A MICRO/MACRO ALGORITHM TO ACCELERATE MONTE CARLO SIMULATION OF STOCHASTIC DIFFERENTIAL EQUATIONS

KRISTIAN DEBRABANT* AND GIOVANNI SAMAEY†

Abstract. We present and analyze a micro/macro acceleration technique for the Monte Carlo simulation of stochastic differential equations (SDEs) in which there is a separation between the (fast) time-scale on which individual trajectories of the SDE need to be simulated and the (slow) time-scale on which we want to observe the (macroscopic) function of interest. The method performs short bursts of microscopic simulation using an ensemble of SDE realizations, after which the ensemble is restricted to a number of macroscopic state variables. The resulting macroscopic state is then extrapolated forward in time and the ensemble is projected onto the extrapolated macroscopic state. We relate the algorithm to existing analytical and numerical closure approximations and provide a first analysis of its convergence in terms of extrapolation time step and number of macroscopic state variables. The effects of the different approximations on the resulting error are illustrated via numerical experiments.

Key words. accelerated Monte Carlo simulation, micro/macro simulation, multiscale simulation, weak approximation, stochastic differential equation

AMS subject classifications. 65C30, 60H35, 65C05, 65C20, 68U20

1. Introduction.

In many applications, one considers a process that is modeled as a stochastic differential equation (SDE), while one is only interested in the time evolution of the expectation of a certain function of the state, i.e., in weak approximation. Consider, for instance, the micro/macro simulation of dilute solutions of polymers [36]. Here, an SDE models the evolution of the configuration of an individual polymer driven by the flow field, and the function of interest is a non-Newtonian stress tensor (the expectation of a function of the polymer configuration). For this type of problem, one often resorts to Monte Carlo simulation [6], i.e., the simulation of a large ensemble of realizations of the SDE, combined with ensemble averaging to obtain an approximation of the quantity of interest at the desired moments in time. For concreteness, we introduce the SDE

$$d\boldsymbol{X}(t) = \boldsymbol{a}(t, \boldsymbol{X}(t))\ dt + \boldsymbol{b}(t, \boldsymbol{X}(t)) \star d\boldsymbol{W}(t), \quad t \in I := [t^0, T], \quad \boldsymbol{X}(t^0) = \boldsymbol{X}_0, \tag{1.1}$$

in which $\boldsymbol{a} : I \times \mathbb{R}^d \to \mathbb{R}^d$ is the drift, $\boldsymbol{b} : I \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$ is the diffusion, and $\boldsymbol{W}(t)$ is an m-dimensional Wiener process. The initial value $\boldsymbol{X}_0$ is independent of $\boldsymbol{W}$ and follows some known distribution $\varphi_0(\boldsymbol{X})$. As usual, (1.1) is an abbreviation of the integral form

$$\boldsymbol{X}(t) = \boldsymbol{X}_0 + \int_{t^0}^{t} \boldsymbol{a}(s, \boldsymbol{X}(s))\ ds + \int_{t^0}^{t} \boldsymbol{b}(s, \boldsymbol{X}(s)) \star d\boldsymbol{W}(s), \quad t \in I.$$

The integral with respect to $\boldsymbol{W}$ can be interpreted, e.g., as an Itô integral with $\star\ d\boldsymbol{W}(s) = d\boldsymbol{W}(s)$ or as a Stratonovich integral with $\star\ d\boldsymbol{W}(s) = \circ\ d\boldsymbol{W}(s)$. The function of interest for the Monte Carlo simulation is defined as the expectation E of a function $\boldsymbol{f}(\boldsymbol{X}(t))$,

*Department of Computer Science, K.U. Leuven, Celestijnenlaan 200A, 3001 Leuven, Belgium (Kristian.Debrabant@cs.kuleuven.be).

†Department of Computer Science, K.U. Leuven, Celestijnenlaan 200A, 3001 Leuven, Belgium (Giovanni.Samaey@cs.kuleuven.be). GS is a Postdoctoral Fellow of the Research Foundation – Flanders (FWO).

$$\bar{\boldsymbol{f}}(t) = \mathrm{E}\, \boldsymbol{f}(\boldsymbol{X}(t)).$$

The numerical properties of Monte Carlo simulations have been analyzed extensively in the literature. We mention studies on the order of weak convergence of explicit [7, 12, 33, 34, 44, 52] and implicit [8, 9, 11, 33, 35] time discretizations of SDE (1.1), the investigation of stability [22, 23, 25, 53, 60], and techniques for variance reduction [19, 20, 48]. For more references, we refer to [33]. Also for strong approximation, there has been a growing interest in the study of numerical methods for stiff SDEs [1–5, 33, 45, 46, 58, 59].

In this paper, we present and analyze a *micro/macro* acceleration technique for the Monte Carlo simulation of SDEs of the type (1.1) in which there exists a time-scale separation between the (fast) time-scale on which individual trajectories of the SDE need to be simulated and the (slow) time-scale on which the function $\bar{\boldsymbol{f}}(t)$ evolves. We use the simulation of a dilute polymer solution as an illustrative example. The *microscopic* level is defined via an ensemble $\boldsymbol{\mathcal{X}} \equiv (\boldsymbol{X}_j)_{j=1}^J$ of J realizations evolving according to equation (1.1); the *macroscopic* level will be defined by a set of L *macroscopic state variables* $\boldsymbol{U} \equiv (U_l)_{l=1}^L$, with $U_l(t) = \mathrm{E}\, u_l(\boldsymbol{X}(t))$, for some appropriately chosen functions u_l. The method exploits the separation in time scales by combining short bursts of *microscopic* simulation with the SDE (1.1) with a *macroscopic* extrapolation step, in which only the macroscopic state $\boldsymbol{U}$ is extrapolated forward in time. One time step of the algorithm can be written as follows: (1) microscopic *simulation* of the ensemble using the SDE (1.1); (2) *restriction*, i.e., extraction of (an estimate of) the macroscopic state (or macroscopic time derivative); (3) forward in time *extrapolation* of the macroscopic state; and (4) *projection* of the ensemble that was available at the end of the microscopic simulation onto the extrapolated macroscopic state. Remark that the resulting method is fully explicit as soon as the microscopic simulation is explicit, and that the method can readily be implemented as a higher order method by an appropriate choice of the extrapolation.

The proposed method is motivated by the development of recent generic multiscale techniques, such as *equation-free* [31,32] and *heterogeneous multiscale* methods [14,15]. In these methods, one makes the modeling assumption that a model describing the evolution of the chosen macroscopic state variables exists, but cannot be obtained in closed form. For instance, in coarse projective integration [18, 31, 32], a *coarse time-stepper* is constructed to emulate the evolution of this unavailable macroscopic model. One coarse time step involves: (1) *lifting*, i.e., the creation of an ensemble of appropriate initial conditions for the microscopic model, conditioned upon the current macroscopic state; (2) microscopic *simulation* using the SDE (1.1); and (3) *restriction* to a macroscopic state or time derivative. The macroscopic state is then extrapolated forward in time, after which the procedure is repeated. In the heterogeneous multiscale method (HMM), the equivalent of the lifting, resp. restriction, operator is called reconstruction, resp. compression.

In the equation-free/HMM setting, the definition of the lifting/reconstruction operator is strongly problem-dependent; see, e.g., [62, 64, 67] for singularly perturbed systems of ordinary differential equations (ODEs), or [43, 61] for a lattice Boltzmann method. To regard the coarse time-stepper as a time discretization of an (unavailable) closed macroscopic model, the result of the lifting operator should only depend on the desired macroscopic state, not on the previous microscopic state. In the context of Monte Carlo simulation of SDEs, using the ensemble $\boldsymbol{\mathcal{X}}$ of J realizations evolving according to (1.1), the lifting operator maps a macroscopic state $\boldsymbol{U}$ to a corresponding state for each of the realizations. In [54], a constrained simulation was introduced that

evolves an ensemble of SDE realizations under the constraint that the macroscopic state variables remain constant. This constrained simulation was then performed until the ensemble reached an equilibrium distribution; the corresponding microscopic state was taken as the result of the lifting.

Due to the constrained simulation, part (if not all) of the computational gain that is obtained in coarse projective integration by extrapolating forward in time is lost during each lifting step. The micro/macro acceleration technique of this paper reduces the computational cost by replacing the lifting operator by a projection operator; i.e., the constrained simulation is eliminated. As a consequence, however, the interpretation of the method as a discretization of an unavailable closed macroscopic equation is lost. The purpose of the present paper is to identify all sources of numerical error of the resulting method for the Monte Carlo simulation of SDEs of the type (1.1), and to analyze convergence. The main contributions are the following:

- From a numerical analysis viewpoint, we study convergence of the proposed micro/macro acceleration method in the absence of statistical error. Specifically, we discuss how the error that is introduced during the projection depends on the number of macroscopic state variables, and how the deterministic error depends on the extrapolation time step. Additionally, we also comment on the effects of the extrapolation on the statistical error.
- From a practical viewpoint, we provide numerical results for a nontrivial test case, showing the interplay between the different sources of numerical error. We illustrate the effects of the choice of macroscopic state variables, as well as the dependence of the numerical error on the chosen extrapolation strategy.

A stability analysis of the proposed method will be given in a separate publication.

For singularly perturbed systems of ODEs, it has already been proposed to only project the last available microscopic state onto the extrapolated macroscopic state, see, e.g., [16]. In that work, however, one still assumes that a closed macroscopic equation exists, and uses a strong time-scale separation to show quick relaxation (healing) of the projected microscopic state towards its conditional equilibrium. In this paper, we do not require such a strong time-scale separation, nor do we make use of the existence of an approximate macroscopic model.

The remainder of the paper is organized as follows. In Section 2, we first introduce the mathematical setting of the paper. We discuss the necessary assumptions on the SDE (1.1) for our Monte Carlo setting, and introduce the illustrative example that will be used for the numerical experiments. As an alternative to Monte Carlo simulation, one might try to approximate the evolution of the function of interest via a low-dimensional set of closed ODEs for a few macroscopic state variables. Section 3 overviews the recent closure approximations (both analytical and numerical). Subsequently, in Section 4, we propose a modified algorithm that will be the focus of this paper. Section 5 provides some results on the projection operator, whereas the extrapolation operator is discussed in Section 6. We provide a general convergence result in Section 7. Section 8 provides numerical illustrations, which are chosen to illuminate the properties of the proposed method. We conclude in Section 9, where we also outline some directions for future research.

2. Mathematical setting. Let us first introduce in detail the notations that we will use (Subsection 2.1), as well as the illustrative example that will be considered throughout the paper (Subsection 2.2).

2.1. Notations. We first define the appropriate function spaces. Let $C^r_P(\mathbb{R}^d, \mathbb{R})$ denote the space of all $g \in C^r(\mathbb{R}^d, \mathbb{R})$ fulfilling that there are constants $\tilde{C} > 0$ and

$\kappa > 0$ such that $|\partial^i_{\boldsymbol{x}} g(\boldsymbol{x})| \leq \tilde{C}(1 + \|\boldsymbol{x}\|^\kappa)$ for any partial derivative of order $i \leq r$ and all $\boldsymbol{x} \in \mathbb{R}^d$. Further, let $g \in C_P^{q,r}(I \times \mathbb{R}^d, \mathbb{R})$ if $g(\cdot, \boldsymbol{x}) \in C^q(I, \mathbb{R})$, $g(t, \cdot) \in C^r(\mathbb{R}^d, \mathbb{R})$ for all $t \in I$ and $\boldsymbol{x} \in \mathbb{R}^d$, and $|\partial^i_{\boldsymbol{x}} g(s, \boldsymbol{x})| \leq \tilde{C}(1 + \|\boldsymbol{x}\|^\kappa)$ holds for $0 \leq i \leq r$ uniformly with respect to $s \in [t^0, t]$ and for all $\boldsymbol{x} \in \mathbb{R}^d$ [33, 44].

We consider the SDE (1.1). Besides the exact solution $\boldsymbol{X}(t)$ of SDE (1.1) starting from $\boldsymbol{X}(t^0) = \boldsymbol{X}_0$, we also introduce the exact solution of an auxiliary initial value problem for the SDE: the solution of the SDE starting from an initial value $\tilde{\boldsymbol{x}}$ at time $\tilde{t}$ will be denoted as $\boldsymbol{X}^{\tilde{t},\tilde{\boldsymbol{x}}}$. With this notation, we state the following definition:

Definition 2.1 (Uniform weak continuity of the SDE). *Consider a class of random variables. The SDE* (1.1) *is called uniformly weakly continuous for this class if, for all $g \in C_P^{0,2}(I \times \mathbb{R}^d, \mathbb{R})$, there exist constants $\Delta t_0 > 0$ and C such that*

$$|\operatorname{E} g\big(s, \boldsymbol{X}^{t^-,\boldsymbol{Z}}(t^- + \Delta t)\big) - \operatorname{E} g(s, \boldsymbol{Z})| \leq C\Delta t \tag{2.1}$$

holds for all initial values $\boldsymbol{Z}$ in the considered class, $\Delta t \in [0, \Delta t_0]$, and all $t^- \in [t^0, T - \Delta t]$, $s \in I$.

Next, we discretize equation (1.1) in time with step size δt and denote the numerical approximation $\boldsymbol{Y}^k = \boldsymbol{Y}(t^k) \approx \boldsymbol{X}(t^k)$ with $t^k = t^0 + k\delta t$. Leaving the concrete choice of discretization undecided for now, we introduce a short-hand notation for an abstract one step discretization scheme,

$$\boldsymbol{Y}^{k+1} = \phi(t^k, \boldsymbol{Y}^k; \delta t), \qquad \boldsymbol{Y}^0 = \boldsymbol{X}_0, \qquad k \geq 0. \tag{2.2}$$

In the Monte Carlo setting, we are interested in weak approximation of the SDE (1.1). We define the weak order of consistency of the discretization ϕ as follows (compare [44]):

Definition 2.2 (Weak consistency of SDE discretization). *Assume that for all $g \in C_P^{0,2(p_\phi+1)}(I \times \mathbb{R}^d, \mathbb{R})$ there exists a $C_g \in C_P^0(\mathbb{R}^d, \mathbb{R})$ such that*

$$|\operatorname{E} g\big(s, \phi(t, \boldsymbol{x}; \delta t)\big) - \operatorname{E} g\big(s, \boldsymbol{X}^{t,\boldsymbol{x}}(t + \delta t)\big)| \leq C_g(\boldsymbol{x})\, \delta t^{p_\phi+1}$$

is valid for $\boldsymbol{x} \in \mathbb{R}^d$, $s \in I$, and t, $t + \delta t \in [t^0, T]$. Then the one step method ϕ is called weakly consistent of order p_ϕ.

Since we can only simulate a finite number of realizations of approximations of (1.1) (via its discretization (2.2)), we also need to approximate the expectation E by an empirical mean $\widehat{\operatorname{E}}$. Using the ensemble $\boldsymbol{\mathcal{Y}} \equiv (\boldsymbol{Y}_j)_{j=1}^J$ of realizations of (2.2), the empirical mean $\widehat{\operatorname{E}}$ is defined as

$$\widehat{\operatorname{E}} f(\boldsymbol{\mathcal{Y}}) = \frac{1}{J} \sum_{j=1}^{J} f(\boldsymbol{Y}_j).$$

The numerical integration scheme for the ensemble $\boldsymbol{\mathcal{Y}}$ will be denoted as

$$\boldsymbol{\mathcal{Y}}^{k+1} = \phi_{\boldsymbol{\mathcal{Y}}}(t^k, \boldsymbol{\mathcal{Y}}^k; \delta t) = \Big(\phi(t^k, \boldsymbol{\mathcal{Y}}_j^k; \delta t)\Big)_{j=1}^J. \tag{2.3}$$

The total error of a Monte Carlo simulation of (1.1) consists of a *deterministic* error due to the time discretization, and a *statistical* error due to the finite number of realizations. In general, for a given tolerance, a Monte Carlo simulation will be most efficient if the statistical and deterministic error are balanced. However, as for stiff systems of ODEs, one might encounter situations in which the time step δt cannot be

increased because of stability problems. The acceleration method that we will propose in Section 4 will prove to be particularly useful to accelerate Monte Carlo simulations in such situations.

REMARK 2.3 (Probability density functions). *One can equivalently describe the process* (1.1) *via an advection-diffusion equation, also known as Fokker–Planck equation (see, e.g., [51]), which in the Itô case takes the form*

$$\partial_t \varphi = -\nabla_{\boldsymbol{X}} \left(\boldsymbol{a}\, \varphi \right) + \frac{1}{2} \nabla_{\boldsymbol{X}} \cdot \left[\nabla_{\boldsymbol{X}} \cdot \left(\boldsymbol{b}^T \boldsymbol{b}\, \varphi \right) \right], \tag{2.4}$$

and which describes the evolution of the probability density function $\varphi(t, \boldsymbol{x})$ *of* $\boldsymbol{X}(t)$*, starting from the initial density* $\varphi(t^0, \boldsymbol{x}) = \varphi_0(\boldsymbol{x})$*. The functional of interest then becomes*

$$\bar{\boldsymbol{f}}(t) = \int \boldsymbol{f}(\boldsymbol{x}) \varphi(t, \boldsymbol{x}) d\boldsymbol{x}. \tag{2.5}$$

For simulation purposes, however, the Monte Carlo algorithm is generally preferred, due to the possibly high number of dimensions of the Fokker–Planck equation.

REMARK 2.4 (Spatial dimension). *While the model* (1.1) *can have arbitrary dimension, we will consider a one-dimensional version in the numerical illustrations for ease of visualization. In that case the state vector reduces to a scalar,* $X(t)$*. Whenever we consider the one-dimensional case, all bold typesetting in equation* (1.1) *will be removed.*

2.2. A motivating model problem: FENE dumbbells. To illustrate the behavior of the proposed numerical methods, we will consider the micro/macro simulation of the evolution of immersed polymers in a solvent. Here, one models the evolution of the configuration of a polymer ensemble via an SDE of the type (1.1), driven by the flow field, for each of the individual polymers. This results in a polymer configuration distribution at each spatial point. This microscopic model is coupled to a Navier–Stokes equation for the solvent, in which the effect of the immersed polymers is taken into account via a non-Newtonian stress tensor. We refer to [26,36,38] for an introduction to the literature on this subject.

