Basis operator bispectral analysis

D. A. Baver,[1] P. W. Terry,[1] and C. Holland[2]
[1]*Department of Physics, University of Wisconsin, Madison, Wisconsin 53706, USA*
[2]*Center for Energy Research, University of California, San Diego, California 92093, USA*

A new procedure for calculating model coefficients from fluctuation data for fully developed turbulence is derived. This procedure differs from previous related methods in that it is constructed in a spatial rather than spectral representation. This has a number of advantages, such as reduced data set requirements, ability to represent spatially inhomogeneous systems such as the ones with curvature or zonal flows, and ability to use data from experimental diagnostics with limited spatial resolution. In this method, the model equation is represented as a linear superposition of linear and nonlinear differential operators. The coefficients of this superposition are calculated using a least-squares method. This method has been tested on simulations of fully developed two dimensional turbulence and compared to previous methods. © *2009 American Institute of Physics*.
[DOI: 10.1063/1.3093844]

I. INTRODUCTION

Bispectral analysis is a term used to describe a variety of analysis techniques. In plasma physics, it typically refers to methods that exploit three-wave correlations, also known as the bispectrum. The most common methods of this type use the bispectrum along with some normalizing factors to estimate the strength of three-wave coupling.[1–6] While these methods are simple and easy to apply, they fail to exploit the full potential of analyzing the bispectrum. A second, less commonly used approach uses the bispectrum along with other statistical moments (up to fourth order) to calculate not only coupling strengths but linear growth rates and frequencies as well.[7–10] It is this last class of algorithms that we will be looking at in this paper.

In theory, fully quantitative methods of bispectral analysis offer a number of advantages over other methods of turbulence analysis. First of all, because it directly infers the coefficients of the underlying turbulence equations rather than examines their effects, it invokes a one-to-one correspondence between source equations and output solutions. This contrasts with comparisons between experiment and simulation or experiment and analytic theory, in which more than one set of growth rates could produce a similar-looking spectrum. As a result, it becomes possible to obtain detailed information about the functional form of the driving instability.[11] Second, bispectral analysis offers more detailed information about nonlinear mode coupling than other approaches. This means that one does not assume *a priori* the structure of the nonlinearity. Third, it permits the examination of nonlinear instability and damping, which manifest themselves as finite-amplitude-induced changes to the growth rate,[12] and the nonlinear decorrelation rate. Measurement of these quantities offers access to key effects such as nonlinear damping of zonal modes[13] and the scaling of the turbulent decorrelation rate with global quantities such as the magnetic field.[14] Fourth, by combining these pieces of information, one can infer the actual equations governing the turbulence.

In practice, the utility of these methods tends to be hampered by the requirements they place on measurements. First, there must be measurements corresponding to each of the dynamically active linear eigenmodes to be analyzed. For many important types of plasma turbulence, dynamics are essentially two dimensional (2D) due to the effect of the magnetic field. However, many diagnostics capable of accessing the core are only capable of one dimensional fluctuation measurements. This creates considerable uncertainty in the inferred growth rates.[11] In addition, if there are multiple interacting fields, this can result in more than one dynamically important branch in the linear dispersion relation. This necessitates the measurement of more than one field. Both of these factors increase not only the amount of data that needs to be gathered but also the complexity of the experimental apparatus required to gather it. Second, one must have data not only for the eigenmodes of interest but also for every mode they nonlinearly interact with. This can be a problem if a diagnostic only covers a small portion of the plasma. Third, until the development of the basis function method,[15] bispectral analysis required long-time-series data. In addition to these problems related to experimental measurements, bispectral methods are limited because they essentially fit coefficients to an analytic formula. Thus, any effect that is not included in the analytic formula cannot be inferred. In traditional spectral algorithms, this included things such as geometric effects.

In this paper we present a new class of bispectral algorithms designed to solve some of these fundamental problems. By operating in a spatial rather than spectral regime, we will be able to make calculations that depend only on local dynamics. This allows us to account for a variety of spatially inhomogeneous behavior: curvature of a cylindrical coordinate system, zonal flows, and variations in underlying gradients driving the turbulence. It is also well suited to analysis of experimental data in which the diagnostics only have a limited number of data channels in one direction since

we only need data from a small neighborhood around each point of interest.

The new method retains many of the advantages of the existing basis function method, such as ability to work with multiple coupled fields and ability to work with short time series data. This comes at the price of sharing many of its disadvantages. As we will see, it is necessary to make some assumptions about the system to be analyzed in order to construct our operator basis. This means there is a risk of "confirming" false assumptions if sufficient care is not taken in interpreting the results. However, these problems can be minimized by using as large an operator set as is practical and by using quality-of-fit estimation to look for signs that a particular set of operators is insufficient.

II. BACKGROUND

In order to understand the basis operator method, it is helpful to first understand the spectral methods that it is related to. We will therefore begin with a discussion of prior art in the field of bispectral analysis. The line of research that we are interested in begins with the Ritz method developed by Ritz and colleagues.[7–9] It continues with the modified Ritz method developed by Kim and co-workers.[10,11] The most recent addition to this line consists of the basis function method developed by Baver and Terry.[15]

In each of these methods, the idea is to assume that the turbulence observed obeys some predefined equation with undetermined coefficients, then solve for these coefficients using statistical correlations in the data set. Both the Ritz and modified Ritz methods use the same basic turbulence model and notation, so we will discuss that notation here. The basis function method uses a different model, which will be discussed in that section. The generic turbulence equation used in the former two methods is as follows:

$$Y_k = L_k X_k + \sum_{\substack{k_1 \geq k_2 \\ k=k_1+k_2}} Q_k^{k_1,k_2} X_{k_1} X_{k_2}, \tag{1}$$

where $X_k = \phi(k,t)$ and $Y_k = \phi(k,t+\tau)$, with ϕ as the measured fluctuating quantity as a function of wavenumber and time. From its form, it is evident that this equation is a difference-equation representation in the temporal domain of a first-order-in-time nonlinear partial differential equation. The coefficient L_k determines the growth rate of the turbulence, with $\gamma_k \approx (|L_k|^2 - 1)/\tau$, and the coefficient $Q_k^{k_1,k_2}$ is the coefficient of the quadratic nonlinearity responsible for nonlinear transfer under model conditions. The coefficients L and Q are what these algorithms attempt to calculate.

