J Sci Comput (2008) 37: 189–201
DOI 10.1007/s10915-008-9242-4

Modified Optimal Prediction and its Application to a Particle-Method Problem

Alina Chertock · David Gottlieb · Alex Solomonoff

Received: 29 October 2007 / Revised: 7 March 2008 / Accepted: 19 August 2008 /
Published online: 10 September 2008

Abstract The paper is concerned with system reduction by statistical methods and, in particular, by the optimal prediction method introduced in (Chorin, A.J., Hald, O.H., Kupferman, R., Optimal prediction with memory, Phys. D 166:239–257, 2002). The optimal prediction method deals with problems that possess large and small scales and uses the conditional expectation to model the influence of the small scales on the large ones.

In the current paper, we develop a different variant of the optimal prediction method as well as introduce and compare several approximations of this method. We apply the original and modified optimal prediction methods to a system of ODEs obtained from a particle method discretization of a hyperbolic PDE and demonstrate their performance in a number of numerical experiments.

Keywords Optimal prediction · Particle methods · Multiscale computations

1 Introduction

Many physical systems of scientific or engineering interest can have very large numbers of degrees of freedom. Usually we are not interested in all of these millions of variables, we just want to know a handful of statistical quantities.

The situation can be further explained as follows: in the numerical solution of nonlinear partial differential equations (PDEs), there are always (small) scales that can not be resolved

A. Chertock (✉)
Department of Mathematics, North Carolina State University, Raleigh, NC 27695, USA
e-mail: chertock@math.ncsu.edu

D. Gottlieb · A. Solomonoff
Division of Applied Mathematics, Brown University, Providence, RI 02912, USA

D. Gottlieb
e-mail: dig@dam.brown.edu

A. Solomonoff
e-mail: alex.solomonoff@yahoo.com

because there are not enough computational (memory/running time) to simulate those scales. In many important applications, however, the unresolved small scales may be important: very often it is not that these scales themselves needed to be resolved, but their influence on the large scales has to be accounted for. In most cases the impact of the small scales on the large ones is crucial in understanding the phenomenon. Therefore, we would like ways of approximating the interesting features of a system without keeping all the variables and doing all the work.

Reduction schemes attempt to take a very large dynamical system and reduce its size by taking into account the effects of a large subset of variables (called unresolved) on a small subset (resolved) without explicitly computing their evolution. This large subset is usually thought of as consisting of relatively uninteresting variables at the extreme end of some scale, e.g., fast or small, and thus the goal is to find a dynamical system that has a substantially smaller number of degrees of freedom, yet still modeling interesting aspects of the original system with acceptable accuracy.

In [7, 8] and a number of related papers [9, 12], an optimal prediction (OP) framework, which can be regarded as a type of order reduction scheme, has been recently developed. The OP method is based on the Mori-Zwanzig (M-Z) identity, which transforms a system of ordinary differential equations (ODEs) to a form amenable to simplifying approximations of systems of ODEs. The M-Z equation yields, in principle, exact reduction methods, but often hard to use (see, e.g., [11, 13, 16, 17, 21]), so in [7, 8] an approximation of the M-Z equation termed the t-model was presented. The t-model of a system of ODEs is a smaller system of ODEs that approximates the original system in a useful sense. It has been applied to Euler equations in [14] and to the Burgers equation in [1]. Its higher order versions have been recently constructed in [19] and used to compute the rate of energy decay for the Taylor-Green vortex problems [20].

In the current work, we have developed a different variant of the optimal prediction (OP) method. Several approximations of this method are introduced, one of them is similar to the aforementioned t-model. In Sect. 2, we summarize the former results of the OP method. We introduce the M-Z formulation and define the projection operator as the conditional expectation. In Sect. 3, we present our version of obtaining an equation for the resolved scales. We also derive simplified formulations and compare them. In Sect. 4, we show how the OP method can be applied to a systems of ODEs that arises in numerical solutions of PDEs by particle methods. To this end we consider a simple model problem, which nevertheless will give new insight into many important details of a reduced model derivation. In Sect. 4.2, we demonstrate the performance and accuracy of our approach in a number of numerical experiments. Finally, a brief overview of particle methods is given in Appendix.

2 The Optimal Prediction Method and the t-model

In this section, we briefly describe the OP method. For a detail derivation and analysis of the method we refer the reader to [7–9] and the references therein.

We start with the system of ODEs

$$\boldsymbol{\phi}_t = \mathbf{f}(\boldsymbol{\phi}) \tag{2.1}$$

that will be the focus of our study. Here, $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_n) \in \mathbb{R}^n$ is a vector function of a time variable t and $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ is a given function. We assume that n is large enough that the work of integrating (2.1) in time is inconveniently large. The system may be obtained from discretization of PDE's, in which case n tends to infinity.

The system (2.1) is supplemented by initial conditions, and part of the Mori-Zwanzig framework is the fact that $\boldsymbol{\phi}$ depends on them, so we have

$$\boldsymbol{\phi} = \boldsymbol{\phi}(t, \mathbf{x}) \quad \text{and} \quad \boldsymbol{\phi}(t = 0, \mathbf{x}) = \mathbf{x} \tag{2.2}$$

and $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ just as $\boldsymbol{\phi}$ is.