In this paper, we consider only the Monte Carlo simulation of the microscopic model, leaving the coupling with the Navier–Stokes equations for future work. We eliminate the spatial dependence by considering the microscopic equations along characteristics, i.e., in a Lagrangian frame. In general, a microscopic model describes an individual polymer as a series of beads, connected by nonlinear springs, resulting in a coupled system of SDEs for the position of each of the beads. In the simplest case, that we will also use as an illustrative example here, one represents the polymers as non-interacting dumbbells, connecting two beads by a spring that models intramolecular interaction. The state of the polymer chain is described by the end-to-end vector $\boldsymbol{X}(t)$ that connects both beads, and whose evolution is modelled using the non-dimensionalised SDE

$$d\boldsymbol{X}(t) = \left[\boldsymbol{\kappa}(t)\, \boldsymbol{X}(t) - \frac{1}{2\mathrm{We}} \boldsymbol{F}\big(\boldsymbol{X}(t)\big) \right] dt + \frac{1}{\sqrt{\mathrm{We}}} d\boldsymbol{W}(t), \tag{2.6}$$

where $\boldsymbol{\kappa}(t)$ is the velocity gradient of the solvent, We is the Weisenberg number, and $\boldsymbol{F}$ is a spring force, here considered to be finitely extensible nonlinearly elastic (FENE),

$$\boldsymbol{F}(\boldsymbol{X}) = \frac{\boldsymbol{X}}{1 - \|\boldsymbol{X}\|^2/\gamma}, \tag{2.7}$$

with γ a non-dimensional parameter that is related to the maximal polymer length. The resulting non-Newtonian stress tensor is given by the Kramers' expression,

$$\boldsymbol{\tau}_p(t) = \frac{\epsilon}{\mathrm{We}}\left(\mathrm{E}\left(\boldsymbol{X}(t) \otimes \boldsymbol{F}\big(\boldsymbol{X}(t)\big)\right) - \mathbf{Id}\right), \tag{2.8}$$

in which ϵ represents the ratio of polymer and total viscosity; see [38] for details and further references. This model, which is of the type (1.1), takes into account Stokes drag (due to the solvent velocity field), intramolecular elastic forces, and Brownian motion (due to collisions with solvent molecules). The functional of interest in the Monte Carlo simulation is $\bar{\boldsymbol{f}}(t) = \boldsymbol{\tau}_p(t)$.

Equations (2.6) and (2.7) ensure that the length of the end-to-end vector, $\|\boldsymbol{X}\|$, cannot exceed the maximal value $\sqrt{\gamma}$ [29]. However, a naive explicit discretization scheme might yield spring lengths beyond this maximal value. This can be avoided via an accept-reject strategy, see, e.g., [49, Section 4.3.2]. Here, for each polymer, the state after each time step is rejected if the calculated polymer length exceeds $\sqrt{(1-\sqrt{\delta t})\gamma}$, and a new random number is tried until acceptance. To prevent the distribution of the approximation process to be heavily influenced by this, the microscopic time-step has to be chosen small enough. Alternatively, one can use an implicit method [49], which alleviates the time-step restriction. However, even for implicit SDE discretizations, the maximal time step is limited when coupling the Monte Carlo simulation with a discretization of the Navier–Stokes equations for the solvent. This is due to the fact that the coupling between the Monte Carlo and Navier–Stokes parts is, in most existing work, done explicitly in time, creating an additional stability constraint due to the coupling. (Some notable exceptions are given in [37, 56].) For this coupled simulation, one would extrapolate both the Monte Carlo and the Navier–Stokes part simultaneously. This will be done in future work. Here, we simply conclude that, for this model problem, the required time step for a stable SDE (or coupled) simulation is indeed small compared to the time scale of the evolution of the stress.

As in, e.g., [30], we will consider a one-dimensional version in the numerical illustrations; the stress tensor then reduces to a scalar $\tau_p(t)$. As time discretization, we will use the explicit Euler–Maruyama scheme, combined with an accept-reject strategy.

3. Macroscopic closure approximations.

3.1. Analytical closure approximations. Due to the possibly high computational cost of Monte Carlo simulation, another route has been followed in the literature, in which one derives an approximate macroscopic model to describe the system; see, e.g., [24, 27, 30, 41, 42, 55] for derivations of macroscopic closures for FENE dumbbell models. In these approaches, one considers a number L of macroscopic state variables, $\boldsymbol{U} = (U_l)_{l=1}^{L}$, which are defined as expectations of scalar functions u_l of the state $\boldsymbol{X}$ and time t,

$$U_l(t) = \mathrm{E}\, u_l\big(t, \boldsymbol{X}(t)\big). \tag{3.1}$$

Remark 3.1 (Choice of macroscopic state variables). *The choice of the functions u_l is problem-dependent and will be specified with the numerical illustrations for the examples considered in this text. For the exposition in this section, it may be helpful to think about the standard moments of the distribution in a one-dimensional setting,*

i.e., $u_l(t,x) = x^l$. *Note that by allowing* u_l *to depend directly on time* t, *centralized moments* $u_1(t,x) = x$, $u_l(t,x) = (x - U_1(t))^l$ *for* $l \geq 2$, *for instance, can also be considered.*

The goal is to obtain a closed system of L evolution equations,

$$\frac{d\boldsymbol{U}(t)}{dt} = \mathcal{H}\big(\boldsymbol{U}(t)\big), \tag{3.2}$$

for the state variables $\boldsymbol{U}$, complemented with a constitutive equation,

$$\bar{\boldsymbol{f}}(t) = T\big(\boldsymbol{U}(t)\big), \tag{3.3}$$

for the observable $\bar{\boldsymbol{f}}$ of interest as a function of these macroscopic state variables. We briefly illustrate this approach in a one-dimensional setting; the exposition closely follows [42]. Using Itô calculus, one can easily obtain an equation of state for the macroscopic state variables. For Itô SDEs (1.1) it is of the form

$$\begin{aligned}\frac{dU_l(t)}{dt} &= \mathrm{E}\left(\frac{du_l(t,X)}{dt}\right)\\ &= \underbrace{\mathrm{E}\left(\frac{\partial u_l(t,X)}{\partial t}\right)}_{U_l^a} + \underbrace{\mathrm{E}\left(a(t,X)\,\frac{\partial u_l(t,X)}{\partial x}\right)}_{U_l^b} + \frac{1}{2}\underbrace{\mathrm{E}\left(\frac{\partial^2 u_l(X)}{\partial x^2} b(t,X)^2\right)}_{U_l^c},\end{aligned}$$

in which three new macroscopic state variables U_l^a, U_l^b, and U_l^c appear that, in general, are not functions of the initially chosen macroscopic state variables $(U_l)_{l=1}^L$. One can write evolution equations for these new state variables, which in turn will create additional state variables; this procedure typically goes on endlessly. At some point, one has to stop, and try to approximate the state variables for which no evolution equation is available by writing them as a function of other (already available) state variables. By adding such *closure relations*, one obtains an explicit, but approximate, closed system of macroscopic evolution equations.

The predominant way of deriving closure relations is to approximate the probability density function φ of $\boldsymbol{X}$ by a canonical density $\varphi_{\boldsymbol{U}}$ (in polymer dynamics often called *canonical [configuration] distribution function*), which is determined using only the macroscopic state variables $\boldsymbol{U}$ (typically low-order moments of the distribution), and to compute the remaining macroscopic variables appearing in the evolution equations for $\boldsymbol{U}$ via averaging with respect to $\varphi_{\boldsymbol{U}}$. For the FENE dumbbell problem, several closures have been proposed. In [41,42], approximate closures are obtained by restricting the space of admissible probability density functions to linear combinations of L *canonical basis functions* and considering the set of macroscopic state variables to be determined by the first L even moments of the distribution. (The odd moments vanish due to symmetry, as long as the initial distribution is symmetric.) We refer to [41] for more details on the one-dimensional setting and [42] for the general three-dimensional case. A related approach is described in [13, 27, 66].

In [28], a *quasi-equilibrium* approach is proposed, based on thermodynamical considerations; although the method has been formulated for the FENE dumbbell case, it is applicable to general SDEs of the type (1.1). The method first defines an entropy, and then requires the canonical (configuration) distribution to maximize this entropy, imposing the constraint that the macroscopic state is given. While this method results in a uniquely defined closed macroscopic model, it is often impossible to obtain this model analytically in closed form; several algorithms have been presented to simulate the evolution of the quasi-equilibrium model numerically [54, 65].

3.2. The coarse time-stepper: a numerical closure approach. One particular numerical method to simulate the quasi-equilibrium model is given in [54], also considering the FENE dumbbell problem. Consider the SDE (1.1), and define a set of L macroscopic state variables $\boldsymbol{U}(t) = (U_l(t))_{l=1}^L$, see equation (3.1). The procedure performs a numerical integration of a closed model (3.2) for the macroscopic state variables, *without deriving this model in analytic form.* To this end, a three-step numerical procedure is proposed that approximates a time-stepper of this approximate macroscopic model.

Two operators are introduced to make the transition between macroscopic and microscopic variables: a *lifting operator*,

$$\mathcal{L} : \boldsymbol{U} \mapsto \boldsymbol{\mathcal{Y}} = \mathcal{L}(\boldsymbol{U}), \tag{3.4}$$

which maps a macroscopic state onto a microscopic ensemble of J microscopic realizations, and the converse *restriction operator*,

$$\mathcal{R} : \boldsymbol{\mathcal{Y}} \mapsto \boldsymbol{U} = \mathcal{R}(\boldsymbol{\mathcal{Y}}). \tag{3.5}$$

Together, these operators have the effect of introducing the closure approximation.

As a restriction operator, we can readily replace the expectation E by the empirical mean $\widehat{\mathrm{E}}$,

$$U_l(t) = \mathcal{R}_l\big(\boldsymbol{\mathcal{Y}}(t)\big) = \widehat{\mathrm{E}} u_l\big(t, \boldsymbol{\mathcal{Y}}(t)\big). \tag{3.6}$$

In the lifting step, we need to create an ensemble of realizations corresponding to a given macroscopic state; the distribution that is sampled by this ensemble corresponds to the canonical density $\varphi_{\boldsymbol{U}}$ defined in Subsection 3.1. The numerical lifting procedure that was proposed in [54] for the FENE dumbbell model generates an ensemble taken from the conditional equilibrium distribution of (1.1) under the constraint that the macroscopic state is kept fixed. The method is based on a constrained simulation of the SDE (1.1). The easiest way to perform this constrained simulation is to take an unconstrained time-step for the polymer ensemble, followed by a projection onto the constraint [40]. When performing this constrained simulation at time t^*, we have:

$$\begin{cases} \boldsymbol{\mathcal{Y}}^{k_c+1} = \phi_{\boldsymbol{\mathcal{Y}}}(t^*, \boldsymbol{\mathcal{Y}}^{k_c}; \delta t) + \displaystyle\sum_{l=1}^{L} \lambda_l \nabla_{\boldsymbol{\mathcal{Y}}} \mathcal{R}_l(\boldsymbol{\mathcal{Y}}^{k_c}), \\ \text{with } \Lambda = \{\lambda_l\}_{l=1}^L \text{ such that } \mathcal{R}_l(\boldsymbol{\mathcal{Y}}^{k_c+1}) = U_l(t^*) \text{ for } l = 1, \ldots, L, \end{cases} \tag{3.7}$$

where the subscript on the counter $k_c = 0, \ldots, K_c - 1$ has been introduced to emphasize that the simulation is constrained, and physical time is frozen. The result of the lifting operator is then defined as the ensemble $\boldsymbol{\mathcal{Y}}^{K_c}$ for a sufficiently large time index K_c, for which the ensemble has approximately reached an equilibrium distribution,

$$\mathcal{L}(\boldsymbol{U}) = \boldsymbol{\mathcal{Y}}^{K_c}. \tag{3.8}$$

This lifting is closely related to the quasi-equilibrium approximation [54].

During the constrained simulation, an accept-reject strategy is applied on the combined evolution and projection operation, i.e., if the state of a polymer would become unphysical during projection, we reject the trial move in the evolution step and repeat the time step for this polymer, after which the projection of the ensemble is tried again.

The numerical closure algorithm can be written as follows [54]:

ALGORITHM 3.2 (Coarse time-stepper). *Given an initial condition for the macroscopic state variables $\boldsymbol{U}(t^*)$ at time t^*, one time step of the coarse time-stepper consists of a three-step procedure:*

(i) Lifting, *i.e., the creation of an initial ensemble $\boldsymbol{\mathcal{Y}}(t^*) = \mathcal{L}(\boldsymbol{U}(t^*))$ for the microscopic model, consistently with the macroscopic state $\boldsymbol{U}(t^*)$ at t^*.*

(ii) Simulation *using the numerical integration scheme* (2.3) *over a time interval $[t^*, t^* + K\delta t]$, where δt is the microscopic step size and K is the number of (microscopic) time steps, to get $\boldsymbol{\mathcal{Y}}(t^* + K\delta t)$: for $k = 0, \ldots, K-1$,*

$$\boldsymbol{\mathcal{Y}}\big(t^* + (k+1)\delta t\big) = \phi_{\boldsymbol{\mathcal{Y}}}\big(t^* + k\delta t, \boldsymbol{\mathcal{Y}}(t^* + k\delta t); \delta t\big).$$

(iii) Restriction, *i.e., observation (estimation) of the macroscopic state at $t^* + K\delta t$,*

$$\boldsymbol{U}(t^* + K\delta t) = \mathcal{R}\big(\boldsymbol{\mathcal{Y}}(t^* + K\delta t)\big),$$

as well as of an approximation of the function of interest, $\hat{\boldsymbol{f}}(t^* + K\delta t) = \widehat{\mathrm{E}}\boldsymbol{f}\big(\boldsymbol{\mathcal{Y}}(t^* + K\delta t)\big)$.

The numerical closure introduces an error in the evolution of the function of interest $\bar{\boldsymbol{f}}(t)$ due to the restriction onto a finite number of macroscopic state variables. Indeed, in the absence of statistical error, the coarse time-stepper approximates $\bar{\boldsymbol{f}}(t)$ in the limit $K_c \to \infty$ as

$$\hat{\boldsymbol{f}}(t^* + K\delta t) = \int \boldsymbol{f}(\boldsymbol{x})\varphi_{\boldsymbol{U}(t^*+K\delta t)}(\boldsymbol{x})d\boldsymbol{x}, \qquad i = 0, \ldots, n,\ k = 0, \ldots, K. \tag{3.9}$$

Clearly, the difference between $\varphi_{\boldsymbol{U}(t^*+K\delta t)}(\boldsymbol{x})$ and $\varphi(t^* + K\delta t, \boldsymbol{x})$ (with φ defined in Remark 2.3) depends on the choice of macroscopic state variables $\boldsymbol{U}$ in a way that is (in the limit $K_c \to \infty$) independent of the coarse time step. However, the assumption is that this error can be made arbitrarily small by considering sufficient macroscopic state variables (see also Subsection 5.1).

REMARK 3.3. *With a slight abuse of notation, we have also denoted the microscopic ensemble by the symbol $\boldsymbol{\mathcal{Y}}$ that was introduced to denote the time-discretized approximation of $\boldsymbol{\mathcal{X}}$. Note, however, that, for the coarse time-stepper, the approximation errors that are present in $\boldsymbol{\mathcal{Y}}$ are a consequence of both the time integration and statistical errors of the Monte Carlo simulation, as well as modeling errors through the numerically imposed closure.*

We also introduce the idealized operators $\overline{\mathcal{L}}$ and $\overline{\mathcal{R}}$ obtained in the limit $J \to \infty$, eliminating statistical error. Rather than acting on ensembles of configurations, these operators are defined directly on the random variables. More precisely, the restriction operator $\overline{\mathcal{R}}$ reduces a random variable $\boldsymbol{Z}$ to macroscopic state variables,

$$\overline{\mathcal{R}}(\boldsymbol{Z}) = \big(\overline{\mathcal{R}}_l(\boldsymbol{Z})\big)_{l=1}^{L} \text{ with } \overline{\mathcal{R}}_l(\boldsymbol{Z}) = U_l = \mathrm{E}\, u_l(\boldsymbol{Z}) \text{ for } l = 1, \ldots, L. \tag{3.10}$$

Likewise, the lifting operator $\overline{\mathcal{L}}$ maps a macroscopic state $\boldsymbol{U}^*$ onto a random variable $\boldsymbol{Z}^*$, which is distributed according to the probability density function $\varphi_{\boldsymbol{U}^*}$,

$$\boldsymbol{Z}^* = \overline{\mathcal{L}}(\boldsymbol{U}^*) \text{ with } \overline{\mathcal{R}}(\boldsymbol{Z}^*) = \boldsymbol{U}^*. \tag{3.11}$$

The procedure above can be useful to study the extent at which a closure based on a given set of macroscopic state variables is able to capture the macroscopic behavior of the system. It should, however, be emphasized that the method is by no means a computationally viable alternative to a fully microscopic simulation. (On the contrary, the associated computational cost is even much higher, due to the required constrained simulation in each lifting step.)

3.3. Coarse projective integration. In *coarse projective integration*, simulation using the coarse time-stepper is accelerated by extrapolating a (restricted) macroscopic state forward in time. The extrapolated macroscopic state is then lifted back to a corresponding microscopic state, and the simulation is continued.

We first introduce some notation. Let $I^{\Delta t} = \{t^0, t^1, \ldots, t^N\}$ be a (macroscopic) time discretization with $t^0 < t^1 < \ldots < t^N = T$ of the time interval I with step sizes $\Delta t_n = t^{n+1} - t^n$ for $n = 0, 1, \ldots, N-1$. We also introduce the discrete time instances $t^{n,k} = t^n + k\delta t$ that are defined on a microscopic grid, and, correspondingly, the discrete approximation $\boldsymbol{U}^{n,k} \approx \boldsymbol{U}(t^{n,k})$ and $\boldsymbol{U}^n \approx \boldsymbol{U}(t^n)$, and similarly for $\boldsymbol{\mathcal{Y}}$ and $\boldsymbol{Y}$. Clearly, $(\cdot)^n = (\cdot)^{n,0}$.

One step of the complete algorithm reads:

ALGORITHM 3.4 (Coarse projective integration).

