

# Optimal prediction of underresolved dynamics

ALEXANDRE J. CHORIN, ANTON P. KAST, AND RAZ KUPFERMAN

Department of Mathematics, Lawrence Berkeley National Laboratory, Mail Stop 50A-2152, 1 Cyclotron Road, Berkeley, CA 94720

**ABSTRACT** A method is presented for computing the average solution of problems that are too complicated for adequate resolution, but where information about the statistics of the solution is available. The method involves computing average derivatives by interpolation based on linear regression, and an updating of a measure constrained by the available crude information. Examples are given.

## Outline of Goal and Method

There are many problems in science whose solution is described by a set of differential equations, but where the solution of these equations is so complicated that it cannot be found in practice, even numerically, because it cannot be resolved properly. An accurate numerical solution requires that the problem be well resolved, i.e., that enough variables (“degrees of freedom”) be retained in the calculation to represent all relevant features of the solution. Well known examples where good resolution cannot be achieved in practice include turbulence and various problems in statistical physics and in economics. We consider in the present paper what one should do in underresolved problems, i.e., problems where good resolution has not been achieved.

There exists a large literature on the solution of underresolved problems, describing a wide variety of problem-specific tools; for example, in the case of turbulence there are modeling methods and methods for “large-eddy simulation” (1). All these methods involve some assumptions about the relation between what one can effectively compute and the “invisible” degrees of freedom that cannot be properly represented. Obviously, nothing can be done without some information in addition to what can be computed. In the present paper we assume that the additional information consists of explicit information about a measure preserved by the differential equations. In many problems of interest this kind of information is available, but it does not seem to have been fruitfully utilized previously. The key special case of turbulence will be treated in detail elsewhere, where it will be shown that it falls within the class of problems to which our methods apply.

The situation of interest is as follows: Consider a differential equation of the form

$$u_t = R(u), \quad [1]$$

where  $t$  is the time and  $R(u) = R(u, \partial u / \partial x, \dots)$  is a (generally nonlinear) function of its arguments. Assume in addition that a measure on the space of solutions of Eq. 1 is invariant under the flow induced by Eq. 1, and that we know what it is. We denote averages with respect to this measure by angle brackets:  $\langle \cdot \rangle$ . We further assume that we cannot resolve  $u$ , but that we can find some information about  $u$  at a small number of mesh points; the information could consist of point values but it is more reasonable physically and mathematically to assume that what one has is “filtered” values, as would be indeed produced

by a real physical measurement. We thus assume that what we have at the mesh points are the values  $\bar{u}_\alpha$  defined by:

$$\bar{u}_\alpha(t) = \int G(x - x_\alpha) u(x, t) dx, \quad [2]$$

where  $\alpha$  is an index on the mesh. The kernel  $G$  can represent, for example, spatial averaging. The coarse grid data in Eq. 2 specify, at every moment in time, a subset of functions (the functions that are consistent with these observations). We denote averages over this restricted subset by angle brackets with a subscript corresponding to the observations:  $\langle \cdot \rangle_{\bar{u}}$ . We call the measure on the restricted subset the “constrained measure.” If the problem is underresolved, the measure is carried by a nontrivial set of functions. It is important to note that the constrained measure is not invariant; indeed, if the invariant measure we start with is ergodic, the constrained measure tends in time to the unconstrained measure; for example, if we demanded that initially all the functions assumed given values at the grid points, there is no reason to believe that the solutions of the differential equations that evolve from these data would still take on the very same values at the mesh points at later times.