A. Ritz method

The first method for quantitatively estimating such information in plasma was developed by Ritz and co-workers[7–9] in the late 1980s. A review of this method is presented in the introduction of an article on the modified Ritz method by Kim *et al.*[10]

The Ritz method solves for the growth rates and transfer functions by expanding the model equation in a series of moment equations, multiplying by X_k^* and $X_{k_1}^* X_{k_2}^*$, respec-

tively. The fourth order moments in this series are approximated as products of second order moments. This approximation, which is commonly used in analytic turbulence theory, is known as the Millionshchikov approximation and is derived by assuming a nearly Gaussian distribution in the fluctuating quantities. This results in the following equations:

$$\langle Y_k X_k^* \rangle = L_k \langle X_k X_k^* \rangle + \sum_{\substack{k_1 \geq k_2 \\ k=k_1+k_2}} Q_k^{k_1,k_2} \langle X_{k_1} X_{k_2} X_k^* \rangle, \tag{2}$$

$$\langle Y_k X_{k_1}^* X_{k_2}^* \rangle = L_k \langle X_k X_{k_1}^* X_{k_2}^* \rangle + Q_k^{k_1,k_2} \langle |X_{k_1} X_{k_2}|^2 \rangle. \tag{3}$$

Using these equations, Ritz *et al.* proceeded to solve for L_k, yielding the following:

$$L_k = \frac{\langle X_k^* Y_k \rangle - \sum_{\substack{k_1 \geq k_2 \\ k=k_1+k_2}} \dfrac{\langle X_k^* X_{k_1} X_{k_2} \rangle \langle Y_k X_{k_1}^* X_{k_2}^* \rangle}{\langle |X_{k_1} X_{k_2}|^2 \rangle}}{\langle X_k^* Y_k \rangle - \sum_{\substack{k_1 \geq k_2 \\ k=k_1+k_2}} \dfrac{|\langle X_k^* X_{k_1} X_{k_2} \rangle|^2}{\langle |X_{k_1} X_{k_2}|^2 \rangle}}. \tag{4}$$

This method has some significant disadvantages when applied to measured fluctuation data. In particular, as noted by Ritz, it can yield unphysically large damping coefficients at all wavenumbers. This problem arises because the method does not take into account nonideal fluctuations, that is to say, deviations of the data from the physics described by the model equation. Such deviations can arise from noise, measurement error, or interactions of the fluctuating quantities with physical effects outside of the scope of the model. This issue is addressed by the modified Ritz method.

B. Modified Ritz method

This method was developed by Kim *et al.*[10] and is described in more detail in that article. The modified Ritz method applies to the generic turbulence equation [Eq. (1)] and additionally assumes that each of the measured spectra can be divided into an ideal and nonideal spectra,

$$X_k = \beta_k + X_k^{ni}, \quad Y_k = \alpha_k + Y_k^{ni}, \tag{5}$$

where (X_k, Y_k) are the measured spectra at time t, $t+\Delta t$, respectively; (β_k, α_k) are the ideal spectra at the same times; and (X_k^{ni}, Y_k^{ni}) are the nonideal spectra.

From this Kim *et al.* derived moment equations for the ideal and nonideal spectra and then drop all cross terms involving the nonideal spectrum. Unfortunately, this approach results in equations to which the Millionshchikov approximation cannot be applied, thus entailing the increased computational cost of calculating the fourth order moments. Moreover, since the Q's in Eqs. (2) and (3) refer to different pairs k_1 and k_2, a matrix notation is needed to represent the equations for the third and fourth order moments. This is solved by introducing the following notation:

$$\mathbf{Q} = (Q_l^{(l+2i)/2,(l-2i)/2}),$$

$$\mathbf{A} = (\langle X_{(l+2i)/2} X_{(l-2i)/2} X_l^* \rangle),$$

$$\mathbf{B} = (\langle X_{(l+2i)/2} X_{(l-2i)/2} Y_l^* \rangle),$$

$$F = (\langle X_{(l+2i)/2} X_{(l-2i)/2} X_{(l+2j)/2} X_{(l-2j)/2} \rangle),$$

where l is the index of the mode wavenumber k ($k=f(l)$, where f is a linear function), and i and j are indices of the resulting tensors.

An additional constraint is required to solve this system. This constraint is supplied by the assumption that the turbulence is steady state, that is, $\langle \alpha_k \alpha_k^* \rangle = \langle \beta_k \beta_k^* \rangle$. This allows us to obtain the following expressions for L_k:

$$L_k = \frac{\langle Y_k X_k^* \rangle - (\mathbf{B}^*)^T \cdot F^{-1} \cdot \mathbf{A}}{\langle \beta_k \beta_k^* \rangle - (\mathbf{A}^*)^T \cdot F^{-1} \cdot \mathbf{A}}, \tag{6}$$

$$L_k = \frac{\langle \alpha_k \alpha_k^* \rangle - (\mathbf{B}^*)^T \cdot F^{-1} \cdot \mathbf{B}}{\langle X_k Y_k^* \rangle - (\mathbf{A}^*)^T \cdot F^{-1} \cdot \mathbf{B}}. \tag{7}$$

Combining these gives a formula for the growth rate γ,

$$\gamma_k = \frac{(\mathbf{A}^*)^T \cdot F^{-1} \cdot \mathbf{A} - (\mathbf{B}^*)^T \cdot F^{-1} \cdot \mathbf{B}}{\langle \beta_k \beta_k^* \rangle - (\mathbf{A}^*)^T \cdot F^{-1} \cdot \mathbf{A}}. \tag{8}$$

This method produces more accurate fits than the Ritz method, but a long-time series is still required to produce an accurate fit. This limitation was the primary motivation for the development of the basis function method.

C. Basis function method

The basis function method was developed recently to address some of the limitations of the previous methods of bispectral analysis.[15] Its main purpose is to reduce the number of pairs of time steps (realizations) required to get a numerically well-posed solution and to generally increase the accuracy of the fit. The key realization here is that the modified Ritz method is, in fact, related to the technique of least-squares fitting, and is therefore sensitive to the number of degrees of freedom in the solution. In the basis function method, this sensitivity is addressed by artificially constraining the solution so as to enforce some condition that permits physically realistic solutions, while reducing the number of variables needed to describe that solution. This reduces the size of the input data set needed to yield a unique and well-posed solution.

The method employed to impose this constraint is to limit the nonlinear terms to an incomplete set of basis functions. In doing so, one permits nonlinearities with a smooth functional form but prohibits more irregular nonlinear forms that could, in principle, be extracted to fit the data. Smooth nonlinearities can be represented as linear superpositions of members of the basis set, and with a sufficient number of basis functions, one can represent a wide variety of nonlinear coupling functions while still having far fewer degrees of freedom than required to describe the linear terms.

In addition, the basis function method is constructed in such a way as to be easily generalizable to multiple fields. The previous methods could also be theoretically generalized in that sense, but the basis function method, being more accurate, can better deal with the complications presented by multifield data sets.

