An Explicitly Solvable Energy-Conserving Algorithm for Pitch-Angle Scattering in Magnetized Plasmas

Yichen Fu,[1,2] Xin Zhang,[1,2] and Hong Qin[1,2,*]

[1] *Princeton Plasma Physics Laboratory, Princeton University*

[2] *Department of Astrophysical Sciences, Princeton University*

Abstract

We develop an Explicitly Solvable Energy-Conserving (ESEC) algorithm for the Stochastic Differential Equation (SDE) describing the pitch-angle scattering process in magnetized plasmas. The Cayley transform is used to calculate both the deterministic gyromotion and stochastic scattering, affording the algorithm to be explicitly solvable and exactly energy conserving. An unusual property of the SDE for pitch-angle scattering is that its coefficients diverge at the zero velocity and do not satisfy the global Lipschitz condition. Consequently, when standard numerical methods, such as the Euler-Maruyama (EM), are applied, numerical convergence is difficult to establish. For the proposed ESEC algorithm, its energy-preserving property enables us to overcome this obstacle. We rigorously prove that the ESEC algorithm is order 1/2 strongly convergent. This result is confirmed by detailed numerical studies. For the case of pitch-angle scattering in a magnetized plasma with a constant magnetic field, the numerical solution is benchmarked against the analytical solution, and excellent agreements are found.

Keywords: Plasmas, Collisions, Stochastic differential equation, Structure-preserving algorithm, Monte Carlo method

* hongqin@princeton.edu

I. INTRODUCTION

Coulomb collision between charged particles is an important physical process in plasmas. For most systems, it is dominated by the accumulative effects of small-angle scatterings. In the Fokker-Planck (FP) equation for particle distribution functions, it is described by the Landau collision operator [1],

$$C_{ab}[f_a, f_b] = -\frac{q_a^2 q_b^2 \ln\Lambda}{8\pi\epsilon_0^2 m_a}\frac{\partial}{\partial \boldsymbol{v}}\cdot\int\left(\frac{u^2\boldsymbol{I}-\boldsymbol{u}\boldsymbol{u}}{u^3}\right)\left[\frac{f_a(\boldsymbol{v})}{m_b}\frac{\partial f_b(\boldsymbol{v}')}{\partial \boldsymbol{v}'}-\frac{f_b(\boldsymbol{v}')}{m_a}\frac{\partial f_a(\boldsymbol{v})}{\partial \boldsymbol{v}}\right]\mathrm{d}^3\boldsymbol{v}', \quad (1)$$

where C_{ab} denotes the collision experienced by particle species a colliding with particle species b, $q_{a(b)}, m_{a(b)}$ are the charge and mass of species $a(b)$, $\ln\Lambda$ is the Coulomb logarithm, $\boldsymbol{u} = \boldsymbol{v} - \boldsymbol{v}'$, and $u = |\boldsymbol{u}|$ is the Euclidean norm of $\boldsymbol{u}$. Due to the nonlinearity in distribution functions, FP equations with the Landau collision operators are difficult to solve either analytically or numerically. Fortunately, in many scenarios, simplifications are possible when the system has a well-defined separation of ordering. For the case of a low-density species propagating in a Maxwellian background species, the distribution of the background species $f_b(\boldsymbol{v})$ can be taken to be time-independent, and the collision operator C_{ab} can be written as a linear function of the distribution of the propagating species f_a. When the mass ratio m_a/m_b is small, the lead order of C_{ab} reduces to the Lorentz collision operator modeling the pitch-angle scattering process,

$$\mathcal{L}[f_a] = \frac{1}{2}\frac{\partial}{\partial \boldsymbol{v}}\cdot\left[\left(\frac{v^2\boldsymbol{I}-\boldsymbol{v}\boldsymbol{v}}{v^3}\right)\cdot\frac{\partial f_a}{\partial \boldsymbol{v}}\right], \quad (2)$$

where time has been normalized by the collision frequency ν_{ei} , and $\boldsymbol{v}$ by the thermal velocity. Pitch-angle scattering is an important physical process in plasma physics for several reasons. It is the dominant process for tokamak current drive [2–4] and plays an important role in the runaway electron dynamics [4–6]. For first-principles-based simulation studies, a reliable algorithm for pitch-angle scattering is needed. In addition, equations of a similar structure appear in the models of stochastic Landau-Lifshitz-Gilbert dynamics for magnetization [7], where similar algorithmic issues and challenges exist. Th algorithm that we develop in the present study can be applied to simulate these processes in condensed matter physics as well.

There are two major types of numerical methods to simulate the Coulomb collision, the continuum methods [8–11] and Monte-Carlo methods [12–20]. One commonly used

Monte-Carlo method is directly calculating the binary collisions between random pairs of particles [12, 13]. The particles are scattered elastically within spatially localized cells. But this method may not be computational efficient when the number of particles is large [21]. Another approach is to solve the corresponding Stochastic Differential Equations (SDE), i.e., the Langevin equations, of the FP equation [14–20]. In this approach, the SDE corresponding to the FP equation needs to be derived.

In the present study, we are concerned with the physics of pitch-angle scattering of electrons in a magnetized plasma, which can be equivalently described by the FP equation,

$$\frac{\partial f_e(\boldsymbol{v},t)}{\partial t} = -\big[\boldsymbol{v}\times\boldsymbol{B}(t)\big]\cdot\frac{\partial f_e}{\partial \boldsymbol{v}} + \mathcal{L}[f_e(\boldsymbol{v},t)], \tag{3}$$

or the corresponding Ito SDE,

$$\mathrm{d}\boldsymbol{v} = \left(\boldsymbol{v}\times\boldsymbol{B}(t) - D(v)\frac{\boldsymbol{v}}{v^2}\right)\mathrm{d}t + \sqrt{D(v)}\left(\boldsymbol{I} - \frac{\boldsymbol{v}\boldsymbol{v}}{v^2}\right)\cdot\mathrm{d}\boldsymbol{W}. \tag{4}$$

In Eq. (4), $\boldsymbol{W}(t)$ is the three dimensional Wiener process, $D(v) = 1/v$ is the pitch-angle diffusion coefficient, and $\boldsymbol{B}$ has been normalized by e, m_e, and ν_{ei}. We will focus on developing an effective algorithm for solving Eq. (4).

An unusual property of Eq. (4) is that its coefficients are not globally Lipschitz continuous and diverge at $\boldsymbol{v} = 0$. However, for a given initial condition $\boldsymbol{v}(t_0) = \boldsymbol{v}_0 \neq 0$, it can be easily proved that [see Eq. (43)] the exact solution of Eq. (4) preserves the norm of $\boldsymbol{v}$, i.e., $v = v_0 \neq 0$. Consequently, the singularity at $\boldsymbol{v} = 0$ is irrelevant for exact solutions.

One can apply either the standard Euler-Maruyama (EM) algorithm or a higher-order method, such as the Milstein algorithm [22], to Eq. (4). However, in Cartesian coordinates, these SDE methods do not typically conserve particle energy due to the truncation error, accumulation of which can lead to numerical instabilities [23] and divergent numerical solutions. Such accumulation in errors in particle energy may have disastrous consequences in physics implications.

For pitch-angle scattering, a common method to ensure energy conservation in SDE-based simulations is to cast Eq. (4) in the spherical coordinates (v, φ, θ), where v is the velocity magnitude, φ the azimuthal angle, and θ the polar (pitch) angle [17]. In practice, the $(v, \varphi, \mu = \cos\theta)$ coordinates are often used instead [19, 24, 25]. To simplify the discussion,

let's consider the unmagnetized case here. In the (v, φ, μ) coordinates, Eq. (4) is

$$\mathrm{d}\mu = -D_a(v)\mu \mathrm{d}t + \sqrt{D_a(v)(1-\mu^2)}\mathrm{d}W_\mu, \tag{5}$$

$$\mathrm{d}\varphi = \sqrt{\frac{D_a(v)}{1-\mu^2}}\mathrm{d}W_\varphi, \tag{6}$$

$$\mathrm{d}v = 0, \tag{7}$$

where $D_a = 1/v^3$ is the angular diffusion coefficient, and W_μ and W_φ are independent Wiener processes. It is obvious that only Eqs. (5) and (6) need to be solved, and the energy is automatically conserved. Here, μ should be bounded in the interval of $[-1, 1]$, but standard numerical schemes, such as the Euler-Maruyama or the Milstein scheme, do not preserve this bound [17]. When $|\mu| > 1$, the coefficient of Eq. (5) becomes imaginary and unphysical, and empirical modifications are required [26]. Furthermore, when μ is close to 1, which does happen regularly, the coefficient of Eq. (6) diverges, and the algorithm cannot correctly calculate the dynamics of φ. In a magnetized plasma with a constant, strong magnetic field, one can ignore this failure by choosing φ to be the gyrophase, and ignoring the dynamics of φ based on the gyrokinetic ordering. However, this strategy is invalid if the magnetic field is weak or depends on space and/or time, or if we are interested in the physics depending the gyrophase, such as the cyclotron waves.

Beside this shortcoming, there is another mathematical difficulty associated with the SDE for the pitch-angle scattering. In Eqs. (4), (5) and (6), both $D(v)$ and $1/\sqrt{1-\mu^2}$ are not globally Lipschitz continuous and diverge in the neighborhood of $v = 0$ and $|\mu| = 1$, respectively. However, the classical proof of convergence for the EM scheme requires that both the drift and the diffusion coefficients are globally Lipschitz continuous and have linear growth bounds [22, 27]. Many studies on the convergence of numerical schemes, especially for the EM method, have been carried out for SDEs with non-global Lipschitz coefficients [28, 29]. In these studies, the coefficients of the SDEs are not Lipschitz continuous at infinity, which is different from the case of the pitch-angle scattering. We are not aware of any existing proof of convergence for these numerical schemes when applied to Eqs. (4), (5), and (6). In practical simulations, if we are only interested in the statistical properties of the ensemble of particles when using the Monte-Carlo method, this problem could be easily overlooked. Upon closer examinations of individual sample paths, however, one may find that a small fraction of them diverges due to v drifting towards zero. This will be demonstrated numerically in

Sec. IV. Ad hoc remedies include regularizing the coefficient when v is smaller than a critical v_c [26], the value of which can only be determined empirically without a systematic approach. Other times the stochastic term is simply ignored based on other physical considerations [17] when v is small.

