

# Basis function multifield bispectral deconvolution analysis

D. A. Baver and P. W. Terry

*Department of Physics, University of Wisconsin, Madison, Wisconsin 53706*

(Received 8 September 2004; accepted 8 December 2004; published online 24 March 2005)

A different procedure for calculating linear and nonlinear coefficients of model systems for fully developed turbulence is derived. This procedure can be applied to systems with multiple interacting fields; in the single-field case the linear coefficients consist of mode frequencies and growth rates. This method differs from previous methods in the use of a limited set of functions or basis set from which the nonlinear terms in the turbulence equation are approximated in a series expansion. The algorithm is derived from this assumption using a least squares approach. This approach has been tested on simulations of fully developed two-dimensional turbulence and compared to previous methods. It is able to reconstruct coefficients with several significant figures precision and offers excellent noise rejection capabilities, and is moreover able to operate using tiny data sets compared to those required by previous methods. © 2005 American Institute of Physics.

[DOI: 10.1063/1.1854156]

## I. INTRODUCTION

Bispectral deconvolution analysis is a term applied to a class of algorithms which seek to determine the growth rates and nonlinear coupling coefficients of a turbulent system from the third-order statistical moments in its measured fluctuation spectrum.<sup>1-4</sup> In theory, this offers a variety of advantages over other methods of turbulence analysis. This is important because growth rates are convolved into spectra in a nontrivial way. Researchers can now obtain detailed information about the functional form of the driving instability,<sup>5</sup> which can then be directly compared to theory. Bispectral deconvolution analysis also isolates nonlinear effects and processes that enter the spectrum. These include nonlinear instability and stability, which manifest themselves as finite-amplitude-induced changes to the growth rate<sup>6</sup> and the nonlinear decorrelation rate. Measurement of these quantities offers access to key effects such as nonlinear damping of zonal flows<sup>7</sup> and the scaling of the turbulent decorrelation rate with global quantities such as the magnetic field.<sup>8</sup>

In practice, the application of bispectral analysis to experiment is limited by a number of factors. First, experiments usually fail to provide sufficient data channels for meaningful application of such algorithms. For instance, microscale fluctuations in tokamak experiments are essentially two-dimensional (2D), represented by a power spectrum in poloidal and radial wave number, while diagnostic techniques, particularly those capable of accessing the plasma core, have been limited to measurement of variation with respect to a single wave number. This adds significant uncertainties in inferred growth rates,<sup>5</sup> and has driven the development of data analysis techniques that will allow determination of both poloidal and radial wave number variations.<sup>9</sup> Second, bispectral deconvolution analysis has heretofore been restricted to single field models for instability-driven fluctuation dynamics.<sup>10</sup> Such models are sometimes referred to as  $i\delta$  models for the way in which instability is incorporated into the model. They suffer from the limitation that the growth rate is a fixed function of wavenumber that enters the model

as a linear coefficient. Hence these models cannot capture nonlinear instability or other anomalous adjustments to the instability rate or threshold, both of which occur in multifield models.<sup>6,11</sup> Third, and in practice most important, existing bispectral analysis algorithms display high levels of noise sensitivity, which must be compensated for by using large data sets, in particular long time series, to average out random noise. The requirement for long time series seriously limits the usefulness of such algorithms to plasma experiments because the short duration of each experimental run or shot means that the shot may be over before sufficient data can be gathered; moreover, even if sufficient data can be gathered it is likely that plasma parameters may have changed significantly during the course of data sampling, such that the bispectral algorithm will be attempting to infer a single growth rate from a time series during which several different growth rates occurred. All of these factors conspire to produce algorithm output which is problematic for interpreting an experiment.

In this paper we present a different class of bispectral analysis algorithms which solve the latter two problems. These algorithms feature vastly improved data utilization efficiency and reduced noise sensitivity, allowing application to short time series data. This in turn not only permits a typical experiment to produce adequate data for analysis, but permits analysis of transient properties of turbulence, for instance analysis of systems where growth rates are themselves dynamical variables. Moreover, the algorithm can be extended to multiple fields (i.e., density, potential, temperature, etc.) with minimal additional complexity. This potentially permits the algorithm to detect certain forms of nonlinear instability, and in general will provide more detailed information about the underlying turbulent dynamics than a single-field model. Because mode plasma motions involve multiple fields, these algorithms will ultimately provide a closer comparison with theory.

This class of algorithms achieves this great efficiency at the price of requiring additional assumptions about the be-

havior of the system it is to analyze. As such, great care must be taken to avoid “confirming” erroneous assumptions. This problem can be dealt with by allowing the algorithm degrees of freedom which it is not expected to exercise. A noninfinitesimal value returned in a corresponding output channel would indicate an error in the assumptions. On the other hand, the ability to utilize such corroborating information to improve the efficiency of the algorithm permits the development of an entire class of customized algorithms designed to capture the dynamics of specific types of turbulence or to correct for minor deficiencies in an experimental diagnostic. The algorithm we will be discussing is designed to operate under relatively simple conditions. As such, it is to be considered as a starting point for further development.

## II. BACKGROUND

In this section we discuss the prior art in the field of bispectral analysis. This consists of the Ritz method developed by Ritz *et al.*<sup>1-3</sup> and the modified Ritz method developed by Kim *et al.*<sup>4,5</sup>

For the purposes of discussing these methods, it is useful to first go over the notation employed in these papers. For reasons which will be discussed later we will be using different notation for the newer techniques, however, the structure of the resulting algorithms is nevertheless similar.

