Fluctuation-response relations for multitime correlations

Gregory L. Eyink
Department of Mathematics, University of Arizona, Tucson, Arizona 85721
(Received 14 May 1999)

We show that time-correlation functions of arbitrary order for any random variable in a statistical dynamical system can be calculated as higher-order response functions of the mean history of the variable. The response is to a ''control term'' added as a modification to the master equation for statistical distributions. The proof of the relations is based upon a variational characterization of the generating functional of the time correlations. The same fluctuation-response relations are preserved within moment closures for the statistical dynamical system, when these are constructed via the variational Rayleigh-Ritz procedure. For the two-time correlations of the moment variables themselves, the fluctuation-response relation is equivalent to an ''Onsager regression hypothesis'' for the small fluctuations. For correlations of higher order, there is a further effect in addition to such linear propagation of fluctuations present instantaneously: the dynamical generation of correlations by nonlinear interaction of fluctuations. In general, we discuss some physical and mathematical aspects of the *Ansätze* required for an accurate calculation of the time correlations. We also comment briefly upon the computational use of these relations, which is well suited for automatic differentiation tools. An example will be given of a simple closure for turbulent energy decay, which illustrates the numerical application of the relations.

PACS number(s): 05.10.−a, 05.45.−a, 05.20.Gg, 05.40.−a

I. INTRODUCTION

It is well known that, in statistical equilibrium systems, there are very useful relations between two-time correlation functions and mean response functions [1,2]. The best-known form of this relation gives the two-time correlation function in terms of a response function of the solution of the microscopic equation of motion to an imposed infinitesimal perturbation, when the response is averaged over the equilibrium ensemble. These relations are often called ''fluctuation-dissipation relations'' but we prefer the term *fluctuation-response relation* (FRR) as being more descriptive. A similar relation has been shown to hold arbitrarily far from thermodynamic equilibrium in stochastic dynamical systems described by nonlinear Langevin equations [3]. In this case, however, the response is to a forcing term added into the Fokker-Planck equation rather than to the dynamical equation for individual realizations. The validity of this form of the theorem depends upon a correct coupling of the force, which, unfortunately, requires a knowledge of the steady-state invariant measure. This latter fact makes the generalized theorem quite difficult to apply in practice.

It is the purpose of this work to prove a far-reaching generalization of the fluctuation-response relation. Our version of the theorem holds for any (time-dependent) Markov process described by a master equation for the distribution function in phase space:

$$\partial_t \mathcal{P}(\mathbf{x},t) = \hat{L}(t)\mathcal{P}(\mathbf{x},t). \tag{1.1}$$

We include in our discussion the limiting case of the Liouville equation for a deterministic dynamical system. Our theorem is more similar to that in [3], since it considers the response to a driving or ''control'' term added into the master equation (1.1) rather than to the equation for individual realizations. However, in contrast to that result, the coupling of our control field does not require any knowledge of the steady-state measure and is quite easy to write down explicitly. Most importantly, all multitime correlations of any finite order are obtained as higher-order response functions to the same control field. Furthermore, the statistics of the system need not be those of thermal equilibrium or even stationary in time. The proof of the relations is based upon a variational characterization of the generating functional for the time-correlation functions, which was established in previous work [4]. Here we shall give a reformulation of that result which is of interest in its own right, as it considerably simplifies and streamlines the analysis in the old work.

The FRR we derive is, however, prohibitively difficult to apply when Eq. (1.1) describes a spatially extended system with many degrees of freedom. In such cases the master equation is a partial differential equation (PDE) in a huge number of variables, far too many to permit a direct numerical solution. The practical use of the FRR in this context will depend upon the employment of moment-closure approximations. As we shall show, the FRR remains valid within the moment closures when these are formulated variationally via the Rayleigh-Ritz method proposed in [4]. We shall review here the Rayleigh-Ritz approximation, providing some new derivations of the old results in addition to establishing the FRR's for the moment closures.

The plan of this paper is as follows. In Sec. II we discuss the variational approach to statistical dynamics. The treatment here will be different in several points and provides significant simplifications of that in [4]. We then employ the variational apparatus to establish the FRR's. In Sec. III we review the Rayleigh-Ritz formulation of moment closure and the FRR's in that approximation. We also discuss there the physical significance of the FRR's, relating that for the two-time correlation to a linear ''regression hypothesis.'' The closure FRR for the $(n+1)$-time correlations, with $n>1$, contains also an effect from the nonlinear terms in the closure, namely, the creation of correlations by interaction of fluctuations. Finally, some physical and mathematical properties required of the probability density function (PDF) models employed for moment closure will be discussed. Section IV concerns numerical aspects, in particular, computationally efficient and accurate methods for computing the derivatives required in the FRR's. A simple example of a turbulence closure will be used to illustrate the numerical issues. The last section, Sec. V, contains our conclusions from this work. An Appendix is also included which summarizes the results of the closure FRR diagrammatically in terms of Feynman-type graphs with propagators and vertices generated from the closure.

II. VARIATIONAL FORMULATION AND FLUCTUATION-RESPONSE RELATIONS

A. Variational approach to statistical dynamics

Suppose that $\mathbf{X}(t)$ is a (vector-valued) Markov process, whose distribution $\mathcal{P}(\mathbf{x},t)$ at time t is governed by the forward Kolmogorov equation or master equation

$$\partial_t \mathcal{P}(\mathbf{x},t) = \hat{L}(t)\mathcal{P}(\mathbf{x},t), \tag{2.1}$$

with $\hat{L}(t)$ the instantaneous Markov generator. The random process governed by a stochastic differential equation is a particular example, for which the generator is the Fokker-Planck operator. This includes the degenerate case of a deterministic dynamics, for which the generator is the Liouville operator. Observables, or random variables, $\mathcal{A}(\mathbf{x},t)$ evolve under the corresponding backward Kolmogorov equation

$$\partial_t \mathcal{A}(\mathbf{x},t) = -\hat{L}^*(t)\mathcal{A}(\mathbf{x},t), \tag{2.2}$$

in which $\hat{L}^*(t)$ is the adjoint operator of $\hat{L}(t)$ with respect to the canonical bilinear form on $L^\infty \times L^1$, i.e., $\langle \mathcal{A},\mathcal{P}\rangle := \int d\mathbf{x}\, \mathcal{A}(x)\mathcal{P}(\mathbf{x})$. The backward and forward Kolmogorov equations may be simultaneously obtained as Euler-Lagrange equations for stationarity of the action functional

$$\Gamma[\mathcal{A},\mathcal{P}] := \int_{t_i}^{t_f} dt\ \ \langle \mathcal{A}(t), [\partial_t - \hat{L}(t)]\mathcal{P}(t)\rangle \tag{2.3}$$

when varied over $\mathcal{P}(t)\in L^1$ with initial condition $\mathcal{P}(t_i) = \mathcal{P}_0$ and $\mathcal{A}(t)\in L^\infty$ with final condition $\mathcal{A}(t_f)\equiv 1$. For details, see [4].

Let $\mathbf{Z}(t) := \boldsymbol{\mathcal{Z}}(\mathbf{X}(t),t)$ be a random variable for the system given by the continuous function $\boldsymbol{\mathcal{Z}}(\mathbf{x},t)$. Then, the cumulant generating functional $W_Z[\mathbf{h}]$ is defined as

$$W_Z[\mathbf{h}] = \ln\left\langle \exp\left(\int_{t_i}^{t_f} dt\, \mathbf{h}^\mathsf{T}(t)\mathbf{Z}(t)\right)\right\rangle. \tag{2.4}$$

The nth-order multitime cumulants of $\mathbf{Z}(t)$ are obtained from $W_Z[\mathbf{h}]$ by functional differentiation with respect to the "test history" $\mathbf{h}(t)$:

$$C_{i_1\dots i_n}(t_1,\dots,t_n) = \left.\frac{\delta^n W_Z[\mathbf{h}]}{\delta h_{i_1}(t_1)\cdots \delta h_{i_n}(t_n)}\right|_{\mathbf{h}=\mathbf{0}}. \tag{2.5}$$

It is not hard to check from its definition (2.4) that $W_Z[\mathbf{h}]$ is a convex functional of $\mathbf{h}$. The Legendre dual of this functional is defined to be the effective action of $\mathbf{Z}(t)$:

$$\Gamma_Z[\mathbf{z}] = \sup_{\mathbf{h}}\{\langle \mathbf{h},\mathbf{z}\rangle - W_Z[\mathbf{h}]\}, \tag{2.6}$$

with $\langle \mathbf{h},\mathbf{z}\rangle := \int dt\, \mathbf{h}^\mathsf{T}(t)\mathbf{z}(t)$. It is a generating functional of so-called irreducible correlation functions of $\mathbf{Z}(t)$:

$$\Gamma_{i_1\dots i_n}(t_1,\dots,t_n) = \left.\frac{\delta^n \Gamma_Z[\mathbf{z}]}{\delta z_{i_1}(t_1)\cdots \delta z_{i_n}(t_n)}\right|_{\mathbf{z}=\overline{\mathbf{z}}}. \tag{2.7}$$

The functional derivatives here are evaluated at the mean history $\overline{\mathbf{z}(t)} := \langle \mathbf{Z}(t)\rangle$. It is not hard to check from the definition (2.6) that $\Gamma_Z[\mathbf{z}]$ is a convex, non-negative functional with a unique global minimum (equal to zero) at the mean history $\mathbf{z}=\overline{\mathbf{z}}$.

