A Backward Monte Carlo method for fast-ion-loss simulations

Filippo Zonta,[1] Lucia Sanchis,[1] Eero Hirvijoki,[1] and the ASDEX Upgrade Team[a]

[1] *Department of Applied Physics, Aalto University, P.O. Box 11100, 00076 AALTO, Finland*

(Dated: November 4, 2021)

This paper presents a novel scheme to improve the statistics of simulated fast-ion loss signals and power loads to plasma-facing components in fusion devices. With the so-called Backward Monte Carlo method, the probabilities of marker particles reaching a chosen target surface can be approximately traced from the target back into the plasma. Utilizing the probabilities as *a priori* information for the well-established Forward Monte Carlo method, statistics in fast-ion simulations are significantly improved. For testing purposes, the scheme has been implemented to the ASCOT suite of codes and applied to a realistic ASDEX Upgrade configuration of beam-ion distributions.

I. INTRODUCTION

In fusion devices, the fast-ion population plays a crucial role in the plasma performance as it provides one of the primary sources of plasma heating. This is of special importance for future devices such as ITER, where fusion-born alpha particles are to ensure the self-sufficiency of the burning plasma. Moreover, a degradation in the fast-ion confinement can lead to substantial damage in the plasma-facing components in the form of incident heat fluxes, risking the integrity of the device. The behaviour of fast ions therefore calls for careful assessment, especially due to the sensitivity of fast-ion physics to complex 3D magnetic background credited to the relatively long mean-free-path and slowing-down times of fast ions [1–3]. In this regard, it is of vital interest to predict the behaviour and the possible confinement losses of such particles with tested and validated models.

The standard approach to fast-ion-loss simulations is based on the so-called forward Monte Carlo method (FMC): a population of fast-ions is initialized and then followed until particles either reach thermal energies or hit a first-wall element. The method, although well-established and straightforward to implement, is commonly burdened with statistics problems which may lead to low accuracy in analyzing small targets. Often only a fraction of the initialized markers contribute to the signal that is being calculated, such as the number of ions lost to a fast-ion-loss detector attached to the first wall.

As a remedy, the so-called Backward Monte Carlo (BMC) algorithm can be adopted. While perhaps less known within the fast-ion fusion-physics community, Backward Monte Carlo schemes are commonly used, e.g., in quantitative finance to price options [4]. In plasma physics, the method has been successfully applied to compute escape probabilities of runaway electrons in the velocity space [5] and the closely related idea of studying the adjoint of the Fokker-Planck equation linked to the stochastic particle motion has been extensively used in studying current drive (see, e.g., [6]). The idea of applying it to fast-ion simulations has, however, been floated only recently [7]. In the BMC approach, instead of starting with a given birth marker population, the probability of a marker reaching a specific area is

[a] See the author list of H. Meyer et al 2019 Nucl. Fusion 59 112014

tracked backwards in time starting at the region of interest. The backwards calculation of the probability involves an iterative process in which, during every step, points in phase-space are advanced according to their forward equations of motion and the statistical spread from the Monte Carlo collision operator is taken into account deterministically. Mathematically, the forward and backward methods provide identical results for the probability.

In the present work, we propose a novel 5D guiding-center BMC-based scheme for simulation of fast-ion losses, integrated into the existing and established Monte Carlo orbit-following code ASCOT [8, 9]. Instead of directly utilizing the escape probability obtained with the BMC-scheme, the computed probability is used as *a priori* information for importance sampling of marker particles from the birth distribution. This combined scheme we refer to as BMC-IS and it enables an optimally efficient FMC execution, mitigating the error from numerical diffusion unavoidably present in the standard BMC-scheme, an issue that was demonstrated in [7] and also discussed here. The proposed BMC-IS scheme alleviates the statistics problem of the standard FMC scheme, providing a more accurate algorithm to estimate the fast-ion fluxes on small targets. We expect this tool to be especially useful in accelerating computational studies regarding heat loading of any particular wall element or estimating the velocity-space distribution of fast ions observed in fast-ion-loss detectors.

The rest of this paper is organized as follows. In section II, the theory behind fast-ion simulations is briefly reviewed and the connection to stochastic differential equations of Itô kind established. Section III first introduces the Backward Monte Carlo scheme with a simple toy problem, then generalizes it for the Itô processes describing the fast-ions, and finally explains the discretization of the equations involved and how they are used to develop a numerical scheme. The section also discusses the issue of numerical diffusion, how to minimize it, how to avoid it entirely via combining the BMC and FMC methods via importance sampling, and briefly also the memory and CPU optimizations used in the present implementation. Section IV then provides numerical examples applying the novel scheme to a realistic ASDEX Upgrade case and compares the results to the standard FMC-scheme. Finally, in section V, the main results are summarized and future prospects discussed.

II. THEORETICAL BACKGROUND FOR FAST-ION SIMULATIONS

The distribution function of charged particles in a magnetized plasma is traditionally studied with the Vlasov equation coupled to the Landau collision operator. The collisional evolution only affects the velocity part of the distribution function and can be expressed in the Fokker-Planck form

$$\left.\frac{\partial f(\boldsymbol{z},t)}{\partial t}\right|_{\mathrm{c}} = -\frac{\partial}{\partial \boldsymbol{v}} \cdot \left(\boldsymbol{K}(\boldsymbol{z},t) f(\boldsymbol{z},t)\right) + \frac{\partial}{\partial \boldsymbol{v}} \frac{\partial}{\partial \boldsymbol{v}} : \left(\mathbf{D}(\boldsymbol{z},t) f(\boldsymbol{z},t)\right), \tag{1}$$

where $\boldsymbol{z} = (\boldsymbol{x}, \boldsymbol{v})$ is a phase space point and the coefficients $\boldsymbol{K}$ and $\boldsymbol{D}$ are functionals of the distribution functions of different plasma species f_β. In studying fast ions, the background plasma distributions are typically approximated as Maxwellians and the self-collisions neglected. The kinetic equation (1) then becomes linear, and the collisional evolution of an individual sample particle from f obeys the following stochastic differential equation of Itô kind

$$d\boldsymbol{v}_t^c = \boldsymbol{K}\left(\boldsymbol{z}_t, t\right) dt + \boldsymbol{\sigma}\left(\boldsymbol{z}_t, t\right) \cdot d\boldsymbol{W}_t, \tag{2}$$

where $d\boldsymbol{W}_t$ represent the differential of a standard Wiener process $\boldsymbol{W}_t \sim \mathcal{N}(\boldsymbol{0}, t\mathbf{I})$ with $\mathcal{N}$ being a multivariate normal distribution. The coefficients $\boldsymbol{K}$ and $\boldsymbol{\sigma}$ have the explicit expressions in Cartesian coordinates

$$K^i(\boldsymbol{x}, \boldsymbol{v}, t) = \nu v^i, \tag{3}$$

$$\sigma^{ij}(\boldsymbol{x}, \boldsymbol{v}, t) = \sqrt{2D_\parallel}\frac{v^i v^j}{|\boldsymbol{v}|^2} + \sqrt{2D_\perp}\left(\delta^{ij} - \frac{v^i v^j}{|\boldsymbol{v}|^2}\right), \tag{4}$$

with $D_\parallel$, $D_\perp$ and ν defined in terms of the function $\Psi(y) = \pi^{-1/2}\exp(-y^2) + (y + 1/(2y))\,\mathrm{erf}(y)$ and evaluated at the particle speed normalized to the thermal velocities of the species β, i.e., at $y_\beta = v/\sqrt{2T_\beta/m_\beta}$. The explicit forms are given by

