Efficient scheme for parametric fitting of data in arbitrary dimensions

Ning-Ning Pang,[1,*] Wen-Jer Tzeng,[2,†] and Hisen-Ching Kao[1]

[1]*Department of Physics, National Taiwan University, Taipei, Taiwan*
[2]*Department of Physics, Tamkang University, Tamsui, Taipei, Taiwan*

We propose an efficient scheme for parametric fitting expressed in terms of the Legendre polynomials. For continuous systems, our scheme is exact and the derived explicit expression is very helpful for further analytical studies. For discrete systems, our scheme is almost as accurate as the method of singular value decomposition. Through a few numerical examples, we show that our algorithm costs much less CPU time and memory space than the method of singular value decomposition. Thus, our algorithm is very suitable for a large amount of data fitting. In addition, the proposed scheme can also be used to extract the global structure of fluctuating systems. We then derive the exact relation between the correlation function and the detrended variance function of fluctuating systems in arbitrary dimensions and give a general scaling analysis.

PACS number(s): 02.50.−r, 05.40.−a, 05.45.Tp, 02.60.−x

I. INTRODUCTION

Various physical systems in Nature, from the scale of atomic clusters to the scale of oceans, consist of macroscopic structures. Some examples are molecular-beam-epitaxial growth at high temperature [1,2], pulse-current electrochemical deposition [3], cultivated tumor growth [4], and the temperature and salinity distribution in the world oceans [5]. Moreover, many time series data also show macroscopic trends, such as the clustering features of seismic sequences [6], the interfiring time intervals between neural action potentials [7], and the heart-rate fluctuations during sleep [8]. When performing various experiments, people try to condense the data and extract the essence by fitting data to appropriate parametric models. Without knowledge of the origin of macroscopic structures, the polynomial function is commonly chosen to be the fitting function. The values of the coefficients of the polynomial function are determined by a least-squares fit to the data set.

In the literature, the most commonly used methods for the least-squares fit are the method of normal equations and the method of singular value decomposition (SVD) [9–11]. However, each of these two methods has some drawbacks. (1) Although the ideas behind the method of normal equations are very straightforward and natural, the analytical solutions of the normal equations are hardly attainable except in some simple cases. In addition, the numerical solutions of the normal equations suffer seriously from the roundoff errors. (2) Among the available methods in the literature, the SVD method is known to be the most accurate method for a least-squares fit. However, this method takes a relatively long computation time and large memory space. Thus, to fit a large amount of data points to a high-degree polynomial in high dimensions, the SVD method is in practice not feasible. Since the need for data fitting to polynomial functions is very

common, we are strongly motivated to construct a fitting algorithm, which is relatively efficient and accurate for processing large data sets in arbitrary dimensions.

Notably, fluctuating systems with the formation of macroscopic structures have drawn much attention in recent years. Some examples are electrochemical deposition, physiology signals, and the mosaic structure of DNA sequences [3,7,8,12–14]. The common feature of these systems is the existence of long-range positive (or negative) correlation between the fluctuations. The scaling exponents can then be used to categorize the universality classes of these systems. People usually adopt the polynomial approximation to obtain the best guess at the global structures of such systems. Since these fluctuating systems usually contain large amounts of data points, our proposed fitting scheme is very suitable for extracting the global structures of these systems effectively. In addition, in our proposed algorithm, the fitting parameters are explicitly expressed in a systematic way. It will be very helpful for further analytical study of the scaling analysis of fluctuations.

Interestingly, in the field of computational molecular biology, Peng *et al.* [12–14] have proposed a "detrended fluctuation analysis" (DFA) method to determine the long-range correlation in DNA sequences but exclude the effect caused by the mosaic structure. Its basic algorithm [12–14] is briefly given as follows. The first step is to extract the trend (or the macroscopic structure), which is expressed in terms of a qth-degree polynomial with the coefficients determined by a numerical least-squares fit. The second step is to calculate the corresponding qth-degree detrended variance function and analyze its scaling behaviors. Although the DFA method has been widely accepted as a standard analysis scheme in the study of fluctuating systems with the formation of global structures, there still exist two unresolved issues in this method. (1) The DFA method is based on the following conjecture. For any fluctuating system in arbitrary dimensions, the detrended variance function can successfully suppress the influence of trends (or global structures) and, at the same time, retain the scaling behaviors originating from the stochastic nature of the system. Although this conjecture is intuitively correct, it still lacks a rigorous proof. (2) The explicit and exact relations among the correlation functions, the

———————
*Author to whom correspondence should be addressed. nnp@phys.ntu.edu.tw
†Author to whom correspondence should be addressed. wjtzeng@mail.tku.edu.tw

variance functions, and the detrended variance functions of fluctuating systems in arbitrary dimensions have not yet been derived. We are strongly motivated to address the above two issues rigorously. Note that, in our proposed fitting scheme, all the fitting parameters of the global structure are explicitly expressed in terms of Legendre polynomials. By employing this expression, we are able to analytically obtain the exact relation among the correlation functions, the variance functions, and the detrended variance functions. If these relations are obtained, one just needs to know the information about the correlation functions and then the other two quantities can be easily derived. Hence, a large amount of computation time can be saved. Furthermore, by using these obtained exact relations, we can then give a rigorous proof of the intuitive conjecture of the DFA method.

The outline of this paper is as follows. In Sec. II, we briefly review the concept of the general least-squares fit and its two commonly used methods: normal equations and the SVD method. In Sec. III, by employing the Legendre polynomials, we propose a scheme for parametric fitting and discuss its merits. In Sec. IV, we use three examples to numerically demonstrate the merits of our proposed scheme over the old methods. In Sec. V, the exact relation between the correlation function and the detrended variance function for fluctuating systems with formation of global structures is derived. We then present a scaling analysis applicable for any fluctuating systems. In Sec. VI, a brief summary is given.

II. GENERAL LEAST-SQUARES FIT AND THE METHOD OF SINGULAR VALUE DECOMPOSITION

In this section, we will first briefly review the concept of general least-squares fit and its two commonly used methods: normal equations and the SVD method. Then, the merits and the drawbacks of each method will be discussed.

Suppose that one is measuring a variable y as a function of $\mathbf{x}$, a vector in the d-dimensional space. Given a set of data points $\{y(\mathbf{x})\}$, measured within a d-dimensional observation window, one often would like to condense the data and extract the essential information by fitting the data set to a parametric model with adjustable parameters. Without losing generality, the polynomial functions are usually chosen to be the fitting functions and the fit supplies the appropriate coefficients. That is, the fitting function is a $\{q\}$th-degree $[\{q\} \equiv (q_1, \ldots, q_d)]$ polynomial with adjustable parameters $a_{\{n\}}$,

$$\tilde{y}_{\{q\}}(\mathbf{x}) \equiv \sum_{\{n\}=\{0\}}^{\{q\}} a_{\{n\}} \prod_{j=1}^{d} (x_j - \hat{x}_j)^{n_j},$$

with the summation running over the set of indices $0 \le n_i \le q_i$ for all $1 \le i \le d$, and one obtains the appropriate parameter values by minimizing

$$\langle [y(\mathbf{x}) - \tilde{y}_{\{q\}}(\mathbf{x})]^2 \rangle$$

with $\langle \cdots \rangle$ denoting the average over a d-dimensional observation window centered at $\hat{\mathbf{x}}$ and of size $l_1 \times \cdots \times l_d$. For continuous systems, the coefficients $\{a_{\{n\}}\}$ satisfy the following relations:

$$\left\langle y(\mathbf{x}) \prod_{j=1}^{d} (x_j - \hat{x}_j)^{n_j} \right\rangle = \sum_{\{n'\}=\{0\}}^{\{q\}} \left\{ a_{\{n'\}} \prod_{j=1}^{d} \left(\frac{l_j}{2} \right)^{n_j+n'_j} \right. \\ \left. \times \frac{[1 + (-1)^{n_j+n'_j}]}{2(n_j + n'_j + 1)} \right\} \quad (1)$$