There is a useful characterization of the effective action $\Gamma_Z[\mathbf{z}]$ by means of a constrained variation of the action $\Gamma[\mathcal{A},\mathcal{P}]$, which was established in [4]. In fact,

$$\Gamma_Z[\mathbf{z}] = (\text{st.pt.})_{\mathcal{A},\mathcal{P}}\Gamma[\mathcal{A},\mathcal{P}] \tag{2.8}$$

[where (st.pt.) is the stationary point], when varied over the same classes as above, but subject to constraints of fixed overlap

$$\langle \mathcal{A}(t),\mathcal{P}(t)\rangle = 1 \tag{2.9}$$

and fixed expectation

$$\langle \mathcal{A}(t), \hat{\boldsymbol{\mathcal{Z}}}(t)\mathcal{P}(t)\rangle = \mathbf{z}(t) \tag{2.10}$$

for all $t\in[t_i,t_f]$. Note that $\hat{\boldsymbol{\mathcal{Z}}}(t)$ is used to denote the operator (in both L^1 and L^∞) of multiplication by $\boldsymbol{\mathcal{Z}}(\mathbf{x},t)$. The Euler-Lagrange equations for this constrained variation may be obtained by incorporating the expectation constraint (2.10) with a Lagrange multiplier $\mathbf{h}(t)$. The overlap constraint may also be imposed with a Lagrange multiplier $\lambda(t)$, as it was in [4].

However, it turns out to be advantageous to impose Eq. (2.9) through the definitions

$$\mathcal{A}(t) := 1 + [\mathcal{B}(t) - \langle \mathcal{B}(t)\rangle_t] := 1 + \mathcal{C}(t), \tag{2.11}$$

with the final conditions $\mathcal{B}(t_f) = \mathcal{C}(t_f) \equiv 0$. Note that $\langle \mathcal{B}(t)\rangle_t := \langle \mathcal{B}(t),\mathcal{P}(t)\rangle$ is the expectation with respect to the distribution $\mathcal{P}(t)$. Hence, the overlap constraint (2.9) is satisfied when $\mathcal{B}(t)$ is varied independently of $\mathcal{P}(t)$. The variable $\mathcal{C}(t)$ is no longer independent of $\mathcal{P}(t)$, but must satisfy the orthogonality condition $\langle \mathcal{C}(t),\mathcal{P}(t)\rangle = 0$. The expectation constraint must still be implemented by the Lagrange multiplier $\mathbf{h}(t)$. In terms of $\mathcal{B}(t)$ or $\mathcal{C}(t)$ the latter constraint is

$$\begin{aligned}\mathbf{z}(t) &= \langle \boldsymbol{\mathcal{Z}}(t)\rangle_t + [\langle \boldsymbol{\mathcal{Z}}(t)\mathcal{B}(t)\rangle_t - \langle \boldsymbol{\mathcal{Z}}(t)\rangle_t \langle \mathcal{B}(t)\rangle_t] \\ &= \langle \boldsymbol{\mathcal{Z}}(t)\rangle_t + \langle \boldsymbol{\mathcal{Z}}(t)\mathcal{C}(t)\rangle_t.\end{aligned} \tag{2.12}$$

The Euler-Lagrange equations are obtained by varying the action $\Gamma[\mathcal{A},\mathcal{P}]$ over $\mathcal{B}(t),\mathcal{P}(t)$ with $\mathcal{A}(t) = 1 + [\mathcal{B}(t) - \langle \mathcal{B}(t)\rangle_t]$, incorporating the constraint (2.12) with the multiplier $\mathbf{h}(t)$. A straightforward calculation gives

$$\partial_t \mathcal{P}(t) = \hat{L}(t)\mathcal{P}(t) + \mathbf{h}^{\mathsf{T}}(t)[\boldsymbol{\mathcal{Z}}(t) - \langle \boldsymbol{\mathcal{Z}}(t)\rangle_t]\mathcal{P}(t) \tag{2.13}$$

and

$$\partial_t \mathcal{B}(t) + \hat{L}^*(t)\mathcal{B}(t) + \mathbf{h}^{\mathsf{T}}(t)[\boldsymbol{\mathcal{Z}}(t)\mathcal{B}(t) - \langle \boldsymbol{\mathcal{Z}}(t)\rangle_t \mathcal{B}(t) - \langle \mathcal{B}(t)\rangle_t \boldsymbol{\mathcal{Z}}(t)] + \mathbf{h}^{\mathsf{T}}(t)\boldsymbol{\mathcal{Z}}(t) = 0. \tag{2.14}$$

Let us introduce the new operator

$$\hat{L}_{\mathbf{h}}(t) := \hat{L}(t) + \mathbf{h}^{\mathsf{T}}(t)[\hat{\boldsymbol{\mathcal{Z}}}(t) - \langle \boldsymbol{\mathcal{Z}}(t)\rangle_t]. \tag{2.15}$$

The variational equations are written in terms of this operator as

$$\partial_t \mathcal{P}(t) = \hat{L}_{\mathbf{h}}(t)\mathcal{P}(t) \tag{2.16}$$

and

$$\partial_t \mathcal{B}(t) + \hat{L}^*_{\mathbf{h}}(t)\mathcal{B}(t) + \mathbf{h}^{\mathsf{T}}(t)\boldsymbol{\mathcal{Z}}(t)[1 - \langle \mathcal{B}(t)\rangle_t] = 0. \tag{2.17}$$

Using the resulting identity $(d/dt)\langle \mathcal{B}(t)\rangle_t = -\mathbf{h}^{\mathsf{T}}(t)[1 - \langle \mathcal{B}(t)\rangle_t]\langle \boldsymbol{\mathcal{Z}}(t)\rangle_t$, the last equation can be rewritten in terms of $\mathcal{C}(t) = \mathcal{B}(t) - \langle \mathcal{B}(t)\rangle_t$ as

$$\partial_t \mathcal{C}(t) + \hat{L}^*_{\mathbf{h}}(t)\mathcal{C}(t) + \mathbf{h}^{\mathsf{T}}(t)[\boldsymbol{\mathcal{Z}}(t) - \langle \boldsymbol{\mathcal{Z}}(t)\rangle_t] = 0. \tag{2.18}$$

The action functional may be expressed in terms of $\mathcal{C}$, $\mathcal{P}$ as

$$\Gamma[\mathcal{C},\mathcal{P}] = \int_{t_i}^{t_f} dt\, \langle \mathcal{C}(t), [\partial_t - \hat{L}(t)]\mathcal{P}(t)\rangle, \tag{2.19}$$

using $\hat{L}^*(t)1 = 0$. The effective action $\Gamma_Z[\mathbf{z}]$ is then obtained by substituting the solutions of Eqs. (2.16) and (2.18), when the ''control field'' $\mathbf{h}(t)$ is chosen so that Eq. (2.12) reproduces the considered history $\mathbf{z}(t)$. The quantity $\mathbf{z}(t)$ can be seen to be ''controllable'' by $\mathbf{h}(t)$ from Legendre duality. That is, the control $\mathbf{h}[t;\mathbf{z}]$ for a specified $\mathbf{z}(t)$ is obtained from the minimization of the convex function $W_Z[\mathbf{h}] - \langle \mathbf{h}, \mathbf{z}\rangle$ [compare Eq. (2.6)]. Gathering together all of our previous discussion we may state the following proposition.

Proposition. The effective action of the variable $\mathbf{Z}(t)$ is obtained as

$$\Gamma_Z[\mathbf{z}] = \int_{t_i}^{t_f} dt\, \langle \mathcal{C}(t), [\partial_t - \hat{L}(t)]\mathcal{P}(t)\rangle, \tag{2.20}$$

where $\mathcal{C}$, $\mathcal{P}$ satisfy

$$\partial_t \mathcal{P}(t) = \hat{L}_{\mathbf{h}}(t)\mathcal{P}(t) \tag{2.21}$$

and

$$\partial_t \mathcal{C}(t) + \hat{L}^*_{\mathbf{h}}(t)\mathcal{C}(t) + \mathbf{h}^{\mathsf{T}}(t)[\boldsymbol{\mathcal{Z}}(t) - \langle \boldsymbol{\mathcal{Z}}(t)\rangle_t] = 0 \tag{2.22}$$

with initial and final conditions

$$\mathcal{P}(t_i) = \mathcal{P}_0, \quad \mathcal{C}(t_f) = 0, \tag{2.23}$$

and the value of the control field $\mathbf{h}$ is selected to give for all $t \in [t_i, t_f]$

$$\langle \boldsymbol{\mathcal{Z}}(t)\rangle_t + \langle \boldsymbol{\mathcal{Z}}(t)\mathcal{C}(t)\rangle_t = \mathbf{z}(t). \tag{2.24}$$

B. Fluctuation-response relations

It is not accidental that the same notation $\mathbf{h}(t)$ was chosen above for the control field as for the argument of the cumulant-generating functional $W_Z[\mathbf{h}]$. In fact, we shall prove that

$$W_Z[\mathbf{h}] = \int_{t_i}^{t_f} dt\, \mathbf{h}^{\mathsf{T}}(t)\langle \boldsymbol{\mathcal{Z}}(t)\rangle_t, \tag{2.25}$$

using just the solution $\mathcal{P}(t;\mathbf{h})$ of the forward equation (2.21), for the control history $\mathbf{h}(t)$ which appears as the argument of W_Z. The result is obtained by simply substituting the constraint (2.24) into the inverse Legendre transform for W_Z:

$$\begin{aligned} W_Z[\mathbf{h}] &= \int_{t_i}^{t_f} dt\, \mathbf{h}^{\mathsf{T}}(t)\mathbf{z}(t) - \Gamma_Z[\mathbf{z}] \\ &= \int_{t_i}^{t_f} dt \{\mathbf{h}^{\mathsf{T}}(t)[\langle \boldsymbol{\mathcal{Z}}(t)\rangle_t + \langle \boldsymbol{\mathcal{Z}}(t)\mathcal{C}(t)\rangle_t] \\ &\quad - \langle \mathcal{C}(t), (\partial_t - \hat{L})\mathcal{P}(t)\rangle\} \\ &= \int_{t_i}^{t_f} dt [\mathbf{h}^{\mathsf{T}}(t)\langle \boldsymbol{\mathcal{Z}}(t)\rangle_t - \langle \mathcal{C}(t), (\partial_t - \hat{L}_{\mathbf{h}})\mathcal{P}(t)\rangle] \\ &= \int_{t_i}^{t_f} dt\, \mathbf{h}^{\mathsf{T}}(t)\langle \boldsymbol{\mathcal{Z}}(t)\rangle_t. \end{aligned} \tag{2.26}$$

The second term in the third line vanishes by Eq. (2.21).

