ACADEMIC PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Computational Physics 185 (2003) 194–212

JOURNAL OF COMPUTATIONAL PHYSICS

www.elsevier.com/locate/jcp

Numerical integration of stochastic differential equations: weak second-order mid-point scheme for application in the composition PDF method

Renfeng Cao *, Stephen B. Pope

Sibley School of Mechanical and Aerospace Engineering, Cornell University, Ithaca, NY 14853-750, USA

Received 31 January 2002; received in revised form 16 October 2002; accepted 7 November 2002

Abstract

In this paper, a weak second-order accurate mid-point scheme for the stochastic differential equations (SDEs) arising in the composition PDF method for turbulent reactive flows is proposed and tested. The results are compared with two other schemes which are commonly used in the composition PDF method. In contrast to most higher-order schemes for SDEs, the present scheme uses a mid-point, which makes it especially suitable for the implementation of the position-advance fractional step in the composition PDF method. The scheme can also be applied to the PDF method used in conjunction with large eddy simulation (LES), since the SDEs considered in this paper include explicit time dependence of the drift and diffusion coefficients. Test calculations are reported, including a 2D unsteady case, to demonstrate the weak second-order accuracy of the scheme and to compare its performance to that of two first-order schemes. A new methodology for developing higher-order scheme for SDEs (by comparing the moments of the increments of the numerical approximation with those of the exact increments) is used to develop this scheme.

Keywords: Composition PDF method; Stochastic differential equations

1. Introduction

In application to turbulent reactive flows, PDF methods have the advantage of providing full information on the turbulent fluctuations, and hence of avoiding the closure associated with non-linear chemical reactions. The PDF calculations of Xu and Pope [27], Tang et al. [23], and Lindstedt et al. [12] clearly demonstrate the ability of PDF methods to account accurately for strong turbulent–combustion interactions such as local extinction and re-ignition. Background information on PDF methods is provided by Pope [19,21]. Examples of recent composition PDF calculations are provided by Tsai and Fox [24], Mobus et al. [15], James et al. [8], and Nooren et al. [16].

* Corresponding author. Tel.: 1-607-255-9113; fax: 1-607-255-1222.
E-mail address: renfeng@mae.cornell.edu (R. Cao).

0021-9991/02/$ - see front matter
doi:10.1016/S0021-9991(02)00054-2

The modeled PDF transport equations are usually solved numerically by particle methods. For the composition PDF method, there are two different approaches. In the "node-based" approach [18], there is a finite-difference mesh in physical space, and at each node of the mesh the PDF is represented by an ensemble of particles. In the "distributed-particle" approach [19], the particles are continuously distributed in physical space. The node-based approach has the inherent limitation that it has only first-order spatial accuracy; whereas a simple implementation of the distributed-particle approach yields second-order spatial accuracy. Hence, although it is somewhat more difficult to implement, the distributed-particle approach is preferred.

In the distributed-particle implementation of the composition PDF method, there are a large number of particles, each of which has a position $\mathbf{X}(t)$ and a composition $\boldsymbol{\phi}(t)$. The position $\mathbf{X}(t)$ of a particle evolves by the stochastic differential equation (SDE)

$$\mathrm{d}\mathbf{X}(t) = \left(\tilde{\mathbf{U}} + \frac{1}{\langle\rho\rangle}\nabla\Gamma_{\mathrm{T}}\right)\mathrm{d}t + (2\Gamma_{\mathrm{T}}/\langle\rho\rangle)^{1/2}\,\mathrm{d}\mathbf{W}, \tag{1}$$

where $\tilde{\mathbf{U}}$ is the Favre-mean velocity, $\langle\rho\rangle$ is the mean density, Γ_{T} is the turbulent diffusivity, and $\mathbf{W}(t)$ is a vector-valued Wiener process. The coefficients in Eq. (1) are evaluated at the particle location. (Pope [21] provides an introduction to SDEs in the context of PDF methods.) This SDE can be written in the more general form

$$\mathrm{d}\mathbf{X}(t) = \mathbf{D}(\mathbf{X}[t], t)\,\mathrm{d}t + B(\mathbf{X}[t], t)\,\mathrm{d}\mathbf{W}, \tag{2}$$

where $\mathbf{D}(\mathbf{X}[t], t)$ and $B(\mathbf{X}[t], t)$ are the drift and diffusion coefficients, respectively. Both of these coefficients include an explicit time dependence, which must be included for unsteady solutions that arise, for example, in the use of PDF methods in conjunction with LES [2,7].

We are not concerned here with the details of the composition equation, and so we write it simply as

$$\frac{\mathrm{d}\boldsymbol{\phi}(t)}{\mathrm{d}t} = \mathbf{A}(\boldsymbol{\phi}[t], \mathbf{X}[t], t). \tag{3}$$

The function $\mathbf{A}$ on the right-hand side represents the effects of reaction and mixing.

The coupled equations (1) and (3) for $\mathbf{X}(t)$ and $\boldsymbol{\phi}(t)$ can be conveniently integrated numerically (from time t_0 to $t_0 + \Delta t$) by the following splitting scheme:

1. Eq. (1) is integrated for a time $\frac{1}{2}\Delta t$ to obtain $\mathbf{X}^{1/2}$, an approximation to $\mathbf{X}(t_0 + \frac{1}{2}\Delta t)$.
2. Eq. (3), with the right-hand side approximated by $\mathbf{A}(\boldsymbol{\phi}[t], \mathbf{X}^{1/2}, t_0 + \frac{1}{2}\Delta t)$, is integrated for a time Δt, to obtain an approximation to $\boldsymbol{\phi}(t_0 + \Delta t)$.
3. Eq. (1) is integrated from $t_0 + \frac{1}{2}\Delta t$ to $t_0 + \Delta t$ to obtain an approximation to $\mathbf{X}(t_0 + \Delta t)$.

This scheme has two favorable properties. First, if each equation is integrated (separately) with second-order temporal accuracy, then the overall (coupled) scheme is second-order accurate. Second, at each stage, the coefficients in the equations are evaluated at the particle's current location. (On the second step, typically a splitting scheme is used for reaction and mixing, possibly using sub-stepping.)

The objective of this work is to develop a second-order accurate scheme for integrating Eq. (1) that can be used in the splitting scheme described above, i.e., a mid-point scheme.

The construction of higher-order numerical schemes for SDEs is considerably more involved than is the case for ODEs. In recent years there has been much interest in designing numerical methods for stochastic differential equations (SDEs). A comprehensive account of the theory and a compendium of schemes is provided by Kloeden and Platen [10]. Although there are many methods for solving different types of SDEs with different properties, e.g. [1,3–6,9,10,14,17,20,25], there is no higher-order mid-point scheme available for the general class of SDE considered here (Eq. (2)). (Li and Modest [11] try to modify and apply the scheme introduced by Welton and Pope in [25] to the general case, but it has been shown [13] that the modified scheme is inconsistent with the SDE in general.)

The second-order scheme developed here is compared to two simpler first-order schemes. The first is the explicit Euler scheme, which is generally used in PDF methods. The second we refer to as the modified Euler scheme. In the degenerate case of the diffusion coefficient being zero, this method becomes the standard mid-point scheme for ODEs, which is second-order accurate. But in general (with non-zero diffusion coefficient) the modified Euler scheme is first-order accurate.

Depending on the level of statistical errors arising from the finite number of particles, it may or may not be beneficial to use a higher-order scheme, especially if it is computationally expensive. Because of this fact, it is rare to find higher-order schemes being put to practical use for the solution of stochastic differential equations in Monte Carlo methods: second-order accuracy is usually considered a good compromise between efficiency and accuracy [9].

Although motivated by the composition PDF method, the second-order accurate scheme developed here is applicable to any SDE of the form of Eq. (2), and hence it may be useful also in other applications. (Eq. (2) is not the most general SDE possible, since the diffusion coefficient is taken to be a scalar.)

We will begin in Section 2 by describing the three schemes mentioned above for the numerical integration of Eq. (2). These schemes are: the Euler scheme, the modified Euler scheme, and the weak second-order mid-point scheme. All of these schemes are tested and compared (in Section 4) by reference to a 2D unsteady turbulent test flow which is described in Section 3. Finally, conclusions draw from these results in Section 5. A new methodology for developing higher-order scheme for SDEs (by comparing the moments of the increments of the numerical approximation to those of the exact increments) is used to develop this scheme. This methodology, and the analysis of developing the new scheme, are shown in Appendix A.

2. Description of different schemes

2.1. Strong and weak convergence

Following [10], we introduce the definitions of strong and weak convergence and accuracy, and following [25] we argue that weak convergence and accuracy is appropriate in the context of PDF methods.

We consider the solution $\mathbf{X}(T)$ to the general SDE of the form of Eq. (2) after time T from the deterministic initial condition $\mathbf{X}(0) = \mathbf{X}^0$. This can be written as the Ito integral of Eq. (2):

$$\mathbf{X}(T) = \mathbf{X}^0 + \int_0^T \mathbf{D}(\mathbf{X}[t], t)\,\mathrm{d}t + \int_0^T B(\mathbf{X}[t], t)\,\mathrm{d}\mathbf{W}(t). \tag{4}$$

We also consider a numerical approximation to $\mathbf{X}(T)$, denoted by $\mathbf{Y}(T)$, which is obtained by approximately integrating Eq. (2) in a sequence of time steps of size Δt.

