and $\sum_k w_j(k) = 1$. The plots show mean curves and standard deviation bars of 64 random simulations. The signals are 1-D with spectrum slopes splitting at $k = N/32$ ($j = 10.5$). Fewer samples at large scales (small j) result in larger error bars.

TABLE 1

Errors in spectrum slopes of averaged estimations, $s - s^*$, where s is the slope of the original signal, i.e., a k^{-s} power law. Slopes are computed by least squares fits in the range of the second slope, from $k = 1024$ to $k = 8192$, except for the Shannon slope, which is fitted in its nonzero range only. The 1-D signal lengths are $N = 2^{15}$, and the number of simulations is 64.

Signal	N/M	SpESO	Shannon	M/2-best	M-best	LOMP
W(3, 5/3)	4	-0.06	0.58	-0.78	-0.29	1.46
	8	-0.50	1.27	-1.92	-0.78	2.23
	16	-1.06	n/a	n/a	-1.92	2.14
F(3, 5/3)	4	-0.09	0.49	-0.86	-0.26	1.45
	8	-0.26	0.90	-4.10	-0.86	2.23
	16	-0.75	n/a	n/a	-4.10	2.26
W(5/3, 3)	4	0.93	0.54	-1.30	-0.33	2.42
	8	0.76	1.65	-4.84	-1.30	3.33
	16	0.25	n/a	n/a	-4.84	3.28
F(5/3, 3)	4	0.95	0.35	-1.87	-0.27	2.27
	8	0.85	0.87	-4.76	-1.87	3.23
	16	0.46	n/a	n/a	-4.76	3.29

Table 1 gives the error of the estimates of the power law scaling of the energy spectrum over the small scale power law range (i.e., large wavenumber range). These slopes are computed for averages of estimations. This is a crucial quantity character-

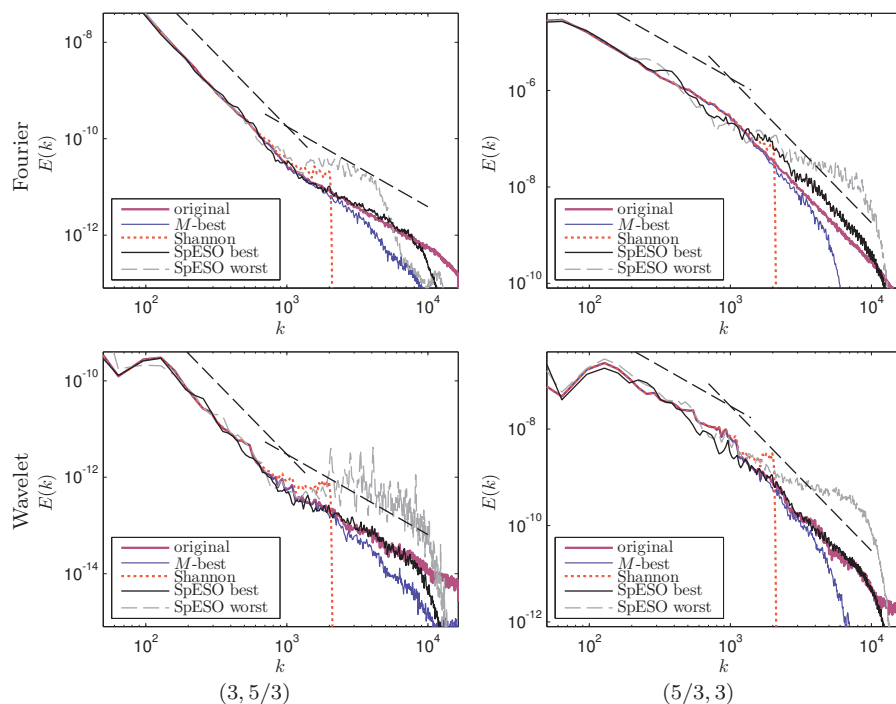


FIG. 4. Representative energy spectrum estimations of 1-D signals, with $N = 2^{15}$ and $N/M = 8$, out of the 16 simulations. The slope of the energy spectrum changes at $k = N/32 = 1024$, smaller than the Nyquist wavenumber $k = N/16 = 2048$. The best and worst cases of SpESO reveal the variation of its approximations.