The relation (2.25) is a compact presentation of the fluctuation-response relations. Let us define the response functional of order n for the variable $\mathbf{Z}(t)$ as

$$R_{i;i_1\cdots i_n}[t;t_1,\ldots,t_n;\mathbf{h}] := \frac{\delta^n \langle \mathcal{Z}_i(t)\rangle_t}{\delta h_{i_1}(t_1)\cdots \delta h_{i_n}(t_n)}[\mathbf{h}], \tag{2.27}$$

where $\langle \boldsymbol{\mathcal{Z}}(t)\rangle_t$ denotes as before the average with respect to the solution $\mathcal{P}(t;\mathbf{h})$ of Eq. (2.21). This functional is causal, i.e., it vanishes if $t < t_i$ for any $i = 1,\ldots,n$, because the average $\langle \boldsymbol{\mathcal{Z}}(t)\rangle_t$ cannot have any dependence upon $\mathbf{h}(t')$ for $t' > t$. The nth-order response function is taken to be the value at $\mathbf{h} = \mathbf{0}$:

$$R_{i;i_1\cdots i_n}(t;t_1\ldots,t_n) := R_{i;i_1\cdots i_n}[t;t_1,\ldots,t_n;\mathbf{h}]|_{\mathbf{h}=\mathbf{0}}. \tag{2.28}$$

We now state the fluctuation-response relations.

Proposition. The $(n+1)$st-order cumulant $C_{i_1\cdots i_{n+1}}(t_1,\ldots,t_{n+1})$ is determined for each integer $n \geqslant 0$ by

$$C_{i_1\cdots i_{n+1}}(t_1,\ldots,t_{n+1}) = \sum_{k=1}^{n+1} R_{i_k;i_1\cdots\widehat{i_k}\cdots i_{n+1}}(t_k;t_1,\ldots,\widehat{t_k},\ldots,t_{n+1}), \tag{2.29}$$

where the caret ''^'' denotes omission of the corresponding expression. Observe that only the one term with $t_k = \max_p t_p$ is actually nonzero in the sum, by causality of the response function.

The proof of the relations is very simple. We recall the formula (2.5) for the cumulant in terms of the generating functional W_Z and the corresponding definition (2.27) of the response functionals. Taking $n+1$ functional derivatives of W_Z in Eq. (2.25), one obtains

$$\frac{\delta^{n+1}W_Z[\mathbf{h}]}{\delta h_{i_1}(t_1)\cdots\delta h_{i_{n+1}}(t_{n+1})} = \sum_{k=1}^{n+1} R_{i_k;i_1\cdots\widehat{i_k}\cdots i_{n+1}}[t_k;t_1,\ldots,\widehat{t_k},\ldots,t_{n+1};\mathbf{h}] + \sum_j \int_{t_i}^{t_f} dt\, h_j(t) R_{j;i_1\cdots i_{n+1}}[t;t_1,\ldots,t_{n+1};\mathbf{h}]. \tag{2.30}$$

This formula is easy to prove by induction upon *n*. Setting $\mathbf{h}=\mathbf{0}$, the last integral term vanishes and one obtains Eq. (2.29).

It is important to point out that the operator $\hat{L}_{\mathbf{h}}$ in Eq. (2.21) has a quite simple and explicit dependence upon the control field $\mathbf{h}(t)$, given in Eq. (2.15). While simple, the coupling does depend upon the solution $\mathcal{P}(t)$ itself, through the average $\langle \boldsymbol{\mathcal{Z}}(t)\rangle_t$. Hence, Eq. (2.21) for $\mathcal{P}(t)$ is actually quadratically nonlinear, unlike the original master equation. Nevertheless, this nonlinearity is exactly that required to preserve the normalization of the solution. The equation may be integrated forward in time to obtain simultaneously $\mathcal{P}(t)$ and $\langle \boldsymbol{\mathcal{Z}}(t)\rangle_t$ as functionals of the control $\mathbf{h}$. The response functions can then be determined by differentiating the results.

III. MOMENT-CLOSURE APPROXIMATIONS

A. Variational formulation of moment closure

The results of the previous section provide a general approach to computation of the multitime cumulants. However, it is obvious that the required integration of the modified master equation (2.21) will be possible only for the simplest of models, with a few degrees of freedom. For spatially extended systems with many degrees of freedom, this integration is totally intractable. The practical employment of the FRR's then depends upon making moment-closure approximations. We shall review here the variational formulation of moment-closure approximation, following essentially the treatment in [4]. However, we shall also introduce some important simplifications, which we comment upon as we proceed.

Moment-closure approximations to the generating functional $\Gamma_Z[\mathbf{z}]$ of irreducible multitime correlations of $\mathbf{Z}(t)$ are obtained by means of the characterization of that functional through the constrained variation in Eq. (2.8). Rather than varying over all $\mathcal{A}\in L^\infty$, $\mathcal{P}\in L^1$, one varies only over finitely parametrized trial functions. The trial functions are constructed from the usual elements of a moment closure: a set of moment functions $M_i(\mathbf{x},t)$, $i=1,...,R$, and a PDF *Ansatz* $\mathcal{P}(\mathbf{x},t;\boldsymbol{\mu})$, which is conveniently parametrized by the mean values that it attributes to the moment functions $\boldsymbol{\mu} \coloneqq \int d\mathbf{x}\,\mathcal{P}(\mathbf{x},t;\boldsymbol{\mu})\mathbf{M}(\mathbf{x},t)$. The left trial function is then taken to be

$$\mathcal{B}(\mathbf{x},t;\boldsymbol{\alpha}) \coloneqq \sum_{i=1}^{R} \alpha_i M_i(\mathbf{x},t). \tag{3.1}$$

Following the discussion in Sec. II A, we have chosen the left trial state in the form (2.11), to incorporate automatically the overlap constraint (2.9). The histories $\boldsymbol{\alpha}(t)$ and $\boldsymbol{\mu}(t)$ are the parameters to be varied over. Substituting the trial forms, one obtains the reduced action

$$\Gamma[\boldsymbol{\alpha},\boldsymbol{\mu}] = \int_{t_i}^{t_f} dt\, \boldsymbol{\alpha}^{\top}(t)[\dot{\boldsymbol{\mu}}(t) - \mathbf{V}(\boldsymbol{\mu}(t),t)] \tag{3.2}$$

with

$$\mathbf{V}(\boldsymbol{\mu},t) \coloneqq \langle(\partial_t + \hat{L}^*)\mathbf{M}(t)\rangle_{\boldsymbol{\mu}}. \tag{3.3}$$

Of course, $\langle\cdot\rangle_{\boldsymbol{\mu}}$ denotes the average with respect to the PDF *Ansatz*. An unconstrained variation of Eq. (3.2) recovers the standard moment-closure equation $\dot{\boldsymbol{\mu}} = \mathbf{V}(\boldsymbol{\mu},t)$.

For the calculation of the action, however, there is the additional expectation constraint (2.10). In terms of the trial parameters, it becomes

$$\mathbf{z}(t) = \boldsymbol{\zeta}(\boldsymbol{\mu}(t),t) + \mathbf{C}_Z(\boldsymbol{\mu}(t),t)\boldsymbol{\alpha}(t). \tag{3.4}$$

Here,

$$\boldsymbol{\zeta}(\boldsymbol{\mu},t) \coloneqq \langle \mathbf{Z}(t)\rangle_{\boldsymbol{\mu}} \tag{3.5}$$

is the *Z* expectation within the PDF *Ansatz* and

$$\mathbf{C}_Z(\boldsymbol{\mu},t) \coloneqq \langle \mathbf{Z}(t)\mathbf{M}^{\top}(t)\rangle_{\boldsymbol{\mu}} - \boldsymbol{\zeta}(\boldsymbol{\mu},t)\boldsymbol{\mu}^{\top} \tag{3.6}$$

is the corresponding *ZM* covariance matrix. It is remarkable that $\boldsymbol{\zeta}(\boldsymbol{\mu},t)$ and $\mathbf{C}_Z(\boldsymbol{\mu},t)$ are the only inputs of the PDF *Ansatz* actually required for the calculation. When the constraint (3.4) is incorporated into the action functional (3.2) by means of a Lagrange multiplier $\mathbf{h}(t)$, the resulting Euler-Lagrange equations are

$$\dot{\boldsymbol{\mu}} = \mathbf{V}(\boldsymbol{\mu},t) + \mathbf{C}_Z^{\top}(\boldsymbol{\mu},t)\mathbf{h}(t) \coloneqq \mathbf{V}_Z(\boldsymbol{\mu},\mathbf{h},t) \tag{3.7}$$

and

$$\dot{\boldsymbol{\alpha}} + \left(\frac{\partial \mathbf{V}_Z}{\partial \boldsymbol{\mu}}\right)^{\top}(\boldsymbol{\mu},\mathbf{h},t)\boldsymbol{\alpha} + \left(\frac{\partial \boldsymbol{\zeta}}{\partial \boldsymbol{\mu}}\right)^{\top}(\boldsymbol{\mu},t)\mathbf{h}(t) = \mathbf{0}. \tag{3.8}$$

These are solved subject to an initial condition $\boldsymbol{\mu}(t_i) = \boldsymbol{\mu}_0$ and a final condition $\boldsymbol{\alpha}(t_f) = \mathbf{0}$. When the solutions of the integrations are substituted into Eq. (3.2), there results a Rayleigh-Ritz approximation $\tilde{\Gamma}_Z[\mathbf{z}]$ to the effective action of

$\mathbf{Z}(t)$. The value $\mathbf{z}(t)$ of the argument is that given by the constraint equation (3.4) for the given value of the control field $\mathbf{h}(t)$.