$$\nu = \sum_\beta \frac{n_\beta c_\beta}{m^2}\left(1 + \frac{m}{m_\beta}\right)\left(\frac{m_\beta}{2T_\beta}\right)^{3/2}\frac{1}{y}\frac{d}{dy}\left(\frac{d^2\Psi(y)}{dy^2} + \frac{2}{y}\frac{d\Psi(y)}{dy}\right)\Bigg|_{y=y_\beta}, \tag{5}$$

$$D_\parallel = \sum_\beta \frac{n_\beta c_\beta}{m^2}\sqrt{\frac{m_\beta}{2T_\beta}}\frac{d^2\Psi(y)}{dy^2}\Bigg|_{y=y_\beta}, \tag{6}$$

$$D_\perp = \sum_\beta \frac{n_\beta c_\beta}{m^2}\sqrt{\frac{m_\beta}{2T_\beta}}\frac{1}{y}\frac{d\Psi(y)}{dy}\Bigg|_{y=y_\beta}, \tag{7}$$

with $c_\beta = q^2 q_\beta^2 \ln\Lambda/(8\pi\epsilon_0^2)$.

In many plasmas relevant for fusion physics, the variations of the electromagnetic fields occur at length scales longer than the gyroradius and time scales slower than the gyrofrequency. In this case, the rapid gyromotion of a charged particle can be decoupled from the guiding-center motion. The sample particles from the distribution f under such conditions still obey an Itô-like equation

$$d\mathbf{z}_t^c = \boldsymbol{\mathcal{K}}dt + \boldsymbol{\Sigma}\cdot d\boldsymbol{W}, \tag{8}$$

where the guiding center coordinates $\boldsymbol{z} = (\boldsymbol{X}, u, \xi)$ are the position, $\boldsymbol{X}$, the magnitude of the velocity, v, and the pitch with respect to the magnetic field direction, $\xi = \boldsymbol{v}\cdot\boldsymbol{b}/v$. Written component-wise, the collision operator (8) takes the form

$$dX_t^i = \sqrt{2D_\mathrm{X}}(\delta^{ij} - b^i b^j)dW_X^i, \tag{9}$$

$$dv_t = K_v dt + \sqrt{2D_\parallel}dW_v, \tag{10}$$

$$d\xi_t = -2\xi_t\frac{D_\perp}{v^2}dt + \sqrt{2\left(1-\xi^2\right)\frac{D_\perp}{v^2}}dW_\xi. \tag{11}$$

The spatial diffusion coefficient D_X is defined by the expression

$$\mathcal{D}_X = \left[\left(D_\parallel - D_\perp\right)\frac{1-\xi^2}{2} + D_\perp\right]\frac{1}{\omega_c^2}, \tag{12}$$

where $\omega_c = qB/m$ is the cyclotron frequency, and the momentum drift K_v is

$$K_v = \sum_\beta \frac{n_\beta c_\beta m_\beta}{2mT_\beta}\left(-\frac{2y}{m_\beta}\frac{d^2\Psi}{dy^2} + \frac{1}{m}\frac{d^3\Psi}{dy^3} + \frac{2}{m}\frac{1}{y}\frac{d^2\Psi}{dy^2}\right)\Bigg|_{y=y_\beta}. \tag{13}$$

For details, the reader is referred to [9, 10]. In this paper, the Backward Monte Carlo scheme is presented by referring to the guiding center coordinates only and hence to the collisional operator equations (9)–(11).

The full equation of motion for the guiding center includes the collisional part as well as the Hamiltonian guiding-center vector field, labelled as $\boldsymbol{U}(\boldsymbol{z}, t)$. The stochastic differential equation to be solved for guiding-center motion is then

$$d\boldsymbol{z}_t = \boldsymbol{U}(\boldsymbol{z}_t, t)dt + d\boldsymbol{z}_t^C. \tag{14}$$

In the FMC approach, ASCOT integrates the Hamiltonian vector field $\boldsymbol{U}$ with an adaptive fourth order Runge-Kutta Cash-Karp scheme, denoted by the map

$$\boldsymbol{z} \mapsto \varphi_{\boldsymbol{U}}^{\Delta t, t}(\boldsymbol{z}), \tag{15}$$

and the collisional part with the Euler-Mariyama scheme

$$\boldsymbol{z} \mapsto \varphi_{\boldsymbol{C}}^{\Delta t, t}(\boldsymbol{z}, \boldsymbol{\mathcal{R}}) = \boldsymbol{z} + \boldsymbol{\mathcal{K}}(\boldsymbol{z}, t)\Delta t + \boldsymbol{\Sigma}(\boldsymbol{z}, t) \cdot \boldsymbol{\mathcal{R}}\sqrt{\Delta t}, \tag{16}$$

where $\boldsymbol{\mathcal{R}} \sim \mathcal{N}(0, \mathbf{I})$ is a random variable following a standard normal multivariate distribution and Δt the time step length.

III. BACKWARD MONTE CARLO METHOD

Throughout the rest of this paper the target domain Ω refers to the subset of the phase space boundary domain, typically a small portion of the wall, which mimics the function of a loss-detector.

If charged particles hit the target domain with a relatively low rate compared to the temporal scales of the system, it is typically difficult for standard FMC methods to reach good statistics for hits and at the same time keep the total number of simulated markers to a reasonable value. Countering this, the idea of the BMC scheme is to reconstruct the probability of the markers to hit the target backwards in time using an iterative algorithm. To do so, the phase space is discretized on a mesh and, at each time step, the particle trajectories, normally randomized with the collision operators, are deterministically branched and composed using conditional probabilities and probability measures.

A. Theory

To intuitively visualize the method, consider the simplified one dimensional problem depicted in Fig. 1, where the aim is to determine the probability of particles to reach a certain end condition in a given time. At the final time $3\Delta t$, the probability is trivially vanishing everywhere except at the target domain, in this case represented by the point $z = 2$. Going backwards in time, assuming, e.g., that a marker at a point has equal probability to jump to any nearest points or to stay put, the probability to reach the target from a specific initial location in the given total time interval is built recursively, utilizing the probabilities of the marker particles to move forward in time. The generalization to a realistic 5D guiding center case follows the same philosophy with some additional subtleties, for example the fact that the particle trajectory is continuous and needs to be suitably discretized, and that boundary conditions are handled correctly.

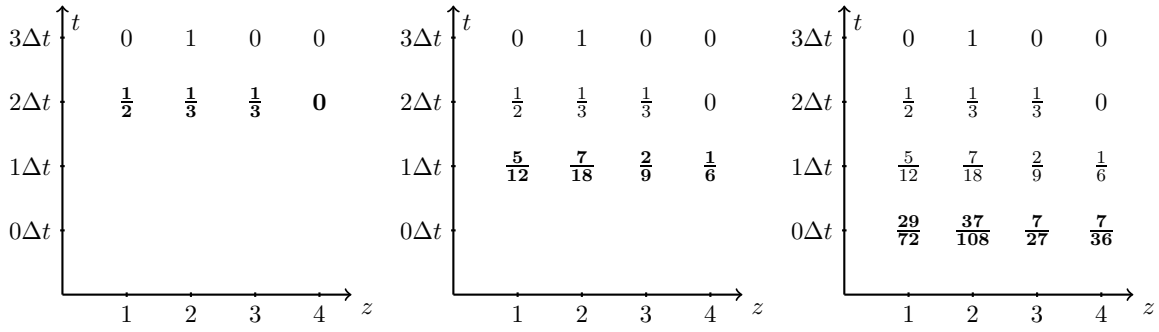


FIG. 1. Backwards evolution of probability for an example, discrete one-dimensional problem. A particle at each discrete point z_i has equal probability to jump to any possible nearest location on the ordered set $\{1, 2, 3, 4\}$, including staying put. At every step, the probability is evolved backwards in time, yet constructed from the probabilities for a marker to move forward in time.