izing turbulent flows and other experimental signals. Those cases that are too bad for a reasonable fit are indicated by “n/a”. SpESO performs better than the $M/2$ -best term approximation in all cases, and better than the M -best term approximation and Shannon sampling (where it is valid) in all but two cases shown in Table 1, namely, both $(5/3, 3)$ cases with sampling ratio $N/M = 4$. LOMP is again noticeably worse than all other methods. Even for the power law scaling at small wavenumbers (not shown), which is well resolved by the Shannon sampling, SpESO still gives results similar to the best $M/2$ -term approximation at high sampling ratios and much better than LOMP.

It is important to note that there is significant stochastic variation in the SpESO energy spectrum estimates. Figure 4 shows the best and worst cases of SpESO as measured by the ℓ^2 norm on a logarithmic scale when the energy spectrum slope changes at a wavenumber two times smaller than the Nyquist wavenumber. Even the worst cases are not much different from the ideal case of Shannon sampling in the low wavenumbers, although they significantly overestimate the energy at high wavenumbers. In addition, the worst SpESO cases seem unable to track the steepening slope at high wavenumber in the $(5/3, 3)$ cases, although they do estimate approximately the correct high wavenumber slope for the $(3, 5/3)$ cases (but at incorrectly high energy levels). In contrast, at all wavenumbers the best SpESO cases estimate the spectra better than even the best M -term estimates. This suggests there is potential to greatly improve the reliability and accuracy of the method if acceptable trials could be determined a priori. Figure 5 shows the variation of the SpESO estimates in two

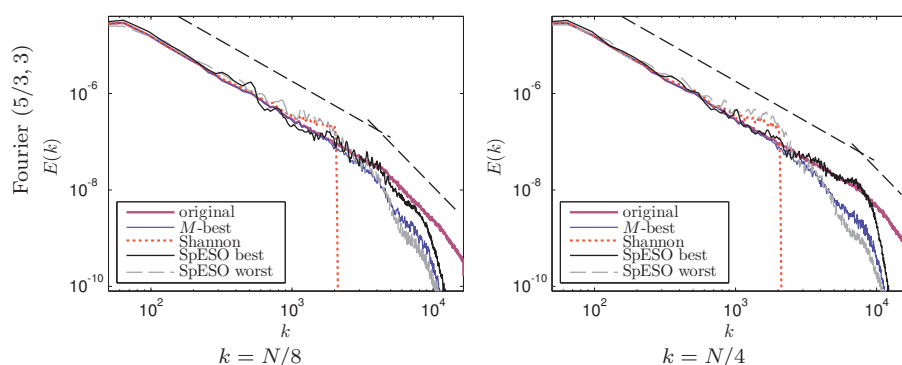


FIG. 5. Representative energy spectrum estimations similar to Figure 4 (upper right), but with the slope changing at $k = N/8 = 4096$ (left) and $k = N/4 = 8192$ (right).

cases where the slope of the energy spectrum changes at wavenumbers four and eight times larger than the Nyquist wavenumber. In this case, the worst SpESO estimate is similar in accuracy to the best M -term approximation (both methods underestimate the energy in the second power law range), while the best SpESO result gives an excellent estimate. The averaged spectra are, however, not as responsive as the best cases.

The averages of estimated energy spectra corresponding to Table 1 are shown in Figure 6. Comparing SpESO to Shannon, SpESO somewhat more accurately estimates the spectra beyond the Nyquist wavenumber, although performance is signal-dependent. For the $(5/3, 3)$ cases, it is unclear to us why the higher ratios of N/M are more accurate than the lower ratios. On average, SpESO is not worse than the best $M/2$ -term approximation.