The above derivation of the moment-closure approximation $\tilde{\Gamma}_Z[\mathbf{z}]$ is equivalent to that in [4] but differs in some important details. The trial states employed in [4] each contained an additional parameter, μ_0 and α_0, with $\overline{\boldsymbol{\mu}} = \mu_0(1,\boldsymbol{\mu})$, $\overline{\boldsymbol{\alpha}} = (\alpha_0, \boldsymbol{\alpha})$. Thus, the trial states employed there may be written, in our present notation as

$$\mathcal{P}(\mathbf{x},t;\overline{\boldsymbol{\mu}}) := \mu_0 \mathcal{P}(\mathbf{x},t;\boldsymbol{\mu}) \tag{3.9}$$

and

$$\mathcal{A}(\mathbf{x},t;\overline{\boldsymbol{\alpha}}) := \sum_{i=0}^{R} \alpha_i M_i(\mathbf{x},t). \tag{3.10}$$

Thus, μ_0 was an arbitrary normalization factor and α_0 was the coefficient of the constant moment function $M_0(\mathbf{x},t) \equiv 1$. With this pair of trial functions, the overlap constraint (2.9) was no longer automatically enforced and needed to be incorporated via a Lagrange multiplier $\lambda(t)$. The resulting Euler-Lagrange equations of the constrained variation for $\overline{\boldsymbol{\mu}}(t), \overline{\boldsymbol{\alpha}}(t)$ then involved both multipliers $\mathbf{h}(t)$ and $\lambda(t)$. See Eqs. (3.93)–(3.95) in [4]. Nevertheless, those equations are equivalent to Eqs. (3.7) and (3.8) above. We shall not give all the details here, but leave it as a relatively simple exercise for the reader to check. We only point out that one may always take $\mu_0 \equiv 1$, without any loss of generality, by absorbing that factor into the coefficients $\boldsymbol{\alpha}$ of the left trial function. It then follows from the 0 component of Eq. (3.93) in [4] that the Lagrange multiplier for the overlap constraint is given explicitly as

$$\lambda(t) = \overline{V}_0(\boldsymbol{\mu},\mathbf{h},t) = \mathbf{h}^{\mathsf{T}}(t)\boldsymbol{\zeta}(\boldsymbol{\mu},t). \tag{3.11}$$

If one uses this result and also uses the constraint equation (3.95) in [4] to eliminate the variable α_0 from the equations, then one derives from Eqs. (3.93)–(3.95) in [4] identically the same equations as Eqs. (3.7) and (3.8) above.

Despite their equivalence to the variational equations in [4], the form in Eqs. (3.7) and (3.8) above is far more convenient. Because of the presence of the multiplier $\lambda(t)$ in both the forward and backward equations (3.93) and (3.94) in [4], those equations posed—apparently—a true initial-final value problem. It was proposed in [4] to solve that boundary-value problem in time with a relaxation method. However, we now see that the forward equation (3.7) is completely uncoupled from the backward equation. It may be integrated forward in time, storing the solution $\boldsymbol{\mu}(t)$ for subsequent input into the backward equation (3.8) for $\boldsymbol{\alpha}(t)$. Efficient numerical algorithms for doing so and then integrating the results to calculate the approximate action have been developed by us and will be discussed in another work.

B. Fluctuation-response relations in closures

A Rayleigh-Ritz approximation $\tilde{W}_Z[\mathbf{h}]$ to the cumulant-generating functional may be introduced by the formal relation

$$\tilde{W}_Z[\mathbf{h}] + \tilde{\Gamma}_Z[\mathbf{z}] = \langle \mathbf{h}, \mathbf{z} \rangle. \tag{3.12}$$

It may be easily checked that this definition is equivalent to the formula

$$\tilde{W}_Z[\mathbf{h}] = \int_{t_i}^{t_f} dt\, \mathbf{h}^{\mathsf{T}}(t)\boldsymbol{\zeta}(\boldsymbol{\mu}(t),t), \tag{3.13}$$

in which $\boldsymbol{\mu}(t)$ is the solution of the forward equation (3.7) for the control history $\mathbf{h}(t)$. The derivation is the exact analog of that of Eq. (2.26). Indeed, it follows that

$$\begin{aligned}
\tilde{W}_Z[\mathbf{h}] &= \int_{t_i}^{t_f} dt\, \mathbf{h}^{\mathsf{T}}(t)\mathbf{z}(t) - \tilde{\Gamma}_Z[\mathbf{z}] \\
&= \int_{t_i}^{t_f} dt \{\mathbf{h}^{\mathsf{T}}(t)[\boldsymbol{\zeta}(\boldsymbol{\mu}(t),t) \\
&\quad + \mathbf{C}_Z(\boldsymbol{\mu}(t),t)\boldsymbol{\alpha}(t)] - \boldsymbol{\alpha}^{\mathsf{T}}(t)[\dot{\boldsymbol{\mu}}(t) - \mathbf{V}(\boldsymbol{\mu}(t),t)]\} \\
&= \int_{t_i}^{t_f} dt\, \{\mathbf{h}^{\mathsf{T}}(t)\zeta(\boldsymbol{\mu}(t),t) \\
&\quad - \boldsymbol{\alpha}^{\mathsf{T}}(t)[\dot{\boldsymbol{\mu}}(t) - \mathbf{V}_Z(\boldsymbol{\mu}(t),\mathbf{h}(t),t)]\} \\
&= \int_{t_i}^{t_f} dt\, \mathbf{h}^{\mathsf{T}}(t)\boldsymbol{\zeta}(\boldsymbol{\mu}(t),t),
\end{aligned} \tag{3.14}$$

where Eq. (3.7) was used to eliminate the second term of the third line.

It is a very attractive feature of the Rayleigh-Ritz approximation scheme that the resulting functionals $\tilde{\Gamma}_Z[\mathbf{z}]$ and $\tilde{W}_Z[\mathbf{h}]$ remain formal Legendre transforms of each other. That is,

$$\mathbf{h}[t;\mathbf{z}] = \frac{\delta \tilde{\Gamma}_Z}{\delta \mathbf{z}(t)}[\mathbf{z}], \quad \mathbf{z}[t;\mathbf{h}] = \frac{\delta \tilde{W}_Z}{\delta \mathbf{h}(t)}[\mathbf{h}]. \tag{3.15}$$

This follows from the discussion in [4], where it was derived by a reduction from the underlying variational formulation of the master equation. It is worthwhile to record here another derivation, which is instead based directly upon the constrained moment equations (3.7) and (3.8). Among other things, this new proof carries over usefully to discrete approximations of the moment equations employed in numerical computations that do not follow directly from an underlying microscopic theory.

The derivation begins by functionally differentiating Eq. (3.13) with respect to $\mathbf{h}(t)$:

$$\begin{aligned}
\frac{\delta \tilde{W}_Z}{\delta \mathbf{h}(t)}[\mathbf{h}] &= \boldsymbol{\zeta}(t) + \int_{t_i}^{t_f} ds \left(\frac{\delta \boldsymbol{\zeta}(s)}{\delta \mathbf{h}(t)}\right)^{\mathsf{T}} \mathbf{h}(s) \\
&= \boldsymbol{\zeta}(t) + \int_{t}^{t_f} ds \left(\frac{\delta \boldsymbol{\mu}(s)}{\delta \mathbf{h}(t)}\right)^{\mathsf{T}} \left(\frac{\partial \boldsymbol{\zeta}}{\partial \boldsymbol{\mu}}(\boldsymbol{\mu},s)\right)^{\mathsf{T}} \mathbf{h}(s).
\end{aligned} \tag{3.16}$$

The abbreviation $\boldsymbol{\zeta}(t) := \boldsymbol{\zeta}(\boldsymbol{\mu}(t),t)$ was introduced and in the second line the chain rule was employed. Causality was also invoked to reset the lower limit of integration. The response matrix $\delta\boldsymbol{\mu}(s)/\delta\mathbf{h}(t)$ that now appears satisfies an equation obtained by functionally differentiating Eq. (3.7) with respect to $\mathbf{h}(t)$:

$$\frac{\delta\dot{\boldsymbol{\mu}}(s)}{\delta\mathbf{h}(t)}=\mathbf{A}(s)\frac{\delta\boldsymbol{\mu}(s)}{\delta\mathbf{h}(t)}+\mathbf{C}_Z^T(\boldsymbol{\mu},s)\,\delta(s-t) \tag{3.17}$$

with the abbreviation

$$\mathbf{A}(s):=\frac{\partial\mathbf{V}_Z}{\partial\boldsymbol{\mu}}(\boldsymbol{\mu}(s),\mathbf{h}(s),s). \tag{3.18}$$