Discretizing time as the regularly spaced ordered set $\{0, \Delta t, ..., m\Delta t, ..., M\Delta t\}$, the function Φ is used to describe the probability of a particle at phase-space position $\boldsymbol{z}$ and time index m to reach the target domain Ω during a time interval $\tau \in [m\Delta t, M\Delta t]$. In other words,

$$\Phi(\boldsymbol{z}, m) = \mathbb{E}\left[P\left(\boldsymbol{z}_\tau; \Omega, M\Delta t\right) \mid \boldsymbol{z}_{m\Delta t} = \boldsymbol{z}\right], \tag{17}$$

where the expression $P(\boldsymbol{z}_\tau; \Omega, M\Delta t)$ denotes a statement of the end condition and could read "a marker starting from the position $\boldsymbol{z}$ at time $m\Delta t$ has intersected the domain Ω boundary before the time $M\Delta t$". For the end point in time, the probability function Φ is clearly the indicator function of the target domain

$$\Phi(\boldsymbol{z}, M) = \mathbf{1}_\Omega(\boldsymbol{z}), \tag{18}$$

as the particles have effectively zero time to reach the end condition, while $\Phi(\boldsymbol{z}, 0)$ gives the probability for markers to reach the end condition during the interval $t \in [0, M\Delta t]$, the quantity one is interested in. For times other than the final one, [5, 7] showed that the probability can be evaluated with the following recursive formula

$$\Phi(\boldsymbol{z}, m-1) = \int_{\mathbb{R}^d} \Phi\left(\varphi^{\Delta t,(m-1)\Delta t}(\boldsymbol{z}, \boldsymbol{r}), m\right) \frac{\exp\left[-\boldsymbol{r}^2/2\right]}{(2\pi)^{d/2}} d\boldsymbol{r}, \tag{19}$$

where $(\boldsymbol{z}, t) \mapsto (\varphi^{\Delta t,t}(\boldsymbol{z}, \boldsymbol{\mathcal{R}}), t + \Delta t)$ is the particle integrator map combining both the Hamiltonian motion $\varphi_{\boldsymbol{U}}^{\Delta t,t}(\boldsymbol{z})$ and the collisional contribution $\varphi_{\boldsymbol{C}}^{\Delta t,t}(\boldsymbol{z}, \boldsymbol{\mathcal{R}})$. The combined map is constructed as the composition $\varphi^{\Delta t,t}(\boldsymbol{z}, \boldsymbol{\mathcal{R}}) = \varphi_{\boldsymbol{C}}^{\Delta t,t}(\varphi_{\boldsymbol{U}}^{\Delta t,t}(\boldsymbol{z}), \boldsymbol{\mathcal{R}})$, moving a particle from a point $\boldsymbol{z}$ and time t to a new point $\varphi^{\Delta t,t}(\boldsymbol{z}, \boldsymbol{\mathcal{R}})$ and time $t + \Delta t$, taking in also the multivariate random numbers $\boldsymbol{\mathcal{R}}$ needed in the Euler-Maryama step (16). In (19), the random sampling of the trajectories normally done with $\boldsymbol{\mathcal{R}}$ is replaced by the deterministic integration over $\boldsymbol{r}$, weighted with the distribution of the random numbers $\boldsymbol{\mathcal{R}}$.

B. Numerical implementation

In order to test the Backward Monte Carlo theory, a numerical scheme has been developed and integrated with the orbit-following code ASCOT, which is able to follow minority

particles, such as fast ions, in realistic plasmas and magnetic field profiles. These properties, along with the extensive use of modern parallelization techniques, such as multithreading, SIMD vectorization, and MPI, constitutes an ideal test bed for the BMC scheme. For the test case, the implementation focuses on time-independent plasma backgrounds only. This effectively drops the t dependency of the diffusion coefficients and integrator maps, and enables pre-computing some of the necessary expressions leading to quick evaluations. Also a time-dependent version has been implemented but, here, only the time-independent case will be discussed explicitly.

The BMC algorithm is structured as follows. First, a phase-space mesh is constructed to represent the function Φ, making sure that it covers the whole domain of interest, including the plasma region and the wall. For now, regular rectangular meshes have been used for simplicity and for software engineering reasons, even though the scheme does not rely on any particular mesh structures and can therefore be generalized directly to unstructured meshes. The mesh nodes are distributed accross all the threads and MPI nodes, and they are further divided into SIMD blocks for vectorization, so that one computational unit will have to compute the probability of a specific subgroup of nodes. Given the fact that the recursive computation of the probability in (19) for a specific time step m depends on the probabilities of other mesh nodes, but only on the $m+1$ time step, each MPI node, at every time step, needs to allocate and have access to the entire probability mesh for two consecutive time steps. This could be relaxed somewhat by implementing domain decomposition but it is left for future work.

After the mesh initialization procedure, the main time loop is initiated, starting from the end time $t = T_{\max}$ and proceeding backwards to the initial time $t = 0$ with steps of Δt. The process is as follows:

1. Markers starting from the mesh nodes $\boldsymbol{z}_i$ are evolved forward in time for one time step Δt, using the adaptive Runge-Kutta Cash-Karp method for the Hamiltonian part $\varphi_{\boldsymbol{U}}^{\Delta t}$. In case of fast ions, the pitch part of the diffusion tensor is dominant, and consequently the stochastic part can be simplified. The diffusion in $\boldsymbol{X}$ and v is neglected, resulting in a simple scheme which first pushes all of the guiding-center coordinates with the mapping

$$\boldsymbol{z}_i \mapsto \varphi_{\boldsymbol{U}}^{\Delta t}(\boldsymbol{z}_i) + \boldsymbol{\mathcal{K}}(\mathbf{z}_i)\Delta t, \tag{20}$$

where $\boldsymbol{\mathcal{K}}$, defined in (8), models the friction part of the collision operator, and then only adds the pitch diffusion according to the map

$$\xi_i \mapsto \xi_i + r\sqrt{2\left(1-\xi_i^2\right)\frac{D_\perp(\boldsymbol{X}_i, v_i,)}{v_i^2}\Delta t}. \tag{21}$$

Here the collisional coefficients are defined in (3)–(7), and r is the parameter which in FMC method would be replaced with a random variable following a normal distribution. We note that it is possible to include also the diffusion in $\mathbf{X}$ and v but here computational performance is favoured on the basis of the pitch diffusion being the dominant stochastic effect for fast ions. For time-independent backgrounds, the map consisting of the steps (20) and (21), is precomputed, stored in the memory, and denoted as

$$\boldsymbol{z}_i \mapsto \varphi^{\Delta t}(\boldsymbol{z}_i, r) \equiv \varphi_i^{\Delta t}(r), \tag{22}$$

Storing of this map is possible as it depends linearly on the parameter r.

