Computer Physics Communications 289 (2023) 108758

Contents lists available at ScienceDirect

Computer Physics Communications

journal homepage: www.elsevier.com/locate/cpc

Weakly convergent stochastic simulation of electron collisions in plasmas ☆

Wentao Wu [a], Jian Liu [a,b,*], Nathaniel J. Fisch [c], Jianyuan Xiao [a], Huishan Cai [a], Zhaoyuan Liu [d,b], Ruili Zhang [e], Yang He [f]

[a] Department of Plasma Physics and Fusion Engineering, School of Nuclear Science and Technology, University of Science and Technology of China, Hefei, Anhui 230026, China
[b] Advanced Algorithm Joint Lab, Shandong Computer Science Center, Qilu University of Technology, Jinan, Shandong 250014, China
[c] Astrophysical Sciences, Princeton University, Princeton, NJ 08540, USA
[d] Shandong Computer Science Center (National Supercomputing Center in Jinan), Qilu University of Technology (Shandong Academy of Sciences), Jinan, Shandong 250014, China
[e] School of Science, Beijing Jiaotong University, Beijing 100044, China
[f] Department of Applied Mathematics, School of Mathematics and Physics, University of Science and Technology Beijing, Beijing 100083, China

ARTICLE INFO

Article history:
Received 16 June 2022
Received in revised form 25 February 2023
Accepted 17 April 2023
Available online 20 April 2023

Keywords:
Stochastic differential equation
Backward runaway
Weak convergence
Lorentz collision operator
Boltzmann equations

ABSTRACT

Collisions between charged particles can be described and solved using Monte Carlo methods in the framework of stochastic differential equations (SDEs). In this paper, we start from an SDE including the extended Lorentz collision operator, which can recover the collisions between a sampling electron and background ions and electrons. On this basis, we construct a second order weakly convergent algorithm (WCA2) to simulate collisional effects of electrons in plasmas. Superseding the Weiner process by a three-point distribution, WCA2 possesses high weakly convergent accuracy as well as low computational costs. The definition and properties of weak convergence are discussed in detail. The weakly convergent order of WCA2 is verified both theoretically and numerically. Through two trial moment functions, we carefully analyze the numerical solutions of the SDE using rigorous statistical tests in the sense of weak convergence. The criteria and practical operations of finding the benchmark solution of SDEs are introduced at length. In order to illustrate the power of WCA2, we apply it to simulate the backward runaways in plasmas, which is a dramatic physical phenomenon. By comparison with the Euler-Maruyama method and the Cadjan-Ivanov method, the advantage and efficiency of WCA2 is exhibited. The backward runaway probability and its dependence on initial conditions are accurately studied using WCA2.

1. Introduction

Collisions among charged particles are essential processes in plasma physics. Collisional effects participate in many critical physical mechanisms in magnetic confinement fusion researches, such as the momentum scattering, the formation of runaway electrons, and the thermalization of fast alpha-particles. Many coefficients in non-ideal magneto-hydrodynamical equations are determined by collisional processes. The solution of the Fokker-Planck equation requires the accurate calculation of its collisional term. Because collisions in plasmas are microscopic many-body interactions, which are extremely sensitive to the initial distribution in the phase space of all species of particles, the direct calculation of particle trajectories under collisions is impractical.

Monte Carlo methods are widely applied to calculate the distribution function and the expectations of physical quantities through random sampling and stochastic simulations. Collisions in plasmas are thus simulated by first solving the corresponding Newtonian equations with random forces for each sampling particle. Then the distribution function and macroscopic physical quantities can be statistically reconstructed. The random force was originally formulated by Paul Langevin [1], and the equation is hereby named after him. It was proposed as an intuitive approach to explain the Brownian motion following the diffusion theory developed by Einstein [2]. The rigorous mathematical framework for the Langevin equation had not been revealed until Wiener and Ito formulated the integral theory of the Wiener process and the stochastic differential equation (SDE) theory [3–7]. According to their theories, the

☆ The review of this paper was arranged by Prof. David W. Walker.

* Corresponding author.
E-mail address: jliuphy@ustc.edu.cn (J. Liu).

https://doi.org/10.1016/j.cpc.2023.108758
0010-4655/

effects of the random force should be interpreted using the Wiener process, i.e., the integral of a white noise Gaussian process, rather than a simple random variable plugged into the ordinary differential equation. Hence, the Langevin equation was reformed as an SDE. And its connection to the diffusion theory was explicitly revealed by the famous Feynman-Kac formula, which boosted interests of research on SDEs. In most applications [8–16], SDEs have to be solved numerically. Various stochastic numerical algorithms, such as the Euler-Maruyama method [8,9,14,17] and the Milstein method [18], have been proposed.

When dealing with collisions in plasma physics, different kinds of SDEs have been engaged. Jones and Manheimer [14,17] independently developed Coulomb collision models in particle-in-cell (PIC) simulations by applying the traditional Langevin approach. Cadjan and Ivanov [10,11] expressed the Lorentz collision operator in a modern SDE form for the first time. Subsequently, Albright developed the quiet direct simulation Monte Carlo (QDSMC) [8,9] technique utilizing the Ito stochastic integral. These methods have been successfully applied to the study of wave-particle interactions [12] and runaway electrons [19,20]. Additionally, Cadjan and Ivanov proposed a numerical method to perform practical computations for SDEs [10,11], referred to as the Cadjan-Ivanov method. Recently, Zhang et al. presented a modern approach to simulate runaway electrons by solving backward SDEs using the Euler-Maruyama method [21].

Above all, the numerical errors in traditional methods of solving the Langevin equation have not been intensively discussed. Traditional schemes for numerically solving SDEs mainly focus on providing accurate solution trajectories in the phase space [22]. As a result, the measurement of the numerical accuracy is evaluated by its deviation from the exact theoretical trajectory for each sampling particle. The convergent criterion on sampling trajectories is rigorously defined as the strong convergence condition. On the other hand, when talking about Monte Carlo simulations, it is more important to pursue an accurate distribution function or corresponding moments rather than the particle trajectory itself. In other words, although the strong convergence condition guarantees the correctness of the distribution function, it is not necessary. Alternatively, a numerical stochastic method for SDEs can be assessed by its statistical behaviors. Its numerical errors are analyzed based on the deviation of the distribution function and the moments of different orders, instead of a single trajectory. This convergent criterion on statistical quantities is defined to be the weak convergence condition. Sometimes, the strong convergence condition is too strict. There exist weak convergent algorithms for SDEs, which are not strong convergent.

According to the above definitions, the order of both the Euler-Maruyama method and the Cadjan-Ivanov method is 0.5 in the sense of strong convergence. Numerical solutions of these two methods converge to exact solutions with numerical errors on the order of $\sqrt{\Delta}$, where Δ denotes the time stepsize of simulations. From the viewpoint of weak convergence, the distribution of numerical solutions of the above two methods converge on the order of Δ. Then we can conclude that the Euler-Maruyama method and the Cadjan-Ivanov method are both 0.5-order strongly convergent and 1-order weakly convergent algorithms. As another example, the Milstein method [23,24] can improve the strongly convergent order to 1, but it is still first order weakly convergent. It indicates that the accuracy of the distribution function does not necessarily increase with the accuracy of sampling trajectories. Therefore, to increase the accuracy of a Monte Carlo simulation, it is more efficient and economic to improve its weakly convergent behaviors directly.

In this paper, we obtain a second order weakly convergent full three-dimensional Monte Carlo algorithm (WCA2) by discretizing the extended Lorentz collision operator in the modern SDE framework. The standard Lorentz operator represents the collision between an electron and background ions, while the extended Lorentz collision operator also includes the collisional effects from background electrons. Considering electrons in plasmas, their governing SDE can be derived from the Boltzmann equation with extended Lorentz collision terms using the Dynkin formula [4,6]. By further assuming the diffusion coefficient matrix to be symmetric, an analytical SDE form can be deduced using Cadjan and Ivanov's decomposition method [11]. In the sense of weak convergence, the Wiener process, which satisfies a normal distribution, in the stochastic integral of the SDE can be replaced by a three-point distribution process [23,25]. This simplification evidently relieves the computation burden when keeping the same weakly convergent performance.

In order to illustrate the power of the weakly convergent algorithm derived here, we consider the backward runaway problem in plasma. Runaway electrons are intensively studied because they can be accelerated to extremely high energy and play key roles to the safe operation of large tokamaks [26,27]. The runaway problem is a perfect example because it sometimes requires a strongly convergent approach and sometimes can be solved using a weakly convergent approach. For example, to capture the important features of the long term behavior of runaway electrons [28–30] or runaway positrons [31], it is important to use a strongly convergent approach to capture details of the long term individual particle trajectories. However, the important features of so-called backward runaways [20] are inherently stochastic processes that lead to a distribution of runaway probability, which entails a statistical treatment. For a stochastic process, the weakly convergent approach suffices. The application of first principle simulations on runaway electrons based on particle sampling methods has explored new dynamical behaviors and complex physical natures in magnetized plasmas [32–35], including the collisionless pitch-angle scattering phenomenon. Unlike normal runaway electrons, the backward runaway happens when a group of fast electrons run in the same direction along the external electric field. Because of extremely low collisional frequency of fast electrons and three-dimensional nature of the collision process, most fast electrons can first be decelerated in the parallel direction and go through the zero parallel velocity with high perpendicular velocities. Then they are accelerated to runaways by the external electric field instead of being stopped. The backward runaway probability is defined as the chance that this kind of fast electrons finally achieves backward runaways. Thus the runaway probability is, by definition, an integral of the final distribution function over certain phase-space region, which is calculated using WCA2 in this paper. The contour of the backward runaway probability in the initial $v_{||} - v_{\perp}$ velocity space is also exhibited.