The equation (3.17) can be solved with a Greens function given by a time-ordered exponential:

$$\frac{\delta\boldsymbol{\mu}(s)}{\delta\mathbf{h}(t)}=T\exp\left(\int_t^s\mathbf{A}(r)dr\right)\mathbf{C}_Z^T(\boldsymbol{\mu},t). \tag{3.19}$$

The adjoint of this propagator appears in the solution of the backward equation (3.8). That equation can be written as

$$\dot{\boldsymbol{\alpha}}(t)+\mathbf{A}^{\mathsf{T}}(t)\boldsymbol{\alpha}(t)+\left(\frac{\partial\boldsymbol{\zeta}}{\partial\boldsymbol{\mu}}\right)^{\mathsf{T}}(\boldsymbol{\mu},t)\mathbf{h}(t)=\mathbf{0}, \tag{3.20}$$

and its solution is

$$\boldsymbol{\alpha}(t)=\int_t^{t_f}ds\,\overline{T}\exp\left(\int_t^s\mathbf{A}^{\mathsf{T}}(r)dr\right)\left(\frac{\partial\boldsymbol{\zeta}}{\partial\boldsymbol{\mu}}\right)^{\mathsf{T}}(\boldsymbol{\mu},s)\mathbf{h}(s). \tag{3.21}$$

Here $\overline{T}\exp[\cdot]$ denotes the anti-time-ordered exponential. If the solution (3.19) for the response matrix is substituted into the second term of Eq. (3.16), and then Eq. (3.21) is employed, it follows that

$$\frac{\delta\widetilde{W}_Z}{\delta\mathbf{h}(t)}[\mathbf{h}]=\boldsymbol{\zeta}(t)+\mathbf{C}_Z(t)\boldsymbol{\alpha}(t)=\mathbf{z}(t), \tag{3.22}$$

using Eq. (3.4). Thus, the second relation in Eq. (3.15) is proved. Of course, the dual first relation is obtained by functionally differentiating Eq. (3.12) with respect to $\mathbf{z}(t)$ and using Eq. (3.22).

The Rayleigh-Ritz approximate generating functional $\widetilde{W}_Z[\mathbf{h}]$ given in Eq. (3.13) retains most of the remarkable features of the exact generating functional $W_Z[\mathbf{h}]$. In particular, its value may be obtained by integrating just the forward moment equation (3.7) for the selected control field $\mathbf{h}(t)$, and then substituting the solution $\boldsymbol{\mu}(t;\mathbf{h})$ into Eq. (3.13). Fluctuation-response relations follow for the approximate cumulants in the same way as before. Just as before, one may define the approximate response function of order n for the variable $\mathbf{Z}(t)$ as

$$\widetilde{R}_{i;i_1\cdots i_n}(t;t_1,\ldots,t_n):=\left.\frac{\delta^n\zeta_i(t)}{\delta h_{i_1}(t_1)\cdots\delta h_{i_n}(t_n)}\right|_{\mathbf{h}=\mathbf{0}}. \tag{3.23}$$

This functional is also causal. By the same argument as before, one obtains the fluctuation-response relations for the Rayleigh-Ritz approximate cumulants generated from $\widetilde{W}_Z[\mathbf{h}]$:

$$\widetilde{C}_{i_1\cdots i_{n+1}}(t_1,\ldots,t_{n+1})=\sum_{k=1}^{n+1}\widetilde{R}_{i_k;i_1\cdots\widehat{i_k}\cdots i_{n+1}}(t_k;t_1,\ldots,\widehat{t_k},\ldots,t_{n+1}). \tag{3.24}$$

This FRR within moment closures is a practical way to compute multitime correlations numerically, as we shall see below.

C. Physical interpretation of the closure FRR

The results are easiest to interpret physically in the case $n=1$. In that case, the FRR deals with the second-order cumulant of $\mathbf{Z}(t)$ or the two-time covariance matrix $\mathbf{C}(t,t_0):=\langle\delta\mathbf{Z}(t)\,\delta\mathbf{Z}^{\mathsf{T}}(t_0)\rangle$ [with $\delta\mathbf{Z}(t):=\mathbf{Z}(t)-\langle\mathbf{Z}(t)\rangle$]. The FRR here states that

$$\widetilde{\mathbf{C}}(t,t_0)=\widetilde{\mathbf{R}}(t,t_0)+[\widetilde{\mathbf{R}}(t_0,t)]^{\mathsf{T}}, \tag{3.25}$$

with $\widetilde{\mathbf{R}}(t,t_0):=[\delta\boldsymbol{\zeta}(t)/\delta\mathbf{h}(t_0)]|_{\mathbf{h}=\mathbf{0}}$ the response matrix. Thus, for $t>t_0$,

$$\begin{aligned}\widetilde{\mathbf{C}}(t,t_0)&=\left.\frac{\delta\boldsymbol{\zeta}(t)}{\delta\mathbf{h}(t_0)}\right|_{\mathbf{h}=\mathbf{0}}=\left(\frac{\partial\boldsymbol{\zeta}}{\partial\boldsymbol{\mu}}\right)(\boldsymbol{\mu},t)\left.\frac{\delta\boldsymbol{\mu}(t)}{\delta\mathbf{h}(t_0)}\right|_{\mathbf{h}=\mathbf{0}}\\&=\left(\frac{\partial\boldsymbol{\zeta}}{\partial\boldsymbol{\mu}}\right)(\boldsymbol{\mu},t)T\exp\left(\int_{t_0}^t\mathbf{A}_*(s)ds\right)\\&\quad\times\mathbf{C}_Z^T(\boldsymbol{\mu},t_0)\end{aligned} \tag{3.26}$$

using Eq. (3.19). We set $\mathbf{A}_*(t):=\mathbf{A}(t)|_{\mathbf{h}=\mathbf{0}}$. Recall also that $\mathbf{C}_Z^T(\boldsymbol{\mu},t_0)=\langle\delta\mathbf{M}(t_0)\,\delta\mathbf{Z}^{\mathsf{T}}(t_0)\rangle$.

We see that the same result can be obtained by making two physically motivated approximations. The first is the slaving hypothesis: that fluctuations of the variable $\mathbf{Z}(t)$ are instantaneously slaved to those of the moment variables $\mathbf{M}(t)$, or $\delta\mathbf{Z}(t)\approx(\partial\boldsymbol{\zeta}/\partial\boldsymbol{\mu})(\boldsymbol{\mu},t)\,\delta\mathbf{M}(t)$. Thus,

$$\langle\delta\mathbf{Z}(t)\,\delta\mathbf{Z}^{\mathsf{T}}(t_0)\rangle\approx\left(\frac{\partial\boldsymbol{\zeta}}{\partial\boldsymbol{\mu}}\right)(\boldsymbol{\mu},t)\langle\delta\mathbf{M}(t)\,\delta\mathbf{Z}^{\mathsf{T}}(t_0)\rangle. \tag{3.27}$$

The second approximation is the regression hypothesis: that fluctuations of the moment variables $\mathbf{M}(t)$ decay on average according to the linearized closure equation $\delta\dot{\mathbf{M}}(t)\approx\mathbf{A}_*(t)\,\delta\mathbf{M}(t)$. Thus,

$$\langle\delta\mathbf{M}(t)\,\delta\mathbf{Z}^{\mathsf{T}}(t_0)\rangle\approx T\exp\left(\int_{t_0}^t\mathbf{A}_*(s)ds\right)\langle\delta\mathbf{M}(t_0)\,\delta\mathbf{Z}^{\mathsf{T}}(t_0)\rangle. \tag{3.28}$$

Together, Eqs. (3.27) and (3.28) lead directly back to the result (3.26).

The special case of the FRR in Eq. (3.25) with $n=1$ and with $\mathbf{Z}(t)$ taken to be the moment variable $\mathbf{M}(t)$ itself was previously derived in [5]. It was already pointed out there that the physical interpretation of the approximate FRR was provided by the regression hypothesis. That result has now been generalized to the case where $\mathbf{Z}(t)\neq\mathbf{M}(t)$ and the additional physical principle in the Rayleigh-Ritz approximation is the slaving hypothesis.

Of course, it is also of interest to consider the physical meaning of the approximations involved for $n>1$. For the next case, $n=2$, the object of interest is the third-order cumulant

$$C_{ijk}(t_2,t_1,t_0):=\langle\delta Z_i(t_2)\delta Z_j(t_1)\delta Z_k(t_0)\rangle. \tag{3.29}$$

We consider, without loss of generality, the case $t_2>t_1>t_0$. Using the FRR and the chain rule twice, we obtain

$$\begin{aligned}\widetilde{C}_{ijk}(t_2,t_1,t_0)&=\left.\frac{\delta^2\zeta_i(t_2)}{\delta h_j(t_1)\delta h_k(t_0)}\right|_{\mathbf{h}=\mathbf{0}}\\&=\frac{\partial^2\zeta_i}{\partial\mu_l\partial\mu_m}(\boldsymbol{\mu},t_2)\left.\frac{\delta\mu_l(t_2)}{\delta h_j(t_1)}\right|_{\mathbf{h}=\mathbf{0}}\left.\frac{\delta\mu_m(t_2)}{\delta h_k(t_0)}\right|_{\mathbf{h}=\mathbf{0}}\\&\quad+\frac{\delta\zeta_i}{\partial\mu_l}(\boldsymbol{\mu},t_2)\left.\frac{\delta^2\mu_l(t_2)}{\delta h_j(t_1)\delta h_k(t_0)}\right|_{\mathbf{h}=\mathbf{0}}.\end{aligned} \tag{3.30}$$

We see that the slaving principle holds in a generalized sense. Now higher-order derivative terms in the Taylor expansion of $\zeta(\boldsymbol{\mu},t)$ appear beyond the leading one.