2. The markers in step 1, although starting from the mesh nodes, in general evolve to arbitrary phase-space position. Hence, interpolation of the probability function Φ is required. In this paper, the probability is represented as a linear interpolant

$$\Phi(\boldsymbol{z}, m) = \sum_i \Phi_i^m \Lambda^i(\boldsymbol{z}) \tag{23}$$

and, with 5D guiding-center coordinates, each degree-of-freedom Φ_i, associated with a mesh node $\boldsymbol{z}_i$, will require information from 32 nearest surrounding mesh nodes. In the end, the probability function Φ at the mesh nodes is evaluated from (19), computing the now one-dimensional integral over pitch scattering numerically with a Gauss-Hermite quadrature of desired order K. This results in an explicit formula for the update rule

$$\Phi_i^{m-1} = \pi^{-1/2} \sum_k w_k \sum_j \Phi_j^m \Lambda^j(\varphi_i^{\Delta t}(\sqrt{2} r_k)), \tag{24}$$

where r_k and w_k are the standard nodes and weights for the Gauss-Hermite integration of the type $\int_{-\infty}^{\infty} \exp(-r^2) f(r) dr \approx \sum_k w_k f(r_k)$. If the trajectory $\boldsymbol{z}_i \mapsto \varphi_i^{\Delta t}(\sqrt{2} r_k)$ hits the wall, then its probability contribution to the summation is recorded as 1 or 0 depending on whether the trajectory hits the desired target or not.

3. At this point, each MPI node is keeping a partial copy of the probability for the current time step. The nodes are then synchronized and the probabilities merged with the MPI built-in function `MPI_Allreduce`.

4. The time-step index m is decreased by one and the process repeated from the step 1.

Pre-computing the values $\varphi_i^{\Delta t}(\sqrt{2} r_k)$, one identifies that (24) constitutes an explicit sparse-matrix–vector multiplication to update Φ_i^m. Pre-computing the maps, the scheme is fast and avoids the expensive repetitive evaluations of the background magnetic field and plasma profiles required in the FMC scheme. Different choices of the quantities to be pre-computed lead to different memory and arithmetic performances, however. Pre-computing the expressions $\varphi_i^{\Delta t}(\sqrt{2} r_k)$ requires $32 \times K$ maps to be stored for each mesh node $\boldsymbol{z}_i$, with K the number of points in the Gauss-Hermite integration. Treating instead the map as $\varphi_i^{\Delta t}(r)$, i.e., allowing it to depend on the r as a linear function, requires storing only 32×2 maps per mesh node and likely leads to higher arithmetic intensity.

C. Numerical diffusion, importance sampling

The introduction of interpolation to evaluate the probability at an arbitrary phase-space point results in small error every time the interpolant is evaluated: markers that normally would continue from the position they end up after a time step, are "rewound" back to the mesh nodes and the probability transmitted via the interpolation. This results in numerical diffusion effect in the probability evolution. The underlying issue is effectively the same as in using the so-called semi-Lagrangian methods to solve pure advection equations. In the current case, the issue is emphasized by the low collisionality, which nevertheless is critical to obtain, e.g., a slowing-down distribution. Given the recursive nature of the algorithm, this

error will accumulate over the time steps and may end up affecting the pure BMC solution, even drastically.

While this issue of numerical diffusion cannot be erased completely, its effect can be reduced by increasing the time step Δt to reduce the number of interpolant evaluations. On the other hand, increasing the time step will decrease the precision of the integrator, i.e., the mapping of the mesh nodes. For example, the pitch update (21) sets an upper limit to the time step of the stochastic part of the Euler-Maryama equation for the numerical error to remain acceptable:

$$\frac{D_\perp}{v^2}\Delta t \lessapprox 1. \tag{25}$$

This limitation has been checked in practical simulations making sure the pitch error is small enough. For example, in the test scenarios that are described in the following section, a time step in the range of $1 \times 10^{-4}s$–$1 \times 10^{-3}s$ usually produces good results.

With such limit set, the precision of the deterministic part, namely the push along the Hamiltonian flow $\varphi_{\boldsymbol{U}}^{\Delta t}$ and the friction term in (20), is then recovered by sub-cycling, i.e. by dividing further the time step into N shorter steps and pushing the markers through the deterministic part only. Since the deterministic part is computed only once per mesh node, this sub-stepping has insignificant impact on the global computational time. In fact, each marker is followed forward in time through its equations of motion for a total N minor time steps. For what has just been said, an optimal time step Δt is expected to exist, being a trade-off between the numerical diffusion error and the error of the stochastic collision term.

If the efforts to mitigate the numerical diffusion are not having sufficient effect, the final possibility, the one this paper exploits, is to use the possibly inaccurate, yet informative estimate of the probability Φ as *a priori* information and to perform importance sampling of the birth marker distribution on the basis of this information. As the BMC solution can only smooth out away from the correct solution, it should not create artificial peaks of probability. As such, it should form a robust way to sample the marker-particle population for initializing the FMC method, which then doesn't suffer from numerical diffusion and can be performed exactly in the same manner as standard FMC simulations. When the BMC and FMC are combined, we refer to the scheme as the Backward Monte Carlo with Importance Sampling or BMC-IS for short. The re-sampling is performed by first evaluating the BMC hit probability at the existent marker position. Each marker of the birth distribution is then assigned a new weight based on the BMC probability, and finally the weights are normalized in order to respect the total desired number of markers. The weights need to be rounded to the closest integer values and the markers are duplicated accordingly. This means that markers with a weight that is too low are completely discarded. The development of a more refined and production ready algorithm able to sample birth markers continuously in phase-space is left for a future study.

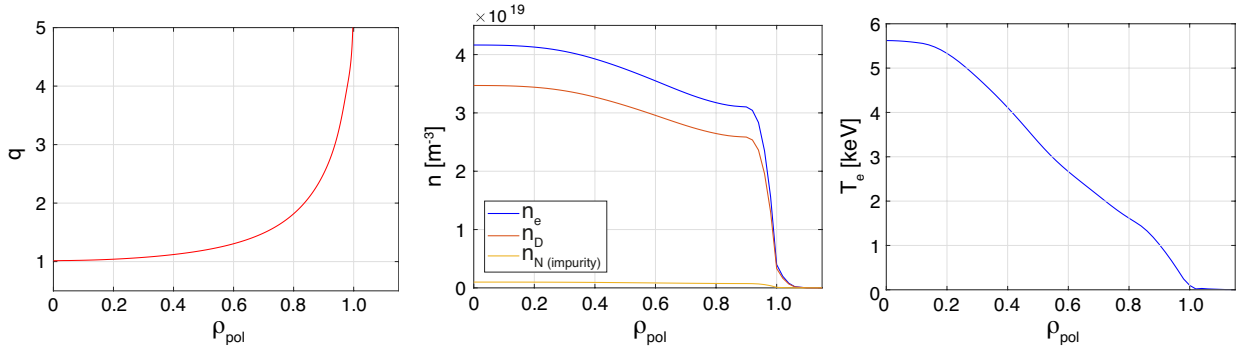
Next we demonstrate how the BMC-IS method can provide significant statistics improvements compared to only performing the standard FMC simulations.