(i) Lift *the macroscopic state to an ensemble of microscopic realizations using the lifting procedure described in Subsection 3.2,* $\boldsymbol{\mathcal{Y}}^n = \mathcal{L}(\boldsymbol{U}^n)$*;*

(ii) Simulate *the microscopic system via* (2.3) *over* K *time steps of size* δt*,*

$$\boldsymbol{\mathcal{Y}}^{n,k} = \phi_{\boldsymbol{\mathcal{Y}}}(t^{n,k-1}, \boldsymbol{\mathcal{Y}}^{n,k-1}; \delta t), \qquad 1 \leq k \leq K,$$

and record the restrictions $\boldsymbol{U}^{n,k} = \mathcal{R}(\boldsymbol{\mathcal{Y}}^{n,k})$*, as well as (an approximation of) the function of interest,* $\hat{\boldsymbol{f}}^{n,k} = \widehat{\mathrm{E}}\boldsymbol{f}(\boldsymbol{\mathcal{Y}}^{n,k})$*;*

(iii) Extrapolate *the macroscopic state* $\boldsymbol{U}$ *from (some of) the time points* $t^{i,k}$*,*$i = 0, \ldots, n$*,* $k = 1, \ldots, K$*, to a new macroscopic state* $\boldsymbol{U}^{n+1}$ *at time* t^{n+1}*.*

While discussion of the extrapolation step is postponed to Section 6, some remarks on the lifting step are in order. The constrained simulation that is performed during each lifting step requires an initial condition that satisfies the constraints. In general, during coarse projective integration, an ensemble is available that corresponds to a nearby macroscopic state; this ensemble can be projected onto the desired macroscopic state exactly as is done in each step of the constrained simulation, see equation (3.7). If no nearby ensemble is available, one can start from an ensemble that is obtained through an analytical closure approximation, see Section 3. It is important to emphasize that in the limit $K_c \to \infty$, the result of the lifting operator is a sample from the *conditional equilibrium density* $\varphi_{\boldsymbol{U}^*}$, so in this limit the initial condition for the constrained simulation does not affect the result.

4. Micro/macro acceleration method.

4.1. Algorithm formulation. Performing coarse projective integration as outlined above is clearly more efficient than just evolving the coarse time-stepper of Algorithm 3.2. However, when comparing to a fully microscopic Monte Carlo simulation, part (if not all) of the computational gain that is obtained by extrapolating in time is lost during the constrained simulation in the lifting step. In this section, we propose a method that aims at drastically reducing the computational cost of a fully microscopic Monte Carlo simulation by limiting the number of constrained time-steps that is performed during the lifting.

In particular, we propose to eliminate the constrained simulations completely and to replace the lifting step by a mere projection onto the extrapolated macroscopic state. As a consequence, the microscopic state after extrapolation depends on both the extrapolated macroscopic state and the microscopic state before extrapolation. Therefore, the clear relation with a numerical closure is lost, and the method can only be interpreted as a micro/macro acceleration technique for Monte Carlo simulation. One step of the micro/macro acceleration method then reads:

ALGORITHM 4.1 (Micro/macro acceleration).

(i) Simulate *the microscopic system over K time steps of size δt,*

$$\boldsymbol{\mathcal{Y}}^{n,k} = \phi_{\boldsymbol{\mathcal{Y}}}(t^{n,k-1}, \boldsymbol{\mathcal{Y}}^{n,k-1}; \delta t), \qquad 1 \le k \le K,$$

and record the restrictions $\boldsymbol{U}^{n,k} = \mathcal{R}(\boldsymbol{\mathcal{Y}}^{n,k})$, as well as an approximation of the function of interest $\hat{\boldsymbol{f}}^{n,k} = \widehat{\mathrm{E}}\boldsymbol{f}(\boldsymbol{\mathcal{Y}}^{n,k})$;

(ii) Extrapolate *the macroscopic state $\boldsymbol{U}$ from (some of) the time points $t^{i,k}$,$i = 0, \ldots, n$, $k = 1, \ldots, K$, to a new macroscopic state $\boldsymbol{U}^{n+1}$ at time t^{n+1},*

$$\boldsymbol{U}^{n+1} = \mathcal{E}\left(\left(\boldsymbol{U}^{i,k}\right)_{i,k=0,0}^{n,K}; (\Delta t_i)_{i=0}^{n}, \delta t \right). \tag{4.1}$$

(iii) Project *the microscopic state $\boldsymbol{\mathcal{Y}}^{n,K}$ onto the extrapolated macroscopic state $\boldsymbol{\mathcal{Y}}^{n+1} = \mathcal{P}(\boldsymbol{\mathcal{Y}}^{n,K}, \boldsymbol{U}^{n+1})$.*

(Again, if the state of a polymer would become unphysical during projection, we reject the trial move in the evolution step and repeat the time step for this polymer, after which the projection of the ensemble is tried again.)

Compared to Algorithm 3.4, the lifting operator $\mathcal{L}$ has been replaced by a projection operator $\mathcal{P}$. Also, the ordering of the steps has changed: Instead of evolving the *macroscopic* state over a time step of size Δt, making use of the microscopic SDE, as in Algorithm 3.4, the micro/macro acceleration method evolves the *microscopic* state over a time step of size Δt using an extrapolation on the macroscopic level.

Clearly, Algorithm 4.1 is computationally much faster than both Algorithm 3.4 and a fully microscopic simulation. However, it introduces, besides the already present microscopic *deterministic error* (due to time-stepping with δt) and *statistical error* (due to the finite number of samples), two additional sources of deterministic error due to (1) *extrapolation* of the macroscopic state forward in time; and (2) *projection* onto the selected macroscopic state variables. Moreover, in the extrapolation step, also the statistical error will be affected. The goal of the following sections is to investigate these additional errors, both analytically and numerically.

4.2. Projection failure and adaptive time-stepping. In the numerical experiments, one might encounter situations in which the distributions evolve on timescales that are not significantly slower than those of the macroscopic functions of interest. In that case, when taking a large extrapolation time step, the extrapolated macroscopic state differs significantly from the state corresponding to the last available polymer ensemble, and it is possible that projection onto the desired macroscopic state fails. Consequently, a failure in the projection can be used as an indication that we should decrease the extrapolation time step. Based on this observation, we propose the following criterion to adaptively determine the macroscopic step size Δt: If the projection fails, we reject the step and try again with a time-step

$$\Delta t_{\mathrm{new}} = \min(\underline{\alpha}\Delta t, K\delta t), \qquad \underline{\alpha} < 1, \tag{4.2}$$

whereas, when the projection succeeds, we accept the step and propose

$$\Delta t_{\mathrm{new}} = \max(\overline{\alpha}\Delta t, \Delta t_{\max}) \tag{4.3}$$

for the next step. If the macroscopic step size $\Delta t_{\mathrm{new}} = K\delta t$, projection becomes trivial (the identity operator), since there is no extrapolation. Note that, when this happens, the criterion will ensure that the larger time steps are tried after the next burst of microscopic simulation.

5. Projection operator. In this section, we discuss in detail the properties of the projection operator. We first give some necessary definitions (Subsection 5.1). We then prove a consistency result in a very special case (Subsection 5.2) and give and illustrate then a conjecture for the general case (Subsection 5.3).

5.1. Notations and definitions. The projection operator $\mathcal{P}$ can be defined in several ways. One option is to perform the same projection as during each time step of the constrained simulation (3.7), i.e.,

$$\begin{cases} \boldsymbol{\mathcal{Y}}^{n+1} = \boldsymbol{\mathcal{Y}}^{n,K} + \sum_{l=1}^{L} \lambda_l \nabla_{\boldsymbol{\mathcal{Y}}} \mathcal{R}_l(\boldsymbol{\mathcal{Y}}^n), \\ \text{with } \Lambda = \{\lambda_l\}_{l=1}^{L} \text{ such that } \mathcal{R}_l(\boldsymbol{\mathcal{Y}}^{n+1}) = U_l^{n+1} \text{ for } l = 1, \ldots, L. \end{cases} \tag{5.1}$$

Then, $\boldsymbol{\mathcal{Y}}^{n+1} = \mathcal{P}(\boldsymbol{\mathcal{Y}}^{n,K}, \boldsymbol{U}^{n+1})$ results after obtaining Λ from a Newton procedure that solves the L-dimensional nonlinear system that defines the constraints. If, instead, we write an implicitly defined gradient $\nabla_{\boldsymbol{\mathcal{Y}}} \mathcal{R}_l(\boldsymbol{\mathcal{Y}}^{n+1})$ in the first line of (5.1), one can show that the resulting ensemble satisfies

$$\boldsymbol{\mathcal{Y}}^{n+1} = \underset{\boldsymbol{\mathcal{Z}}:\ \mathcal{R}(\boldsymbol{\mathcal{Z}}) = \boldsymbol{U}^{n+1}}{\operatorname{argmin}} \frac{1}{2} \|\boldsymbol{\mathcal{Z}} - \boldsymbol{\mathcal{Y}}^{n,K}\|^2. \tag{5.2}$$

For some choices of the macroscopic state variables $\boldsymbol{U}$ one could also consider classical moment matching, as described in [6].

Due to the constraint $\mathcal{R}(\boldsymbol{\mathcal{Y}}^{n+1}) = \boldsymbol{U}^{n+1}$, the elements of $\boldsymbol{\mathcal{Y}}^{n+1}$ generated by (5.1) or (5.2) are in general not independent, and Monte Carlo error estimates are not straightforward to obtain. For example, the Central Limit Theorem is not directly applicable [6]. Therefore, as before, we introduce the operator $\overline{\mathcal{P}}$ obtained in the limit $J \to \infty$; the corresponding restriction operator $\overline{\mathcal{R}}$ was defined in equation (3.10). The projection operator $\overline{\mathcal{P}}$ maps a random variable $\boldsymbol{Z}$ onto a new random variable $\boldsymbol{Z}^*$ that corresponds to a macroscopic state $\boldsymbol{U}^*$,

$$\boldsymbol{Z}^* = \overline{\mathcal{P}}(\boldsymbol{Z}, \boldsymbol{U}^*) \text{ with } \overline{\mathcal{R}}(\boldsymbol{Z}^*) = \boldsymbol{U}^*. \tag{5.3}$$

When analyzing the projection, we will consider the idealized operators $\overline{\mathcal{P}}$ and $\overline{\mathcal{R}}$.

We are now ready to define a few properties that we require to be satisfied for any reasonable projection. First, we require the following property:

DEFINITION 5.1 (Self-consistency property). *A pair $\bar{\mathcal{P}}, \bar{\mathcal{R}}$ of projection and restriction operators is called self-consistent if*

$$\boldsymbol{Z} = \bar{\mathcal{P}}\big(\boldsymbol{Z}, \bar{\mathcal{R}}(\boldsymbol{Z})\big)$$

for any suitable random variable $\boldsymbol{Z}$.

The self-consistency states that a random variable remains unaffected by projection if its macroscopic state is equal to the macroscopic state on which one wants to project. Note that this property typically is not satisfied for the lifting operators discussed in Subsection 3.2, whereas it can easily be checked that it is fulfilled for the operators defined in (5.1) and (5.2). We remark also that Definition 5.1 implies $\bar{\mathcal{P}}^2 = \bar{\mathcal{P}}$, which justifies the use of the term *projection*.

Next, we consider the number of macroscopic state variables L to vary, and define a sequence of vectors of macroscopic state variables $\big(\boldsymbol{U}_{[L]}\big)_{L=1,2,\ldots}$, such that $\boldsymbol{U}_{[L]} = (U_l)_{l=1}^{L}$, i.e., for increasing L, additional macroscopic state variables are added.

The corresponding sequences of projection and restriction operators are denoted as $\left(\overline{\mathcal{P}}_{[L]}\right)_{L=1}^{\infty}$ and $\left(\overline{\mathcal{R}}_{[L]}\right)_{L=1}^{\infty}$, respectively.

Using this notation, we are ready to formulate the definitions of continuity and consistency of the projection step:

DEFINITION 5.2 (Continuity of projection). *Consider a set of random variables and a set of sequences of macroscopic states. A sequence of projection operators $\left(\overline{\mathcal{P}}_{[L]}\right)_{L=1,2,\dots}$ is called continuous for these sets if, for all $g \in C_P^{0,2}(I \times \mathbb{R}^d, \mathbb{R})$, there exists a constant C, depending only on g, such that*

$$|\operatorname{E} g\big(s, \overline{\mathcal{P}}_{[L]}(\boldsymbol{Z}, \boldsymbol{U}_{[L]}^*)\big) - \operatorname{E} g\big(s, \overline{\mathcal{P}}_{[L]}(\boldsymbol{Z}, \boldsymbol{U}_{[L]}^+)\big)| \leq C\|\boldsymbol{U}_{[L]}^* - \boldsymbol{U}_{[L]}^+\| \tag{5.4}$$

holds for all $L \geq 1$, all sequences $\left(\boldsymbol{U}_{[L]}^*\right)_{L=1,2,\dots}$, $\left(\boldsymbol{U}_{[L]}^+\right)_{L=1,2,\dots}$ of macroscopic states and all $\boldsymbol{Z}$ in the considered sets, and all $s \in I$.

DEFINITION 5.3 (Consistency of projection). *Consider a sequence of projection operators $\left(\overline{\mathcal{P}}_{[L]}\right)_{L=1,2,\dots}$. This sequence is called consistent for a class of sequences of triples $\left(\boldsymbol{Z}_{[L]}^*, \boldsymbol{Z}_{[L]}^+, \boldsymbol{U}_{[L]}\right)_{L=1,2,\dots}$ of random variables $\boldsymbol{Z}_{[L]}^*$, $\boldsymbol{Z}_{[L]}^+$, and macroscopic states $(\boldsymbol{U}_{[L]})$, if for all $g \in C_P^{0,2}(I \times \mathbb{R}^d, \mathbb{R})$ there exist constants C_L, with $C_L \to 0$ for $L \to \infty$, and L_0 such that it holds*

$$|\operatorname{E} g\big(s, \overline{\mathcal{P}}_{[L]}(\boldsymbol{Z}_{[L]}^*, \boldsymbol{U}_{[L]})\big) - \operatorname{E} g\big(s, \overline{\mathcal{P}}_{[L]}(\boldsymbol{Z}_{[L]}^+, \boldsymbol{U}_{[L]})\big)| \leq C_L |\operatorname{E} g(s, \boldsymbol{Z}_{[L]}^*) - \operatorname{E} g(s, \boldsymbol{Z}_{[L]}^+)| \tag{5.5}$$

for all $L \geq L_0$ and all sequences of triples $\left(\boldsymbol{Z}_{[L]}^*, \boldsymbol{Z}_{[L]}^+, \boldsymbol{U}_{[L]}\right)_{L=1,2,\dots}$ in the considered class. The possible dependence of the random variables on L is required since the random variables will be considered to have been generated using the micro/macro acceleration algorithm, and therefore depend on L.

Whereas continuity measures (in a weak sense) the difference between the projection of a random variable onto two different macroscopic states, the consistency measures the difference between the projection of two different random variables onto the same macroscopic state.

Definitions 5.1, 5.3, and 2.1 immediately imply the following corollary:

COROLLARY 5.4. *Consider a sequence of (self-consistent) projection operators $\left(\overline{\mathcal{P}}_{[L]}\right)_{L=1,2,\dots}$, and assume that this sequence is consistent for a set of sequences of triples $\left(\boldsymbol{Z}_{[L]}, \boldsymbol{X}^{t^-, \boldsymbol{Z}_{[L]}}(t^*), \overline{\mathcal{R}}_{[L]}\big(\boldsymbol{X}^{t^-, \boldsymbol{Z}_{[L]}}(t^*)\big)\right)_{L=1,2,\dots}$, where $Z_{[L]}$ are random variables, $t^- = t^* - \Delta t$, $\Delta t \in [0, t^* - t^0]$, $t^* \in I$. Suppose further that SDE (1.1) is uniformly weakly continuous for the set of all $\boldsymbol{Z}_{[L]}$. Then for all $g \in C_P^{0,2}(I \times \mathbb{R}^d, \mathbb{R})$ there exist constants $\Delta t_0 > 0$ and C_L, with $C_L \to 0$ for $L \to \infty$, and L_0, such that it holds*

$$|\operatorname{E} g\big(s, \boldsymbol{X}^{t^-, \boldsymbol{Z}_{[L]}}(t^*)\big) - \operatorname{E} g\big(s, \overline{\mathcal{P}}_{[L]}(\boldsymbol{Z}_{[L]}, \boldsymbol{U}_{[L]}^*)\big)| \leq C_L \Delta t \tag{5.6}$$

for all $L \geq L_0$, all $\left(\boldsymbol{Z}_{[L]}, \boldsymbol{X}^{t^-, \boldsymbol{Z}_{[L]}}(t^*), \boldsymbol{U}_{[L]}^* = \overline{\mathcal{R}}_{[L]}\big(\boldsymbol{X}^{t^-, \boldsymbol{Z}_{[L]}}(t^*)\big)\right)_{L=1,2,\dots}$ in the considered set, $\Delta t \in [0, \Delta t_0]$, and all $t^* \in [t^0 + \Delta t, T]$, $s \in I$.

This corollary states that the difference, measured in a weak sense, between the exact distribution at some time instance t^* and a projection from a previous time $t^- = t^* - \Delta t$ onto the exact macroscopic state at t^* vanishes when projecting onto more macroscopic state variables or letting $\Delta t \to 0$.

In the remainder of the text, we will only use the properties stated above. Therefore, one may use any projection operator that would come to mind for a particular problem, as long as these properties are satisfied.

5.2. Particular results for normal distributions. For simplicity of the argument, and without loss of generality, we restrict ourselves to a one-dimensional notation for the rest of this section. In the case of scalar normally distributed random variables, the following result can be proved.

LEMMA 5.5. *Consider a scalar random variable Z and suppose that the macroscopic state variables are given by $u_1(z) = z$ and $u_2(z) = (z - U_1)^2$. Suppose further that in analogy to* (5.2)*, $\overline{\mathcal{P}}$ is given by*

$$\overline{\mathcal{P}}(Z, \boldsymbol{U}^*) = \underset{\overline{\mathcal{R}}(Z^*)=\boldsymbol{U}^*}{\operatorname{argmin}} \frac{1}{2} \operatorname{E}\left((Z^* - Z)^2\right). \tag{5.7}$$

Then, the random variable $\overline{\mathcal{P}}(Z, \boldsymbol{U}^*)$ is also normally distributed with mean and variance given by $\boldsymbol{U}^*$.