Recently, we proposed an Explicitly Solvable Energy-Conserving scheme (ESEC) to solve the pitch-angle scattering SDE in unmagnetized plasmas using Cartesian coordinates [20]. The ESEC scheme applies the Cayley transform [30, 31] to rotate the velocity, which guarantees the exact conservation of energy. It has been demonstrated that the ESEC method has an order 1/2 global strong error, the same as the EM scheme when applied to SDEs satisfying the global Lipschitz condition. Since the calculation is energy-preserving and performed in Cartesian coordinates, the algorithm is not subject to the problems of imaginary and divergent coefficients associated with spherical coordinates. In the present study, we generalize the scheme proposed in Ref. [20] to Eq. (4) in a magnetized plasma, and rigorously prove the strong convergence of the algorithm. We emphasize again that because the coefficients of Eq. (4) are not globally Lipschitz continuous, the standard proof of the strong convergence of the EM method is not applicable. We are not aware of any previous rigorous result on the strong convergence of numerical methods for Eq. (4). The energy-conserving nature of the ESEC method enables us to overcome the difficulty associated with the diverge at the neighborhood of $v = 0$ and establish the strong convergence.

The paper is organized as follows. In Sec. II, we construct the ESEC method for Eq. (4) and discuss its properties. In Sec. III, its strong convergence is rigorously proved. Numerical verification of the convergence rate and energy conservation properties is demonstrated in Sec. IV. Finally, the numerical solution of the pitch-angle scattering in a constant magnetic field is directly compared with the analytical solution in Sec. V.

II. THE EXPLICITLY SOLVABLE ENERGY-CONSERVING ALGORITHM FOR PITCH-ANGLE SCATTERING

As discussed in Sec. I, the physics of electrons undergoing pitch-angle scattering in magnetized plasmas can be described by the SDE (4). We propose the following Explicitly Solvable

and Energy-Conserving (ESEC) one-step method as a numerical algorithm for Eq. (4),

$$\bar{\boldsymbol{v}}_{k+1}^{\mathrm{ES}} - \boldsymbol{v}_k = \boldsymbol{v}_{k+1/2} \times \boldsymbol{B}(t_k)h + \sqrt{D_k}\left(\frac{\boldsymbol{v}_k \times \Delta\boldsymbol{W}}{v_k^2}\right) \times \boldsymbol{v}_{k+1/2}, \tag{8}$$

$$\boldsymbol{v}_{k+1/2} \coloneqq (\boldsymbol{v}_k + \bar{\boldsymbol{v}}_{k+1}^{\mathrm{ES}})/2, \tag{9}$$

where h is the step size in time, $t_{k+1} = t_k + h$, $D_k = D(v_k) = 1/v_k$, and $\Delta\boldsymbol{W} = \boldsymbol{W}(t_{k+1}) - \boldsymbol{W}(t_k) \sim \mathcal{N}(0, \boldsymbol{I}h)$ is a 3D Gaussian random variable. It generalizes the previous algorithm for pitch-angle scattering in unmagnetized plasmas [20]. Taking the dot products of $\boldsymbol{v}_{k+1/2}$ with both sides of Eq. (8), we can see that the ESEC scheme preserves the velocity norm.

The discretization of the Lorentz force is the same as the Boris algorithm [30, 32], and the discretization of the diffusion term can be regarded as a hybrid between the EM method and the midpoint method [27, 33]. Although this scheme is implicit, the value of $\bar{\boldsymbol{v}}_{k+1}^{\mathrm{ES}}$ can be explicitly solved. This is because the implicitness is linear, i.e., the right hand side of Eq. (8) depends on $\bar{\boldsymbol{v}}_{k+1}^{\mathrm{ES}}$ linearly. To solve for $\bar{\boldsymbol{v}}_{k+1}^{\mathrm{ES}}$, we make use of the hat map

$$\boldsymbol{X} = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} \longmapsto \hat{\boldsymbol{X}} \coloneqq \begin{pmatrix} 0 & -X_3 & X_2 \\ X_3 & 0 & -X_1 \\ -X_2 & X_1 & 0 \end{pmatrix} \tag{10}$$

between a vector and a 3×3 skew-symmetric matrix. The cross product between two vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$ can be written as

$$\boldsymbol{X} \times \boldsymbol{Y} = \hat{\boldsymbol{X}}\boldsymbol{Y} = \begin{pmatrix} 0 & -X_3 & X_2 \\ X_3 & 0 & -X_1 \\ -X_2 & X_1 & 0 \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix}. \tag{11}$$

It means that $\hat{\boldsymbol{X}}\boldsymbol{Y}$ is the infinitesimal 3D rotation of $\boldsymbol{Y}$ generated by $\boldsymbol{X}$. The right hand side of Eq. (8) is thus the infinitesimal 3D rotation of $2\boldsymbol{v}_{k+1/2}$ generated by

$$\boldsymbol{M}_k \coloneqq \sqrt{D_k}\left(\frac{\boldsymbol{v}_k \times \Delta\boldsymbol{W}}{2v_k^2}\right) - \frac{1}{2}\boldsymbol{B}(t_k)h. \tag{12}$$

Then the ESEC scheme defined by Eq. (8) can be explicitly solved as

$$\bar{\boldsymbol{v}}_{k+1}^{\mathrm{ES}} = \mathcal{C}(\hat{\boldsymbol{M}}_k)\boldsymbol{v}_k, \tag{13}$$

where

$$\mathcal{C}(\hat{\boldsymbol{M}}_k) \coloneqq (\boldsymbol{I} - \hat{\boldsymbol{M}}_k)^{-1}(\boldsymbol{I} + \hat{\boldsymbol{M}}_k), \tag{14}$$

is the Cayley transform of matrix $\hat{\boldsymbol{M}}_k$.

It is worth mentioning that the deterministic drift term in the Ito SDE (4) does not show up explicitly in the ESEC scheme in Eq. (8). Instead, the drift term is hidden in the implicit part of $\boldsymbol{v}_{k+1/2}$. Since the ESEC scheme is explicitly solvable, if one explicitly calculates the expression of $\boldsymbol{v}_{k+1}$ in terms of $\boldsymbol{v}_k$, the deterministic term is recovered. See Appendix B in Ref. [20] for more details of this calculation. More rigorously, in the next section we will calculate the one-step weak error of the ESEC scheme. It shows that the deterministic drift term is recovered from the implicit part with an error in the order of $\mathcal{O}(h^{3/2})$, which guarantees the strong convergence of the ESEC scheme.

As pointed out in Ref. [33], implicit methods, even when explicitly solvable, for stochastic systems with multiplicative noise may not be acceptable a priori due to its infinite expectation in the one-step approximation. In those situations, a truncation of ΔW is required [33]. Such a modification of ΔW is not necessary for the ESEC scheme. This is because the ESEC scheme preserves the velocity norm, and the expectation of the one-step approximation always exists, i.e., $\mathbb{E}|\bar{\boldsymbol{v}}_{k+1}^{\mathrm{ES}}| = \mathbb{E}|\boldsymbol{v}_0| < \infty$.

III. PROOF OF STRONG CONVERGENCE OF THE ESEC METHOD

In this section, we rigorously prove the strong convergence of the ESEC method for Eq. (4) defined on the time interval $t \in [t_0, T]$ with the initial condition $\boldsymbol{v} = \boldsymbol{v}_0$ at $t = t_0$. Denote by $\boldsymbol{v}(t; \tilde{t}, \tilde{\boldsymbol{v}})$ the analytical solution of the SDE with initial condition $\boldsymbol{v} = \tilde{\boldsymbol{v}}$ at $t = \tilde{t} \in [t_0, T]$, and by $\bar{\boldsymbol{v}}(\tilde{t} + h; \tilde{t}, \tilde{\boldsymbol{v}})$ the one-step approximation from a given value $\boldsymbol{v} = \tilde{\boldsymbol{v}}$ at $t = \tilde{t} \in [t_0, T - h]$ according to a numerical method, for example, Eq. (8). Using the one-step approximation, we recurrently construct the numerical solution $\bar{\boldsymbol{v}}(t_k; t_0, \boldsymbol{v}_0, h)$ at $t_k = t_0 + kh$ for $k = 0, \cdots, N$ with $t_N = T$ and $h = (T - t_0)/N$.

We will prove the following convergence theorem of the ESEC method.

Theorem 1. *For the SDE (4), the numerical solution $\bar{\boldsymbol{v}}(t_k; t_0, \boldsymbol{v}_0, h)$ generated by the ESEC algorithm (8) has order $1/2$ strong error in the time interval of $[t_0, T]$. More precisely, for any $\boldsymbol{v}_0 \in \mathbb{R}^3/\{0\}$, $N \in \mathbb{N}$, and $k = 0, 1, \cdots, N$, the following inequality holds,*

$$\left[\mathbb{E}|\boldsymbol{v}(t_k; t_0, \boldsymbol{v}_0) - \bar{\boldsymbol{v}}(t_k; t_0, \boldsymbol{v}_0, h)|^2\right]^{1/2} \leq K(1 + |\boldsymbol{v}_0|^2)^{1/2} h^{1/2}, \tag{15}$$

where the constant K is independent of h.

As discussed above, the coefficients in Eq. (4) is not globally Lipschitz continuous and diverge at $\boldsymbol{v} = 0$, which makes Theorem 1 difficult to prove directly. To overcome this obstacle, we consider the following modified SDE for pitch-angle scattering,

$$\mathrm{d}\boldsymbol{v}(t) = \left(\boldsymbol{v} \times \boldsymbol{B} - \alpha^2 \boldsymbol{v}\right) \mathrm{d}t + \alpha \left(\boldsymbol{I} v - \frac{\boldsymbol{v}\boldsymbol{v}}{v}\right) \cdot \mathrm{d}\boldsymbol{W}, \tag{16}$$

where α is a non-zero constant. It can be verified (shown in Appendix A) that the coefficients in Eq. (16) are all globally Lipschitz continuous and finite at $v = 0$. They are also bounded by linear growth as $\boldsymbol{v} \to \infty$. We further define a modified ESEC method for Eq. (16) as

$$\bar{\boldsymbol{v}}_{k+1}^{\mathrm{ES}} - \boldsymbol{v}_k = \boldsymbol{v}_{k+1/2} \times \boldsymbol{B}(t_k) h + \alpha \left(\frac{\boldsymbol{v}_k \times \Delta \boldsymbol{W}}{v_k}\right) \times \boldsymbol{v}_{k+1/2}. \tag{17}$$

The modified ESEC method obviously also preserves the velocity norm exactly.

Instead of proving Theorem 1 directly, we first prove in Lemma 1 the convergence of the modified ESEC method (17) for the modified SDE (16), and then show that the lemma implies Theorem 1 for any non-zero initial condition $\boldsymbol{v}_0$.