For comparison, the Euler-Maruyama method and Cadjan-Ivanov method are also applied to simulate the same backward runaway process. Their weakly convergent orders are analyzed using the ordinary least squares (OLS) regression between the logarithm of weakly convergent numerical errors and the logarithm of time steps [20]. The error analyses are performed on the mean parallel velocity and the total energy, respectively, by order evaluation. The high efficiency of WCA2 is verified at the same time. Given the same weakly convergent numerical error, say 10^{-9}, the time cost of WCA2 is only about half of the Cadjan-Ivanov method and 5% of the Euler-Maruyama method.

The remainder of this paper is organized as follows. In Sec. 2, we introduce the nonrelativistic SDE governing the electron dynamics with collisional effects in plasmas, which is a nonrelativistic version of the one used in Refs. [21,36]. In Sec. 3, the WCA2 algorithm is introduced. And it is compared with the Euler-Maruyama method and the Cadjan-Ivanov method in the sense of weak convergence. In Sec. 4, the backward runaway phenomenon

is simulated and the backward runaway probability is calculated using WCA2. Section 5 finally summarizes this work.

2. Collision model of electrons

The distribution function $f(\mathbf{x}, \mathbf{v}, t)$ of electrons in plasmas evolves according to the Boltzmann equation with a proper collision operator, i.e.,

$$\frac{\partial f}{\partial t} + \boldsymbol{v} \cdot \frac{\partial f}{\partial \boldsymbol{x}} + \frac{q}{m}(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}) \cdot \frac{\partial f}{\partial \boldsymbol{v}} = \left(\frac{\partial f}{\partial t}\right)_c. \tag{1}$$

The collision operator on the right-hand side of Eq. (1) describes the collisional effects as a partial derivative of the distribution function $\left(\frac{\partial f}{\partial t}\right)_c$.

The collisional effects on electrons from background ions are described by the Lorentz collision operator [37]

$$\left(\frac{\partial f}{\partial t}\right)_L = \frac{Z_i \Gamma}{2v^3} \frac{\partial}{\partial \mu}\left(1-\mu^2\right)\frac{\partial f}{\partial \mu}, \tag{2}$$

where $\mu = v_{||}/v$ represents the cosine of the angle between velocity and the external electric field, Z_i is the charge number carried by the ith species of ions. The collisional intensity factor Γ is defined as $\Gamma = n_e q^4 \ln\Lambda / 4\pi\epsilon_0^2 m_e^2$, where n_e and m_e are the density and mass of electrons, respectively, q is the charge carried by an electron, $\ln\Lambda$ is the Coulomb logarithm. In the Lorentz collision operator model, the ion-electron mass ratio is assumed to be infinitely large.

If we also consider the contribution of background electrons, the standard Lorentz collision operator should be amended. The electron-electron collision brings two consequences. One is the friction drag, and the other is the elastic pitch angle scattering. For electrons satisfying $v > v_T$, which run much faster than thermal electrons in the plasma, the additional friction term is written as [16,20]

$$\left(\frac{\partial f}{\partial t}\right)_f = \Gamma \frac{\boldsymbol{v}}{v^3} \cdot \frac{\partial f}{\partial \boldsymbol{v}}. \tag{3}$$

The collision term corresponding to the electron-electron pitch angle scattering is

$$\left(\frac{\partial f}{\partial t}\right)_{se} = \frac{\Gamma}{2v^3} \frac{\partial}{\partial \mu}\left(1-\mu^2\right)\frac{\partial f}{\partial \mu}. \tag{4}$$

The complete extended Lorentz collision operator for electrons can be finally written as

$$\begin{aligned}\left(\frac{\partial f}{\partial t}\right)_c &= \left(\frac{\partial f}{\partial t}\right)_L + \left(\frac{\partial f}{\partial t}\right)_f + \left(\frac{\partial f}{\partial t}\right)_{se} \\ &= \Gamma \frac{\boldsymbol{v}}{v^3} \cdot \frac{\partial f}{\partial \boldsymbol{v}} + \frac{\Gamma(1+Z_i)}{2v^3} \frac{\partial}{\partial \mu}\left(1-\mu^2\right)\frac{\partial f}{\partial \mu}.\end{aligned} \tag{5}$$

Then we can write down the corresponding Ito SDE according to the Boltzmann equation in Eq. (1) with the collision term in Eq. (5) as

$$d\boldsymbol{v}(t) = \boldsymbol{\mu}(\boldsymbol{v})\, dt + \sum_{j=1}^{3} \boldsymbol{\sigma}_j(\boldsymbol{v})\, dW^j, \tag{6}$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}_j$ are three-dimensional vector coefficients, $\boldsymbol{v}$ is the solution stochastic process, W^j stands for standard Wiener process. The SDE Eq. (6) is actually the characteristic line equation for Eq. (1). The initial conditions of Eq. (6) are determined by the sampling of the initial distribution function. And the solution of Eq. (1) can be reconstructed from a cluster of stochastic processes governed by Eq. (6), according to Dynkin formula. The coefficient of Eq. (6) should obey

$$\boldsymbol{\sigma}\boldsymbol{\sigma}^T = (Z_i+1)\,\Gamma \cdot \frac{v^2 \boldsymbol{I} - \boldsymbol{v}\boldsymbol{v}}{v^3}, \tag{7}$$

$$\boldsymbol{\mu} = \frac{q\boldsymbol{E}}{m} + \frac{q}{m}\boldsymbol{v} \times \boldsymbol{B} - (2+Z_i)\Gamma\frac{\boldsymbol{v}}{v^3}, \tag{8}$$

where $\boldsymbol{\sigma}$ is a 3 by 3 matrix composed of $\boldsymbol{\sigma}_j$, and $\boldsymbol{\sigma}_j$ denotes the jth column vector of $\boldsymbol{\sigma}$.

It can be directly observed that the right-hand side of Eq. (7) is a positive definite matrix. The matrix $\boldsymbol{\sigma}$ is its square root and can be numerically solved via Cholesky decomposition. By further assuming that $\boldsymbol{\sigma}$ is symmetric, $\boldsymbol{\sigma}$ can be solved in a closed form using Cadjan and Ivanov's decomposition method [10,11] to be

$$\boldsymbol{\sigma} = \sqrt{\frac{(Z_i+1)\,\Gamma}{v}}\left(\boldsymbol{I} - \frac{\boldsymbol{v}\boldsymbol{v}}{v^2}\right). \tag{9}$$

Submitting Eqs. (8) and (9) into Eq. (6), we obtain the desired master Ito SDE

$$\begin{aligned}d\boldsymbol{v} = &\left[\frac{q}{m}\boldsymbol{E} + \frac{q}{m}\boldsymbol{v} \times \boldsymbol{B} - (2+Z_i)\Gamma\frac{\boldsymbol{v}}{v^3}\right]dt \\ &- \sqrt{\frac{(Z_i+1)\Gamma}{v^5}}\boldsymbol{v} \times \boldsymbol{v} \times d\boldsymbol{W}.\end{aligned} \tag{10}$$

This SDE depicts the dynamics of a sampling electron in external electric and magnetic fields with the collisional effects from background ions and electrons. Its left-hand side is the infinitesimal increment of the velocity vector, while its right-hand side contains two terms. The first term comes from the deterministic force exerted by external fields and the friction drag of electron-electron collisions. The second term denotes the pitch-angle scattering effect under random collisional forces. It corresponds to the random force in the Langevin equation. Here, it is written in a handy notation of an Ito integral in the Wiener process.

For the feasibility of numerical treatments, we rewrite the equation to a dimensionless form as

$$d\tilde{\boldsymbol{v}} = \left(\tilde{\boldsymbol{E}} + \tilde{\boldsymbol{v}} \times \tilde{\boldsymbol{B}} - (2+Z_i)\frac{\tilde{\boldsymbol{v}}}{\tilde{v}^3}\right)d\tilde{t} + \sqrt{1+Z_i}\,\tilde{v}^{-5/2}\,\tilde{\boldsymbol{v}} \times \tilde{\boldsymbol{v}} \times d\tilde{\boldsymbol{W}}. \tag{11}$$

We use the critical velocity, also called the Dreicer velocity [26], v_c as the characteristic velocity. Correspondingly the characteristic collision time is then defined as $\tau = v_c^3/\Gamma$. The critical force is defined as $F_c = \Gamma m_e / v_c^2$. The normalized variables become $\tilde{\boldsymbol{v}} = \boldsymbol{v}/v_c$, $\tilde{t} = t/\tau$, $\tilde{\boldsymbol{E}} = q\boldsymbol{E}/F_c$, $\tilde{\boldsymbol{B}} = q\boldsymbol{B}\tau/m_e$, and $\tilde{\boldsymbol{W}} = \boldsymbol{W}(\tau\tilde{t})/\sqrt{\tau}$, respectively. $\tilde{\boldsymbol{W}}$ is also a standard Wiener process similar to $\boldsymbol{W}(t)$.

3. Second-order weakly convergent algorithm

Under the strong convergence condition, traditional Monte Carlo algorithms aim to the accuracy of trajectories of SDEs. Formally speaking, let a stochastic process $\boldsymbol{X}_t$ be a theoretical solution and $\boldsymbol{X}^{\Delta}(t)$ be its numerical solution. If there exist constants C and Δ_0 independent of the time step Δ, for any $\Delta \in (0, \Delta_0)$, the expectation of the norm of the difference between the two solutions at time T satisfies

$$E\left[\left|\boldsymbol{X}^{\Delta}(T) - \boldsymbol{X}_T\right|\right] \le C\Delta^{\gamma}, \tag{12}$$

we can say that X^{Δ} is strongly convergent [23,25] to X_T with order γ. Note that X_t is a stochastic process. At a given time point T, X_T is a random variable, so is $X^{\Delta}(T)$. The expectation operator

$E[\cdot]$ is applied to the norm of their difference to compare these two random variables.