First we consider a particular given sample path of the Wiener process. Given $\mathbf{W}(t)$, there is no randomness in Eq. (4), and hence none in $\mathbf{X}(T)$. The numerical approximation $\mathbf{Y}(T)$ can be obtained by using increments of the given Wiener process; and, for the scheme to be consistent, $\mathbf{Y}(T)$ must converge to $\mathbf{X}(T)$ as Δt tends to zero. The numerical scheme is deemed to be *strong pth-order accurate* if the error $\varepsilon \equiv |\mathbf{X}(T) - \mathbf{Y}(T)|$ is of order Δt^p, i.e.,

$$|\mathbf{X}(T) - \mathbf{Y}(T)| \leqslant C\Delta t^p, \tag{5}$$

for some constant C which may depend on T and $\mathbf{W}(t)$ but is independent of Δt. This definition of strong pth-order accuracy is the same as that used for ODEs. And as with ODEs, it follows that the error made on each step must be of order Δt^{p+1}.

In the second way of regarding Eq. (4), the sample path $\mathbf{W}(t)$ is not known, but is drawn from the distribution of Wiener processes. In this case $\mathbf{X}(T)$ is a random variable. The numerical approximation $\mathbf{Y}(T)$ is obtained using samples of Wiener-process increments, and therefore is also a random variable. In this case, it is meaningful only to consider the convergence of $\mathbf{X}(T)$ and $\mathbf{Y}(T)$ *in distribution*. Such convergence is most conveniently analyzed in terms of means $\langle g(\mathbf{X}[T])\rangle$ of test functions $g(\mathbf{x})$.

For the class of test functions considered, $g(\mathbf{x})$ is bounded, infinitely differentiable, and its behavior as $|\mathbf{x}|$ tends to infinity is such that the existence of means is not in question. Based on these ideas, the numerical scheme is deemed to be *weak pth-order accurate* if the error $\varepsilon_g \equiv |\langle g(\mathbf{X}[T])\rangle - \langle g(\mathbf{Y}[T])\rangle|$ is of order Δt^p, i.e.,

$$|\langle g(\mathbf{X}[T])\rangle - \langle g(\mathbf{Y}[T])\rangle| \leqslant C_g \Delta t^p, \tag{6}$$

where C_g may depend on g and T but is independent of Δt.

In PDF methods, an ensemble of statistically identical particles is tracked for the purpose of estimating mean quantities. Hence it is weak accuracy that is relevant in this context.

2.2. The Euler scheme

Eq. (4) is the Ito integral of Eq. (2). The simplest numerical method to integrate this equation approximately is the explicit Euler scheme in which the solution is advanced from time t_0 to $t_0 + \Delta t$ by

$$\mathbf{X}(t_0 + \Delta t) \approx \mathbf{Y}(t_0 + \Delta t) \equiv \mathbf{X}(t_0) + \mathbf{D}(\mathbf{X}[t_0], t_0)\Delta t + B(\mathbf{X}[t_0], t_0)\Delta t^{1/2}\boldsymbol{\zeta}, \tag{7}$$

where $\mathbf{Y}(t_0 + \Delta t)$ is the numerical approximation to the exact value $\mathbf{X}(t_0 + \Delta t)$, and $\boldsymbol{\zeta}$ is a standardized Gaussian random vector ($\langle\boldsymbol{\zeta}\rangle = 0, \langle\zeta_i\zeta_j\rangle = \delta_{ij}$). This method is only first-order accurate.

2.3. The modified Euler scheme

If the diffusion coefficient is zero ($B = 0$) in Eq. (2), then the stochastic differential equation degenerates to the ordinary differential equation (ODE)

$$\frac{\mathrm{d}\mathbf{X}}{\mathrm{d}t} = \mathbf{D}(\mathbf{X}[t], t). \tag{8}$$

This can be integrated numerically by the second-order accurate mid-point scheme:

$$\mathbf{X}^{1/2} \equiv \mathbf{X}(t_0) + \tfrac{1}{2}\Delta t\mathbf{D}(\mathbf{X}[t_0], t_0), \tag{9}$$

$$\mathbf{X}(t_0 + \Delta t) \approx \mathbf{Y}(t_0 + \Delta t) \equiv \mathbf{X}(t_0) + \Delta t\mathbf{D}\big(\mathbf{X}^{1/2}, t_0 + \tfrac{1}{2}\Delta t\big). \tag{10}$$

In the general case with a strictly positive diffusion coefficient, the modified Euler scheme is defined to be the above mid-point scheme, with a diffusion added to the final step. Thus, the scheme is

$$\mathbf{X}^{1/2} \equiv \mathbf{X}(t_0) + \tfrac{1}{2}\Delta t\mathbf{D}(\mathbf{X}[t_0], t_0), \tag{11}$$

$$\mathbf{X}(t_0 + \Delta t) \approx \mathbf{Y}(t_0 + \Delta t) \equiv \mathbf{X}(t_0) + \mathbf{D}\big(\mathbf{X}^{1/2}, t_0 + \tfrac{1}{2}\Delta t\big)\Delta t + B\big(\mathbf{X}^{1/2}, t_0 + \tfrac{1}{2}\Delta t\big)\Delta t^{1/2}\boldsymbol{\zeta}. \tag{12}$$

For the general case ($B > 0$), this scheme is weak first-order accurate.

2.4. The weak second-order mid-point scheme

In Appendix A, by comparing the moments of the increments of the numerical approximation with those of the exact increments, we develop a weak second-order accurate mid-point scheme for the numerical integration of Eq. (2).

This scheme is defined for the general step from time t_0 to $t_0 + \Delta t$ from the initial condition $\mathbf{X}(t_0) = \mathbf{X}^0$. The result, $\mathbf{Y}(t_0 + \Delta t) = \mathbf{X}^0 + \Delta\mathbf{Y}$, is a weak second-order approximation to $\mathbf{X}(t_0 + \Delta t) = \mathbf{X}^0 + \Delta\mathbf{X}$. This claim is substantiated in Appendix A where it is shown that this scheme satisfies the sufficient conditions for a scheme to be weak second-order accuracy.

The scheme involves one mid-point (i.e., an approximation to $\mathbf{X}(t_0 + \frac{1}{2}\Delta t)$) which is denoted by $\mathbf{M}$ and three independent standardized Gaussian random vectors, denoted by $\boldsymbol{\zeta}$, $\boldsymbol{\xi}$, and $\boldsymbol{\eta}$, which have the following properties:

$$\langle\boldsymbol{\zeta}\rangle = \langle\boldsymbol{\xi}\rangle = \langle\boldsymbol{\eta}\rangle = 0, \tag{13}$$

$$\langle\zeta_i\zeta_j\rangle = \langle\xi_i\xi_j\rangle = \langle\eta_i\eta_j\rangle = \delta_{ij}. \tag{14}$$

A derived zero-mean random tensor N_{ij} is also used in this scheme, which is defined by

$$N_{ij}(\boldsymbol{\eta}) \equiv \eta_i\eta_j - \delta_{ij}. \tag{15}$$

The weak second-order mid-point scheme is defined by

$$\mathbf{M} \equiv \mathbf{X}^0 + \tfrac{1}{2}\Delta t\mathbf{D}^0 + \left(\tfrac{1}{2}\Delta t b_0\right)^{1/2}\boldsymbol{\zeta}, \tag{16}$$

$$\mathbf{X}(t_0 + \Delta t) \approx \mathbf{Y}(t_0 + \Delta t) \equiv \mathbf{X}^0 + \Delta\mathbf{Y}, \tag{17}$$

$$\Delta Y_i \equiv \Delta t D_i^{\mathbf{M}} + \left[\tfrac{1}{2}\Delta t b^{\mathbf{M}}\right]^{1/2}(\xi_i + \eta_i) + \tfrac{1}{2}\Delta t b_{,j}^{\mathbf{M}} N_{ij}(\boldsymbol{\eta}) - \left(\tfrac{1}{2}\Delta t\right)^{3/2} g_{ij}^{\mathbf{M}}(\xi_j + \eta_j), \tag{18}$$

where the superscripts and subscripts "0" and "$\mathbf{M}$" denote quantities evaluated at the initial point $(\mathbf{X}^0, t_0)$, and at the mid-point $(\mathbf{M}, t_0 + \frac{1}{2}\Delta t)$, respectively. The quantity b is defined as the square of the diffusion coefficient

$$b \equiv B^2, \tag{19}$$

and its spatial derivative are denoted by

$$b_{,i} \equiv \frac{\partial b}{\partial x_i}. \tag{20}$$

The final term in Eq. (18) introduces the symmetric second-order tensor g_{ij} which is defined by

$$g_{ij} \equiv B(B_{,i}B_{,j} + B_{,k}B_{,k}\delta_{ij}) - \tfrac{1}{2}B(D_{i,j} + D_{j,i}) = \tfrac{1}{4}b^{-1/2}(b_{,i}b_{,j} + b_{,k}b_{,k}\delta_{ij}) - \tfrac{1}{2}b^{1/2}(D_{i,j} + D_{j,i}), \tag{21}$$

with

$$D_{i,j} \equiv \frac{\partial D_i}{\partial x_j}. \tag{22}$$

The accuracy and efficiency of these schemes are examined in Section 4.

3. Numerical test case

Computations were performed for a sequence of 1D and 2D test cases of increasing complexity in order to measure the numerical error (as a function of Δt) incurred by the three schemes considered.

Test cases with different ratios of the drift term and the diffusion term were investigated, including two extreme cases: the pure drift problem (in which the diffusion term is set to zero) and the pure diffusion problem (in which the drift term is set to zero). These tests confirm that the present scheme is strong second-order accurate for the pure drift problem and weak second-order accurate for the pure diffusion problem.

It suffices to describe here the final case in the sequence, which is an unsteady 2D flow with circular streamlines. The test case is described in this section, and the results of the tests are shown in the following section.