Both the Ritz and modified Ritz methods model turbulent fluctuations measured in a plasma with a generic turbulence equation whose coefficients are to be determined from the measurements:

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

where  $X_k = \phi(k, t)$  and  $Y_k = \phi(k, t + \tau)$ ,  $\phi$  being the measured fluctuating quantity as a function of wave number 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  $Q_k^{k_1, k_2}$  determines the nonlinear transfer rate. The goal of these algorithms is to calculate  $L$  and  $Q$  from fluctuation measurements.

### A. Ritz method

The first method for quantitatively estimating such information in a plasma was developed by Ritz *et al.* in the late 1980s.<sup>1-3</sup> A review of this method is presented in the Introduction of an article on the modified Ritz method by J. S. Kim *et al.*<sup>4</sup>

The Ritz method solves the growth rates and transfer functions by expanding the model equation in a series of moment equations, multiplying by  $X_k^*$  and  $X_{k_1}^* X_{k_2}^*$ , respectively. 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 Millionschikov approximation and is derived by assuming a nearly Gaussian distribution in the fluctuating quantities. This results in the following equations:

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

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

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

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

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 wave numbers. 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.*<sup>4</sup> and is described in more detail in that paper. The modified Ritz method begins by assuming that each of the measured spectra can be divided into an ideal and nonideal spectrum:

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

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

From this Kim *et al.* derive 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 Millionschikov 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  $k_1, k_2$ , a matrix notation is needed to represent the equations for the third- and fourth-order moments. This is solved with 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 wave number  $k$  ( $k=f(l)$ , where  $f$  is a linear function), and  $i, j$  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  $L_k$ :

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

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

Combining these gives a formula for  $\gamma$

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

This method produces more accurate fits than the Ritz method, but a long time series is still required to produce an accurate fit. Also, like the Ritz method, it is derived from a single-field equation and therefore can only fit spectra containing a single fluctuating quantity. These limitations will be addressed by the algorithms presented in this paper.

### III. MULTIFIELD BISPECTRAL ANALYSIS AND LEAST SQUARES DERIVATION

Methods of bispectral analysis have so far been motivated by rough statistical arguments based on the properties of moment equations derived from analytic turbulence theory. In order to apply such methods to dynamical systems with multiple fields, we will need a more rigorous derivation.

We begin by writing down a generic turbulence equation for multiple fields,

$$\hat{Y}_i^k \equiv \tilde{Y}_i^k + Y_i^k = D_{ij}^k X_j^k + \sum_{k'} Q_{ilm}^{k,k'} X_l^{k'} X_m^{k-k'}, \quad (9)$$

where  $Y_i^k = [\psi_i^k(t+\Delta t) - \psi_i^k(t)]/\Delta t$  and  $X_i^k = [\psi_i^k(t+\Delta t) + \psi_i^k(t)]/2$ .  $\psi_i^k(t) \equiv \phi_k(t)$ ,  $n_k(t)$ ,  $P_k(t)$ , etc., with each value of  $i$  corresponding to a particular field.  $\hat{Y}_{ik}$  is the value of  $Y_{ik}$  predicted for a given set of coefficients, and  $\tilde{Y}_{ik}$  is the error in this value.

The parameter  $\tilde{Y}_{ik}$  will then take the place of  $\chi$  such that we will then seek to minimize  $\sum_k |\chi_k|^2$ . Note that our coefficients  $L$  and  $Q$  from the Ritz and modified Ritz methods have been replaced by tensors  $D$  and  $Q$ , reflecting the fact that our model equation permits multiple fields. The growth rate spectrum can be derived by solving the eigenmodes of the matrix  $D$ . Some other changes in notation involve the construction of the coefficients  $X$  and  $Y$ .  $Y$  now represents a derivative rather than a value at a future time step; this is primarily to control the number of terms in our equations.  $X$  is an average of values at different time steps rather than a value at a particular time step. This is done in order to reduce the effect of the time step size  $\Delta t$  on the accuracy with which the derivative is calculated by approximating the value of the

fluctuating quantity at an intermediate time. If  $X$  were calculated at either  $t$  or  $t+\Delta t$  there would be an error in  $Y$  due to the second derivative in the fluctuating quantity, which would scale with  $\Delta t^{-1}$  since there is already a factor of  $\Delta t$  in the denominator. With the centered calculation of  $X$  there is an error due to the second derivative in  $X$ , which is proportional to  $\Delta t^{-2}$ , however, the second derivative error in  $Y$  cancels leaving a third derivative error, which is also proportional to  $\Delta t^{-2}$ . This converts the algorithm from first order to second order accuracy with respect to our time step. This allows a significant improvement in the accuracy of our fits without a significant increase in algorithm complexity.

From this we can use the model equation to write down the error function,

$$\chi^2 = \sum_{i,k,t} \tilde{Y}_i^k \tilde{Y}_i^{-k} = \sum_{i,k,t} |D_{ij}^k X_j^k - Y_i^k + \sum_{k'} Q_{ilm}^{k,k'} X_l^{k'} X_m^{k-k'}|^2. \quad (10)$$

This quantity can be minimized by applying the variational principle, introducing the quantities  $\delta D$  and  $\delta Q$  which are infinitesimal variations in  $D$  and  $Q$ , respectively. For an extremal value the above formula goes from zero to lowest order in the variational quantities, yielding

$$\forall k, i, j: \sum_t \delta D_{ij}^{*k} X_j^{*k} (D_{in}^k X_n^k - Y_i^k + \sum_{k'} Q_{ilm}^{k,k'} X_l^{k'} X_m^{k-k'}) = 0 \quad (11)$$

$$\forall k', k, i, l, m: \sum_t \delta Q_{ilm}^{*k,k'} X_l^{*k'} X_m^{*k-k'} (D_{ij}^k X_j^k - Y_i^k + \sum_{k''} Q_{inp}^{k,k''} X_n^{k''} X_p^{k-k''}) = 0. \quad (12)$$