For random processes, the distribution function and its moments are physically significant. It is more reasonable to assess algorithms based on the numerical errors of the distribution function. Formally, the weak convergence condition can be expressed as

$$\left|E\left[g(\boldsymbol{X}^{\Delta}(T))\right]-E\left[g(\boldsymbol{X}_T)\right]\right|\leq C\Delta^{\beta}, \tag{13}$$

where g is any $2(\beta+1)$ order continuous function $g:\mathbb{R}^d\to\mathbb{R}$, where d is the dimension of $\boldsymbol{X}_t$. The numerical solution $\boldsymbol{X}^{\Delta}$ obeying Eq. (13) is said to be weakly convergent [23] to $\boldsymbol{X}$ at time T with order β. Some simple examples of function g will make this definition more intuitive. For instance, the random process $\boldsymbol{X}(t)=\boldsymbol{v}(t)$ denotes the evolution of stochastic velocity. If we choose $g(\boldsymbol{v})=\boldsymbol{v}$, $E[g(\boldsymbol{v})]$ is just the first moment of distribution function $f(t,\boldsymbol{v})$. The weak convergence condition estimates the numerical error of the expectation of velocity, i.e., the error of the mean velocity. If choosing $g(\boldsymbol{v})=(\boldsymbol{v}-\bar{\boldsymbol{v}})^2$, $E[g(\boldsymbol{v})]$ is actually the second central moment of $f(t,\boldsymbol{v})$. The weak convergence condition checks the expected variance of velocity, which corresponds to the temperature. The convergence of nth moments of $f(t,\boldsymbol{v})$ can be guaranteed in simulations if a βth order weakly convergent algorithm is used with the β in Eq. (13) satisfying $\beta\geq n$.

Strong convergence is a sufficient, but not a necessary condition for weak convergence [23]. Sometimes, the increase of the strongly convergent order does not result in higher weakly convergent order. The Milstein method [23], for example, is first-order strongly convergent, compared to the Euler-Maruyama method with 1/2 strongly convergent order. But they are both first order weakly convergent. In this case, extra computation costs for improving the accuracy of sampling trajectories are invalid for the improvement of physical results. It is more effective to employ algorithms with higher weakly convergent order directly.

We now introduce a second order weakly convergent algorithm, named as WCA2. In order to analyze a weak numerical solution $g(\boldsymbol{X}^{\Delta})$ of an SDE with the general form

$$d\boldsymbol{X}=\boldsymbol{\mu}(\boldsymbol{X})\,dt+\boldsymbol{\sigma}(\boldsymbol{X})\cdot d\boldsymbol{W}, \tag{14}$$

it should be properly expanded into a series. Using Wagner-Platen expansion [4,23], $g(\boldsymbol{X}_\tau)$ can be expressed as

$$g(\boldsymbol{X}_\tau)=\sum_{\alpha\in\mathcal{A}}I_\alpha\left[g_\alpha(\boldsymbol{X}_\rho)\right]_{\rho,\tau}+\sum_{\alpha\in\mathcal{B}(\mathcal{A})}I_\alpha[g_\alpha(\boldsymbol{X}_\cdot)]_{\rho,\tau}, \tag{15}$$

where $I[\cdot]_{\rho,\tau}$ is the Ito integral operator over an interval $[\rho,\tau)$. $\mathcal{A}$ is a hierarchical expansion set, and $\mathcal{B}(\mathcal{A})$ is the corresponding residual set [23]. The integrals in $\mathcal{A}$ are evaluated including the left boundary points of intervals while excluding their right boundary points according to Ito's definition. The dot symbol stands for a certain value in $[\rho,\tau)$, which makes the equality hold. The first term on the right-hand side of Eq. (15) includes terms of a series expanded to a certain order. The second term contains the remainder of the series. Equation (15) is a stochastic analog to the traditional Taylor expansion. If truncating the series by dropping the remainder part, an approximation of the function g is obtained.

The n-tuple $\alpha=(j_1,j_2,\ldots,j_l)$ in the expansion is an l-dimensional multi-index, with each index $j_i\in\{0,1,2,3\}$ when Eq. (14) is a three-dimensional equation. $j_i=0$ corresponds to the time component, and $j_i=1,2,3$ corresponds to each spatial component, respectively. Ito integral with a given multi-index α is calculated recursively as

$$I_\alpha[g]_{\rho,\tau}=\begin{cases}f_\tau & l=0\\ \int_\rho^\tau I_{\alpha-}[g]_{\rho,s}\,ds & l>0 \text{ and } j_l=0\\ \int_\rho^\tau I_{\alpha-}[g]_{\rho,s}\,dW_s^{j_l} & l>0 \text{ and } j_l\neq 0\end{cases}, \tag{16}$$

where $\alpha-=(j_1,j_2,\ldots,j_{l-1})$ is an (l-1)-dimensional multi-index. And W^j is the jth component of $\boldsymbol{W}$.

Define L as the Ito differential operator, which satisfies

$$L^0=\frac{\partial}{\partial t}+\boldsymbol{\mu}\cdot\frac{\partial}{\partial\boldsymbol{x}}+\frac{1}{2}\boldsymbol{\sigma}\boldsymbol{\sigma}^T:\frac{\partial^2}{\partial\boldsymbol{x}\partial\boldsymbol{x}}, \tag{17}$$

$$L^j=\boldsymbol{\sigma}^j\cdot\frac{\partial}{\partial\boldsymbol{x}^j}, \tag{18}$$

where $\boldsymbol{\sigma}^j$ is jth column of matrix $\boldsymbol{\sigma}$. Then the coefficient function of $g(\boldsymbol{X})$ with multi-index α is recursively calculated as

$$g_\alpha=\begin{cases}g & l=0\\ L^{j_1}g_{-\alpha} & l\neq 0\end{cases}, \tag{19}$$

where the (l-1)-dimensional multi-index $-\alpha=(j_2,j_3,\ldots,j_l)$.

When keeping up to second-order terms in hierarchical expansion set, $\mathcal{A}=\{\phi,(j_1),(j_1,j_2)\}$, where ϕ stands for a 0-length list and j_1,j_2 iterate over $\{0,1,2,3\}$. The residual set becomes $\mathcal{B}=\{(j_1,j_2,j_3),\ldots\}$. The expansion with multi-indices in $\mathcal{A}$ is written as

$$\begin{aligned}\boldsymbol{X}_t=&\boldsymbol{X}_0+\boldsymbol{\mu}I_{(0)}+\boldsymbol{\sigma}\cdot\left[I_{(1)},I_{(2)},I_{(3)}\right]^T\\&+\left(\boldsymbol{\mu}\cdot\frac{\partial\boldsymbol{\mu}}{\partial\boldsymbol{x}}+\frac{1}{2}\boldsymbol{\sigma}\boldsymbol{\sigma}^T:\frac{\partial^2}{\partial\boldsymbol{x}\partial\boldsymbol{x}}\boldsymbol{\mu}\right)I_{(0,0)}\\&+\left(\boldsymbol{\mu}\cdot\frac{\partial\boldsymbol{\sigma}}{\partial\boldsymbol{x}}+\frac{1}{2}\boldsymbol{\sigma}\boldsymbol{\sigma}^T:\frac{\partial^2}{\partial\boldsymbol{x}\partial\boldsymbol{x}}\boldsymbol{\sigma}\right)\cdot[I_{(0,1)},I_{(0,2)},I_{(0,3)}]^T\\&+\boldsymbol{\sigma}\cdot\frac{\partial\boldsymbol{\mu}}{\partial\boldsymbol{x}}\cdot[I_{(1,0)},I_{(2,0)},I_{(3,0)}]^T\\&+\boldsymbol{\sigma}\cdot\frac{\partial\boldsymbol{\sigma}}{\partial\boldsymbol{x}}:\begin{bmatrix}I_{(1,1)}&I_{(1,2)}&I_{(1,3)}\\I_{(2,1)}&I_{(2,2)}&I_{(2,3)}\\I_{(3,1)}&I_{(3,2)}&I_{(3,3)}\end{bmatrix}+R,\end{aligned} \tag{20}$$

where ":" denotes tensor contraction. The leading integral terms in $\mathcal{A}$ are explicitly written as

$$I[g]_{0,\tau}=g(\boldsymbol{X}_\tau) \tag{21}$$

$$I_{(0)}[g]_{0,\tau}=\int_0^\tau g(\boldsymbol{X}_s)\,ds \tag{22}$$

$$I_{(1)}[g]_{0,\tau}=\int_0^\tau g(\boldsymbol{X}_s)\,dW_s^1 \tag{23}$$

$$I_{(2)}[g]_{0,\tau}=\int_0^\tau g(\boldsymbol{X}_s)\,dW_s^2 \tag{24}$$

$$I_{(0,0)}[g]_{0,\tau}=\int_0^\tau\int_0^{s_2} g(\boldsymbol{X}_{s_1})\,ds_1\,ds_2 \tag{25}$$

$$I_{(0,1)}[g]_{0,\tau}=\int_0^\tau\int_0^{s_2} g(\boldsymbol{X}_{s_1})\,ds_1\,dW_{s_2}^1 \tag{26}$$

Assuming the time step $\Delta=\tau-\rho$ is sufficiently small, each above integral can be numerically approximated by

$$I_\alpha\left[g_\alpha(X_\rho)\right]_{\rho,\tau}\approx g_\alpha(X_\rho)\,I_\alpha[1]\rho,\tau. \tag{27}$$

Meanwhile, dW^j can be replaced by $\Delta W^j = W_\tau^j - W_\rho^j$, which obeys Gaussian distribution $\mathcal{N}(0,\Delta)$ according to the definition of Wiener process.