A more severe test is to apply SpESO to signals where the change in slope is at a wavenumber higher than the equivalent Nyquist wavenumber for the sampling ratio used. Results for this test are shown in Figure 7 when the slope changes at $k = N/32 = 1024$, while the equivalent Nyquist sampling wavenumber is only $k = 512$. These results show SpESO is still able to estimate the spectra for low wavenumbers, but it is not always reliable for high wavenumbers. At these relatively large sampling ratios SpESO performs very well for the $W(3, 5/3)$ cases (i.e., for intermittent cases when the slope becomes shallower at higher wavenumbers). SpESO performs worst for the $F(5/3, 3)$ cases (not shown but similar to the $W(5/3, 3)$ cases) when the data are statistically nonintermittent.

The experiments for the hot-wire data (see Figure 8) show the ability of SpESO to estimate the spectra beyond the capabilities of the Shannon sampling. However, the estimates are not accurate in the range of the high wavenumber exponential decay of E .

5.3. Results for 2-D Signals. We shall now examine 2-D signals of length $N = n \times n$. For indicated ratio $N/M = 64$, SpESO actually has the ratio 63.3 due to the computational setup. We note that individual 2-D estimates vary much less and are much smoother than the 1-D estimates. The results for Fourier and wavelet synthetic signals are shown in Figures 9 and 10, and the results for signals from the JHU DNS database are shown in Figure 11.

Results for single $(5/3)$ slope synthetic signals are in Figure 9. Clearly, SpESO is able to estimate the spectra accurately for the mid and high wavenumbers much better

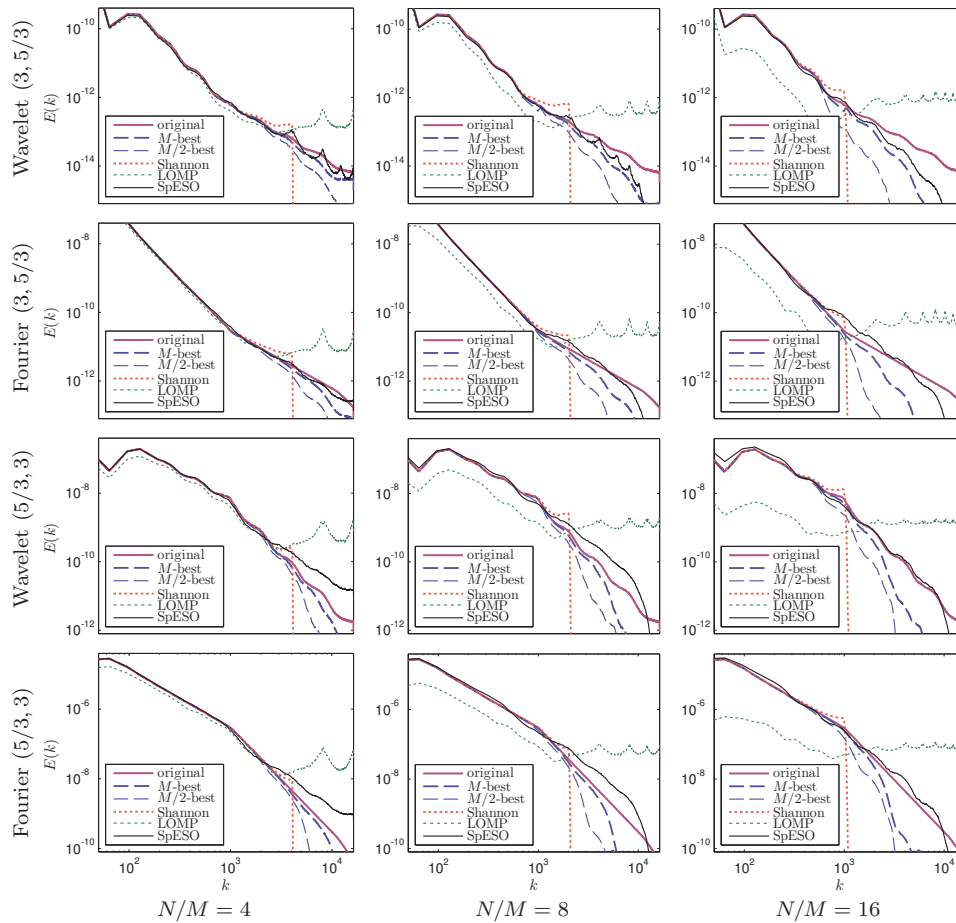


FIG. 6. Logarithmic scale averages of spectrum estimates for 1-D signals and various sampling ratios. Corresponds to Table 1.