IV. TEST CASES FOR THE NEW ALGORITHM

In this section, the BMC scheme is applied to an ASDEX Upgrade (AUG) H-mode discharge where the fast-ion source is introduced by means of the neutral beam injection (NBI) system. In particular, the magnetic equilibrium and kinetic profiles used for this test

correspond to the shot AUG#33143, with a high $\beta_N = 2.4$, low collisionallity $\nu^* = 0.2$ and toroidal magnetic field $B_t = -1.8$T. Figure 2(a) shows the safety factor profile and figure 2(b) presents the density profiles of electrons and ions. For this case, the ion species included are D as the main component and N as the impurity ions using an effective charge of $Z_{eff} = 2$. Figure 2(c) shows the electron temperature profile and it has been assumed that the temperature of ions and electrons is the same.

The AUG NBI system consists of 8 beams providing different injection geometries and energies. For this analysis, the birth distributions of beams Q7 and Q8 were calculated with the BBNBI code [11] assuming a D injection of energy 93 keV and power 2.5 MW for each beam.



a) Safety factor profile b) Density profiles c) Electron temperature profile

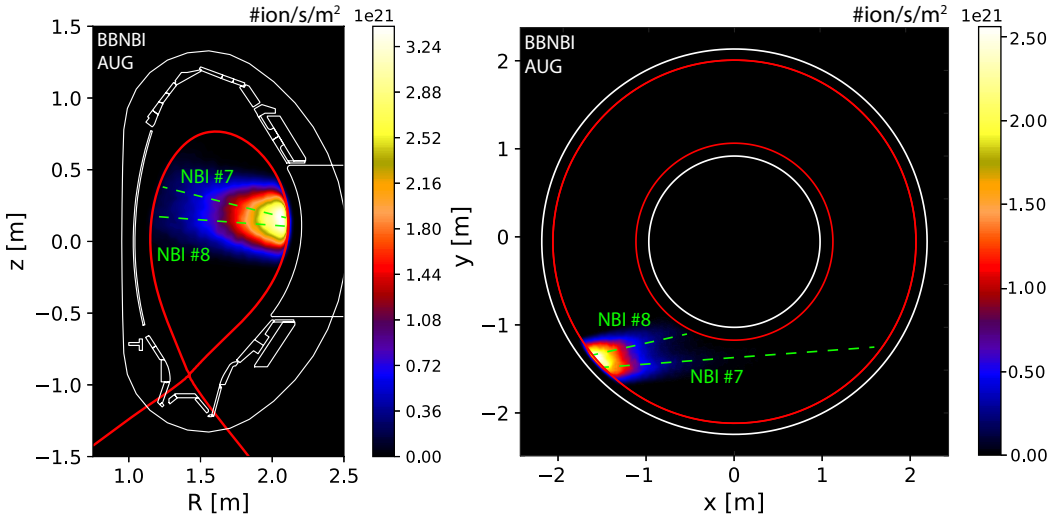
FIG. 2. Plasma profiles including safety factor (a); electron, main ion species and impurity densities (b) and electron density (c) as a function of the normalized poloidal flux coordinate ρ_{pol}.

The birth distribution, corresponding to the neutral beams Q7 and Q8 is illustrated in Fig. 3. In particular, the beam Q8 is tested with a target domain corresponding to the entire wall, while the beam Q7 is tested with a target domain corresponding to a small subset of the wall located near the divertor. Both cases are equally interesting as they allow estimates of the total loss of fast ions and the numerical simulations of specific fast-ion loss detectors. The mesh physical ranges are kept fixed and they are listed in Table I, while the number of mesh nodes has been varied to reveal the effect of the mesh density on the end result.

	min	max
R	0.75m	2.5m
ϕ	0	2π
z	-1.5m	1.5m
$p_\parallel$	-1×10^{-20}kg m s^{-1}	6×10^{-21}kg m s^{-1}
$p_\perp$	1×10^{-24}kg m s^{-1}	1×10^{-20}kg m s^{-1}

TABLE I. Mesh parameters.

For each numerical test, the BMC-scheme is used to compute the target hit probability on a regular rectangular mesh, using a fifth order Gauss-Hermite integration, and the result is compared with the probability generated by the standard FMC-method with markers placed on the same mesh. A convergence analysis is presented at the end of this section to assess precisely how the BMC scheme differs from a true solution and how it converges



a) NBI density (poloidal cross section) b) NBI density (toroidal cross section)

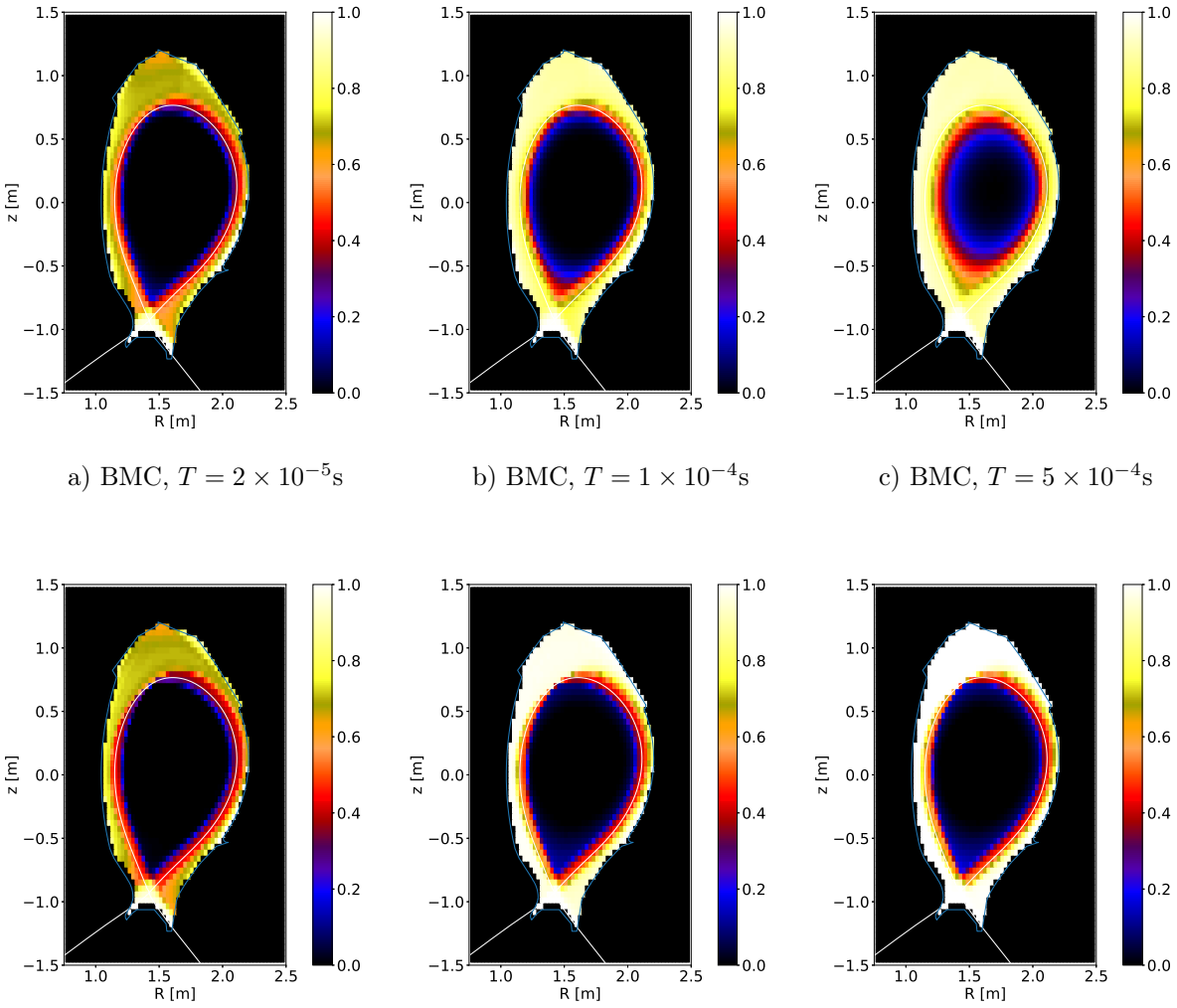
FIG. 3. 2D histograms showing the surface density of the birth distribution of beams Q7 and Q8 projected onto the poloidal (a) and toroidal (b) cross sections. The red lines indicate the separatrix and the white lines the vacuum vessel.