Proof. Denote by

$$\begin{pmatrix} \mu \\ \sigma^2 \end{pmatrix} = \overline{\mathcal{R}}(Z), \quad \begin{pmatrix} \mu^* \\ (\sigma^*)^2 \end{pmatrix} = \boldsymbol{U}^*, \quad \sigma, \sigma^* > 0.$$

Then, (5.7) yields

$$\overline{\mathcal{P}}(Z, \boldsymbol{U}^*) = \pm\sqrt{\frac{(\sigma^*)^2}{\sigma^2}} \left(Z - \mu \pm \mu^* \sqrt{\frac{\sigma^2}{(\sigma^*)^2}} \right). \tag{5.8}$$

Thus, if Z is normally distributed, so is $\overline{\mathcal{P}}(Z, \boldsymbol{U}^*)$. ∎

With the help of this lemma, we can easily show the following two corollaries.

COROLLARY 5.6. *Suppose that the hierarchy of macroscopic state variables is defined using $u_1(z) = z$ and $u_l(z) = (z - U_1)^l$ for $l \geq 2$. Suppose further that $\overline{\mathcal{P}}$ is given by* (5.7)*. Then, the sequence of projection operators is consistent with $C_L = 0$ for $L \geq 2 = L_0$ for all sequences of triples $\left(Z^+_{[L]}, Z^-_{[L]}, \boldsymbol{U}^*_{[L]}\right)_{L=1,2,\ldots}$, where $Z^+_{[L]}$ and $Z^-_{[L]}$ are normally distributed and $\left(\boldsymbol{U}^*_{[L]}\right)_{L=1,2,\ldots}$ are sequences of centralized moment values consistent with normal distributions.*

Proof. We first consider the case $L = 2$. As the normal distribution is uniquely determined by its first two (centralized) moments, Lemma 5.5 implies that for two normally distributed random variables Z_1 and Z_2, $\overline{\mathcal{P}}_{[2]}(Z_1, \boldsymbol{U}^*_{[2]})$ and $\overline{\mathcal{P}}_{[2]}(Z_2, \boldsymbol{U}^*_{[2]})$ are identically distributed, and thus $C_2 = 0$. For the same reason, (5.8) holds also for $L > 2$, and also in this case $C_L = 0$, and the projection is consistent. ∎

COROLLARY 5.7. *Suppose that the hierarchy of macroscopic state variables is defined using $u_1(z) = z$ and $u_l(z) = (z - U_1)^l$ for $l \geq 2$. Suppose further that $\overline{\mathcal{P}}$ is given by* (5.7)*. Then, the sequence of projection operators is continuous for all normally distributed random variables, and sequences of centralized moment values consistent with normal distributions.*

Proof. For $L \geq 2$, Lemma 5.5 implies again that $Z^* = \overline{\mathcal{P}}_{[L]}(Z, \boldsymbol{U}^*_{[L]})$ and $Z^+ = \overline{\mathcal{P}}_{[L]}(Z, \boldsymbol{U}^+_{[L]})$ are normally distributed if Z is normally distributed and $\boldsymbol{U}^*_{[L]}$, $\boldsymbol{U}^+_{[L]}$ are sequences of centralized moment values consistent with normal distributions.

Denoting the corresponding expectations by μ^*, resp. μ^+, and variances by $(\sigma^*)^2$, resp. $(\sigma^+)^2$, it holds for all $f \in C^1_P(\mathbb{R},\mathbb{R})$

$$\begin{aligned}|\operatorname{E} f(s,Z^*) - \operatorname{E} f(s,Z^+)| &= \left| \int_{\mathbb{R}} \big(f(s,\sigma^* z + \mu^*) - f(s,\sigma^+ z + \mu^+)\big) \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz \right| \\ &= \left| \int_{\mathbb{R}} f'(s,\xi_z)[(\sigma^* - \sigma^+)z + \mu^* - \mu^+] \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz \right|,\end{aligned}$$

where $\xi_z \in [\sigma^* z + \mu^*, \sigma^+ z + \mu^+]$. As there exist constants $\tilde{C}$ and $r \in \mathbb{N}$ such that $|f'(s,\xi_z)| \leq \tilde{C}(1 + |\max\{\sigma^*,\sigma^+\}z + \max\{\mu^*,\mu^+\}|^{2r})$, this implies also that the sequence of projection operators is continuous (the corresponding equation (5.4) can be verified similarly in the case $L = 1$). □

Several observations can be made. First, one can obtain a similar result whenever the distributions are defined by a finite number of moments (for instance, a lognormal distribution) by taking this knowledge into account when defining the projection operator. Second, we remark that, if the random variables are normally distributed at all moments in time, this implies that they represent solutions of a linear SDE in the narrow sense,

$$dX(t) = \big(a_1(t)X(t) + a_2(t)\big)\, dt + b(t) \star dW(t), \quad t \in I, \tag{5.9}$$

with normally distributed initial values. For this equation, it is clear that the complete time evolution of the distributions can be completely described by a system of two ODEs for $U_1 = \mu = \operatorname{E} X$ and $U_2 = \sigma^2 = \operatorname{E}\left((X-\mu)^2\right)$, see equation (3.2), namely

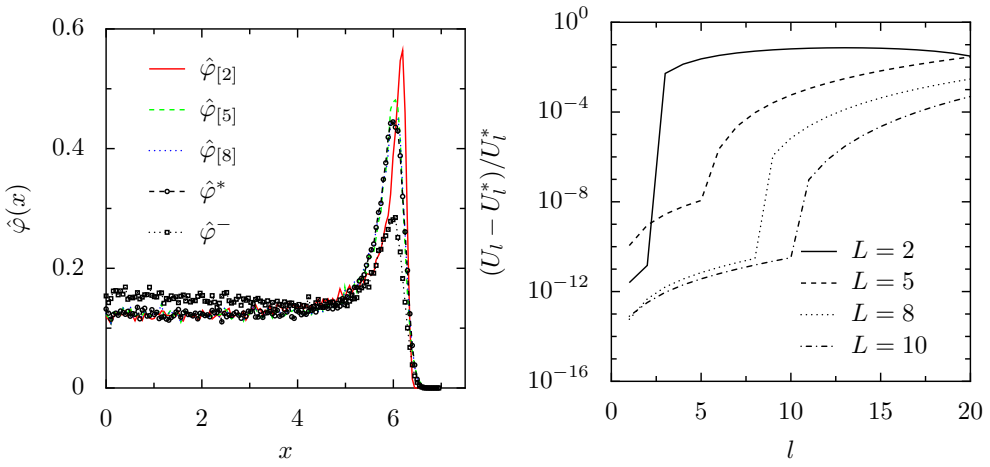
$$\begin{cases} dU_1/dt &= a_1(t)\, U_1 + a_2(t), \\ dU_2/dt &= 2a_1(t)\, U_2 + b(t)^2. \end{cases} \tag{5.10}$$

As a consequence, the projection operator (5.7) corresponds to the reconstruction of the normal distribution corresponding to a given macroscopic state, i.e., the projection operator reduces to a lifting operator as defined in Subsection 3.2. Consequently, in this setting, the coarse projective integration algorithm (Algorithm 3.4), the micro/macro acceleration algorithm (Algorithm 4.1), and a direct discretization of the closed system (5.10) yield equivalent numerical results, albeit at different computational costs.

5.3. Conjecture for general distributions. We now turn to more general distributions. We assume that the distribution of the random variable is uniquely determined by its moments; this is the case if, for instance, the moment generating function $\sum_{i=0}^{\infty} \frac{\operatorname{E}(Z^i)t^i}{i!}$ is bounded in an interval around 0.

In this setting, we propose the following conjecture:

CONJECTURE 5.8. *Consider a sequence of restriction operators* $\left(\overline{\mathcal{R}}_{[L]}\right)_{L=1,2,\ldots}$ *in which the macroscopic state variables corresponding to* $\overline{\mathcal{R}}_{[L]}$ *are defined as the first* L *centralized moments of the distribution, i.e.,* $u_1(z) = z$, $u_l(z) = (z - U_1)^l$, *for* $l = 2,\ldots,L$, *and define the corresponding sequence of projection operators* $\left(\overline{\mathcal{P}}_{[L]}\right)_{L=1,2,\ldots}$ *via* (5.7). *Consider further a set* S *of random variables for which all (centralized) moments of its members exist and uniquely determine the corresponding distribution function, and each moment can be uniformly bounded. Then, the sequence of projection operators is continuous for* S *and all sequences of macroscopic states* $\boldsymbol{U}_{[L]}$, *and consistent for all sequences of triples* $\left(Z^*_{[L]}, Z^+_{[L]}, \boldsymbol{U}_{[L]}\right)_{L=1,2,\ldots}$ *where* $Z^*_{[L]}, Z^+_{[L]} \in S$.



(a) Empirical probability density functions (b) Error of the l-th moment as a function of l

FIG. 5.1. *Results after projecting a prior ensemble of FENE dumbbells onto the first L even centralized moments of a reference ensemble for several values of L. Simulation details are given in the text.*

We illustrate the main properties of the projection operator for general distributions by means of numerical experiments. Below, we consider equation (2.6) in one space dimension, with $\kappa(t) \equiv 2$, $F(X)$ the FENE force (2.7) with $\gamma = 49$, $\mathrm{We} = 1$. We discretize in time with the classical Euler-Maruyama scheme with time step $\delta t = 2 \cdot 10^{-4}$. As the macroscopic state $\boldsymbol{U}_{[L]}$, we consider the first L *even* centralized moments, since for the exact solution, the odd (centralized) moments vanish due to symmetry.

5.3.1. Error dependence on the number of moments. We first simulate $J = 1 \cdot 10^5$ realizations, whose initial states are standard normally distributed, up to time $t^* = 1.15$ and record the microscopic state $\boldsymbol{\mathcal{Y}}^*$ and corresponding macroscopic states $\boldsymbol{U}^*_{[L]}$ for $L = 1, \ldots, 10$, at time t^*, as well as the microscopic state $\boldsymbol{\mathcal{Y}}^-$ at time $t^- = 1$. We then project the ensemble $\boldsymbol{\mathcal{Y}}^-$ onto the macroscopic state $\boldsymbol{U}^*_{[L]}$ and compare the density of $\mathcal{P}_{[L]}(\boldsymbol{\mathcal{Y}}^-, \boldsymbol{U}^*_{[L]})$ with that of $\boldsymbol{\mathcal{Y}}^*$. Since the absolute value of the moments increases quickly with the order of the moment, the residuals in the Newton procedure for the projection are scaled relative to the requested value of the corresponding moment; the Newton iterations are stopped if the norm of the residual is smaller than $1 \cdot 10^{-9}$.

We perform three tests to examine the convergence (in empirical density) of $\mathcal{P}_{[L]}(\boldsymbol{\mathcal{Y}}^-, \boldsymbol{U}^*_{[L]})$ to $\boldsymbol{\mathcal{Y}}^*$ for $L \to \infty$. First, we visually inspect the corresponding empirical probability density functions, see Fig. 5.1(a). Shown are histogram approximations of the empirical density $\hat{\varphi}^-$ of $|\boldsymbol{\mathcal{Y}}^-|$ (the initial condition for the projection), the reference empirical density $\hat{\varphi}^*$ of $|\boldsymbol{\mathcal{Y}}^*|$, and approximations $\hat{\varphi}_{[L]}$ of $\mathcal{P}_{[L]}(\boldsymbol{\mathcal{Y}}^-, \boldsymbol{U}_{[L]})$ for several values of L. The figure visually suggests that, when increasing the number of macroscopic state variables, the reference empirical density gets approximated more accurately. We now take a closer look to the projected ensembles by computing the relative difference between the l-th even empirical moment of the projected ensemble, U_l, and the corresponding empirical moment of the reference ensemble U_l^* as

L	2	3	4	5	6	7	8	9	10
p	0.000	0.000	$1.197 \cdot 10^{-3}$	0.840	0.862	0.999	0.999	0.999	0.999

TABLE 5.1

The p-values of a two-sample Kolmogorov–Smirnov test that compares the reference and projected empirical distributions. Simulation details are in the text.

$(U_l - U_l^*)/U_l^*$. Figure 5.1(b) shows this error as a function of l for different values of the number of macroscopic state variables L.

We make two key observations. First, for $l < L$, the relative difference in the corresponding moment is of the order of the tolerance of the Newton procedure. This is expected, since these are the macroscopic state variables onto which the distribution is projected. We see that this very small error increases nevertheless with l; this can be explained by pointing out that the value of the moments increases very quickly with l, and that the equations in the Newton procedure have been rescaled accordingly. Second, the error in the higher moments ($l > L$) also decreases with increasing L. This indicates that the convergence of $\hat{\varphi}_{[L]}$ to $\hat{\varphi}^*$ for $L \to \infty$ is not only due to the fact that we project onto more moments, but also because the approximation of the higher order moments improves. So far, we have no complete theoretical justification for this observation.

Finally, we compare the reference and projected empirical distributions using a two-sample Kolmogorov–Smirnov test [39]; this classical hypothesis test results in a high p-value (≤ 1) if the two samples are likely to have been drawn from the same probability distribution. The results are shown in Table 5.1. We clearly see a p-value that approaches 1 for increasing L.

5.3.2. Error dependence on the time step. In a next experiment, we again simulate $J = 1{\cdot}10^5$ realizations, whose initial states are standard normally distributed, up to time $t^* = 1.7$, and record the microscopic state $\boldsymbol{\mathcal{Y}}^-$ at time $t^- = 1.5$, as well as the macroscopic states $\boldsymbol{U}_{[L]}(t^- + \Delta t)$ for $L = 3, 4, 5$, and the function of interest $\tilde{\tau}_p(t^- + \Delta t)$ for $\Delta t \in [0, t^* - t^-]$. We then project the ensemble $\boldsymbol{\mathcal{Y}}^-$ onto the macroscopic state $\boldsymbol{U}_{[L]}(t^- + \Delta t)$, obtaining $\mathcal{P}_{[L]}\big(\boldsymbol{\mathcal{Y}}^-, \boldsymbol{U}_{[L]}(t^- + \Delta t)\big)$, and denote the corresponding value of the stress as $\hat{\tau}_p(t^- + \Delta t)$. We record the relative error $|\hat{\tau}_p(t) - \tilde{\tau}_p(t)|/\tilde{\tau}_p(t)$ as a function of Δt. To reduce the statistical error, we report the averaged results of 100 realizations of this experiment. The results are shown in Fig. 5.2. We indeed see the linear increase of the projection error as a function of Δt; notice also, as was shown in the previous experiment, that the projection error decreases with increasing L. From the figure, we conclude that the remaining statistical error is at least lower than 10^{-5}.

6. Extrapolation operator. Next, we need to specify how the extrapolation is performed. Let the order of consistency be defined as follows:

DEFINITION 6.1 (Consistency of extrapolation). *Consider a certain class of sufficiently smooth functions. An extrapolation operator $\mathcal{E}$ is called consistent of order $p_e > 0$ for this class if there exist $\Delta t_0 > 0$ and C such that for all $\Delta t \in [0, \Delta t_0]$, all $n \leq N$ and all functions U in the considered class it holds*

$$\left\| \tilde{U}^{n+1} - U(t^{n+1}) \right\| \leq C \Delta t^{p_e+1}, \tag{6.1}$$

with

$$\tilde{U}^{n+1} = \mathcal{E}\left(\left(U(t^{i,k}) \right)_{i,k=0,0}^{n,K} ; (\Delta t_i)_{i=0}^{n}, \delta t \right).$$

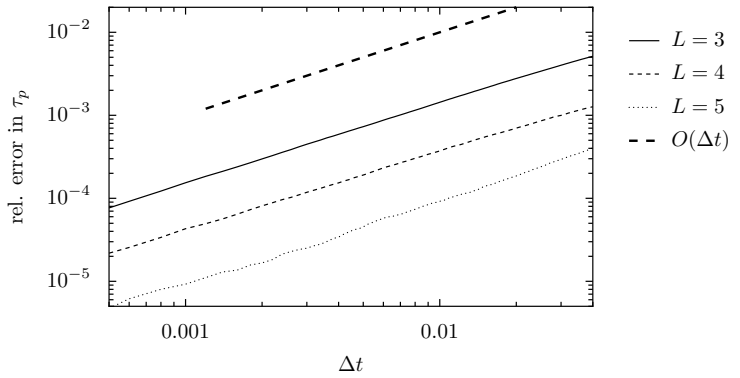


FIG. 5.2. *Error of the stress after projecting a prior ensemble of FENE dumbbells onto the first L even centralized moments of a reference ensemble, as a function of Δt for several values of L. Simulation details are given in the text.*

Further, we will also use the following definition of continuity.

DEFINITION 6.2 (Continuity of extrapolation). *Consider a certain class of sufficiently smooth functions. An extrapolation operator $\mathcal{E}$ consistent of order $p_e > 0$ for this class is called continuous for this class, if there exist $\Delta t_0 > 0$ and C such that for all $\Delta t \in [0, \Delta t_0]$, all $n \leq N$, and all functions U_1, U_2 in the considered class it holds*

$$\begin{aligned} \|\mathcal{E}\left(\left(U_1(t^{i,k})\right)_{i,k=0,0}^{n,K}; (\Delta t_i)_{i=0}^{n}, \delta t\right) - \mathcal{E}\left(\left(U_2(t^{i,k})\right)_{i,k=0,0}^{n,K}; (\Delta t_i)_{i=0}^{n}, \delta t\right)\| \\ \leq C\left(\frac{\Delta t}{\delta t}\right)^{p_e} \sum_{i,k=0,0}^{n,K} \|U_1(t^{i,k}) - U_2(t^{i,k})\|. \end{aligned} \tag{6.2}$$

In the remainder of this section, we consider two extrapolation strategies: projective extrapolation and multistep state extrapolation.