The basis function method begins with a modified model equation,

$$\hat{Y}_i^k \equiv \widetilde{Y}_i^k + Y_i^k = D_{ij}^k X_i^k + \sum_{k'} \alpha_\mu \beta_{\mu i l m}^{k,k'} X_l^{k'} X_m^{k-k'}, \tag{9}$$

where repeated indices indicate summation over the index. The values X and Y in this case are field vectors at times t and $t+\delta t$, with the subscript indicating which field is referenced, i.e., 1 gives density, 2 gives potential, etc., or whatever labeling scheme is appropriate for a particular system. The values $\beta_{\mu i l m}^{k,k'}$ are a predefined basis of functions used to represent nonlinearities. The coefficients α_μ are the amplitudes of the linear superposition of these basis functions, which, along with the linear coupling coefficients D_{ij}^k, are the quantities solved for by this method. This type of model equation is suitable for fluid models with an arbitrary number of equations and quadratic nonlinearities.

Using our model equation to define the error $\widetilde{Y}$, we can now minimize this error using the variational principle. The procedure for doing this is covered in more detail in the original article on the basis function method. The end result is a formula for the coefficients α_μ and D_{ij}^k,

$$D_{lm}^k = (I_{ij,lm}^{k',k} a_{jn}^{k'} - A_{vin}^{*k'} F_{v\mu}^{-1} A_{\mu lm}^k)^{-1} (b_{in}^{k'} - A_{vin}^{*k'} F_{v\mu}^{-1} B_\mu), \tag{10}$$

$$\alpha_\nu = F_{v\mu}^{-1} (B_\mu - D_{ij}^k A_{\mu ij}^k), \tag{11}$$

where $F_{\mu\nu} \equiv \Sigma_{k,k',k',i} \beta_{\mu i l m}^{*k,k'} \beta_{vinp}^{k,k'} \langle X_n^{k'} X_p^{k-k'} X_l^{*k'} X_m^{*k-k'} \rangle$, $A_{\mu i j}^k \equiv \Sigma_{k'} \beta_{\mu i l m}^{*k,k'} \langle X_j^k X_l^{*k'} X_m^{*k-k'} \rangle$, $B_\mu \equiv \Sigma_{k,k',i,j} \beta_{\mu i l m}^{*k,k'} \langle Y_j^k X_l^{*k'} X_m^{*k-k'} \rangle$, $a_{ij}^k \equiv \langle X_i^k X_j^{*k} \rangle$, $b_{ij}^k \equiv \langle Y_i^k X_j^{*k} \rangle$, and $I_{ij,lm}^{k',k} \equiv 1$ if $ij=lm$, $k=k'$, or 0 otherwise, and "$*$" denotes a complex conjugate.

When tested against simulation data, this method is capable of reconstructing model coefficients with an exceptionally high degree of accuracy. It is also capable of producing meaningful fits using very small amounts of data, as few as six time steps in the case of a two-field model and four if there is only a single field. Larger numbers of time steps do improve noise rejection capabilities, but data requirements are still not nearly as large as with the previous methods.

While reducing the need for long-time series is certainly an advantage, the requirement that most hampers the application of this method to experiment is that of 2D spectral data at a resolution appropriate to the physics of the underlying system, i.e., data that are capable of resolving both the spectral range responsible for instability and that responsible for dissipation. While it is common to find experimental diagnostics with sufficient resolution in one direction due to

the common effect of plasma rotation, finding diagnostics that can achieve good resolution in two different directions simultaneously is rather more difficult. It is this problem that forms the motivation for the development of the basis operator method.

III. BASIS OPERATOR METHOD

A. Motivation

In reviewing the aforementioned methods, it becomes evident that all share a common limitation: they all assume that the system being analyzed is susceptible to a spectral representation. In the case of homogeneous turbulence, this assumption is at least approximately valid; while the boundary conditions will not be the same as a periodic box, the behavior within those boundaries will still fit the model equations. However, if the system is inhomogeneous, this assumption will break down completely.

In an inhomogeneous system, there will be terms in the model equation that depend on position. In order to include these terms, one must do one of two things. One option is to represent these inhomogeneous terms in spectral form. The other option is to convert the entire model equation into a purely spatial form. Of these, converting the model equation to spatial form has several advantages. First, it eliminates issues pertaining to boundary conditions, as one can simply not evaluate the resulting algorithm close to boundaries. Second, it eliminates any restriction on what type of inhomogeneities can be included in the model, which proves convenient in the case of cylindrical coordinates. Third, it results in an algorithm of superior computational efficiency. To evaluate nonlinear terms efficiently, it is necessary to use a pseudospectral method, which converts from Fourier space to real space, evaluates nonlinear terms in real space, then converts the answer back to Fourier space. Evaluating linear terms in real space eliminates two steps in this process.

The drawback to this approach is that certain types of linear behavior (such as resonances) may be difficult to model. Also, it necessitates the use of a basis function representation for both linear and nonlinear terms. This means that issues pertaining to basis selection are more serious. Basis representation of the linear terms is necessary because otherwise the resulting algorithm is nonlocal, and thus gives up any computational advantages over a pseudospectral code. While these issues may result in a significantly less accurate fit in some cases, they are a small price to pay for the ability to analyze systems where spectral methods may not work at all.

B. Derivation

The first step in deriving the basis operator method is to make a further modification to the model equation for the basis function method. This is done partly for the purpose of adjusting notation, but mainly in anticipation of the incorporation of a basis function representation for linear terms. The modified model equation is as follows:

$$P^i(\vec{k})\dot{\psi}_k^j = \beta_\nu D_\nu^{ij}(\vec{k})\psi_k^j + \alpha_\mu Q_\mu^{ilm}(\vec{k},\vec{k}')\psi_{k'}^l\psi_{k-k'}^m, \qquad (12)$$

where i, j, l, and m are field labels. The changes in notation compared to Eq. (9) are to avoid confusion later in the derivation. β is now a model coefficient rather than a basis set. This choice is made by analogy since we are already using α as a model coefficient. The role of β in the nonlinear term is now replaced by $Q_\mu^{ilm}(\vec{k},\vec{k}')$. D is now a basis set rather than a model coefficient. The dynamical variables are represented by the field vector ψ rather than by X and Y. This is done because x will soon be used as a spatial coordinate.

In addition to the above changes in notation, a preoperator $P^i(\vec{k})$ has been introduced. This is done by simply multiplying both sides by a function of k, then incorporating the resulting terms on the right hand side into the definition of the basis sets D and Q. The purpose of the preoperator is to avoid basis functions that have reciprocal dependence on k (i.e., vary with some negative power of the wavenumber). This is important because when we transform to real space, any negative powers of k turn into integral rather than differential operators, resulting in integrodifferential equations that are difficult to solve; eliminating these terms allows the model equation to convert to simply a differential equation in real space. One example of a system where this is an issue is the Hasegawa–Wakatani equation, in which the ϕ equation can be written either with a factor of $1/k^2$ on the right hand side or a factor of k^2 on the left hand side. Forcing the equation into a form consistent with Eq. (9) results in the former; the task of converting to real space is much simpler with the latter.