with respect to the mesh density. Once this comparison is done. The information obtained from the BMC calculation is utilized in importance sampling of the birth distribution and FMC performed with respect to the new marker-particle distribution. The statistics of this combined BMC-IS calculation are then compared with the FMC scheme with respect to the number of markers reaching the target and the resulting velocity distributions at the target.

In the examples studied below, the time-independent version of the BMC scheme is used, which, as explained in the previous section, requires the computation of one major time step only and is trivially fast for the remaining steps. As long as the mesh size is reasonable, the running time of the BMC part in the BMC-IS scheme is negligible compared to the main FMC integration.

A. Neutral beam Q7: entire wall target domain

The first example considers the fast-ion birth distribution originating from the Q7 neutral beam, with the target domain being the entire wall. Figure 4 illustrates the probability, averaged over the velocity space and toroidal direction, for a particle to reach the target domain from any phase-space point on the mesh during the given time T. The same probability is then computed with the standard FMC scheme by placing markers on the same mesh and tracking their evolution. The major time step $\Delta t = 4 \times 10^{-6}$s and the number of sub-cycles 40 are chosen for the BMC. The two methods are in good agreement until about $T = 1 \times 10^{-4}$s when the diffusion error in the BMC method eventually appears and critically changes the solution in comparison to the FMC solution. This happens despite the relatively high-mesh density of 64 nodes in each of the five dimensions.



a) BMC, $T = 2 \times 10^{-5}$s b) BMC, $T = 1 \times 10^{-4}$s c) BMC, $T = 5 \times 10^{-4}$s

d) FMC, $T = 2 \times 10^{-5}$s e) FMC, $T = 1 \times 10^{-4}$s f) FMC, $T = 5 \times 10^{-4}$s

FIG. 4. Comparison of the probability from BMC and FMC for different total times T with target domain set as the entire wall. Major time step is $\Delta t = 4 \times 10^{-6}$s with 40 subcycles. Grid size is 64 nodes in all dimensions. The illustrated quantity is the mean value of Φ over the $\phi, p_{\|}$ and $p_{\perp}$ dimensions.

Regardless of the apparent error in the BMC wall-hit probability, the combined approach of BMC-IS does succeed in improving the wall-hit statistics as demonstrated in the Table II and Fig. 5. As expected, the improvement increases with increasing mesh density. Table II also displays the behaviour of the BMC solution as the total time for particles to reach the target is increased: for short periods, the BMC provides an accurate *a priori* information for sampling while for longer periods the information suffers from numerical diffusion and begins to lose its value. From Table II, it is also evident that increasing the mesh density clearly is able to address some of the deterioration as the numerical diffusion is reduced: between $T = 1.2 \times 10^{-5}$s and $T = 5 \times 10^{-4}$s the number of wall hits with a mesh containing

128^5 nodes is reduced by a factor of 4 whereas with a mesh containing only 16^5 nodes a reduction by a factor of 50 occurs.

In Table II, the number of wall hits computed with the FMC shows a saturation tendency from $T = 2 \times 10^{-5}$s. This behaviour can be explained by comparing the total simulation times with the bounce period of the ions, which is in the range of $1.0 - 2.5 \times 10^{-5}\, s$. This comparison indicates that most of the escaping ions are prompt losses and that longer simulation times are required to observe collision effects.

Finally, Figure 6 shows the energy–pitch histogram of markers of the main NBI source hitting the wall for FMC and for the BMC-IS using 64 points per dimension, with a total integration time of $T = 5 \times 10^{-4}$s and a total of 10^5 input markers. The results agree qualitatively, with the FMC simulation resulting in somewhat more noise. It is important to note that even though the BMC estimate for the probability maybe off, this does not pollute the final signal computed with the BMC-IS scheme.

	Mesh elements	T=1.2×10^{-5}s	T=2×10^{-5}s	T=1×10^{-4}s	T=5×10^{-4}s
FMC	—	410	463	464	467
BMC-IS	16^5	13699	7744	1106	685
BMC-IS	32^5	45135	34472	5419	984
BMC-IS	64^5	70047	63545	24130	3759
BMC-IS	128^5	81968	81281	51968	18580

TABLE II. Number of wall hits as function of T for AUG NBI Q7 configuration, using 10^5 input markers. Major time step for BMC is $\Delta t = 4 \times 10^{-6}$s with 40 sub-cycles.

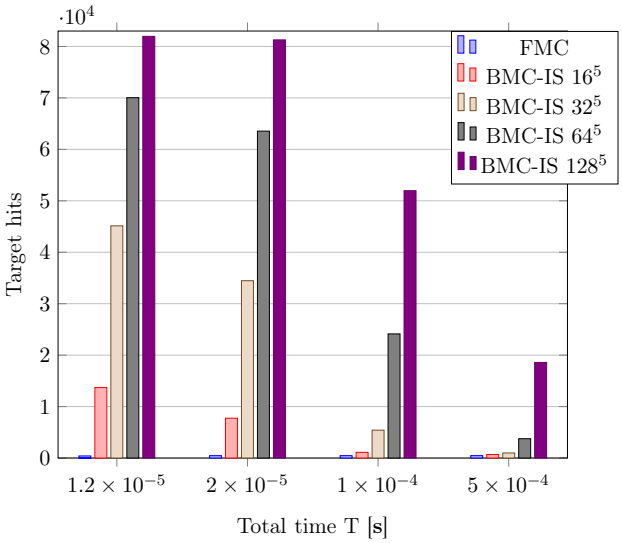
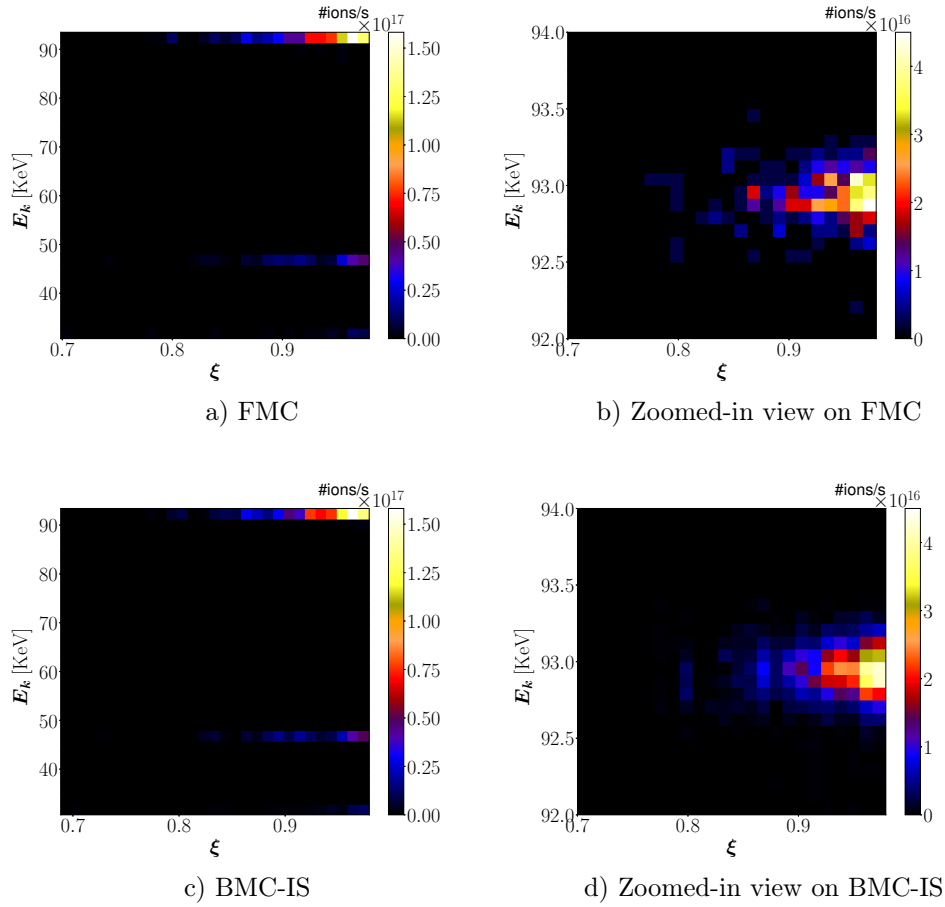