for all $\{n\}$'s. For discrete systems, the coefficients $\{a_{\{n\}}\}$ satisfy the relations

$$\left\langle y(\mathbf{x}) \prod_{j=1}^{d} (x_j - \hat{x}_j)^{n_j} \right\rangle \\ = \sum_{\{n'\}=\{0\}}^{\{q\}} \left\{ a_{\{n'\}} \prod_{j=1}^{d} \left[\frac{1}{l_j} \sum_{x_j=1}^{l_j} \left(x_j - \frac{l_j}{2} \right)^{n_j+n'_j} \right] \right\} \quad (2)$$

for all $\{n\}$'s. In the literature, the collection of Eqs. (1) or (2) are called the normal equations of the least-squares problem. To determine $\{a_{\{n\}}\}$ in Eqs. (1) or (2), one needs to deal with the inverse of a tensor of rank $2d$. Since the above line of thought is very natural and straightforward, one usually comes to solving the normal equations. However, the numerical solution of $\{a_{\{n\}}\}$ directly from Eqs. (2) is very susceptible to roundoff error, and to analytically obtain $\{a_{\{n\}}\}$ from Eqs. (1) or (2) is barely feasible except for some simple cases. Thus, solving normal equations directly is not a recommended way to deal with the least-squares problem.

For discrete systems, Cuyt *et al.* [15] have proposed to reformulate the multivariate data fitting problem in terms of a product of orthogonal polynomial basis, instead of the multinomial basis. For discrete systems with the number of data points (which are sampled at optimal locations) equal to the number of fitting parameters, these authors have theoretically shown that, by employing the fast LU factorization with partial pivoting, both computational cost and the degree of ill conditioning can be greatly improved when using an orthogonal polynomial basis. Note that, in linear algebra, the LU factorization with partial pivoting is a matrix decomposition which writes a matrix as the product of a lower triangular matrix, an upper triangular matrix, and a permutation matrix.

To numerically solve all kinds of least-squares problems, the most often recommended technique in the literature is the SVD method [9,11] for its general and rigorous treatment of ill-conditioned situations. This method is based on the following theorem of linear algebra [9,11]. Any $m \times n$ matrix A, with $m \ge n$, can be decomposed as

$$A = U \cdot \Phi \cdot V^{\mathrm{T}}.$$

The matrices U, Φ, and V are an $m \times n$ column-orthogonal matrix, an $n \times n$ diagonal matrix with elements ϕ_j being positive or zero (the singular value), and an $n \times n$ orthogonal matrix, respectively. Given an $m \times n$ matrix A and an $m \times 1$ vector $\boldsymbol{\beta}$, the $n \times 1$ vector $\boldsymbol{\alpha}$ minimizing $|A \cdot \boldsymbol{\alpha} - \boldsymbol{\beta}|$ is shown to be

$$\boldsymbol{\alpha} = V \cdot \widetilde{\Phi} \cdot U^{\mathrm{T}} \cdot \boldsymbol{\beta}$$

with the elements of the $n \times n$ diagonal matrix $\widetilde{\Phi}$ being $1/\phi_j$ (if $\phi_j \neq 0$) or zero (if $\phi_j = 0$). See Refs. [9,11] for the detailed proof of the above theorem. Subsequently, we give the main steps for implementing the SVD method in parametric fitting of data.

(1) Put the measured data set $\{y(\mathbf{x})\}$ in the $m \times 1$ vector $\boldsymbol{\beta}$ with the elements $\beta_j = y(\mathbf{x}_j)$ and $\{\mathbf{x}_j; \ j=1,\ldots,m\}$ denoting the data points in the $\mathbf{x}$ space.

(2) Let $\{f_i(\mathbf{x}); \ i=1,\ldots,n\}$ denote the basis functions and the fitting function is given by $\Sigma_{i=1}^{n} \alpha_i f_i(\mathbf{x})$.

(3) The $n \times 1$ vector $\boldsymbol{\alpha}$ is then constructed by the fitting parameters $\{\alpha_i\}$ and the $m \times n$ matrix A constructed by $\{f_i(\mathbf{x})\}$ with the elements $A_{ji} = f_i(\mathbf{x}_j)$.

(4) Then the appropriate fitting parameters $\{\alpha_i\}$ are obtained by the minimization of $|A \cdot \boldsymbol{\alpha} - \boldsymbol{\beta}|$. By employing the theorem we just mentioned, one gets $A = U \cdot \Phi \cdot V^{\mathrm{T}}$ and $\boldsymbol{\alpha} = V \cdot \widetilde{\Phi} \cdot U^{\mathrm{T}} \cdot \boldsymbol{\beta}$.

In theory, there should not be any column degeneracy in the matrix A. However, in some cases the matrix A might be ill conditioned; namely, some of the ϕ_j's are so small that their apparent values are probably an artifact of roundoff error. In such cases, the vector $\boldsymbol{\alpha}$ (of fitting parameters) obtained by zeroing the small ϕ_j's is usually much better than the solution with the small ϕ_j's left nonzero. Thus, the SVD method cannot be applied blindly. One must decide at what threshold to zero the small ϕ_j's. If the SVD method is applied correctly, it is the most reliable method in dealing with the least-squares problem. However, it has one significant disadvantage: It usually requires a very long computation time for iteration and a large memory space to store an extra array of size $m \times n$. Thus, for the fitting with a large amount of data points, the SVD method might be practically infeasible.

III. AN ALGORITHM WITH GREAT EFFICIENCY

In this section, we would like to propose a different scheme for parametric fitting. Our proposed scheme is exact for continuous systems. For discrete systems, this scheme is almost as accurate as the SVD method, and it takes much less CPU time and memory space than the SVD method in computation. In the following, we will first give a detailed derivation of our scheme and then discuss its implications and applications.

Recall that, to determine $\{a_{\{n\}}\}$ in Eqs. (1) or (2), one needs to deal with the inverse of a tensor of rank $2d$. For continuous systems, one can actually avoid the difficulty in solving the inverse of a high-rank tensor by choosing an appropriate set of orthogonal polynomial basis $\{f_i(x)\}$, which satisfies the following relation:

$$\langle f_i(x)f_j(x) \rangle \equiv \frac{1}{l} \int_{-l/2}^{l/2} dx \, f_i(x)f_j(x) = 0 \quad \text{for all } i \neq j. \tag{3}$$

In the literature [16], the classical orthogonal polynomials are categorized into three classes, Jacobi-like, Laguerre-like,

and Hermite-like polynomials, with the intervals of orthogonality being $[-1,1]$, $[0,\infty)$, and $(-\infty,\infty)$, respectively. Obviously, Laguerre-like and Hermite-like polynomials do not meet the required relation Eq. (3). In general, Jacobi-like polynomials $\{J_i^{(\mu,\nu)}(x)\}$ satisfy the orthogonal relation $\int_{-1}^{1} dx \, H^{(\mu,\nu)}(x)J_i^{(\mu,\nu)}(x)J_j^{(\mu,\nu)}(x) = 0$ for all $i \neq j$, with the weight function $H^{(\mu,\nu)}(x) = (1-x)^{\mu}(1+x)^{\nu}$ and the parameters $\mu, \nu > -1$. An important subclass of Jacobi-like polynomials is the Gegenbauer polynomials, which correspond to $\{J_i^{(\mu,\nu)}(x)\}$ with $\mu = \nu > -1$. Gegenbauer polynomials consist of two important subclasses: Legendre polynomials and two types of Chebyshev polynomials, corresponding to $\{J_i^{(\mu,\nu)}(x)\}$ with $\mu = \nu = 0$ and $\mu = \nu = \pm 1/2$, respectively. Clearly, the set of Legendre polynomials has the weight function simply 1 and will satisfy the required relation Eq. (3) under appropriate scaling. Hence, the set of Legendre polynomials

$$P_i(x) = \sum_{j=0}^{[i/2]} (-1)^j \frac{(2i-2j-1)!!}{(2j)!!(i-2j)!} x^{i-2j} \quad \text{for } i=1,2,\ldots \tag{4}$$

with the orthogonal property $\int_{-1}^{1} dx \, P_i(x)P_{i'}(x) = \frac{2\delta_{i,i'}}{2i+1}$ is a perfect candidate to be chosen.