Once these changes are made, we can now transform the model equation into real space. To do this, we first make a Taylor expansion in k_x and k_y for each of the k-dependent terms (P,D,Q) in Eq. (12). Once this is done, each power of k_x or k_y is replaced by an appropriate order derivative in that direction. For instance, suppose the model equation has the following structure as a polynomial in k:

$$(1+k^2)\dot{\phi}_k = ik_y\phi_k + \sum_{k'}(k\times k')(k'^2)\phi_{k'}\phi_{k-k'}. \qquad (13)$$

Then the resulting spatial equation is as follows:

$$(1-\nabla^2)\dot{\phi}(\vec{x}) = \partial_y\phi(\vec{x}) + \partial_x\nabla^2\phi\partial_y\phi - \partial_y\nabla^2\phi\partial_x\phi. \qquad (14)$$

When Eq. (12) has been converted along these lines, the resulting generic spatial model equation is as follows:

$$P^i(\dot{\psi}^j;\vec{x}) = \beta_\nu D_\nu^{ij}(\psi^j;\vec{x}) + \alpha_\mu Q_\mu^{ilm}(\psi^l,\psi^m;\vec{x}), \qquad (15)$$

where $D_\nu^{ij}(\psi^j;\vec{x})$, $Q_\mu^{ilm}(\psi^l,\psi^m;\vec{x})$, and $P^i(\dot{\psi};\vec{x})$ are now differential operators rather than functions of k, and ψ^j is the value of the field vector at location $\vec{x}$. From here we can define a variance to be minimized, equal to the mismatch between the two sides of Eq. (15),

$$\chi^i = \beta_\nu D_\nu^{ij}(\psi^j;\vec{x}) - P^i(\dot{\psi};\vec{x}) + \alpha_\mu Q_\mu^{ilm}(\psi^l,\psi^m;\vec{x}). \qquad (16)$$

Much as with the original basis function method, we then derive a formula for the coefficients in this equation by minimizing the variance using the variational principle. This results in the following moment equations:

$$\delta\beta_\lambda\{D_\lambda^{in}(\psi^n;\vec{x})[\beta_\nu D_\nu^{ij}(\psi^j;\vec{x}) - P^i(\dot{\psi}^j;\vec{x})$$
$$+ \alpha_\mu Q_\mu^{ilm}(\psi^l,\psi^m;\vec{x})]\} = 0, \tag{17}$$

$$\delta\alpha_\lambda\{Q_\lambda^{inp}(\psi^n,\psi^p;\vec{x})[\beta_\nu D_\nu^{ij}(\psi^j;\vec{x}) - P^i(\dot{\psi}^j;\vec{x})$$
$$+ \alpha_\mu Q_\mu^{ilm}(\psi^l,\psi^m;\vec{x})]\} = 0. \tag{18}$$

These equations are then used to construct moment equations for α and β. This results in the following:

$$\beta_\nu a_{\nu\lambda} - b_\lambda + \alpha_\mu A_{\mu\lambda}^T = 0, \tag{19}$$

$$\beta_\nu A_{\nu\lambda} - B_\lambda + \alpha_\mu F_{\mu\lambda} = 0, \tag{20}$$

where $a_{\nu\lambda} = \langle D_\nu^{ij}(\psi^j;\vec{x})D_\lambda^{il}(\psi^l;\vec{x})\rangle$, $b_\lambda = \langle P^i(\dot{\psi}^j;\vec{x})D_\lambda^{il}(\psi^l;\vec{x})\rangle$, $A_{\nu\lambda} = \langle D_\nu^{ij}(\psi^j;\vec{x})Q_\lambda^{ilm}(\psi^l,\psi^m;\vec{x})\rangle$, $B_\lambda = \langle P^i(\dot{\psi}^j;\vec{x})Q_\lambda^{ilm}(\psi^l,\psi^m;\vec{x})\rangle$, $F_{\mu\lambda} = \langle Q_\nu^{ilm}(\psi^l,\psi^m;\vec{x})Q_\lambda^{inp}(\psi^n,\psi^p;\vec{x})\rangle$, and T denotes transpose.

From these equations, we are able to solve for the coefficients α and β,

$$\alpha_\mu = -F_{\mu\lambda}^{-1}(\beta_\nu A_{\nu\lambda} - B_\lambda), \tag{21}$$

$$\beta_\nu = (a_{\nu\lambda} - A_{\mu\lambda}^T F_{\mu\kappa}^{-1} A_{\nu\kappa})^{-1}(b_\lambda - A_{\mu\lambda}^T F_{\mu\kappa}^{-1} B_\kappa). \tag{22}$$

Once we have the values of α and β, these may be substituted into Eq. (15) in order to determine the equation obeyed by the turbulent system in question.

Constructing an algorithm to calculate the values of α and β from these formulas is a straightforward procedure. The precise algorithm which results from this, however, depends on the choice of basis operators.

C. Basis operator selection

In order for the solutions in Eqs. (21) and (22) to be useful, we must first define the terms P, D, and Q in Eq. (15). These are the basis operators on which this algorithm is based. There are many possible ways to define these operators, so we will concentrate on the definitions used in the algorithm applied in Sec. IV. The construction of these consists of three parts: the elementary differential operators used to calculate local derivatives, the arrangement of these elementary operators to form basis operators, and the indexing of the basis operators.

The elementary differential operators are the mechanism used in this algorithm to calculate spatial derivatives. In a typical experimental or simulation data set, the spatial derivatives required by Eq. (15) are not directly available. Therefore, we must calculate derivatives using finite difference methods.

To calculate spatial derivatives at a given point, we begin by sampling values within an $n \times m$ "patch" of the data set, centered at the point we want to find derivatives at. Since the algorithm uses statistical ensembles of correlations between different order derivatives to solve for the coefficients α and β, this patch will therefore be scanned across the data set. Within this patch, we can approximate spatial derivatives with matrix operations,

$$\partial_x^p \partial_y^q \psi^i(x_{(m+1)/2}, y_{(n+1)/2}) = M_{pj}^{(m)} M_{ql}^{(n)} \psi^i(x_j, y_l), \tag{23}$$

where the matrix $M^{(m)}$ is the elementary differential operator of the order of m. The subscript j ranges from 1 to m and the subscript l ranges from 1 to n, so $(m+1)/2$ and $(n+1)/2$ define the center of the patch. The differential operators in the x direction may or may not be different from the operators in the y direction; this depends on what type of data is being analyzed, and whether or not $m=n$. In all of the cases described in Sec. IV, the patch is square and operators of the same type are used in both directions. Note that the maximum order derivative that can be calculated by this method depends on the size of the patch. Thus, a patch of width one can only calculate zeroth order derivatives, a patch of width three yields up to second order derivatives, a width of five yields fourth order, and so on.

There are several ways to construct these operators, but only two are used in Sec. IV for testing the algorithm. The first of these is designed for more general use, particularly for tests against spectral simulations. It is also the more suitable of the two for analyzing experimental data. This method is optimized to isolate different order derivatives as much as possible given a particular size patch. The price paid for this is that the solution uses the entire patch to calculate any nonzero order derivative. In some cases, this may mean ascribing to the data set longer range spatial correlations than actually exist. However, in most cases this is not a significant drawback.