While the particle method and SDE integrations are performed in Cartesian coordinates (x_1, x_2), the flow is naturally described in polar coordinates (r, θ), with $x_1 = r\cos\theta$ and $x_2 = r\sin\theta$. The coefficients in the SDE Eq. (1) are $\langle\rho\rangle$, $\tilde{\mathbf{U}}$, and Γ_{T}. The density is taken to be unity. The velocity field corresponds to an unsteady flow with circular streamlines,

$$\tilde{U}_1(x_1, x_2, t) = -u_\theta(r, t)\sin\theta, \tag{23}$$

$$\tilde{U}_2(x_1, x_2, t) = u_\theta(r, t)\cos\theta, \tag{24}$$

where (following Yeung and Pope [28]) the circumferential velocity is specified as

$$u_\theta(r, t) = \begin{cases} r[1 - 3(\frac{r}{\pi})^2 + 3(\frac{r}{\pi})^4 - (\frac{r}{\pi})^6](1 + \cos 4t)/2 & \text{for } r \leqslant \pi, \\ 0 & \text{otherwise.} \end{cases} \tag{25}$$

The turbulent diffusivity is specified to be

$$\Gamma_{\mathrm{T}}(x_1, x_2, t) = \left(1 + \tfrac{1}{2}\sin r\right)(1 + \cos 2t)/2. \tag{26}$$

Each particle in the computation has a Cartesian position $\mathbf{X}(t) = \{X_1(t), X_2(t)\}$, and a corresponding radial position

$$R(t) \equiv |\mathbf{X}(t)| = [X_i(t)X_i(t)]^{1/2}. \tag{27}$$

In addition, each particle has a scalar composition $\phi(t)$ which is conserved, i.e., $\mathrm{d}\phi/\mathrm{d}t = 0$. The initial composition $\phi(0)$ is set deterministically based on the particle's position as

$$\phi(0) = \left(\tfrac{1}{2}\pi^2 - 2\right)\exp\left[-\tfrac{1}{4}R(0)^2\right]. \tag{28}$$

The mean composition is defined by

$$\Phi(r, t) \equiv \langle\phi(t)|R(t) = r\rangle, \tag{29}$$

i.e., as the mean of $\phi(t)$ conditioned on the radial location of the particle.

It is readily deduced (e.g., from the composition PDF equation) that, for the flow considered, the mean composition evolves by the standard diffusion equation

$$\frac{\partial\Phi}{\partial t} = \frac{1}{r}\frac{\partial}{\partial r}\left[\Gamma_{\mathrm{T}}\frac{\partial\Phi}{\partial r}\right], \tag{30}$$

with the initial condition corresponding to Eq. (28) and the boundary conditions $(\partial\Phi/\partial r)_{r=0} = 0$ and $\Phi(\infty, t) = 0$. An extremely accurate numerical solution to this equation was obtained using

Crank–Nicolson [22], followed by Richardson extrapolation. This is taken as the "exact" solution against which the particle-method solutions are compared. (Eq. (28)) was chosen so that Eq. (30) has an analytic solution for the case $\Gamma_T = \text{constant}$, which is an earlier test case.)

Fig. 1 shows the evolution of the mean composition $\Phi(r,t)$ computed from Eq. (30). We take $T = 1$ to be the reference time at which to measure the numerical errors (and thus to be the duration of the particle-method computations). It may be seen from Fig. 1, that at $t = 1$, the mean composition is negligible beyond $r = 5$, say.

The particle method computations are performed in a square box of side 6π with periodic boundary conditions. Given that the velocity is zero beyond $r = \pi$, and that Φ is very small beyond $r = 5$, this box (of half-side $3\pi \approx 9.4$) is sufficiently large so that the solution $\Phi(r,t)$ (for $r \leqslant 5$, $t \leqslant 1$) corresponds to that in an infinite domain.

Each simulations is performed with 40,960 particles. To specify the initial conditions, the solution domain is divided into square cells by a uniform 64×64 grid, and then 10 particles are placed randomly in each cell with a uniform distribution. In order to reduce the statistical error in estimates of means obtained from the particles, for a given case (i.e., given scheme and value of Δt), 4000 statistically identical but independent simulations are performed, to yield effectively $4000 \times 40,960 \approx 1.6 \times 10^8$ particles per case.

In order to extract information on the mean composition, the radial coordinate is divided into cells of length $\Delta r = 3\pi/32$, so that the kth cell ($k \geqslant 1$) extends from $r = (k-1)\Delta r$ to $r = k\Delta r$. In the x_1–x_2 plane, the kth cell corresponds to an annulus of width Δr, and outer radius $k\Delta r$. For a given case at time t, let $N_k(t)$ be the number of particles in cell k, and let $S_k(t)$ be the sum of the particle values of ϕ in the cell. Then

$$\hat{\Phi}_k(t) \equiv S_k(t)/N_k(t), \tag{31}$$

i.e., the arithmetic mean of ϕ of particles in the cell, is an approximations to the cell mean

$$\Phi_k(t) \equiv \frac{\int_{(k-1)\Delta r}^{k\Delta r} r\Phi(r,t)\,\mathrm{d}r}{\int_{(k-1)\Delta r}^{k\Delta r} r\,\mathrm{d}r}. \tag{32}$$

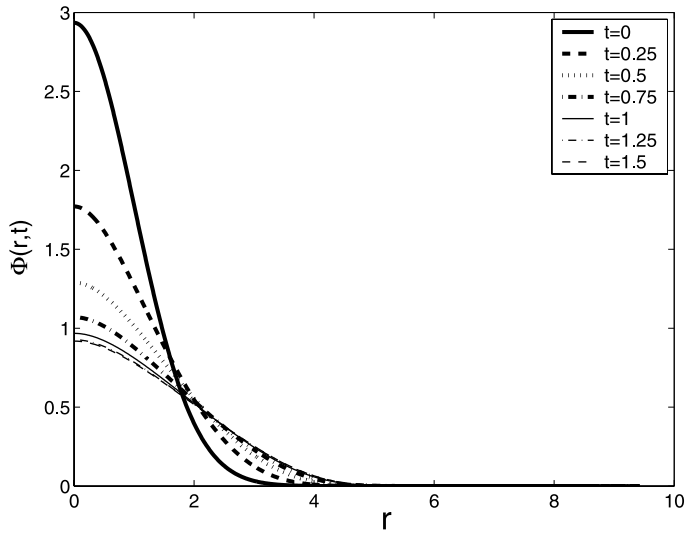

Fig. 1. Evolution of the mean composition $\Phi(r,t)$ obtained from the solution to Eq. (30).

The numerical error at the reference time $T(t = 1)$ is then defined as

$$\varepsilon_k \equiv |\hat{\boldsymbol{\Phi}}_k(T) - \boldsymbol{\Phi}_k(T)|. \tag{33}$$

In general, in a PDF particle method, the numerical error (e.g., ε_k) arises from: bias; statistical error; spatial truncation error; and temporal truncation error [26]. In the present case the bias is zero (since the particles are independent), and there is no spatial truncation error (since the coefficients, e.g., $\tilde{\mathbf{U}}$, are evaluated exactly at particle locations, not interpolated from a grid). The large effective number of particles ($\approx 1.6 \times 10^8$) is chosen so that the statistical error is small, but it is not insignificant compared to the temporal truncation error (for small Δt with the second-order scheme). It is important, therefore, to construct confidence intervals.

The confidence intervals shown in the figures below are 95% confidence intervals based on the assumption that $\hat{\boldsymbol{\Phi}}_k(T)$ (Eq. (31)) is normally distributed. Given the large number of particles involved ($\approx 1.6 \times 10^8$), ensemble means can be expected to be normally distributed and this was confirmed directly for a few cases. Thus the upper and lower confidence intervals for $\hat{\boldsymbol{\Phi}}_k(T)$ are given by

$$\hat{\boldsymbol{\Phi}}_k^{(\mathrm{u})}(t) = \hat{\boldsymbol{\Phi}}_k(t) + K_{\alpha/2}\hat{\boldsymbol{\sigma}}_k/\sqrt{N_k(t)}, \tag{34}$$

$$\hat{\boldsymbol{\Phi}}_k^{(\mathrm{l})}(t) = \hat{\boldsymbol{\Phi}}_k(t) - K_{\alpha/2}\hat{\boldsymbol{\sigma}}_k/\sqrt{N_k(t)}. \tag{35}$$

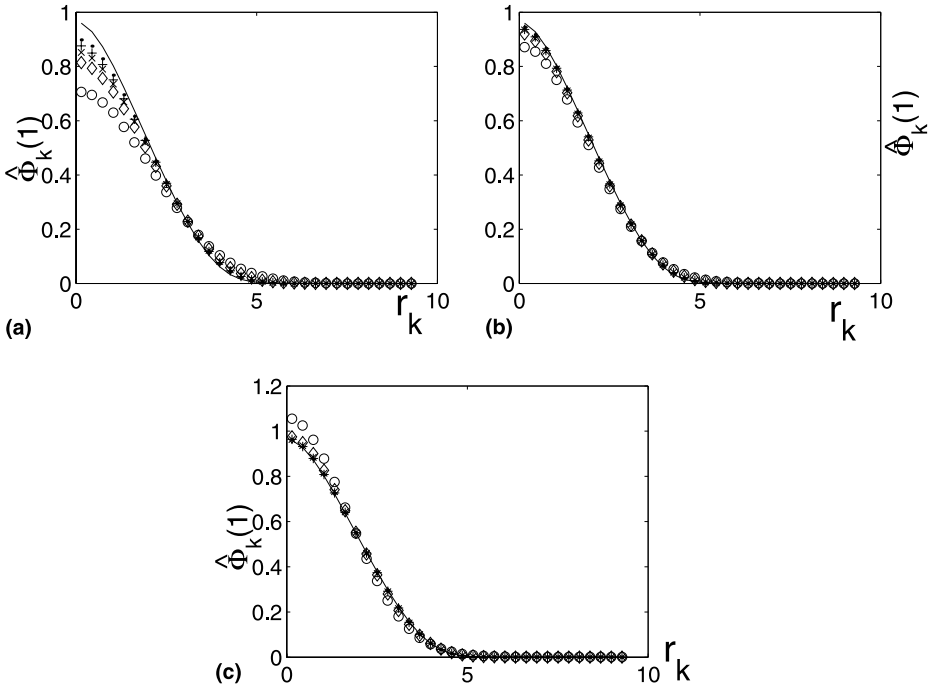

Fig. 2. The convergence of different schemes: (a) Euler scheme; (b) modified Euler scheme; (c) second-order scheme. Computed cell mean composition $\hat{\boldsymbol{\Phi}}_k(1)$ against cell radius r_k at $t = 1$ for different values of time step: ○, $\Delta t = 1$; ◇, $\Delta t = 1/2$; ×, $\Delta t = 1/3$; +, $\Delta t = 1/4$; ·, $\Delta t = 1/6$; and solid lines: exact value $\boldsymbol{\Phi}_k(1)$.