The next piece of the framework is partitioning $\boldsymbol{\phi}$ and $\mathbf{x}$ into a "resolved" part $\widehat{\boldsymbol{\phi}}, \widehat{\mathbf{x}} \in \mathbb{R}^m$ (with $m \ll n$) and an "unresolved" part $\widetilde{\boldsymbol{\phi}}, \widetilde{\mathbf{x}} \in \mathbb{R}^{n-m}$, so that

$$\boldsymbol{\phi} = (\widehat{\boldsymbol{\phi}}, \widetilde{\boldsymbol{\phi}}), \qquad \mathbf{x} = (\widehat{\mathbf{x}}, \widetilde{\mathbf{x}}). \tag{2.3}$$

An alternative way of looking at the splitting is to assume that the solution contain large and small scales and we are interested in resolving only the large scale part of the solution with taking into account the effect of the small scales on the large ones. If the ODE (2.1) is derived from a partial differential equation, then typically the high-frequency parts of the solution correspond to the unresolved variables, and low-frequency parts to the resolved variables. We will use high modes/low modes interchangeably with resolved/unresolved variables.

A convenient way of deriving the OP method is to use the Liouville operator that leads to a linear version of (2.1). The Liouville operator $\mathcal{L}$ associated with the ODE (2.1) is defined by

$$\mathcal{L} = \sum_j f_j(\mathbf{x}) \frac{\partial}{\partial x_j} = \mathbf{f} \cdot \nabla_{\mathbf{x}}. \tag{2.4}$$

It can be also defined as

$$\mathcal{L} = \sum_j f_j(\boldsymbol{\phi}) \frac{\partial}{\partial \phi_j} = \mathbf{f} \cdot \nabla_{\boldsymbol{\phi}}. \tag{2.5}$$

For the proof that the two definitions are equivalent we refer the reader to [7, p. 125].

The Liouville equation is then given by

$$u_t = \mathcal{L}u,$$

where $u : \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}$, which is a *linear* first-order PDE whose characteristics are the solutions of (2.1). If we now let $u(t = 0, \mathbf{x}) = u_0(\mathbf{x}) = x_j$, j takes any value, then

$$u(t, \mathbf{x}) = \phi_j(t, \mathbf{x}).$$

This important result requires a somewhat lengthy proof, which can be found in, e.g., [7].

The M-Z formula involves a time invariant projection operator $\mathcal{P}$ (which will be defined later) that projects every function to its large (resolved) scales. Having the operators $\mathcal{P}$ and $\mathcal{Q} = I - \mathcal{P}$, one can obtain the equation for the large scales only:

$$\frac{\partial \mathcal{P}u}{\partial t} = \mathcal{P}e^{\mathcal{L}t}\mathcal{P}\mathcal{L}u_0 + \int_0^t \mathcal{P}e^{\mathcal{L}(t-\tau)}\mathcal{P}\mathcal{L}e^{\mathcal{Q}\mathcal{L}\tau}\mathcal{Q}\mathcal{L}u_0\, d\tau. \tag{2.6}$$

This yields a reduced equation with fewer variables, though it is much more complicated than the original ODE (2.1).

In order to derive an efficient method for system reduction, the integral (the *memory term*) on the right-hand-side (RHS) of (2.6) has to be simplified. An approximation to (2.6), known as the t-model, was introduced in [8] and [9]. Assuming that initially u_0 can be

replaced by $\mathcal{P}u_0$ and expanding the memory term to first order in t, one can obtain that for short times

$$\int_0^t \mathcal{P}e^{\mathcal{L}(t-\tau)}\mathcal{P}\mathcal{L}e^{\mathcal{Q}\mathcal{L}\tau}\mathcal{Q}\mathcal{L}u_0\,d\tau \sim t\mathcal{P}\mathcal{L}\mathcal{Q}\mathcal{L}\mathcal{P}u_0.$$

Furthermore, assuming that for short times $e^{\mathcal{L}t}$ can be approximated by $e^{\mathcal{Q}\mathcal{L}t} \sim (I + \mathcal{Q}\mathcal{L}t)$ and replacing u_0 by $u(t)$ in the first term on the RHS of (2.6) and by $(I - \mathcal{Q}\mathcal{L}\mathcal{P})u(t)$ in the second term, one gets:

$$\frac{\partial \mathcal{P}u}{\partial t} \sim \mathcal{P}(I + \mathcal{Q}\mathcal{L}t)\mathcal{P}\mathcal{L}\mathcal{P}(I - \mathcal{Q}\mathcal{L}t)u + t\mathcal{P}\mathcal{L}\mathcal{Q}\mathcal{L}\mathcal{P}u.$$

Since $\mathcal{P}\mathcal{Q} = 0$, one obtains following the t-model:

$$\xi_t = \mathcal{P}\mathcal{L}\mathcal{P}\xi + t\mathcal{P}\mathcal{L}\mathcal{Q}\mathcal{L}\xi, \tag{2.7}$$

where ξ approximates $\mathcal{P}u$.

Finally one needs to define the operator $\mathcal{P}$. The original M-Z framework deals with this issue by using a probability distribution on $\mathbf{x}$, that is, making $\mathbf{x}$ a random vector, and then considering the conditional expectations of quantities, conditioned on $\widehat{\mathbf{x}}$. Recall that if the probability on $\mathbf{x}$ has a density function $q(\mathbf{x}) = q(\widehat{\mathbf{x}}, \widetilde{\mathbf{x}})$, then the conditional expectation of $\widetilde{\mathbf{x}}$ with respect to $\widehat{\mathbf{x}}$ is

$$E(\widetilde{\mathbf{x}}|\widehat{\mathbf{x}}) = \frac{\int \widetilde{\mathbf{x}}\, q(\widetilde{\mathbf{x}}, \widehat{\mathbf{x}})\, d\widetilde{\mathbf{x}}}{\int q(\widetilde{\mathbf{x}}, \widehat{\mathbf{x}})\, d\widetilde{\mathbf{x}}} \tag{2.8}$$

and the conditional expectation of a function $g(\mathbf{x}) = g(\widehat{\mathbf{x}}, \widetilde{\mathbf{x}})$ with respect to $\widehat{\mathbf{x}}$ is

$$E(g|\widehat{\mathbf{x}}) = \frac{\int g(\widehat{\mathbf{x}}, \widetilde{\mathbf{x}}) q(\widetilde{\mathbf{x}}, \widehat{\mathbf{x}})\, d\widetilde{\mathbf{x}}}{\int q(\widetilde{\mathbf{x}}, \widehat{\mathbf{x}})\, d\widetilde{\mathbf{x}}}. \tag{2.9}$$

So, for example, $E(\boldsymbol{\phi}(t, \widehat{\mathbf{x}}, \widetilde{\mathbf{x}})|\widehat{\mathbf{x}})$ depends on $\widehat{\mathbf{x}}$ and t, but not on $\widetilde{\mathbf{x}}$.