6.1. Projective extrapolation. The first approach that we consider, proposed in [17], is to extrapolate $\boldsymbol{U}$ to time t^{n+1} using only (some of) the time points $t^{n,k}$, $k = 0, \ldots, K$, i.e., by using the sequence of points obtained in the last burst of microscopic simulation. If this *coarse projective extrapolation* is based on the interpolating polynomial of degree p_e (with $p_e \leq K$) through the parameter values at times $t^{n,k}$, $k = K - p_e, \ldots, K$, we obtain

$$\boldsymbol{U}^{n+1} = \sum_{s=0}^{p_e} l_s(\alpha_n) \boldsymbol{U}^{n,K-s}, \tag{6.3}$$

with

$$\alpha_n = \frac{\Delta t_n}{\delta t} - K, \tag{6.4}$$

and the Lagrange polynomials

$$l_s(\alpha) = \frac{\alpha(\alpha+1)\cdots(\alpha+p_e)}{s!(p_e-s)!(-1)^s(\alpha+s)}. \tag{6.5}$$

EXAMPLE 6.3 (Coarse projective forward Euler). *The simplest, first order version of the above method is called coarse projective forward Euler. In this case, the procedure can be rewritten as*

$$\boldsymbol{U}^{n+1} = \boldsymbol{U}^{n,K} + (\Delta t_n - K\delta t)\overline{\boldsymbol{\mathcal{H}}}^n, \qquad \overline{\boldsymbol{\mathcal{H}}}^n = \frac{\boldsymbol{U}^{n,K} - \boldsymbol{U}^{n,K-1}}{\delta t}. \tag{6.6}$$

The procedure described above is reminiscent of a Taylor method [17]. Consequently, time integration based on this extrapolation will resemble a Taylor method when repeatedly extrapolating forward in time, and the global deterministic error will be dominated by a term of the form $C\Delta t^{p_e}$ (assuming $\delta t \ll \Delta t$), as results from an accuracy analysis of coarse projective integration for deterministic microscopic models [63].

To assess qualitatively the influence of coarse projective extrapolation on the statistical error, we apply it to the linear test equation

$$dX(t) = aX(t)\ dt + bdW(t). \tag{6.7}$$

Application of the one step method ϕ to (6.7) yields

$$Y^{n,k} = \tilde{R}_\phi(a, \delta t, \eta^{n,k-1})Y^{n,k-1} + \tilde{S}_\phi(a, b, \delta t, \eta^{n,k-1}),$$

where $\tilde{R}_\phi$ and $\tilde{S}_\phi$ are functions depending on ϕ and $\eta^{n+1,k-1}$ are (vectors of) the i. i. d. random variables used by ϕ. In the following, we assume that $\tilde{R}_\phi(a, \delta t, \eta^{n,i})$ is independent of $\eta^{n,i}$ and can be written as $\tilde{R}_\phi(a, \delta t, \eta^{n,i}) = R_\phi(a\delta t)$; this holds, e.g., for typical Runge-Kutta methods. The above assumptions imply

$$\begin{aligned} Y^{n,k} &= R_\phi(a\delta t)^k Y^{n,0} + \sum_{i=0}^{k-1} \tilde{S}_\phi(a, b, \delta t, \eta^{n,i}) R_\phi(a\delta t)^{k-i-1}, \\ \mathrm{E}\, Y^{n,k} &= R_\phi(a\delta t)^k\, \mathrm{E}\, Y^{n,0} + \mathrm{E}\, \tilde{S}_\phi(a, b, \delta t, \eta^{n,0}) \sum_{i=0}^{k-1} R_\phi(a\delta t)^i. \end{aligned}$$

If we now apply the extrapolation step (6.3), we obtain

$$\begin{aligned} \mathrm{E}\, Y^{n+1,0} &= \sum_{s=0}^{p_e} l_s(\alpha)\, \mathrm{E}\, Y^{n,K-s} \\ &= \underbrace{\left(\sum_{s=0}^{p_e} l_s(\alpha) R_\phi(a\delta t)^{K-s} \right)}_{=:R_{\mathrm{E}}(a\delta t)} \mathrm{E}\, Y^{n,0} + \mathrm{E}\, \tilde{S}_\phi(a, b, \delta t, \eta^{n,0}) \sum_{s=0}^{p_e} l_s(\alpha) \sum_{i=0}^{K-s-1} R_\phi(a\delta t)^i, \end{aligned} \tag{6.8}$$

with α and $l_s(\alpha)$ given by (6.4) and (6.5). In analogy to (6.8) we obtain also

$$\widehat{\mathrm{E}} Y^{n+1,0} = R_{\mathrm{E}}(a\delta t)\widehat{\mathrm{E}} Y^{n,0} + \sum_{s=0}^{p_e} l_s(\alpha) \sum_{i=0}^{K-s-1} R_\phi(a\delta t)^i \widehat{\mathrm{E}} \tilde{S}_\phi(a, b, \delta t, \eta^{n,i}).$$

Thus $\mathrm{E}\, \widehat{\mathrm{E}} Y^{n+1,0} = \mathrm{E}\, Y^{n+1,0}$, but

$$\begin{aligned} \mathrm{Var}\, \widehat{\mathrm{E}} Y^{n+1,0} =& \frac{1}{J} R_{\mathrm{E}}(a\delta t)^2\, \mathrm{Var}\, Y^{n,0} \\ &+ \frac{1}{J}\, \mathrm{Var}\, \tilde{S}_\phi(a, b, \delta t, \eta^{n,0}) \sum_{i=0}^{K-1} R_\phi(a\delta t)^{2i} \left(\sum_{s=0}^{\min\{p_e, K-1-i\}} l_s(\alpha) \right)^2. \end{aligned} \tag{6.9}$$

The second summand in (6.9) behaves as $\delta t \alpha^{2p_e}/J$ for large $\alpha = \Delta t/\delta t - K$. Assuming $\Delta t \gg \delta t$, this results in an amplification of the statistical error with a factor $\Delta t^{p_e}/\delta t^{p_e}$ during extrapolation. A natural question is then: how many realizations $\tilde{J}$ are needed to obtain the same variance using a fully microscopic simulation? In that case, we have

$$\begin{aligned}\operatorname{Var} \widehat{\mathrm{E}} Y^{n,\alpha+K} =& \frac{1}{\tilde{J}} R_\phi(a\delta t)^{2(\alpha+K)} \operatorname{Var} Y^{n,0} \\ &+ \frac{1}{\tilde{J}} \operatorname{Var} \tilde{S}_\phi(a,b,\delta t,\eta^{n,0}) \sum_{i=0}^{\alpha+K-1} R_\phi(a\delta t)^{2i}. \end{aligned} \tag{6.10}$$

Thus, for large α, the required number of realizations for a full microscopic simulation is smaller by a factor $1/\alpha^{2p_e-1}$, i.e, $\tilde{J} \sim \frac{J}{\alpha^{2p_e-1}}$, whereas the computational costs per realization increases by a factor α. This means that for large α and $p_e = 1$, the computational cost of the micro/macro acceleration technique with coarse projective extrapolation is similarly to that of a full microscopic simulation for a given variance. For large α and $p_e > 1$, coarse projective extrapolation is even more expensive than a full microscopic simulation.

To reduce statistical error, it has been proposed to use a chord based approximation, for instance, using $\boldsymbol{U}^{n,K-K_1}$ for the time derivative estimate instead of $\boldsymbol{U}^{n,K-1}$ in equation (6.6) [50]. Instead of taking Lagrange polynomials in equation (6.3), we then have

$$\boldsymbol{U}^{n+1} = \sum_{s=0}^{K} l_s(\alpha_n) \boldsymbol{U}^{n,K-s}, \tag{6.11}$$

in which $l_K(\alpha) = 1 + \frac{\alpha}{K-K_1}$, $l_{K_1}(\alpha) = -\frac{\alpha}{K-K_1}$, and $l_s(\alpha) = 0$ otherwise (see Example 6.3) reduces the variance by a factor $1/(K-K_1)$. However, the conclusion on the computational cost remains the same.

6.2. Multistep state extrapolation. Because of the amplification of statistical error, we look into alternative extrapolation strategies. One approach, proposed in [57,64], is to extrapolate $\boldsymbol{U}$ to time t^{n+1} using only (some of) the time points $t^{i,K}$, $i = 1, \ldots, n$, i.e., by using *the last point of each sequence of microscopic simulations*, instead of a sequence of points from the last microscopic simulation. If this *multistep state extrapolation method* is based on the interpolating polynomial of degree p_e through the parameter values at times $t^{i,K}$, $i = n - p_e, \ldots, n$, and we assume equidistant coarse time steps Δt, we obtain

$$\boldsymbol{U}^{n+1} = \sum_{s=0}^{p_e} l_s(\beta) \boldsymbol{U}^{n-s,K}, \tag{6.12}$$

where

$$\beta = \frac{\alpha}{\alpha + K} \tag{6.13}$$

is the fraction of the interval Δt over which we extrapolate, and α and l_s are defined as in (6.4) and (6.5). Note that such an extrapolation strategy requires a separate starting procedure.

For a detailed comparison of the accuracy and stability properties of acceleration of the numerical integration of ODEs using projective extrapolation and multistep

state extrapolation, we refer to [63, 64]. Here, we only remark that, while the local error using multistep state extrapolation only differs by a factor two with respect to the local error of projective extrapolation, the global error is affected quite significantly. This is due to the fact that, when increasing Δt, also the time derivative estimate itself is taken over a larger time interval. It has been shown in [57, 63] that this results in an error constant (see, e.g., [21, Section III.2]) of the form $C\alpha\Delta t^{p_e}$. This amplification effect will be illustrated in Subsection 6.3.

Let us now look into the statistical error, again using the linear test equation (6.7). One extrapolation step of coarse multistep state extrapolation yields

$$\widehat{\mathrm{E}}Y^{n+1,0} = \sum_{s=0}^{p_e} l_s(\beta)\widehat{\mathrm{E}}Y^{n-s,K}.$$

Thus again $\mathrm{E}\,\widehat{\mathrm{E}}Y^{n+1,0} = \mathrm{E}\,Y^{n+1,0}$, but now

$$\operatorname{Var}\widehat{\mathrm{E}}Y^{n+1,0} \leq \frac{1}{J}\max_{s=0}^{p_e}\operatorname{Var}Y^{n-s,K}\left(\sum_{s=0}^{p_e}|l_s(\beta)|\right)^2.$$

As $\beta < 1$, the last factor can be bounded (independently of α). Consequently, the amplification of statistical error during the extrapolation does not depend on α, whereas the corresponding computational costs per simulation path are reduced by a factor α compared to a full microscopic simulation.

6.3. Numerical illustration. We now provide a numerical result to illustrate the effects of extrapolation on the deterministic and statistical error. To avoid effects of the projection step, we consider the linear equation (5.9) with $a_2(t) = -a_1(t) = b(t) \equiv 1$, for which we know that macroscopic evolution closes in terms of the first two moments of the distribution. This microscopic SDE is discretized using an Euler-Maruyama scheme with $\delta t = 2\cdot 10^{-4}$. We consider 500 realizations of a computational experiment with $J = 1000$ SDE realizations. As an initial condition, we sample from a standard normal distribution. We compare the sample mean behavior and sample standard deviation of a full microscopic simulation (which we will call the reference simulation) with the micro/macro acceleration algorithm using $\Delta t = 1\cdot 10^{-3}$, $2\cdot 10^{-3}$, $4\cdot 10^{-3}$, and $8\cdot 10^{-3}$. We denote by $\tilde{\tau}_p(t)$ the approximation to the function of interest calculated from one realization of the reference simulation using J SDE realizations, and by $\hat{\tau}_p(t)$ the function of interest obtained via one realization of the micro/macro acceleration technique. As extrapolation techniques, we use first order projective extrapolation and first and second order multistep state extrapolation.

Figure 6.1 shows the results for first order projective extrapolation. The left figure clearly shows, as expected, that the deterministic error grows with increasing Δt. However, from the right figure follows that, for an individual realization of the experiment, the error is dominated by the statistical error. The zoom shows that, for small t, the sample standard deviation grows linearly as a function of time, with a slope that is larger for larger Δt. This is in agreement with the theoretical result on the local propagation of statistical error.

Next, we look at first order multistep state extrapolation, for which the results are shown in Fig. 6.2. The left figure indicates that, when comparing with projective extrapolation, the deterministic error grows much more rapidly with increasing Δt. On the right, we see that, while the sample standard deviation is larger than for the reference simulation, the sample standard deviation does not depend crucially on Δt,

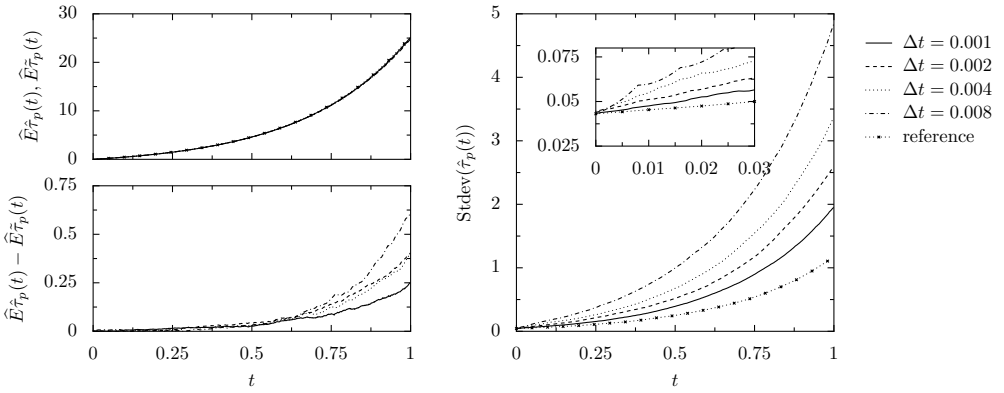


FIG. 6.1. *Results of micro/macro acceleration of the linear equation* (5.9) *using* $L = 2$ *moments and projective extrapolation for different values of the time step* Δt*, as well as a full microscopic (reference) simulation. Top left: evolution of the sample means of the stresses* $\tilde{\tau}_p$ *and* $\hat{\tau}_p$*. Bottom left: deterministic error on* $\hat{\tau}_p$*. Right: evolution of the sample standard deviation of* $\hat{\tau}_p$*. Simulation details are given in the text.*

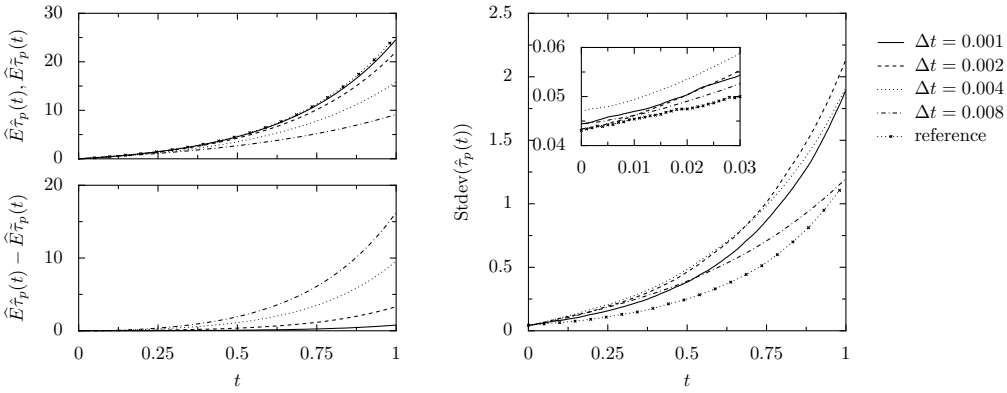


FIG. 6.2. *Results of micro/macro acceleration of the linear equation* (5.9) *using* $L = 2$ *moments and multistep state extrapolation for different values of the time step* Δt*, as well as a full microscopic (reference) simulation. Top left: evolution of the sample means of the stresses* $\tilde{\tau}_p$ *and* $\hat{\tau}_p$*. Bottom left: deterministic error on* $\hat{\tau}_p$*. Right: evolution of the sample standard deviation of* $\hat{\tau}_p$*. Simulation details are given in the text.*

as is also expected from the analysis of the local propagation of statistical error. Note that the lower sample standard deviation for $\Delta t = 8 \cdot 10^{-3}$ is related to the fact that $\hat{\tau}_p(t)$ itself is much lower as a consequence of the large deterministic error (see left figure). The zoom shows that, for small t, the statistical error grows linearly as a function of time, with a slope that is independent of Δt, and is identical to the slope for the reference simulation. These results are in agreement with the theoretical result on the local propagation of statistical error.

Finally, we consider second order multistep state extrapolation. The results are shown in Fig. 6.3. The left figure shows that the deterministic error is much better than for the first order version. However, the behavior of the statistical error is more intriguing. When zooming in to the behavior for small t, we observe that, over a short time interval, the sample standard deviation using the micro/macro acceleration technique increases at the same rate as the sample standard deviation in the reference

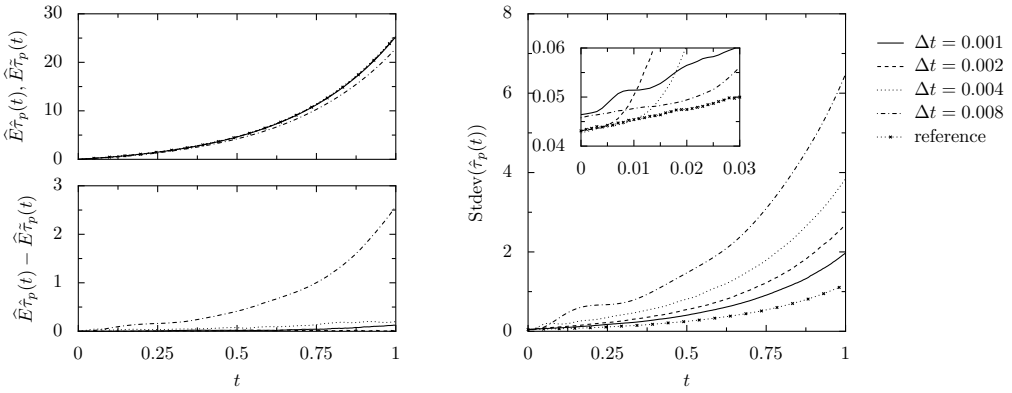


FIG. 6.3. *Results of micro/macro acceleration of the linear equation* (5.9) *using* $L = 2$ *moments and multistep state extrapolation for different values of the time step* Δt*, as well as a reference fully microscopic simulation. Top left: evolution of the sample means of the stresses* $\tilde{\tau}_p$ *and* $\hat{\tau}_p$*. Bottom left: deterministic error on* $\hat{\tau}_p$*. Right: evolution of the sample standard deviation of* $\hat{\tau}_p$*. Simulation details are given in the text.*

simulation. This corresponds to the theoretical result on the local propagation of statistical error. However, on longer time scales, the sample standard deviation for large t grows rapidly, and seems to be larger for larger Δt. As such, second order multistep state extrapolation behaves similarly to first order projective integration on long time scales. We suspect that the loss of this favorable error propagation is due to accumulation effects. This is indicated by the fact that the length of the time interval on which the local theoretical results is observed is longer for larger values of Δt. Indeed, the figure indicates that the effects of accumulated statistical errors start to appear *after a given number of extrapolations*, independently of the size of the extrapolation step. This behavior requires additional analysis.