Lemma 1. *For the analytical solution $\boldsymbol{v}(t; t_0, \boldsymbol{v}_0)$ of Eq. (16) and the numerical solution $\bar{\boldsymbol{v}}(t; t_0, \boldsymbol{v}_0, h)$ generated by Eq. (17) in the time interval of $[t_0, T]$, the following results hold:*

(i) The one-step approximation $\bar{\boldsymbol{v}}(t+h; t, \tilde{\boldsymbol{v}}, h)$ has order 1 strong error, i.e., for arbitrary $t \in [t_0, T-h]$ and $\tilde{\boldsymbol{v}} \in \mathbb{R}^3$,

$$\left[\mathbb{E}|\boldsymbol{v}(t+h; t, \tilde{\boldsymbol{v}}) - \bar{\boldsymbol{v}}(t+h; t, \tilde{\boldsymbol{v}}, h)|^2\right]^{1/2} \leq K(1 + |\tilde{\boldsymbol{v}}|^2)^{1/2} h. \tag{18}$$

(ii) The one-step approximation $\bar{\boldsymbol{v}}(t+h; t, \tilde{\boldsymbol{v}}, h)$ has order 3/2 weak error, i.e., for arbitrary $t \in [t_0, T-h]$ and $\tilde{\boldsymbol{v}} \in \mathbb{R}^3$,

$$\left|\mathbb{E}\left[\boldsymbol{v}(t+h; t, \tilde{\boldsymbol{v}}) - \bar{\boldsymbol{v}}(t+h; t, \tilde{\boldsymbol{v}}, h)\right]\right| \leq K(1 + |\tilde{\boldsymbol{v}}|^2)^{1/2} h^{3/2}. \tag{19}$$

(iii) The approximation $\bar{\boldsymbol{v}}(t_k; t_0, \boldsymbol{v}_0)$ has order 1/2 strong error in the entire time interval, i.e., for any $\boldsymbol{v}_0 \in \mathbb{R}^3$ and $t_k = t_0 + hk$ $(k = 1, ..., N)$,

$$\left[\mathbb{E}|\boldsymbol{v}(t_k; t_0, \boldsymbol{v}_0) - \bar{\boldsymbol{v}}(t_k; t_0, \boldsymbol{v}_0, h)|^2\right]^{1/2} \leq K(1 + |\boldsymbol{v}_0|^2)^{1/2} h^{1/2}. \tag{20}$$

Here, the constant K is independent of h and $\boldsymbol{v}_0$.

Proof. The proof starts from the Ito-Taylor expansion (a.k.a. the Wagner-Platen expansion) of Eq. (16), based on which (i) and (ii) will be proved. We than prove (iii) by applying

Theorem 1.1 in Ref. [27], which states that for a method with order p_1 strong error and order p_2 weak error in one step, if $p_1 \geq 1/2$ and $p_2 \geq p_1 + 1/2$, then the numerical scheme has order $(p_1 - 1/2)$ strong error on the entire time interval.

We rewrite Eq. (16) using Cartesian indices,

$$\mathrm{d}v_i = \mu_i \mathrm{d}t + \sigma_{ir}\mathrm{d}W_r, \qquad \mu_i = \epsilon_{ijk}v_j B_k - \alpha^2 v_i, \quad \sigma_{ir} = \alpha\left(\delta_{ir}v - v_i v_r/v\right) = \alpha\epsilon_{ijk}\epsilon_{jlr}v_l v_k/v, \tag{21}$$

where δ_{ij} is the Kronecker delta, ϵ_{ijk} is the Levi-Civita symbol, and repeated indices are summed over.

When $\boldsymbol{v}(t)$ is a solution of Eq. (21), by Ito formula, any sufficiently smooth function $f(t, \boldsymbol{v})$ can be written as

$$f(t, \boldsymbol{v}(t)) = f(t_0, \boldsymbol{v}_0) + \int_{t_0}^{t} \Lambda_r f(t_1, \boldsymbol{v}(t_1))\mathrm{d}W_r(t_1) + \int_{t_0}^{t} L f(t_1, \boldsymbol{v}(t_1))\mathrm{d}t_1, \tag{22}$$

$$\Lambda_r := \sigma_{ir}\frac{\partial}{\partial x_i}, \quad L := \frac{\partial}{\partial t} + \mu_i\frac{\partial}{\partial x_i} + \frac{1}{2}\sigma_{ir}\sigma_{jr}\frac{\partial^2}{\partial x_i \partial x_j}. \tag{23}$$

Consecutively applying Eq. (22) to the coefficients inside the integral leads to the Ito-Taylor expansion with integral type remainders.

The Ito-Taylor expansion for $v_i(t)$ itself is

$$v_i(t+h) = v_i(t) + \alpha\epsilon_{ijk}v_j(t)B_k(t)h - \alpha^2 v_i(t)h + \alpha\epsilon_{ijk}\epsilon_{jlr}v_l(t)v_k(t)\Delta W_r/v(t) + \sum_{\lambda=1}^{4}\rho_{i,\lambda}, \tag{24}$$

where $\Delta W_r = \int_t^{t+h}\mathrm{d}W_r \sim \mathcal{N}(0, h)$. The first three terms of the expansion is the Euler-Maruyama approximation of Eq. (16); the forth term is the four integral remainders,

$$\begin{aligned} \rho_{i,1} &= \int_t^{t+h}\int_t^{t_1} L\mu_i(t_2)\mathrm{d}t_2\mathrm{d}t_1, \quad \rho_{i,2} = \int_t^{t+h}\int_t^{t_1} \Lambda_m\mu_i(t_2)\mathrm{d}W_m(t_2)\mathrm{d}t_1, \\ \rho_{i,3} &= \int_t^{t+h}\int_t^{t_1} L\sigma_{ir}(t_2)\mathrm{d}t_2\mathrm{d}W_r(t_1), \quad \rho_{i,4} = \int_t^{t+h}\int_t^{t_1} \Lambda_m\sigma_{ir}(t_2)\mathrm{d}W_m(t_2)\mathrm{d}W_r(t_1), \end{aligned} \tag{25}$$

where $\mu_i(t) \equiv \mu_i(\boldsymbol{v}(t), t)$ and $\sigma_{ir}(t) \equiv \sigma_{ir}(\boldsymbol{v}(t), t)$.

Since μ_i and σ_{ir} are well-behaved, it can be verified that functions $L\mu_i$, $\Lambda_m\mu_i$, $L\sigma_{ir}$, and $\Lambda_m\sigma_{ir}$ all have linear growth bound and are finite when $v \to 0$. Thus, the estimation of the magnitude of all the remainders can be immediately found using Lemma 2.2 in Ref. [27].

In the index notation, we denote by $|v_i| = (v_x^2 + v_y^2 + v_z^2)^{1/2}$ the Euclidean norm of the vector $\boldsymbol{v}$. The following mean-squared estimation holds,

$$\mathbb{E}|\rho_{i,\lambda}|^2 \le K(1 + |v_i(t)|^2)h^{p_\lambda}, \tag{26}$$

where $p_1 = 4$, $p_2 = p_3 = 3$, $p_4 = 2$, and K is a constant independent of h and $v_i(t)$. In addition, since $\rho_{i,2}$, $\rho_{i,3}$, $\rho_{i,4}$ are Ito integrals, we have

$$\mathbb{E}\rho_{i,2} = \mathbb{E}\rho_{i,3} = \mathbb{E}\rho_{i,4} = 0. \tag{27}$$

(i) One-step strong error.

Denote the one-step approximation by $\bar{\boldsymbol{v}}(t+h) \equiv \bar{\boldsymbol{v}}(t+h; t, \boldsymbol{v}(t))$, and let $\Delta W_r = \xi_r\sqrt{h}$ where $\xi_r \sim \mathcal{N}(0,1)$. The modified ESEC method Eq. (17) can be written using the Cartesian indices as

$$\begin{aligned}
\bar{v}_i(t+h) - v_i(t) =& \epsilon_{ijk}\frac{v_j(t) + \bar{v}_j(t+h)}{2}B_k(t)h + \alpha\epsilon_{ijk}\epsilon_{jlr}v_l(t)\sqrt{h}\xi_r\frac{v_k(t) + \bar{v}_k(t+h)}{2v(t)}, \\
=& \epsilon_{ijk}M_j\sqrt{h}\left[v_k(t) + \bar{v}_k(t+h)\right], \qquad (28) \\
M_j :=& \alpha\frac{\epsilon_{jlr}v_l(t)\xi_r}{2v(t)} - \frac{1}{2}B_j(t)\sqrt{h}. \qquad (29)
\end{aligned}$$

The difference between the one-step approximation (28) and the exact solution (24) is

$$\begin{aligned}
\bar{v}_i(t+h) - v_i(t+h) &= \epsilon_{ijk}M_j\sqrt{h}\big[\bar{v}_k(t+h) - v_k(t)\big] + \alpha^2 v_i(t)h - \sum_\lambda \rho_{i,\lambda}, \\
&= \epsilon_{ijk}\epsilon_{kpq}M_jM_ph\left[v_q(t) + \bar{v}_q(t+h)\right] + \alpha^2 v_i(t)h - \sum_\lambda \rho_{i,\lambda}, \qquad (30)
\end{aligned}$$

where in the second line, we have used Eq. (28) again to replace the term inside the square bracket. The one-step strong error is given by $\mathbb{E}|\bar{v}_i(t+h) - v_i(t+h)|^2$. Using the inequality $|X_i + Y_i|^2 \le 2|X_i|^2 + 2|Y_i|^2$, we can estimate the mean-square bound of the RHS of Eq. (30) term by term. Because $|\epsilon_{ijk}X_jY_k| \le |X_j| \cdot |Y_k|$, the mean square of the first term is

$$\begin{aligned}
R_{s,1} :=& \mathbb{E}\big|\epsilon_{ijk}\epsilon_{kpq}M_jM_ph\left[v_q(t) + \bar{v}_q(t+h)\right]\big|^2 \le h^2\mathbb{E}\big[|M_j|^4|v_q(t) + \bar{v}_q(t+h)|^2\big] \qquad (31) \\
&\le 4h^2|v_q(t)|^2\mathbb{E}\big[|M_j|^4\big], \qquad (32)
\end{aligned}$$

where use is made of $|v_q(t) + \bar{v}_q(t+h)| \le |v_q(t)| + |\bar{v}_q(t+h)| = 2|v_q(t)|$ in the second line. The bound for $|M_j|$ can be calculated as follows.

$$\left|\alpha\frac{\epsilon_{jlr}v_l(t)\xi_r}{2v(t)} - \frac{1}{2}B_j(t)\sqrt{h}\right| \le \frac{\alpha}{2}\left|\frac{\epsilon_{jlr}v_l(t)\xi_r}{v(t)}\right| + \frac{\sqrt{h}}{2}|B_j(t)| \le \frac{1}{2}(\alpha|\xi_r| + \sqrt{h}|B_j(t)|), \tag{33}$$

where we have used $|v_l(t)/v(t)| = 1$. Therefore,

$$\begin{aligned} R_{s,1} &\leq \frac{h^2}{4}|v_q(t)|^2\mathbb{E}\left[(\alpha|\xi_r| + \sqrt{h}|B_j(t)|)^4\right] \\ &= \frac{h^2}{4}|v_q(t)|^2\left[\alpha^4\mathbb{E}|\xi_r|^4 + \mathcal{O}(\sqrt{h})\right] \leq K\,|v_q(t)|^2h^2. \end{aligned} \tag{34}$$

Here, the boundedness of $B(t)$ has been used and K is a constant independent of h and $v_q(t)$.