We express

$$\widetilde{y}_{\{q\}}(\mathbf{x}) = \sum_{\{n\}=\{0\}}^{\{q\}} c_{\{n\}} \prod_{j=1}^{d} P_{n_j}\left(\frac{2(x_j - \hat{x}_j)}{l_j}\right). \tag{5}$$

The minimization of $\langle [y(\mathbf{x}) - \widetilde{y}_{\{q\}}(\mathbf{x})]^2 \rangle$ then leads to

$$c_{\{n\}} = \left\langle y(\mathbf{x}) \prod_{j=1}^{d} (2n_j + 1) P_{n_j}\left(\frac{2(x_j - \hat{x}_j)}{l_j}\right) \right\rangle. \tag{6}$$

For illustration, we explicitly list some coefficients:

$$c_{00} = \langle y(\mathbf{x}) \rangle,$$

$$c_{10} = 6 \langle y(\mathbf{x}) z_1 \rangle,$$

$$c_{11} = 36 \langle y(\mathbf{x}) z_1 z_2 \rangle,$$

$$c_{20} = \frac{5}{2} \langle y(\mathbf{x})(12z_1^2 - 1) \rangle,$$

$$c_{21} = 15 \langle y(\mathbf{x})(12z_1^2 - 1)z_2 \rangle,$$

$$c_{30} = 7 \langle y(\mathbf{x})(20z_1^3 - 3z_1) \rangle,$$

with the dimensionless quantity $z_j \equiv (x_j - \hat{x}_j)/l_j$. Hence, we obtain

$$\widetilde{y}_{\{q\}}(\mathbf{x}) = \sum_{\{n\}=\{0\}}^{\{q\}} \left\langle y(\mathbf{x}) \prod_{j=1}^{d} (2n_j + 1) P_{n_j}\left(\frac{2(x_j - \hat{x}_j)}{l_j}\right) \right\rangle$$
$$\times \prod_{i=1}^{d} P_{n_i}\left(\frac{2(x_i - \hat{x}_i)}{l_i}\right).$$

In the following analysis, the symbol $[q]$ is used to denote the union of the sets of indices $\{n\} \equiv \{n_1, n_2, \ldots, n_d\}$ with

TABLE I. A list of the values of numerical fitting parameters for example 1 in Sec. IV. The values are obtained from the SVD method and our algorithm proposed in Sec. III, respectively. 128×128 data points are used for fitting. The ratios of the CPU time and the memory space cost for our method to those for the SVD method are 0.025 and 0.09, respectively.

	a_{00}	a_{10}	a_{01}	a_{20}	a_{11}	a_{02}	a_{30}	a_{21}	a_{12}	a_{03}
Exact value	5	1	-4	3	1	-2	-1	0	-2	0
SVD	5.000	1.000	-4.000	3.000	1.000	-2.000	-1.000	0.0000	-2.000	0.0000
Our method	5.000	0.9993	-4.001	2.999	0.9999	-1.999	-0.9991	0.0000	-1.999	0.0014

$|\{n\}| \equiv n_1 + n_2 + \cdots + n_d \leq q$ and all n_i being nonnegative integers. Thus,

$$\tilde{y}[q](\mathbf{x}) = \sum_{|\{n\}| \leq q} \left\langle y(\mathbf{x}) \prod_{j=1}^{d} (2n_j + 1) P_{n_j}\left(\frac{2(x_j - \hat{x}_j)}{l_j}\right) \right\rangle$$
$$\times \prod_{i=1}^{d} P_{n_i}\left(\frac{2(x_i - \hat{x}_i)}{l_i}\right). \tag{7}$$

The above result can be viewed as the original $y(\mathbf{x})$ convolved by some specific filter, and this result is exact for continuous systems in arbitrary dimensions.

In addition, for a fluctuating system with the formation of global structure, the above result $\tilde{y}[q](\mathbf{x})$ can also be used to express the global structure. In Sec. V, by employing the explicit expression of $\tilde{y}[q](\mathbf{x})$, we will undertake a detailed analytical study of the correlation function and the detrended variance function of fluctuating systems.

Next, for numerical parametric fitting of discrete systems, we need to take a continuum approximation of the original data set in order to apply the above scheme. Let $\{y(j_1, \ldots, j_d)\}$ denote the original discrete data set within a d-dimensional observation window of grid size $l_1 \times \cdots \times l_d$ (i.e., $j_i = 1, 2, \ldots, l_i$). This data set is approximated as a continuous step function and rescaled to the range $\prod_{i=1}^{d}[-1, 1]$, with the continuous step function $z(\mathbf{x})$ given by $z(\mathbf{x}) = y(j_1, \ldots, j_d)$ for $\frac{j_i - l_i/2 - 1}{l_i/2} \leq x_i < \frac{j_i - l_i/2}{l_i/2}$ and $i = 1, 2, \ldots, d$. Following the scheme for continuous systems, the fitting function is then expressed as

$$\tilde{z}_{\{q\}}(\mathbf{x}) = \sum_{\{n\}=\{0\}}^{\{q\}} c_{\{n\}} \prod_{i=1}^{d} P_{n_i}(x_i).$$

The parametric fitting corresponds to finding the set of coefficients $c_{\{n\}}$ minimizing

$$\int_{-1}^{1} \cdots \int_{-1}^{1} d\mathbf{x} [z(\mathbf{x}) - \tilde{z}_{\{q\}}(\mathbf{x})]^2.$$

It leads to

$$c_{\{n\}} = \int_{-1}^{1} \cdots \int_{-1}^{1} d\mathbf{x}\, z(\mathbf{x}) \prod_{i=1}^{d} \frac{2n_i + 1}{2} P_{n_i}(x_i) \tag{8}$$

for discrete systems. If one is interested in the correspondence between the coefficients $\{c_{\{n\}}\}$ and $\{a_{\{n\}}\}$, it can be easily derived through the relation

$$\sum_{\{n\}=\{0\}}^{\{q\}} c_{\{n\}} \prod_{i=1}^{d} P_{n_i}(x_i) = \sum_{\{n\}=\{0\}}^{\{q\}} a_{\{n\}} \prod_{i=1}^{d} x_i^{n_i}. \tag{9}$$

The merit of our scheme in numerical parametric fitting is efficiency, i.e., taking much less CPU time and memory space than the SVD method. In the following, we will give an estimation of the complexity of computation time and memory occupancy for both the SVD method and our scheme. Let m, n, and d denote the total number of data points, the number of fitting parameters, and the dimensionality of the system, respectively. The main computation time of the SVD method is spent on the execution of Householder reductions, of which the time span is about $O(mn^2)$. By contrast, in our scheme, the operation of the inner product between the data and each basis function costs the CPU time of $O(m)$. There are n basis functions in total. Hence, the computation time complexity of our method is about $O(mn)$. The complexity of memory occupancy for the SVD method is $O(n(m+n))$, which can be easily derived from the main steps of the SVD method given in Sec. II. By contrast, the complexity of memory space for our fitting scheme is $O(m + m^{1/d}n^{1/d})$ [about $O(m)$ for $d \geq 2$], which can be obtained from Eq. (8). We will demonstrate this merit of our method through a few numerical examples in the next section. In addition, the SVD method cannot be applied blindly; namely, one needs to deal with the subtlety of ill-conditioned matrices with great care. In contrast, our scheme is generally applicable everywhere with straightforward execution. Thus, for a large amount of data fitting, our scheme is obviously superior to the traditional SVD method.

IV. NUMERICAL DEMONSTRATION

In this section, we will use three examples to demonstrate the merits of our proposed algorithm in numerical parametric fitting.