Our goal is to calculate averages of the solutions with respect to the constrained measure; these averages represent what one can calculate on the crude grid, properly averaged over the “invisible” degrees of freedom that cannot be represented on that grid. Given the constrained measure and the filtered values, the mean and the moments of the solution can be found at all points by interpolation (equivalent to linear regression, ref. 2), and therefore the mean derivatives of these quantities at the computational points can be found, for all practical purposes without error. The remaining problem is to characterize the evolution of the constrained measure so that the mean solution (and any moments that may be needed) can be advanced in time. Our assumption is that the constrained measure remains the invariant measure constrained by  $n$  filters, where  $n$  is the number of grid points. The filters therefore have to change in time. The formulas below will allow us to relate quantities whose evolution can be calculated to the parameters that determine the evolving filters. In the present paper we shall simplify the problem of finding the evolving filters by assuming that the filters are determined by the evolution of the mean solution (and not, for example, by the evolution of the higher moments of the solutions). This is equivalent to assuming that the equation

$$\frac{d\bar{u}_\alpha}{dt}(t) = \left\langle \int G(x - x_\alpha) R(u) dx \right\rangle_{\bar{u}(t)} \quad [3]$$

is a good approximation to the real evolution of the  $\bar{u}_\alpha(t)$  given by Eqs. 1 and 2. Of course, the validity of this assumption depends in particular on a good choice of filters  $G$ . In the present paper, we assume furthermore that the measure is either Gaussian or approximately Gaussian, in a sense specified below.

The two keys to success are: (i) averaging with respect to the right constrained measure and (ii) updating the constraints as the solutions evolve. Ingredient (i) was already used numeri-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

© 1998 by The National Academy of Sciences 0027-8424/98/954094-5\$2.00/0  
PNAS is available online at <http://www.pnas.org>.

cally in ref. 3. A number of interesting attempts have been made over the years to “fill in” data from coarse grids in difficult computations so as to enhance accuracy without refining the grid (see, e.g., refs. 4 and 5), but without our two key ingredients the usefulness of the earlier methods is necessarily limited.

We proceed as follows: First we present some elementary but important results on constrained Gaussian probabilities; then we explain and apply our scheme in the special cases of linear and nonlinear Schrödinger equations. Note that in the nonlinear case the invariant measure is not Gaussian. We also explain why these are significant test models. Because these model equations have some simplifying features, we also sketch a more general methodology.

**Gaussian Distributions: Conditional Expectation Under Affine Constraints**

We start by describing how to calculate expectation values of functions of normally distributed variables when the variables satisfy constraints of the affine form. Let  $\mathbf{u} = (u_1, \dots, u_n)^T$  be a real vector of jointly normal random variables; it has a probability density  $F(\mathbf{u})$  of the form,

$$P(s_1 < u_1 \leq s_1 + ds_1, \dots, s_n < u_n \leq s_n + ds_n) = F(\mathbf{s}) ds_1 \dots ds_N = Z^{-1} \exp\left(-\frac{1}{2} s_i a_{ij} s_j + b_i s_i\right) ds_1 \dots ds_N, \quad [4]$$

where  $Z$  is the appropriate normalization factor, repeated indices imply summation, the  $n \times n$  matrix  $A$  with entries  $a_{ij}$  is symmetric, and its inverse  $A^{-1}$ , assumed to exist, is the pairwise covariance matrix whose elements are

$$a_{ij}^{-1} = \text{Cov}[u_i, u_j] \equiv \langle u_i u_j \rangle - \langle u_i \rangle \langle u_j \rangle, \quad [5]$$

where the brackets denote averaging with respect to the probability density, and the vector  $\mathbf{b}$  with components  $b_i$  is related to the pointwise expectation values by

$$a_{ij}^{-1} b_j = \langle u_i \rangle. \quad [6]$$

The distribution is fully determined by the  $n$  means and by the  $\frac{1}{2}n(n + 1)$  independent elements of the covariance matrix; therefore, all expectation values of observables  $\mathcal{O}(\mathbf{u})$  can be expressed in terms of these parameters. In particular, all higher-order moments are given by Wick’s theorem (e.g., ref. 6).

Next, assume that the random vector  $\mathbf{u}$  satisfies a set of affine constraints of the form,

$$g_{\alpha i} u_i = \bar{u}_{\alpha} \quad \alpha = 1, \dots, m < n, \quad [7]$$

where the index  $\alpha$  enumerates the constraints; the matrix  $G$ , whose entries are  $g_{\alpha i}$ , is the discrete analog of the continuous kernel  $G(\cdot)$  introduced in Eq. 2. To distinguish between the vector space of random variables,  $(u_1, \dots, u_n)$ , and the vector space of constraints,  $(\bar{u}_1, \dots, \bar{u}_m)$ , we use Roman and Greek letter indices, respectively.