The estimation of the mean square of the second term in Eq. (30) is straightforward; the estimation of that of the third term is given by Eq. (26). Putting the estimations of all three terms together proves Eq. (18), i.e., the one-step strong error of the modified ESEC scheme is at least order 1.

(ii) One-step weak error.

Next, we calculate the one-step weak error $|\mathbb{E}[\bar{v}_i(t+h) - v_i(t+h)]|$. Plugging the expression of $\bar{v}_i(t+h)$ from Eq. (28) into the RHS of Eq. (30) again, we get

$$\begin{aligned} \bar{v}_i(t+h) - v_i(t+h) = \quad & 2\epsilon_{ijk}\epsilon_{kpq}M_jM_phv_q(t) + \alpha^2v_i(t)h \\ & + \epsilon_{ijk}\epsilon_{kpq}\epsilon_{qmn}M_jM_pM_mh^{3/2}[v_n(t) + \bar{v}_n(t+h)] \\ & - \sum_\lambda \rho_{i,\lambda}. \end{aligned} \tag{35}$$

With the intention to apply the inequality $|\mathbb{E}(X+Y)| \leq |\mathbb{E}(X)| + |\mathbb{E}(Y)|$, we estimate the mean of the RHS of Eq. (35) line by line. Since $|\mathbb{E}X| \leq \mathbb{E}|X|$, the bound of the mean for the second line is

$$\begin{aligned} R_{w,2} &:= \left|\mathbb{E}\left[\epsilon_{ijk}\epsilon_{kpq}\epsilon_{qmn}M_jM_pM_mh^{3/2}[v_n(t) + \bar{v}_n(t+h)]\right]\right| \\ &\leq \mathbb{E}\left|\epsilon_{ijk}\epsilon_{kpq}\epsilon_{qmn}M_jM_pM_mh^{3/2}[v_n(t) + \bar{v}_n(t+h)]\right| \\ &\leq h^{3/2}\mathbb{E}\left[|M_j|^3|v_n(t) + \bar{v}_n(t+h)|\right] \leq \frac{h^{3/2}}{4}|v_n(t)|\mathbb{E}\left[(\alpha|\xi_r| + \sqrt{h}|B_j(t)|)^3\right] \\ &= \frac{h^{3/2}}{4}|v_n(t)|\left[\alpha^3\mathbb{E}|\xi_r|^3 + \mathcal{O}(\sqrt{h})\right] \leq K|v_n(t)|h^{3/2}. \end{aligned} \tag{36}$$

In the third line, $\rho_{i,2}$, $\rho_{i,3}$, $\rho_{i,4}$ vanish when taking expectation. The bound of the mean for $\rho_{i,1}$ can be estimated by the Cauchy-Schwarz inequality,

$$|\mathbb{E}\rho_{i,1}| \leq \mathbb{E}|\rho_{i,1}| \leq \sqrt{\mathbb{E}|\rho_{i,1}|^2} \leq K(1 + |v_i(t)|^2)^{1/2}h^2, \tag{37}$$

For the first line, plugging in the definition of M_j and after some algebra, we have

$$\begin{aligned}
R_{w,1} :=& \mathbb{E}L_1 \\
L_1 :=& 2\epsilon_{ijk}\epsilon_{kpq}M_jM_phv_q(t) + \alpha^2 v_i(t)h \\
=& 2\alpha^2\epsilon_{ijk}\epsilon_{kpq}\frac{\epsilon_{jlr}v_l(t)\xi_r}{2v(t)}\frac{\epsilon_{pmn}v_m(t)\xi_m}{2v(t)}v_q(t)h + \alpha^2 v_i(t)h \\
&- \alpha\frac{h^{3/2}}{2v(t)}\epsilon_{ijk}\epsilon_{kpq}\Big[\epsilon_{jlr}v_l(t)\xi_r B_p(t) + \epsilon_{pmn}v_m(t)\xi_n B_j(t)\Big]v_q(t) \\
&+ \frac{h^2}{2}\epsilon_{ijk}\epsilon_{kpq}B_j(t)B_p(t)v_q(t).
\end{aligned} \tag{38}$$

In Eq. (38), the third line vanishes under expectation because $\mathbb{E}\xi_r = 0$; the fourth line is on the order of $\mathcal{O}(h^2)$. To calculate the mean of the first term in the second line, notice that $\mathbb{E}(\xi_i\xi_j) = \delta_{ij}$, and we have

$$\mathbb{E}\left[2\alpha^2(\epsilon_{ijk}\epsilon_{jlr})(\epsilon_{kpq}\epsilon_{pmn})\frac{v_l(t)\xi_r}{2v(t)}\frac{v_m(t)\xi_n}{2v(t)}v_q(t)h\right] \tag{39}$$

$$=\alpha^2(\delta_{kl}\delta_{ir} - \delta_{kr}\delta_{il})(\delta_{qm}\delta_{kn} - \delta_{qn}\delta_{km})\left[\frac{v_l(t)v_m(t)v_q(t)}{2v^2(t)}\right]\mathbb{E}(\xi_r\xi_n)h \tag{40}$$

$$= -\alpha^2 v_i(t)h, \tag{41}$$

where use is made of $\delta_{ii} = 3$. Notice that the expectation, $-\alpha^2 v_i(t)h$, is exactly the deterministic drift in the Ito-Taylor expansion in Eq. (24) and cancels out the second term in the second line of Eq. (38). Hence, the second line in Eq. (38) vanishes under expectation, and we have

$$R_{w,1} = \mathcal{O}(h^2). \tag{42}$$

Combining all the estimations for terms on the RHS of Eq. (35) proves Eq. (19), i.e., the one-step weak error of the modified ESEC method is at least order 3/2.

(iii) Strong convergence of the ESEC method.

We have proved that the strong and weak errors in one step between the modified ESEC method in Eq. (17) and the solution of Ito SDE Eq. (16) are on order 1 and order 3/2, respectively. In addition, the coefficients of Eq. (16) are globally Lipschitz continuous (see Appendix B) and have linear growth bound. Based on Theorem 1.1 in Ref. [27], the numerical solutions constructed by Eq. (17) converge to exact solutions of Eq. (16) with order 1/2 strong error on the entire time interval.

This completes the proof of Lemma 1. □

We now prove Theorem 1.

Proof. Notice that for both Eq. (4) and Eq. (16), the velocity norm is preserved by the solution, which can be verified by applying the Ito formula [22] to v,

$$\mathrm{d}v = \left[\frac{\partial v}{\partial t} + \boldsymbol{\mu}\cdot\frac{\partial v}{\partial \boldsymbol{v}} + \frac{1}{2}\,\mathrm{Tr}\left(\boldsymbol{\sigma}^{\mathrm{T}}\frac{\partial^2 v}{\partial \boldsymbol{v}\partial \boldsymbol{v}}\boldsymbol{\sigma}\right)\right]\mathrm{d}t + \frac{\partial v}{\partial \boldsymbol{v}}\cdot\boldsymbol{\sigma}\cdot\mathrm{d}\boldsymbol{W} = 0, \tag{43}$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ are the drift and diffusion coefficients in either Eq. (4) or Eq. (16). Thus any analytical solution $\boldsymbol{v}(t)$ of Eq. (4) with initial condition $\boldsymbol{v}(t_0) = \boldsymbol{v}_0 \neq 0$ is also the solution of Eq. (16) with the same initial condition if we choose the constant $\alpha = \sqrt{D(v_0)}/v_0$.

On the other hand, both the ESEC method given by Eq. (8) for Eq. (4) and the modified ESEC method by Eq. (17) for Eq. (16) preserve the velocity norm exactly. For any given initial condition $\boldsymbol{v}_0 \neq 0$, the numerical solution generated by the ESEC method (8) for Eq. (4) is exactly the same as that by the modified ESEC method (17) for Eq. (16), if we choose $\alpha = \sqrt{D(v_0)}/v_0$.

According to Lemma 1, the modified ESEC method (17) generates numerical solutions converging to the exact solutions of Eq. (16) with order 1/2 strong error. Combining these facts, we conclude that the ESEC method (8) generates numerical solutions converging to the exact solutions of Eq. (4) with order 1/2 strong error.

This completes the proof of Theorem 1 □

Two comments are in order. First, in Lemma 1, the constant K is independent of initial condition $\boldsymbol{v}_0$, but depends on the constant α. In order to make Eq. (4) and Eq. (16) equivalent, the choice of α depends on the initial condition, which makes the constant K in Theorem 1 dependent on initial conditions. Second, the proof relies on the fact that the velocity norm v is conserved for both analytical solutions and numerical solutions. For the standard EM method implemented in Cartesian coordinates, the numerical solution does not preserve the velocity norm, and as a consequence the convergence cannot be established. In Sec. V A, divergent sample paths of the EM method will be demonstrated.

IV. NUMERICAL VERFICATION OF CONVERGENCE

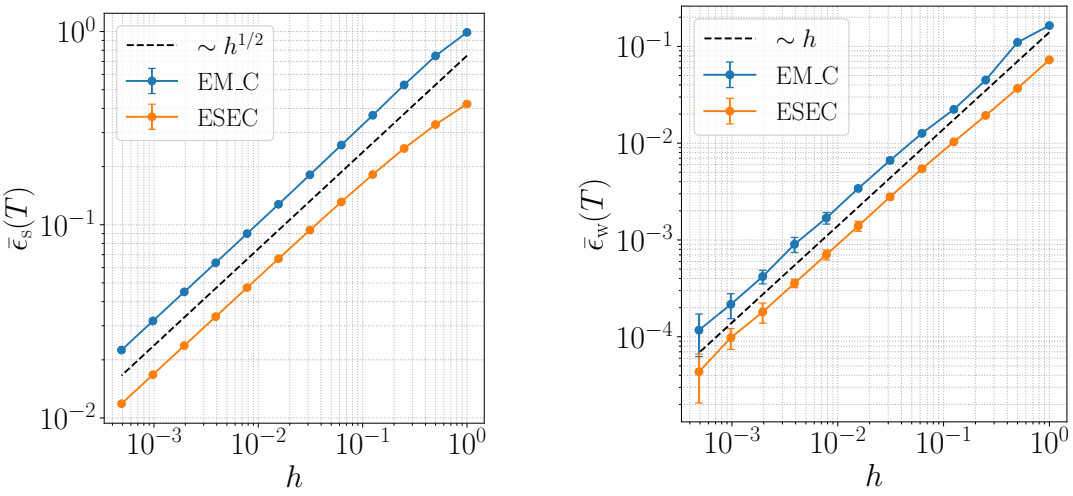
A. Strong and weak convergence

In this and the following sections, we compare the ESEC scheme and the EM scheme implemented in Cartesian coordinates (labeled by EM_C). Since the analytical solution of Eq. (4) is unknown, it is not feasible to calculate the strong and weak errors according to Eqs. (18) and (19). Thus we define the following relative errors between different step sizes h_l and h_{l+1},

$$\bar{\epsilon}_{\mathrm{s}}(t; t_0, \boldsymbol{v}_0, h_l) := \left[\mathbb{E}|\bar{\boldsymbol{v}}(t; t_0, \boldsymbol{v}_0, h_{l+1}) - \bar{\boldsymbol{v}}(t; t_0, \boldsymbol{v}_0, h_l)|^2\right]^{1/2}, \tag{44}$$

$$\bar{\epsilon}_{\mathrm{w}}(t; t_0, \boldsymbol{v}_0, h_l) := \left|\mathbb{E}\left[\bar{\boldsymbol{v}}(t; t_0, \boldsymbol{v}_0, h_{l+1}) - \bar{\boldsymbol{v}}(t; t_0, \boldsymbol{v}_0, h_l)\right]\right|. \tag{45}$$

It is easy to prove that for a numerical scheme with order p_{w} weak error and order p_{s} strong error, $\bar{\epsilon}_{\mathrm{s/w}} \sim \mathcal{O}(h_l^{p_{\mathrm{s/w}}})$ as long as $h_{l+1} < h_l$ (see Appendix B).