Here $N_k(t)$ is the number of particles in cell k; $K_{\alpha/2}$ is a constant corresponding to the confidence interval $1-\alpha$ (e.g., $K_{\alpha/2}=1.96$ for $\alpha=0.05$, corresponding to a 95% confidence interval); and $\hat{\sigma}_k$ is the estimation of the SD of $\hat{\Phi}_k(t)$ which is obtained by

$$\hat{\sigma}_k=\frac{\sum_{i=1}^{N_k(t))}(\phi_i-\hat{\Phi}_k(t))^2}{N_k(t)}. \tag{36}$$

4. Results and discussion

Not all of the calculation results can be shown here because of space limitations. The conclusions, however, are supported by the results given. All of the figures shown in this section are based on the three schemes described in Section 2 and the test case described in Section 3.

4.1. Accuracy

Fig. 2 shows the cell mean compositions $\hat{\Phi}_k(1)$ (computed) and $\Phi_k(1)$ (exact), plotted against the cell-center radius $r_k\equiv(k-\frac{1}{2})\Delta r$, for different schemes and for different time steps. Figs. 2(a)–(c) are the Euler scheme, the modified Euler scheme, and the second-order scheme, respectively. The overall converging tendency may be seen in this plot, and the following observations can be made. First, the second-order scheme and the modified Euler scheme clearly have small errors than the Euler scheme. Second, the error changes sign at different places for different schemes. Third, the values of Φ_k are very small beyond $r=5$.

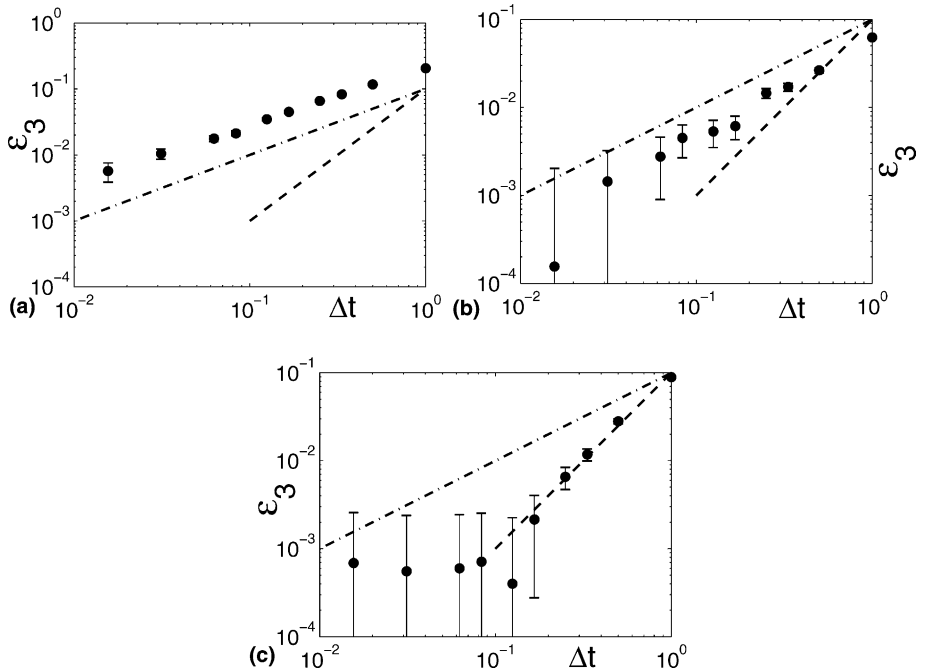

Fig. 3. The absolute error $\varepsilon_k=|\hat{\Phi}_k(1)-\Phi_k(1)|$ for different values of Δt at the third cell. (a) Euler scheme; (b) modified Euler scheme; (c) second-order scheme. Dash dotted line: line with slope 1; dashed line: line with slope 2; ● ε_3; and error bar: 95% confidence interval.

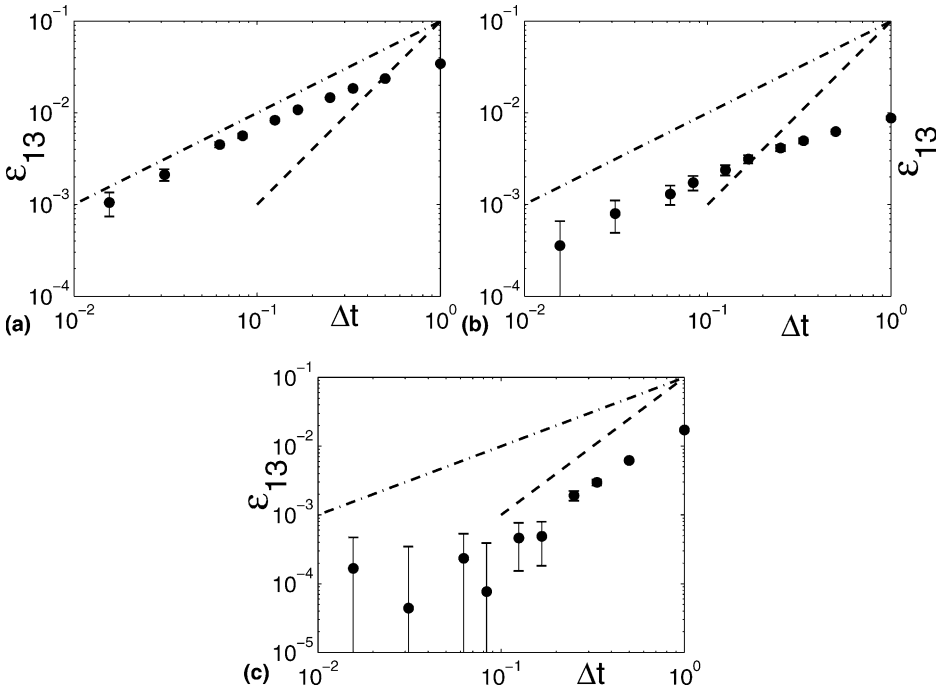

Fig. 4. The absolute error $\varepsilon_k = |\hat{\Phi}_k(1) - \Phi_k(1)|$ for different values of Δt at the 13th cell. (a) Euler scheme; (b) modified Euler scheme; (c) second-order scheme. Dash dotted line: line with slope 1; dashed line: line with slope 2; ● ε_{13}; and error bar: 95% confidence interval.

To determine the consistency and the order of accuracy, we plot the errors ε_k (Eq. (33)) against Δt on a log–log plot so that the slope of these test points represents the order of accuracy of the scheme. For these plots we select the representative cells $k = 3$ and 13, which are removed from the locations where the error changes sign, and at which the value of $\Phi_k(1)$ is at least two orders higher than the statistical error. (The exact values are $\Phi_3(1) = 0.8729$ and $\Phi_{13}(1) = 0.1032$, respectively. The 95% confidence interval is of order 10^{-3}).

The errors $\varepsilon_3(1)$ and $\varepsilon_{13}(1)$ are shown in Figs. 3 and 4 together with their 95% confidence intervals. Both figures support the same conclusions. It may be seen that, for both the Euler and modified Euler schemes, the error $\varepsilon_k(1)$ increases linearly with Δt (i.e., with slope 1 in the figures), showing that these schemes are first-order accurate. However, compared to the Euler scheme, the magnitude of the errors in the modified Euler scheme is significantly smaller. (Note that different scales are used in the plots for these two schemes.) In contrast, it may be seen that $\varepsilon_k(1)$ increases as Δt^2 (i.e., slope 2 in the figures) for the weak second-order scheme, confirming that it is indeed second-order accurate. (The plateau observed for small Δt reflects the statistical precision of the test; the confidence interval extends to zero.)

4.2. Efficiency

We now compare the efficiency and other characteristics of different schemes.

The most efficient algorithm is that which produces the numerical solution to a specified accuracy at a minimum computational cost. In the composition PDF method, the overall computational cost C for a scheme to obtain the result within a specified error tolerance ε can be defined as

$$C \equiv C_{\mathbf{X}} + C_{\phi} = N(\varepsilon)(c_{\mathbf{X}} + c_{\phi}), \tag{37}$$

Table 1
Number of time steps $N_k(\varepsilon)$ required to yield a relative error $\varepsilon_k/\Phi_k(1)$ less than the relative-error tolerance ε (for the standard test case for cell k)

$N_k(\varepsilon)$	Euler scheme	Modified Euler scheme	Second-order scheme
$N_3(0.1)$	3	1	1
$N_3(0.01)$	40	5	3
$N_3(0.001)$	476	45	7
$N_{13}(0.1)$	6	1	1
$N_{13}(0.01)$	64	22	5
$N_{13}(0.001)$	520	180	10

where the $C_{\mathbf{X}}$ and C_{ϕ} are the overall costs for the position-advance and composition calculations, respectively; $c_{\mathbf{X}}$ and c_{ϕ} are the corresponding costs per step; and $N(\varepsilon)$ is the number of time steps required for the scheme to satisfy the specified error tolerance ε.