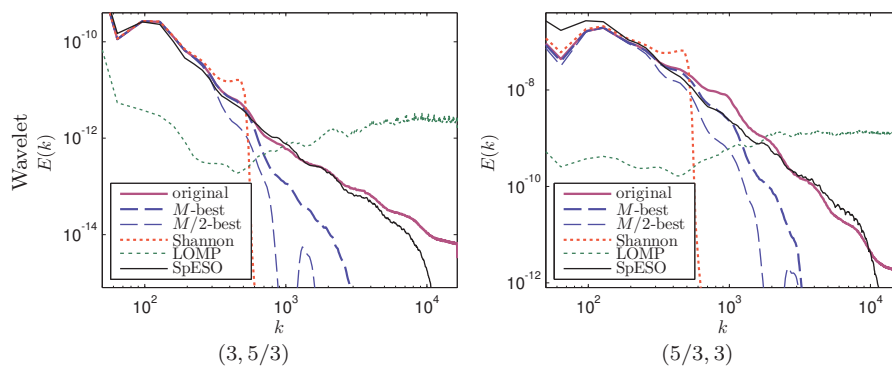


FIG. 7. Logarithmic scale averages of spectrum estimates for 1-D signals for a high sampling ratio where the slope of the energy spectrum changes at a wavenumber higher than the Nyquist wavenumber. The signal lengths are $N = 2^{15}$, the sampling ratio is $N/M = 32$, and the number of simulations is 64. The slope changes at $k = N/32 = 1024$, which is larger than the Nyquist sampling wavenumber $k = 512$ corresponding to the sampling ratio 32.

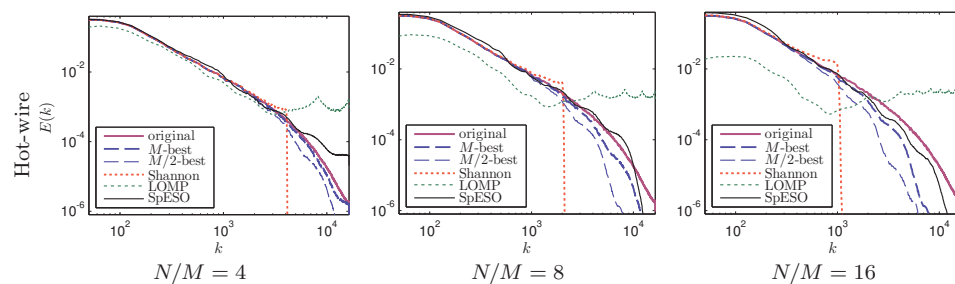


FIG. 8. Logarithmic scale averages of spectrum estimates for a 1-D hot-wire measurement. The signal length is $N = 2^{15}$, and the number of simulations is 64.