According to definition, a Gaussian random variable ΔW^j should be used to calculate all the integrals I_α. But in the sense of weak convergence, ΔW can be approximated using other random variables, say $\Delta\hat{W}$. The difference of moments caused by replacing ΔW with $\Delta\hat{W}$ should be of $\beta+1$ order to preserve the weak order [23]. The weak approximation of the integrals on multi-indices has the following moment properties related to dW,

$$E[I_{(1)}[1]] = 0\,, \tag{28}$$

$$E[I_{(1,0)}[1]] = \int E[W_s]ds = 0\,, \tag{29}$$

$$E[I_{(0,1)}[1]] = E\left[\int dW_s\right] = 0\,, \tag{30}$$

$$E[I_{(1,1)}[1]] = E\left[\int\int dW_s dW_s\right] = \int ds = \Delta\,, \tag{31}$$

$$E[I_{(1,1,1,1)}[1]] = E\left[\int\cdots\int dW_s^4\right] = 3\Delta^2\,. \tag{32}$$

For second order weakly convergent algorithms, the general requirement of $\Delta\hat{W}$ can be specifically written to

$$\begin{aligned}&\left|E[\Delta\hat{W}]-0\right| + \left|E[\Delta\hat{W}^2]-\Delta\right| + \left|E[\Delta\hat{W}^3]-0\right| \\ &+ \left|E[\Delta\hat{W}^4]-3\Delta^2\right| + \left|E[\Delta\hat{W}^5]-0\right| \le K\Delta^{2+1}.\end{aligned} \tag{33}$$

Any random variable satisfying Eq. (33) can support second order weakly convergent algorithms. There is a large freedom to choose $\Delta\hat{W}$ in practical computations.

We then choose a simple three-point distribution of random variables as $\Delta\hat{W}^j$. The probability distribution of $\Delta\hat{W}^j$ obeys

$$P\left(\Delta\hat{W}^j = \pm\sqrt{3\Delta}\right) = \frac{1}{6}, \qquad P\left(\Delta\hat{W}^j = 0\right) = \frac{2}{3}. \tag{34}$$

We can directly verify that the three-point distribution meets the requirement of the second order weak convergence condition by directly using Eq. (33). However, it is much easier and economic to generate the random variable $\Delta\hat{W}$ and use it in computations practically. The choice of the three-point distribution simplifies stochastic simulations evidently when maintaining the same weakly convergent order.

An explicit second-order weak convergence algorithm can thus be constructed [23]. For convenience we define two groups of supporting vector variables

$$\begin{aligned}\overline{R}_\pm^j &= X_n + \mu(X_n)\Delta \pm \sigma^j(X_n)\sqrt{\Delta}, \\ \overline{U}_\pm^j &= X_n \pm \sigma^j(X_n)\sqrt{\Delta},\end{aligned} \tag{35}$$

where μ and σ^j are the same as in Eq. (14). Using these supporting variables, numerical multiple integrals of the first and second order are constructed as

$$\begin{aligned}\Upsilon_{c1} = \frac{1}{4}\sum_{j=1}^{3}\Bigg[&\sigma^j(\overline{R}_+^j) + \sigma^j(\overline{R}_-^j) + 2\sigma^j(X_n) \\ &+ \sum_{\substack{r=1\\ r\neq j}}^{3}\left(\sigma^j(\overline{U}_+^r) + \sigma^j(\overline{U}_-^r) - 2\sigma^j(X_n)\right)\Bigg]\Delta\hat{W}^j,\end{aligned} \tag{36}$$

$$\begin{aligned}\Upsilon_{c2} = \frac{\Delta^{-\frac{1}{2}}}{4}\sum_{j=1}^{3}\Bigg[&\left(\sigma^j(\overline{R}_+^j) - \sigma^j(\overline{R}_-^j)\right)\left(\left(\Delta\hat{W}^j\right)^2 - \Delta\right) \\ &+ \sum_{\substack{r=1\\ r\neq j}}^{3}\left(\sigma^j(\overline{U}_+^r) - \sigma^j(\overline{U}_-^r)\right)\left(\Delta\hat{W}^j\Delta\hat{W}^r + V_{r,j}\right)\Bigg],\end{aligned} \tag{37}$$

where

$$V = -\Delta\mathbf{I} + \Xi = \begin{pmatrix}-\Delta & 0 & 0\\ 0 & -\Delta & 0\\ 0 & 0 & -\Delta\end{pmatrix} + \begin{pmatrix}0 & \xi_{12} & \xi_{13}\\ -\xi_{12} & 0 & \xi_{23}\\ -\xi_{13} & -\xi_{23} & 0\end{pmatrix} \tag{38}$$

is a stochastic matrix, which is the sum of a diagonal matrix $-\Delta\mathbf{I}$ and an antisymmetric matrix Ξ. The entries ξ_{ij} of Ξ are random variables which take the value Δ or $-\Delta$ with the same probability.

Finally, the explicit form of WCA2 can be written out as

$$\overline{\Upsilon} = X_n + \mu(X_n)\Delta + \sum_{j=1}^{3}\sigma^j(X_n)\Delta\hat{W}^j, \tag{39}$$

$$X_{n+1} = X_n + \frac{1}{2}\left(\mu\left(\overline{\Upsilon}\right) + \mu(X_n)\right)\Delta + \Upsilon_{c1} + \Upsilon_{c2}. \tag{40}$$

The first two terms on the right-hand side of Eq. (40) are actually equivalent to the deterministic Euler method using a predictor-corrector scheme. The other two terms come from Wagner-Platen expansion with time step Δ. Υ_{c1} yields the first-order integral term of the Ito expansion, and Υ_{c2} improves the weak accuracy of the final results to second order.

When using weakly convergent methods, samples of the distribution function $f(x,t)$ are calculated. The statistical behavior of these samples is assumed to obey the original partial differential equation, i.e., Eq. (1). For a given function $g(x,t)$, its expectation is evaluated by

$$E_f[g(x,T)] = E[g(X(T))] = \lim_{M\to\infty}\frac{1}{M}\sum_{i=1}^{M}g(X_i(T)), \tag{41}$$

where $X_i(T)$ is the ith independent sample for the random variable $X(T)$ at time T. Two kinds of errors arise in practical estimation of the expectation, i.e., the systematic error and statistical error. The systematic error comes from the discretization of time. For finite time step Δ, the systematic error is

$$e_s = E[g(X^\Delta(T))] - E[g(X(T))], \tag{42}$$

which is restricted by weak order condition Δ^β for βth order weak order algorithms. The statistical error results from the finite discrete sampling. For finite sample number N, the statistical error is

$$e_N = \frac{1}{N}\sum_{i=1}^{N}g(X_i^\Delta(T)) - E[g(X^\Delta(T))], \tag{43}$$

which is a random variable. As the sample number N approaches infinity, the distribution of e_N asymptotically converges to a Gaussian distribution with zero mean, and its variance obeys

$$var(e_n) = var\left(\frac{1}{N}\sum_{i=1}^{N}g(X_i^\Delta(T))\right) = \frac{1}{N}var\left(g(X^\Delta(T))\right), \tag{44}$$

according to the famous large number law. The noise level of g is actually the standard deviation, i.e., the square root of its variance, which decreases with $1/\sqrt{N}$ as N goes to infinity.

4. Simulations of backward runaways

4.1. Benchmark solution

The exact theoretical solutions of the distribution function and moments of different orders cannot be obtained in most cases. Benchmark solutions are then required to evaluate the numerical errors of weakly convergent algorithms. A common way is to use a numerical solution with an extremely small time step as the benchmark solution in traditional numerical analysis. But it is not enough to the analysis of weak convergence for numerical algorithms for SDE and statistical quantities. In this subsection, we introduce a more rigorous treatment of benchmark solutions. To make sure that the benchmark solution is accurate enough in the sense of weak convergence, we propose two general criteria for it. Firstly, the benchmark solution does not change with the decrease of time step. Secondly, the benchmark solution does not depend on the choice of numerical method.

Different from deterministic ODEs, numerical solutions of SDEs are random variables at given a time. For practical purpose, hypothetical tests should be applied to verify whether a numerical solution meets the benchmark criteria. Because expectations of two normally distributed variables cannot be distinguished at a certain confidence level, the first criterion could be decomposed into two related tests. One is to detect whether the candidate benchmark solution is normally distributed, which can be realized by performing the Shapiro test. The other is to check the identity of expectations of two numerical solutions with different given time steps. Variances of solutions may differ with different time steps. The Welch's unequal variance t-test can undertake this task. If both null hypotheses cannot be rejected, we cannot tell the difference between two solutions, and the first criterion is guaranteed.

To satisfy the second criterion, at least two numerical SDE algorithms should be performed to produce candidate benchmark solutions. To compare the solution of different numerical methods, the analysis of variance (ANOVA) technique is used. In addition, normality tests are still required as well. If all solutions are normally distributed and the F-test fails to reject the null hypotheses, the second criterion is satisfied.

4.2. Analysis of weak convergence

The weakly convergent order of a stochastic algorithm is formally defined by Eq. (13). It can also be practically analyzed and confirmed via the regression between numerical errors of statistical quantities and time steps in a particular simulation case. In this subsection we use WCA2, Euler-Maruyama method, and Cadjan-Ivanov method, respectively, to simulate a backward runaway system for comparison. We look into two continuous trial functions, i.e., $g(\boldsymbol{v}) = v_x$ and $g(\boldsymbol{v}) = |v|^2$, of the stochastic velocity $\boldsymbol{v}$. The numerical solution of Eq. (11) with time step Δ at time T is denoted by $\boldsymbol{v}_T^{\Delta}$. The expectation $E[g(\boldsymbol{v}_T^{\Delta})]$ is calculated by statistically averaging $g(\boldsymbol{v}_T^{\Delta})$ over all sampling processes. Simulations with the same physical setup can be repeated using different time steps. Corresponding numerical errors of $E[g(\boldsymbol{v}_T^{\Delta})]$ are recorded by referring to the benchmark solution. Finally, by performing the ordinary least square (OLS) method, the weakly convergent order is calculated as the slope of the regression curve in the plot of the numerical errors versus time steps Δ (both with logarithmic scale).

Give a time step, the mean and variance of $E[g(\boldsymbol{v}_T^{\Delta})]$ are calculated in batches according to the following procedure, see Algorithm 1.