The operator $\mathcal{P}$ is then defined as

$$\mathcal{P}g = E(g(\mathbf{x})|\widehat{\mathbf{x}}) \tag{2.10}$$

for any function $g(\mathbf{x})$. Note that $\mathcal{P}g \neq g(\widehat{\mathbf{x}}, E(\widetilde{\mathbf{x}}|\widehat{\mathbf{x}}))$. Notice also that the operator $\mathcal{P}$ is linear:

$$\mathcal{P}(\alpha g(\mathbf{x}) + \beta h(\mathbf{x})) = \alpha \mathcal{P}g(\mathbf{x}) + \beta \mathcal{P}h(\mathbf{x})$$

and it is idempotent:

$$\mathcal{P}\left(\mathcal{P}g(\mathbf{x})\right) = \mathcal{P}g(\mathbf{x}).$$

3 A Modified Optimal Prediction Method

In this section, we present a new derivation of the t-model equation for the resolved scales as well as its modification and a higher-order version.

Let us consider here a linear equation

$$u_t = \mathcal{L}u \tag{3.1}$$

with an initial condition

$$u(t=0)=u_0 \tag{3.2}$$

and with suitable boundary conditions, which we assume to be incorporated in the form of the operator $\mathcal{L}$. The precise form of the operator $\mathcal{L}$ is of no interest for the discussion to follow. However, since it is sometimes convenient to have a specific model, the reader is encouraged to think about function u being a function of a time variable t and spatial variable $\mathbf{x}$, that is $u=u(t,\mathbf{x})$, and $\mathcal{L}$ being, for instance, the Liouville operator (2.4).

An essential assumption in our discussion is that the solution u possesses large and small scales. We are interested in resolving the large scales accurately, but in order to get the correct information we need to consider the effect of the small scales on the large one.

Let $\mathcal{P}$ be a time independent projection operator that projects (at least initially) the solution into the part that contains the large scales and let, as before, $\mathcal{Q}=I-\mathcal{P}$, which, in the same way, projects the solution into its small scales. For the discussion here $\mathcal{P}$ is a general projection operator. Multiplying both sides of (3.1) by $\mathcal{P}$ and using the fact that $\mathcal{P}^2=\mathcal{P}$, we can rewrite (3.1) as

$$(\mathcal{P}u)_t=\mathcal{P}\mathcal{L}\mathcal{P}(\mathcal{P}u)+\mathcal{P}\mathcal{L}\mathcal{Q}(\mathcal{Q}u), \tag{3.3}$$

and similarly we can obtain an equation for $\mathcal{Q}u$ in the form

$$(\mathcal{Q}u)_t=\mathcal{Q}\mathcal{L}\mathcal{Q}(\mathcal{Q}u)+\mathcal{Q}\mathcal{L}\mathcal{P}(\mathcal{P}u). \tag{3.4}$$

The first term in RHS of (3.3), $\mathcal{P}\mathcal{L}\mathcal{P}$, contains the impact of the large scales, while the second one, $\mathcal{P}\mathcal{L}\mathcal{Q}$, is the impact of the small scales on the development of the large scales and therefore should not be neglected. The RHS of (3.4) is composed of the contribution of the small scales, $\mathcal{Q}\mathcal{L}\mathcal{Q}$ and the effect of the large scales, $\mathcal{Q}\mathcal{L}\mathcal{P}$.

To derive an equation for the large scales Pu, we solve (3.4) for $\mathcal{Q}u$:

$$\mathcal{Q}u(t)=e^{\mathcal{Q}\mathcal{L}\mathcal{Q}t}(\mathcal{Q}u_0)+\int_0^t e^{\mathcal{Q}\mathcal{L}\mathcal{Q}(t-\tau)}\mathcal{Q}\mathcal{L}\mathcal{P}u(\tau)\,d\tau \tag{3.5}$$

and substitute it into (3.3) to obtain:

$$(\mathcal{P}u)_t=\mathcal{P}\mathcal{L}\mathcal{P}(\mathcal{P}u)+\mathcal{P}\mathcal{L}\mathcal{Q}e^{\mathcal{Q}\mathcal{L}\mathcal{Q}t}(\mathcal{Q}u_0)+\mathcal{P}\mathcal{L}\mathcal{Q}\int_0^t e^{\mathcal{Q}\mathcal{L}\mathcal{Q}(t-\tau)}\mathcal{Q}\mathcal{L}\mathcal{P}u(\tau)\,d\tau. \tag{3.6}$$

This formula is exact and involves only the large scales $\mathcal{P}u(t)$ at time $t>0$ ($\mathcal{Q}u_0$ depends only on the initial conditions). The integral (the *memory term*) in the RHS of (3.6) models the effect of the small scales on the large ones being an important component of the model as it was mentioned above.