A. Example 1: A polynomial of degree 3 in two dimensions

In this example, we will draw the data set from a known polynomial function and compare fitting results from the SVD method and our method. The function we choose is a polynomial of degree 3 in two dimensions,

$$\tilde{y}(\mathbf{x}) = 5 + x_1 - 4x_2 + 3x_1^2 + x_1 x_2 - 2x_2^2 - x_1^3 - 2x_1 x_2^2$$
$$\equiv \sum_{0 \leq i+j \leq 3} a_{ij} x_1^i x_2^j.$$

Table I gives the result of parametric fitting of example 1

with the total grids being 128×128 points. The third row gives the values of $\{a_{ij}\}$ obtained by the SVD method. The fourth row gives the values of $\{a_{ij}\}$ obtained by our algorithm proposed in Sec. III with the transformation of coefficients according to Eq. (9). In this example, the ratios of the CPU time and the memory space cost for our method to those for the SVD method are 0.025 and 0.09, respectively. Our method is obviously much more efficient than the SVD method in computation. The relative deviation between the fitting parameter value obtained by our method and the exact value is less than 0.09%. Subsequently, we numerically test how the fitting results improve as the number of grid points for fitting increases. We find that the relative deviation is less than 0.01% if the total grids are more than 1000×1000 points.

B. Example 2: Interfacial growth

Next, let us consider an example in interfacial growth phenomena [1,2]. The growth process is described by a stochastic partial differential equation in $2+1$ dimensions with spatiotemporally correlated noise:

$$\partial_t h(\mathbf{x},t) = -\nu \nabla^4 h(\mathbf{x},t) + \eta(\mathbf{x},t), \tag{10}$$

where $h(\mathbf{x},t)$ denotes the interface height at position $\mathbf{x}$ and time t, and $\eta(\mathbf{x},t)$ represents Gaussian-distributed noise of zero mean and power-law-decaying correlation

$$\overline{\eta(\mathbf{x},t)\eta(\mathbf{x}',t')} = D|\mathbf{x}-\mathbf{x}'|^{2\rho-2}|t-t'|^{2\theta-1}$$

with $0 \leq \rho < 1/2$ and $0 \leq \theta < 1/2$. Here, the overbar denotes the ensemble average. In the right-hand side (RHS) of Eq. (10), the first term accounts for the surface diffusion of deposits to a nearby position with lower chemical potential and the second term accounts for the shot noise. The above growth equation with white noise corresponds to the famous Herring-Mullins equation [17]. It is well known that many experiments can be well described by the Herring-Mullins equation; for example, the growth of amorphous Si films by thermal evaporation with a low substrate temperature [18], Pt sputter deposited on glass at room temperature [19], and the epitaxial growth of Si on a Si(111) substrate [20]. In contrast, the electrochemical noise is known to be spatiotemporally correlated [21,22] and thus the pulse-current electrochemical
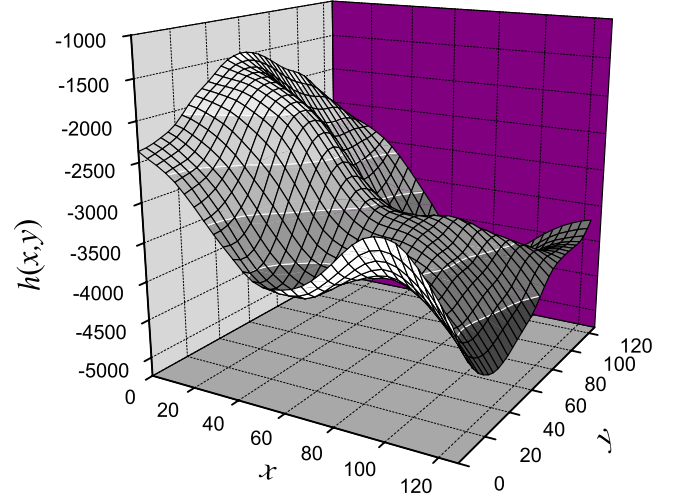

FIG. 1. (Color online) $(2+1)$-dimensional interface configuration in a local observation window of 128×128 lattice sites, numerically generated from Eq. (10) with the parameters $\rho = \theta = 0.3$, $D = \nu = 1$, the lateral system size 4096×4096 lattice sites, and time $t = 1.8 \times 10^6$ time steps.

deposition process is better described by Eq. (10) with spatiotemporally correlated noise.

Through numerically generating the interface configuration governed by Eq. (10) with spatiotemporally correlated noise, we observe that the interface configuration gradually develops large mounds as the growth time increases. Figure 1 shows a typical interface morphology governed by Eq. (10) after a long growth time. Specifically, the values of the parameters in Eq. (10) used to generate Fig. 1 are as follows: the noise correlation exponents $\rho = \theta = 0.3$, the parameters $D = \nu = 1$, the growth time $t = 1.8 \times 10^6$ time steps, and the lateral system size 4096×4096 lattice sites with periodic boundary conditions. Figure 1 plots the $(2+1)$-dimensional interface configuration in a local observation window of 128×128 lattice sites. We take the data set of Fig. 1 as our second example for parametric fitting. The fitting function we choose is a polynomial of degree 4 in two dimensions, $\tilde{y}(\mathbf{x}) = \sum_{0 \leq i+j \leq 4} a_{ij} x_1^i x_2^j$. The local observation window is rescaled to $[-1,1] \times [-1,1]$. Then, the values of the interface heights are rescaled with the same scaling factor along the substrate direction. Table II gives the result of parametric

TABLE II. A list of the values of numerical fitting parameters for example 2 in Sec. IV. The values are obtained from the SVD method and our algorithm proposed in Sec. III, respectively. 128×128 data points are used for fitting. The ratios of the CPU time and the memory space cost for our method to those for the SVD method are 0.02 and 0.06, respectively.

	a_{00}	a_{10}	a_{01}	a_{20}	a_{11}	a_{02}	a_{30}	a_{21}
SVD	-51.93	-19.12	-17.01	35.38	1.562	-45.32	-3.422	15.15
Our method	-51.93	-19.13	-17.00	35.36	1.561	-45.28	-3.413	15.14

	a_{12}	a_{03}	a_{40}	a_{31}	a_{22}	a_{13}	a_{04}
SVD	12.32	10.49	-30.90	15.23	14.31	-17.53	-35.57
Our method	12.32	10.49	-30.87	15.22	14.30	-17.52	-35.52

TABLE III. A list of the values of numerical fitting parameters for example 3 in Sec. IV. The values are obtained from the SVD method and our algorithm proposed in Sec. III, respectively. 355×342 data points are used for fitting. The ratios of the CPU time and the memory space cost for our method to those for the SVD method are 0.014 and 0.043, respectively.

	a_{00}	a_{10}	a_{01}	a_{20}	a_{11}	a_{02}	a_{30}
SVD	18.97	0.1292	25.45	-0.5036	-8.384	-21.49	-1.424
Our method	18.97	0.1300	25.45	-0.4981	-8.385	-21.49	-1.429
	a_{21}	a_{12}	a_{03}	a_{40}	a_{31}	a_{22}	a_{13}
SVD	-5.004	-15.80	-37.56	-1.223	2.650	5.418	4.684
Our method	-4.999	-15.80	-37.54	-1.229	2.651	5.417	4.684
	a_{04}	a_{50}	a_{41}	a_{32}	a_{23}	a_{14}	a_{05}
SVD	7.544	0.6889	-0.3428	1.101	8.573	15.56	19.97
Our method	7.541	0.6932	-0.3486	1.101	8.572	15.56	19.95

fitting of example 2 with the total grids being 128×128 sites. The second row in each box of Table II gives the value of $\{a_{ij}\}$ obtained by the SVD method. The third row gives the values of $\{a_{ij}\}$ obtained by our algorithm with the transformation of coefficients according to Eq. (9). In this example, the ratios of the CPU time and the memory space cost for our method to those for the SVD method are 0.02 and 0.06, respectively. The relative deviation between the fitting parameter value obtained by our method and that by the SVD method is less than 0.1% except for the coefficient a_{30} (whose relative deviation is 0.26%). Again, this discrepancy will diminish as the number of grid points increases. We find that the relative deviation will be less than 0.01% if the total grids are more than 1000×1000 points.