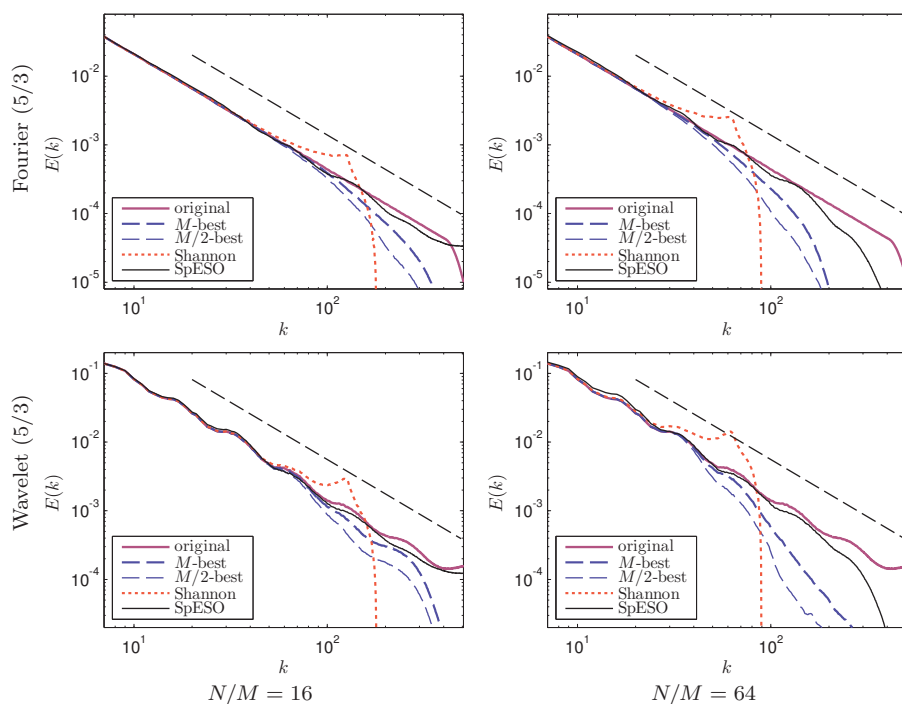


FIG. 9. Logarithmic scale averages of spectrum estimates for single-slope 2-D signals. The signal lengths are $N = 1024^2 = 2^{20}$, and the number of simulations is 16.

than Shannon sampling. The SpESO estimates are closer overall to the exact results than the best M -term approximations. This is surprising because the best M -term approximation is expected to give the upper bound on the accuracy of SpESO since it uses all data, and then reconstructs with the best M largest wavelet coefficients. This suggests that best M -term approximations are not necessarily optimal for estimating nonlinear functions of the data and that carefully designed CS methods may be a better choice even if all data are available for analysis.

As in one dimension, we test our method for synthetic signals with a change in slope (Figure 10). Apart from the wavelet (5/3,3) case, the SpESO slope estimates are at least on a par with the $M/2$ -best. SpESO can predict a change in a spectrum slope at the Nyquist wavenumber, which is not possible using Shannon sampling.

A significant range of the energy spectra for the JHU DNS data has an exponential

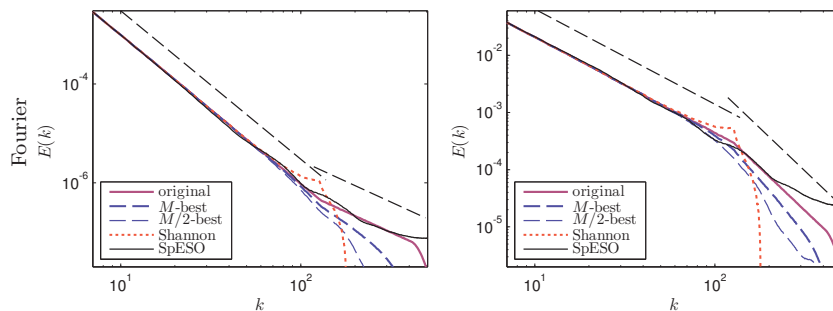


FIG. 10. Logarithmic scale averages of spectrum estimates for 2-D signals. The signal lengths are $N = 1024^2 = 2^{20}$, the sampling ratio is $N/M = 16$, and the number of simulations is 16.

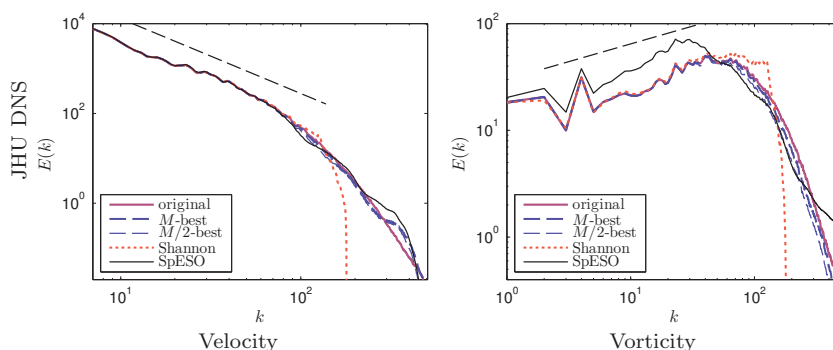


FIG. 11. Logarithmic scale averages of spectrum estimates for 2-D DNS signals. The signal lengths are $N = 1024^2$, the sampling ratio is $N/M = 16$, and the number of simulations is 16.

decay (see Figure 11), and in this range the best-term approximations are indeed better than SpESO—but are not far from each other in the velocity case. However, when applied to an equivalent vorticity field with a positive power law slope at small wavenumbers, SpESO captures the correct scaling, but overestimates the energy by a significant amount.