This allows us to write down a set of statistical moment equations,

$$\forall k, i, n: D_{ij}^k \langle X_j^k X_n^{*k} \rangle - \langle Y_i^k X_n^{*k} \rangle + \sum_{k'} Q_{ilm}^{k,k'} \langle X_n^{*k} X_l^{k'} X_m^{k-k'} \rangle = 0, \quad (13)$$

$$\forall k, k', i, l, m: D_{ij}^k \langle X_j^k X_l^{*k'} X_m^{*k-k'} \rangle - \langle Y_i^k X_l^{*k'} X_m^{*k-k'} \rangle + \sum_{k''} Q_{inp}^{k,k''} \langle X_n^{*k''} X_p^{k-k''} X_l^{*k'} X_m^{*k-k'} \rangle = 0. \quad (14)$$

These equations can be written in more compact form by applying the following definitions:

$$F_{nplm}^{k,k'} \equiv \langle X_n^{*k'} X_p^{k-k'} X_l^{*k'} X_m^{*k-k'} \rangle, \quad (15)$$

$$A_{jlm}^{k,k'} \equiv \langle X_j^k X_l^{*k'} X_m^{*k-k'} \rangle, \quad (16)$$

$$B_{jlm}^{k,k'} \equiv \langle Y_j^k X_l^{*k'} X_m^{*k-k'} \rangle, \quad (17)$$

$$a_{ij}^k \equiv \langle X_i^k X_j^{*k} \rangle, \quad (18)$$

$$b_{ij}^k \equiv \langle Y_i^k X_j^{*k} \rangle. \quad (19)$$

Our moment equations now take on the form

$$\forall k, i, n: D_{ij}^k a_{jn}^k - b_{in}^k + \sum_{k'} Q_{ilm}^{k,k'} A_{nlm}^{*k,k'} = 0, \quad (20)$$

$$\forall k, k', i, l, m: D_{ij}^k A_{jlm}^{k,k'} - B_{ilm}^{k,k'} + \sum_{k''} F_{nplm}^{k,k',k''} Q_{inp}^{k,k''} = 0. \quad (21)$$

Solving Eq. (21) allows us to solve for  $Q$ ,

$$Q_{inp}^{k,k'} = (F_{nplm}^{k,k',k''})^{-1} (B_{ilm}^{k,k'} - D_{ij}^k A_{jlm}^{k,k'}). \quad (22)$$

Substituting this into Eq. (20) yields a formula for  $D$ :

$$D_{lm}^k = [I_{ij,lm}^{k,k'} a_{jn}^{k'} - A_{vin}^{*k'} (F_{\nu\mu})^{-1} A_{\mu lm}^k]^{-1} [b_{in}^{k'} - A_{vin}^{*k'} (F_{\nu\mu})^{-1} B_{\mu}]. \quad (23)$$

#### IV. BASIS FUNCTION ANALYSIS

As we note from the preceding section, development of bispectral analysis to date has tended towards the ideal of a least squares method. At this point we might ask, what further improvements can be made? The notion of a weighted least squares comes to mind. However, because of the way our error function was defined in the preceding section, each mode is minimized individually so a weighting function will have no effect on the final solution.

A closer examination of the least squares derivation of bispectral analysis shows that it suffers from one of the perennial problems of least squares optimization: unnecessary degrees of freedom. When a least squares optimization is presented with degrees of freedom which have no physical meaning, it will attempt to improve the quality of fit by exploiting these degrees of freedom. This results in the problem of fitting to noise: errors in the original data are carried over to the solution without being averaged out. Another problem is an increase in the size of the data set required to get any solution at all; if the number of data points is less than the number of degrees of freedom, the solution is degenerate and many possible solutions provide an equally satisfactory fit from an algorithm point of view, albeit most of them are unsatisfactory interpretations of the data. An algorithm such as that derived in the preceding section will run into problems with noninvertible matrices or with near noninvertible matrices which avoid zero eigenvalues due to noise but produce anomalously large values for some entries.

Conventional bispectral analysis has many unnecessary degrees of freedom because it permits each and every interacting triplet to have, in principle, a different and independent strength of nonlinear coupling. In systems that are highly inhomogenous, or which involve a limited number of modes, this may be appropriate. In most turbulent systems, however, we can safely assume that the nonlinearities will have some recognizable functional form in which the nonlinear interaction strengths for all of the interacting triples can be written down as some function of  $k, k'$ , and  $k-k'$ . All of the nonlinearities that have ever been proposed in turbulence theory combined represent only a tiny fraction of the available function space permitted by conventional bispectral analysis.

In basis function bispectral analysis we remove most of the degrees of freedom in the nonlinearity by representing its

coefficients as a linear superposition of an incomplete set of basis functions. In the limit as our basis set approaches completeness this approach converges to conventional bispectral analysis, albeit by a formula which is cumbersome to execute as an algorithm. In practice, it is the incompleteness of our basis set which motivates the algorithm: by choosing which functions are or are not allowed to participate in the basis set, we have a method of applying *a priori* assumptions about the behavior of our system so as to make our algorithm more efficient in finding a solution which does not violate those assumptions.