C. Example 3: Annual mean temperature for the world ocean on 0.25° grid

Now let us consider an example in the world ocean temperature distribution. It is well known that the world oceans contain large-scale permanent or semipermanent structures (such as the Pacific Ocean current, the Atlantic Ocean current, and the Gulf Stream), which have great influence on the global climatology. The data set of annual, seasonal, and monthly mean temperature and salinity for the world ocean on a 0.25° latitude and longitude grid is given in the World Ocean Database 2001, which is provided by the National Oceanographic Data Center under the NOAA Satellite and Information Service [23]. This data set is obtained by a special method, the objective analysis technique. (The interested reader may refer to Ref. [5] for a detailed description.)

It is well known that both increasing temperature and decreasing salinity can act to decrease local density. The 0.25° grid climatological mean values of temperature and salinity for the annual, seasonal, and monthly time periods represent mean oceanographic characteristics and can be used to test the validity of various ocean circulation models. Hence, ocean climatologists usually use a polynomial approximation to obtain a best guess at the probable structure of the ocean climatological mean field. As the third example to demonstrate the merits of our fitting algorithm, we draw from the data set of the 0.25° latitude and longitude grid annual mean temperature from World Ocean Database 2001 [23]. The area chosen for this example is the central region of the Pacific Ocean from 179.75°E to 91.5°W longitude and 71.75°S to 13.75°N latitude, at the level of 125 m depth from the sea surface. The total grid points are 355×342 points. The rise and fall of the temperature are due to an ocean current passing through this region. The fitting function we choose is a polynomial of degree 5 in two dimensions, $\tilde{y}(\mathbf{x}) = \Sigma_{0 \le i+j \le 5} a_{ij} x_1^i x_2^j$. The latitude and longitude grids for fitting are rescaled to $[-1, 1] \times [-1, 1]$. Table III gives the result of parametric fitting of example 3. The second row in each column of Table III gives the values of $\{a_{ij}\}$ obtained by the SVD method. The third row gives the values of $\{a_{ij}\}$ obtained by our algorithm with the transformation of coefficients according to Eq. (9). In this example, the ratios of the CPU time and the memory space cost for our method to those for the SVD method are 0.014 and 0.043, respectively. We observe that the reduction of the computation time for our method (over the SVD method) becomes more significant as the number of fitting parameters increases. In this example, the relative deviation between the fitting parameter value obtained by our method and that by the SVD method is larger than in examples 1 and 2, because the total number of fitting parameters in this example is much larger than those in examples 1 and 2. The relative deviation between the value obtained by our method and that by the SVD method is less than 0.6% except for the coefficients a_{20} and a_{41} (whose relative deviations are 1.09% and 1.7%, respectively).

In words, our algorithm for numerical fitting is much more efficient than the traditional SVD method. In addition, the SVD method needs to take a large memory space to store a matrix of size $m \times n$ (with m being the total number of data points and n being the total number of fitting parameters) and to decompose the matrix into a product of several matrices. Thus, for a large amount of data fitting, the SVD method may become infeasible in practice. In contrast, our method

does not take large memory space and can always be handled just by a desktop PC. The only drawback of our method in numerical fitting is that it is an approximation. However, the precision of our fitting results can be greatly enhanced as the total number of data points increases. Hence, we believe that our method is the best algorithm for fitting with a large number of data points.

V. THE CORRELATION FUNCTION, THE VARIANCE FUNCTION, AND THE DETRENDED VARIANCE FUNCTION

Since fluctuating systems with the formation of global structures are widely observed in Nature, let us focus on such systems in this section. The most important statistical quantities in this class of systems are the correlation function, the variance function, and the detrended variance function. In the following, we plan to explicitly obtain the exact relations among these statistical quantities. Let $y(\mathbf{x})$ represent a random function of $\mathbf{x}$, with $\mathbf{x}$ being a vector in a d-dimensional space. The correlation function is defined as

$$G(\mathbf{r}) \equiv \overline{[y(\mathbf{x}) - y(\mathbf{x} + \mathbf{r})]^2}$$

with the overbar denoting the ensemble average. The variance function within a d-dimensional observation window is defined as

$$w^2(l_1, \ldots, l_d) \equiv \overline{\langle [y(\mathbf{x}) - \langle y(\mathbf{x}) \rangle]^2 \rangle} \qquad (11)$$

with $\langle \cdots \rangle$ and the overbar denoting the spatial average over an observation window of size $\prod_{i=1}^{d} l_i$ and the ensemble aver-

age, respectively. In addition, by using the scheme we proposed in Sec. III, one can easily extract the macroscopic structure of the fluctuating system within the observation window. Let the $[q]$th-degree polynomial $\tilde{y}[q](\mathbf{x})$, Eq. (7), represent the $[q]$th-degree global structure function of the fluctuating system. Then, the $[q]$th-degree detrended variance function within a d-dimensional observation window is defined as

$$w^2[q](l_1, \ldots, l_d) \equiv \overline{\langle [y(\mathbf{x}) - \tilde{y}[q](\mathbf{x})]^2 \rangle}. \qquad (12)$$

With some calculation, we first obtain the relation between the correlation function $G(\mathbf{r})$ and the variance function $w^2(l_1, \ldots, l_d)$ as

$$w^2(l_1, \ldots, l_d) = \left(\prod_{i=1}^{d} \frac{2}{l_i^2} \int_0^{l_i} dr_i (l_i - r_i) \right) \frac{1}{2} G(\mathbf{r}). \qquad (13)$$

By using Eqs. (5), (7), (11), and (12), the $[q]$th-degree detrended variance function is related to the variance function as follows:

$$w^2[q](l_1, \ldots, l_d) = w^2(l_1, \ldots, l_d) - \sum_{|\{n\}|=1}^{q} c_{\{n\}}^2 \left(\prod_{j=1}^{d} \frac{1}{2n_j + 1} \right). \qquad (14)$$

Subsequently, by employing Eq. (6), one gets

$$\overline{c_{\{n\} \neq \{0\}}^2} = \prod_{j=1}^{d} \left[\frac{(2n_j + 1)^2}{l_j^2} \int_{-l_j/2}^{l_j/2} dx_j \int_{-l_j/2}^{l_j/2} dx_j' P_{n_j}\left(\frac{2x_j}{l_j} \right) P_{n_j}\left(\frac{2x_j'}{l_j} \right) \right] \overline{y(\mathbf{x})y(\mathbf{x}')}$$

$$= -\frac{1}{2} \prod_{j=1}^{d} \left[\frac{(2n_j + 1)^2}{l_j^2} \int_{-l_j/2}^{l_j/2} dx_j \int_{-l_j/2 - x_j}^{l_j/2 - x_j} dr_j P_{n_j}\left(\frac{2x_j}{l_j} \right) P_{n_j}\left(\frac{2x_j + 2r_j}{l_j} \right) \right] G(\mathbf{r}). \qquad (15)$$

By employing the properties that the Legendre polynomials are either even or odd functions of their arguments and the correlation function $G(\mathbf{r})$ is an even function of all r_j, we obtain

$$\overline{c_{\{n\} \neq \{0\}}^2} = -\frac{1}{2} \prod_{j=1}^{d} \left\{ \frac{2(2n_j + 1)^2}{l_j^2} \int_0^{l_j} dr_j \left[\int_{-l_j/2}^{l_j/2 - r_j} dx_j P_{n_j}\left(\frac{2x_j}{l_j} \right) \right. \right.$$

$$\left. \left. \times P_{n_j}\left(\frac{2x_j + 2r_j}{l_j} \right) \right] \right\} G(\mathbf{r}). \qquad (16)$$