It may be seen, then, that the relative efficiency depends on three parameters: $N(\varepsilon)$, $c_{\mathbf{X}}$, and c_{ϕ}. While $N(\varepsilon)$ and $c_{\mathbf{X}}$ depend on the position-advance scheme, c_{ϕ} does not.

We can consider the two extreme cases in which the ratio $c_{\phi}/c_{\mathbf{X}}$ is very large and very small. The first case corresponds, for example, to the composition being composed of many reactive species, so that the computational effort is dominated by the composition equation, i.e., $C \approx N(\varepsilon)c_{\phi}$. Thus, for this case, the cost is simply proportional to the number of steps required, $N(\varepsilon)$. Table 1 shows the number of steps required by the different schemes to achieve relative errors of 10%, 1%, and 0.1% at the two reference cells ($k = 3$ and 13). It may be seen that the modified Euler scheme requires fewer steps than the Euler scheme by a factor between 3 and 10. Since both of these schemes are first-order accurate, this conclusion holds for all error levels. The second-order scheme becomes more advantageous as the error tolerance is decreased. Compared to the modified Euler scheme, the second-order scheme requires fewer steps by a factor of 2–4 for a 1% error tolerance, and by a factor of 6–18 for a 0.1% tolerance.

For the second case, in which the composition-advance work C_{ϕ} is negligible, the overall cost is $C \approx N(\varepsilon)c_{\mathbf{X}}$. To compare the cost of different schemes, we need therefore to assess their relative costs per step ($c_{\mathbf{X}}$). Table 2 compares the attributes of the three schemes (and a scheme of Kloeden and Platen [10] which is discussed below). Shown in the table are: the number of random numbers required per step; and the number of coefficient evaluations required per step, which are sub-divided into local (e.g., b_0), first spatial derivative (e.g., $b^{\mathbf{M}}_{,i}$), second spatial derivatives (e.g., $b^{\mathbf{M}}_{,ij}$) and temporal derivatives (e.g., $\dot{b}_0$). The cost per step $c_{\mathbf{X}}$ is estimated as being proportional to the sum of the number of random numbers and coefficient

Table 2
Attributes of different schemes

Per step	Euler scheme	Modified Euler scheme	Second-order mid-point scheme	Kloeden scheme
Random numbers	3	3	9	6
Local coefficients	4	7	8	4
First spatial derivatives	0	0	12	12
Second spatial derivatives	0	0	0	9
First temporal derivatives	0	0	0	4
Sum	7	10	29	35
Relative cost, $c_{\mathbf{X}}$	1	1.4	4.1	5

The computational cost per step $c_{\mathbf{X}}$ is estimated as being proportional to the sum of the number of random number and coefficient evaluations required per step.

Table 3
Relative computational work $C_{\mathbf{X}} = N(\varepsilon)c_{\mathbf{X}}$ required to yield a relative error $\varepsilon_k/\Phi_k(1)$ less than the relative-error tolerance ε (for the standard test case for cell k)

Cell, k	Relative error tolerance, ε	Euler scheme	Modified Euler scheme	Second-order scheme
3	0.1	3	1.4	4.1
3	0.01	40	7	12.3
3	0.001	476	63	28.7
13	0.1	6	1.4	4.1
13	0.01	64	30.8	20.5
13	0.001	520	252	41

evaluations per step. For the Euler, modified Euler and second-order schemes, the values of $c_{\mathbf{X}}$ are estimated to be in the ratios 1:1.4:4.1.

Table 3 shows the estimate of $C_{\mathbf{X}}$, obtained by multiplying $N(\varepsilon)$ from Table 1 by $c_{\mathbf{X}}$ from Table 2. It may be seen that the modified Euler scheme requires less work than the Euler scheme by a factor of 2–7. The relative cost of the modified Euler scheme and the second-order scheme depends on the specified error tolerance. With a 1% error tolerance the schemes have comparable cost, with the second-order scheme becoming advantageous as the error tolerance is decreased.

In the case of complex grids in which the derivatives are relatively expensive to evaluate in terms of computational cost, the implementation cost of the second-order scheme may be significantly higher than in the case considered above, in which the computational cost of each numerical operation (derivatives, random numbers, etc.) are taken to be comparable.

5. Conclusions

We consider the stochastic differential equations in the form of Eq. (2) arising in the composition PDF method (as Eq. (1)). A weak second-order mid-point scheme has been developed for the numerical integration of these equations. The scheme consists of Eq. (18) for the increment in the process, which involves coefficients evaluated at one mid-point, and three independent standardized Gaussian random vectors. In contrast to most higher-order schemes for SDEs, this scheme uses a mid-point, which makes it especially appropriate for the implementation of the position-advance fractional step in the composition PDF method. The scheme can also be applied to the PDF method used in conjunction with large eddy simulation (LES), since the SDEs considered in this paper include explicit time dependence of the drift and diffusion coefficients.

The second-order mid-point scheme has been tested in a sequence of 1D and 2D test cases of increasing complexity in order to test its consistency and weak second-order accuracy. The final test case, i.e., an unsteady 2D turbulent flow with circular stream line, is described in Section 3 and the results are shown in Section 4. The results are compared with Euler scheme and modified Euler scheme which are commonly used in the composition PDF methods. The analysis of developing this scheme is shown in Appendix A.

The Euler scheme is the simplest scheme and can be used in simple applications, or if the work of programming is the primary consideration.

The modified Euler scheme is better than the Euler scheme for most applications based on the following reasons. First, it is a mid-point scheme and therefore suitable for the splitting scheme in the composition PDF method. Second, the overall calculation cost of the modified Euler scheme is much less than that of the Euler scheme (for given accuracy). Finally, it is also easy to be implemented since only a slight modification need to be made on the Euler scheme.

For complex applications (e.g., with many reactive species) or if a small error tolerance (e.g., $\leqslant 0.1\%$) is required, then the importance of the computational cost increases so that the second-order scheme becomes a good choice since its overall calculation cost is much less than that of the first-order scheme in this situation.

The weak second-order mid-point scheme proposed in the paper has several advantages when compared with other weak second-order scheme available now for the SDEs in the form of Eq. (2). First, it is mid-point scheme which is particularly appropriate for the PDF method. Second, it does not require the evaluation of the temporal derivatives for the coefficients (which are difficult to get in the PDF method). Third, only the first spatial derivatives are needed. A second-order scheme for the SDEs in the form of Eq. (2) is briefly stated in Appendix B. This scheme is not appropriate for the applications of the composition PDF method because of the above reasons.

In the development of this new scheme, a new methodology is used by comparing the moments of the increments of the numerical approximation with those of the exact increments. This method provides a different way of developing higher-order schemes for the stochastic differential equations. It is illustrated in Appendix A by using it to develop the second-order scheme proposed in this paper.

Acknowledgement

This work was supported by Department of Energy, Grant No. DE-FG02-90ER.

Appendix A. Development of the weak second-order scheme

In this section, we present the details of the development of the weak second-order mid-point scheme (Eqs. (16)–(18)) described in Section 2.4. The present analysis is based on the contents of Sections 1 and 2, including notation introduced there.

A.1. Sufficient conditions for weak second-order accuracy

In this Section, we consider the numerical integration of the the general SDEs in the form of Eq. (2). A sufficient condition for weak pth-order accuracy is that the error incurred on each step be of order Δt^{p+1}. To make this precise we consider the general step from time t_0 to time $t_0+\Delta t$ with the deterministic initial condition $\mathbf{X}(t_0)=\mathbf{Y}(t_0)=\mathbf{X}^0$. Then the requirement is

$$\left|\left\langle g(\mathbf{X}[t_0+\Delta t])|\mathbf{X}(t_0)=\mathbf{X}^0\right\rangle-\left\langle g(\mathbf{Y}[t_0+\Delta t])|\mathbf{Y}(t_0)=\mathbf{X}^0\right\rangle\right|\leqslant \mathbf{C}_g\Delta t^{p+1}. \tag{A.1}$$

This requirement can be re-expressed in terms of the increments in $\mathbf{X}$ and $\mathbf{Y}$ which are defined by

$$\Delta\mathbf{X}\equiv\mathbf{X}(t_0+\Delta t)-\mathbf{X}^0 \tag{A.2}$$

and

$$\Delta\mathbf{Y}\equiv\mathbf{Y}(t_0+\Delta t)-\mathbf{Y}^0=\mathbf{Y}(t_0+\Delta t)-\mathbf{X}^0. \tag{A.3}$$

A Taylor series expansion then yields

$$g(\mathbf{X}[t_0+\Delta t])=g(\mathbf{X}^0+\Delta\mathbf{X})=g^0+g^0_{,i}\Delta X_i+\tfrac{1}{2!}g^0_{,ij}\Delta X_i\Delta X_j\cdots, \tag{A.4}$$

where $g^0_{,i}$ denotes $\partial g/\partial x_i$ evaluated at $\mathbf{X}^0$, etc. Substituting Eq. (A.4) for $g(\mathbf{X}[t_0+\Delta t])$ and the analogous expansion for $g(\mathbf{Y}[t_0+\Delta t])$ into Eq. (A.1), we obtain the sufficient condition for weak pth-order accuracy

$$\left|g^0_{,i}(\langle\Delta X_i\rangle-\langle\Delta Y_i\rangle)+\tfrac{1}{2!}g^0_{,ij}(\langle\Delta X_i\Delta X_j\rangle-\langle\Delta Y_i\Delta Y_j\rangle)\cdots\right|\leqslant \mathbf{C}_g\Delta t^{p+1}. \tag{A.5}$$

It is shown below that for n being a positive even integer (including zero), the nth (i.e., even) moments of $\Delta\mathbf{X}$ are of order $\Delta t^{n/2}$, and the $(n+1)$th (i.e., odd) moments are of order $\Delta t^{1+n/2}$. In particular,

$$\langle \Delta X_i \rangle = \mathcal{O}(\Delta t), \tag{A.6}$$

$$\langle \Delta X_i \Delta X_j \rangle = \mathcal{O}(\Delta t), \tag{A.7}$$

$$\langle \Delta X_i \Delta X_j \Delta X_k \rangle = \mathcal{O}(\Delta t^2), \tag{A.8}$$

$$\langle \Delta X_i \Delta X_j \Delta X_k \Delta X_l \rangle = \mathcal{O}(\Delta t^2), \tag{A.9}$$

and

$$\langle \Delta X_i \Delta X_j \Delta X_k \Delta X_l \Delta X_m \rangle = \mathcal{O}(\Delta t^3). \tag{A.10}$$

Thus sufficient conditions for the numerical scheme to be weak second-order accurate are: that the first four moments of $\Delta\mathbf{Y}$ approximate those of $\Delta\mathbf{X}$ with an error of order Δt^3; and that the fifth and higher moments be at most of order Δt^3.