Our goal is to calculate conditional expectation values, i.e., averages over the functions that satisfy the constraints; formally,

$$\langle \mathcal{O}(\mathbf{u}) \rangle_{\bar{\mathbf{u}}} = \frac{\int (\prod_{i=1}^n du_i) \mathcal{O}(\mathbf{u}) F(\mathbf{u}) \prod_{\alpha=1}^m \delta(g_{\alpha j} u_j - \bar{u}_{\alpha})}{\int (\prod_{i=1}^n du_i) F(\mathbf{u}) \prod_{\alpha=1}^m \delta(g_{\alpha j} u_j - \bar{u}_{\alpha})}, \quad [8]$$

where the left-hand side introduces a notation for a constrained average and  $F(\mathbf{u})$  is the properly normalized probability density (Eq. 4). We shall use the following three lemmas:

LEMMA 1. *The conditional expectation of the variable  $u_i$  obeys an affine relation,*

$$\langle u_i \rangle_{\bar{\mathbf{u}}} = q_{i\alpha} \bar{u}_{\alpha} + c_i, \quad [9]$$

where the  $n \times m$  matrix  $Q$  with entries  $q_{i\alpha}$  and the  $n$ -vector  $\mathbf{c}$  with entries  $c_i$  are

$$Q = (A^{-1} G^T)(G A^{-1} G^T)^{-1} \quad [10]$$

$$\mathbf{c} = A^{-1} \mathbf{b} - (A^{-1} G^T)(G A^{-1} G^T)^{-1}(G A^{-1} \mathbf{b}).$$

LEMMA 2. *The conditional covariance matrix has entries*

$$\text{Cov}[u_i, u_j]_{\bar{\mathbf{u}}} = \langle u_i u_j \rangle_{\bar{\mathbf{u}}} - \langle u_i \rangle_{\bar{\mathbf{u}}} \langle u_j \rangle_{\bar{\mathbf{u}}} = a_{ij}^{-1} - [(A^{-1} G^T)(G A^{-1} G^T)^{-1}(G A^{-1})]_{ij}. \quad [11]$$

LEMMA 3. *Wick’s theorem holds for constrained expectations, namely,*

$$\left\langle \prod_{p=1}^P (u_{i_p} - \langle u_{i_p} \rangle_{\bar{\mathbf{u}}}) \right\rangle_{\bar{\mathbf{u}}} = \begin{cases} 0 & P \text{ odd} \\ \sum_{\text{perm}} \text{Cov}[u_{i_1}, u_{i_2}]_{\bar{\mathbf{u}}} \dots \text{Cov}[u_{i_{p-1}}, u_{i_p}]_{\bar{\mathbf{u}}} & P \text{ even} \end{cases} \quad [12]$$

where the summation is over all possible pairings of the  $P$  coordinates that are in the list.

Lemmas 1 and 2 can be deduced from standard linear regression theory. Lemma 3 can be proved by noting that a delta function is the limit of a narrow Gaussian function. As a result, the projection of a Gaussian measure on the subspace of functions that satisfy the constraints can be viewed as approximately Gaussian, hence satisfying Wick’s theorem; an appropriate limit can then be taken.

**A Linear Schrödinger Equation**

We demonstrate our method by applying it to two equations of Schrödinger type. We chose these problems because we feel that a nonlinear Schrödinger equation is a suitable one-dimensional cartoon of the Euler/Navier-Stokes problem that we are most interested in: it is Hamiltonian and nonlinear. The more popular cartoon, the Burgers equation, will be analyzed elsewhere; its peculiar properties (dominance of the solution by shocks and the need for a driving noise term to obtain an invariant measure) introduce added complications whose analysis does not fit within a short introductory paper.