In this case, we set the number of samples in each batch to $N = 100{,}000$ and the number of batches to $M = 30$. Here, we use a factor K to represent the time stepsize of the logarithm by defining

Data: Divide $N \times M$ samples into M batches with N samples in each batch. Determine the continuous trial function $g(\boldsymbol{v})$, end time T, and time step Δ.
Result: $E[g(\boldsymbol{v}_T^{\Delta})]$ and its variance
for $m \leftarrow 0$ **to** $M-1$ **do**
  for $n \leftarrow 0$ **to** $N-1$ **do**
    Generate the nth sample of the stochastic process $\boldsymbol{W}$ in the mth batch;
    Solve the SDE for the sample and obtain the numerical solution $\boldsymbol{v}_T^{\Delta}(n)$;
  end
  calculate the expectation in the mth batch by $E[g(\boldsymbol{v}_T^{\Delta})](m) = \sum_n g[\boldsymbol{v}_T^{\Delta}(n)]/N$;
end
calculate the mean and variance of the m quantities $E[g(\boldsymbol{v}_T^{\Delta})](m)$;

Algorithm 1: Procedure of calculating $E[g(\boldsymbol{v}_T^{\Delta})]$ and corresponding variance.

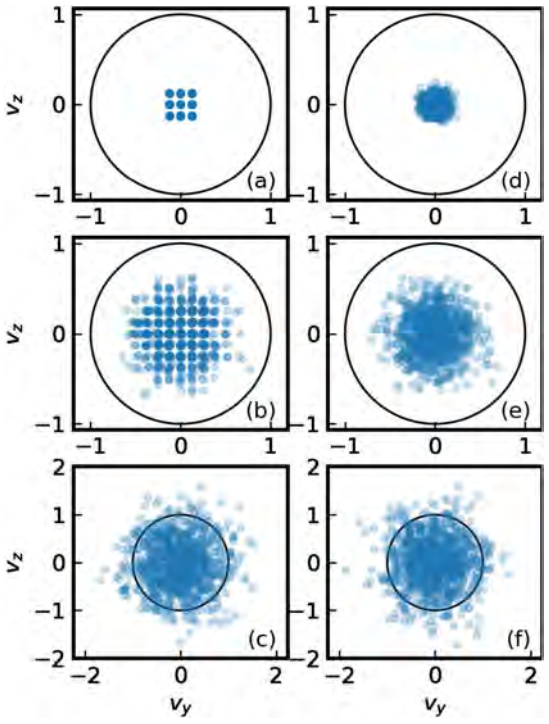

Fig. 1. The distribution evolution of sampling runaways in the velocity space. The time step is $K = -log_2(\Delta) = 7$, i.e., $\Delta = 2^{-7}$. The numerical results using WCA2 at $\tilde{t} = \Delta, 10\Delta, 50\Delta$ are plotted in the left column in subfigures (a), (b) and (c). The numerical results using the Euler-Maruyama method at $\tilde{t} = \Delta, 10\Delta, 50\Delta$ are plotted in the right column in subfigures (d), (e) and (f).

$K = -log_2(\Delta)$. The largest time step is set to $\Delta = 1.0$. We carry out simulations by scanning K from 0 to 7, that is the time step from $\Delta = 2^0$ to $\Delta = 2^{-6}$, with totally 8 different time steps. Other parameters are set as follows. The end time is set to $\tilde{t} = 1.0$, the charge number of ions $Z_i = 1$, the external electric field is along the x-axis as $\tilde{\boldsymbol{E}} = \{1, 0, 0\}$, and the initial velocity of all sampling electrons is set to $\tilde{\boldsymbol{v}} = \{3, 0, 0\}$. The external magnetic field is set to be zero.

The simulated distributions of sampling electrons in the velocity space at different time, using WCA2 and the Euler-Maruyama method, are plotted in Fig. 1. The distributions in the velocity space using different numerical methods evolve differently at the beginning, see subfigures (a) and (d) in Fig. 1. Note that WCA2 only guarantees the accuracy of the expectations of stochastic functions. The numerical trajectories of sampling runaway electrons are governed by a much simpler three-point distribution in WCA2, instead of the smoother but much more complex Gaussian distribution. The simulated distribution using WCA2 splits into nine grids in

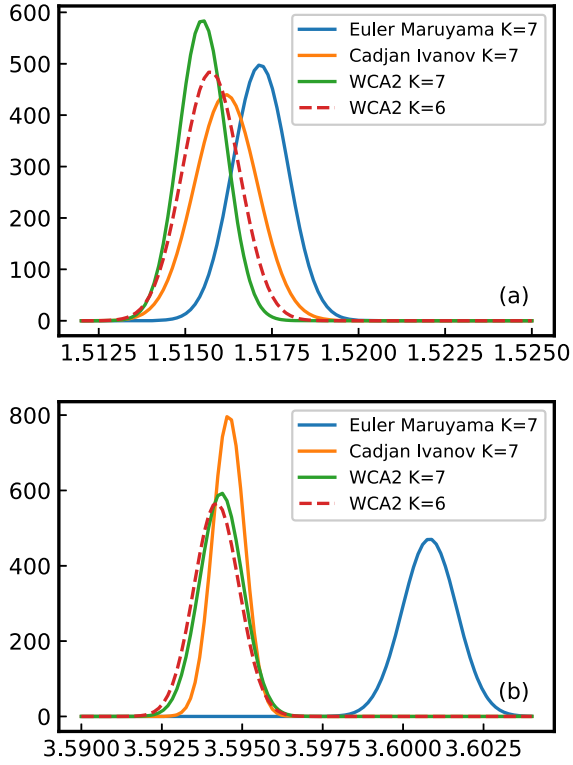

Fig. 2. The plots of distributions of trial functions using different numerical methods. (a) The plot of the distribution of the x-component of velocity $v_x^{\Delta}(T)$ using the Euler-Maruyama method (blue), the Cadjan-Ivanov method (orange), and WCA2 (green) with $K = 7$, and WCA2 with $K = 6$ (dashed red). (b) The plot of the energy $|v^{\Delta}|^2(T)$ using the Euler-Maruyama method (blue), the Cadjan-Ivanov method (orange), and WCA2 (green) with $K = 7$, and WCA2 with $K = 6$ (dashed red). The solution of the Euler-Maruyama method deviates from other distributions clearly. (For interpretation of the colors in the figures, the reader is referred to the web version of this article.)

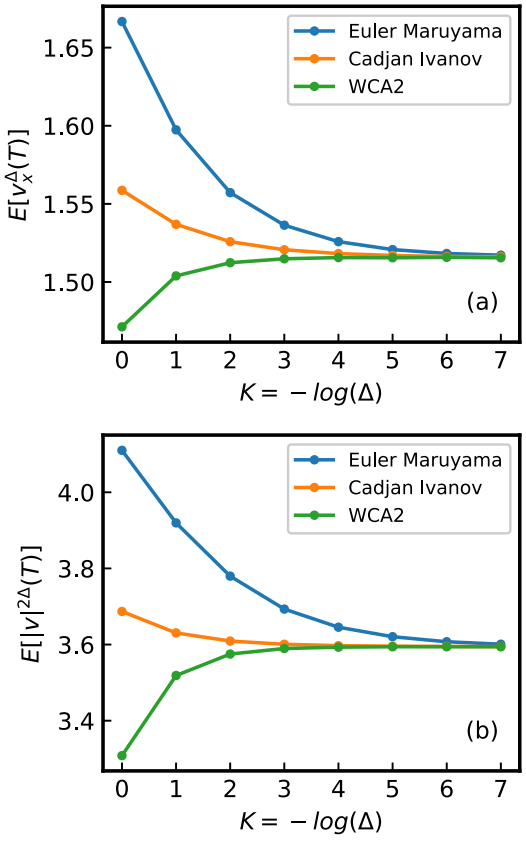

Fig. 3. The plots of expectations of trail functions using different numerical methods with different time steps. (a) The plot of the averaged x-component of velocity $E[v_x^{\Delta}(T))]$ versus the time-step factor $K = -log_2(\Delta)$ using the Euler-Maruyama method (blue), the Cadjan-Ivanov method (orange), and WCA2 (green). (b) The plot of averaged energy $E[|v^{\Delta}|^2(T)]$ versus the time-step factor $K = -log_2(\Delta)$ using the Euler-Maruyama method (blue), the Cadjan-Ivanov method (orange), and WCA2 (green). It is evident that with the decrease of time step, the expectations for three different methods converge.

the $v_y - v_z$ plane at first. But WCA2 still accurately simulates all the moments of orders less than 2. As time goes on, numerical distributions using two different numerical methods converge, see subfigures (c) and (f) in Fig. 1.

The weak convergence of WCA2 can also be verified. We examine the trial functions $g(\boldsymbol{v}) = v_x$, which is the electron velocity parallel to the external electric field, and $g(\boldsymbol{v}) = |v|^2 = v_x^2 + v_y^2 + v_z^2$, which represents the energy of runaway electrons. Because of the symmetry of the problem, the expectations of v_y and v_z are both 0.

The expectations of $g(\boldsymbol{v}_T^{\Delta})$ versus the time-step factor K, using WCA2, the Euler-Maruyama method, and Cadjan-Ivanov method, are plotted in Fig. 3. All three numerical methods converge to the same expectation with the decrease of time step, i.e., the increase of K. The first moment $E[v_x^{\Delta}]$ is plotted in Panel (a), and the second moment $E[|v^{\Delta}|^2]$ is plotted in Panel (b).

To show the weakly convergent order of each algorithm, the benchmark solution is required. The solution of $E[g(\boldsymbol{v}_T^{\Delta})]$ using WCA2 with $K = 7$ is chosen to be the potential candidate. The normal distribution of trail functions as stochastic quantities is verified, see Fig. 2. Fig. 2 exhibits the plots of distributions of trial functions using different numerical methods. The normality of these distributions is formally tested using Shapiro tests, and the statistical values are listed in Table 1. According to the listed data, all distributions in Fig. 2 are qualified to be normally distributed.