(a) Strong errors (b) Weak errors

Figure 1: Strong (a) and weak (b) errors for EM_C and ESEC schemes. Reference lines with order 1/2 and 1 are shown as dashed lines. The initial condition is $\boldsymbol{v}_0 = (0, 0, 1)$ and the magnetic field is $\boldsymbol{B} = (0, 0, 1)$. The error bars in strong errors are too small to be visible. The sample size is 5×10^6.

With the definition and fact above, the relative strong and weak errors at the end time $t = T = 1$ for $h_l = 2^{-l}$ ($l = 1 \cdots 12$) are plotted in Fig. 1. Figure 1a shows that both the ESEC and EM_C methods have order 1/2 strong error, but the strong convergence for the

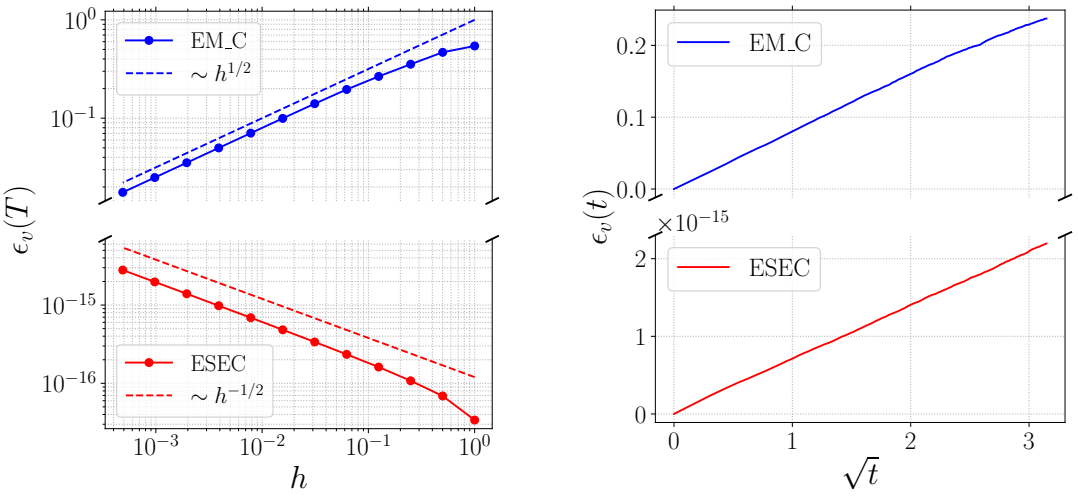
EM_C method cannot be rigorously established due to the violation of the global Lipschitz condition. Although we did not establish the weak order of the two schemes for Eq. (4) in Sec. III, Fig. 1b shows that the both schemes have order 1 weak error in the entire time interval. In addition, the ESEC scheme has smaller scale constants for both strong and weak errors than the EM_C scheme.

B. Conservation of the velocity norm

To compare the conservation of the velocity norm v, we define the following strong error of velocity magnitude,

$$\epsilon_v(t; t_0, \boldsymbol{v}_0, h) := \left[\mathbb{E}|\bar{v}(t; t_0, \boldsymbol{v}_0, h) - v_0|^2\right]^{1/2}, \tag{46}$$

where $\bar{v}(t; t_0, \boldsymbol{v}_0, h)$ is the numerically calculated velocity norm at time t with initial condition $\boldsymbol{v}_0$ at $t = t_0$ and step size h.



(a) Strong errors of v as a function of the time-step size h at the fixed end time t.

(b) Strong errors of v as a function of $\sqrt{t}$ with a fixed time-step size h.

Figure 2: The strong errors of velocity norm. The vertical axes are separated into two parts because the error of the two schemes are different by more than 10 orders of magnitude. The initial condition is $\boldsymbol{v}_0 = (0, 0, 1)$ and the magnetic field is $\boldsymbol{B} = (0, 0, 1)$.

Figure 2a shows the strong errors of the velocity norm ϵ_v at the fixed end time $t = T = 3$ as a function of $h \in [2^{-12}, 2^{-1}]$. For the EM_C method, $\epsilon_v \sim h^{1/2}$, consistent with its usual

strong error. For the ESEC method, due to its energy-preserving property, the magnitude of ϵ_v is the same as the machine precision which is 10^{-15}. The interesting result that $\epsilon_v \sim h^{-1/2}$ can be understood as follows. Because of machine error, the value of $\delta v = v - v_0$ performs a random walk at each time step. For a calculation with total number time steps $L = T/h$, we have $|\delta v(T)| \sim L^{1/2} \sim h^{-1/2}$. Shown in Fig. 2b are the strong errors of the velocity norm with a fixed time step $h = 0.01$ for $t \in [0, 10]$. We can see that for both schemes $\epsilon_v \sim h^{1/2}$. This is because the errors at different time steps are independent and the accumulated errors behave like random walks. The difference between the two methods is that the error for the EM_C method is its truncation error, but that for the ESEC method is the machine error.

In Fig. 2b, we observe that at $t = 10$, the strong error of the velocity norm for the EM_C method is more than 20% of its initial value. This type of error in the velocity norm is deleterious. For some sample paths, it can push v towards zero, where the drift and diffusion coefficients are singular, which in turn leads to the divergence of the sample paths. Such sample paths will be numerically demonstrated in the next section. To avoid this divergence, one method is to regularize the coefficients of the SDE. For example, one can modify the coefficients when v is smaller than a chosen critical value v_c [26] (see Appendix C 1 for more details).

Another choice is to switch to spherical coordinates. Since the velocity magnitude is a constant of motion according to Eq. (6), it conserved exactly without machine error. In this sense, the EM method implemented in spherical coordinates has an even better energy-conservation property than the ESEC method, if one is not concerned with the singularity at $\mu = \pm 1$ for Eq. (6). In practice, regularization techniques in spherical coordinates are still necessary for the dynamics of azimuthal angle described by Eq. (6) (see Appendix C 2 for more details). In comparison, modifications of the governing equations for the purpose of avoiding numerical singularity are not necessary for the ESEC method.

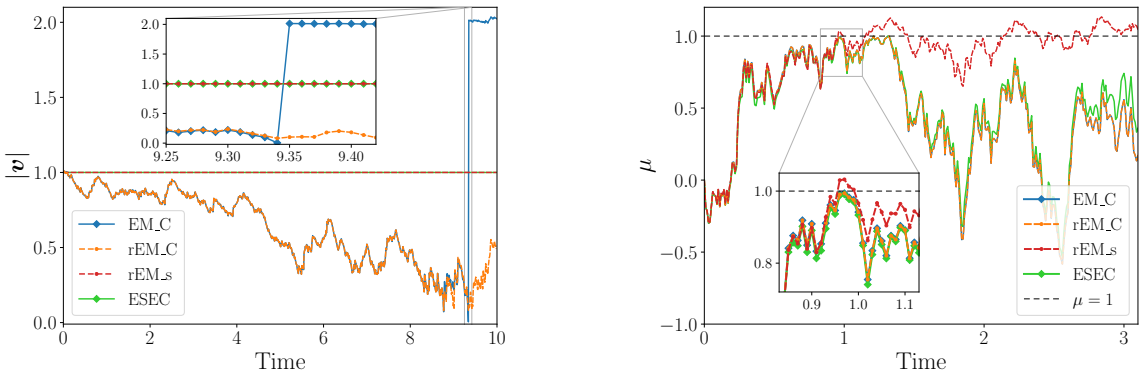
V. PITCH-ANGLE SCATTERING IN A CONSTANT MAGNETIC FIELD

In this section we benchmark our algorithm against an analytical solvable problem. Consider the electron-ion collisions of a magnetized plasma in a constant magnetic field. The electrons have a uniform initial velocity $\boldsymbol{v}_0$ and the ions are cold and stationary. In the limit of $m_e/m_i \to 0$, the electron trajectory can be described by Ito SDE (4), and the

time evolution of the electron distribution function f_e is described by the FP equation (3). For comparison, the problem will be calculated using the ESEC algorithm, the normal EM algorithm in Cartesian coordinates (refer to 'EM_C'), the regularized EM algorithm in Cartesian coordinates (refer to 'rEM_C', see Appendix C 1), and the regularized EM algorithm in spherical coordinates (refer to 'rEM_s', see Appendix C 2).

A. Single-electron trajectories

Firstly, we compare the electron trajectory calculated by different algorithms with the same underlying Wiener process. The method to transform the same Wiener process to different coordinates is discussed in Appendix C 3. Sample paths of v and μ calculated with different algorithms are shown in Fig. 3.



(a) Sample paths of $v(t)$.

(b) Sample paths of $\mu(t)$.

Figure 3: Sample paths calculated with the standard EM in Cartesian coordinates (blue), the regularized EM in Cartesian coordinates (orange), the regularized EM in spherical coordinates (red), and the ESEC algorithm (green). In Fig. (a), the results calculated by 'ESEC' and 'rEM_s' algorithm coincide with each other since they both conserve the velocity magnitude. The dashed line in (b) marks $\mu = 1$.

Fig. 3a shows the sample paths of $v(t)$ from four algorithms. For the EM_C algorithm, the sample path is divergent when $v \to 0$. At $t = 9.34$, v is smaller than 0.005 due to the accumulation of truncation errors over many time steps, but one time step later, v jumps to a value larger than 2. After regularization, the sample path no longer diverges, but

the velocity magnitude still deviates from its original value. On the other hand, for the sample paths given by the ESEC algorithm or the EM algorithm in spherical coordinates, the velocity magnitudes remain constant throughout the whole calculation.