We start with a linear Schrödinger equation written as a pair of real equations:

$$p_t = -q_{xx} + \mu^2 q, \quad [13]$$

$$q_t = +p_{xx} - \mu^2 p,$$

where  $p, q : [0, 2\pi) \times [0, \infty) \rightarrow \mathbb{R}$ ,  $\mu$  is a given constant, and periodic boundary conditions are assumed. This system of equations can be derived from a Hamiltonian,

$$H[p, q] = \frac{1}{2} \int_0^{2\pi} dx [(p_x)^2 + (q_x)^2 + \mu^2(p^2 + q^2)], \quad [14]$$

where  $p(x, t)$  and  $q(x, t)$  are the canonically conjugate variables.

This system has an invariant measure, the canonical measure, depending on a temperature that we set equal to one, so that the invariant measure has the density distribution  $F[p, q] = \exp(-H[p, q])$ . The quadratic form of the Hamiltonian makes the canonical distribution Gaussian, completely speci-

fied by the average functions  $\langle p(x) \rangle$  and  $\langle q(x) \rangle$ , which are zero by symmetry, and by the covariance functions,

$$\begin{aligned} \text{Cov}[p(x), p(y)] &= \text{Cov}[q(x), q(y)] = \sum_{k=-\infty}^{\infty} \frac{e^{ik(x-y)}}{k^2 + \mu^2} \\ \text{Cov}[p(x), q(y)] &= \text{Cov}[q(x), p(y)] = 0. \end{aligned} \tag{15}$$

The measure is absolutely continuous with respect to a Wiener measure, and the sample functions in this ensemble are continuous but not differentiable with probability one, so that the evolution of a single solution in time is hard to calculate.

We pick initial data from the invariant ensemble, but then assume that measurement has revealed, at each of  $m$  points, the values of two sets of filtered quantities that can be thought of as local averages of the solution:

$$\begin{aligned} \bar{p}_\alpha(t) &= \int_0^{2\pi} dx p(x, t) G(x - x_\alpha), \\ \bar{q}_\alpha(t) &= \int_0^{2\pi} dx q(x, t) G(x - x_\alpha), \end{aligned} \tag{16}$$

where  $\alpha = 1, \dots, m$ ,  $x_\alpha = 2\pi\alpha/m$  are the points about which the averages are evaluated, and the function  $G(\cdot)$  is a Gaussian filter (i.e., a function of Gaussian shape, not a random variable) of width  $\sigma$ . Once we know the values of these filtered values, the constrained average is defined. We wish of course to evaluate the future means of the solutions in the ensemble that has been constrained in this way, without calculating any particular solutions and using only the information at the  $m$  given points.

Given the values of the filters, formulas (Eqs. 9 and 11) allow us to find the mean of the solution in the ensemble conditioned by these filtered values; a short manipulation yields

$$\langle q(x) \rangle_{\bar{\mathbf{p}}, \bar{\mathbf{q}}} = r_\alpha(x) s_{\alpha\beta}^{-1} \bar{q}_\beta, \tag{17}$$

where

$$\begin{aligned} r_\alpha(x) &= \int_0^{2\pi} dy \text{Cov}[q(x), q(y)] G(y - x_\alpha) \\ s_{\alpha\beta} &= \int_0^{2\pi} dy r_\alpha(y) G(y - x_\beta), \end{aligned} \tag{18}$$

and the  $s_{\alpha\beta}^{-1}$  are the  $\alpha, \beta$  elements of the matrix that is the inverse of the matrix whose entries are the  $s_{\alpha\beta}$ . The result of this procedure is shown in Fig. 1, where we plot the mean of the functions that satisfy the constraints that have been imposed at a set of 10 points. The width of the filtering kernel is  $\sigma = 2\pi/m$  (the distance between the grid points). Note that we know exactly the mean of the functions in the constrained ensemble, and we know it everywhere.