Deciding what basis functions to include or not to include is not an entirely trivial matter, as different systems under study will have different inherent behaviors, hence will require different and possibly customized algorithms to analyze with optimal efficiency. A larger basis set avoids the problem of erroneous assumptions about system behavior, whereas a smaller basis set results in higher data efficiency and algorithm speed. One compromise is to include a limited number of basis functions which blatantly violate our *a priori* assumptions; these functions act as a bellwether, indicating when a system is behaving in an unexpected manner. If one of these functions returns a coefficient which is significantly above noise levels, we can respond by reoptimizing the algorithm to accommodate this information. For instance, if we expect an isotropic nonlinearity, maintaining several anisotropic functions allows us to determine if anisotropic terms exist, and if they do a larger set of anisotropic functions can be used to determine their nature.

#### A. Derivation

We begin by applying the definitions from Sec. III, but to a different model equation:

$$\hat{Y}_i^k \equiv \tilde{Y}_i^k + Y_i^k = D_{ij}^k X_j^k + \sum_{k'} \alpha_{\mu} \beta_{\mu ilm}^{k,k'} X_l^{k'} X_m^{k-k'}. \quad (24)$$

The values  $\beta_{\mu ilm}^{k,k'}$  are a predefined basis of functions which we use to represent our nonlinearities by projecting the nonlinearity onto the basis set, yielding coefficients  $\alpha_{\mu}$ . Representing our nonlinearities as linear superpositions of predefined functions is the fundamental difference between this and previous forms of bispectral analysis. By choosing a basis set which captures the behavior of nonlinearities which are likely to arise while using far less functions than the number of possible interacting triplets, we dramatically reduce the number of degrees of freedom in our system. This means our algorithm is much less likely to fit to noise, and since calculation of the nonlinear terms uses data from the entire spectrum of our input data, the calculation of the coefficients  $\alpha_{\mu}$  is likely to be much less sensitive to noise since even a single pair of time steps contains a substantial amount of data with which to average out the effects of random noise.

Given this model equation, we can now proceed using the variational principle as we did in Sec. III. This derivation follows a slightly different course, however, because our free parameters in the nonlinearity have fewer indices:

$$\forall k, i, j: \sum_t \delta D_{ij}^{*k} X_j^{*k} (D_{in}^k X_n^k - Y_i^k + \sum_{k'} \alpha_{\nu} \beta_{vilnm}^{k,k'} X_l^{k'} X_m^{k-k'}) = 0, \quad (25)$$

$$\forall \mu: \sum_{t,k,i} \delta \alpha_{\mu}^* \beta_{\mu ilm}^{*k,k'} X_l^{*k'} X_m^{*k-k'} (D_{ij}^k X_j^k - Y_i^k + \sum_{k'} \alpha_{\nu} \beta_{vinp}^{k,k'} X_n^{k'} X_p^{k-k'}) = 0. \quad (26)$$

We can then write down the resulting moment equations in compact notation by applying the following definitions:

$$F_{\mu\nu} \equiv \sum_{k,k',k'',i} \beta_{\mu ilm}^{*k,k'} \beta_{vinp}^{k,k''} \langle X_n^{k''} X_p^{k-k''} X_l^{*k'} X_m^{*k-k'} \rangle, \quad (27)$$

$$A_{\mu ij}^k \equiv \sum_{k'} \beta_{\mu ilm}^{*k,k'} \langle X_j^k X_l^{*k'} X_m^{*k-k'} \rangle, \quad (28)$$

$$B_{\mu ij}^k \equiv \sum_{k'} \beta_{\mu ilm}^{*k,k'} \langle Y_j^k X_l^{*k'} X_m^{*k-k'} \rangle, \quad (29)$$

$$B_{\mu} = \sum_k B_{\mu ii}^k, \quad (30)$$

$$a_{ij}^k \equiv \langle X_i^k X_j^{*k} \rangle, \quad (31)$$

$$b_{ij}^k \equiv \langle Y_i^k X_j^{*k} \rangle. \quad (32)$$

Our moment equations now take on the form

$$\forall k, i, n: D_{ij}^k a_{jn}^k - b_{in}^k + \sum_{\mu} \alpha_{\mu} A_{\mu in}^{*k} = 0, \quad (33)$$

$$\forall \mu: D_{ij}^k A_{\mu ij}^k - B_{\mu} + \sum_{\nu} F_{\mu\nu} \alpha_{\nu} = 0. \quad (34)$$

Solving Eq. (34) allows us to solve for  $\alpha$ :

$$\alpha_{\nu} = F_{\nu\mu}^{-1} (B_{\mu} - D_{ij}^k A_{\mu ij}^k). \quad (35)$$

Substituting this into Eq. (33) yields a formula for  $D$ ,

$$D_{lm}^k = (a_{ij,lm}^{k',k} - A_{vin}^{*k'} F_{\nu\mu}^{-1} A_{\mu lm}^k)^{-1} (b_{in}^{k'} - A_{vin}^{*k'} F_{\nu\mu}^{-1} B_{\mu}). \quad (36)$$

The inverted terms in parentheses nominally form a sixth rank tensor. This can be inverted, however, by ordering the indices to form an ordinary second rank tensor:

$$M_{IJ} \equiv M_{in,lm}^{k',k}, \quad \text{where } I = f(i, n, k'), \quad J = f(l, m, k).$$

With this taken under consideration, Eq. (36) represents the formula which our algorithm will be solving to calculate  $D$  and Eq. (35) represents the formula which our algorithm will be solving to calculate  $\alpha$  once  $D$  has been found.

## B. Basis function selection

The algorithm described so far differs from the algorithm described in the preceding section by the choice of an incomplete set of basis functions for the nonlinear terms. In the case where the basis set is complete, the two algorithms are equivalent. Since this algorithm relies on removing degrees of freedom from the solution, the solution will there-

fore depend in a nontrivial way on which degrees of freedom are removed. As such, the choice of the basis set plays an important role in the quality of the results.