It can also be examined that the solution from WCA2 is almost the same when $K = 6$ and $K = 7$, see Fig. 3. Meanwhile, the solutions from all three numerical methods are the same when $K = 7$. We now quantitatively demonstrate that the two solutions using WCA2 with $K = 6$ and $K = 7$ are sufficiently close, by performing Welch's unequal variance t-tests. The statistical data and corresponding p-values are listed in Table 2. Solutions from the Euler-Maruyama method and the Cadjan-Ivanov method are listed as well, but the statistical results show that they cannot meet the criteria.

According to Table 2, when the time step decreases from $K = 6$ to $K = 7$, the p-value from WCA2 of two trial moment functions are all above 50%. In term of null-hypothetical test, we cannot reject the null hypothesis that the two solutions are the same. In other words, the candidate benchmark solution is indeed close enough to the solution with $K = 6$.

Finally, we compare the solutions from three different numerical methods with the time-step factor $K = 7$. The equality of the mean of solutions from different methods is tested using the ANOVA technique. ANOVA is a statistical tool to test whether given independent normally distributed datasets are identical. The advantage of ANOVA is that it can be applied to more than two datasets at once. Since the normality of the datasets has already been verified using Shapiro test, we perform ANOVA directly. The ANOVA analysis results are listed in Table 3.

According to Table 3, numerical solutions of $E[v_x^{\Delta}(T)]$ and $E[|v^{\Delta}|^2(T)]$ from three different methods are not identical. In Fig. 2, the solution from the Euler-Maruyama method converges so slow that it produces an outlier solution when $K = 7$. So we change the test strategy and abandon the Euler-Maruyama method. We only focus on the Cadjan-Ivanov method and WCA2. Referring

Table 1
Shapiro normality tests of the trial moment functions $E[v_x^{\Delta}(T))]$ and $E[|v^{\Delta}|^2(T)]$ obtained from the Euler-Maruyama method, the Cadjan-Ivanov method and WCA2 when $K = 6$ and $K = 7$.

Expectation	Methods	$K = 6$		$K = 7$	
		W-statistic	p-value	W-statistic	p-value
$E[g_1(v_T^{\Delta})]$	Euler-Maruyama	0.9288509488	0.0457747914	0.9448584914	0.1229689941
	Cadjan-Ivanov	0.9723019600	0.6038923264	0.9418374896	0.1019422486
	WCA2	0.9591118097	0.2939443886	0.9595355392	0.3014061153
$E[g_2(v_T^{\Delta})]$	Euler-Maruyama	0.9620262384	0.3486383259	0.9357385635	0.0698712692
	Cadjan-Ivanov	0.9853706956	0.9433261156	0.9483121037	0.1523194462
	WCA2	0.9637868404	0.3855869174	0.9749118686	0.6801327467

Table 2
Welch's unequal variance tests of the moment functions $E[v_x^{\Delta}(T))]$ and $E[|v^{\Delta}|^2(T)]$ using the Euler-Maruyama method, the Cadjan-Ivanov method and WCA2 when $K = 6$ and $K = 7$. The statistics t, degree of freedom ν, and the p-value are listed in the table. The p-values of WCA2 are all above 50%.

Expectation	Methods	$K = 6$			$K = 7$			t	ν	p-value (two-sided)
		M	Mean	Std	M	Mean	Std			
$E[v_x^{\Delta}(T)]$	Euler-Maruyama	30	1.5181471154	0.0010572291	30	1.5167099144	0.0009388375	-5.4738626822	57.2007133978	0.0000010189
	Cadjan-Ivanov	30	1.5161602266	0.0011350644	30	1.5155411902	0.0008255942	-2.3751147326	52.9744270096	0.0211956264
	WCA2	30	1.5155835766	0.0011201136	30	1.5154497894	0.0011551838	-0.4477535215	57.9449611347	0.6560006370
$E[\|v^{\Delta}\|^2(T))]$	Euler-Maruyama	30	3.6075405807	0.0011471987	30	3.6011830706	0.0007982998	-24.4960634146	51.7508692636	0.0000000000
	Cadjan-Ivanov	30	3.5952511325	0.0005412185	30	3.5945317579	0.0007720412	-4.1554688610	52.4304111506	0.0001202918
	WCA2	30	3.5944935081	0.0008078345	30	3.5943761437	0.0008143028	-0.5510099036	57.9963114240	0.5837434234

Table 3
ANOVA results of $E[v_x^{\Delta}(T)]$ and $E[|v^{\Delta}|^2(T)]$ at $K = 7$ using the Cadjan-Ivanov method and WCA2, with and without the Euler-Maruyama method. According to Fig. 2 the solution of $E[|v^{\Delta}|^2(T)]$ using the Euler-Maruyama method is an outlier, this statistics confirms this result. By removing the Euler-Maruyama method from comparison, the test p-value can be largely increased.

Expectation	Methods Set		Sum Squares	Degree of Freedom	Mean Squares	F-statistic	p-value
$E[v_x^{\Delta}(T)]$	Euler-Maruyama and	Treatments	0.0000296219	2	0.0000148109	14.8238525409	0.0000028839
	Cadjan-Ivanov and	Error	0.0000869241	87	0.0000009991		
	WCA2	Total	0.0001165460	89			
$E[v_x^{\Delta}(T)]$	Cadjan-Ivanov and	Treatments	0.0000001253	1	0.0000001253	0.1201698569	0.7301048951
	WCA2	Error	0.0000604817	58	0.0000010428		
		Total	0.0000606070	59			
$E[\|v^{\Delta}\|^2(T)]$	Euler-Maruyama and	Treatments	0.0009059843	2	0.0004529922	700.0487590997	0.0000000000
	Cadjan-Ivanov and	Error	0.0000562965	87	0.0000006471		
	WCA2	Total	0.0009622808	89			
$E[\|v^{\Delta}\|^2(T)]$	Cadjan-Ivanov and	Treatments	0.0000003632	1	0.0000003632	0.5666709067	0.4546294465
	WCA2	Error	0.0000371781	58	0.0000006410		
		Total	0.0000375413	59			

Table 4
Ordinary Least Squares regression over logarithms of numerical errors and logarithms of time steps using the Euler-Maruyama method, the Cadjan-Ivanov and WCA2 on $E[v_x^{\Delta}(T)]$ and $E[|v^{\Delta}|^2(T)]$, respectively. Slopes of the Euler-Maruyama method and the Cadjan-Ivanov method are about 1.0 and the slope of WCA2 is around 2.0.

Expectation	Method	slope mean	slope std.	slope t-statistics	p-value	R^2
$E[v_x^{\Delta}(T)]$	Euler-Maruyama	-0.9833	0.011	-93.483	0.000	0.999
	Cadjan-Ivanov	-1.0038	0.018	-56.783	0.000	0.998
	WCA2	-1.7603	0.087	-20.265	0.000	0.990
$E[\|v^{\Delta}\|^2(T)]$	Euler-Maruyama	-0.9091	0.020	-45.462	0.000	0.997
	Cadjan-Ivanov	-1.1313	0.042	-27.137	0.000	0.993
	WCA2	-2.0024	0.027	-75.199	0.000	0.999

to the data in Table 3, it is clear that solution from the Cadjan-Ivanov method and WCA2 cannot be distinguished when $K = 7$.

Passing all the above tests, we confidently conclude that the candidate solution from WCA2 with $K = 7$ meets all the defined criteria and can be used as the benchmark solution.

With the help of the benchmark solution, the weakly convergent error for all numerical methods with different time steps can be calculated. The exact weakly convergent order of algorithms is obtained by OLS regression. The order is just the negative slope of the regression line by the logarithms of errors versus K. The regression lines using different methods are plotted in Fig. 4, while detailed statistical data are listed in Table 4.

The numerical error of expectation of the first-order continuous function $E[v_x\Delta(T)]$ and the second-order continuous function $E[|v^{\Delta}|^2(T)]$ are plotted in subplots Fig. 4a and Fig. 4b, respectively. It can be observed that the numerical errors decrease with time step. The weakly convergent order of each algorithm is carefully described using the slope values in Table 4. The Euler-Maruyama method and the Cadjan-Ivanov method have the slope value of 0.98 and 1.00, respectively, for $E[v_x^{\Delta}(T)]$, while WCA2 has a slope

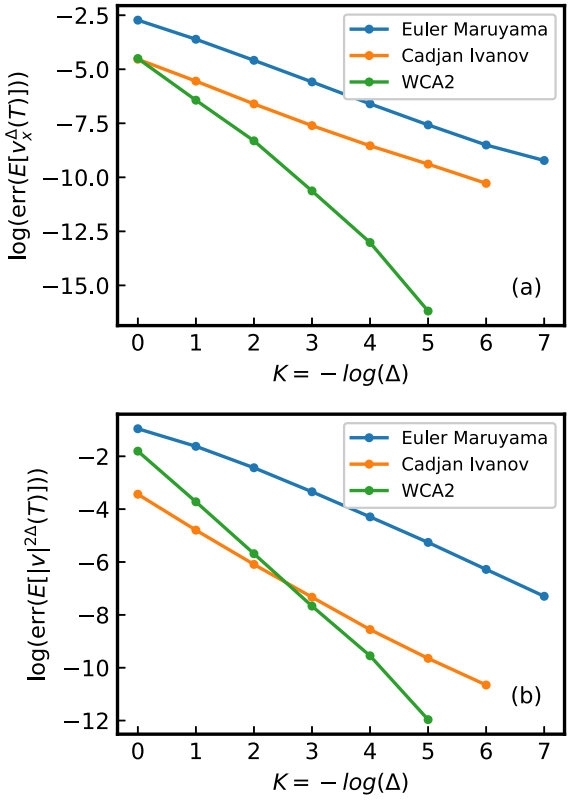

Fig. 4. Plots of weakly convergent errors of $E[v_x^{\Delta}(T)]$ (a) and $E[|v^{\Delta}|^2(T)]$ (b) with logarithmic scale versus $K = -log_2(\Delta)$ from the Euler-Maruyama method (blue), the Cadjan-Ivanov method (orange) and WCA2 (green).

value of 1.76. For $E[|v^{\Delta}|^2(T)]$, the Euler-Maruyama method and the Cadjan-Ivanov method have the slope value of 0.92 and 1.21, respectively, while WCA2 has a slope valued of 2.00. As expected, the weakly convergent order of WCA2 is around 2.0. The weakly convergent order of the Euler-Maruyama method and the Cadjan-Ivanov method is 1.0, which agrees with their theoretical order. When $K > 3$, WCA2 behaves much better than the other two methods.