In the present linear problem, if one knows the mean of the functions in the constrained ensemble, one can readily evaluate the mean time derivative at the grid points by applying the differential operator to the mean function. However, one cannot integrate this mean derivative in time and expect to obtain the mean future, because at a later time the ensemble that will have evolved from the initial constrained ensemble will not be the same as the initial ensemble. The assumption in the whole approach is that the ensemble prepared initially by placing conditions (the values of the filters) on the functions sampled from the invariant ensemble evolves into an ensemble that can still be represented as a subset of the ensemble that carries the invariant measure, with a measure conditioned by

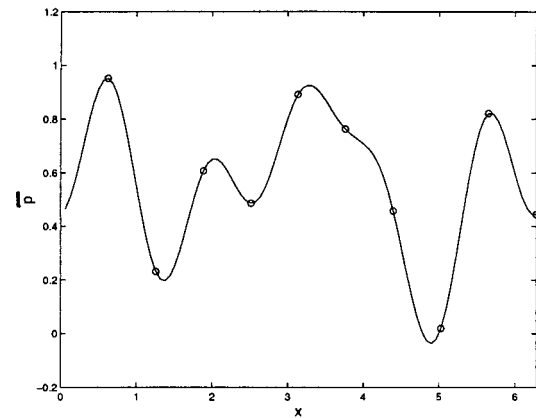


FIG. 1. The interpolating function. Given 10 constraints,  $\bar{p}_\alpha$  (open circles), the mean of the functions in the constrained ensemble is reconstructed (solid line).

new filters  $G$  and filtered values  $\bar{\mathbf{p}}$  and  $\bar{\mathbf{q}}$ . Here we are making the further assumption that it is sufficient to advance in time only mean values of  $\bar{\mathbf{p}}$  and  $\bar{\mathbf{q}}$ . Hence,

$$\begin{aligned} \frac{d\bar{p}_\alpha}{dt} &\approx \left\langle \frac{d\bar{p}_\alpha}{dt} \right\rangle_{\bar{\mathbf{p}}, \bar{\mathbf{q}}} = - \int_0^{2\pi} dx G(x - x_\alpha) \\ &\times \left[ \frac{\partial^2}{\partial x^2} \langle q(x) \rangle_{\bar{\mathbf{p}}, \bar{\mathbf{q}}} - \mu^2 \langle q(x) \rangle_{\bar{\mathbf{p}}, \bar{\mathbf{q}}} \right]. \end{aligned} \tag{19}$$

Using Eqs. 17 and 15 we find the more detailed expression:

$$\left\langle \frac{d\bar{p}_\alpha}{dt} \right\rangle_{\bar{\mathbf{p}}, \bar{\mathbf{q}}} = -d_{\alpha\beta} h_{\beta\gamma}^{-1} \bar{q}_\gamma, \tag{20}$$

with

$$d_{\alpha\beta} = \sum_{k=-\infty}^{\infty} e^{ik(x_\alpha - x_\beta)} e^{-k^2\sigma^2} \quad h_{\beta\gamma} = \sum_{k=-\infty}^{\infty} \frac{e^{ik(x_\beta - x_\gamma)}}{k^2 + \mu^2} e^{-k^2\sigma^2}. \tag{21}$$

An analogous equation with opposite sign is obtained for the mean time derivative of the vector  $\bar{\mathbf{q}}$ . We now have a closed system of equations for the estimation of the mean values of  $\bar{\mathbf{p}}$  and  $\bar{\mathbf{q}}$  in the future. We emphasize that the last assumption whereby it is sufficient to advance in time the mean values of  $\bar{\mathbf{p}}$  and  $\bar{\mathbf{q}}$  is nonessential and unlikely to be true in general (see the concluding section). The domain of validity of the more general assumption remains to be determined.

In Fig. 2 we plot the mean evolution of five filtered values  $\bar{p}_\alpha(t)$  for randomly selected initial values of  $\bar{p}_\alpha(0)$  and  $\bar{q}_\alpha(0)$ . The filtering function  $G(\cdot)$  has a width (defined above) equal to the distance between mesh points. In this linear problem the evolution of the mean of the filtered values can be calculated exactly simply by using Green's function for the problem to advance the mean solution and hence advance the values of the filters. In Fig. 2 the results of an exact calculation of the filtered values are represented by open circles, and the results obtained by integrating the set of 10 ordinary differential equations (Eq. 20) are represented by dashed lines. The two results are practically indistinguishable.