Substituting Eqs. (13) and (16) into Eq. (14), we consequently obtain the exact relation between the $[q]$th-degree detrended variance function and the correlation function as

$$w^2[q](l_1, \ldots, l_d) = \frac{1}{2} \prod_{j=1}^{d} \left(\frac{1}{l_j^2} \int_0^{l_j} dr_j \right) G(\mathbf{r}) K[q](\mathbf{r}) \qquad (17)$$

with the kernel given by

$$K[q](\mathbf{r}) = \sum_{|\{n\}| \leq q} \left\{ \prod_{j=1}^{d} \left[2(2n_j + 1) \int_{-l_j/2}^{l_j/2 - r_j} dx_j P_{n_j}\left(\frac{2x_j}{l_j} \right) \right. \right.$$

$$\left. \left. \times P_{n_j}\left(\frac{2x_j + 2r_j}{l_j} \right) \right] \right\}. \qquad (18)$$

As an illustration, let us consider the case with $d = 2$. We explicitly list out $K[0](\mathbf{r})$ to $K[3](\mathbf{r})$:

$$K[0](\mathbf{r}) = \prod_{j=1}^{2} (2l_j - 2r_j),$$

$$K[1](\mathbf{r}) = K[0](\mathbf{r}) + \sum_{j=1}^{2} \left(2l_j - 6r_j + 4\frac{r_j^3}{l_j^2} \right)(2l_{j'} - 2r_{j'}),$$

$$K[2](\mathbf{r}) = K[1](\mathbf{r}) + \sum_{j=1}^{2} \left(2l_j - 10r_j + 20\frac{r_j^3}{l_j^2} - 12\frac{r_j^5}{l_j^4} \right)$$
$$\times (2l_{j'} - 2r_{j'}) + \prod_{j=1}^{2} \left(2l_j - 6r_j + 4\frac{r_j^3}{l_j^2} \right),$$

$$K[3](\mathbf{r}) = K[2](\mathbf{r}) + \sum_{j=1}^{2} \left[\left(2l_j - 14r_j + 56\frac{r_j^3}{l_j^2} - 84\frac{r_j^5}{l_j^4} + 40\frac{r_j^7}{l_j^6} \right) \right.$$
$$\times (2l_{j'} - 2r_{j'}) + \left(2l_j - 10r_j + 20\frac{r_j^3}{l_j^2} - 12\frac{r_j^5}{l_j^4} \right)$$
$$\left. \times \left(2l_{j'} - 6r_{j'} + 4\frac{r_{j'}^3}{l_{j'}^2} \right) \right],$$

with $j' \equiv j \pmod{2} + 1$.

Since many fluctuating systems are modeled by continuous stochastic partial differential equations, the obtained explicit relations among the correlation function, the variance function, and the detrended variance function can greatly help the analytical investigation of such systems. In addition, the direct numerical calculation of the high-order detrended variance function from its definition, Eq. (12), is very CPU-time consuming. Now, with the explicit relations among correlation function, the variance function, and the detrended variance function being derived, one just needs to obtain the correlation function first and then the detrended variance function can be obtained easily without the cost of large CPU time.

To numerically verify the obtained relations [Eqs. (17) and (18)], we perform some numerical simulation on two-dimensional fractal surfaces as follows. We first numerically generate $y_{\mathrm{sto}}(x_1,x_2)$, which is the fractional Brownian field characterized by the Hurst exponent α [24]. The variance function of the two-dimensional fractional Brownian surface $w^2(l_1,l_2) \equiv \langle [y_{\mathrm{sto}}(x_1,x_2) - \langle y_{\mathrm{sto}}(x_1,x_2) \rangle]^2 \rangle$ is expected to scale as $(l_1^2 + l_2^2)^\alpha$ with $0 < \alpha < 1$. Without losing generality, we choose cases with the Hurst exponent $\alpha = 0.3$, 0.5, and 0.7 in our simulation. To check the accuracy of $y_{\mathrm{sto}}(x_1,x_2)$, we draw a log-log plot (Fig. 2) of $w^2(l_1,l_2)$ vs $l_1^2 + l_2^2$ for the numerically generated fractional Brownian surfaces. The whole fractal domain is 4096×4096 lattice sites. The variance function is first calculated within an observation window of size $l_1 \times l_2$ and then a sliding average is taken over the whole fractal domain. The excellence of the data fitting in Fig. 2 confirms the accuracy of $y_{\mathrm{sto}}(x_1,x_2)$ representing the fractional Brownian field. Subsequently, we numerically impose a global structure function

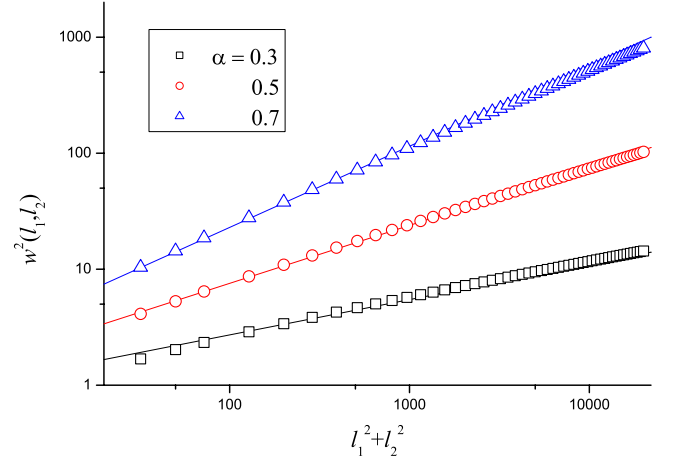FIG. 2. (Color online) Log-log plot of the variance function $w^2(l_1,l_2)$ for two-dimensional fractional Brownian surfaces vs $l_1^2 + l_2^2$, in which (l_1,l_2) denote the side lengths of the observation window. The whole system size is 4096×4096 lattice sites. The data refer to surfaces with Hurst exponents $\alpha = 0.3$, 0.5, and 0.7, represented by the square, circle, and triangle symbols, respectively. The solid lines are drawn as guides to the eye.

$$y_{\mathrm{global}}(x_1,x_2) = 4\left(\frac{x_1}{10^3}\right) + 3\left(\frac{x_1}{10^3}\right)\left(\frac{x_2}{10^3}\right) - 2\left(\frac{x_1}{10^3}\right)\left(\frac{x_2}{10^3}\right)^2$$
$$- \left(\frac{x_1}{10^3}\right)^3 \tag{19}$$

in the same fractal domain, i.e., $x_i = 1, 2, \ldots, 4096$. Consequently, the two-dimensional surfaces are defined as $y(x_1,x_2) = y_{\mathrm{global}}(x_1,x_2) + y_{\mathrm{sto}}(x_1,x_2)$. We then numerically calculate the $[q]$th-degree detrended variance function $w^2[q]$ in two ways: The first way is the direct calculation from its definition, Eq. (12), and the second way is through the derived relation, Eqs. (17) and (18). Figure 3 clearly shows that $R[q]$, the ratio of the values of $w^2[q]$ obtained in these two ways, always falls in the range $[0.99, 1.01]$ for various values of the Hurst exponent α and the detrended degree q, if the side length of the observation window (l) is larger than 24 lattice sites. Hence, the derived explicit relation between the correlation function and the detrended variance function, Eqs. (17) and (18), is numerically reconfirmed. Moreover, one can employ Eqs. (17) and (18) to numerically obtain the values of $w^2[q]$, and the reduction of computation time is remarkable. For example, the ratio of the computation time to obtain $w^2[q]$ (with $l = 4, 8, \ldots, 100$ and $q = 1, 2, 3$) from Eqs. (17) and (18) to that from Eq. (12) is about 0.012.