To inspect and compare the variance of the three methods, Fig. 2 provides a qualitative analysis. Regardless of the heights and positions of peaks, the standard deviations (width) of all distributions are of the same level. Table 2 quantitatively reflects the statistical errors of different methods. In this table, when checking $E[v_x^{\Delta}(T)]$ for $K = 6$, the three numerical methods provide similar standard deviations of approximately 0.0011. For $K = 7$, all standard deviations decrease to approximately 0.0010. For $E[|v^{\Delta}|^2(T)]$, disregarding the unreliable Euler-Maruyama result, Cadjan-Ivanov and WCA2 yield similar standard deviations of approximately 0.0008. In summary, the standard deviation measures the noise level. With the same sample number, three algorithms of different weakly convergent orders have similar statistical (noise) errors.

We also compare the time consumption and computation efficiency of the three methods. We perform the benchmark computation on an Intel i7 cpu machine with 64G RAM on Linux system. When reducing the numerical error of $E[v_x^{\Delta}(T)]$ to the level of 10^{-9}, the Euler-Maruyama method costs 26.374 s, and the Cadjan-Ivanov method costs 3.411 s. By contrast WCA2 only uses 1.510 s. When reducing the numerical error of $E[|v^{\Delta}|^2(T)]$ to the same level, the Euler-Maruyama method cost 95.94 s, the Cadjan-Ivanov method costs 6.796 s, and the second-order method costs 5.780 s. As a second-order weakly convergent algorithm, WCA2 behaves better when requiring higher accuracy, compared with the Euler-Maruyama method and the Cadjan-Ivanov method.

4.3. Runaway probability

The mechanism of backward runaways is similar to that of standard runaway electrons, but with their initial velocities opposite to the electric force. Because the electrons move with velocities larger than the critical velocity and opposite to the electric force initially, they are dragged to slow down at the beginning. If the velocity of electrons drops down to the thermal velocity v_T, the extended Lorentz collision operator is no longer valid [19]. The collision frequency is assumed to a constant [19] and the electron is assumed to be thermalized. The distribution of the velocity of background thermal electrons obeys normal distribution in the velocity space. When the parallel velocity is slowed down to zero, most backward runaways still have large perpendicular velocity to make sure they are not stopped. They are then accelerated by the external field in the parallel direction and reach runaway state. The collisional pitch-angle scattering may also provide sufficient perpendicular velocity for backward runaways [13,19] to avoid the stopping phenomenon. On the other hand, if the external electric field is large enough, the electric force exerted on electrons can easily overcome the collisional friction. Large amount of electrons will be readily accelerated to runaways. In the formation of backward runaways, the friction drag and pitch-angle scattering compete with each other. They are both caused by collisions and belong to stochastic processes. There is a probability for each backward moving electron with a specific initial velocity to runaway. This probability is called runaway probability. Given the initial condition, it is defined as the number of samples, which are not thermalized and finally runaway, divided by the total sample number.

When the velocity of an electron goes below the critical velocity, it is labeled as stopped. For stopped electrons, Eq. (11) is no longer valid. The stopped electron is then treated as the thermalized background electrons. We did not consider the secondary electron emission problem here. Once an electron is stopped, we assume that it cannot become excited again.

At the end time T, the runaway probability P_r is formally defined as

$$P_r = E\left[I_r(\boldsymbol{v}_T)\right], \tag{45}$$

where $I_r(\boldsymbol{v})$ is an indicator function obeying

$$I_r(\boldsymbol{v}) = \begin{cases} 1 & v \geq v_c, \\ 0 & \text{otherwise}. \end{cases} \tag{46}$$

The species of background ions is assumed to be only proton which has $Z_i = 1$. The electric field is uniform in space as $\tilde{E} = (1, 0, 0)$, and no magnetic field exists. The initial velocity of sampling electrons is set to $\tilde{v}_{||} = 3$ and $\tilde{v}_{\perp} = 0$. For this setup, electrons can only gain their perpendicular velocities from stochastic pitch-angle scattering. As a dramatic phenomenon, the pitch-angle scattering resists the friction drag and the deceleration of external field at the same time, to save the electrons from being stopped. The time step is set to $\Delta = 0.01$. The total simulation time is $\tilde{t} = 30$.

We solve the SDE Eq. (11) using WCA2 and obtain 30 batches of solutions with 100,000 samples in each batch. The snapshots of samples in a typical batch are plotted in Fig. 5. The spread of samples in the velocity space are plotted at three different time points $\tilde{t} = 1$, $\tilde{t} = 5$, and $\tilde{t} = 20$. The horizontal and vertical axes denote the parallel velocity $\tilde{v}_{\parallel}$ and the vertical velocity $\tilde{v}_{\perp}$, respectively. Orange points denote those samples with velocities less than v_c. We freeze the state of stopped electrons in the velocity plots at

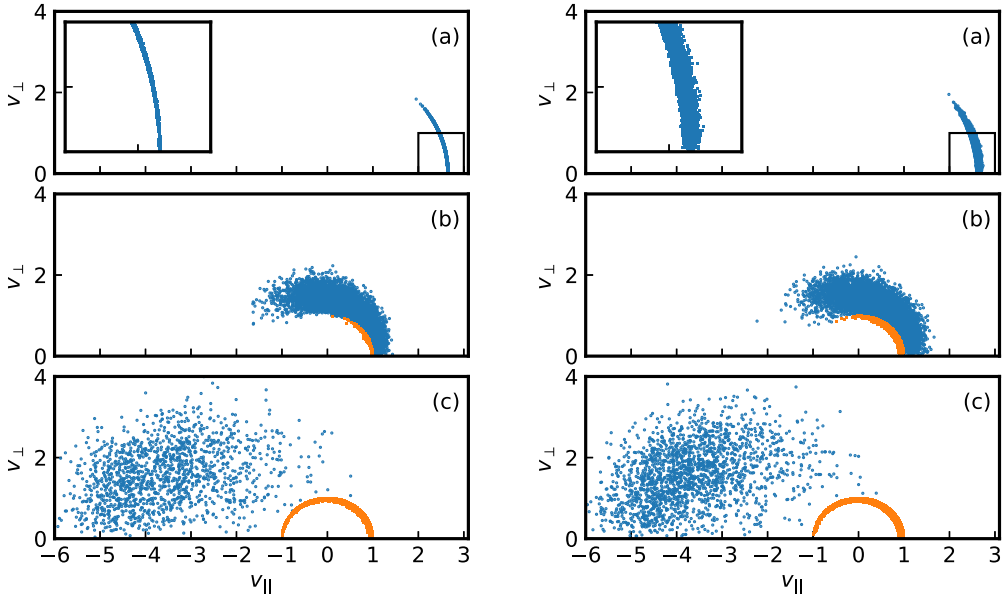

Fig. 5. Snapshots of scattering plot of 100,000 samples at 3 different time points, (a) $\tilde{t} = 1$, (b) $\tilde{t} = 5$, and (c) $\tilde{t} = 20$. For comparison, the left column is the result using WCA2, and the right column is the result using standard Euler-Maruyama algorithm. All samples start from $\tilde{v}_{||} = 3$ and $\tilde{v}_\perp = 0$. Orange points denote the samples which are decelerated to be slower than normalized Deicer velocity $\tilde{v}_c = 1$ and thus stopped. We freeze the state of stopped electrons in the velocity plots at the moment when they dropped below $\tilde{v}_c$. So the distribution of orange points display how these electrons are stopped. Because of the collisional scattering, some samples achieve sufficiently high perpendicular velocity to runaway (see blue points), while others are stopped due to the collisional friction and the external electric force. We also zoom out part of the plot in subfigures (a) to show the detailed electron distribution near $\tilde{v}_\perp = 0$. The result using Euler-Maruyama algorithm suffers a wider spread of $\tilde{v}_{||}$ due to its larger numerical error.

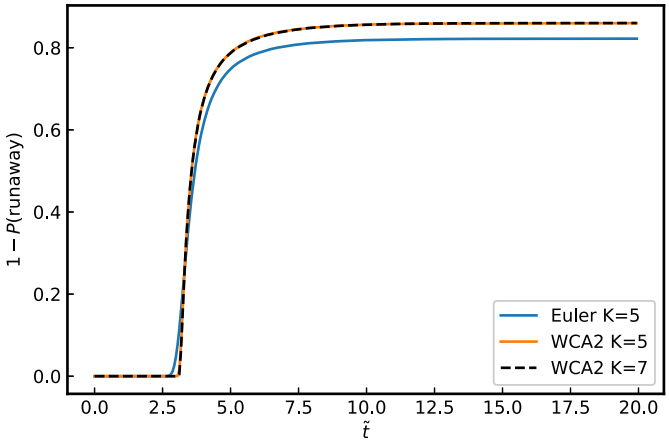

Fig. 6. The evolution of probability of being stopped, i.e., one minus the probability of backward runaway, with respect to simulation time. For comparison, simulations are carried out using WCA2 with K=5 (blue) and K=7 (black dashed), and the standard Euler-Maruyama algorithm (orange), respectively. Taking the curve calculated by WCA2 with K=7 as the benchmark solution, the poorer accuracy of Euler-Maruyama can be directly observed.

the moment when they dropped below $\tilde{v}_c$. So the distribution of orange points display how these electrons are stopped. On the contrary, blue points are samples which eventually runaway.