### A Nonlinear Hamiltonian System

We now consider a nonlinear generalization of the method demonstrated in the preceding section. We want to exhibit the power of our method by comparing the solutions that it yields with exact solutions; in the nonlinear case, exact solutions of problems with random data are hard to find, and we resort to

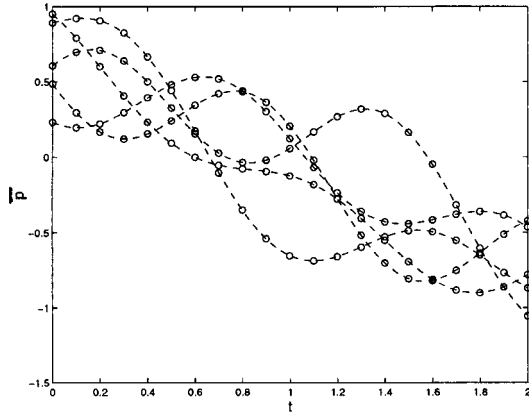


FIG. 2. Mean evolution of the vector  $\bar{p}$  for  $m = 5$ . The initial values of the coarse variables were chosen randomly. The open circles represent the exact solution, whereas the dashed lines were obtained by integrating equations 20.

a stratagem. Even though our method applies to the full nonlinear partial differential equation, we study a finite dimensional system of  $2n$  ordinary differential equations that is formally an approximation of a nonlinear Schrödinger equation:

$$\begin{aligned} \frac{dq_i}{dt} &= -\frac{p_{i-1} - 2p_i + p_{i+1}}{\Delta^2} + p_i^3 \\ \frac{dp_i}{dt} &= +\frac{q_{i-1} - 2q_i + q_{i+1}}{\Delta^2} - q_i^3 \end{aligned} \quad i = 1, \dots, n, \quad [22]$$

where  $\Delta = 1/n$  is a mesh size. The approximation is formal because the solutions of the differential equation are not smooth enough to ensure convergence. We consider a periodic system so that  $q_0 = q_n$  and  $q_{n+1} = q_1$  (with analogous conditions on  $p$ ). In the calculations below,  $n = 16$ .

The Hamiltonian of this system is

$$H[\mathbf{p}, \mathbf{q}] = \frac{1}{2} \sum_{i=1}^n \left[ \left( \frac{q_{i+1} - q_i}{\Delta} \right)^2 + \left( \frac{p_{i+1} - p_i}{\Delta} \right)^2 + \frac{1}{2} (q_i^4 + p_i^4) \right]. \quad [23]$$

Starting from this discrete system, we shall use a discrete version of the method above to deduce equations for a reduced set of variables, which will reproduce the mean solution of the full set of equations, and we shall be able to show explicitly that the algorithm does what it should.

As in the linear case, we assume that the canonical distribution has a temperature equal to one. Note that now the Hamiltonian is not a quadratic function and the invariant measure is not Gaussian. The lemmas of section 2 cannot be used without approximation. The equilibrium distribution  $P(\mathbf{p}, \mathbf{q}) = e^{-H[\mathbf{p}, \mathbf{q}]}$  is approximated by a Gaussian distribution, chosen so as to yield the same means and covariances as the original distribution. Symmetry considerations imply that  $\langle p_i \rangle = \langle p_j \rangle = \langle p_i q_j \rangle = 0$ . Translation and reflection invariance, on the other hand, imply that  $\langle p_i p_j \rangle = \langle q_i q_j \rangle = f_{|i-j|}$ . Thus, the approximate distribution depends on a single vector  $f_{|i-j|}$  that can be calculated, for example, by a straightforward Monte Carlo simulation.

We define a subset of  $2m$  coarse variables ( $m \ll n$ ) that represent local averages of the vectors  $\mathbf{p}$  and  $\mathbf{q}$ :

$$\bar{p}_\alpha = g_{\alpha i} p_i \quad \text{and} \quad \bar{q}_\alpha = g_{\alpha i} q_i, \quad [24]$$

where the index  $\alpha = 1, \dots, m$  enumerates the coarse variables. We consider the case  $m = 2$ ; thus, we replace a set of