Fig. 3b shows the sample paths of $\mu(t)$. By definition, μ is always less than 1, i.e., $|\mu| \leq 1$. If the numerical calculation is done in Cartesian coordinates, this condition is automatically satisfied, which is indeed the case for the EM_C, rEM_C, and ESEC algorithms in Fig. 3b. However, if the sample path is calculated in the spherical coordinates, this condition is not guaranteed. Due to the intrinsic singularity of the spherical coordinate system, unphysical behaviors may occur at $|\mu| \to 1$. In Fig. 3b, the sample path given by rEM_s deviates from those by the other three algorithms around $t = 0.95$ when μ is close to 1. Then it exceeds 1 and becomes drastically different from the others. It is worth mention that the behavior of μ calculated in spherical coordinates when $|\mu| \to 1$ depends on the implementation and regularization techniques. For example, one can avoid $|\mu| > 1$ by directly using the polar angle θ instead of $\cos\theta$ [17, 34], and the SDE becomes $\mathrm{d}\theta = \frac{1}{2}D_a \cot\theta\,\mathrm{d}t + \sqrt{D_a}\mathrm{d}W_\theta$. However, the singularities at $\theta = 0, \pi$ still exist because the coefficient of $\mathrm{d}t$ in the SDE is proportional to $\cot\theta$. Detailed study comparing different implementation and regularization methods in spherical coordinates is beyond the scope of this paper.

B. Electron distribution functions

From the single-electron trajectories, we see that using the EM method to calculate the pitch-angle scattering brings about a few challenges. In Cartesian coordinates, the EM method does not conserve the energy and the governing differential equation does not satisfied the global Lipschitz condition due to the singularity at $v = 0$. If one adopts the EM method in spherical coordinates, then the energy is a constant of motion according to Eq. (7), but $\mu = \pm 1$ now become singularities for Eq. (6). This issue can be clearly demonstrated when these different algorithms are applied to calculate the distribution function and compared with the analytical solution.

Assume that $\boldsymbol{B} = (0, 0, B)$ is in the $\hat{z}$ direction. In the velocity spherical coordinates

(v, θ, φ), Eq. (3) becomes

$$\frac{\partial}{\partial t} f_e(\boldsymbol{v}, t) = B \frac{\partial f_e}{\partial \varphi} + \mathcal{L}[f_e], \tag{47}$$

$$\mathcal{L} = \frac{1}{2v^3} \left[\frac{1}{\sin\theta} \frac{\partial}{\partial \theta} \left(\sin\theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial \varphi^2} \right]. \tag{48}$$

The initial condition is $f_{e0} = \delta(\boldsymbol{v} - \boldsymbol{v}_0)$, $\boldsymbol{v}_0 = (v_0, \theta_0, \varphi_0)$. Since Eq. (47) is a linear equation in the spherical coordinates, it can be solved in terms of the spherical harmonics [20]. The solution of the initial value problem is

$$f_e(\boldsymbol{v}, t) = \frac{\delta(v - v_0)}{v_0^2 \sin\theta_0} \sum_{l=0}^{+\infty} \sum_{m=-l}^{l} \overline{Y_l^m(\theta_0, \varphi_0)} Y_l^m(\theta, \varphi + Bt) \exp\left[-\frac{1}{2v_0^3} l(l+1) t \right], \tag{49}$$

where $\bar{Y}_l^m$ is the complex conjugate of Y_l^m.

From the analytical solution, we can draw several physical insights. Firstly, the coefficients of all the spherical harmonics decay exponentially in time, except for the $l = 0$ component. In the limit of $t \to \infty$, f_e becomes uniform on the $v = v_0$ sphere. Secondly, the magnetic field only influences the dynamics of the azimuthal angle. The azimuthally averaged quantities, such as those of the parallel and perpendicular velocities, evolve in the exactly same way as in the un-magnetized case [20].

We now solve the same problem numerically using the ESEC algorithm and the regularized EM algorithm in both Cartesian and spherical coordinates. Then we benchmark the numerical solution against the analytical solution (49). As an example, we choose $\boldsymbol{v}_0 = (1, 0, 0)$ and $\boldsymbol{B} = (0, 0, 1)$. In the spherical coordinates, the initial conditions for Eq. (4) are $v_0 = 1$, $\mu_0 = 0$, and $\varphi_0 = 0$. The distribution function is numerically calculated from 10^5 sample paths with time-step size $h = 0.01$. The distributions in spherical coordinates at different times are shown by the histogram in Fig. 4, while the analytical solutions are shown by the solid lines. The azimuthal angular distributions calculated by all three algorithms are almost identical, while the other two distributions show some interesting differences. The distribution over μ spreads out from $\mu_0 = 0$ and gradually becomes uniform between ± 1. However, in Fig. 4h, the distribution calculated in spherical coordinates gradually leaks out from the boundary $|\mu| = 1$, which leads to unphysical results. In comparison, the distributions calculated in Cartesian coordinates do not have this problem. On the other hand, for the distribution in velocity magnitude, the ESEC algorithm and the EM algorithm in spherical coordinates retain the initial delta function at $v = 1$ for all time, which is the

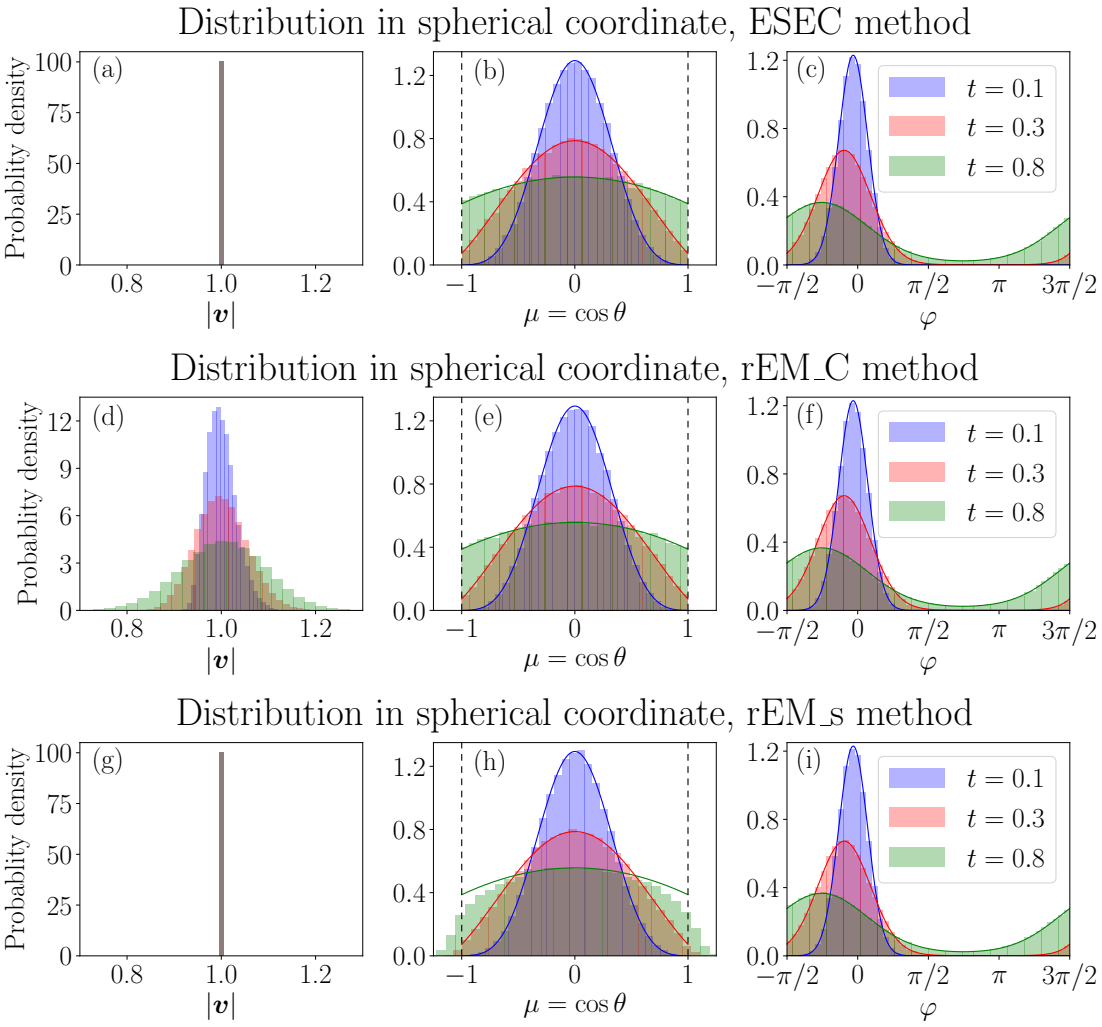


Figure 4: Distributions in v, $\mu = \cos\theta$ and φ, calculated by the ESEC algorithm, the EM algorithm in Cartesian coordinates (rEM_C) and spherical coordinates (rEM_s). The histograms show the distributions obtained numerically by the three algorithms at different times. The solid lines are the analytical solution obtained at each time by integrating Eq. (49) and adding up all components with $l \leq 50$. The dashed lines mark $\mu = \pm 1$.

same as the analytical solution. In contrast, although the distribution calculated by the EM algorithm in Cartesian coordinates is centered around $v = 1$, it numerically diffuses over time, which reflects the fact that the EM algorithm in Cartesian coordinates does not conserve energy.

VI. CONCLUSIONS AND DISCUSSIONS

In this paper, we have constructed an Explicitly Solvable Energy-Conserving (ESEC) algorithm for the SDE (4) describing the pitch-angle scattering process in magnetized plasmas. We have rigorously proved and numerically verified the order 1/2 strong convergence of the algorithm. The algorithm has also been benchmarked using the analytical solution for the pitch-angle scattering in a constant magnetic field.

In Cartesian coordinates, the coefficients of the Eq. (4) are singular at $v = 0$. In spherical coordinates, the coefficients of the Eq. (5), (6) are also singular at $|\mu| = 1$. These coefficients do not satisfy the global Lipschitz condition. When standard numerical methods, such as the Euler-Maruyama, are applied, numerical convergence is difficult to establish, and unphysical results may show up. For the proposed ESEC algorithm, the Cartesian-coordinates-based calculation and its energy-preserving property enable us to overcome these obstacles.

We would like to emphasize that pitch-angle scattering is not a physics process that only appears in the zero mass-ratio limit. For collisions between two species with small but finite mass ratio, such as those between electrons and ions in fusion plasmas, the collisional effects will include frictional slowing down and diffusion in the incoming (parallel) direction. Similarly, external electrical field will accelerate particles. In these circumstances, pitch-angle scattering is still an important, if not the most important, component of the collisional process, and the algorithmic difficulty associated with divergence at small v remains. Thus, the pitch-angle scattering component can still be integrated by the ESEC algorithm, which can automatically avoid the numerical divergence at small v due to the perpendicular dynamics. Specifically, the operator-splitting method [28] can be applied to separately calculate the pitch-angle scattering and other dynamics at each time step.