In order to focus on the dynamical aspects of the Rayleigh-Ritz approximation for $n=2$, let us consider now just the special case $\mathbf{Z}(t)=\mathbf{M}(t)$, so that $\boldsymbol{\zeta}(\boldsymbol{\mu},t)=\boldsymbol{\mu}$. We shall show that the FRR for $n=2$ [and $\mathbf{Z}(t)=\mathbf{M}(t)$] implies that

$$\begin{aligned}&\widetilde{C}_{ijk}(t_2,t_1,t_0)\\&\quad=\int_{t_1}^{t_2}dt\,E_{ip}(t_2,t)\frac{\partial^2V_p}{\partial\mu_q\partial\mu_r}(\boldsymbol{\mu},t)\widetilde{C}_{rj}(t,t_1)\widetilde{C}_{qk}(t,t_0)\\&\quad\quad+E_{ip}(t_2,t_1)\frac{\partial C_{jp}}{\partial\mu_q}(\boldsymbol{\mu},t_1)\widetilde{C}_{qk}(t_1,t_0).\end{aligned} \tag{3.31}$$

We have introduced the propagator of the linearized closure dynamics:

$$\mathbf{E}(t,t'):=T\exp\left(\int_{t'}^{t}\mathbf{A}_*(s)ds\right). \tag{3.32}$$

The result (3.31) is obtained by taking the second functional derivative with respect to $\mathbf{h}(t_1)$ of the first-order response functional

$$\widetilde{\mathbf{R}}[t_2,t_0;\mathbf{h}]=T\exp\left(\int_{t_0}^{t_2}\mathbf{A}(s)ds\right)\mathbf{C}(\boldsymbol{\mu},t_0) \tag{3.33}$$

using the simple identity (a continuous ''product rule'' of functional differentiation)

$$\begin{aligned}&\frac{\delta}{\delta h_j(t_1)}T\exp\left(\int_{t_0}^{t_2}\mathbf{A}(t)dt\right)\\&\quad=\int_{t_1}^{t_2}dt\,T\exp\left(\int_{t}^{t_2}\mathbf{A}(s)ds\right)\frac{\delta\mathbf{A}(t)}{\delta h_j(t_1)}T\exp\left(\int_{t_0}^{t}\mathbf{A}(s)ds\right),\end{aligned} \tag{3.34}$$

and then setting $\mathbf{h}=\mathbf{0}$.

Some physical insight into the result (3.31) of the closure FRR can be obtained by rederiving the result in a more heuristic way. Let us make a *nonlinear regression hypothesis*: that the fluctuations evolve, in general, according to the full closure dynamics. By a Taylor expansion to quadratic order, one then obtains

$$\begin{aligned}\delta\dot{M}_i(t)&=A_{*,ij}(t)\delta M_j(t)+\frac{1}{2}\frac{\partial^2V_i}{\partial\mu_q\partial\mu_r}(\boldsymbol{\mu},t)\delta M_q(t)\delta M_r(t)\\&\quad+O(\delta M^3).\end{aligned} \tag{3.35}$$

This equation can be rewritten in integral form as

$$\begin{aligned}\delta M_i(t_2)&=E_{ip}(t_2,t_1)\delta M_p(t_1)\\&\quad+\frac{1}{2}\int_{t_1}^{t_2}dt\,E_{ip}(t_2,t)\frac{\partial^2V_p}{\partial\mu_q\partial\mu_r}(\boldsymbol{\mu},t)\delta M_q(t)\delta M_r(t)\\&\quad+O(\delta M^3).\end{aligned} \tag{3.36}$$

Multiplying by $\delta M_j(t_1)\delta M_k(t_0)$ and averaging then yields the formula

$$\begin{aligned}&\langle\delta M_i(t_2)\delta M_j(t_1)\delta M_k(t_0)\rangle\\&\quad\approx\frac{1}{2}\int_{t_1}^{t_2}dt\,E_{ip}(t_2,t)\frac{\partial^2V_p}{\partial\mu_q\partial\mu_r}(\boldsymbol{\mu},t)\\&\quad\quad\times\langle\delta M_q(t)\delta M_r(t)\delta M_j(t_1)\delta M_k(t_0)\rangle+E_{ip}(t_2,t_1)\\&\quad\quad\times\langle\delta M_{kp}(t_1)\delta M_j(t_1)\delta M_k(t_0)\rangle.\end{aligned} \tag{3.37}$$

Next, one can approximate the remaining correlations by discarding all cumulants of higher order than third (and disconnected terms single-time in t). Then,

$$\begin{aligned}&\langle\delta M_q(t)\delta M_r(t)\delta M_j(t_1)\delta M_k(t_0)\rangle\\&\quad\approx\widetilde{C}_{qj}(t,t_1)\widetilde{C}_{rk}(t,t_0)+(q\leftrightarrow r),\end{aligned} \tag{3.38}$$

which relation, substituted into the first term of Eq. (3.37), yields precisely the first term of Eq. (3.31). Likewise,

$$\begin{aligned}&\langle\delta M_p(t_1)\delta M_j(t_1)\delta M_k(t_0)\rangle\\&\quad\approx-C_{pl}(t_1)C_{jm}(t_1)\Gamma_{lmq}(t_1)\widetilde{C}_{qk}(t_1,t_0),\end{aligned} \tag{3.39}$$

where $\Gamma_{lmq}(\boldsymbol{\mu},t_1)$ is the instantaneous third-order irreducible correlation in the PDF *Ansatz*. Substitution into the second term of Eq. (3.37) yields the corresponding term of Eq. (3.31) if we assume that

$$\frac{\partial C_{jp}}{\partial\mu_q}(\boldsymbol{\mu},t_1)=-C_{pl}(t_1)C_{jm}(t_1)\Gamma_{lmq}(t_1) \tag{3.40}$$

instantaneously at time t_1.

Thus, the two terms in Eq. (3.31) have quite different physical interpretations. The first integral term represents triple correlations dynamically generated at intermediate times $t_2>t>t_1$ from the correlations propagating in from times t_1 and t_0, which subsequently then relax to time t_2 according to the linearized closure dynamics. The second derivative $\partial^2V_p/\partial\mu_q\partial\mu_r$ can be interpreted as an ''interaction vertex,'' due to the nonlinear terms in the closure equation, by means of which the fluctuations interact. The second term, on the other hand, is not produced by any interaction of the fluctuations. It represents a triple correlation present instantaneously in the closure PDF *Ansatz* which is simply propagated in time by the linearized dynamics. Both of these terms, as well as the terms for cases $n>2$, can be expressed in terms of suitable diagrams. These involve propagators from the linearized dynamics and vertices from both the nonlinear terms in the dynamics and the higher-order correlations in the instantaneous PDF *Ansatz* (see the Appendix). Of course, these are not ''bare'' Feynman diagrams, for the Rayleigh-Ritz approximation is highly nonperturbative and the propagators and vertices represent ''dressed'' objects resulting from statistical closure.

The single-time relation (3.40) that we derived heuristically is, in fact, a necessary condition for consistency of the Rayleigh-Ritz approximation to the triple correlation. Only if it is true will the three-time correlation in Eq. (3.31) coincide along the diagonal $t_2=t_1=t_0=t$ with the value $C_{ijk}(\boldsymbol{\mu},t)$ calculated from the single-time PDF *Ansatz* $\mathcal{P}(\mathbf{x};\boldsymbol{\mu},t)$ that is input into the Rayleigh-Ritz calculation. Indeed, if we *assume* that Eq. (3.40) holds, then the second term of Eq. (3.31) can be rewritten as

$$\tilde{C}^{(2)}_{ijk}(t_2,t_1,t_0)=-\tilde{C}_{il}(t_2,t_1)C_{jm}(t_1)\Gamma_{lmq}(t_1)\tilde{C}_{qk}(t_1,t_0), \tag{3.41}$$

from which it is manifest that $\tilde{C}_{ijk}(t,t,t)=C_{ijk}(\boldsymbol{\mu},t)$. Relations such as Eq. (3.40) are also very important in other contexts within the Rayleigh-Ritz method. For example, a relation equivalent to Eq. (3.40), or

$$\Gamma_{ijk}(\boldsymbol{\mu},t)=\frac{\partial\Gamma_{ij}}{\partial\mu_k}(\boldsymbol{\mu},t), \tag{3.42}$$

was employed in [5] to prove an H theorem at quadratic order.

However, the relations (3.40) and (3.42) are not automatically true for an arbitrary PDF *Ansatz*. They might be taken as definitions of the triple correlations within the Rayleigh-Ritz approximation, which, we should remember, has available to it directly from the single-time PDE *Ansatz* only the mean $\zeta(\boldsymbol{\mu},t)=\boldsymbol{\mu}$ and the covariance $\mathbf{C}_Z(\boldsymbol{\mu},t)=\mathbf{C}(\boldsymbol{\mu},t)$. However, the triple correlators $C_{ijk}(\boldsymbol{\mu},t)$ and $\Gamma_{ijk}(\boldsymbol{\mu},t)$ are symmetric in their indices i, j, k, while the definitions through Eqs. (3.40) and (3.42) need not have such symmetry. Thus, such a definition may not be consistent. Fortunately, it is possible to construct the closure to ensure that Eqs. (3.40) and (3.42) hold, by employing an exponential PDF *Ansatz*, such as those previously developed for Boltzmann kinetic equations in transport theory [6] (see also [5]). Within such a closure scheme the single-time irreducible correlators are all obtained from a generating function:

$$\Gamma_{i_1\cdots i_n}(\boldsymbol{\mu},t)=\frac{\partial^n H}{\partial\mu_{i_1}\cdots\mu_{i_n}}(\boldsymbol{\mu},t). \tag{3.43}$$

In fact, $H(\boldsymbol{\mu},t)$ is just the *relative entropy*. Because of Eq. (3.43), the consistency condition (3.42), as well as all higher-order ones, can be automatically ensured by constructing the closure via an exponential PDF *Ansatz*. Thus, this closure methodology has a special relation with the Rayleigh-Ritz approximation scheme. This will be discussed in detail elsewhere [7].