Equation (3.6) cannot be solved easily and therefore should be approximated. Following a similar venue to [9], we first assume that $\mathcal{Q}u_0$ is small and the operator $e^{\mathcal{Q}\mathcal{L}\mathcal{Q}t}$ is stable (does not amplify the initial values), and thus the second term in the RHS of (3.6), $\mathcal{P}\mathcal{L}\mathcal{Q}e^{\mathcal{Q}\mathcal{L}\mathcal{Q}t}(\mathcal{Q}u_0)$, can be neglected. Neglecting this term and using the trapezoidal rule for the memory term yields:

$$(\mathcal{P}u)_t\approx\mathcal{P}\mathcal{L}\mathcal{P}(\mathcal{P}u)+\frac{t}{2}\mathcal{P}\mathcal{L}\mathcal{Q}\big[e^{\mathcal{Q}\mathcal{L}\mathcal{Q}t}\mathcal{Q}\mathcal{L}\mathcal{P}u_0+\mathcal{Q}\mathcal{L}\mathcal{P}u(t)\big]. \tag{3.7}$$

Equation (3.7) can be simplified further. To first order in t, one can approximate $e^{\mathcal{Q}\mathcal{L}\mathcal{Q}t}$ in (3.7) by I and also replace u_0 by $u(t)$ (to avoid an obvious growth in time even when the

exact solution does not grow). The approximation obtained is the aforementioned t-model (2.7).

In problems where the important effect is the impact of the large scales on the small ones (the term $\mathcal{QLP}$) and not the dynamic of the small scales themselves (described by $\mathcal{QLQ}$), the term $e^{\mathcal{QLQ}t}$ decays fast (this will be shown explicitly in Sect. 4 for our particular application). Neglecting $e^{\mathcal{QLQ}t}$ in (3.7), we obtain a modification of the t-model:

$$\xi_t = \mathcal{PLP}\xi + \frac{t}{2}\mathcal{PLQLP}\xi, \tag{3.8}$$

where, as before, by ξ we denote an approximation to $\mathcal{P}u$. We expect (2.7) to be better for short times and (3.8) to be better for longer times. This is supported by numerical experiments reported in Sect. 4.2.

To obtain a higher-order method containing t^2 (the next term in expation in t), we approximate $e^{\mathcal{QLQ}t}$ by the first order Taylor expansion $I + t\mathcal{QLQ}$ and u_0 by $u(t) - tu_t(t)$:

$$e^{\mathcal{QLQ}t}\mathcal{QLP}u_0 \sim [I + t\mathcal{QLQ}t]\mathcal{QLP}[u - tu_t]. \tag{3.9}$$

Expanding the RHS of (3.9) to first order in t and using $(\mathcal{P}u)_t \sim \mathcal{PLP}u$, yields

$$e^{\mathcal{QLQ}t}\mathcal{QLP}u_0 \sim \mathcal{QLP}u + t[\mathcal{QLQLP} - \mathcal{QLPLP}]u.$$

Finally, the t^2-approximation is

$$\xi_t = \mathcal{PLP}\xi + t\mathcal{PLQLP}\xi + \frac{t^2}{2}\mathcal{PLQL}(I - 2\mathcal{P})\mathcal{LP}\xi \tag{3.10}$$

with ξ approximating $\mathcal{P}u$.

4 Application to a Particle-Method Problem

In this section, we demonstrate how the reduction method, described in Sect. 3, can be applied for a system of ODEs that arises in numerical solutions of PDEs by particle methods. For this purpose, we consider a model problem

$$w_t + (r^2 w)_r = 0, \quad w(r,0) = w_0(r), \quad r \in \mathbb{R},\ t > 0, \tag{4.1}$$

where $w = w(r,t)$ is a function of a space variable r and time variable t. We derive a particle system of ODEs for (4.1) and show how to reduce its dimensionality by using the OP method, as well as perform a number of numerical experiments. Despite its apparent simplicity, this example offers a clear view of how the reduced model for a particle system can be obtained in general. A brief description of particle methods is given in Appendix.

The particle-method way of solving problem (4.1) is to first approximate the initial condition $w_0(r)$ by a linear combination of Dirac distributions

$$w_0(r) \approx w_0^N(r) = \sum_{i=1}^{N} \omega_i(0)\delta(r - r_i(0)) \tag{4.2}$$

for some set (r_i, ω_i) of points (particles) $r_i \in \mathbb{R}$ and their weights $\omega_i \in \mathbb{R}$ ($\omega_i(0)$ can be thought of as an integral of $w_0(r)$ over a neighborhood of the point $r_i(0)$). Then, an approximate solution w^N to (4.1) is sought in the form

$$w^N(r,t) = \sum_{i=1}^{N} \omega_i(t)\delta(r - r_i(t)), \tag{4.3}$$

where the evolution of the weights ω_i and the locations r_i is described by the system of ODEs:

$$\begin{cases} \dfrac{dr_i(t)}{dt} = r_i^2(t), \\ \dfrac{d\omega_i(t)}{dt} = 0, \end{cases} \tag{4.4}$$

with initial values $(r_i(0), \omega_i(0))$, see [18] and Appendix.

Note that the equations for each r_i are identical and independent and that the ω_is are all constant, so from now on we will omit the dependence on t for the weights, that is, $\omega_i := \omega_i(t) \equiv \omega_i(0)$ and consider the system

$$\frac{dr_i(t)}{dt} = r_i^2(t), \quad i = 1, \ldots, N, \tag{4.5}$$

for individual particles only. Equation (4.5) can be easily solved analytically and its solution is given by

$$r_i(t) = \frac{1}{r_i^{-1}(0) - t}, \quad i = 1, \ldots, N,\ t > 0, \tag{4.6}$$

so if $r_i(0) < 0$, (4.6) is valid for all t. In this particular example, the exact solution of the particle system is available, in general, however, the system of ODEs obtained after a particle discretization is to be solved numerically, and at final time, the solution $w(r,t)$ is to be recovered from the computed approximation $w^N(r,t)$ (the details are discussed in Appendix).