In the following two sections: the coefficients in the series expansions for the moments of $\Delta\mathbf{X}$ are determined; and it is shown that the moments of $\Delta\mathbf{Y}$ given by the weak second-order accurate scheme (Eqs. (16)–(18)) of Section 2.4 satisfy the sufficient conditions given in this section.

A.2. Moments of ΔX

To simplify the notation, and without loss of generality, we consider the step from $t = 0$ to $t = \Delta t$ (i.e., $t_0 = 0$), from the deterministic initial condition $\mathbf{X}(0) = 0$. (The step from t_0 to $t_0 + \Delta t$ with initial condition $\mathbf{X}(t_0) = \mathbf{X}^0$ can be transformed into this simpler form.) The increment $\Delta\mathbf{X}$ over the time step is simply

$$\Delta\mathbf{X} \equiv \mathbf{X}(\Delta t) - \mathbf{X}(0) = \mathbf{X}(\Delta t). \tag{A.11}$$

The moments of $\mathbf{X}(t)$ can be determined from its PDF which is denoted by $f(\mathbf{x}; t)$. More generally, means of functions of $\mathbf{X}(t)$ can be obtained from

$$\langle Q(\mathbf{X}[t]) \rangle = \int Q(\mathbf{x}) f(\mathbf{x}; t)\, \mathrm{d}\mathbf{x}, \tag{A.12}$$

where the integration is over all $\mathbf{x}$. Expanding f in a Taylor series about $t = 0$ yields

$$\langle Q(\mathbf{X}[\Delta t]) \rangle = \int Q(\mathbf{x}) \left\{ f_0 + \Delta t \dot{f}_0 + \frac{1}{2}\Delta t^2 \ddot{f}_0 \right\} \mathrm{d}\mathbf{x} + \mathcal{O}(\Delta t^3), \tag{A.13}$$

where $\dot{f}_0$ denotes $\partial f/\partial t$ evaluated at $t = 0$, etc.

Now $\mathbf{X}(t)$ evolves by the SDE Eq. (2), so $f(\mathbf{x}; t)$ evolves by the corresponding Fokker–Planck equation ([10])

$$\dot{f} \equiv \frac{\partial f}{\partial t} = -\frac{\partial}{\partial x_i}[f D_i(\mathbf{x}, t)] + \frac{\partial^2}{\partial x_i \partial x_i}\left[\frac{1}{2} f b(\mathbf{x}, t)\right], \tag{A.14}$$

from the initial condition

$$f_0(\mathbf{x}) \equiv f(\mathbf{x}; 0) = \delta(\mathbf{x}), \tag{A.15}$$

where

$$b(\mathbf{x},t) \equiv B(\mathbf{x},t)^2. \tag{A.16}$$

In view of this initial condition, the first term in Eq. (A.13) is

$$\int Q(\mathbf{x}) f_0 \,\mathrm{d}\mathbf{x} = Q_0 \equiv Q(0). \tag{A.17}$$

The coefficient of the second term in the series equation (A.13) is

$$\begin{aligned}\int Q(\mathbf{x})\dot{f}_0 \,\mathrm{d}\mathbf{x} &= \int Q(\mathbf{x})\left\{ -\frac{\partial}{\partial x_i}[fD_i] + \frac{\partial^2}{\partial x_i x_i}\left[\frac{1}{2} f b\right]\right\}_0 \mathrm{d}\mathbf{x} = \int \left\{ fD_i \frac{\partial Q}{\partial x_i} + \frac{1}{2} f b \frac{\partial^2 Q}{\partial x_i \partial x_i}\right\}_0 \mathrm{d}\mathbf{x} \\ &= D_i^0 Q_{,i}^0 + \frac{1}{2} b_0 Q_{,ii}^0,\end{aligned} \tag{A.18}$$

where a subscript or superscript "0" indicates evaluation at $\mathbf{x} = 0$.

The third term in the series involves $\ddot{f}$ which is obtained by differentiating Eq. (A.14)

$$\ddot{f} = -\frac{\partial}{\partial x_i}\left[\dot{f} D_i + f\dot{D}_i\right] + \frac{1}{2}\frac{\partial^2}{\partial x_i \partial x_i}\left[\dot{f} b + f\dot{b}\right], \tag{A.19}$$

and then substituting Eq. (A.14) for $\dot{f}$. When the result is substituted for the third term in Eq. (A.13) and integrated, the resulting coefficient is

$$\begin{aligned}\int Q(\mathbf{x})\frac{1}{2}\ddot{f}_0 \,\mathrm{d}x = {} & \frac{1}{2} D_j^0 D_{i,j}^0 Q_{,i}^0 + \frac{1}{2} D_j^0 D_i^0 Q_{,ij}^0 + \frac{1}{2}\dot{D}_i^0 Q_{,i}^0 + \frac{1}{4} b_0 (D_{i,jj}^0 Q_{,i}^0 + 2D_{i,j}^0 Q_{,ij}^0 + D_i^0 Q_{,ijj}^0) \\ & + \frac{1}{4} D_j^0 (b_{,j}^0 Q_{,ii}^0 + b_0 Q_{,iij}^0) + \frac{1}{4}\dot{b}_0 Q_{,ii}^0 + \frac{1}{8} b_0 (b_{,jj}^0 Q_{,ii}^0 + 2b_{,j}^0 Q_{,iij}^0 + b_0 Q_{,iijj}^0).\end{aligned} \tag{A.20}$$

The series expansions for moments of $\mathbf{X}(t)$ are obtained by substituting the corresponding moments of $\mathbf{x}$ for Q. For example, the series for $\langle X_p X_q\rangle$ is obtained by substituting $Q = x_p x_q$. In this case, Q and its first derivatives,

$$(x_p x_q)_{,i} = x_p \delta_{qi} + x_q \delta_{pi}, \tag{A.21}$$

make no contribution since they are zero at the origin ($\mathbf{x} = 0$). The second-derivative terms,

$$(x_p x_q)_{,ij} = \delta_{pi}\delta_{qj} + \delta_{qi}\delta_{pj}, \tag{A.22}$$

are non-zero for $\mathbf{x} = 0$ and hence lead to non-trivial contributions. Third and higher derivatives are zero. So, in general, contributions to an mth moment arise solely from terms containing an mth derivative of Q. Since the highest derivative of Q in the first three terms of the series is a fourth derivative, it follows immediately that fifth and higher moments have no contributions from these terms and hence are at most of order Δt^3.

Explicit evaluation of the series (from Eqs. (A.13), (A.17), (A.18), and (A.20)) for the first four moments of $\Delta\mathbf{X} = \mathbf{X}(\Delta t)$ yields

$$\langle \Delta X_p\rangle = \Delta t D_p^0 + \Delta t^2\left\{\tfrac{1}{2} D_j^0 D_{p,j}^0 + \tfrac{1}{4} b_0 D_{p,jj}^0 + \tfrac{1}{2}\dot{D}_p^0\right\} + \mathcal{O}(\Delta t^3), \tag{A.23}$$

$$\langle \Delta X_p \Delta X_q\rangle = \Delta t b_0 \delta_{pq} + \Delta t^2\left\{D_p^0 D_q^0 + \tfrac{1}{2} b_0 (D_{p,q}^0 + D_{q,p}^0) + \tfrac{1}{2}\delta_{pq} D_j^0 b_{,j}^0 + \tfrac{1}{4} b_0 b_{,jj}^0 \delta_{pq} + \tfrac{1}{2}\dot{b}_0 \delta_{pq}\right\} + \mathcal{O}(\Delta t^3), \tag{A.24}$$

$$\langle \Delta X_p \Delta X_q \Delta X_r \rangle = \Delta t^2 \Big\{ b_0 (D_p^0 \delta_{qr} + D_q^0 \delta_{pr} + D_r^0 \delta_{pq}) + \tfrac{1}{2} b_0 (b_{,p}^0 \delta_{qr} + b_{,q}^0 \delta_{pr} + b_{,r}^0 \delta_{pq}) \Big\} + \mathcal{O}(\Delta t^3), \tag{A.25}$$

$$\langle \Delta X_p \Delta X_q \Delta X_r \Delta X_s \rangle = \Delta t^2 b_0^2 (\delta_{pq}\delta_{rs} + \delta_{pr}\delta_{qs} + \delta_{ps}\delta_{qr}) + \mathcal{O}(\Delta t^3). \tag{A.26}$$

A.3. Moments of ΔY

The weak second-order mid-point scheme Eqs. (16)–(18) proposed in Section 2.4 is analyzed by determining the moments of $\Delta \mathbf{Y}$. It is shown that these agree with those of $\Delta \mathbf{X}$ (Eqs. (A.23)–(A.26)) to order Δt^2, thus establishing the scheme's weak second-order accuracy.