IV. THE FRR IN NUMERICAL COMPUTATION

A. Numerical differentiation

We have studied some of the properties of a PDF *Ansatz* required for a physically accurate and mathematically consistent approximation of time correlations via the FRR's. Another important issue is the feasibility and accuracy of the FRR's for use in numerical computations. Except in special circumstances, it will not be possible to employ the FRR's analytically and numerical solution on the computer will be required. We have seen that the FRR's give the time correlations by (functionally) differentiating the solutions of the modified closure equation (3.7) with respect to the control field **h**. The numerical problem is to compute the required derivatives. It is well known that finite-difference approximations of derivatives are inherently numerically unstable, because the decrease in differentiation step $\Delta\mathbf{h}$ needed to reduce truncation error must cause the round-off error in finite-precision arithmetic to grow. If the FRR's are to be a useful computational tool, better numerical differentiation methods must be devised.

Fortunately, this problem has been encountered and solved in the context of other dynamical problems. One of the main application areas is *sensitivity analysis*, in which the sensitivity of the solution of a dynamical equation to changes in the initial data or to parameters in the equation is required [8]. By ''sensitivity'' we mean just the Jacobian derivative matrix of the solution vector with respect to the parameter vector (or higher-order derivatives). Our problem is exactly of this form, in which the ''sensitivities'' required are those of the solution of the modified closure equation with respect to the added control field **h**. The numerical techniques that yield accurate derivatives in sensitivity analysis depend upon solving additional dynamical equations for the derivatives themselves. There are two general techniques for doing so, depending upon the time order of propagating derivatives: the ''forward mode'' and the ''reverse mode.'' These two techniques have already been illustrated in the context of our earlier discussion. Equation (3.17) for the response matrix is equivalent to

$$\frac{\delta\dot{\boldsymbol{\mu}}(t)}{\delta\mathbf{h}(t_0)}=\mathbf{A}(t)\,\frac{\delta\boldsymbol{\mu}(t)}{\delta\mathbf{h}(t_0)} \tag{4.1}$$

integrated forward in time with initial data

$$\left.\frac{\delta\boldsymbol{\mu}(t)}{\delta\mathbf{h}(t_0)}\right|_{t=t_0}=\mathbf{C}_Z^T(\boldsymbol{\mu},t_0). \tag{4.2}$$

Substituting the result into Eq. (3.16) gives the derivative $(\delta\tilde{W}_Z/\delta\mathbf{h}(t_0))[\mathbf{h}]$. This illustrates the forward mode. Alternatively, one may compute the same derivative by integrating the adjoint equation (3.20) for $\boldsymbol{\alpha}(t)$ backward in time and then substituting the result into the formula

$$\frac{\delta \widetilde{W}_Z}{\delta \mathbf{h}(t_0)}[\mathbf{h}] = \boldsymbol{\zeta}(t) + \mathbf{C}_Z(t)\boldsymbol{\alpha}(t). \tag{4.3}$$

This illustrates the reverse mode. Such equations may be developed for arbitrary derivatives and, implemented numerically, they yield accurate and stable approximations.

Not only are these approaches numerically efficient but they can also be largely automated. Software for ''automatic differentiation'' is now becoming widely available; see [9]. Such tools directly generate from source code for numerical computation of the solution of the dynamical equation a corresponding code for the computation of its derivative. There is no need to compute required input derivatives, such as $A_{ij} = \partial V_i/\partial \mu_j$, by hand. Furthermore, it easy to compute ''sensitivities'' with respect to new perturbations, such as those corresponding to a new class of variables $\mathbf{Z}(t)$ of interest, without requiring extensive recalculations. The method has been tested and proved successful in application to real-life codes for PDE's employed in fluid dynamics and elsewhere. The availability of such software greatly enhances the attractiveness of the FRR's as a computational method to calculate time correlations.

B. A simple example

To illustrate the computational use of the FRR, we shall consider a simple closure for the decay of homogeneous, isotropic turbulence governed by the Navier-Stokes dynamics. This closure was originally employed by Kolmogorov to predict the mean energy decay. It was employed within the Rayleigh-Ritz formalism in [10] to predict the two-time correlation of the energy fluctuations. It should be emphasized that this closure omits a physical effect that is very important in the decay of energy fluctuations: their relaxation by turbulent diffusion in space. To see such effects, one must construct the closure not just for the kinetic energy at a single point (say, the origin) but with the kinetic energy at *all* space points as closure variables. In that case, the closure equations contain ''eddy viscosity'' terms, which are an important linear relaxation mechanism for fluctuations. Such improvements have been investigated and tested against simulation data in [11]. However, the one-moment closure is adequate for our purpose, which is to study the utility of the FRR for numerical computations. The main merit of the closure is that the Rayleigh-Ritz two-time correlation is given analytically; see Eq. (4.4) in [10]. This provides an objective basis of comparison for numerical results. We shall consider here only the two-time correlations, i.e., the case $n=1$.

The closure we consider has just one moment function, the kinetic energy $K = \frac{1}{2}v^2$ at a single point in space. The moment average μ is here denoted E. It obeys the equation

$$\dot{E}(t) = -\Lambda_m E^p(t) \tag{4.4}$$

in which Λ_m, p are suitable real constants. See [10] for details. The single-time covariance $C(t) := \langle [\delta K(t)]^2 \rangle$ is given in the closure by

$$C(E;t) = \tfrac{2}{3}E^2, \tag{4.5}$$

which follows from assuming a Gaussian one-point velocity distribution. Thus, the perturbed closure equation becomes here

$$\dot{E} = V(E) + h(t)C(E), \tag{4.6}$$

in which $V(E)$ and $C(E)$ are given by Eqs. (4.4) and (4.5), respectively. Now the FRR states that the two-time correlation $C(t,t_0) := \langle \delta K(t)\delta K(t_0)\rangle$ is given in the Rayleigh-Ritz approximation by

$$\widetilde{C}(t,t_0) = \left.\frac{\delta E(t)}{\delta h(t_0)}\right|_{h=0} \tag{4.7}$$

for $t > t_0$.

The right-hand side of Eq. (4.7) has been calculated by us numerically, in two different ways. The first method is based upon the observation that

$$\left.\frac{\delta E(t)}{\delta h(t_0)}\right|_{h=0} = \left.\frac{\partial E}{\partial h}(t;h)\right|_{h=0}, \tag{4.8}$$

where $E(t;h)$ is the solution of the closure equation for the modified initial datum

$$E(t_0;h) := E_0 + hC(E_0;t_0). \tag{4.9}$$

The closure equations were numerically integrated with a fourth-order Runge-Kutta scheme with time step $\Delta t = 10^{-3}$, in double-precision arithmetic, but with initial datum given by Eq. (4.9) for the two small values $h = \pm 10^{-6}$. The derivative (4.8) of the solution at later times t was then estimated by the symmetric, second-order finite-difference approximation to the derivative. The second method for calculating the functional derivative in Eq. (4.7) is to solve the analog of Eq. (4.1) with initial datum $[\delta E(t)/\delta h(t_0)]|_{t=t_0} = C(E_0,t_0)$, together with the closure equation itself. These were both integrated numerically by the same Runge-Kutta code as before. The matrix $\mathbf{A}(t)$ that appears in Eq. (4.1) (here, a 1×1 matrix) is given analytically by

$$A_*(t) = -p\Lambda_m E^{p-1}(t) \tag{4.10}$$

and it was input directly into the code. Hence, the only errors in the functional derivative calculated by this second method, as in the solution of the closure dynamics for the means, are the fourth-order truncation errors and the round-off errors.

We show in Fig. 1 the correlation $C(t,t_0)$ calculated by the FRR, compared with the analytical Rayleigh-Ritz solution given in Eq. (4.4) of [10]. Both the values calculated by the finite-difference approximation (method 1) and the adjoined equation for the Jacobian (method 2) are shown. As may be seen, these agree perfectly, both with each other and with the exact result. This is not unexpected in our example, considering the high-order accuracy of our approximations and the double-precision arithmetic. In other cases, it is hard to assess *a priori* the accuracy of the finite-difference approximation, for which the optimal discretization step size $\Delta\mathbf{h}_{\rm opt}$ may depend upon both space and time in an unknown manner.

It was shown in [5] that, in general, the same two-time correlations provided by the FRR are also given by a linear

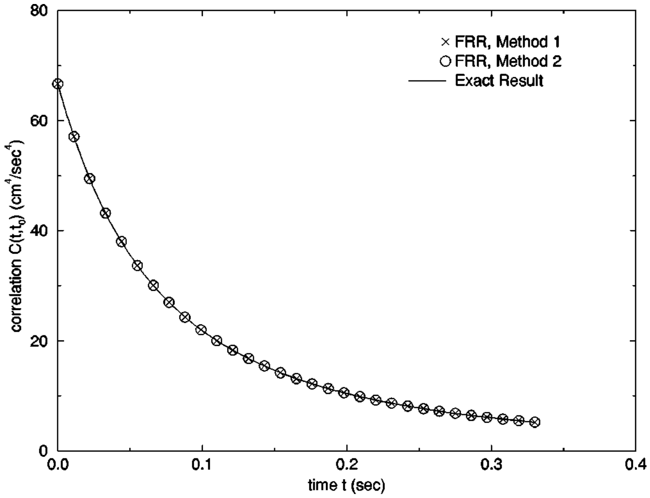

FIG. 1. Comparison of two-time correlations. FRR calculations vs exact result. Initial energy at $t_0=0$ was set to $E_0=10\,\mathrm{cm}^2/\mathrm{s}^2$. The closure parameters in the notation of Ref. [10] are $m=2$, $A=0.001\,\mathrm{cm}^5/\mathrm{s}^2$, and Kolmogorov constant $\alpha=2$.