7. Convergence results. Using the above, we are now ready to give a definition of convergence for the proposed algorithm:

DEFINITION 7.1. *Consider a sequence of restriction operators* $\left(\overline{\mathcal{R}}_{[L]}\right)_{L=1,2,\ldots}$ *and a sequence of projection operators* $\left(\overline{\mathcal{P}}_{[L]}\right)_{L=1,2,\ldots}$*, and denote the corresponding numerical approximation process obtained by using* $\overline{\mathcal{R}}_{[L]}$ *and* $\overline{\mathcal{P}}_{[L]}$ *in Algorithm 4.1 by* $\boldsymbol{Y}_{[L]}$*, and the maximum step size by* Δt*,* $\Delta t = \max_{n=1}^{N} \Delta t_n$*. The accelerated micro/macro Monte Carlo simulation is then called weakly convergent to the solution* $\boldsymbol{X}$ *of SDE* (1.1) *as* $\Delta t \to 0$ *and* $L \to \infty$ *at any time* $t \in I^{\Delta t}$ *with time order* p *if for each* $f \in C_P^{2(p+1)}(\mathbb{R}^d, \mathbb{R})$ *there exist constants* $\Delta t_0 > 0$*,* L_0*,* C_L*, and* $\tilde{C}_L$*, with* $C_L \to 0$ *for* $L \to \infty$*, such that*

$$|\,\mathrm{E}\, f\big(\boldsymbol{Y}_{[L]}(t)\big) - \mathrm{E}\, f\big(\boldsymbol{X}(t)\big)| \leq C_L + \tilde{C}_L(\Delta t)^p \tag{7.1}$$

holds for all $t \in I^{\Delta t}$*, all* $L \geq L_0$*, and all* $\Delta t \in [0, \Delta t_0]$*.*

We first discuss convergence when extrapolation is performed as in coarse projective integration (Subsection 7.1). Due to the multistep nature of the extrapolation, proving convergence for the multistep state extrapolation method is more involved; Subsection 7.2 contains a result for a linear SDE.

7.1. Convergence using projective extrapolation. The following theorem generalizes the theorem for the convergence of one step methods due to Milstein

(see [44, 47] or also [10]).

THEOREM 7.2. *Suppose the following conditions hold:*

(i) The coefficient functions $\boldsymbol{a}(\boldsymbol{x})$ and $\boldsymbol{b}^i(\boldsymbol{x})$ (where $\boldsymbol{b}^i$ denotes the i-th column of $\boldsymbol{b}$) are continuous, satisfy a Lipschitz condition with respect to $\boldsymbol{x}$, and belong to $C_P^{p+1,2(p+1)}(I\times\mathbb{R}^d,\mathbb{R})$, $i=1,\dots,m$. For non Itô SDEs, we require in addition that b^i is differentiable and that also $b^{i'}b^i$ satisfies a Lipschitz condition and belongs to $C_P^{p,2(p+1)}(I\times\mathbb{R}^d,\mathbb{R})$, $i=1,\dots,m$.

(ii) For sufficiently large r the moments $\mathrm{E}(\|\boldsymbol{Y}_{[L]}^{n,k}\|^{2r})$ exist for $k=0,\dots,K$ and $n=0,1,\dots,N$ and are uniformly bounded with respect to L, N.

(iii) SDE (1.1) is uniformly weakly continuous for the set of all $\boldsymbol{Y}_{[L]}^{n,K}$.

(iv) The one step method ϕ is weakly consistent of order p_ϕ.

(v) The sequence of (self-consistent) projection operators is continuous for the numerical approximation process and all sequences of macroscopic states, and consistent for all sequences of triples $\Big(\boldsymbol{Y}_{[L]}^{n,K},\boldsymbol{X}^{t^{n,K},\boldsymbol{Y}_{[L]}^{n,K}}(t^{n+1,0}),\overline{\mathcal{R}}_{[L]}(\boldsymbol{Y}_{[L]}^{n+1,0})\Big)_{L=1,2,\dots}$.

(vi) The extrapolation is consistent of order $p_e\geq 1$ and continuous for the class of all functions $U(t)=\overline{\mathcal{R}}_{[L]}\big(\boldsymbol{X}^{t^{i,k},\boldsymbol{Y}_{[L]}^{i,k}}(t)\big)$.

Then, the micro/macro acceleration algorithm with projective extrapolation is weakly convergent with time order $p=\min\{p_e,p_\phi\}$.

Proof. Let $g(s,\boldsymbol{x}):=\mathrm{E}\left(f\big(\boldsymbol{X}(t^{n+1})\big)|\boldsymbol{X}(s)=\boldsymbol{x}\right)$ for $s\in I$, $\boldsymbol{x}\in\mathbb{R}^d$, and $t^{n+1}\in I^{\Delta t}$ with $s\leq t^{n+1}$. Due to condition (i) $g\in C_P^{0,2(p+1)}$ [44]. Therefore, the consistency of ϕ implies that g satisfies

$$|\,\mathrm{E}\,g\big(s,\boldsymbol{X}^{t,\boldsymbol{x}}(t+\delta t)\big)-\mathrm{E}\,g\big(s,\phi(t,\boldsymbol{x};\delta t)\big)|\leq C_g(\boldsymbol{x})\,\delta t^{p_\phi+1} \tag{7.2}$$

uniformly w. r. t. $s\in[t^0,t^{n+1}]$ for some $C_g\in C_P^0(\mathbb{R}^d,\mathbb{R})$. Analogously, Corollary 5.4 and the continuity of the projection imply that g satisfies

$$\begin{aligned}
&|\,\mathrm{E}\,g\big(s,\boldsymbol{X}^{t^{i,K},\boldsymbol{Y}_{[L]}^{i,K}}(t^{i+1,0})\big)-\mathrm{E}\,g\big(s,\boldsymbol{Y}_{[L]}^{i+1,0}\big)|\\
&\leq|\,\mathrm{E}\,g\big(s,\boldsymbol{X}^{t^{i,K},\boldsymbol{Y}_{[L]}^{i,K}}(t^{i+1,0})\big)-\mathrm{E}\,g\Big(s,\overline{\mathcal{P}}_{[L]}\Big(\boldsymbol{Y}_{[L]}^{i,K},\overline{\mathcal{R}}_{[L]}\big(\boldsymbol{X}^{t^{i,K},\boldsymbol{Y}_{[L]}^{i,K}}(t^{i+1,0})\big)\Big)\Big)|\\
&+|\,\mathrm{E}\,g\Big(s,\overline{\mathcal{P}}_{[L]}\Big(\boldsymbol{Y}_{[L]}^{i,K},\overline{\mathcal{R}}_{[L]}\big(\boldsymbol{X}^{t^{i,K},\boldsymbol{Y}_{[L]}^{i,K}}(t^{i+1,0})\big)\Big)\Big)-\mathrm{E}\,g\Big(s,\overline{\mathcal{P}}_{[L]}\big(\boldsymbol{Y}_{[L]}^{i,K},\overline{\mathcal{R}}_{[L]}(\boldsymbol{Y}_{[L]}^{i+1,0})\big)\Big)|\\
&\leq C_L\Delta t+C_1\|\overline{\mathcal{R}}_{[L]}\big(\boldsymbol{X}^{t^{i,K},\boldsymbol{Y}_{[L]}^{i,K}}(t^{i+1,0})\big)-\overline{\mathcal{R}}_{[L]}(\boldsymbol{Y}_{[L]}^{i+1,0})\|.
\end{aligned}$$

The last summand can be expanded as follows:

$$\begin{aligned}
&\|\overline{\mathcal{R}}_{[L]}\big(\boldsymbol{X}^{t^{i,K},\boldsymbol{Y}_{[L]}^{i,K}}(t^{i+1,0})\big)-\overline{\mathcal{R}}_{[L]}(\boldsymbol{Y}_{[L]}^{i+1,0})\|\\
\leq&\|\overline{\mathcal{R}}_{[L]}\big(\boldsymbol{X}^{t^{i,K},\boldsymbol{Y}_{[L]}^{i,K}}(t^{i+1,0})\big)-\overline{\mathcal{R}}_{[L]}\big(\boldsymbol{X}^{t^{i,0},\boldsymbol{Y}_{[L]}^{i,0}}(t^{i+1,0})\big)\|\\
&+\|\overline{\mathcal{R}}_{[L]}\big(\boldsymbol{X}^{t^{i,0},\boldsymbol{Y}_{[L]}^{i,0}}(t^{i+1,0})\big)-\mathcal{E}\left(\Big(\overline{\mathcal{R}}_{[L]}\big(\boldsymbol{X}^{t^{i,0},\boldsymbol{Y}_{[L]}^{i,0}}(t^{i,k})\big)\Big)_{k=0}^K;\Delta t_i,\delta t\right)\|\\
&+\|\mathcal{E}\left(\Big(\overline{\mathcal{R}}_{[L]}\big(\boldsymbol{X}^{t^{i,0},\boldsymbol{Y}_{[L]}^{i,0}}(t^{i,k})\big)\Big)_{k=0}^K;\Delta t_i,\delta t\right)-\mathcal{E}\left(\Big(\overline{\mathcal{R}}_{[L]}\big(\boldsymbol{Y}_{[L]}^{i,k}\big)\Big)_{k=0}^K;\Delta t_i,\delta t\right)\|.
\end{aligned}$$

Let $\tilde{g}_{[L]}(s,\boldsymbol{x};t):=\left(\mathrm{E}\left(g_l\big(\boldsymbol{X}(t)\big)|\boldsymbol{X}(s)=\boldsymbol{x}\right)\right)_{l=1}^L$ for $s\in I$, $\boldsymbol{x}\in\mathbb{R}^d$, and $t\in I^{\Delta t}$ with $s\leq t$. With this definition, the continuity and consistency of the extrapolation imply

$$
\begin{aligned}
&\|\overline{\mathcal{R}}_{[L]}\big(\boldsymbol{X}^{t^{i,K},\boldsymbol{Y}^{i,K}_{[L]}}(t^{i+1,0})\big) - \overline{\mathcal{R}}_{[L]}(\boldsymbol{Y}^{i+1,0}_{[L]})\| \\
\le&\|\overline{\mathcal{R}}_{[L]}\big(\boldsymbol{X}^{t^{i,K},\boldsymbol{Y}^{i,K}_{[L]}}(t^{i+1,0})\big) - \overline{\mathcal{R}}_{[L]}\big(\boldsymbol{X}^{t^{i,0},\boldsymbol{Y}^{i,0}_{[L]}}(t^{i+1,0})\big)\| + C_2(\Delta t)^{p_e+1} \\
&+ C_3\left(\frac{\Delta t}{\delta t}\right)^{p_e}\sum_{k=1}^{K}\|\overline{\mathcal{R}}_{[L]}\big(\boldsymbol{X}^{t^{i,0},\boldsymbol{Y}^{i,0}_{[L]}}(t^{i,k})\big) - \overline{\mathcal{R}}_{[L]}\big(\boldsymbol{Y}^{i,k}_{[L]}\big)\| \\
=&\|\operatorname{E}\tilde{g}_{[L]}\big(t^{i,K},\boldsymbol{Y}^{i,K}_{[L]};t^{i+1,0}\big) - \operatorname{E}\tilde{g}_{[L]}\big(t^{i,K},\boldsymbol{X}^{t^{i,0},\boldsymbol{Y}^{i,0}_{[L]}}(t^{i,K});t^{i+1,0}\big)\| + C_2(\Delta t)^{p_e+1} \\
&+ C_3\left(\frac{\Delta t}{\delta t}\right)^{p_e}\sum_{k=1}^{K}\|\left(\operatorname{E} g_l\big(\boldsymbol{X}^{t^{i,0},\boldsymbol{Y}^{i,0}_{[L]}}(t^{i,k})\big) - \operatorname{E} g_l\big(\boldsymbol{Y}^{i,k}_{[L]}\big)\right)_{l=1}^{L}\|,
\end{aligned}
$$

where we made also use of $\boldsymbol{X}^{t^{i,0},\boldsymbol{Y}^{i,0}_{[L]}}(t^{i+1,0}) = \boldsymbol{X}^{t^{i,K},\boldsymbol{X}^{t^{i,0},\boldsymbol{Y}^{i,0}_{[L]}}(t^{i,K})}(t^{i+1,0})$. Due to the consistency of ϕ, altogether we obtain

$$
|\operatorname{E} g\big(s,\boldsymbol{X}^{t^{i,K},\boldsymbol{Y}^{i,K}_{[L]}}(t^{i+1,0})\big) - \operatorname{E} g\big(s,\boldsymbol{Y}^{i+1,0}_{[L]}\big)| \le C_L\Delta t + \tilde{C}_L(\Delta t)^{p_e}\delta t^{p_\phi} + \tilde{C}_L\delta t^{p_\phi+1} \tag{7.3}
$$

uniformly w. r. t. $s \in [t^0, t^{n+1}]$ for some constants C_L and $\tilde{C}_L$ with $C_L \to 0$ for $L \to \infty$. For ease of notation, in the following we will neglect the L-dependency of Y. Then

$$
\begin{aligned}
&\operatorname{E} f\big(\boldsymbol{X}^{t^0,X_0}(t^{n+1})\big) - \operatorname{E} f\big(\boldsymbol{Y}^{n+1,0}\big) \\
=&\sum_{i=0}^{n}\sum_{k=0}^{K-1}\Big(\operatorname{E} f\big(\boldsymbol{X}^{t^{i,k},\boldsymbol{Y}^{i,k}}(t^{n+1})\big) - \operatorname{E} f\big(\boldsymbol{X}^{t^{i,k+1},\boldsymbol{Y}^{i,k+1}}(t^{n+1})\big)\Big) \\
&+\sum_{i=0}^{n-1}\Big(\operatorname{E} f\big(\boldsymbol{X}^{t^{i,K},\boldsymbol{Y}^{i,K}}(t^{n+1})\big) - \operatorname{E} f\big(\boldsymbol{X}^{t^{i+1,0},\boldsymbol{Y}^{i+1,0}}(t^{n+1})\big)\Big) \\
&+\operatorname{E} f\big(\boldsymbol{X}^{t^{n,K},\boldsymbol{Y}^{n,K}}(t^{n+1})\big) - \operatorname{E} f\big(\boldsymbol{Y}^{n+1,0}\big).
\end{aligned}
$$

Making use of $\boldsymbol{X}^{t^{i,k},\boldsymbol{Y}^{i,k}}(t^{n+1}) = \boldsymbol{X}^{t^{i,k+1},\boldsymbol{X}^{t^{i,k},\boldsymbol{Y}^{i,k}}(t^{i,k+1})}(t^{n+1})$ and combining the last two summands we obtain

$$
\begin{aligned}
&\operatorname{E} f\big(\boldsymbol{X}^{t^0,X_0}(t^{n+1})\big) - \operatorname{E} f\big(\boldsymbol{Y}^{n+1,0}\big) \\
=&\sum_{i=0}^{n}\sum_{k=0}^{K-1}\Big(\operatorname{E} f\big(\boldsymbol{X}^{t^{i,k+1},\boldsymbol{X}^{t^{i,k},\boldsymbol{Y}^{i,k}}(t^{i,k+1})}(t^{n+1})\big) - \operatorname{E} f\big(\boldsymbol{X}^{t^{i,k+1},\boldsymbol{Y}^{i,k+1}}(t^{n+1})\big)\Big) \\
&+\sum_{i=0}^{n}\Big(\operatorname{E} f\big(\boldsymbol{X}^{t^{i,K},\boldsymbol{Y}^{i,K}}(t^{n+1})\big) - \operatorname{E} f\big(\boldsymbol{X}^{t^{i+1,0},\boldsymbol{Y}^{i+1,0}}(t^{n+1})\big)\Big).
\end{aligned}
$$

Using $\boldsymbol{X}^{t^{i,K},\boldsymbol{Y}^{i,K}}(t^{n+1}) = \boldsymbol{X}^{t^{i+1,0},\boldsymbol{X}^{t^{i,K},\boldsymbol{Y}^{i,K}}(t^{i+1,0})}(t^{n+1})$ and the definition of g this implies

$$
\begin{aligned}
&\operatorname{E} f\big(\boldsymbol{X}^{t^0,X_0}(t^{n+1})\big) - \operatorname{E} f\big(\boldsymbol{Y}^{n+1,0}\big) \\
=&\sum_{i=0}^{n}\sum_{k=0}^{K-1}\Big(\operatorname{E} g\big(t^{i,k+1},\boldsymbol{X}^{t^{i,k},\boldsymbol{Y}^{i,k}}(t^{i,k+1})\big) - \operatorname{E} g\big(t^{i,k+1},\boldsymbol{Y}^{i,k+1}\big)\Big) \\
&+\sum_{i=0}^{n}\Big(\operatorname{E} f\big(\boldsymbol{X}^{t^{i+1,0},\boldsymbol{X}^{t^{i,K},\boldsymbol{Y}^{i,K}}(t^{i+1,0})}(t^{n+1})\big) - \operatorname{E} f\big(\boldsymbol{X}^{t^{i+1,0},\boldsymbol{Y}^{i+1,0}}(t^{n+1})\big)\Big)
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{i=0}^{n} \sum_{k=0}^{K-1} \Big(\mathrm{E}\, g\big(t^{i,k+1}, \boldsymbol{X}^{t^{i,k},\boldsymbol{Y}^{i,k}}(t^{i,k+1})\big) - \mathrm{E}\, g\big(t^{i,k+1}, \boldsymbol{Y}^{i,k+1}\big)\Big) \\
&\quad + \sum_{i=0}^{n} \Big(\mathrm{E}\, g\big(t^{i+1,0}, \boldsymbol{X}^{t^{i,K},\boldsymbol{Y}^{i,K}}(t^{i+1,0})\big) - \mathrm{E}\, g\big(t^{i+1,0}, \boldsymbol{Y}^{i+1,0}\big)\Big).
\end{aligned}
$$