Note that when the parallel dynamics is not negligible, the physics of energy conservation is no longer solely represented by the invariance of the velocity norm anymore. As a consequence, the rigorous proof of convergence presented in this paper can not be directly applied. This is a topic for future investigation.

Other future works include the generalization of the ESEC algorithm to more complex and realistic situations. The current paper only considers the particle distribution in the velocity-space. When the background fields are spatially non-uniform, the spatial distribution should also be included. The drift and diffusion coefficients in the SDEs may explicitly depend on

the distribution function. For example, the nonlinear Fokker-Plank equation can be solved using methods based on distribution dependent SDEs [35–37]. Furthermore, the algorithm developed can also be used as a part of the Particle-In-Cell method [38]. We plan to use the algorithm to study physics problems involving interaction between particles and self-consistent electromagnetic fields, e.g., the runaway electrons dynamics coupled with plasma instabilities during a tokamak disruption. Another direction for research is to extend the algorithm to high orders. In this regard, the methods used to develop high-order structure-preserving algorithms for both deterministic differential equations [39–42] and SDEs [33, 43–47] could potentially be adapted.

ACKNOWLEDGMENTS

The authors are grateful to the anonymous referee for the suggestion of the comparison between our algorithm and the EM algorithm in spherical coordinates. X. Zhang would like to thank Francesca Poli and Nicolas Lopez for helpful suggestions. H. Qin thanks Tom Tyranowski and Yajuan Sun for fruitful discussions. This work is supported by US DOE (DE-AC02-09CH11466).

Appendix A: The global Lipschitz condition of coefficients

A function $f(\boldsymbol{x})$ defined on $\boldsymbol{x} \in \mathbb{R}^d$ satisfies the global Lipschitz condition, or is globally Lipschitz continuous, means that for all $\boldsymbol{x} \in \mathbb{R}^d$ and $\boldsymbol{y} \in \mathbb{R}^d$, there exist a constant K independent of $\boldsymbol{x}, \boldsymbol{y}$ such that

$$|f(\boldsymbol{x}) - f(\boldsymbol{y})| \leq K|\boldsymbol{x} - \boldsymbol{y}|. \tag{A1}$$

For the coefficients in Eq. (16), if $\boldsymbol{B}(t)$ is bounded for all t, terms $\boldsymbol{v} \times \boldsymbol{B}$, $\boldsymbol{v}$, $\boldsymbol{I}v$ satisfy the global Lipschitz condition uniformly for all t. The term $\boldsymbol{v}\boldsymbol{v}/v$ consists of two types of functions,

$$f_1(\boldsymbol{x}) = \frac{x_1^2}{\sqrt{x_1^2 + x_2^2 + x_3^2}}, \quad f_2(\boldsymbol{x}) = \frac{x_1 x_2}{\sqrt{x_1^2 + x_2^2 + x_3^2}}. \tag{A2}$$

We can verify that both types are globally Lipschitz continuous as follows.

Since f_1 or f_2 is differentiable in $\mathbb{R}^3/\{0\}$, when $\boldsymbol{x}$ is not proportional to $\boldsymbol{y}$, Eq. (A1) can be proved by the mean value theorem, which states that

$$f(\boldsymbol{x}) - f(\boldsymbol{y}) = (\boldsymbol{x} - \boldsymbol{y}) \cdot \nabla f(\boldsymbol{z}), \tag{A3}$$

where $\boldsymbol{z}$ is a point in the segment between $\boldsymbol{x}$ and $\boldsymbol{y}$. By the Cauchy-Schwarz inequality, we have

$$|f(\boldsymbol{x}) - f(\boldsymbol{y})| \leq |\nabla f(\boldsymbol{z})| \cdot |\boldsymbol{x} - \boldsymbol{y}|. \tag{A4}$$

The gradients of f_1 and f_2 are bounded,

$$|\nabla f_1|^2 = 2 - \frac{x_1^4 + 2(x_2^2 + x_3^2)^2}{(x_1^2 + x_2^2 + x_3^2)^2} \leq 2, \quad |\nabla f_2|^2 = 1 - \frac{3x_1^2 x_2^2 + (x_1^2 + x_2^2)x_3^2 + x_3^4}{(x_1^2 + x_2^2 + x_3^2)^2} \leq 1. \tag{A5}$$

Therefore, they satisfy the global Lipschitz condition.

When $\boldsymbol{x}$ is proportional to $\boldsymbol{y}$, the segment between them passes through the undifferentiable point 0, and the mean value theorem can not be applied directly. In this case, we could write $\boldsymbol{y} = \kappa \boldsymbol{x}$ with some constant κ, and estimate the difference as

$$\begin{aligned} |f_i(\boldsymbol{x}) - f_i(\kappa \boldsymbol{x})| = |(1 - |\kappa|) f_i(\boldsymbol{x})| &= |(1 - |\kappa|) x_i| \cdot \left| \frac{x_1}{\sqrt{x_1^2 + x_2^2 + x_3^2}} \right| \\ &\leq |(1 - |\kappa|)\boldsymbol{x}| = \left| \frac{1 - |\kappa|}{1 - \kappa} \right| |\boldsymbol{x} - \kappa \boldsymbol{x}| \leq |\boldsymbol{x} - \boldsymbol{y}|, \end{aligned} \tag{A6}$$

where $i = 1, 2$.

Therefore, functions f_1 and f_2, as well as all coefficients in Eq. (16), are globally Lipschitz continuous.

Appendix B: Numerical evaluations of strong and weak errors

Assuming relative to the analytical solution $\boldsymbol{v}(t_k; t_0, \boldsymbol{v}_0)$, the approximation $\bar{\boldsymbol{v}}(t_k; t, \boldsymbol{v}_0, h)$ has order p_s strong error and order p_w weak error in the entire time interval, then for two different time steps h_l and h_{l+1} with $h_{l+1} < h_l$, we have

$$\begin{aligned} \bar{\epsilon}_{\mathrm{s}}^2 =& \mathbb{E}|\bar{\boldsymbol{v}}(t; t_0, \boldsymbol{v}_0, h_{l+1}) - \bar{\boldsymbol{v}}(t; t_0, \boldsymbol{v}_0, h_l)|^2 \\ \leq& \mathbb{E}|\boldsymbol{v}(t; t_0, \boldsymbol{v}_0) - \bar{\boldsymbol{v}}(t; t_0, \boldsymbol{v}_0, h_{l+1})|^2 + \mathbb{E}|\boldsymbol{v}(t; t_0, \boldsymbol{v}_0) - \bar{\boldsymbol{v}}(t; t_0, \boldsymbol{v}_0, h_l)|^2 \\ \leq& (1 + \mathbb{E}|\boldsymbol{v}_0|^2)(K_l^2 h_l^{2p_s} + K_{l+1}^2 h_{l+1}^{2p_s}) \leq K(1 + \mathbb{E}|\boldsymbol{v}_0|^2) h_l^{2p_s}. \end{aligned} \tag{B1}$$

Thus, $\bar{\epsilon}_{\rm s}$ is order $\mathcal{O}(h_l^{p_s})$. Similarly for the weak error, we have

$$\begin{aligned}\bar{\epsilon}_{\rm w} &= |\mathbb{E}\left[\bar{\boldsymbol{v}}(t;t_0,\boldsymbol{v}_0,h_{l+1}) - \bar{\boldsymbol{v}}(t;t_0,\boldsymbol{v}_0,h_l)\right]| \\ &\le |\mathbb{E}\left[\boldsymbol{v}(t;t_0,\boldsymbol{v}_0) - \bar{\boldsymbol{v}}(t;t_0,\boldsymbol{v}_0,h_{l+1})\right]| + |\mathbb{E}\left[\boldsymbol{v}(t;t_0,\boldsymbol{v}_0) - \bar{\boldsymbol{v}}(t;t_0,\boldsymbol{v}_0,h_l)\right]| \\ &\le (1+\mathbb{E}|\boldsymbol{v}_0|^2)^{1/2}(K_l h_l^{p_w} + K_{l+1}h_{l+1}^{p_s}) \le K(1+\mathbb{E}|\boldsymbol{v}_0|^2)^{1/2}h_l^{p_w}. \end{aligned} \tag{B2}$$

Thus, $\bar{\epsilon}_{\rm w}$ is order $\mathcal{O}(h_l^{p_w})$.

Appendix C: Euler-Maruyama scheme in different coordinates

In this appendix we briefly describe the Euler-Maruyama (EM) scheme for pitch-angle scattering in both Cartesian and spherical coordinate systems. Different regularization techniques used in each coordinate system will be introduced. The method for comparing sample paths with the same underlying Wiener process but calculated in different coordinate systems will also be discussed.

1. EM scheme in Cartesian coordinates

In Cartesian coordinates, the Ito SDE in Eq. (4) can be re-written as:

$$\mathrm{d}\boldsymbol{v} = \Big(\boldsymbol{v}\times\boldsymbol{B} - \mathcal{F}(v)\hat{\boldsymbol{v}}\Big)\mathrm{d}t + \mathcal{D}(v)(\boldsymbol{I}-\hat{\boldsymbol{v}}\hat{\boldsymbol{v}})\cdot\mathrm{d}\boldsymbol{W}, \tag{C1}$$

where $\hat{\boldsymbol{v}} = \boldsymbol{v}/v$ is the unit vector parallel to $\boldsymbol{v}$, $\mathcal{F}(v) = D(v)/v$, $\mathcal{D}(v) = \sqrt{D(v)}$, and $D(v) = 1/v$. When $v\to 0$, the scalar functions $\mathcal{F}$ and $\mathcal{D}$ are divergent, while the rest are finite. Similar to Ref. [26], for a chosen critical velocity v_c, the regularized function $\mathcal{F}_{\rm r}(v)$ is defined as

$$\mathcal{F}_{\rm r}(v) := \begin{cases} \mathcal{F}(v), & v > v_c, \\ \dfrac{\mathcal{F}'(v_c)}{2v_c}(v^2-v_c^2) + \mathcal{F}(v_c), & v\le v_c, \end{cases} \tag{C2}$$

where $\mathcal{F}' = \mathrm{d}\mathcal{F}/\mathrm{d}v$. The regularized function $\mathcal{D}_{\rm r}(v)$ is similarly defined. It is clear that $\mathcal{D}_{\rm r}(v)$ and $\mathcal{F}_{\rm r}(v)$ are C^1-continuous and regular at $v=0$. The regularized EM scheme in Cartesian coordiantes (rEM_C) is

$$\boldsymbol{v}_{k+1}^{\rm rEM\_C} - \boldsymbol{v}_k = \Big(\boldsymbol{v}_k\times\boldsymbol{B}_k - \mathcal{F}_{\rm r}(v_k)\hat{\boldsymbol{v}}_k\Big)h + \mathcal{D}_{\rm r}(v_k)(\boldsymbol{I}-\hat{\boldsymbol{v}}_k\hat{\boldsymbol{v}}_k)\cdot\Delta\boldsymbol{W}. \tag{C3}$$

For the simulation conducted in Sec. V, we set $v_c = 0.2$.