4.1 Derivation of the Reduced Model

Following the discussion in Sect. 3, we denote by $\mathbf{x} = (x_1, \ldots, x_N)$ the initial locations of particles, that is, $x_i = r_i(0),\ i = 1, \ldots, N$, and split the set of degrees of freedom into the large (resolved) and small (unresolved) scales. We assume that the number of particles is even, $N = 2M$, and define for each $j = 1, 2, \ldots, M$:

$$\widehat{r}_j = r_{2j-1} + r_{2j}, \tag{4.7}$$

$$\widetilde{r}_j = r_{2j-1} - r_{2j}, \tag{4.8}$$

and correspondingly $\widehat{\mathbf{r}}(0) = \widehat{\mathbf{x}}$ and $\widetilde{\mathbf{r}}(0) = \widetilde{\mathbf{x}}$. Then $\widehat{\mathbf{r}} = (\widehat{r}_1, \ldots, \widehat{r}_M)$ is the resolved degree of freedom, containing the large scales, and $\widetilde{\mathbf{r}} = (\widetilde{r}_1, \ldots, \widetilde{r}_M)$ the unresolved one, containing the small scales, which satisfy the following system of equations:

$$\begin{cases} \dfrac{d\widehat{r}_j}{dt} = \dfrac{1}{2}\left(\widehat{r}_j^2 + \widetilde{r}_j^2\right), \quad j = 1, 2, \ldots, M, \\ \dfrac{d\widetilde{r}_j}{dt} = \widehat{r}_j\widetilde{r}_j. \end{cases} \tag{4.9}$$

The Liouville operator corresponding to (4.9) is

$$\mathcal{L} = \sum_{j=1}^{M} \left[\frac{1}{2}\left(\widehat{r}_j^2 + \widetilde{r}_j^2\right) \frac{\partial}{\partial \widehat{r}_j} + \widehat{r}_j \widetilde{r}_j \frac{\partial}{\partial \widetilde{r}_j} \right]. \tag{4.10}$$

Taking into account (4.6), we now assume that the particles $x_i,\ i = 1, \ldots, N$ are initially uniformly distributed on the interval $[-1, 0]$. This puts pairs of particles (x_i, x_{i+1}) on a square in the lower left quadrant. In addition, we would like these pairs to have "neighbor relationships", so we require $|x_i - x_{i+1}| < \sigma,\ i = 1, \ldots, N-1$, for some small positive number $\sigma = O(1/N)$. This restricts the pairs (x_i, x_{i+1}) to be uniformly distributed on a hexagonal region.

The corresponding boundaries on the initial values $\widehat{x}_j$ and $\widetilde{x}_j$ then become

$$\widehat{x}_j \in [-2, 0], \quad |\widetilde{x}_j| < \alpha(\widehat{x}_j), \quad j = 1, \ldots, M, \tag{4.11}$$

with

$$\alpha(\widehat{x}_j) = \min(2 + \widehat{x}_j, -\widehat{x}_j, \sigma). \tag{4.12}$$

Following the original M-Z framework and [7, 8], we define the operator $\mathcal{P}$ as the conditional expectation in terms of $\widehat{\mathbf{x}}$. By symmetry, the first moment vanishes,

$$\mathcal{P}\widetilde{\mathbf{x}} = E(\widetilde{\mathbf{x}}|\widehat{\mathbf{x}}) = 0, \tag{4.13}$$

but the second moment, which we denote by $\boldsymbol{\beta}(\widehat{\mathbf{x}}) := E(\widetilde{\mathbf{x}}^2|\widehat{\mathbf{x}})$, is nonzero. Indeed, if $\boldsymbol{\beta}(\widehat{\mathbf{x}}) = (\beta(\widehat{x}_1), \ldots, \beta(\widehat{x}_M))$, then

$$\beta(\widehat{x}_j) = \frac{1}{2\alpha(\widehat{x}_j)} \int_{-\alpha(\widehat{x}_j)}^{\alpha(\widehat{x}_j)} \widetilde{x}_j^2 \, d\widetilde{x}_j = \frac{1}{3}\alpha^2(\widehat{x}_j), \quad j = 1, \ldots, M. \tag{4.14}$$

Let now $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_M)$ be an approximation to the large scales $\widehat{\mathbf{r}}$ and $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_M)$ to the small scales $\widetilde{\mathbf{r}}$. Using (4.10) and (4.14), the Mori-Zwanzig functions that appear in the RHS of (2.7), (3.8), and (3.10) can now be computed as follows:

$$\mathcal{L}\xi_j = \frac{1}{2}(\xi_j^2 + \eta_j^2), \tag{4.15}$$

$$\mathcal{P}\mathcal{L}\xi_j = \frac{1}{2}\left(\xi_j^2 + \beta(\xi_j)\right), \tag{4.16}$$

$$\mathcal{Q}\mathcal{L}\xi_j = \frac{1}{2}\left(\eta_j^2 - \beta(\xi_j)\right), \tag{4.17}$$

$$\mathcal{L}\mathcal{Q}\mathcal{L}\xi_j = \xi_j \eta^2 - \beta'(\xi_j)\frac{\xi_j^2 + \eta_j^2}{4}, \tag{4.18}$$

$$\mathcal{P}\mathcal{L}\mathcal{Q}\mathcal{L}\xi_j = \xi_j \beta(\xi_j) - \beta'(\xi_j)\frac{\xi_j^2 + \beta(\xi_j)}{4}. \tag{4.19}$$