As before, to simplify the notation and without loss of generality, we take $t_0 = 0$ and $\mathbf{X}^0 = 0$. In the evaluation of the moments of $\Delta \mathbf{Y}$, we will use the first and second moments of $\mathbf{M}$ in many places. From the definition of $\mathbf{M}$ (Eq. (16)) with $\mathbf{X}^0 = 0$, these moments are determined to be:

$$\langle M_i \rangle = \tfrac{1}{2} \Delta t D_i^0 = \mathcal{O}(\Delta t), \tag{A.27}$$

$$\langle M_i M_j \rangle = \tfrac{1}{2} \Delta t b_0 \delta_{ij} + \tfrac{1}{4} \Delta t^2 D_i^0 D_j^0 = \tfrac{1}{2} \Delta t b_0 \delta_{ij} + \mathcal{O}(\Delta t^2). \tag{A.28}$$

Thus we have

$$\langle M_i \Delta t \rangle = \tfrac{1}{2} \Delta t^2 D_i^0 = \mathcal{O}(\Delta t^2). \tag{A.29}$$

A.3.1. First moments

The mean of ΔY_p (Eq. (18)) is:

$$\langle \Delta Y_p \rangle = \Delta t \Big\langle D_p^{\mathbf{M}} \Big\rangle + \Big\langle \big[\tfrac{1}{2} \Delta t b^{\mathbf{M}}\big]^{1/2} \Big\rangle \langle (\xi_p + \eta_p) \rangle + \Big\langle \tfrac{1}{2} \Delta t b_{,i}^{\mathbf{M}} \Big\rangle \langle N_{ip}(\boldsymbol{\eta}) \rangle - \Big\langle \big(\tfrac{1}{2}\Delta t\big)^{3/2} g_{pi}^{\mathbf{M}} \Big\rangle \langle (\xi_i + \eta_i) \rangle. \tag{A.30}$$

Note that the last three terms on the right-hand side can be decomposed as the product of two means because $\mathbf{M}$ is independent of $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$; and these terms then vanish, because the means of $\boldsymbol{\xi}$, $\boldsymbol{\eta}$, and N_{ip} are zero. The remaining term is

$$\langle \Delta Y_p \rangle = \Delta t \Big\langle D_p^{\mathbf{M}} \Big\rangle. \tag{A.31}$$

The coefficient $\langle D_p^{\mathbf{M}} \rangle$ can be expressed in terms of the moments of $\mathbf{M}$ by expanding $\langle D_p^{\mathbf{M}} \rangle$ in a Taylor series

$$\Big\langle D_p^{\mathbf{M}} \Big\rangle = D_p^0 + D_{p,i}^0 \langle M_i \rangle + \tfrac{1}{2} D_{p,ij}^0 \langle M_i M_j \rangle + \dot{D}_p^0 \big(\tfrac{1}{2}\Delta t\big) + \mathcal{O}(\Delta t^2). \tag{A.32}$$

Substituting Eqs. (A.27) and (A.28) into the above two equations yields

$$\langle \Delta Y_p \rangle = \Delta t D_p^0 + \Delta t^2 \Big(\tfrac{1}{2} D_j^0 D_{p,j}^0 + \tfrac{1}{4} b_0 D_{p,jj}^0 + \tfrac{1}{2} \dot{D}_p^0 \Big) + \mathcal{O}(\Delta t^3), \tag{A.33}$$

which agrees with Eq. (A.23) for $\langle \Delta X_p \rangle$ to order Δt^2.

A.3.2. Second moments

Several results needed in the evaluation of the second moments of $\Delta \mathbf{Y}$ are:

$$\langle \eta_i N_{jk} \rangle = \langle \eta_i \eta_j \eta_k \rangle - \langle \eta_i \rangle \delta_{jk} = 0, \tag{A.34}$$

$$\langle N_{ip}N_{jq}\rangle = \langle(\eta_i\eta_p - \delta_{ip})(\eta_j\eta_q - \delta_{jq})\rangle = \langle\eta_i\eta_j\eta_p\eta_q\rangle - \langle\eta_i\eta_p\rangle\delta_{jq} - \langle\eta_j\eta_q\rangle\delta_{ip} + \delta_{ip}\delta_{jq}$$
$$= \delta_{ij}\delta_{pq} + \delta_{ip}\delta_{jq} + \delta_{iq}\delta_{jp} - \delta_{ip}\delta_{jq} - \delta_{jq}\delta_{ip} + \delta_{ip}\delta_{jq} = \delta_{ij}\delta_{pq} + \delta_{iq}\delta_{jp}, \tag{A.35}$$

$$\langle b^{\mathbf{M}}\rangle = b_0 + b^0_{,j}\langle M_j\rangle + \tfrac{1}{2}b^0_{,ij}\langle M_iM_j\rangle + \dot{b_0}\left(\tfrac{1}{2}\Delta t\right) + \mathcal{O}(\Delta t^2)$$
$$= b_0 + \left(\tfrac{1}{2}b^0_{,j}D^0_j + \tfrac{1}{4}b^0_{,jj}b_0 + \tfrac{1}{2}\dot{b_0}\right)\Delta t + \mathcal{O}(\Delta t^2), \tag{A.36}$$

$$\left\langle D^{\mathbf{M}}_p D^{\mathbf{M}}_q\right\rangle = D^0_pD^0_q + \mathcal{O}(\Delta t), \tag{A.37}$$

$$\left\langle b^{\mathbf{M}}_{,i}b^{\mathbf{M}}_{,j}\right\rangle = b^0_{,i}b^0_{,j} + \mathcal{O}(\Delta t). \tag{A.38}$$

From Eq. (18) and Eqs. (A.34)–(A.38), we obtain

$$\langle\Delta Y_p\Delta Y_q\rangle = \Delta t^2\left\langle D^{\mathbf{M}}_pD^{\mathbf{M}}_q\right\rangle + \tfrac{1}{2}\Delta t\langle b^{\mathbf{M}}\rangle 2\delta_{pq} + \tfrac{1}{4}\Delta t^2\left\langle b^{\mathbf{M}}_{,i}b^{\mathbf{M}}_{,j}\right\rangle(\delta_{ij}\delta_{pq} + \delta_{iq}\delta_{jp})$$
$$- \tfrac{1}{2}\Delta t^2\left\langle b^{1/2}_{\mathbf{M}}\left[g^{\mathbf{M}}_{qp} + g^{\mathbf{M}}_{pq}\right]\right\rangle + \mathcal{O}(\Delta t^3)$$
$$= \Delta t^2D^0_pD^0_q + \Delta t\delta_{pq}\left(b_0 + b^0_{,i}\tfrac{1}{2}\Delta tD^0_i + b^0_{,ii}\tfrac{1}{4}\Delta tb_0 + \tfrac{1}{2}\dot{b_0}\Delta t\right) + \tfrac{1}{4}\Delta t^2(b^0_{,i}b^0_{,i}\delta_{pq} + b^0_{,q}b^0_{,p})$$
$$- \tfrac{1}{2}\Delta t^2\left\langle b^{1/2}_{\mathbf{M}}\left[g^{\mathbf{M}}_{qp} + g^{\mathbf{M}}_{pq}\right]\right\rangle + \mathcal{O}(\Delta t^3). \tag{A.39}$$

By construction of the scheme, the symmetric second-order tensor g_{ij} (Eq. (21)) is defined so that Eq. (A.39) agrees with Eq. (A.24) to order Δt^2.

A.3.3. Third moments

When third moments of $\Delta\mathbf{Y}$ are formed from Eq. (18), we have the following observations:

- Since $\langle(\xi_p + \eta_p)(\xi_q + \eta_q)(\xi_r + \eta_r)\rangle = 0$, the products of the square of the second term with the fourth term equals zero.
- The leading-order terms are of order Δt^2 and arise from products of the square of second term with the first and third terms.
- Other terms are at most of order Δt^3.

The correlations of $\boldsymbol{\eta}$ and $N_{ij}(\boldsymbol{\eta})$ that arise can be evaluated as, for example:

$$\langle(\xi_p + \eta_p)(\xi_q + \eta_q)N_{ri}(\boldsymbol{\eta})\rangle = \langle\eta_p\eta_qN_{ir}\rangle = \langle\eta_p\eta_q(\eta_i\eta_r - \delta_{ir})\rangle = \langle\eta_p\eta_q\eta_i\eta_r\rangle - \langle\eta_p\eta_q\rangle\delta_{ir}$$
$$= \delta_{pq}\delta_{ir} + \delta_{pi}\delta_{qr} + \delta_{pr}\delta_{qi} - \delta_{pq}\delta_{ir} = \delta_{pi}\delta_{qr} + \delta_{pr}\delta_{qi}. \tag{A.40}$$

Thus, we can obtain the third moments of $\Delta\mathbf{Y}$:

$$\langle\Delta Y_p\Delta Y_q\Delta Y_r\rangle = \tfrac{1}{2}\Delta t^2\left(\left\langle b^{\mathbf{M}}D^{\mathbf{M}}_r\right\rangle 2\delta_{pq} + \left\langle b^{\mathbf{M}}D^{\mathbf{M}}_p\right\rangle 2\delta_{qr} + \left\langle b^{\mathbf{M}}D^{\mathbf{M}}_q\right\rangle 2\delta_{pr}\right) + \tfrac{1}{4}\Delta t^2\left\langle b^{\mathbf{M}}b^{\mathbf{M}}_{,j}\right\rangle(\langle\eta_p\eta_qN_{rj}\rangle$$
$$+ \langle\eta_q\eta_rN_{pj}\rangle + \langle\eta_p\eta_rN_{qj}\rangle)$$
$$= \Delta t^2\left\{b_0(D^0_p\delta_{qr} + D^0_q\delta_{pr} + D^0_r\delta_{pq}) + \tfrac{1}{2}b_0(b^0_{,p}\delta_{qr} + b^0_{,q}\delta_{pr} + b^0_{,r}\delta_{pq})\right\} + \mathcal{O}(\Delta t^3), \tag{A.41}$$

which agrees with the third moments of $\Delta\mathbf{X}$ (Eq. (A.25)) to order Δt^2.