Next, we would like to address the following issue. In the formalism of the DFA method proposed by Peng et al. [12–14], the detrended variance function is designed to retain the scaling behaviors due to the stochastic nature of the systems and to exclude the influence of global structures. Although this design is intuitively correct and has been justified by numerical examples in the literature, it still lacks a rigorous analytical proof. In the following, by undertaking a detailed analysis of the kernel function $K[q](\mathbf{r})$, we will give a rigorous proof of the above issue. We first recast $y(\mathbf{x})$ as

$y(\mathbf{x}) = y_{\text{global}}(\mathbf{x}) + y_{\text{sto}}(\mathbf{x})$. $y_{\text{global}}(\mathbf{x})$ represents the global structure, expected to be continuous and smooth. $y_{\text{sto}}(\mathbf{x})$ represents the part from some stochastic nature of the system relative to the global structure. From the definition of the correlation function $G(\mathbf{r})$, $G(\mathbf{r})$ can also be separated into two parts $G_{\text{global}}(\mathbf{r}) + G_{\text{sto}}(\mathbf{r})$. It is physically reasonable to assume that $G_{\text{global}}(\mathbf{r})$ is analytic and thus can be expressed in terms of a power series expansion. Since the correlation function is a "difference" correlation function, we further

expect that $G_{\text{global}}(\mathbf{r})$ is an even function of all r_j and thus $G_{\text{global}}(\mathbf{r}) = \Sigma_{|\{n\}|=1}^{\infty} b_{\{n\}} \Pi_{j=1}^{d} r_j^{2n_j}$. In most physical situations, the dominant term of $G_{\text{sto}}(\mathbf{r})$ is proportional to $r^{2\alpha}$ with a characteristic scaling exponent α, which is usually not an integer.

Subsequently, the kernel $K[q](\mathbf{r})$ is rewritten as $K[q](\mathbf{r}) \equiv \Sigma_{|\{n\}| \le q}[\Pi_{i=1}^{d} F_{n_i}(r_i)]$. After some calculation, we obtain the following property of $F_{n_i}(r_i)$, with m_i being any nonnegative number (including noninteger):

$$\int_0^{l_i} dr_i\, r_i^{2m_i} F_{n_i}(r_i) = \begin{cases} 0, & m_i = 1, 2, \ldots, n_i - 1, \\ \dfrac{l_i^{2m_i+2}(-1)^{n_i}(2n_i+1)\Gamma(m_i+1)^2}{(2m_i+1)\Gamma(m_i-n_i+1)\Gamma(m_i+n_i+2)} & \text{otherwise.} \end{cases} \qquad (20)$$

Then, the above relation is employed to compute $\int_{\Pi_{i=1}^{d}[0,l_i]} K[q](\mathbf{r}) r^{2\mu} d\mathbf{r}$.

(i) For μ a positive integer,

$$\int_{\Pi_{i=1}^{d}[0,l_i]} K[q](\mathbf{r}) r^{2\mu} d\mathbf{r} = \sum_{|\{n\}| \le q} \sum_{|\{m\}|=\mu} \frac{\mu!}{\prod_{i=1}^{d}(m_i!)} \left(\prod_{i=1}^{d} \int_0^{l_i} F_{n_i}(r_i) r_i^{2m_i} dr_i \right)$$

$$= \sum_{|\{n\}| \le q} \sum_{|\{m\}|=\mu} \frac{\mu!}{\prod_{i=1}^{d}(m_i!)} \prod_{i=1}^{d} \begin{cases} 0, & m_i = 1, 2, \ldots, n_i - 1, \\ \dfrac{l_i^{2m_i+2}(-1)^{n_i}(2n_i+1)\Gamma(m_i+1)^2}{(2m_i+1)\Gamma(m_i-n_i+1)\Gamma(m_i+n_i+2)} & \text{otherwise.} \end{cases}$$

With some calculation and by induction, it can be shown that the above integral is equal to 0 for $\mu = 1, 2, \ldots, q$. In addition, in many physical situations, one usually takes the observation window with equal side lengths (i.e., $l_i = l$ for all i) and thus the above integral is proportional to $l^{2\mu+2d}$, for $\mu = q + 1, q + 2, \ldots$.

(ii) For μ a noninteger, the calculation becomes extremely complicated and no general closed form can be attained. Since most physical situations have the substrate dimension $d = 2$, we choose to undertake the calculation for $d = 2$ with the side lengths of the observation window equal to l. First, consider the integral

$$H_\mu(l, n_1, n_2) \equiv \int_{l \times l} d\mathbf{r}\, r_1^{n_1} r_2^{n_2} (r_1^2 + r_2^2)^\mu$$

$$= \frac{l^{n_1+n_2+2\mu+2}}{2(n_1+n_2+2\mu+2)} \sum_{i=1}^{2} B\left(\frac{n_i+1}{2}, \right.$$

$$\left. -\mu - \frac{n_i+1}{2}; \frac{1}{2} \right) \qquad (21)$$

with the incomplete Beta function [16]

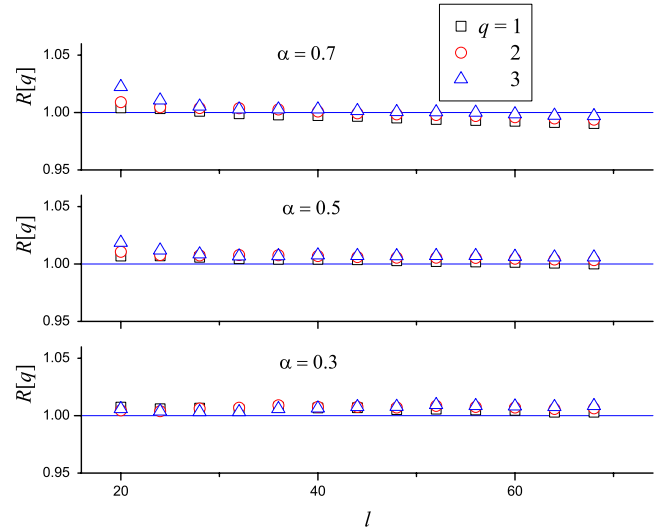$$B(\zeta, \xi; x) \equiv \int_0^x t^{\zeta-1}(1-t)^{\xi-1} dt.$$

FIG. 3. (Color online) $R[q]$ represents the ratio of the value of the $[q]$th-degree detrended variance function $w^2[q]$ calculated from Eqs. (17) and (18) to that calculated from Eq. (12). The side lengths (l_1, l_2) of the observation window are set as $l_1 = l_2 = l = 20, 24, \ldots, 68$. The whole system size is 4096×4096 lattice sites. Note that the two-dimensional surfaces are set as $y(x_1, x_2) = y_{\text{global}}(x_1, x_2) + y_{\text{sto}}(x_1, x_2)$, with y_{global} given by Eq. (19) and y_{sto} being the fractional Brownian surfaces shown in Fig. 2. The solid lines are drawn as guides to the eye.

By employing the above integral and the expansion formula for the Legendre polynomial, we then derive

$$
\int_{l \times l} d\mathbf{r} \, F_{n_1}(r_1) F_{n_2}(r_2)(r_1^2 + r_2^2)^\mu
$$

$$
= 4l^2 H_\mu(l,0,0) + 4 \prod_{i=1}^{2} \left(\sum_{j_i=0}^{n_i} \frac{(2n_i+1)(-n_i)_{j_i}(1+n_i)_{j_i}}{(2j_i+1)(j_i!)^2 l^{2j_i}} \right)
$$

$$
\times H_\mu(l, 2j_1+1, 2j_2+1) - 4(2n_2+1)
$$

$$
\times \sum_{k=0}^{n_2} \frac{(-n_2)_k(1+n_2)_k}{(2k+1)(k!)^2 l^{2k-1}}
$$

$$
\times H_\mu(l,0,2k+1) - 4(2n_1+1) \sum_{j=0}^{n_1} \frac{(-n_1)_j(1+n_1)_j}{(2j+1)(j!)^2 l^{2j-1}}
$$

$$
\times H_\mu(l, 2j+1, 0) \tag{22}
$$

with $(n)_j \equiv n(n-1)\cdots(n-j+1)$. Subsequently, by using Eqs. (21) and (22) and the following relations for the incomplete Beta function [16]:

$$
B(\xi, \zeta; x) = \frac{\Gamma(\xi)\Gamma(\zeta)}{\Gamma(\xi+\zeta)} - B(\zeta, \xi; 1-x), \tag{23}
$$