The runaway probability at a given time can be calculated as the fraction of unstopped electrons and is a function of time, see Fig. 6. According to Fig. 6, the runaway probability has reached a steady value far before the end time $\tilde{t} = 20$. The runaway probability defined in Eq. (45) as an expectation can be calculated using the average of the function $I_r(\boldsymbol{v}_T)$. Note that the indicator function takes the value of one only when the sample runs away, this definition is equivalent to the fraction of runaways in total samples. By calculating the fraction of runaways, the runaway probability is obtained. The trick for weakly convergent methods here is that we no longer check whether a single trajectory is correctly calculated, which is critical in strongly convergent approaches. Each single sampling trajectory is not required to be accurate enough to predict whether a single electron sample is exactly stopped or not. We only demand that the statistical results are accurate enough, which means the total count of stopped electrons is precise, although the judgment for each single sample electron might be inaccurate.

Using WCA2 and Euler-Maruyama method with time step $K = 5$, the runaway probability of electrons starting from different initial velocities is simulated. We simulate one batch for each initial velocity with $N = 100{,}000$ samples in each batch. The criterion for runaway electrons is $\tilde{v}_{||} < -2.0$. Each round of simulation runs until each sampling particle is either stopped or runaway. The parallel velocity $\tilde{v}_{||}$ ranges from -2.0 to 8.0, and the perpendicular velocity ranges from 0.0 to 5.0. The contour of runaway probability in the initial velocity space is plotted in Fig. 7. It is evident that for a given value of $v_{||}$, the runaway probability increases with $v_\perp$. And for a given value of $v_\perp$, the runaway probability increases with the absolute value of $v_{||}$. Almost all electrons having initial $v_{||}$ much larger than v_c runaway successfully. The normal forward runaway situation is also included in Fig. 7.

5. Conclusion and discussion

Starting from the Boltzmann equation of electrons with the extended Lorentz collision operator, we introduce an SDE to simulate nonrelativistic electrons under collision effects. Instead of focusing on the accuracy of sampling trajectories, we emphasize the importance and efficiency of weak convergence. A second order weakly convergent algorithm WCA2 is proposed. We also provide detailed instructions of finding benchmark solutions for weakly convergent algorithms using sufficient statistical techniques. The weakly convergent orders of different algorithms are confidently verified. The accuracy, efficiency, and economy of WCA2 are fully analyzed and revealed. The backward runaway process of electrons and backward runaway probability are simulated.

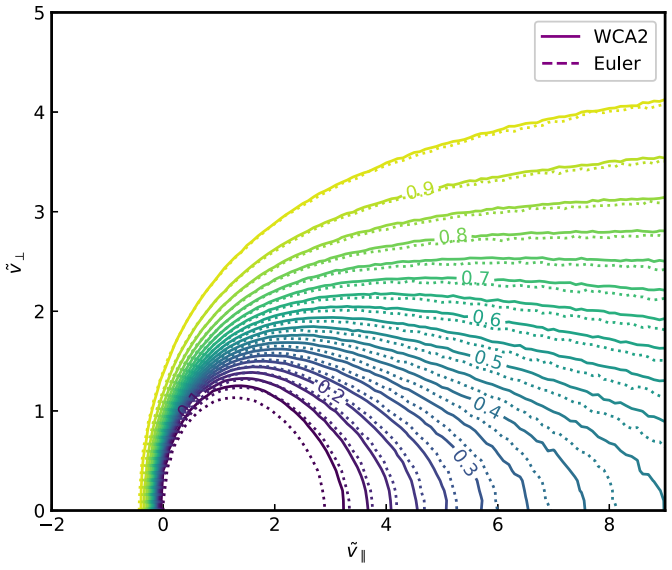

Fig. 7. The contour plots of backward runaway probability using WCA2 (solid curves) and Euler-Maruyama method (dashed curves) are simulated, respectively, with time step $K = 5$ and 100,000 samples for each initial value. The criterion for runaway electrons is $\tilde{v}_\parallel < -2.0$, which is too high to be stopped. The direction of external electric field is defined as the positive direction. Each round of simulation runs until each sampling particle is either stopped or runaway. It is obvious that the numerical error of Euler-Maruyama method has brought significant inaccuracy to the simulation results.

The SDE handled in this work is nonrelativistic. In our future research, we will consider weak convergent algorithms for relativistic SDEs. Structure-preserving algorithms, such as symplectic algorithms [38] and volume-preserving algorithms [39], will be developed for SDEs in the scope of weak convergence. Noise reduction techniques, such as importance sampling, will be used to reduce statistical errors in weakly convergent algorithms. The advantage of weak convergence will be further explored. In addition, the electron system with a source will also be considered in the framework of SDEs and corresponding numerical methods. Numerical SDE methods, including WCA2, will be further applied to study physical progress in fusion plasmas, geophysics, and space physics.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This research is supported by the National Natural Science Foundation of China (Grant Nos. 11901564, 11775222), the National Magnetic Confinement Fusion Energy Research and Development Program of China (2019YFE03090100), the Key Research Program of Frontier Sciences CAS (QYZDB-SSW-SYS004), and the Geo-Algorithmic Plasma Simulator (GAPS) Project.

References

[1] P. Langevin, Cr. Hebd. Acad. Sci. 146 (1908) 530.
[2] A. Einstein, Ann. Phys. 19 (1906) 371.
[3] K. Itô, Diffusion Processes, Wiley Online Library, 1974.
[4] F.C. Klebaner, Introduction to Stochastic Calculus with Applications, vol. 57, World Scientific, 2005.
[5] B. Øksendal, Stochastic Differential Equations, Springer, 2003.
[6] I. Karatzas, Brownian Motion and Stochastic Calculus, vol. 113, Springer, 1991.
[7] D.J. Higham, SIAM Rev. 43 (2001) 525.
[8] B.J. Albright, W. Daughton, D.S. Lemons, D. Winske, M.E. Jones, Phys. Plasmas 9 (2002) 1898.
[9] B.J. Albright, D. Winske, D.S. Lemons, W. Daughton, M.E. Jones, IEEE Trans. Plasma Sci. 31 (2003) 19.
[10] M.G. Cadjan, M.F. Ivanov, J. Plasma Phys. 61 (1999) 89.
[11] M.G. Cadjan, M.F. Ivanov, Phys. Lett. A 236 (1997) 227.
[12] F. Castejon, S. Eguilior, Plasma Phys. Control. Fusion 45 (2003) 159.
[13] I. Fernández-Gómez, J. Martín-Solís, R. Sánchez, Phys. Plasmas 19 (2012) 102504.
[14] W.M. Manheimer, M. Lampe, G. Joyce, J. Comput. Phys. 138 (1997) 563.
[15] M.S. Rosin, L.F. Ricketson, A.M. Dimits, R.E. Caflisch, B.I. Cohen, J. Comput. Phys. 274 (2014) 140.
[16] I. Shkarofsky, M. Shoucri, V. Fuchs, Comput. Phys. Commun. 71 (1992) 269.
[17] M.E. Jones, D.S. Lemons, R.J. Mason, V.A. Thomas, D. Winske, J. Comput. Phys. 123 (1996) 169.
[18] G. Mil'shtejn, Theory Probab. Appl. 19 (1975) 557.
[19] N.J. Fisch, Rev. Mod. Phys. 59 (1987) 175.
[20] C.F.F. Karney, N.J. Fisch, Phys. Fluids 29 (1986) 180.
[21] G. Zhang, D. del Castillo-Negrete, Phys. Plasmas 24 (2017) 092511.
[22] Y. Wang, J. Liu, H. Qin, Z. Yu, Y. Yao, Comput. Phys. Commun. 220 (2017) 212.
[23] P. Kloeden, Numerical Solution of Stochastic Differential Equations, vol. 23, Springer-Verlag Berlin Heidelberg, 1992.
[24] G.N. Milstein, M.V. Tretyakov, Stochastic Numerics for Mathematical Physics, Springer, 2004.
[25] P.E. Kloeden, E. Platen, J. Stat. Phys. 66 (1992) 283.
[26] H. Dreicer, Phys. Rev. 115 (1959) 238.
[27] J. Wesson, Tokamaks, vol. 149, Oxford University Press, 2011.
[28] J. Liu, Y. Wang, H. Qin, Nucl. Fusion 56 (2016) 064002.
[29] X. Guan, H. Qin, N.J. Fisch, Phys. Plasmas 17 (2010) 092502.
[30] Y. Wang, H. Qin, J. Liu, Phys. Plasmas 23 (2016) 062505.
[31] J. Liu, H. Qin, N.J. Fisch, Q. Teng, X. Wang, Phys. Plasmas 21 (2014) 393.
[32] J. Liu, Y. Wang, H. Qin, Nucl. Fusion 56 (2016) 064002.
[33] Y. Wang, H. Qin, J. Liu, Phys. Plasmas 23 (2016) 062505.
[34] J. Liu, H. Qin, Y. Wang, G. Yang, J. Zheng, Y. Yao, Y. Zheng, Z. Liu, X. Liu, arXiv preprint, arXiv:1611.02362, 2016.
[35] C. Liu, H. Qin, E. Hirvijoki, Y. Wang, J. Liu, Nucl. Fusion 58 (2018) 106018.
[36] G. Papp, M. Drevlak, T. Fulop, P. Helander, Nucl. Fusion 51 (2011).
[37] P. Helander, D.J. Sigmar, Collisional Transport in Magnetized Plasmas, Cambridge University Press, 2005.
[38] G.N. Milstein, Y.M. Repin, M.V. Tretyakov, SIAM J. Numer. Anal. 39 (2002) 2066.
[39] H. Qin, S.X. Zhang, J.Y. Xiao, J. Liu, Y.J. Sun, W.M. Tang, Phys. Plasmas 20 (2013) 084503.