Substituting (4.15)–(4.19) into (2.7), (3.8), and (3.10) for each $j = 1, \ldots, M$, we obtain the following t-model

$$\frac{d\xi_j}{dt} = \frac{1}{2}\left(\xi_j^2 + \beta(\xi_j)\right) + t\left(\xi_j \beta(\xi_j) - \frac{1}{4}\beta'(\xi_j)\left(\xi_j^2 + \beta(\xi_j)\right)\right), \tag{4.20}$$

the modified t-model

$$\frac{d\xi_j}{dt} = \frac{1}{2}\big(\xi_j^2 + \beta(\xi_j)\big) + \frac{t}{2}\left(\xi_j\beta(\xi_j) - \frac{1}{4}\beta'(\xi_j)(\xi_j^2 + \beta(\xi_j))\right), \tag{4.21}$$

and the t^2-model

$$\begin{aligned}\frac{d\xi_j}{dt} = {} & \frac{1}{2}\big(\xi_j^2 + \beta(\xi_j)\big) + t\left(\xi_j\beta(\xi_j) - \frac{1}{4}\beta'(\xi_j)\big(\xi_j^2 + \beta(\xi_j)\big)\right) \\ & + \frac{t^2}{2}\left[\big(\xi_j - \beta'(\xi_j)\big)\left(\xi_j\beta(\xi_j) - \frac{1}{4}\beta'(\xi_j)\big(\xi_j^2 + \beta(\xi_j)\big)\right) + \frac{1}{5}\beta^2(\xi_j)(1 - \beta''(\xi_j))\right].\end{aligned} \tag{4.22}$$

respectively.

Remark It should be observed that, in the case under consideration, $e^{\mathcal{Q}\mathcal{L}\mathcal{Q}t}$ decays with time, as it has been predicted in Sect. 3. Indeed, note that $\mathcal{Q}\mathcal{L}\mathcal{Q}\eta = \xi\eta$ and therefore $\frac{d\eta}{dt} = \mathcal{Q}\mathcal{L}\mathcal{Q}\eta = \xi\eta$. Since $\xi < 0$, then $\eta = e^{\mathcal{Q}\mathcal{L}\mathcal{Q}t}\eta(0)$ decreases as time grows.

4.2 Numerical Examples

The goal of this section is to demonstrate the performance of the three different reduction schemes derived in Sect. 4.1 for the particle method problem (4.1), (4.4).

To numerically compute the approximation $\boldsymbol{\xi}$, we integrate (4.20), (4.21), and (4.22) in time using the standard fourth-order Runge-Kutta ODE solver. For the initial condition, we took $x_j = x_{j-1} + \Delta x_j$, where Δx_j is a random number between 0 and $1/N$, $j = 1, \ldots, N$, and N is the total number of particles. Tables 1 and 2 show a convergence study for the t-model, the modified t-model and the t^2-model. The relative L_2-error at times $t = 30$ (Table 1) and $t = 60$ (Table 2) and the convergence rates are presented for each one of the models. The convergence rate is computed as

$$\ln\left(\frac{\|\widehat{\mathbf{r}} - \boldsymbol{\xi}^{N_1}\|_2}{\|\widehat{\mathbf{r}} - \boldsymbol{\xi}^{N_2}\|_2}\right) \Big/ \ln\left(\frac{N_2}{N_1}\right),$$

where $\widehat{\mathbf{r}}$ is the exact solution, $\boldsymbol{\xi}^N$ is the numerical solution for the system of N particles, and $\|\widehat{\mathbf{r}} - \boldsymbol{\xi}^N\|_2$ is the discrete L_2-norm of the error. As one can observe, the relative errors for the modified t-model and the t^2-model are smaller than the corresponding errors for the original t-model, but in all the models, the computed convergence rate is 2.

Table 1 Relative L_2-error computed at time $t = 30$

Number of particles	t-model		Modified t-model		t^2-model	
	Error	Rate	Error	Rate	Error	Rate
$N = 100$	7.519×10^{-3}		2.380×10^{-3}		2.065×10^{-3}	
$N = 200$	1.873×10^{-3}	2.00	5.942×10^{-4}	2.00	5.185×10^{-4}	1.99
$N = 400$	4.679×10^{-4}	2.00	1.484×10^{-4}	2.00	1.296×10^{-4}	2.00
$N = 800$	1.170×10^{-4}	2.00	3.712×10^{-5}	2.00	3.247×10^{-5}	2.00
$N = 1600$	2.924×10^{-5}	2.00	9.280×10^{-6}	2.00	8.117×10^{-6}	2.00

Table 2 Relative L_2-error computed at time $t = 60$

Number of particles	t-model		Modified t-model		t^2-model	
	Error	Rate	Error	Rate	Error	Rate
$N = 100$	3.054×10^{-2}		9.856×10^{-3}		8.927×10^{-3}	
$N = 200$	7.530×10^{-3}	2.02	2.446×10^{-2}	2.01	2.267×10^{-3}	1.98
$N = 400$	1.876×10^{-3}	2.00	6.102×10^{-4}	2.00	5.687×10^{-4}	2.00
$N = 800$	4.686×10^{-4}	2.00	1.525×10^{-4}	2.00	1.424×10^{-4}	2.00
$N = 1600$	1.171×10^{-4}	2.00	3.813×10^{-5}	2.00	3.562×10^{-5}	2.00