In this method, the goal is to construct a matrix such that multiplying it by a sequence of data values returns the Taylor series coefficients for that sequence. This can be found by taking a matrix that converts Taylor series coefficients to data values and inverting it. Therefore, if we construct a matrix consisting of powers of displacement from center and invert it, we will get the matrix we are looking for. In more specific terms,

$$M_{ij}^{(m)} = \left[\frac{(i-h)^{j-1}}{(j-1)!}\right]^{-1}, \tag{24}$$

where $m=2h+1$ and i and j ranged from 1 to m. Applying this formula for $m=7$ yields the following matrix:

	$\psi(x_{-3})$	$\psi(x_{-2})$	$\psi(x_{-1})$	$\psi(x_0)$	$\psi(x_1)$	$\psi(x_2)$	$\psi(x_3)$
∂_x^0	0	0	0	1	0	0	0
∂_x^1	-0.0166	0.15	-0.75	0	0.75	-0.15	0.0166
∂_x^2	0.0111	-0.15	1.5	-2.722	1.5	-0.15	0.0111
∂_x^3	0.125	-1	1.625	0	-1.625	1	-0.125
∂_x^4	-0.166	2	-6.5	9.333	-6.5	2	-0.166
∂_x^5	-0.5	2	-2.5	0	2.5	-2	0.5
∂_x^6	1	-6	15	-20	15	-6	1

$$(25)$$

The other method is designed for tests against simulation data generated via a finite difference method. Its purpose is to match the structure of differential operators commonly used in simulation codes. Since what order derivatives are used in the simulation is not known, this method is designed so that the matrix M_{li}^{m-2} can be found from the matrix M_{li}^m by simply removing excess rows and columns. That way, if we overestimate the order derivatives in the code, the algorithm will simply return near-zero values for the coefficients associated with these surplus operators. In other words, the algorithm does not assume longer range spatial correlations than what is in the simulation. The price paid for this is that it does not isolate different order derivatives. Thus, the value returned for a second order derivative will also depend on the fourth order derivative, the value returned for third order will also depend on fifth order, and so on. This results in a less accurate calculation of physical dispersion, but more accurately matches the numerical dispersion found in most simulation codes.

In this approach, the matrix $M^{(m)}$ is constructed iteratively. Starting with the unit matrix $M^{(1)}$, we add rows to the bottom and columns to either side. The elements that are in both new rows and new columns are initialized with values of 1 for odd rows and $\pm\frac{1}{2}$ for even rows. We then calculate the response of each row to derivatives of lower order and subtract a multiple of the appropriate previous row to remove this response. This yields the following formulas:

$$M_{ij}^{(1)} = 1,$$

$$M_{ij}^{(m)} = M_{i,j-1}^{(m-2)} + N_{ij}^{(m)}$$
$$- \sum_{n=1}^{m-2} M_{n,j-1}^{(m-2)} \frac{\sum_{l=1}^m N_{il}^{(m)}(l-h+1)^{n-1}}{\sum_{p=1}^{m-2} M_{n,p}^{(m-2)}(p-h-1)^{n-1}},\qquad (26)$$

$$N_{ij}^{(m)} = \begin{cases} 1 & \text{if } i=m \text{ and } j=1 \text{ or } m \\ \frac{1}{2} & \text{if } i=m-1 \text{ and } j=m \\ -\frac{1}{2} & \text{if } i=m-1 \text{ and } j=1 \\ 0 & \text{otherwise,} \end{cases}$$

where $m=2h+1$ and i and j ranged from 1 to m. Note that m can only take odd values with this method, so a definition of the matrix M_{ij}^0 is not needed. Applying these formulas for $m=7$ yields the following:

	$\psi(x_{-3})$	$\psi(x_{-2})$	$\psi(x_{-1})$	$\psi(x_0)$	$\psi(x_1)$	$\psi(x_2)$	$\psi(x_3)$
∂_x^0	0	0	0	1	0	0	0
∂_x^1	0	0	-0.5	0	0.5	0	0
∂_x^2	0	0	1	-2	1	0	0
∂_x^3	0	-0.5	1	0	-1	0.5	0
∂_x^4	0	1	-4	6	-4	1	0
∂_x^5	-0.5	2	-2.5	0	2.5	-2	0.5
∂_x^6	1	-6	15	-20	15	-6	1

$$(27)$$

Once the elementary differential operators have been constructed, we can now define the operators D, Q, and P in terms of them. The linear operators D_μ are the simplest, as it is generally not necessary to exclude any possible operators for the sake of computational or data use efficiency. Thus, all we have do is take every derivative that can be calculated for a given size patch and assign it an index. Since no possible terms are excluded, this basis set makes no assumptions about the underlying equations except that it can be adequately approximated by a finite number of terms in a Taylor expansion.

The nonlinear operators are rather more complicated, as there are more possible operators than it is possible or practical to implement in the code. Much as with the original basis function method, this is where assumptions about the underlying physics of the system under study are brought to bear. In the version of the algorithm used in this paper, nine types of nonlinearities are permitted,

$$Q_1 = (k \times k')\psi_{k'}^{(1)}\psi_{k-k'}^{(2)} = \partial_x\psi_1\partial_y\psi_2 - \partial_y\psi_1\partial_x\psi_2,$$

$$Q_2 = (k \times k')(k'^2)\psi_{k'}^{(1)}\psi_{k-k'}^{(2)} = \partial_x\nabla^2\psi_1\partial_y\psi_2 - \partial_y\nabla^2\psi_1\partial_x\psi_2,$$

$$Q_3 = (k \times k')(k^2 k'^2)\psi_{k'}^{(1)}\psi_{k-k'}^{(2)}$$
$$= \nabla^2(\partial_x\nabla^2\psi_1\partial_y\psi_2 - \partial_y\nabla^2\psi_1\partial_x\psi_2),$$

$$Q_4 = (k \times k')(k^2)\psi_{k'}^{(1)}\psi_{k-k'}^{(2)} = \nabla^2(\partial_x\psi_1\partial_y\psi_2 - \partial_y\psi_1\partial_x\psi_2),$$

$$Q_5 = (k \times k')(k^4)\psi_{k'}^{(1)}\psi_{k-k'}^{(2)} = \nabla^4(\partial_x\psi_1\partial_y\psi_2 - \partial_y\psi_1\partial_x\psi_2),$$

$$(28)$$

$$Q_6 = (k \times k')(k'^2 k_x) \psi_{k'}^{(1)} \psi_{k-k'}^{(2)}$$

$$= \partial_x (\partial_x \nabla^2 \psi_1 \partial_y \psi_2 - \partial_y \nabla^2 \psi_1 \partial_x \psi_2),$$

$$Q_7 = (k \times k')(k'^2 k_y) \psi_{k'}^{(1)} \psi_{k-k'}^{(2)}$$

$$= \partial_y (\partial_x \nabla^2 \psi_1 \partial_y \psi_2 - \partial_y \nabla^2 \psi_1 \partial_x \psi_2),$$

$$Q_8 = (k \times k')(k_x) \psi_{k'}^{(1)} \psi_{k-k'}^{(2)} = \partial_x (\partial_x \psi_1 \partial_y \psi_2 - \partial_y \psi_1 \partial_x \psi_2),$$

$$Q_9 = (k \times k')(k_y) \psi_{k'}^{(1)} \psi_{k-k'}^{(2)} = \partial_y (\partial_x \psi_1 \partial_y \psi_2 - \partial_y \psi_1 \partial_x \psi_2).$$