Thus, (7.2) and (7.3) imply

$$
\begin{aligned}
&|\, \mathrm{E}\, f\big(\boldsymbol{X}^{t^0,X_0}(t^{n+1})\big) - \mathrm{E}\, f\big(\boldsymbol{Y}_{[L]}(t^{n+1})\big)| \\
&\le \sum_{i=0}^{n} \sum_{k=0}^{K-1} \mathrm{E}\, C_g(\boldsymbol{Y}_{[L]}^{i,k}) \delta t^{p_\phi+1} + \sum_{i=0}^{n} \Big(C_L \Delta t + \tilde{C}_L (\Delta t)^{p_e} \delta t^{p_\phi} + \tilde{C}_L \delta t^{p_\phi+1} \Big),
\end{aligned}
$$

which yields together with condition (ii) the desired convergence. □

7.2. A result using multistep state extrapolation. For multistep state extrapolation, the analysis is complicated by the multistep nature of the method. In this subsection we therefore restrict ourselves to consider linear SDEs (5.9) with normally distributed initial values and restriction operator

$$
\overline{\mathcal{R}}(Y) = \begin{pmatrix} \mathrm{E}\, Y \\ \mathrm{Var}\, Y \end{pmatrix}. \tag{7.4}
$$

If we assume now equidistant coarse time steps and that the extrapolation step is given by (6.12), then we obtain

$$
\mathrm{E}\, \boldsymbol{Y}^{n+1,0} = \sum_{s=0}^{p_e} l_s(\beta)\, \mathrm{E}\, \boldsymbol{Y}^{n-s,K}, \quad \mathrm{Var}\, \boldsymbol{Y}^{n+1,0} = \sum_{s=0}^{p_e} l_s(\beta)\, \mathrm{Var}\, \boldsymbol{Y}^{n-s,K}
$$

with l_s and β given in (6.5) and (6.13). Application of the one step method ϕ to (5.9) yields

$$
\begin{aligned}
\boldsymbol{Y}^{n+1,K} &= \hat{R}_\phi(a_1, t^{n+1,K-1}, \delta t, \eta^{n+1,K-1}) \boldsymbol{Y}^{n+1,K-1} \\
&\quad + \hat{S}_\phi(a_1, a_2, b, t^{n+1,K-1}, \delta t, \eta^{n+1,K-1}) \\
&= \prod_{k=0}^{K-1} \hat{R}_\phi(a_1, t^{n+1,k}, \delta t, \eta^{n+1,k}) \boldsymbol{Y}^{n+1,0} \\
&\quad + \sum_{k=0}^{K-1} \hat{S}_\phi(a_1, a_2, b, t^{n+1,k}, \delta t, \eta^{n+1,k}) \prod_{i=k}^{K-1} \hat{R}_\phi(a_1, t^{n+1,i}, \delta t, \eta^{n+1,i}),
\end{aligned} \tag{7.5}
$$

where $\hat{R}_\phi$ and $\hat{S}_\phi$ are functions depending on ϕ, similar to the stability function in the deterministic case, and $\eta^{n+1,k}$ are (vectors of) i. i. d. random variables used by ϕ. Assuming that $\hat{R}_\phi(a_1, t^{n+1,k}, \delta t, \eta^{n,k})$ is independent of $\eta^{n,k}$, which holds, e.g., for typical Runge-Kutta methods, we obtain the multistep formulas

$$
\begin{aligned}
\mathrm{E}\, \boldsymbol{Y}^{n+1,K} &= \prod_{k=0}^{K-1} \hat{R}_\phi(a_1, t^{n+1,k}, \delta t) \sum_{s=0}^{p_e} l_s(\beta)\, \mathrm{E}\, \boldsymbol{Y}^{n-s,K} \\
&\quad + \sum_{k=0}^{K-1} \mathrm{E}\, \hat{S}_\phi(a_1, a_2, b, t^{n+1,k}, \delta t, \eta^{n+1,k}) \prod_{i=k}^{K-1} \hat{R}_\phi(a_1, t^{n+1,i}, \delta t),
\end{aligned} \tag{7.6}
$$

$$\operatorname{Var} \boldsymbol{Y}^{n+1,K} = \prod_{k=0}^{K-1} \hat{R}_\phi(a_1, t^{n+1,k}, \delta t)^2 \sum_{s=0}^{p_e} l_s(\beta) \operatorname{Var} \boldsymbol{Y}^{n-s,K} + \sum_{k=0}^{K-1} \operatorname{Var} \hat{S}_\phi(a_1, a_2, b, t^{n+1,k}, \delta t, \eta^{n+1,k}) \prod_{i=k}^{K-1} \hat{R}_\phi(a_1, t^{n+1,i}, \delta t)^2. \quad (7.7)$$

As due to the consistency of ϕ we have $\hat{R}_\phi(a_1, t^{n+1,k}, 0) = 1$ and

$$\operatorname{E} \hat{S}_\phi(a_1, a_2, b, t^{n+1,k}, 0, \eta^{n+1,k}) = \operatorname{Var} \hat{S}_\phi(a_1, a_2, b, t^{n+1,k}, 0, \eta^{n+1,k}) = 0,$$

the corresponding characteristic polynomial is given for both equations by

$$P(\xi; \beta, p_e) = \xi^{p_e+1} - \sum_{s=0}^{p_e} l_s(\beta) \xi^s.$$

As in the deterministic case (theory of linear multistep methods) we then obtain the following theorem.

THEOREM 7.3. *Assume that all roots ξ of $P(\xi; \beta, p_e) = 0$ lie within the unit circle and that all roots with absolute value one are simple. If then the one-step method is weakly consistent of order p_ϕ and given by* (7.5) *with R_ϕ independent of η, then coarse multistep state extrapolation with restriction operator* (7.4) *and extrapolation given by* (6.12) *is convergent of order $p = \min\{p_e, p_\phi\}$ for linear SDEs* (5.9) *with normally distributed initial values.*

8. Numerical results. In this section, we provide further numerical results. We first illustrate the dependence of the local error in one accelerated time step on the step size (Subsection 8.1). Subsequently, we perform a number of long-term simulations (Subsection 8.2).

Below, we consider equation (2.6) in one space dimension, with $F(X)$ the FENE force (2.7) with $\gamma = 49$, We $= 1$. As the velocity field, we choose $\kappa(t) = 2 \cdot (1.1 + \sin(\pi t))$, and we impose standard normally distributed initial states. We discretize in time with the classical Euler-Maruyama scheme with time step $\delta t = 2 \cdot 10^{-4}$. As before, the macroscopic state $\boldsymbol{U}_{[L]}$ consists of the first L *even* centralized moments. We study the micro/macro acceleration algorithm with first order projective extrapolation, and with first and second order multistep state extrapolation. In all cases, we perform $K = 1$ microscopic steps before extrapolation.

8.1. Local error. We simulate $J = 1 \cdot 10^5$ realizations up to time $t^* = 1.6$, and record the microscopic state $\boldsymbol{\mathcal{Y}}^-$ at time $t^- = 1.4$, as well as the macroscopic states $\boldsymbol{U}_{[L]}(t^- + \Delta t)$ for $L = 3, 4, 5$, and the approximated function of interest $\tilde{\tau}_p(t^- + \Delta t)$ for $\Delta t \in [0, t^* - t^-]$. We then extrapolate from time t^- to time $t = t^- + \Delta t$, and project the ensemble $\boldsymbol{\mathcal{Y}}^-$ onto $\boldsymbol{U}_{[L]}(t^- + \Delta t)$. Subsequently, we compute the corresponding value of the stress as $\hat{\tau}_p(t^- + \Delta t)$. We record the relative error with respect to the reference solution, $|\hat{\tau}_p(t) - \tilde{\tau}_p(t)|/\tilde{\tau}_p(t)$, as a function of Δt. To reduce the statistical error, we report the averaged results of 50 realizations of this experiment. Then, the statistical error has an order of magnitude of about 10^{-5}. The results are shown in Fig. 8.1.

For projective integration, we clearly see a first order behavior as a function of Δt; this is a consequence of the amplification of the statistical error during projective extrapolation. Note that, due to the presence of three competing sources of errors (extrapolation, projection, and statistical error), which may be of opposite signs, the

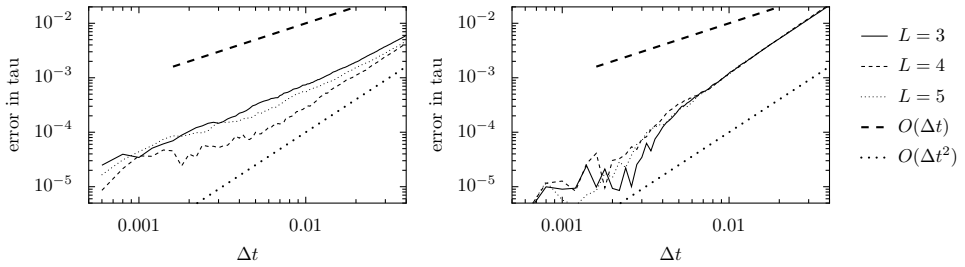


FIG. 8.1. *Error of the stress after extrapolating and projecting a prior ensemble of* $J = 1 \cdot 10^5$ *FENE dumbbells onto the first* L *even centralized moments of a reference ensemble, as a function of* Δt *for several values of* L*. Displayed is the result averaged over* 50 *realizations of the experiment. Left: First order projective extrapolation. Right: First order multistep state extrapolation. Simulation details are given in the text.*

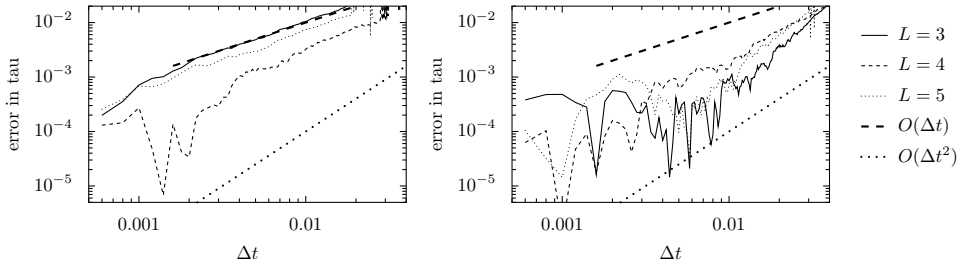


FIG. 8.2. *Error of the stress after extrapolating and projecting a prior ensemble of* $J = 1 \cdot 10^3$ *FENE dumbbells onto the first* L *even centralized moments of a reference ensemble, as a function of* Δt *for several values of* L*. Displayed is the result averaged over* 20 *realizations of the experiment. Left: First order projective extrapolation. Right: First order multistep state extrapolation. Simulation details are given in the text.*

effect of extrapolating more macroscopic state variables is not so clearly visible as in Fig. 5.2. Note that for $L = 4$ the statistical and projection errors seem to be a bit lower, such that the second order behavior of the extrapolation error is already apparent for the largest displayed time steps.

For multistep state extrapolation, the situation is slightly different. Here, we see that, for modest gains (small Δt), the statistical error remains more or less unaffected. For $\Delta t > 2 \cdot 10^{-3}$ (gain factor 10), however, the error increases as Δt^2, as a consequence of the large extrapolation error. Note that, as soon as the extrapolation error dominates, the error appears to be independent of the number of moments used.

To emphasize the effect of the statistical error, we repeat the experiment using $J = 1000$ realizations (averaged over 20 realizations of the experiment). The results are shown in Fig. 8.2. Compared to Fig. 8.1, we see qualitatively the same behavior. Again, the error of projective extrapolation increases linearly (but it is now an order of magnitude larger), while, for multistep state extrapolation, larger gains appear to be possible since the statistical error now dominates for a wider range of extrapolation step sizes Δt.

8.2. Long-term simulation. We now turn to a long-term simulation, and compare the behavior of the sample mean and sample standard deviation of a full microscopic simulation (which we will call the reference simulation) with the micro/macro acceleration algorithm. We consider 500 realizations of a computational experiment

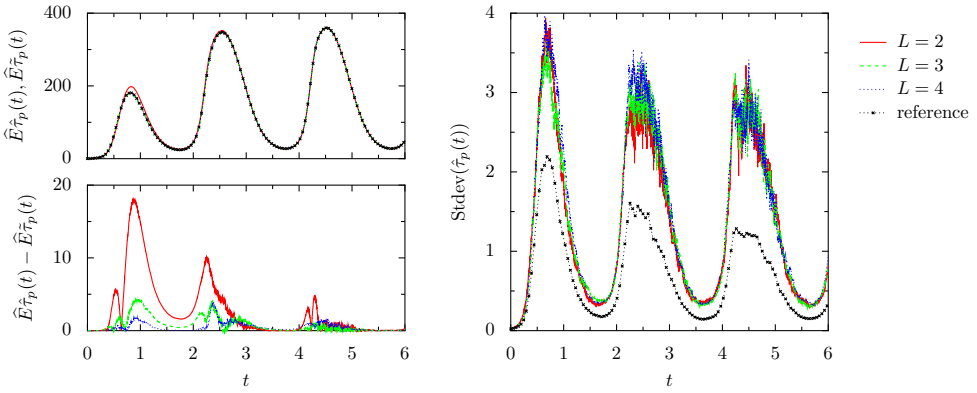


FIG. 8.3. *Results of micro/macro acceleration of the FENE model* (2.6) *using* $\Delta t = 1 \cdot 10^{-3}$ *and projective extrapolation for different numbers* L *of macroscopic state variables, as well as a full microscopic (reference) simulation. Top left: evolution of the sample means of the stresses* $\tilde{\tau}_p$ *and* $\hat{\tau}_p$*. Bottom left: deterministic error on* $\hat{\tau}_p$*. Right: evolution of the sample standard deviation of* $\hat{\tau}_p$*. Simulation details are given in the text.*

with $J = 1000$ SDE realizations. We denote by $\tilde{\tau}_p(t)$ the approximation to the function of interest calculated from one realization of the reference simulation using J SDE realizations, and by $\hat{\tau}_p(t)$ the function of interest obtained via one realization of the micro/macro acceleration technique. As extrapolation techniques, we use first order projective extrapolation and second order multistep state extrapolation. (For the setup in this example, the deterministic error of first order multistep state extrapolation is too high to be considered further.)

Figure 8.3 (top left) shows the evolution of the stress as a function of time. We see that the simulation exhibits a periodic behavior, with a fast increase of the stress followed by a relaxation. During the fast increase of the stress, we observed that the projection operator fails for time steps $\Delta t > 1 \cdot 10^3$, whereas larger accelerations are possible during the relaxation. Therefore, we will use adaptive macroscopic time steps, as outlined in Subsection 4.2.

In a first experiment, we use $\Delta t = 1 \cdot 10^{-3}$ and vary the number L of macroscopic state variables. The results are shown in Fig. 8.3. We make two main observations. First, the deterministic error decreases with increasing L, whereas the sample standard deviation is independent of L. (The different lines in the plot are nearly indistinguishable.) Note also that the variance on the sample standard deviation is quite large in this example. Second, from Fig. 8.3, we see that the error of the micro/macro acceleration algorithm with respect to the reference simulation also decreases as a function of time, until it reaches a level of the order of the statistical error. This behavior can be attributed to the fact that, in this example, the macroscopic behavior of the system on long time scales is determined by only a few macroscopic state variables. The results for multistep state extrapolation (not shown) in terms of L are similar.

In a second experiment, we fix $L = 3$ and consider varying Δt. (This experiment and its conclusions closely resemble the one in Subsection 6.3.) The results are shown in Fig. 8.4. The left figure again shows that the deterministic error grows with increasing Δt, whereas the right figure illustrates that the sample standard deviation is larger for larger Δt. For second order multistep state extrapolation, we obtain Fig. 8.5. The behavior of the sample standard deviation is similar to the linear case. When zooming in to the behavior for small t, we observe that, over a short time interval, the sam-

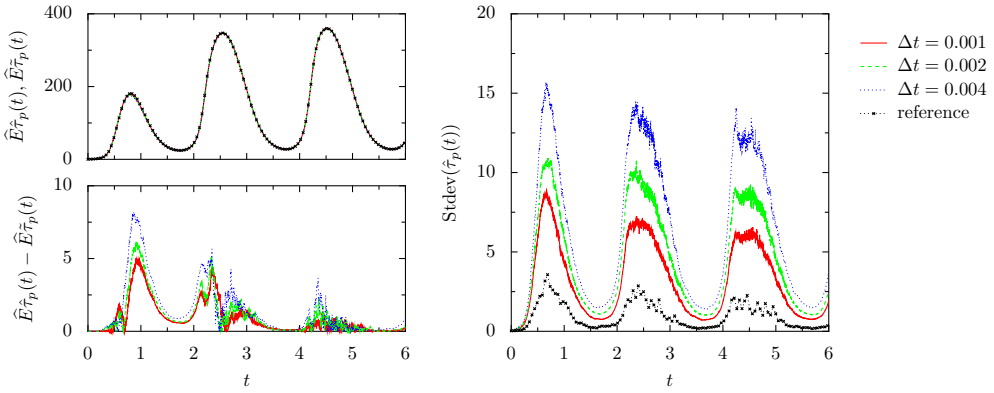


FIG. 8.4. *Results of micro/macro acceleration of the FENE model* (2.6) *using* $L = 3$ *moments and projective extrapolation for different values of* Δt*, as well as a full microscopic (reference) simulation. Top left: evolution of the sample means of the stresses* $\tilde{\tau}_p$ *and* $\hat{\tau}_p$*. Bottom left: deterministic error on* $\hat{\tau}_p$*. Right: evolution of the sample standard deviation of* $\hat{\tau}_p$*. Simulation details are given in the text.*

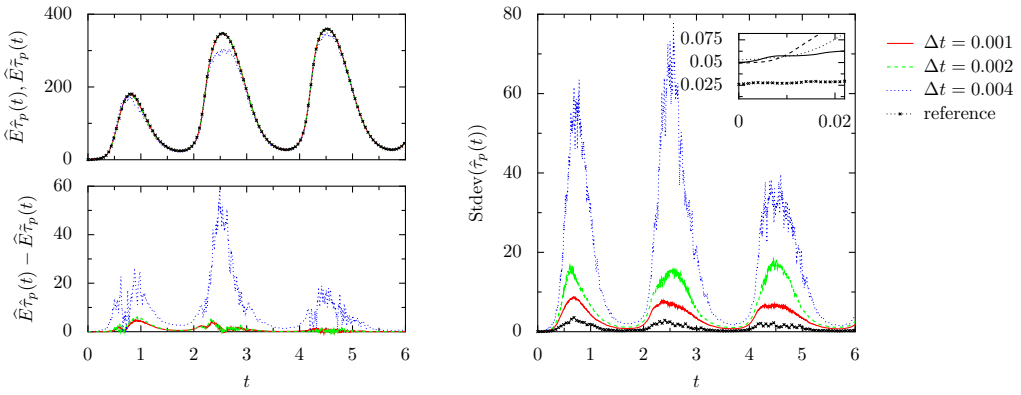


FIG. 8.5. *Results of micro/macro acceleration of the FENE model* (2.6) *using* $L = 3$ *moments and second order multistep state extrapolation for different numbers* L *of macroscopic state variables, as well as a full microscopic (reference) simulation. Top left: evolution of the sample means of the stresses* $\tilde{\tau}_p$ *and* $\hat{\tau}_p$*. Bottom left: deterministic error on* $\hat{\tau}_p$*. Right: evolution of the sample standard deviation of* $\hat{\tau}_p$*. Simulation details are given in the text.*

ple standard deviation using the micro/macro acceleration technique increases at the same rate as the sample standard deviation in the reference simulation, which corresponds to the theoretical result on the local propagation of statistical error. However, on longer time scales, we again see that the sample standard deviation for large t grows rapidly, and seems to be larger for larger Δt, due to accumulation effects. Note that, in this case, second order multistep state extrapolation even behaves worse than first order projective integration on long time scales for sufficiently large Δt.