A.3.4. Fourth moments

To order Δt^2, the fourth moment of $\Delta\mathbf{Y}$ arises entirely from the second term in Eq. (18), and is:

$$\begin{aligned}\langle \Delta Y_p \Delta Y_q \Delta Y_r \Delta Y_s \rangle &= \left\langle \tfrac{1}{2}\Delta t b^{\mathbf{M}} \right\rangle^2 \langle (\xi_p + \eta_p)(\xi_q + \eta_q)(\xi_r + \eta_r)(\xi_s + \eta_s) \rangle \\ &= \Delta t^2 b_0^2 (\delta_{pq}\delta_{rs} + \delta_{pr}\delta_{qs} + \delta_{ps}\delta_{qr}) + \mathcal{O}(\Delta t^3),\end{aligned} \tag{A.42}$$

in agreement with Eq. (A.26) to order Δt^2.

A.3.5. Higher moments

It may be observed that although $\Delta\mathbf{Y}$ (Eq. (18)) contains contributions of order $\Delta t^{1/2}$ and $\Delta t^{3/2}$, the expansions for the moments of $\Delta\mathbf{Y}$ are in integer powers of Δt. This is because the terms in $\Delta t^{1/2}$ and $\Delta t^{3/2}$ are multiplied by zero-mean Gaussian random variables, and hence only contribute to moments when multiplied by themselves or each other an even number of times.

It follows immediately that fifth and higher moments of $\Delta\mathbf{Y}$ are at most of order Δt^3. Thus we have verified that the scheme proposed in this paper satisfies the sufficient conditions to be weak second-order accurate.

Appendix B. One step weak second-order scheme [10]

In this section is a brief description of the weak second-order scheme proposed by Kloeden and Platen [10] for the SDEs which have the form of Eq. (2). The notation is the same as defined in previous sections.

Following [10], consider the general step from time t_0 to $t_0 + \Delta t$ from the initial condition $\mathbf{X}(t_0) = \mathbf{X}^0$. Recall Eq. (17) of Section 2.4,

$$\mathbf{X}(t_0 + \Delta t) \approx \mathbf{Y}(t_0 + \Delta t) \equiv \mathbf{X}^0 + \Delta\mathbf{Y}, \tag{B.1}$$

where $\mathbf{Y}(t_0 + \Delta t)$ is the numerical approximation to the exact value $\mathbf{X}(t_0 + \Delta t)$. Then the one step weak second-order scheme is defined by (from Eqs. (14.2.7), (10.1.1), and (10.1.3) of [10]):

$$\begin{aligned}\Delta Y_i = {} & D_i \Delta t + \tfrac{1}{2}\Delta t^2 \left(\dot{D}_i + D_j D_{i,j} + \tfrac{1}{2} B^2 D_{i,jj} \right) + \Delta t^{1/2} \left[B + \tfrac{1}{2}\Delta t (\dot{B} + D_j B_{,j} + \tfrac{1}{2} B^2 B_{,jj}) \right] \xi_i \\ & + \tfrac{1}{2}\Delta t^{3/2} B D_{i,j} \xi_j + \tfrac{1}{2}\Delta t B B_{,j} (\xi_i \xi_j + V_{ij}),\end{aligned} \tag{B.2}$$

where all coefficients are evaluated at $(\mathbf{X}^0, t_0)$. As previously, ξ_i are independent standardized Gaussians and the two-point distributed random variables V_{ij} are defined by (from Eqs. (14.2.8)–(14.2.10) of [10]):

$$P(V_{ij} = \pm 1) = \tfrac{1}{2} \quad \text{for } i < j, \tag{B.3}$$

$$V_{ij} = -1 \quad \text{for } i = j, \tag{B.4}$$

$$V_{ij} = -V_{ji} \quad \text{for } i > j. \tag{B.5}$$

References

[1] H.P. Breuer, U. Dorner, F. Petruccione, Numerical integration methods for stochastic wave function equations, Comp. Phys. Commun. 132 (2000) 30–43.

[2] P.J. Colucci, F.A. Jaberi, P. Givi, S.B. Pope, Filtered density function for large eddy simulation of turbulent reacting flows, Phys. Fluids 10 (1998) 499–515.
[3] H.A. Forbert, S.A. Chin, Fourth-order algorithm for solving the multi-variable Langevin equation and the Kramers equation, Phys. Rev. E 63 (2000) 016703.
[4] D.C. Haworth, S.B. Pope, A second-order monte carlo method for the solution of the ito stochastic differential equation, Stochastic Anal. Appl. 4 (1986) 151–186.
[5] E. Hershkovitz, A fourth-order numerical integrator for stochastic Langevin equations, J. Chem. Phys. 108 (1998) 9253–9258.
[6] E. Hershkovitz, R. Hernandez, Fast numerical integrator for stochastic differential equations with nonstationary multiplicative noise, J. Phys. Chem. A 105 (2001) 2687–2693.
[7] F.A. Jaberi, P.J. Colucci, S. James, P. Givi, S.B. Pope, Filtered mass density function for large-eddy simulation of turbulent reacting flows, J. Fluid Mech. 401 (1999) 85–121.
[8] S. James, M.S. Anand, M.K. Razdan, S.B. Pope, In situ detailed chemistry calculations in combustor flow analysis, J. Eng. Gas Turbines Power 123 (2001) 747–756.
[9] J. Qiang, S. Habib, Second-order stochastic Leapfrog algorithm for multiplicative noise Brownian motion, Phys. Rev. E 62 (2000) 7430–7437.
[10] P.E. Kloeden, E. Platen, Numerical Solution of Stochastic Differential Equations, Springer, Berlin, 1992.
[11] G. Li, M.F. Modest, An effective particle tracing scheme on structured/unstructured grids in hybrid finite volume/PDF Monte Carlo methods, J. Comp. Phys. 173 (2001) 187–207.
[12] R.P. Lindstedt, S.A. Louloudi, E.M. Vaos, Joint scalar probability density function modeling of pollutant formation in piloted turbulent jet diffusion flames with comprehensive chemistry, Proc. Combust. Inst. 28 (2000) 149–156.
[13] J.P. Minier, R. Cao, S.B. Pope, Comment on the article an effective particle tracing scheme on structured/unstructured grids in hybrid finite volume/PDF Monte Carlo methods by Li and Modest, J. Comp. Phys. (submitted).
[14] T. Misawa, A Lie algebraic approach to numerical integration of stochastic differential equations, Siam. J. Sci. Comput. 23 (2001) 866–890.
[15] H. Mobus, P. Gerlinger, D. Bruggemann, Comparison of Eulerian and Lagrangian Monte Carlo PDF methods for turbulent diffusion flames, Combust. Flame 124 (2001) 519–534.
[16] P.A. Nooren, H.A. Wouters, T.W.J. Peeters, D. Roekaerts, U. Maas, D. Schmidt, Monte Carlo PDF modeling of a turbulent natural-gas diffusion flame, Combust. Theory Model. 1 (1997) 79–96.
[17] W.P. Petersen, Some experiments on numerical simulations of stochastic differential equations and a new algorithm, J. Comp. Phys. 113 (1994) 75–81.
[18] S.B. Pope, A Monte Carlo method for the PDF equations of turbulent reactive flow, Combust. Sci. Technol. 25 (1981) 159–174.
[19] S.B. Pope, PDF methods for turbulent reactive flows, Prog. Energy Combust. Sci. 11 (1985) 119–192.
[20] S.B. Pope, Particle method for turbulent flows: integration of stochastic model equations, J. Comp. Phys. 117 (1995) 332–349.
[21] S.B. Pope, Turbulent Flows, Cambridge University Press, Cambridge, UK, 2000.
[22] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, Numerical Recipes in Fortran 77: The Art of Scientific Computing, Cambridge University Press, Cambridge, 1992.
[23] Q. Tang, J. Xu, S.B. Pope, PDF calculations of local extinction and NO production in piloted-jet turbulent methane/air flames, Proc. Combust. Inst. 28 (2000) 133–139.
[24] K.C. Tsai, R.O. Fox, The BMC/GIEM model for micro mixing in non-premixed turbulent reacting flows, IEC Res. 37 (1998) 2131–2141.
[25] W.C. Welton, S.B. Pope, PDF model calculations of compressible turbulent flows using smoothed particle hydrodynamics, J. Comp. Phys. 134 (1997) 150–168.
[26] J. Xu, S.B. Pope, Assessment of numerical accuracy of PDF/Monte Carlo methods for turbulent reacting flows, J. Comp. Phys. 152 (1999) 192–230.
[27] J. Xu, S.B. Pope, PDF calculations of turbulent non-premixed flames with local extinction, Combust. Flame 123 (2000) 281–307.
[28] P.K. Yeung, S.B. Pope, An algorithm for tracking fluid particles in numerical simulations of homogeneous turbulence, J. Comp. Phys. 79 (1988) 373–416.