In each of these cases, the superscripts 1 and 2 in wave-number space and corresponding subscripts in real space correspond to field indices, i.e., $\psi_1 = \phi$, and $\psi_2 = n$. In cases where the nonlinearity is symmetric in k', only cases where the field indices are unequal are allowed since otherwise these terms will cancel when summed over all possible values of k'. If the nonlinearity is not symmetric in k', then all possible values of the field indices are allowed. This results in a total of 42 different nonlinearities, each of which is assigned an index and a corresponding coefficient α_μ. For the simulation results analyzed in this article, these nonlinearities are more than sufficient since only a handful are actually used in the simulations themselves. The extra nonlinearities serve to measure the completeness of the basis set, as significant amplitude in these coefficients indicates an inaccurate fit that might be improved by adding more terms. However, for application to experiment, where the structure of the nonlinearity is not as well known, additional terms are advised.

The preoperator is a tool for dealing with particular types of systems of equations. Certain common turbulence models (such as Hasegawa–Wakatani, as employed in Sec. IV) have dispersion relations that are singular at zero wavenumber. This is due to differential operators applied to the time derivative of the fluctuating quantities (the left hand side of the equation), as opposed to the fluctuating quantities themselves (the right hand side of the equation). As a result, if we attempt to represent such a system by a model equation that does not contain derivatives on the left hand side, the linear terms cannot be represented by a series expansion in wavenumber. However, if we differentiate both sides of the equation, a series expansion becomes once again possible.

How the resulting operators are labeled is not important in a homogeneous system; the order must be remembered when interpreting the results, but interchanging both coefficients and their corresponding operators does not change the resulting inferred equation. However, in inhomogeneous systems it is useful to create additional labels for operators applied to different parts of the data set. For instance, we might create a set of linear differential operators that are identical in their functional form, but are set to zero except at a single radial position. Each of these operators corresponds to spatial derivatives at a specific radial position, so the corresponding coefficients will give the corresponding derivatives as a function of radius. This capability is useful for studying systems with shear flow or other inhomogeneities that are of particular interest.

IV. TEST RESULTS

In order to test this algorithm, we apply it to simulations of fully developed turbulence. Three tests are listed in this section, all of which consist of 2D two-field models. The first test employs a trapped electron mode (TEM) model. This model is described in more detail in the basis function paper.[15] The other two employ a local Hasegawa–Wakatani model.[16] The two Hasegawa–Wakatani simulations differ in that one of them contains additional terms to account for shear flow, thus introducing a spatial inhomogeneity.

A. Comparison of procedure with spectral basis function method

The first test is a comparison of fit results between the basis function (original) and basis operator (new) methods. This is applied to data from a collisionless TEM simulation; these particular data were chosen mainly for convenience, as it was readily available from previous related work. It also used sufficiently short time steps to permit accurate estimation of time derivatives. Moreover, having intimate familiarity with the simulation model meant that the quality of fit could be evaluated accurately. Since these data were already used for tests of the basis function method, its suitability for bispectral analysis was known in advance.

Since these data were from a spectral simulation, applying the basis operator method to it requires that it first be Fourier transformed to real space. This was done to convert the 33×33 spectral data into 64×64 spatial data. Further conversion of the output was required in order to compare to the basis function results. The basis operator method outputs a list of coefficients for its various operators, which represent a series expansion of the underlying dispersion relation with respect to wavenumber. This type of information is difficult to interpret without familiarity with the operator set in question and is therefore of little value in interpreting the results. To solve this problem, the basis operator coefficients were used to reconstruct a spectral dispersion relation. This could then be compared directly to the output of the basis function method, which also produces coefficients which are functions of wavenumber.

Because the TEM dispersion relation does not contain a singularity at zero, the preoperator was not used. The differential operator set is given by Eq. (25), as constructed for a 7×7 patch. Nonlinear operators are given in Eq. (28).

Some of the results of this comparison are shown in Fig. 1. The horizontal axis shows wavenumber in the y direction with $k_x = 0$. The vertical axis shows one of the eight components of the linear coupling matrix (four real, four imaginary for a two-field system). The comparison shown is between the basis operator method and the prior basis function method. Since the results from the basis function method have already been demonstrated to be accurate to several significant figures,[15] the basis function result can be taken as a proxy for the actual model coefficients; a line for the actual model coefficients would simply overlap with the line for the spectral result and is therefore omitted for clarity. The plot shows some variation between the two results. This is to be expected since the model coefficients include artificial damp-
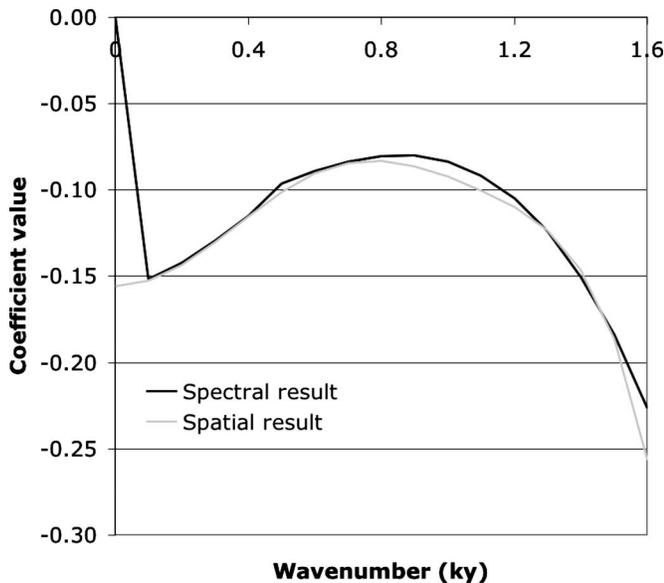
FIG. 1. Real part of phi-phi coefficient from a comparison of spectral and spatial versions of the basis function method on TEM simulation data for $kx=0$.

FIG. 2. ϕ-n coefficient, imaginary part, from a test of basis operator method on Hasegawa–Wakatani simulation data at $kx=0$.

ing that does not have a smooth functional form. Moreover, the TEM dispersion relation does not have an exact polynomial representation, so the finite number of terms permitted by a 7×7 patch (up to sixth power in wavenumber) means the fit will not be exact. Nonetheless, the fit displayed is close enough to confirm that the basis operator method is able to achieve a useful degree of accuracy.