6. Conclusions. The compressive sampling spectrum estimation by sparse optimization (SpESO) method proposed in this paper shows potential for energy spectrum estimation of signals with power law decay. At this stage SpESO is experimental, a proof of concept, without rigorous proofs of convergence or error bounds. Nevertheless, we have derived mathematical estimates for the performance of parts of SpESO for “good” initial guesses and ideal signals in section 4 and tested it numerically on a wide variety of representative synthetic, experimental, and DNS turbulence signals in one and two dimensions in section 5.

The 2-D cases appear more promising than the 1-D cases, probably due to the dimensionality reduction, or to different measurement matrix, or both. The results show that the estimates generated by SpESO distribute errors more evenly over the full range of wavenumbers than traditional Shannon sampling or best-term wavelet approximations. They also correctly predict the power law scaling of the energy spectrum at wavenumbers higher than those that can be captured with Shannon sampling (which is limited by the Nyquist frequency).

Most interestingly, SpESO typically performs better than a best-term wavelet

approximation using the same number of coefficients. This is surprising because best-term wavelet approximations require the wavelet transform of the entire data set and then select the largest M terms (i.e., it uses complete information about the signal to build its approximation from a nonlinear filter of the wavelet coefficients of the data). In contrast, the SpESO method samples only M data points, between 4 and 64 times fewer samples than with the best-term wavelet approximations.

Both SpESO and QOMOMP have several tunable parameters and many possible variations. We do not suggest the method, as it is, should be immediately used in applications. However, by tuning of parameters and estimation of errors, it might be a practically useful method in those cases where high accuracy is less important than minimizing the number of samples, or where obtaining a complete set of evenly spaced measurements at the Nyquist rate is not possible. In particular, it could be used for estimating the energy spectrum of 3-D or 2-D turbulent flows at very high Reynolds numbers where sampling at the Nyquist rate is impossible. For example, energy spectrum estimation of atmospheric flow at a Reynolds number $Re \sim 10^{10}$ would require $\sim 10^{22}$ samples to fully characterize its energy spectrum if sampled in three dimensions at the Nyquist rate. Even a 1-D measurement would require $\sim 10^7$ samples, which may be impractical in some cases.

It should be straightforward to extend SpESO to three dimensions, and it could be tested with measurement matrices more appropriate for field or laboratory experiments. The same approach could be used to estimate other nonlinear functions of compressively sampled data, such as the scaling of high order exponents of turbulence structure functions ζ_p , which require extremely large data sets to characterize properly for $p > 10$. SpESO could be optimized further by improving the performance of the sparsity system, for example, by using wavelet packets instead of orthogonal wavelets.

This paper has shown that it is possible to design a CS-based energy spectrum estimation method that performs much better than the existing LOMP or Shannon sampling approaches, even in the case where the signal is not sparse in Fourier space. In fact, a CS-based method can perform at least as well as, and often better than, a best-term wavelet approximation that requires full sampling of the signal.