Fortunately, the functional form of the nonlinear coupling between modes is relatively well understood and restricted to a relatively limited set of possibilities compared to the functional form of the linear terms, the latter being dependent on the details of the instability mechanism involved. As a result, we can use insights from analytic plasma turbulence theories to develop a basis set that accurately approximates the nonlinearities found in the system being analyzed.

In the results that follow, three types of basis functions are employed representing three nonlinear forms: simple nonlinearities, series nonlinearities, and control nonlinearities. The simple nonlinearities are nonlinear terms found in our equations whose functional form can be reduced to a single function which does not depend on the linear coefficients of the equations, multiplied by a single nonlinear coefficient. These nonlinearities can be introduced directly as single functions. Series nonlinearities are terms whose functional form does depend on equation's linear coefficients, hence cannot be represented in as a single function without prior knowledge of the coefficients being solved for. Since no single function can represent such a nonlinearity for all possible combinations of linear coefficients, they are instead approximated by performing a series expansion in some set of functions which converge to the desired function. Control nonlinearities are nonlinear terms which do not occur in the equations, nor can be represented as linear combinations of terms in the equations. These are included to verify that the basis set is adequate for the problem it is applied to. If this condition is met the control nonlinearities will return small values whose magnitude is determined by the amount of error in our data set, that is to say, they will return zero within the limits of the algorithm's accuracy. If they return larger values, this indicates that more functions must be added for an accurate fit to emerge.

It is important to note that the control nonlinearities are intended to serve a diagnostic purpose rather than improving the overall quality of fit; their purpose is solely to test the appropriateness of a particular basis function to a particular data set. As a result, it is not necessary to include every possible nonlinearity in the control set. This is because the difference between the actual nonlinearity and the best fit from the noncontrol nonlinearities (simple and series) is expected to have a random orientation with respect to the control set. The probability that a random vector will be purely perpendicular to a substantial set of linearly independent vectors to within acceptable tolerances is extremely small, hence if the components of the nonlinear interaction vector parallel to all of the control nonlinearities are close to zero to within noise levels, we can therefore be reasonably certain that our choice of basis is adequate.

## V. RESULTS

In order to test the algorithm derived in the preceding section, we apply it to simulation data of fully developed turbulence. By comparing the coefficients derived from the

algorithm to the coefficients used in the simulation, we are able to assess the accuracy with which our method can reconstruct linear growth rates. We also add random noise to some data sets in order to assess the ability of the algorithm to reconstruct parameters from nonideal data.

### A. Test of the procedure with 2D two-field turbulence

Our first numerical experiment employs a trapped electron mode (TEM) model to generate two-field data. The model employed is a pseudospectral code, and fields in this model represent potential and electron density. As it is a two-field model, our linear coefficient matrix has four complex components which can be represented by eight real components. The equations that are used to generate the data are as follows:

$$\begin{aligned} & \left( \frac{\partial \phi_k}{\partial t} + \gamma_k \phi_k \right) (1 - \sqrt{\epsilon} + k^2 \rho_s^2) - \sqrt{\epsilon} \nu_{eff} (\tilde{n}_k - \phi_k) \\ & + ik_y \nu_D \phi_k [1 - \sqrt{\epsilon} (1 + \alpha \eta_e)] - \rho_s^2 c_s \sum_{k'} (k' \\ & \times k \cdot \hat{z}) k'^2 \phi_{k'} \phi_{k-k'} = 0, \end{aligned} \quad (37)$$

$$\begin{aligned} & \frac{\partial \tilde{n}_k}{\partial t} + \gamma_k \tilde{n}_k + \nu_{eff} (\tilde{n}_k - \phi_k) + ik_y \nu_D (1 + \alpha \eta_e) \phi_k \\ & + c_s \rho_s \sum_{k'} (k' \times k \cdot \hat{z}) \tilde{n}_{k'} \phi_{k-k'} = 0. \end{aligned} \quad (38)$$

In these  $\sqrt{\epsilon}$  is the trapped particle fraction,  $\nu_{eff}$  is the electron detrapping rate,  $\rho_s$  is ion gyroradius,  $\alpha$  is a numerical factor equal to  $\frac{3}{2}$ ,  $\eta_e$  is the ratio of temperature to density scale lengths,  $\nu_D$  is the diamagnetic drift frequency, and  $c_s$  is the sound speed. The factor  $\gamma_k$  is an artificial damping term introduced to enable saturation on a grid scale too small to incorporate physically realistic damping mechanisms. In this simulation,  $\gamma_k$  is hyperviscous at high  $k$  and forms a Gaussian at low  $k$ .

The version of the algorithm applied to these data employed the following basis set:

$$\beta_{1,\phi\phi n}^{k,k'} = (k \times k' \cdot \hat{z}),$$

$$\beta_{2,n\phi n}^{k,k'} = (k \times k' \cdot \hat{z}),$$

$$\beta_{3,\phi\phi\phi}^{k,k'} = (k \times k' \cdot \hat{z}) k'^2,$$

$$\beta_{4,\phi\phi\phi}^{k,k'} = (k \times k' \cdot \hat{z}) k'^2 k^2,$$

$$\beta_{5,n\phi\phi}^{k,k'} = (k \times k' \cdot \hat{z}) k'^2,$$

$$\beta_{6,\phi n n}^{k,k'} = (k \times k' \cdot \hat{z}) k'^2,$$

$$\beta_{7,n n n}^{k,k'} = (k \times k' \cdot \hat{z}) k'^2,$$

$$\beta_{8,\phi\phi\phi}^{k,k'} = (k \times k' \cdot \hat{z}) k'^2 k_x,$$

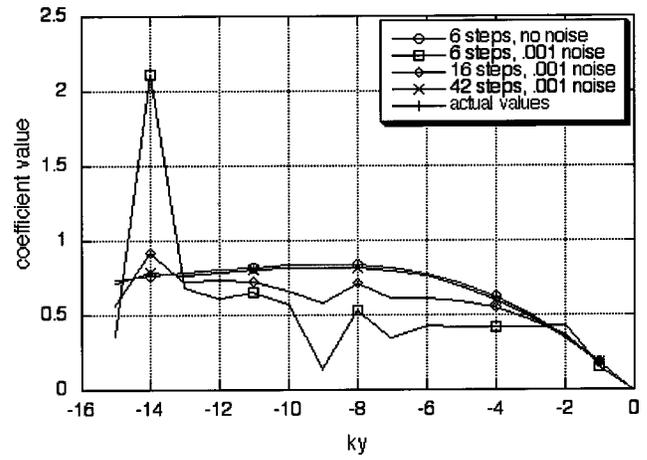


FIG. 1. Bispectral output vs simulation input. Simulation data are from a TEM simulation. Displayed data are for the imaginary part of the  $\phi$ - $\phi$  component of the coefficient matrix for modes with  $k_x=0$ .

$$\beta_{9,\phi\phi\phi}^{k,k'} = (k \times k' \cdot \hat{z}) k'^2 k_y,$$

$$\beta_{10,\phi\phi\phi}^{k,k'} = (k \times k' \cdot \hat{z}) k_x,$$

$$\beta_{11,n\phi n}^{k,k'} = (k \times k' \cdot \hat{z}) k_y,$$

$$\beta_{11+l,\phi\phi\phi}^{k,k'} = (k \times k' \cdot \hat{z}) k'^2 \sin\left(\frac{l\phi k^2}{k_{max}^2}\right), \quad l = 1, 15. \quad (39)$$

In this basis set, function 2 is a simple nonlinearity chosen to represent the  $E \times B$  nonlinearity in our model. Functions 3, 4, and 12–26 form a series nonlinearity which represents the polarization drift nonlinearity in our model. A series representation is necessary because of the functional dependence of this nonlinearity on  $(1 - \sqrt{\epsilon} + k^2 \rho_s^2)$ , which cannot be represented as a single function without prior knowledge of the value of  $\sqrt{\epsilon}$ . The remaining terms (1,5–11) are control functions.

Figure 1 shows the results of one of those components over a select set of wave numbers. The wave numbers are at  $k_x=0$  for different values of  $k_y$ , and the coefficient shown is the imaginary part of the  $\phi$ - $\phi$  component (potential self-reaction) of the coefficient matrix. This particular component is dominated by the diamagnetic drift frequency with dispersive terms due to finite Larmor radius effects.

Marked with circles and bars and situated at the top of Fig. 1 for  $k_y$  values of  $-4$  through  $-12$  are the lines for the actual values calculated from the input parameters of the simulation and the algorithm output for three realizations (pairs of time steps) of data with no added noise. These lines coincide with each other, indicating excellent agreement between algorithm output and simulation input. In this case, the algorithm was able to reconstruct the original coefficients to several significant figures' precision. That it was able to do so with only three realizations is significant when compared to conventional bispectral algorithms, which require thousands of realizations.

The other three lines are fits involving varying numbers of realizations, ranging from 3 to 21, in which the data

sample has been contaminated with white noise. The noise in this case is weighted in proportion to the average fluctuation level at each wave number, and for these cases its average magnitude is 0.1% of the fluctuation level.<sup>12</sup> As we can see, this causes a significant reduction in the accuracy of the fit for three realizations, however, with a modest increase in the number of realizations the quality of fit rapidly improves.

## B. Test of the procedure with 2D one-field turbulence

In our next numerical experiment, we seek to compare the accuracy of the basis function form of bispectral analysis to previous forms of bispectral analysis, in this case the Ritz method. Since the Ritz method applies only to single-field data, we employ a single-field version of the basis function method. The source of our simulation data is a somewhat more generic single-field spectral code; since we are only interested in the accuracy with which our algorithm can reconstruct the simulation coefficients, our growth and frequency spectra are chosen more or less arbitrarily within the constraint of achieving a stable saturated state rather than attempting to model any particular physical system. The simulation code does employ two nonlinearities (density and vorticity advection).

The equations used to generate the data are loosely based on the Terry–Horton equations<sup>10</sup> and are as follows:

$$\begin{aligned} \frac{\partial n}{\partial t} = & ik_y v_D n - \gamma_k n + \sqrt{\epsilon} d_k k_y^2 n + L_1 \frac{1}{2} \sum_{k'} (k \times k' \cdot \hat{z})(k_y \\ & - k'_y) n_{k'} n_{k-k'} + L_2 \sum_{k'} (k \times k' \cdot \hat{z}) k'^2 n_{k'} n_{k-k'}. \end{aligned} \quad (40)$$

The coefficients  $L_1$  and  $L_2$  are strengths assigned to the nonlinearities, which represent  $E \times B$  and polarization drift nonlinearities, respectively. The coefficient  $d_k$  controls in what parts of the spectrum this term is active. It is set to a constant value in a rectangular region of  $k$  space and is zero elsewhere. The term  $\gamma_k$  is artificial damping, as with the TEM model, however, since we are interested here in providing data to test a diagnostic algorithm we are simply interested in generating a stable saturated state, not in modeling any particular physics. As a result, in certain parts of the spectrum this term is given a negative value and this in turn constitutes the main source of free energy input in this simulation. It is hyperviscous at high  $k$  but is given a constant negative (driving) value at intermediate  $k$ .