2. EM scheme in spherical coordinates

Assuming the magnetic field B is constant and in the $\hat{z}$-direction, the Ito SDE in spherical coordinates can be written as:

$$\mathrm{d}v = 0, \tag{C4}$$

$$\mathrm{d}\mu = -D_a(v)\mu\,\mathrm{d}t + \sqrt{D_a(v)(1-\mu^2)}\,\mathrm{d}W_\mu, \tag{C5}$$

$$\mathrm{d}\varphi = -B\,\mathrm{d}t + \sqrt{\frac{D_a(v)}{1-\mu^2}}\,\mathrm{d}W_\varphi, \tag{C6}$$

where v is the velocity magnitude, $\mu = \cos\theta$ is the cosine of the polar angle, φ is the azimuthal angle, and $D_a(v) = 1/v^3$. From Eq. (C4), we know that v and $D_a(v)$ are exactly constant. When $|\mu| = 1$, the coefficient $\sqrt{1-\mu^2}$ is not Lipschitz continuous. In addition, when $|\mu|$ is too close to 1, i.e., when the particle is too close to north and south pole, the numerical calculation may lead to $|\mu| > 1$ and imaginary $\sqrt{1-\mu^2}$. Therefore, some regularization methods have to be applied at $|\mu| \to 1$.

Here we choose the regularization method used in Ref. [26], which modified the coefficient $\sqrt{1-\mu^2}$ to:

$$\mathcal{M}(\mu) = \begin{cases} \sqrt{1-\mu^2}, & |\mu| < \mu_c, \\ \sqrt{1-\mu_c^2}\exp\left[-\left(|\mu| - \mu_c\right) S\left(\mu_c\right)\right], & |\mu| \geq \mu_c, \end{cases} \tag{C7}$$

where $S(\mu_c) = \mu_c/(1-\mu_c^2)$, and μ_c is a chosen critical value. We see that $\mathcal{M}$ is positive and C^1-continuous for all $\mu \in \mathbb{R}$. To ensure the Einstein relation, the drift term in Eq. (C5) is also modified to $D_a\mathcal{M}\mathcal{M}'$, where $\mathcal{M}' = \mathrm{d}\mathcal{M}/\mathrm{d}\mu$. The regularized EM scheme in spherical coordinates (rEM_s) is:

$$\mu_{k+1}^{\mathrm{rEM\_s}} - \mu_k = D_a\mathcal{M}(\mu_k)\mathcal{M}'(\mu_k)\Delta t + \sqrt{D_a}\mathcal{M}(\mu_k)\Delta W_\mu, \tag{C8}$$

$$\varphi_{k+1}^{\mathrm{rEM\_s}} - \varphi_k = B\Delta t + \frac{\sqrt{D_a}}{\mathcal{M}(\mu_k)}\Delta W_\varphi. \tag{C9}$$

For the simulation in Sec. V, we set $\mu_c = 0.9$. One thing to notice is that although this regularization method prevents imaginary value, it does not prevent $|\mu|$ from exceeding 1, which is shown in Sec. V A.

3. Comparing sample paths in different coordinates

Usually in the numerical calculations, $\Delta \boldsymbol{W} = (\Delta W_x, \Delta W_y, \Delta W_z)$ or $(\Delta W_v, \Delta W_\mu, \Delta W_\varphi)$ are generated as Gaussian random variables directly. However, if we want to compare sample paths calculated in different coordinates but with the same underlying Wiener process, for example in Sec. V A, we need to generate $\Delta \boldsymbol{W}$ in one coordinate system, then transform it to the other coordinates.

Assume we have a three dimensional Wiener process $\boldsymbol{W}(t)$ in Cartesian coordinates, but we want to calculate the sample path in spherical coordinates. At each time step t_k, $\Delta \boldsymbol{W} = \boldsymbol{W}(t_{k+1}) - \boldsymbol{W}(t_k)$ and the sample is at $\boldsymbol{v}_k = (v_k, \mu_k, \varphi_k)$. Then $\Delta \boldsymbol{W}$ in spherical coordinates is given by:

$$\begin{pmatrix} \Delta W_r \\ \Delta W_\mu \\ \Delta W_\varphi \end{pmatrix} = \begin{pmatrix} \sqrt{1-\mu_k^2}\cos\varphi_k & \sqrt{1-\mu_k^2}\sin\varphi_k & \mu_k \\ -\mu_k\cos\varphi_k & -\mu_k\sin\varphi_k & \sqrt{1-\mu_k^2} \\ -\sin\varphi_k & \cos\varphi_k & 0 \end{pmatrix} \begin{pmatrix} \Delta W_x \\ \Delta W_y \\ \Delta W_z \end{pmatrix}. \tag{C10}$$

Here we see the factor $\sqrt{1-\mu^2}$ again, which can be regularized using the same method described in the previous section. The inverse transform from spherical to Cartesian coordinates can be performed similarly.

---

[1] L. Landau, Phys. Z. Sowjetunion **10**, 154 (1936).

[2] N. J. Fisch, Physical Review Letters **41**, 873 (1978).

[3] N. J. Fisch, Reviews of Modern Physics **59**, 175 (1987).

[4] C. F. F. Karney and N. J. Fisch, Physics of Fluids **29**, 180 (1986).

[5] M. N. Rosenbluth and S. V. Putvinski, Nuclear Fusion **37**, 1355 (1997).

[6] A. H. Boozer, Physics of Plasmas **22**, 032504 (2015).

[7] M. d'Aquino, C. Serpico, G. Coppola, I. Mayergoyz, and G. Bertotti, Journal of applied physics **99**, 08B905 (2006).

[8] L. Chacón, D. Barnes, D. Knoll, and G. Miley, Journal of Computational Physics **157**, 618 (2000).

[9] E. Yoon and C. Chang, Physics of Plasmas **21**, 032503 (2014).

[10] M. Kraus and E. Hirvijoki, Physics of Plasmas **24**, 102311 (2017).

[11] E. Hirvijoki, M. Kraus, and J. W. Burby, Metriplectic particle-in-cell integrators for the Landau collision operator (2018), arXiv:1802.05263 [physics.comp-ph].

[12] T. Takizuka and H. Abe, Journal of computational physics **25**, 205 (1977).

[13] K. Nanbu, Physical Review E **55**, 4642 (1997).

[14] W. M. Manheimer, M. Lampe, and G. Joyce, Journal of Computational Physics **138**, 563 (1997).

[15] M. Cadjan and M. Ivanov, Journal of plasma physics **61**, 89 (1999).

[16] M. Sherlock, Journal of Computational Physics **227**, 2286 (2008).

[17] D. S. Lemons, D. Winske, W. Daughton, and B. Albright, Journal of Computational Physics **228**, 1391 (2009).

[18] B. I. Cohen, A. M. Dimits, A. Friedman, and R. E. Caflisch, IEEE transactions on plasma science **38**, 2394 (2010).

[19] A. M. Dimits, B. I. Cohen, R. E. Caflisch, M. Rosin, and L. Ricketson, Journal of Computational Physics **242**, 561 (2013).

[20] X. Zhang, Y. Fu, and H. Qin, Phys. Rev. E **102**, 033302 (2020).

[21] J. P. Verboncoeur, Plasma Physics and Controlled Fusion **47**, A231 (2005).

[22] P. E. Kloeden and E. Platen, *Numerical solution of stochastic differential equations*, Vol. 23 (Springer Science & Business Media, 2013).

[23] D. Lemons, J. Lackman, M. Jones, and D. Winske, Physical Review E **52**, 6855 (1995).

[24] M. N. Rosenbluth, W. M. MacDonald, and D. L. Judd, Physical Review **107**, 1 (1957).

[25] A. H. Boozer and G. Kuo-Petravic, The Physics of Fluids **24**, 851 (1981).

[26] M. Rosin, L. Ricketson, A. M. Dimits, R. E. Caflisch, and B. I. Cohen, Journal of Computational Physics **274**, 140 (2014).

[27] G. N. Milstein and M. V. Tretyakov, *Stochastic numerics for mathematical physics* (Springer Science & Business Media, 2013).

[28] D. J. Higham, X. Mao, and A. M. Stuart, SIAM Journal on Numerical Analysis **40**, 1041 (2002).

[29] M. Hutzenthaler, A. Jentzen, and P. E. Kloeden, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences **467**, 1563 (2011).

[30] H. Qin, S. Zhang, J. Xiao, J. Liu, Y. Sun, and W. M. Tang, Physics of Plasmas **20**, 084503 (2013).

[31] Z. Shang, (2013), private communication.

[32] J. P. Boris, in *Proc. Fourth Conf. Num. Sim. Plasmas* (1970) pp. 3–67.

[33] G. N. Milstein, Y. M. Repin, and M. V. Tretyakov, SIAM Journal on Numerical Analysis **40**, 1583 (2002).

[34] D. R. Brillinger, A particle migrating randomly on a sphere, in *Selected Works of David Brillinger*, edited by P. Guttorp and D. Brillinger (Springer New York, New York, NY, 2012) pp. 73–87.

[35] V. Barbu, M. Röckner, *et al.*, Annals of Probability **48**, 1902 (2020).

[36] F.-Y. Wang, Stochastic Processes and their Applications **128**, 595 (2018).

[37] T. D. Frank, *Nonlinear Fokker-Planck equations: fundamentals and applications* (Springer Science & Business Media, 2005).

[38] E. Allen and H. Victory Jr, Physica A: Statistical Mechanics and its Applications **209**, 318 (1994).

[39] Y. He, Y. Sun, J. Liu, and H. Qin, Journal of Computational Physics **281**, 135 (2015).

[40] Y. He, Y. Sun, J. Liu, and H. Qin, Journal of Computational Physics **305**, 172 (2016).

[41] Y. He, Y. Sun, R. Zhang, Y. Wang, J. Liu, and H. Qin, Physics of Plasmas **23**, 092109 (2016).

[42] Y. He, Z. Zhou, Y. Sun, J. Liu, and H. Qin, Physics Letters A **381**, 568 (2017).

[43] J. Deng, C. Anton, and Y. S. Wong, Communications in Computational Physics **16**, 169 (2014).

[44] P. Wang, J. Hong, and D. Xu, Communications in Computational Physics **21**, 237 (2016).

[45] J. Hong, L. Sun, and X. Wang, SIAM Journal on Numerical Analysis **55**, 3006 (2017).

[46] W. Zhou, J. Zhang, J. Hong, and S. Song, Journal of Computational and Applied Mathematics **325**, 134 (2017).

[47] D. D. Holm and T. M. Tyranowski, BIT Numerical Mathematics **58**, 1009 (2018).