Langevin equation. That model may be only formal, since the covariance of its noise term need not be positive. In any case, the numerical use of the Langevin model to calculate the two-time correlations is far less efficient than the numerical use of the FRR. Not only must the stochastic equation be integrated for a large enough number of realizations $N\gg 1$, but also in each realization random number generators must be called in each step of the time integration. Furthermore, the individual realizations governed by the Langevin dynamics will be far less smooth in space and time than averages over the ensemble, and thus much smaller space and time discretization steps $\Delta\mathbf{x}$ and Δt will be required. The computational expense of using the Langevin model is, thus, far higher than for the FRR and is not to be recommended. The stochastic equation is useful only for conceptual purposes. The FRR, on the other hand, is quite efficient because it takes full advantage of the increased regularity and stability of statistically averaged quantities. It is really a ‘‘thermodynamic approach’’ to calculating the time correlations and not a ‘‘statistical-mechanical’’ method.

V. CONCLUSIONS

In this paper we have reviewed and simplified the variational approach to statistical dynamics proposed in [4]. As a main result, we derived a general fluctuation-response relation for arbitrary multitime correlations. We demonstrated that the FRR's are preserved in a moment-closure approximation by the Rayleigh-Ritz method. We discussed the physical significance of the closure FRR's in terms of various intuitive hypotheses: slaving, regression (linear and nonlinear) of fluctuations. We also discussed computationally efficient and accurate methods for computing the derivatives required in the FRR's.

Many interesting problems can be investigated with the present methods. These include temporal multiscaling in turbulence [12,13], aging phenomena in glassy relaxation [14,15], transition rate theory in chemical kinetics [16], and Lagrangian statistics of advected scalar reactants [17]. The FRR should also hold for quantum systems, governed by quantum Liouville or master equations. Rayleigh-Ritz methods could provide a tractable means to compute multitime statistics in quantum field theory and in the quantum many-body problem.

ACKNOWLEDGMENTS

The author wishes to thank F. Alexander, C. D. Levermore, and J. Restrepo for valuable conversations and suggestions that contributed to this work.

APPENDIX: DIAGRAMMATIC RULES

We sketch concisely here the diagrammatic rules that follow from the closure FRR for the multitime cumulants or connected correlation functions. The derivation generalizes that for the three-time cumulant in Sec. III C. It is advantageous to specify the latest time to be t_n and the earliest time t_0. Then, $(n+1)$-time correlations for $t_f=t_n>t_{n-1}>\cdots>t_t>t_0=t_i$ are obtained by successive functional differentiation of expression (3.33) for $\widetilde{\mathbf{R}}[t_n,t_0;\mathbf{h}]$ with respect to $\mathbf{h}(t_1),\ldots,\mathbf{h}(t_{n-1})$, using the ‘‘product rule’’ (3.34). We recall from Eq. (3.18) that

$$\mathbf{A}(t):=\frac{\partial\mathbf{V}}{\partial\boldsymbol{\mu}}(\boldsymbol{\mu},t)+\mathbf{h}^{\top}(t)\,\frac{\partial\mathbf{C}}{\partial\boldsymbol{\mu}}(\boldsymbol{\mu},t)$$

and employ the chain rule to calculate successive derivatives. We also assume that the single-time irreducible correlations have the entropy as a generating function, as in Eq. (3.43).

The terms that result may be associated with graphs. The lines in the graphs terminating at times t, t' (internal or external) are given by the covariance function $\widetilde{C}_{ij}(t,t')$. If $t=t'$, then $\widetilde{C}_{ij}(t,t)=C_{ij}(\boldsymbol{\mu},t)$. The vertices are of two types. For each integer $r\geq 2$ there are $(r+1)$-fold vertices of the form

$$W_i^{j_1\cdots j_r}(s):=\Gamma_{im}(s)\,\frac{\partial^r V_m}{\partial\mu_{j_1}\cdots\mu_{j_r}}(\boldsymbol{\mu},s)$$

and

$$-\Gamma_{j_1\cdots j_{r+1}}(t)=-\frac{\partial^{r+1}H}{\partial\mu_{j_1}\cdots\mu_{j_{r+1}}}(\boldsymbol{\mu},t),$$

where the latter is just minus the single-time, irreducible $(r+1)$st-order correlator. The minus sign appears because of the fact that $\mathbf{C}=\boldsymbol{\Gamma}^{-1}$ and thus

$$\frac{\partial\mathbf{C}}{\partial\boldsymbol{\mu}}=-\mathbf{C}\,\frac{\partial\boldsymbol{\Gamma}}{\partial\boldsymbol{\mu}}\mathbf{C}.$$

One may replace the W vertices with $V_m^{j_1\cdots j_r}(s):=(\partial^r V_m/\partial\mu_{j_1}\cdots\mu_{j_r})(\boldsymbol{\mu},s)$ if the propagator line $\widetilde{C}_{ki}(s',s)$ entering the i node of the W vertex is replaced by a linear propagator $E_{km}(s',s)$ entering the m node of the V vertex.

The following rules apply

(i) The graphs that appear are all tree graphs with the times $t_n,t_{n-1},\ldots,t_1,t_0$ terminating the external lines. The trees are rooted at time t_n and branch up to earlier times, with the times nonincreasing as one ascends the tree.

(ii) Each vertex must be linked to at least one of the early external times $t_{n-1},\ldots,t_0$ directly by a propagator line.

(iii) The Γ-type vertices are all evaluated at an early external time $t_{n-1},\ldots,t_0$, which is determined as the latest time reached by any branch starting upward from that vertex and passing only through Γ-type vertices.

(iv) The *W*-type vertices (or the *V*-type) are evaluated at internal times *s* that are integrated over the largest possible subrange of $t_n > s > t_0$ consistent with the rule of nonincreasing times ascending the tree.

Because of rule (ii), it is clear that there are only finitely many graphs contributing to each $(n+1)$-time cumulant function, with vertices of at most $(n+1)$st order appearing. The finite sum of all the contributions from these graphs gives the FRR result for the cumulant function. Thus, it is clear that this graphical representation is not a perturbation expansion into Feynman diagrams, since the latter contain closed loops and infinitely many terms. The propagators and vertices here are all ''dressed objects'' and the representation is nonperturbative.

When all the external times are equal, $t_n = \ldots = t_0 = t$, then the graphical representation simplifies considerably. There are then no *V*- or *W*-type vertices, because the integration range over each internal time *s* has shrunk to zero. Furthermore, all of the Γ-type vertices are now evaluated at the same time *t*. In fact, the resulting graphical expansion is just that of the well-known representation of the single-time $(n+1)$st-order cumulant $C_{i_1 \ldots i_{n+1}}(t)$ as a sum over tree diagrams with single-time irreducible correlations $\Gamma_{i_1 \ldots i_{r+1}}(t)$ as vertices and second-order correlators $C_{ij}(t)$ on the internal and external lines. Thus, we obtain a proof for any order $(n+1)$ that, along the diagonal $t_n = \ldots = t_0 = t$ in time, $\tilde{C}_{i_1 \ldots i_{n+1}}(t,\ldots,t) = C_{i_1 \ldots i_{n+1}}(\boldsymbol{\mu},t)$.

[1] R. H. Kraichnan, Phys. Rev. **113**, 1181 (1959).
[2] U. Deker and F. Haake, Phys. Rev. A **11**, 2043 (1975).
[3] R. Graham, Z. Phys. B: Condens. Matter **26**, 397 (1977).
[4] G. L. Eyink, Phys. Rev. E **54**, 3419 (1996).
[5] G. L. Eyink, Phys. Rev. E **58**, 6975 (1998).
[6] C. D. Levermore, J. Stat. Phys. **83**, 1021 (1996).
[7] G. L. Eyink and C. D. Levermore (unpublished).
[8] D. G. Cacuci, J. Math. Phys. **22**, 2794 (1981); **22**, 2803 (1981).
[9] C. Bischof *et al.*, Sci. Prog. **1**, 1 (1992).
[10] G. L. Eyink, Phys. Rev. E **56**, 5413 (1997).
[11] G. L. Eyink and A. J. Wray, in *Studying Turbulence Using Numerical Simulation Databases-VII*, 1998 CTR Summer Program, edited by D. Moin and W. C. Reynolds (NASA Ames and Stanford University, Stanford, 1998); available at http://www-fpc.standford.edu/CTR/SP98.html.
[12] L. Biferale *et al.*, e-print chao-dyn/9804035.
[13] V. L'vov, E. Podivilov, and I. Procaccia, Phys. Rev. E **55**, 7030 (1997).
[14] L. F. Cugliandolo and J. Kurchan, Prog. Theor. Phys. **126**, 407 (1998).
[15] J.-P. Bouchad *et al.*, e-print chao-dyn/9702070.
[16] J. M. Sancho, A. H. Romero, and K. Lindenberg, J. Chem. Phys. **109**, 9888 (1998).
[17] R. O. Fox, Phys. Fluids **9**, 2364 (1997).