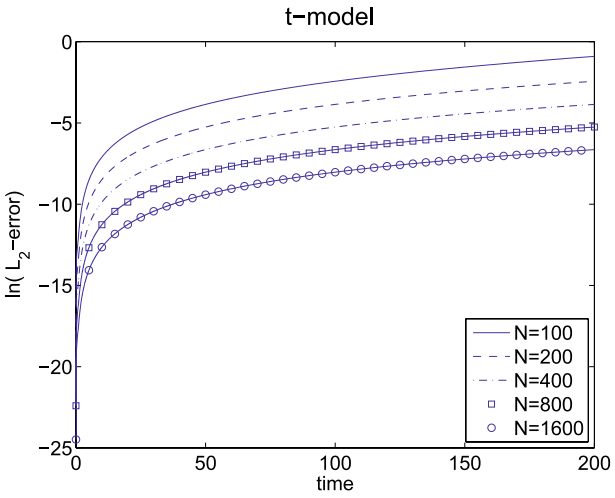

Fig. 1 The relative L_2-error as a function of time for the t-model

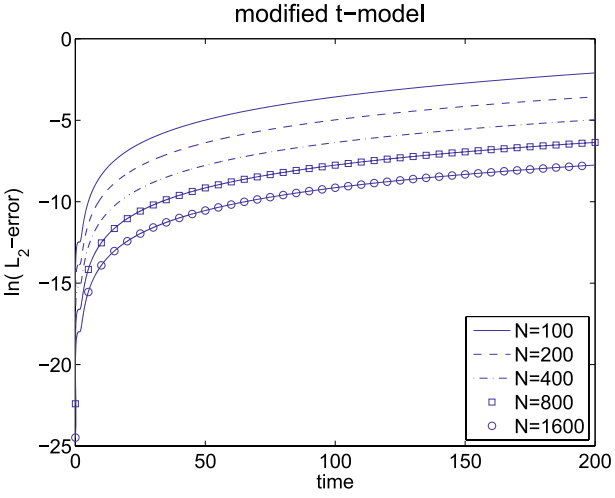

Fig. 2 The relative L_2-error as a function of time for the modified t-model

In order to highlight the difference in the quality of the obtained results, we also plot in Figs. 1, 2, and 3 the relative error as a function of time for the t-model, the modified t-model

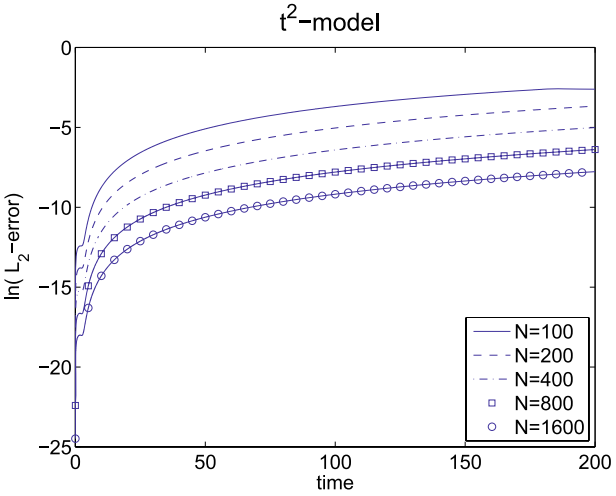

Fig. 3 The relative L_2-error as a function of time for the t^2-model

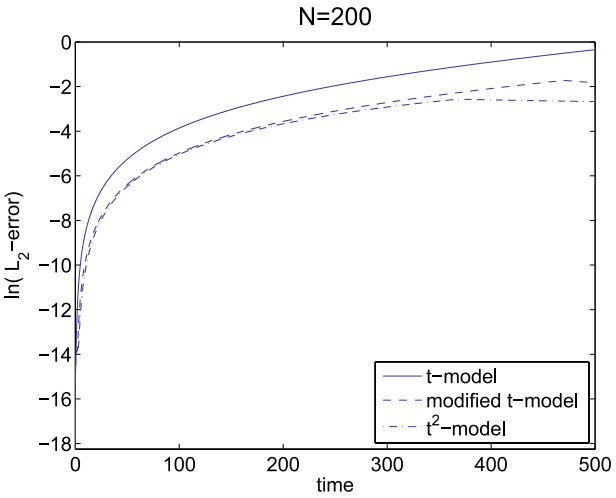

Fig. 4 The relative L_2-error as a function of time for three different models

and the t^2-model, respectively. For comparison, in Fig. 4, we show the relative errors computed with $N = 200$ particles for all three reduced models. As one can see, both the modified t-model and the t^2-model outperform the original t-model in a long time integration, while for short times the original t-model is slightly better than the modified one.

Acknowledgements We would like to thank Prof. A.J. Chorin for useful discussions and for bringing the optimal prediction method to our attention.

The research of A. Chertock was supported in part by NSF Grants DMS-0410023 and DMS-0712898. The research of D. Gottlieb was supported in part by AFOSR Grants FA9550-05-1-0123 and FA9550-07-1-0544, DOE Grant DE-FG02-98ER25346 and NSF Grant DMS-0510345. The research of A. Solomonoff was supported in part by AFOSR Grants FA9550-05-1-0123 and FA9550-07-1-0544 and DARPA/AFOSR Grant FA9550-05-0108.

Appendix: Particle Methods—An Overview

Here we briefly describe a particle method for a one-dimensional (1-D) linear transport equations. For more detail derivation, we refer the reader to [18].

Let us consider an equation of the form

$$w_t + (a(r,t)w)_r + a_0(r,t)w = f(r,t), \quad r \in \mathbb{R},\ t > 0 \tag{A.1}$$

subject to an initial condition:

$$w(r,0) = w_0(r), \quad r \in \mathbb{R}. \tag{A.2}$$

Here w is a function of a time variable t and spatial variable r and a, a_0, f and w_0 are given functions. We look for the solution of the initial-value problem (A.1), (A.2) as a linear combination of Dirac distributions,

$$w^N(r,t) = \sum_{i=1}^{N} \omega_i(t)\delta(r - r_i(t)), \tag{A.3}$$

where N is a total number of particles, $r_i(t)$ and $\omega_i(t)$ are the location and the weight of the ith particle at time t, respectively.