FIG. 5. Histogram of the number of wall hits as recorded in Table II.

B. Neutral beam Q8: small target domain

The second test considers the fast-ion birth distribution originating from the Q8 neutral beam, and the target domain consists of a small strip of the wall near the divertor, having



a) FMC b) Zoomed-in view on FMC

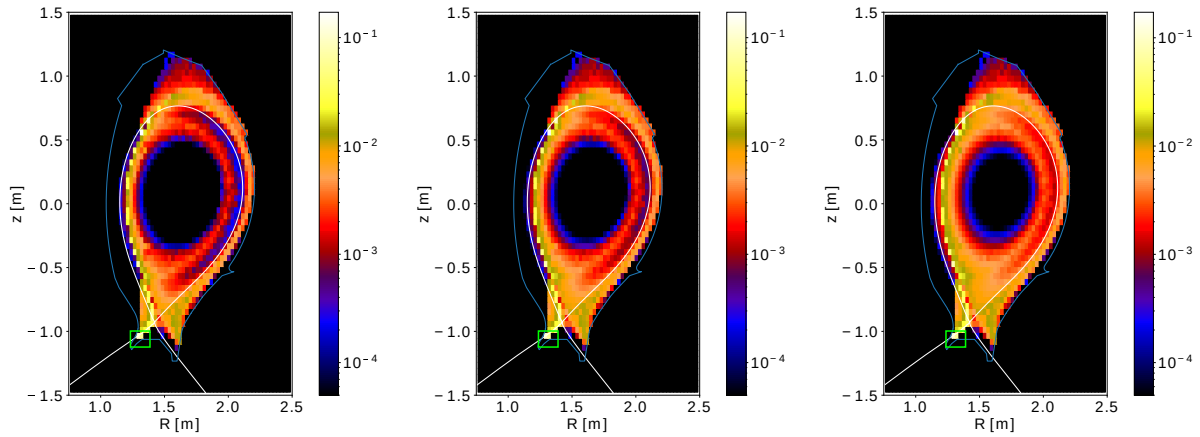
c) BMC-IS d) Zoomed-in view on BMC-IS

FIG. 6. Energy–pitch histogram for the FMC and BMC-IS schemes with the AUG NBI Q7 configuration. The number of input markers used is 10^5 in both cases with total time $T = 5 \times 10^{-4}s$. BMC-IS is using a 64^5 mesh and major time step $\Delta t = 4 \times 10^{-6}s$

coordinates defined by the intervals $R \in [1.2343\text{m}, 1.2355\text{m}]$ and $z \in [-1.1026\text{m}, -1.1073\text{m}]$, and spanning the entire torus. The purpose of this test is to illustrate what happens when the FMC scheme has difficulties accumulating enough target hits and the limits of the BMC-IS scheme are tested by extending the simulation time significantly.

For the Q8 case, the target hit probability from the BMC scheme is illustrated in Fig. 7, up to a total integration time $T = 2 \times 10^{-2}$s. The simulations were run with a major time $\Delta t = 8 \times 10^{-4}$s with 8000 syb-cycling steps to compute the deterministic motion, and the mesh includes 64 nodes in each coordinate direction, similarly as in the earlier Q7 case. In contrary to the Q7 case, here the significantly longer total time prevented performing the direct computation of the probability with the pure FMC method: the computational costs would have been too high for running a number of markers in the order of 10^{10} possibly up to $2 \times^{-2}$ s with a time step of 1×10^{-7}s.

The recorded hit-rate statistics are summarized in Table III and visualized in Fig. 8. The BMC-IS provides larger number of recorded target hits than the pure FMC, for all mesh densities. Similarly as in the case of Q7, the number of target hits again decreases with respect to increasing the total time for the particles to reach the target. The advantage from the BMC-IS scheme nevertheless remains significant. The reduction over time is not as



a) BMC, $T = 5 \times 10^{-3}$s b) BMC, $T = 1 \times 10^{-2}$s c) BMC, $T = 2 \times 10^{-2}$s

FIG. 7. Comparison of the probability from BMC and FMC for different total times T with a small target domain, positioned inside the green box. Major time step is $\Delta t = 8 \times 10^{-4}$s with 8000 sub-cycles. Grid size is 64 nodes in all dimensions. The illustrated quantity is the mean value of Φ over the $\phi, p_{\|}$ and $p_{\perp}$ dimensions.

severe as in the Q7 case, which is likely credited to there being more losses from neoclassical physics than in the case of Q7 with a significantly shorter total simulation time and losses being primarily from the first orbits. This is reflected also in the number of target hits recorded by the FMC now clearly increasing over time.

	Mesh elements	T=5×10^{-3}s	T=1×10^{-2}s	T=2×10^{-2}s
FMC	-	11	26	65
BMC-IS	16^5	517	395	348
BMC-IS	32^5	837	575	562
BMC-IS	64^5	1488	1071	885
BMC-IS	128^5	5217	3445	2128

TABLE III. Number of hits on the small target as a function of T for FMC and BMC-IS in AUG NBI Q8 configuration, using 10^5 input markers. Major time step is $\Delta t = 8 \times 10^{-4}$s with 8000 sub-cycles.

Although we do not have the direct comparison of the target hit probabilities produced with BMC and FMC available due to the extensive computational cost of the FMC, the quality of the BMC-IS scheme can still be assessed from the energy–pitch histograms of the recorded target hits, similarly as was performed for the Q7 case. The energy–pitch histograms produced with 10^5 input markers in this manner are illustrated in Figure 9. From the colors in the figure, one can observe that the FMC effectively manages to record only one or two markers per histogram cell (see the zoomed-in view), essentially making it rather difficult to consider the results to enjoy statistical accuracy. This is made more apparent by the zoomed-in view. The BMC-IS, on the other hand, clearly performs better, also agreeing qualitatively with the results from the FMC.