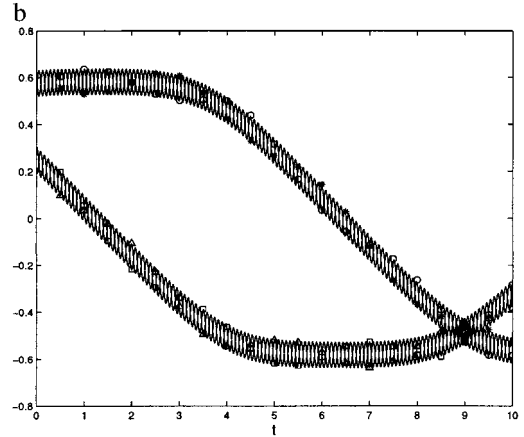
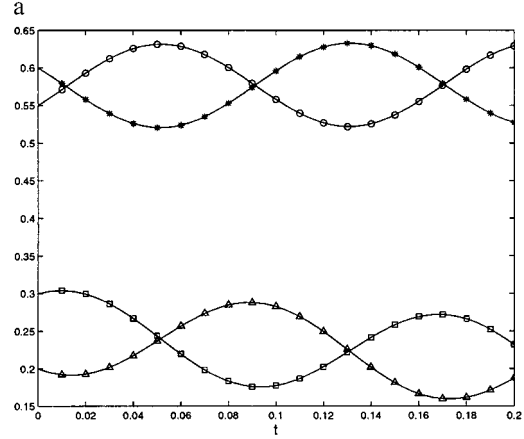


FIG. 3. Mean evolution of the coarse variables,  $\bar{p}_1$  ( $\Delta$ ),  $\bar{p}_2$  ( $\square$ ),  $\bar{q}_1$  ( $*$ ), and  $\bar{q}_2$  ( $\circ$ ). The symbols represent the mean evolution of an ensemble of initial conditions generated by means of a Monte Carlo algorithm. The solid lines are the solutions obtained by integrating the mean time derivatives (Eq. 28). The two graphs show different time intervals: (a)  $t = 0 - 0.2$ ; (b)  $t = 0 - 10$ .

32 variables by a reduced set of 4 variables. The filtering matrix  $G$  with entries  $g_{\alpha i}$  has a Gaussian profile:

$$g_{\alpha i} = B \exp \left[ -\frac{(x_i - \bar{x}_\alpha)^2}{\sigma^2} \right], \quad [25]$$

where  $B$  is a normalization factor,  $x_i = i/n$ ,  $\bar{x}_\alpha = \alpha/m$ , and  $\sigma$  is the width of the filter. In the calculations below we took  $\sigma = 0.25$ , i.e., the filter averages over an interval of the order of the distance between the coarse data points.

Given the constrained distribution, the mean time derivative of the coarse variables can be calculated. For example,

$$\left\langle \frac{d\bar{p}_\alpha}{dt} \right\rangle_{\bar{\mathbf{p}}, \bar{\mathbf{q}}} = \frac{1}{\Delta^2} g_{\alpha i} [\langle q_{i-1} \rangle_{\bar{\mathbf{p}}, \bar{\mathbf{q}}} - 2\langle q_i \rangle_{\bar{\mathbf{p}}, \bar{\mathbf{q}}} + \langle q_{i+1} \rangle_{\bar{\mathbf{p}}, \bar{\mathbf{q}}}] - g_{\alpha i} \langle q_i^3 \rangle_{\bar{\mathbf{p}}, \bar{\mathbf{q}}}. \quad [26]$$

The first term on the right-hand side can readily be calculated by the interpolation formula (Lemma 1). For the second term, which involves the expected value of a cubic function of  $q_i$ , we use the Gaussian approximation and apply Wick's theorem,

$$\langle q_i^3 \rangle_{\bar{\mathbf{p}}, \bar{\mathbf{q}}} = 3\langle q_i^2 \rangle_{\bar{\mathbf{p}}, \bar{\mathbf{q}}} \langle q_i \rangle_{\bar{\mathbf{p}}, \bar{\mathbf{q}}} - 2\langle q_i \rangle_{\bar{\mathbf{p}}, \bar{\mathbf{q}}}^3. \quad [27]$$