Another measure of the quality of fit is to take the time derivative predicted by substituting the results in Eqs. (21) and (22) into Eq. (15), then take the difference between that and the actual derivatives in the data set. This approach is generally useful for any form of basis function or basis operator method because it allows one to determine whether the basis set is adequate; if the basis set is too incomplete (if there are important processes not represented by one or more basis functions or operators), then the measured quality of fit (rms variance) will be large. If there are too many basis functions relative to the size of the data set, then the rms variance will be small, but the results still appear noisy.

In the case of the above comparison, the actual measured variance is 1.4×10^{-3}. This indicates a relatively accurate fit.

B. Tests on homogeneous Hasegawa–Wakatani turbulence

The second test employs data from a simulation of Hasegawa–Wakatani turbulence. This model was chosen based on its similarity to turbulence predicted to exist in the controlled shear decorrelation experiment,[17] which is a candidate for potential application of the basis operator method to experimental data.

Because the Hasegawa–Wakatani model is singular at small wavenumber, we must use the preoperator in this comparison. The specific type of preoperator places a ∇^2 operator on the ϕ terms and no operator on the n terms. Also, because the basis function method has never been applied to
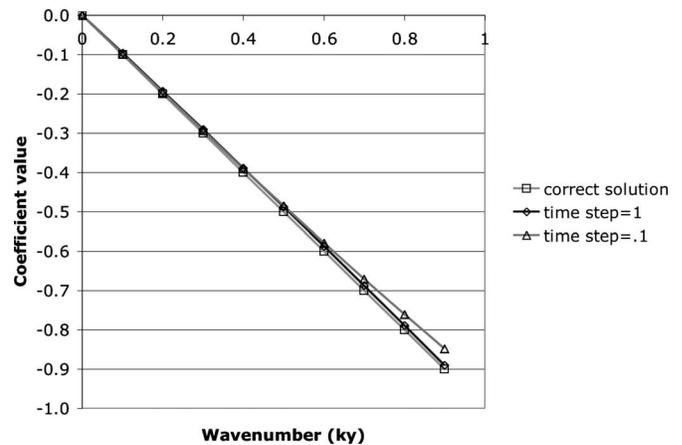
Hasegawa–Wakatani simulations, we cannot compare the two methods. Comparison therefore is between the output from the basis operator method and the coefficients of the original model equations. As with the previous comparison, both of these are converted into spectral dispersion relations for clarity. In addition, while converting the results into spectral form we also correct for numerical dispersion resulting from the size of time intervals in the data set.

The results of this comparison are shown in Figs. 2 and 3. These graphs each compare one of the eight components of the linear coupling matrix, but as we can see, the quality of fit is significantly different for different components. In the case of Fig. 2, the fit is very close to theoretical predictions. In the case of Fig. 3, the fit is significantly less accurate. Naturally, the question arises as to why this is the case.

As with the TEM comparison, a quality of fit estimate was used on these data. In the long-time-step case, this gave a result of 6.9×10^{-4}. This is a somewhat better fit than the TEM case and is moreover orders of magnitude smaller than the error evident from the plot. Thus, we can rule out random error. The question is what type of systematic error is at work here?
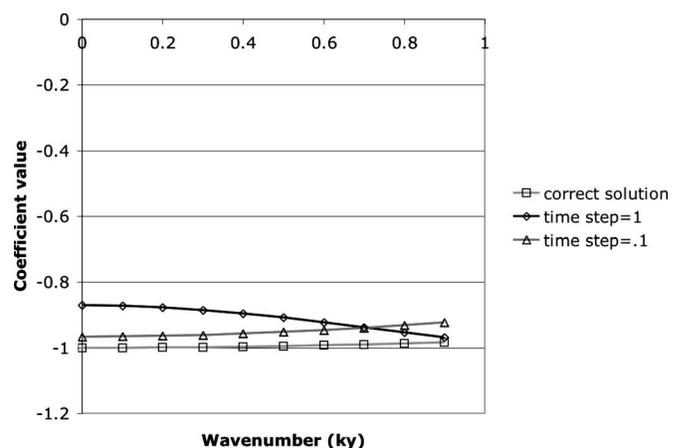

FIG. 3. n-n coefficient, real part, from a test of basis operator method on Hasegawa–Wakatani simulation data at $kx=0$.

As it turns out, there is a very significant difference between the data set used in the TEM case and the data set used in the Hasegawa–Wakatani case. That difference is that the TEM data were sampled at shorter time intervals. Shorter time intervals means that time derivatives, which in this algorithm are calculated by a first-order accurate differencing method, are therefore more accurate. In order to test the effect of time intervals on accuracy of fit, the algorithm was applied to a second data set with shorter time intervals. The results of this form the third line on the graphs in Figs. 2 and 3. The effect of this is ambiguous at high wavenumber but yields a vastly superior fit at low wavenumber. Moreover, the second derivative has the correct sign. Thus we conclude that shorter time steps are the key to resolving this problem and producing results more like those in the TEM case.

C. Tests on inhomogeneous Hasegawa–Wakatani turbulence

The third test demonstrates one of the more interesting capabilities of the basis operator method. Because basis operators are calculated locally (within a patch of limited spatial size), it is therefore possible for the linear coefficients from the solution to have different values at different spatial locations. This, in turn, permits the method to diagnose spatial variation in the underlying physics of a turbulent system, such as variations in mean temperature and density gradients (or more specifically their effect on turbulence) and flow shear, including zonal flows.

This type of capability is implemented by changing the scheme used to label coefficients. Low-order linear basis operators (ones containing less than third order total derivatives) are assigned different coefficients for each value of radius. Linear operators of higher order, as well as nonlinear operators, are averaged over the entire data set as usual in order to aid the algorithm in rejecting noise.

The test of this capability involves applying this method to data from a Hasegawa–Wakatani simulation involving externally applied flow shear. The flow shear in question has the specific functional form,

$$V_y^0 = - U^0[1 + \alpha \cos(2\pi f^0 t)]\sin(2\pi x/L_x), \tag{29}$$

where $U^0 = 1.202\,64$, $f^0 = 0.1$, and $\alpha = 0.1$. The flow described by the above equations is of course an $m = 1$ poloidal shear. The time varying components are results of conditions in the original simulation which are not specifically adapted to this type of data analysis; this complication must be considered in interpreting the results.

As with the previous Hasegawa–Wakatani test, we must use a preoperator. The preoperator used here is of the same type as that described above. Unlike the previous cases, the results are not converted into spectral form for interpretation. This is because our expected result is primarily spatial in nature, so a spectral representation does not add clarity. Moreover, the external poloidal flow is represented by a single specific operator coefficient, specifically, the ∂_y operator.