Figure 2 shows the results from this comparison. Both algorithms are applied to the same data set, and the output growth rates for  $k_x=0$  and different values of  $k_y$  are compared.

The version of the algorithm applied to these data employed the following basis set:

$$\begin{aligned} \beta_1^{k,k'} &= (k \times k' \cdot \hat{z}) k'^2, \\ \beta_2^{k,k'} &= (k \times k' \cdot \hat{z}) k_y. \end{aligned} \quad (41)$$

This basis consists entirely of simple nonlinearities found in our model. This was chosen out of expedience since

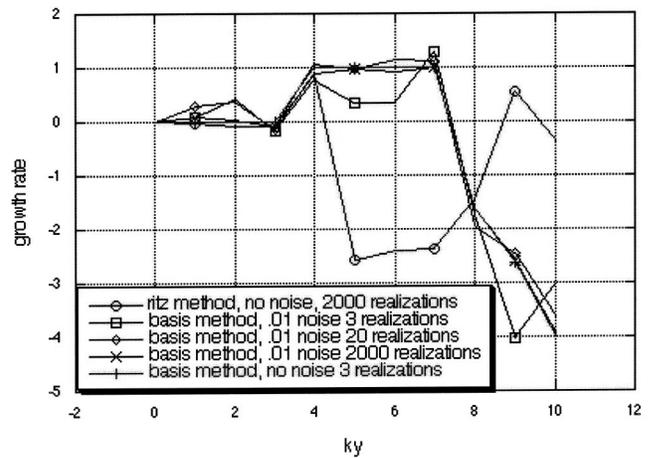


FIG. 2. Comparison of basis function and Ritz bispectral methods on one-field data. Displayed are growth rates for modes with  $k_x=0$ .

in this test we are interested in a noise sensitivity comparison, not a proof of the validity of the algorithm. This does result in a small basis set, which may artificially reduce noise sensitivity somewhat compared to the algorithm with a more realistic basis set, however, for the purposes of a general comparison we can accept this limitation.

The basis function method was tested for three realizations without noise and for between 3 and 2000 realizations with 1% noise. Compared to this are the results from the Ritz method for 2000 realizations and no noise. The basis function method produces accurate fits to several significant figures in the absence of noise. Adding noise causes some deviations from the correct output, but the overall quality of fit is still good even for three realizations.

This particular data set proved intractable to the Ritz method, possibly on account of violating the Millionshchikov approximation. This can be seen by the poor quality of fit for  $k_y > 4$ .

## C. Test of the procedure with time-varying coefficients

As we saw in the preceding sections, the basis function method is capable of reconstructing linear coefficients from fully developed turbulence data using extremely short time series data. This opens up an interesting possibility that did not exist previously in the history of bispectral analysis: the analysis of turbulent systems with time-varying coefficients.

Turbulence with time-varying coefficients can occur whenever the physics determining the coefficients of the turbulence equations are changing in time. This can be due to varying pressure or temperature gradients, such as might occur during  $L$ - to  $H$ -mode transitions and vice versa, sporadic transport events such as sawtooth crashes or ELM (edge-localized modes) or general adjustment of such gradients due to the turbulence itself. It can also be due to nonlinear instability in which the growth rate changes as the turbulence reaches finite amplitude.

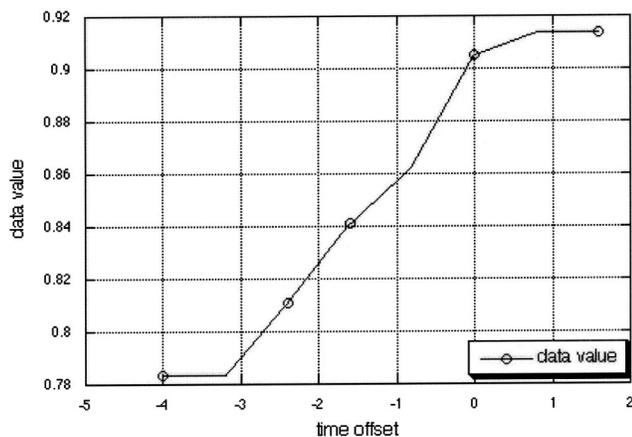


FIG. 3. Bispectral output of imaginary part of  $\phi$ - $\phi$  component of the linear coefficient matrix at  $k_x=0$ ,  $k_y=1.3$  over different time windows. Time offset is measured from an abrupt transition in  $v_D$ . Each time window contains 32 realizations beginning with the offset time.

Introducing time-varying coefficients presents some potential problems for a bispectral algorithm. Because there is no exact fit of constant coefficients to a data set generated by time-varying coefficients, the fit which the algorithm actually produces will have a large residual  $\chi$  squared. This is similar to the problem presented by attempting to solve a system with extreme levels of added noise. As such, the actual result from such a fit will depend in a nontrivial way on the structure of the algorithm, and thus the quality of its performance cannot be easily extrapolated from the much simpler case of static coefficients.

To test the capacity of the algorithm to handle time-varying coefficients, we use our TEM model to construct a simple case of this. The simulation is run until it reaches saturation with one value of the diamagnetic drift velocity. A second simulation using a different drift velocity is run using the final values from the first simulation as starting values, and the data from the two simulations is spliced together to form a single data set. This data set now contains an abrupt transition in the diamagnetic velocity, giving certain coefficients a step-function dependence in time.