$$
B(\zeta, \xi; x) = \frac{(1-x)^\xi}{\zeta} \sum_{j=0}^{\infty} \frac{(\xi+\zeta)_j}{(\zeta+1)_j} x^{j+\zeta}, \tag{24}
$$

we finally obtain

$$
\int_{l \times l} K[q](\mathbf{r})(r_1^2 + r_2^2)^\mu d\mathbf{r} = \sum_{n_1+n_2 \leq q} \int_0^l dr_1 \int_0^l dr_2 F_{n_1}(r_1) F_{n_2}(r_2)(r_1^2 + r_2^2)^\mu
$$

$$
= l^{2\mu+4} \left[\frac{2^\mu(q^2+2q+2)}{\mu+1} \sum_{\nu=0}^{\infty} \frac{(-\mu)_\nu}{2^\nu(1/2)_{\nu+1}} + 2^\mu \sum_{\nu=0}^{\infty} \frac{(-\mu)_\nu}{2^\nu} \sum_{n_1+n_2 \leq q} (2n_1+1)(2n_2+1) \right.
$$

$$
\times \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \frac{(-n_1)_j(1+n_1)_j(-n_2)_k(1+n_2)_k}{(2j+1)(2k+1)(j+k+\mu+2)(j!\,k!)^2} \left(\frac{1}{(k+1)_{\nu+1}} + \frac{1}{(j+1)_{\nu+1}} \right)
$$

$$
\left. - 2^{\mu+2} \sum_{\nu=0}^{\infty} \frac{(-\mu)_\nu}{2^\nu} \sum_{n=0}^{q} (2n+1)(q-n+1) \sum_{k=0}^{\infty} \frac{(-n)_k(1+n)_k}{(2k+1)(2k+2\mu+3)(k!)^2} \left(\frac{1}{(k+1)_{\nu+1}} + \frac{1}{(1/2)_{\nu+1}} \right) \right]
$$

$$
\propto l^{2\mu+4}. \tag{25}
$$

By substituting the expansion of $G(\mathbf{r})[\equiv G_{\text{global}}(\mathbf{r}) + G_{\text{sto}}(\mathbf{r})]$ into Eq. (17) and employing the properties of the kernel functions we just derived, we rigorously show that, by raising the degree of the detrended variance function, the contribution due to the global structure can be successfully suppressed. The above result is applicable in any dimensions. In addition, at least in two dimensions, we rigorously show that the detrended variance functions do retain the scaling exponent α (if it is not an integer) originating from the stochastic nature of the fluctuating system.

VI. CONCLUSION

In the physical sciences, data fitting is a very important step for researchers to condense the data and extract the essential information from experiments. Given a set of data points, one usually fits the data set to a parametric model. Without losing generality, polynomial functions are commonly chosen to be the fitting functions and the fit gives the appropriate parameter values. In addition, many fluctuating systems in Nature consist of global structures; examples include molecular-beam-epitaxial growth, electrochemical deposition, the clustering features of seismic sequences, the climatological temperature and salinity distribution of the world oceans, physiology signals, and the mosaic structure of DNA sequences. People usually adopt a polynomial approximation to obtain a best guess at the global structures of such systems. In the literature, the method of normal equations and the SVD method are two commonly used methods for numerical fitting. However, each of these two methods has some drawbacks. The method of normal equations is rather susceptible to roundoff errors. Although the SVD method is very accurate, it requires a relatively long computation time and large memory space. Thus, the SVD method is not practically feasible for fitting a large amount of data. In this work, by employing Legendre polynomials, we propose a different algorithm for parametric fitting. Our proposed scheme is exact for continuous systems. For numerical fitting, our proposed schemes takes much less CPU time and memory space than the traditional SVD method in computation. Although our proposed scheme is an approximation for numerical fitting of discrete systems, its accuracy can be greatly enhanced as the number of data points increases. We

believe that our algorithm is the best method for fitting a large amount of data points. Furthermore, for fluctuating systems with formation of global structures, we explicitly derive the exact relations among the correlation function, the variance function, and the detrended variance function in arbitrary dimensions. The obtained relations can greatly help further analytical and numerical study of such systems. In addition, by undertaking a detailed analysis of the kernel functions, we rigorously show that the detrended variance functions indeed retain the scaling behaviors due to the stochastic nature of the system and exclude the influence of global structures on the scaling behaviors, verifying the intuitive design of the DFA method. All our results are generally applicable in arbitrary dimensions.

[1] J. Krug, Adv. Phys. **46**, 139 (1997).
[2] T. Halpin-Healy and Y.-C. Zhang, Phys. Rep. **254**, 215 (1995).
[3] M. Saitou, Phys. Rev. B **66**, 073416 (2002).
[4] A. Brú, J. M. Pastor, I. Fernaud, I. Brú, S. Melle, and C. Berenguer, Phys. Rev. Lett. **81**, 4008 (1998).
[5] T. Boyer, S. Levitus, H. Garcia, R. A. Locarnini, C. Stephens, and J. Antonov, Int. J. Climatol. **25**, 931 (2005).
[6] L. Telesca, V. Cuomo, V. Lapenna, and M. Macchiato, Geophys. Res. Lett. **28**, 4323 (2001).
[7] S. Bahar, J. W. Kantelhardt, A. Neiman, H. H. A. Rego, D. F. Russell, L. Wilkens, A. Bunde, and F. Moss, Europhys. Lett. **56**, 454 (2001).
[8] A. Bunde, S. Havlin, J. W. Kantelhardt, T. Penzel, J.-H. Peter, and K. Voigt, Phys. Rev. Lett. **85**, 3736 (2000).
[9] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes* (Cambridge University Press, New York, 1992).
[10] C. Radhakrishna Rao, *Linear Statistical Inference and Its Applications* (Wiley, New York, 1973).
[11] G. E. Forsythe, M. A. Malcolm, and C. B. Moler, *Computer Methods for Mathematical Computations* (Prentice-Hall, Englewood Cliffs, NJ, 1977).
[12] C.-K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger, Phys. Rev. E **49**, 1685 (1994).
[13] S. V. Buldyrev, A. L. Goldberger, S. Havlin, R. N. Mantegna, M. E. Matsa, C.-K. Peng, M. Simons, and H. E. Stanley, Phys. Rev. E **51**, 5084 (1995).
[14] K. Hu, P. C. Ivanov, Z. Chen, P. Carpena, and H. E. Stanley, Phys. Rev. E **64**, 011114 (2001).
[15] A. Cuyt, R. B. Lenin, S. Becuwe, and B. Verdonk, IEEE Trans. Microwave Theory Tech. **54**, 2265 (2006).
[16] J. Spanier and K. B. Oldham, *An Atlas of Functions* (Springer-Verlag, Berlin, 1987); A. P. Prudnikov, Y. A. Brychkov, and O. I. Marichev, *Integrals and Series* (Gordon and Breach, New York, 1986).
[17] W. W. Mullins, J. Appl. Phys. **28**, 333 (1957).
[18] H.-N. Yang, Y.-P. Zhao, G.-C. Wang, and T.-M. Lu, Phys. Rev. Lett. **76**, 3774 (1996).
[19] J.-H. Jeffries, J. K. Zuo, and M. M. Craig, Phys. Rev. Lett. **76**, 4931 (1996).
[20] H.-N. Yang, G.-C. Wang, and T.-M. Lu, Phys. Rev. Lett. **73**, 2348 (1994); Phys. Rev. B **50**, 7635 (1994).
[21] B. Röseler and C. A. Schiller, Mater. Corros. **52**, 413 (2001).
[22] W. Hachem, F. Desbouvries, and P. Loubaton, IEEE Trans. Signal Process. **50**, 651 (2002).
[23] The website for the National Oceanographic Data Center under the NOAA Satellite and Information Service is http://www.nodc.noaa.gov/
[24] A. Carbone, Phys. Rev. E **76**, 056703 (2007).