The results of this comparison are shown in Fig. 4. The smooth line is the average poloidal flow, as given by Eq.
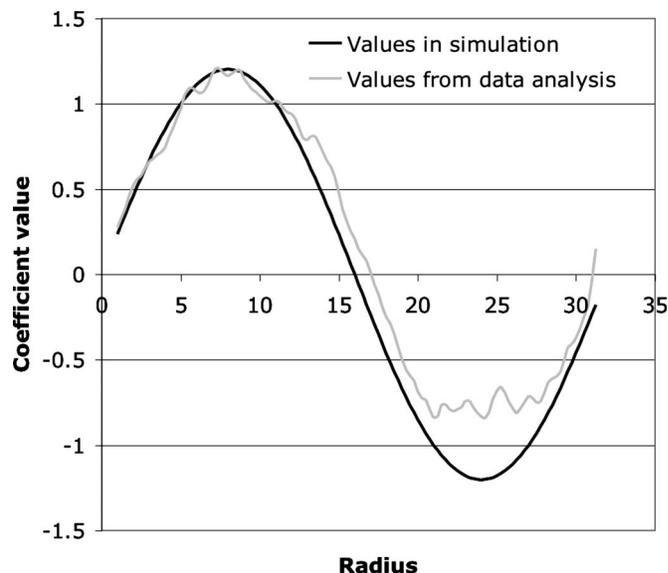

FIG. 4. *n-n dy* coefficient from a Hasegawa–Wakatani simulation with flow shear as a function of radius.

(29). The other line is the output from the basis operator method. Three rows on either end of the simulation region do not have solutions due to the finite size of the differentiating patch.

As we can see from this graph, the basis operator method was successfully able to infer qualitative features of the external flow. This is an impressive result considering that the algorithm was averaging over several periods of flow modulation, so no exact fit was possible.

V. SUMMARY AND DISCUSSION

We have derived a new procedure for estimating the parameters governing turbulence from experimental data. This is done by first representing our model equation as a superposition of differential operators. Then we apply a least-squares minimization to determine the coefficients of that superposition. This procedure can be applied to systems with multiple interacting fields without loss of generality.

We have also tested this procedure against simulation data. These tests demonstrate the capabilities of this algorithm under a variety of conditions. They also demonstrate that the fits this algorithm produces are not precise, particularly in comparison with the previous basis function method. However, this is to be expected since the reduced linear basis set in this algorithm does not permit a solution which exactly matches the parameters in the simulation. Given an increased linear basis set, it is possible that this problem may go away. In addition, there may be problems that limit quality of fit, such as time step size in the original data. Thus, the algorithm may be unsuited to analyze the specific data sets analyzed, but other data sets from similar simulations might produce significantly better results.

The basis operator method inherits many of the problems of the basis function method. The problem of artificial constraints on the solution due to the choice of basis functions is potentially more serious than in the basis function method

due to the fact that a low-quality solution will tend to appear superficially accurate; this is because all solutions from the basis operator method have a smooth functional form for their linear terms, whereas poor solutions from the basis function method appear noisy. However, this problem is largely solved by the addition of a quality of fit diagnostic. A poor fit will always have a high chi squared, and this indicates that additional basis operators need to be added to the model. Too many basis operators could theoretically result in fitting to noise. Currently the method tends to overconstrain rather than underconstrain the model. In addition to these problems, the basis operator creates an additional potential problem: high-order spatial derivatives are sensitive to noise. How much of a problem this will be remains unknown at the present.

Unlike the basis function method, the basis operator method is far more flexible in its application. An example of this is the fact that it was practical to apply it to a 128 $\times$ 128 data set, a task that would have been very computation intensive for a spectral technique. Also, the ability to analyze a system with flow shear is something that a spectral method would not be able to do accurately. In these cases, some loss of accuracy is a small price to pay for the ability to apply any sort of detailed quantitative analysis. By converting the concepts of the basis function method to a spatial rather than spectral representation, we have created an algorithm that can be used to produce quantitative insights in real turbulent systems rather than their idealized theoretical analogs. As such, we have brought this class of algorithms closer to practical and widespread application. Analysis of experimental data by this approach is currently underway and should yield results in the near future.

[1] Y. C. Kim and E. J. Powers, IEEE Trans. Plasma Sci. **7**, 120 (1979).
[2] L. Rezeau, G. Belmont, B. Gueret, and B. Lembege, J. Geophys. Res. **102**, 24387, DOI:10.1029/97JA01994 (1997).
[3] J. Blecki, R. Wronowski, M. Parrot, and S. P. Savin, Acta Geophysica **55**, 459 (2007).
[4] Y. Nagashima, K. Itoh, S.-I. Itoh, A. Fujisawa, K. Hoshino, Y. Takase, M. Yagi, A. Ejiri, K. Ida, K. Shinohara, K. Uehara, Y. Kusama, and the JFT-2M Group, Plasma Phys. Controlled Fusion **48**, S1 (2006).
[5] K. Itoh, Y. Nagashima, S.-I. Itoh, P. H. Diamond, A. Fujisawa, M. Yagi, and A. Fukuyama, Phys. Plasmas **12**, 102301 (2005).
[6] R. A. Moyer, G. R. Tynan, C. Holland, and M. J. Burin, Phys. Rev. Lett. **87**, 135001 (2001).
[7] Ch. P. Ritz and E. J. Powers, Physica D **20**, 320 (1986).
[8] Ch. P. Ritz, E. J. Powers, R. W. Miksad, and R. S. Solis, Phys. Fluids **31**, 3577 (1988).
[9] Ch. P. Ritz, E. J. Powers, and R. D. Bengtson, Phys. Fluids B **1**, 153 (1989).
[10] J. S. Kim, R. D. Durst, R. J. Fonck, E. Fernandez, A. Ware, and P. W. Terry, Phys. Plasmas **3**, 3998 (1996).
[11] J. S. Kim, R. J. Fonck, R. D. Durst, E. Fernandez, P. W. Terry, S. F. Paul, and M. C. Zarnstorff, Phys. Rev. Lett. **79**, 841 (1997).
[12] D. A. Baver, P. W. Terry, R. Gatto, and E. Fernandez, Phys. Plasmas **9**, 3318 (2002).
[13] P. W. Terry, R. Gatto, and D. A. Baver, Phys. Rev. Lett. **89**, 205001 (2002).
[14] P. W. Terry, E. Fernandez, and A. S. Ware, Astrophys. J. **504**, 821 (1998).
[15] D. A. Baver and P. W. Terry, Phys. Plasmas **12**, 042303 (2005).
[16] A. Hasegawa and M. Wakatani, Phys. Rev. Lett. **50**, 682 (1983).
[17] C. Holland, J. H. Yu, A. James, D. Nishijima, M. Shimada, N. Taheri, and G. R. Tynan, Phys. Rev. Lett. **96**, 195002 (2006).