Considering a weak formulation of the problem and substituting (A.3) into (A.1)–(A.2) results in the following system of ODEs for the locations of particles, $r_i(t)$, and their weights, $\omega_i(t)$:

$$\begin{cases} \dfrac{dr_i(t)}{dt} = a(r_i(t),t), \\ \dfrac{d\omega_i(t)}{dt} + a_0(r_i(t),t)\omega_i(t) = \gamma(r_i(t),t), \end{cases} \tag{A.4}$$

where $\gamma_i(t)$ reflects the contribution of the source term f. The initial positions of the particles, $r_i(0)$, and the weights, $\omega_i(0)$, are chosen to give a high-order approximation to the initial datum according to (A.3). The latter can be done, for instance, in the sense of measures on $\mathbb{R}$. Given a test function $\psi \in C_0^0(\mathbb{R})$, the inner product, $(w_0(r), \psi(r))$, should be approximated by

$$(w^N(r), \psi(r)) = \int_{\mathbb{R}} u_0(r)\psi(r)\,dr \approx \sum_{i=1}^{N} \omega_i(0)\psi(r_i(0)).$$

In other words, the constants $\{\omega_i(0)\}$, should be determined by solving the standard numerical quadrature problem. For example, a midpoint quadrature is then given by setting $\omega_i(0) = w_0(r_i(0))\Delta r$, where $\Delta r > 0$ is an initial uniform distance between the particles.

The system (A.4) should, in general, be solved numerically and, the final time, the solution $w(r,t)$ is to be recovered from its particle approximation. A commonly used way of computing the point values of the numerical solution is a regularization of (A.3), which is usually performed by a convolution product with so-called "cut-off function" $\zeta(r)$ that after a proper scaling takes into account the initial tightness of the particle discretization, namely

$$w_\varepsilon^N(r,t) = (w^N * \zeta_\varepsilon)(r,t) = \sum_{i=1}^{N} \omega_i(t)\zeta_\varepsilon(r - r_i(t)), \tag{A.5}$$

and the function $\zeta_\varepsilon(r)$ is taken as a smooth approximation of the δ-function which satisfies

$$\zeta_\varepsilon(r) = \frac{1}{\varepsilon}\zeta\left(\frac{r}{\varepsilon}\right) \quad \text{and} \quad \int_{\mathbb{R}} \zeta(r)\,dr = 1. \tag{A.6}$$

There is an extensive discussion in the literature on the selection of the cut-off function and its relation to the accuracy of particle methods, see, e.g. [6, 10, 15, 18] and references therein. We also refer the reader to [2–5], where several different strategies for recovering point values of the solution were suggested.

References

1. Bernstein, D.: Optimal prediction for Burgers equations. Multi. Mod. Sim. **6**(1), 27–52 (2007)
2. Chertock, A., Kurganov, A.: On a hybrid finite-volume particle method, M2AN. Math. Model. Numer. Anal. **38**, 1071–1091 (2004)
3. Chertock, A., Kurganov, A.: On a practical implementation of particle methods. Appl. Numer. Math. **56**, 1418–1431 (2006)
4. Chertock, A., Kurganov, A.: Computing multivalued solutions of pressureless gas dynamics by deterministic particle methods. Commun. Comput. Phys. **5**, 565–581 (2009)
5. Chertock, A., Kurganov, A., Petrova, G.: Finite-volume-particle methods for models of transport of pollutant in shallow water. J. Sci. Comput. **27**, 189–199 (2006)
6. Chorin, A.J.: Numerical study of slightly viscous flow. J. Fluid Mech. **57**(4), 785–796 (1973)
7. Chorin, A.J., Hald, O.H.: Stochastic Tools in Mathematics and Science. Springer, New York (2005)
8. Chorin, A.J., Hald, O.H., Kupferman, R.: Optimal prediction with memory. Phys. D **166**(3–4), 239–257 (2002)
9. Chorin, A.J., Stinis, P.: Problem reduction, renormalization, and memory. Commun. Appl. Math. Comput. Sci. **1**, 1–27 (2005)
10. Cottet, G.-H., Koumoutsakos, P.D.: Vortex Methods. Cambridge University Press, Cambridge (2000)
11. Fick, E., Sauermann, G.: In: The Quantum Statistics of Dynamic Processes. Springer Series in Solid-State Sciences, vol. 86. Springer, Berlin (1990)
12. Givon, D., Hald, O.H., Kupferman, R.: Existence proof for orthogonal dynamics and the Mori-Zwanzig formalism. Israel J. Math. **199**, 279–316 (2004)
13. Grabert, H.: In: Projection Operator Techniques in Nonequilibrium Statistical Mechanics. Springer Series in Solid-State Sciences, vol. 95. Springer, Berlin, New York (1982)
14. Hald, O.H., Stinis, P.: Optimal prediction and the rate of decay of solutions of the Euler equations in two and three dimensions. Proc. Natl. Acad. Sci. **104**(16), 6527–6532 (2007)
15. Hald, O.H.: Convergence of Vortex Methods, Vortex Methods and Vortex Motion, pp. 33-58. SIAM, Philadelphia (1991)
16. Mori, H.: Transport, collective motion and Brownian motion. Prog. Theor. Phys. **33**, 423–450 (1965)
17. Nordholm, S., Zwanzig, R.: A systematic derivation of exact generalized Brownian motion theory. J. Stat. Phys. **13**, 347–371 (1975)
18. Raviart, P.-A.: In: An Analysis of Particle Methods, Numerical Methods in Fluid Dynamics (Como, 1983). Lecture Notes in Math., vol. 1127, pp. 243–324. Springer, Berlin (1985)
19. Stinis, P.: Higher order Mori-Zwanzig models for the Euler equations. Tech. report, Technical Report LBNL-60899 (2006)
20. Taylor, G.I., Green, A.E.: Mechanism of the production of small eddies from large ones. Proc. R. Soc. Lond. Ser. A **158**, 499 (1937)
21. Zwanzig, R.: Nonlinear generalized Langevin equation. J. Stat. Phys. **9**, 215–220 (1973)