Figure 3 shows the results for a particular value of  $k_x$  and  $k_y$  when the basis function method is applied to this data set. The basis function method is applied to a series of 32-realization time windows, each starting at a progressively later point in the data set. The x-axis shows the starting time of the window, with the transition in the coefficient value occurring at  $t=0$ . Each window therefore contains different numbers of time steps before and after the transition.

As we can see from the figure, the algorithm returns the actual values from the simulation, as we would expect from our experience in the static case. As the time window crosses the transition, the returned coefficient varies smoothly from the initial value to the final value. This is remarkably close to the ideal behavior of a weighted average between the two cases. This is a promising result, since the ability to return an average value over a time window is good enough to analyze a system with time-varying coefficients, provided the width

of the time window is smaller than the time scale over which the coefficients vary. In the event that the time window is longer than the time scale of variation, the algorithm simply filters out the high frequency components of the signal. Moreover, this is much better performance than we could have expected from a system with random noise added to achieve an equivalent residual  $\chi$  squared.

## VI. SUMMARY AND DISCUSSION

We have derived a different procedure for estimating linear growth rates and nonlinear transfer rates of turbulent systems from experimental data. This procedure is obtained by first representing the nonlinear coefficients as a superposition of a limited set of functions, thus reducing their degrees of freedom. An algorithm can then be derived from this model equation using a least squares minimization. This procedure permits analysis of turbulence with multiple interacting fields without loss of generality.

This method has been tested against simulation data to measure its ability to reconstruct model coefficients. Through these tests it is also compared to previous methods for estimating growth rates and nonlinear transfer rates from data. In these tests this method produces vastly improved performance. It is capable of producing accurate fits using nearly three orders of magnitude less data, even in the presence of modest levels of noise. While a quantitative reduction in noise sensitivity has not yet been verified, the capacity to fit using such small data sets in the presence of noise strongly suggests this is the case. Its quality of fit in the absence of noise is also superior, recovering simulation coefficients to several significant figures accuracy.

These improvements in the capabilities of bispectral analysis permit its application to more complicated turbulent dynamics than would have previously been possible. Because it can be applied to multiple fields without loss of generality, it is not necessary to assume that a single field is sufficient to characterize observed turbulence. The capacity to apply this approach to multiple fields is in practice limited by the availability of correlated multifield data. In the event that such data becomes available this approach provides a tool for extracting different types of information from it, and moreover the availability of such a data analysis technique provides an incentive for the development of such diagnostics.

This method is also capable of analyzing transient behavior. This permits the study of turbulent processes in transient phenomena, such as ELMs, sawtooth crashes, transitions between different confinement regimes, and so forth. It also permits the study of nonlinear instability. Because it can be applied to multiple fields, this permits the detection of certain types of nonlinear instability. Regardless of the mechanism of the instability such a phenomenon can be detected by observing the variation of growth rates with turbulence fluctuation level.

As of this time, this method has not been applied to actual experimental data. The quality of reconstructed

growth rates from experimental data therefore remains untested. This is mainly due to the limitations of the underlying model equation it assumes to accurately represent physical turbulence, particularly after it has been filtered by an applied measurement apparatus. Some potential problems include inhomogeneities due to rational surfaces and magnetic shear, essential turbulent dynamics at a scale below the resolution of the measurement device, and differences between the experimental and model coordinate system. The latter can be addressed by preprocessing data to fit a common coordinate system or by careful choice of basis functions to account for the probable form of nonlinear coupling in the experimental coordinates; however, the latter two require more serious adjustments which are beyond the scope of this method.

In systems where the above problems are not so serious as to detract significantly from the quality of fit, this method offers great promise as a means of data analysis for experimentally measured turbulence. Moreover, the approach employed to derive this method is fairly generic and may be applied to more complicated systems by suitable adjustments to the form of the underlying model equations.

## ACKNOWLEDGMENT

This work was supported by the U.S. Department of Energy under Grant No. DE-FG02-89ER-53291.

- <sup>1</sup>Ch. P. Ritz and E. J. Powers, *Physica D* **20**, 320 (1986).
- <sup>2</sup>Ch. P. Ritz, E. J. Powers, R. W. Miksad, and R. S. Solis, *Phys. Fluids* **31**, 1577 (1988).
- <sup>3</sup>Ch. P. Ritz, E. J. Powers, and R. D. Bengtson, *Phys. Fluids B* **1**, 153 (1989).
- <sup>4</sup>J. S. Kim, R. D. Durst, R. J. Fonck, E. Fernandez, A. Ware, and P. W. Terry, *Phys. Plasmas* **3**, 3998 (1996).
- <sup>5</sup>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).
- <sup>6</sup>D. A. Baver, P. W. Terry, and R. Gatto, *Phys. Plasmas* **9**, 3318 (2002).
- <sup>7</sup>P. W. Terry, R. Gatto, and D. A. Baver, *Phys. Rev. Lett.* **89**, 205001 (2002).
- <sup>8</sup>P. W. Terry, E. Fernandez, and A. S. Ware, *Astrophys. J.* **504**, 821 (1998).
- <sup>9</sup>M. Jakubowski, "Measurement of velocity turbulence and application to zonal flow detection in tokamak plasmas," Ph.D. thesis, University of Wisconsin-Madison, 2003.
- <sup>10</sup>P. W. Terry and W. Horton, *Phys. Fluids* **25**, 491 (1982).
- <sup>11</sup>P. W. Terry, A. S. Ware, and D. E. Newman, *Phys. Plasmas* **1**, 3974 (1994).
- <sup>12</sup>A. Chekhlov, S. A. Orszag, S. Sukoriansky, B. Galperin, and I. Staroselsky, *Physica D* **98**, 321 (1996).