REFERENCES

- [1] J. ALAM, N. K.-R. KEVLAHAN, AND O. V. VASILYEV, *Scaling of space-time modes with Reynolds number in two-dimensional turbulence*, J. Fluid Mech., 570 (2007), pp. 217–226.
- [2] J. ALMEIDA, J. PRIOR, AND M. B. PLENIO, *Computation of 2-D Spectra Assisted by Compressed Sampling*, preprint, arXiv:1207.2404, 2012.
- [3] F. ANSELMET, Y. GAGNE, E. J. HOPFINGER, AND R. A. ANTONIA, *High-order velocity structure functions in turbulent shear flows*, J. Fluid Mech., 140 (1984), pp. 63–89.
- [4] D. DONY ARIANANDA, G. LEUS, AND Z. TIAN, *Multi-coset sampling for power spectrum blind sensing*, in Proceedings of the 17th International Conference on Digital Signal Processing (DSP), 2011.
- [5] R. BENZI, L. BIFERALE, A. CRISANTI, G. PALADIN, M. VERGASSOLA, AND A. VULPIANI, *A random process for the construction of multifractal fields*, Phys. D, 65 (1993), pp. 352–358.
- [6] T. BLUMENSATH, *Compressed Sensing with Nonlinear Observations and Related Nonlinear Optimisation Problems*, preprint, arXiv:1205.1650, 2012.
- [7] E. CANDÈS, *The restricted isometry property and its implications for compressed sensing*, C. R. Math. Acad. Sci. Paris, 346 (2008), pp. 589–592.
- [8] E. CANDÈS AND J. ROMBERG, *Sparsity and incoherence in compressive sampling*, Inverse Problems, 23 (2007), pp. 969–985.
- [9] E. CANDÈS, J. ROMBERG, AND T. TAO, *Robust uncertainty principles: Exact signal reconstruction from highly incomplete Fourier information*, IEEE Trans. Inform. Theory, 52 (2006),