All these terms can be computed explicitly by using Lemmas

1 and 2. The mean time derivative of  $\bar{p}_1$  is

$$\left\langle \frac{d\bar{p}_1}{dt} \right\rangle_{\bar{\mathbf{p}}, \bar{\mathbf{q}}} = 19.5(\bar{q}_2 - \bar{q}_1) + [1.50\bar{q}_1^3 - 0.88\bar{q}_1^2\bar{q}_2 + 0.27\bar{q}_1\bar{q}_2^2 + 0.11\bar{q}_2^3 + 0.009\bar{q}_1 + 0.001\bar{q}_2]. \quad [28]$$

The term in square brackets is the “gaussianized” approximation to the cubic term in Eq. 26. The equation for  $\bar{p}_2$  is obtained by substituting  $1 \leftrightarrow 2$ ; the equations for  $\bar{q}_1, \bar{q}_2$  are obtained by the transformation  $\bar{p} \rightarrow \bar{q}$  and  $\bar{q} \rightarrow -\bar{p}$ .

We now test the accuracy of the prediction of the mean evolution of the vectors  $\bar{\mathbf{p}}$  and  $\bar{\mathbf{q}}$ , obtained by the integration of equations (Eq. 28). In contrast to the linear example, we do not have an exact mean solution for comparison. Instead, we generated numerically an ensemble of initial conditions that sampled the constrained canonical ensemble. We employed a standard Metropolis algorithm and sampled  $10^4$  states. Each state was then evolved in time using a fourth-order Runge–Kutta solver, and by averaging over the evolved ensemble we were able to compute the mean evolution of the solution of the system of 32 equations, which we then compared with the solution of the reduced system. The possibility of making this comparison is the reason we started with a discrete system rather than the full partial differential equation.

A comparison between the true mean and the mean produced by our averaging/updating procedure is shown in Fig. 3 *a* and *b*. The two graphs show different intervals of time; the coarse initial conditions are  $\bar{p}_1 = 0.2, \bar{p}_2 = 0.3, \bar{q}_1 = 0.6$ , and  $\bar{q}_2 = 0.55$ .

### Discussion

We have shown by examples that one can calculate average solutions accurately on a crude grid. The examples were simple but not trivial; in particular, they were complicated enough so

that the calculation of even a single sample solution was difficult. The main tools were interpolation based on regression and the characterization of constrained measures by time-dependent filters. In more complete papers to come we plan to show that the method generalizes to more complicated situations. In the Gaussian case, one can advance in time not only the mean solution but also its moments and covariances. As more information is updated, one can update not only the values of the filtered variables, but also parameters that determine the structure of the filters (i.e., one does not have to keep  $G$  in Eq. 2 fixed). In non-Gaussian problems, one can perform local “gaussianization” rather than global “gaussianization,” as above, and describe a broader range of measures. Furthermore, it should be obvious from the discussion that one does not need complete information about the measures, but only the covariances and low-order moments on scales smaller than the distance between mesh points. This partial information is available in a broad range of situations (7).

We are grateful to Prof. G. I. Barenblatt and to Dr. D. Nolan for many fruitful discussions. This work was supported by the Applied Mathematical Sciences Subprogram of the Office of Energy Research, under Contract Number DE-AC03-76SF-00098.

1. Lesieur, M., Comte, P. & Zinn-Justin, J., eds. (1996) *Computational Fluid Mechanics, Les Houches Summer School of Theoretical Physics, Session 54* (Elsevier, Amsterdam).
2. Papoulis, A. (1991) *Probability, Random Variables, and Stochastic Processes* (McGraw–Hill, New York).
3. Chorin, A. J. (1973) *Math. Comp.* **27**, 1–15.
4. Scotti, A. & Meneveau, C. (1997) *Phys. Rev. Lett.* **78**, 867–870.
5. McComb, W. D. & Watt, H. G. (1990) *Phys. Rev. Lett.* **65**, 3281–3284.
6. Kleinert, H. (1989) *Gauge Fields in Condensed Matter* (World Scientific, Teaneck, NJ).
7. Barenblatt, G. I. & Chorin, A. J. (1998) *Proc. Symp. Appl. Math.* **54**, 1–25.