9. Conclusions and outlook. We presented and analyzed a micro/macro acceleration technique for the Monte Carlo simulation of stochastic differential equations (SDEs) in which short bursts of simulation using an ensemble of microscopic SDE realizations are combined with an extrapolation of an estimated macroscopic state forward in time. The method is designed for problems in which the required time

step for each realization of the SDE is small compared to the time scales on which the function of interest evolves. For such systems, one often needs to take a very small microscopic time step, which results in a deterministic error that is much smaller than the statistical error.

We showed that the proposed procedure converges in the absence of statistical error, provided the projection operator satisfies a number of natural conditions, and we introduced a projection operator that satisfies these conditions for Gaussian random variables. We also conjectured that this projection operator is suitable for general distributions, and provided numerical evidence to support this conjecture. Concerning the statistical error, a local analysis of projective extrapolation shows that the amplification of statistical error depends on the ratio α of macroscopic (extrapolation) and microscopic (simulation) time steps, while this is not the case for multistep state extrapolation. Numerical evidence, however, suggests that, when using higher order multistep state extrapolation, accumulation of statistical error over macroscopic time scales may nevertheless induce an α-dependent statistical error.

This paper has not focused on quantifying the computational gains that can be expected from this method. It is clear from the description that the method will be more efficient when there is a bigger separation in time scales between the microscopic and macroscopic levels. The numerical examples in this text do not exhibit such a strong time-scale separation; they were mainly chosen for their ability to clearly illustrate the effects of the different sources of numerical error. However, some conclusions on efficiency can nevertheless be drawn. For a given required variance on the solution, the computational cost using first order projective extrapolation is comparable to that of a full simulation, since the former requires more SDE realizations due to the α-dependent amplification of statistical error. For first order multistep state extrapolation, extrapolation without such a drastic amplification of statistical error is possible, at the cost of an amplified deterministic error. This can be acceptable if the macroscopic function of interest changes slowly compared to the time step of the SDE.

We note that, for the model problem of coupled micro/macro simulation of dilute polymer solutions, the amplification of statistical error using projective extrapolation need not be dramatic: while the computational cost of a micro/macro accelerated simulation and a full microscopic simulation are comparable for given variance, this is no longer true when coupling this Monte Carlo simulation to a PDE for the solvent. Indeed, when extrapolating the complete coupled system forward in time, a computational gain is obtained since the PDE for the solvent also does not need to be simulated on the whole time domain.

In future work, we will study stability and propagation of statistical error on long time scales. The numerical experiments indicate that these issues can be studied in a linear setting. Another open question is for which distributions the conjecture can be proved. From an algorithmic point of view, this work raises questions on the adaptive/automatic selection of all method parameters (number of moments to extrapolate, macroscopic time step, number of SDE realizations) to ensure a reliable computation with minimal computational cost. Also, a numerical comparison with other approaches, such as implicit approximations, could be envisaged.

Acknowledgments. This work was supported by the Research Council of the K.U. Leuven through grant OT/09/27 and by the Interuniversity Attraction Poles Programme of the Belgian Science Policy Office through grant IUAP/V/22. The scientific responsibility rests with its authors.

REFERENCES

[1] A. Abdulle and S. Cirilli, *S-ROCK: Chebyshev methods for stiff stochastic differential equations*, SIAM J. Sci. Comput., 30 (2008), pp. 997–1014.

[2] A. Abdulle and T. Li, *S-ROCK methods for stiff Itô SDEs*, Commun. Math. Sci., 6 (2008), pp. 845–868.

[3] K. Burrage and T. Tian, *Stiffly accurate Runge–Kutta methods for stiff stochastic differential equations*, Comput. Phys. Commun., 142 (2001), pp. 186–190.

[4] ———, *Predictor-corrector methods of Runge–Kutta type for stochastic differential equations*, SIAM J. Numer. Anal., 40 (2002), pp. 1516–1537 (electronic).

[5] ———, *Implicit stochastic Runge–Kutta methods for stochastic differential equations*, BIT, 44 (2004), pp. 21–39.

[6] R. E. Caflisch, *Monte Carlo and quasi-Monte Carlo methods*, in Acta numerica, 1998, vol. 7 of Acta Numer., Cambridge Univ. Press, Cambridge, 1998, pp. 1–49.

[7] K. Debrabant, *Runge-Kutta methods for third order weak approximation of SDEs with multidimensional additive noise*, BIT, 50 (2010), pp. 541–558.

[8] K. Debrabant and A. Kværnø, *B-series analysis of stochastic Runge-Kutta methods that use an iterative scheme to compute their internal stage values*, SIAM J. Numer. Anal., 47 (2008/09), pp. 181–203.

[9] ———, *B-series analysis of iterated Taylor methods*, Report 1-2010, Martin Luther University Halle-Wittenberg, Institute for Mathematics, Jan. 2010.

[10] K. Debrabant and A. Rössler, *Continuous weak approximation for stochastic differential equations*, J. Comput. Appl. Math., 214 (2008), pp. 259–273.

[11] ———, *Diagonally drift-implicit Runge-Kutta methods of weak order one and two for Itô SDEs and stability analysis*, Appl. Numer. Math., 59 (2009), pp. 595–607.

[12] ———, *Families of efficient second order Runge-Kutta methods for the weak approximation of Itô stochastic differential equations*, Appl. Numer. Math., 59 (2009), pp. 582–594.

[13] Q. Du, C. Liu, and P. Yu, *FENE dumbbell model and its several linear and nonlinear closure approximations*, Multiscale Modeling and Simulation, 4 (2006), pp. 709–731.

[14] W. E and B. Engquist, *The heterogeneous multi-scale methods*, Communications in Mathematical Sciences, 1 (2003), pp. 87–132.

[15] W. E, B. Engquist, X. Li, W. Ren, and E. Vanden-Eijnden, *Heterogeneous multiscale methods: A review*, Communications in Computational Physics, 2 (2007), pp. 367–450.

[16] I. Fatkullin and E. Vanden-Eijnden, *A computational strategy for multiscale systems with applications to lorenz 96 model*, Journal of Computational Physics, 200 (2004), pp. 605–638.

[17] C. W. Gear and I. G. Kevrekidis, *Projective methods for stiff differential equations: Problems with gaps in their eigenvalue spectrum*, SIAM Journal on Scientific Computing, 24 (2003), pp. 1091–1106.

[18] C. W. Gear, I. G. Kevrekidis, and C. Theodoropoulos, *"Coarse" integration/bifurcation analysis via microscopic simulators: micro-Galerkin methods*, Computers and Chemical Engineering, 26 (2002), pp. 941–963.

[19] M. B. Giles, *Improved multilevel Monte Carlo convergence using the Milstein scheme*, in Monte Carlo and quasi-Monte Carlo methods 2006, Springer, Berlin, 2008, pp. 343–358.

[20] ———, *Multilevel Monte Carlo path simulation*, Oper. Res., 56 (2008), pp. 607–617.

[21] E. Hairer, S. P. Nørsett, and G. Wanner, *Solving ordinary differential equations I: Nonstiff problems*, vol. 8 of Spinger Series in Computational Mathematics, Springer Verlag, second revised edition ed., 1993.

[22] D. B. Hernández and R. Spigler, *A-stability of Runge-Kutta methods for systems with additive noise*, BIT, 32 (1992), pp. 620–633.

[23] ———, *Convergence and stability of implicit Runge–Kutta methods for systems with multiplicative noise*, BIT, 33 (1993), pp. 654–669.

[24] M. Herrchen and H. C. Öttinger, *A detailed comparison of various FENE dumbbell models*, Journal of Non-Newtonian Fluid Mechanics, 68 (1997), pp. 17–42.

[25] D. J. Higham, *Mean-square and asymptotic stability of the stochastic theta method*, SIAM J. Numer. Anal., 38 (2000), pp. 753–769 (electronic).

[26] M. A. Hulsen, A. P. G. van Heel, and B. H. A. A. van den Brule, *Simulation of viscoelastic flows using Brownian configuration fields*, Journal of Non-Newtonian Fluid Mechanics, 70 (1997), pp. 79–101.

[27] Y. Hyon, Q. Du, and C. Liu, *An enhanced macroscopic closure approximation to the micro-macro FENE models for polymeric materials*, Multiscale Modeling and Simulation, 7 (2008), pp. 978–1002.

[28] P. Ilg, I. V. Karlin, and H. C. Öttinger, *Canonical distribution functions in polymer dynamics. (I). dilute solutions of flexible polymers*, Physica A, 315 (2002), pp. 367–385.
[29] B. Jourdain and T. Lelièvre, *Mathematical analysis of a stochastic differential equation arising in the micro-macro modelling of polymeric fluids*, in Probabilistic Methods in Fluids Proceedings of the Swansea 2002 Workshop, I. Davies, N. Jacob, A. Truman, O. Hassan, K. Morgan, and N. Weatherill, eds., World Scientific, 2003, pp. 205–223.
[30] R. Keunings, *On the Peterlin approximation for finitely extensible dumbbells*, Journal of Non-Newtonian Fluid Mechanics, 68 (1997), pp. 85–100.
[31] I. G. Kevrekidis, C. W. Gear, J. M. Hyman, P. G. Kevrekidis, O. Runborg, and C. Theodoropoulos, *Equation-free, coarse-grained multiscale computation: enabling microscopic simulators to perform system-level tasks*, Communications in Mathematical Sciences, 1 (2003), pp. 715–762.
[32] I. G. Kevrekidis and G. Samaey, *Equation-free multiscale computation: Algorithms and applications*, Annual Review on Physical Chemistry, 60 (2009), pp. 321–344.
[33] P. E. Kloeden and E. Platen, *Numerical solution of stochastic differential equations*, vol. 21 of Applications of Mathematics, Springer-Verlag, Berlin, 2 ed., 1999.
[34] Y. Komori, *Weak second-order stochastic Runge–Kutta methods for non-commutative stochastic differential equations*, J. Comput. Appl. Math., 206 (2007), pp. 158–173.
[35] ———, *Weak first- or second-order implicit Runge-Kutta methods for stochastic differential equations with a scalar Wiener process*, J. Comput. Appl. Math., 217 (2008), pp. 166–179.
[36] M. Laso and H. C. Öttinger, *Calculation of viscoelastic flow using molecular models: the CONNFFESSIT approach*, Journal of Non-Newtonian Fluid Mechanics, 47 (1993), pp. 1–20.
[37] M. Laso, J. Ramírez, and M. Picasso, *Implicit micro-macro methods*, Journal of Non-Newtonian Fluid Mechanics, 122 (2004), pp. 215–226.
[38] C. Le Bris and T. Lelièvre, *Multiscale modelling of complex fluids: A mathematical initiation*, in Multiscale Modeling and Simulation in Science, B. Engquist, P. Lötstedt, and O. Runborg, eds., vol. 66 of Lecture Notes in Computational Science and Engineering, Springer, 2009.
[39] J. Lehn and H. Wegmann, *Einführung in die Statistik*, Teubner Studienbücher Mathematik. [Teubner Mathematical Textbooks], B. G. Teubner, Stuttgart, 1985.
[40] T. Lelièvre, C. L. Bris, and E. Vanden-Eijnden, *Analysis of some discretization schemes for constrained stochastic differential equations*, Comptes Rendus Mathematique, 346 (2008), pp. 471–476.
[41] G. Lielens, P. Halin, I. Jaumain, R. Keunings, and V. Legat, *New closure approximations for the kinetic theory of finitely extensible dumbbells*, Journal of Non-Newtonian Fluid Mechanics, 76 (1998), pp. 249–279.
[42] G. Lielens, R. Keunings, and V. Legat, *The FENE-L and FENE-LS closure approximations to the kinetic theory of finitely extensible dumbbells*, Journal of Non-Newtonian Fluid Mechanics, 87 (1999), pp. 179–196.
[43] R. Mei, L. S. Luo, P. Lallemand, and D. d'Humières, *Consistent initial conditions for lattice boltzmann simulations*, Computers & Fluids, 35 (2006), pp. 855–862.
[44] G. N. Milstein, *Numerical integration of stochastic differential equations*, vol. 313 of Mathematics and its Applications, Kluwer Academic Publishers Group, Dordrecht, 1995. Translated and revised from the 1988 Russian original.
[45] G. N. Milstein, E. Platen, and H. Schurz, *Balanced implicit methods for stiff stochastic systems*, SIAM J. Numer. Anal., 35 (1998), pp. 1010–1019 (electronic).
[46] G. N. Milstein, Y. M. Repin, and M. V. Tretyakov, *Numerical methods for stochastic systems preserving symplectic structure*, SIAM J. Numer. Anal., 40 (2003), pp. 1583–1604 (electronic).
[47] G. N. Milstein and M. V. Tretyakov, *Stochastic numerics for mathematical physics.*, Scientific Computation. Berlin: Springer. ixx, 594 p., 2004.
[48] N. J. Newton, *Variance reduction for simulated diffusions*, SIAM J. Appl. Math., 54 (1994), pp. 1780–1805.
[49] H. C. Öttinger, *Stochastic processes in polymeric fluids*, Springer-Verlag, Berlin, 1996. Tools and examples for developing simulation algorithms.
[50] R. Rico-Martínez, C. W. Gear, and I. G. Kevrekidis, *Coarse projective kmc integration: forward/reverse initial and boundary value problems*, Journal of Computational Physics, 196 (2004), pp. 474–489.
[51] H. Risken, *The Fokker-Planck equation: Methods of solution and applications*, Springer Verlag, 1996.
[52] A. Rössler, *Second order Runge–Kutta methods for Stratonovich stochastic differential equa-*

tions, BIT, 47 (2007), pp. 657–680.

[53] Y. Saito and T. Mitsui, *Stability analysis of numerical schemes for stochastic differential equations*, SIAM J. Numer. Anal., 33 (1996), pp. 2254–2267.

[54] G. Samaey, T. Lelièvre, and V. Legat, *A numerical closure approach for kinetic models of polymeric fluids: exploring closure relations for fene dumbbells*, Computers & Fluids, (2010. (In press.)).

[55] R. Sizaire, G. Lielens, I. Jaumain, and R. Keunings, *On the hysteretic behaviour of dilute polymer solutions in relaxation following extensional flow*, Journal of Non-Newtonian Fluid Mechanics, 82 (1999), pp. 233–253.

[56] M. Somasi and B. Khomami, *Linear stability and dynamics of viscoelastic flows using time-dependent stochastic simulation techniques*, Journal of Non-Newtonian Fluid Mechanics, 93 (2000), pp. 339–362.

[57] B. P. Sommeijer, *Increasing the real stability boundary of explicit methods*, Comput. Math. Appl., 19 (1990), pp. 37–49.

[58] T. Tian and K. Burrage, *Implicit Taylor methods for stiff stochastic differential equations*, Appl. Numer. Math., 38 (2001), pp. 167–185.

[59] ———, *Two-stage stochastic Runge–Kutta methods for stochastic differential equations*, BIT, 42 (2002), pp. 625–643.

[60] Á. Tocino, *Mean-square stability of second-order Runge–Kutta methods for stochastic differential equations*, J. Comput. Appl. Math., 175 (2005), pp. 355–367.

[61] P. Van Leemput, W. Vanroose, and D. Roose, *Mesoscale analysis of the equation-free constrained runs initialization scheme*, Multiscale Modeling and Simulation, 6 (2007), pp. 1234–1255.

[62] C. Vandekerckhove, I. G. Kevrekidis, and D. Roose, *An efficient Newton–Krylov implementation of the constrained runs scheme for initializing on a slow manifold*, Journal on Scientific Computing, 39 (2009), pp. 167–188.

[63] C. Vandekerckhove and D. Roose, *Accuracy analysis of acceleration schemes for stiff multiscale problems*, J. Comput. Appl. Math., 211 (2008), pp. 181–200.

[64] C. Vandekerckhove, D. Roose, and K. Lust, *Numerical stability analysis of an acceleration scheme for step size constrained time integrators*, J. Comput. Appl. Math., 200 (2007), pp. 761–777.

[65] H. Wang, K. Li, and P. Zhang, *Crucial properties of the moment closure model FENE-QE*, Journal of Non-Newtonian Fluid Mechanics, 150 (2008), pp. 80–92.

[66] P. Yu, Q. Du, and C. Liu, *From micro to macro dynamics via a new closure approximation to the FENE model of polymeric fluids*, Multiscale Modeling and Simulation, 3 (2005), pp. 895–917.

[67] A. Zagaris, C. W. Gear, T. J. Kaper, and I. G. Kevrekidis, *Analysis of the accuracy and convergence of equation-free projection to a slow manifold*, ESAIM: Mathematical Modelling and Numerical Analysis, 43 (2009), pp. 757–784.