- pp. 489–509.
- [10] E. J. CANDÈS AND J. ROMBERG, *Practical signal recovery from random projections*, in Wavelets XI, M. Papadakis, A. F. Laine, and M. A. Unser, eds., Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 5914, SPIE, Bellingham, WA, 2005.
 - [11] E. J. CANDÈS AND J. ROMBERG, *Quantitative robust uncertainty principles and optimally sparse decompositions*, Found. Comput. Math., 6 (2006), pp. 227–254.
 - [12] E. J. CANDÈS, J. K. ROMBERG, AND T. TAO, *Stable signal recovery from incomplete and inaccurate measurements*, Comm. Pure Appl. Math., 59 (2006), pp. 1207–1223.
 - [13] E. J. CANDÈS AND T. TAO, *Decoding by linear programming*, IEEE Trans. Inform. Theory, 51 (2005), pp. 4203–4215.
 - [14] I. DAUBECHIES, *Orthonormal bases of compactly supported wavelets*, Comm. Pure Appl. Math., 41 (1988), pp. 909–996.
 - [15] M. A. DAVENPORT, P. T. BOUFONOS, M. B. WAKIN, AND R. G. BARANIUK, *Signal processing with compressive measurements*, IEEE J. Selected Topics Signal Process., 4 (2010), pp. 445–460.
 - [16] G. DE STEFANO, D. E. GOLDSTEIN, O. V. VASILYEV, AND N. K. R. KEVLAHAN, *Towards Lagrangian dynamic SGS model for SCALES of isotropic turbulence*, in Direct and Large-Eddy Simulation VI, E. Lamballais, R. Friedrich, B. J. Geurts, and O. Metais, O, eds., Springer, 2006, pp. 175–182.
 - [17] D. L. DONOHO, M. VETTERLI, R. A. DEVORE, AND I. DAUBECHIES, *Data compression and harmonic analysis*, IEEE Trans. Inform. Theory, 44 (1998), pp. 2435–2476.
 - [18] D. L. DONOHO, *Compressed sensing*, IEEE Trans. Inform. Theory, 52 (2006), pp. 1289–1306.
 - [19] M. FARGE, N. KEVLAHAN, V. PERRIER, AND E. GOIRAND, *Wavelets and turbulence*, Proc. IEEE, 84 (1996), pp. 639–669.
 - [20] M. FARGE, K. SCHNEIDER, AND N. K.-R. KEVLAHAN, *Non-Gaussianity and coherent vortex simulation for two-dimensional turbulence using an adaptive orthogonal wavelet basis*, Phys. Fluids, 11 (1999), pp. 2187–2201.
 - [21] S. FOUCART AND H. RAUHUT, *A Mathematical Introduction to Compressive Sensing*, Birkhäuser, 2013.
 - [22] U. FRISCH, *Turbulence*, Cambridge University Press, Cambridge, UK, 1996.
 - [23] J. FRÖHLICH AND K. SCHNEIDER, *Numerical simulation of decaying turbulence in an adaptive wavelet basis*, Appl. Comput. Harmon. Anal., 3 (1996), pp. 393–397.
 - [24] C. LA AND M. N. DO, *Signal reconstruction using sparse tree representations*, in Wavelets XI, M. Papadakis, A. F. Laine, and M. A. Unser, eds., Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 5914, SPIE, Bellingham, WA, 2005, pp. 273–283.
 - [25] G. LEUS AND D. D. ARIANANDA, *Power spectrum blind sampling*, IEEE Signal Process. Lett., 18 (2011), pp. 443–446.
 - [26] M. A. LEXA, M. E. DAVIES, AND J. S. THOMPSON, *Compressive and Noncompressive Power Spectral Density Estimation from Periodic Nonuniform Samples*, preprint, arXiv:1110.2722, 2011.
 - [27] Y. LI, E. PERLMAN, M. WAN, Y. YANG, C. MENEVEAU, R. BURNS, S. CHEN, A. SZALAY, AND G. EYINK, *A public turbulence database cluster and applications to study Lagrangian evolution of velocity increments in turbulence*, J. Turbulence, 9 (2008).
 - [28] S. G. MALLAT, *Multiresolution approximations and wavelet orthonormal bases of $L^2(\mathbf{R})$* , Trans. Amer. Math. Soc., 315 (1989).
 - [29] S. G. MALLAT, *A Wavelet Tour of Signal Processing: The Sparse Way*, 3rd ed., Elsevier Computer Science Library, Academic Press, Amsterdam, 2009.
 - [30] D. NEEDELL AND J. A. TROPP, *CoSaMP: Iterative signal recovery from incomplete and inaccurate samples*, Appl. Comput. Harmon. Anal., 26 (2008), pp. 301–321.
 - [31] H. NOBACH, E. MÜLLER, AND C. TROPEA, *Refined reconstruction techniques for LDA data analysis*, in Proceedings of the 8th International Symposium on Applications of Laser Techniques to Fluid Mechanics, Lisbon, Portugal, 1996.
 - [32] E. PERLMAN, R. BURNS, Y. LI, AND C. MENEVEAU, *Data exploration of turbulence simulations using a database cluster*, in Proceedings of the 2007 ACM/IEEE Conference on Supercomputing, SC '07, ACM, New York, 2007, pp. 23:1–23:11.
 - [33] V. PERRIER, T. PHILIPOVITCH, AND C. BASDEVANT, *Wavelet spectra compared to Fourier spectra*, J. Math. Phys., 36 (1995), pp. 1506–1519.
 - [34] J. ROMBERG, *Compressive sensing by random convolution*, SIAM J. Imaging Sci., 2 (2009), pp. 1098–1128.
 - [35] K. SCHNEIDER AND O. V. VASILYEV, *Wavelet methods in computational fluid dynamics*, Ann. Rev. Fluid Mech., 42 (2010), pp. 473–503.

- [36] J. A. TROPP AND A. C. GILBERT, *Signal recovery from random measurements via orthogonal matching pursuit*, IEEE Trans. Inform. Theory, 53 (2007), pp. 4655–4666.
- [37] J. A. TROPP, M. B. WAKIN, M. F. DUARTE, D. BARON, AND R. G. BARANIUK, *Random filters for compressive sampling and reconstruction*, in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2006, pp. III-872–III-875.
- [38] R. VENKATARAMANI AND Y. BRESLER, *Further results on spectrum blind sampling of 2D signals*, in Proceedings of the 1998 International Conference on Image Processing, Vol. 2, IEEE, Piscataway, NJ, 1998, pp. 752–756.
- [39] J. WANG, *Generalized orthogonal matching pursuit*, IEEE Trans. Signal Process., 60 (2012), pp. 6202–6216.
- [40] J. WANG AND B. SHIM, *On the recovery limit of sparse signals using orthogonal matching pursuit*, IEEE Trans. Signal Process., 60 (2012), pp. 4973–4976.