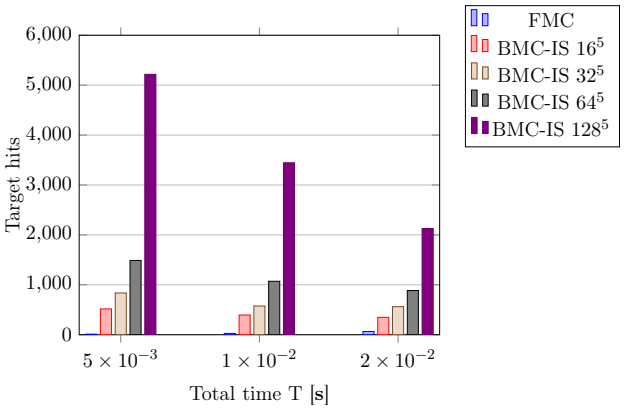


FIG. 8. Histogram of the wall hits recorded in Table III.

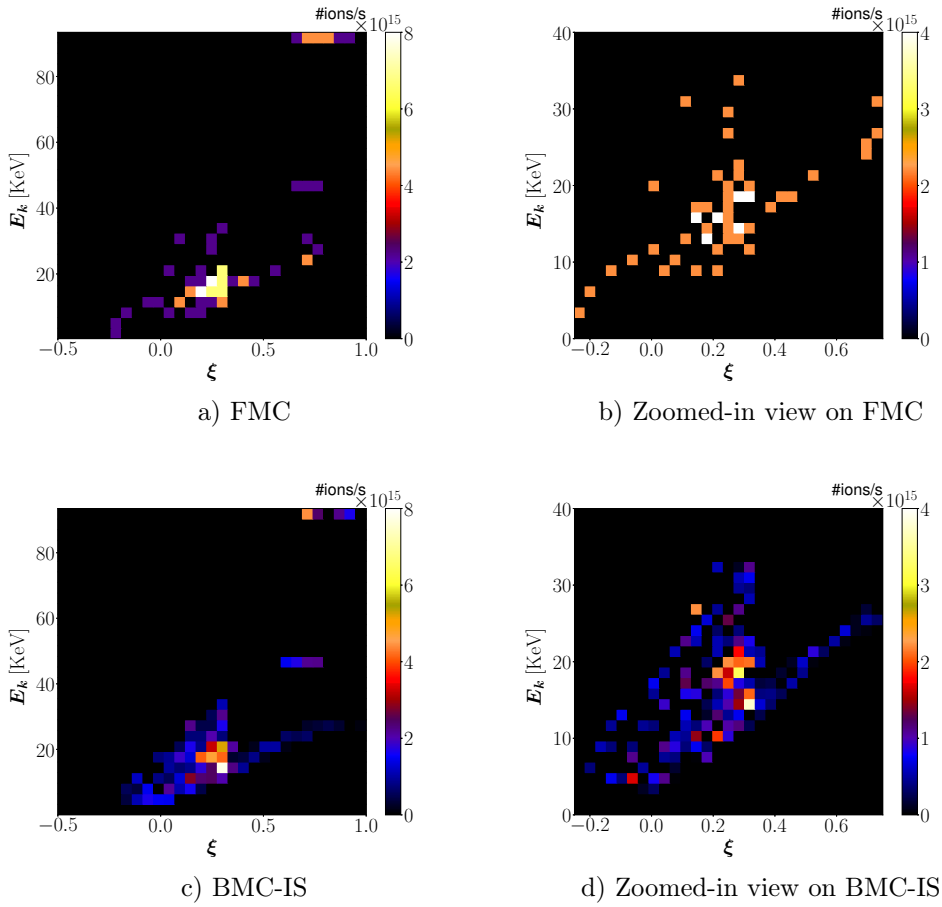


FIG. 9. Energy–pitch histogram for the FMC and BMC-IS schemes with the AUG NBI Q8 configuration. The number of input markers used is 10^5 in both cases with total time $T = 2 \times 10^{-2} s$. BMC-IS is using a 64^5 mesh and major time step $\Delta t = 8 \times 10^{-4} s$

C. Convergence test

In order to check quantitatively the consistency of the BMC scheme, the target hit probability has been computed for different mesh densities by recursively refining an initial sparse mesh, so that all the nodes of the initial mesh are common to the subsequent ones. The mean error of each mesh can be computed by comparing the mean absolute difference on the common nodes. The initial mesh is also small enough that the probability can be computed easily to high precision also with the FMC scheme which can therefore be used as a reference solution. Figure 10 shows the resulting mean error of each mesh refinement after 5 major time steps with respect to a benchmark FMC solution. The error is consistently reduced by increasing the mesh density, indicating that the probability from the BMC calculation appears to converge consistently towards the reference solution.

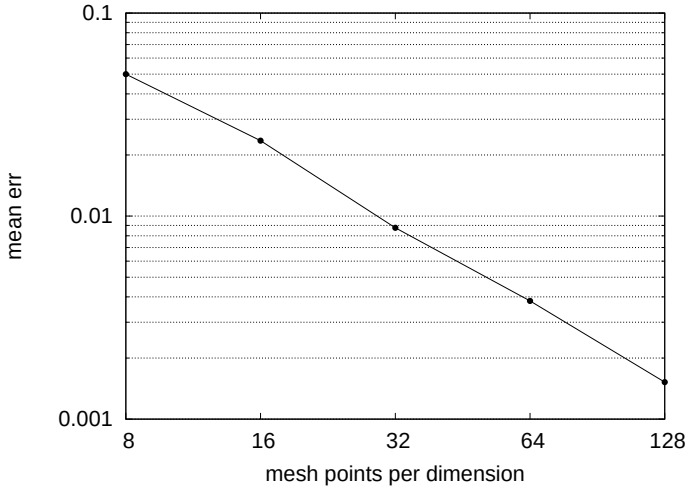


FIG. 10. Convergence test for the BMC scheme.

V. SUMMARY AND CONCLUSIONS

This paper introduced a novel Backward Monte Carlo scheme to improve statistics in simulations of fast-ion losses, especially to small targets. The method was implemented within the ASCOT suite of codes and tested in realistic AUG configurations using NBI particle distributions. In the test cases considered, the approximate target hit probability from the fast Backward Monte Carlo calculation was used to perform importance sampling of the birth marker population. Using the re-sampled markers in Forward Monte Carlo simulations then significantly improved the statistics of target hits in comparison to not using the information from the Backward Monte Carlo calculation. These encouraging results are anticipated to improve further if meshes, that adapt to target and plasma geometry, are explored.

Although not studied here, the model can incorporate the effects of neoclassical tearing modes without any changes and even the Alfvénic perturbations if the time-dependent version is used. In principle, any Fokker-Planck operator could be studied within the BMC formalism. Future work includes, e.g., investigating the possibilities of incorporating ICRH operators into the Backward Monte Carlo framework.

ACKNOWLEDGMENTS

The work of F.Z, L.S, and E.H was supported by the Academy of Finland grant Nos. 320058, 324759, and 315278 respectively. The authors wish to acknowledge CSC – IT Center for Science, Finland, for computational resources. This work has been partially carried out within the framework of the EUROfusion Consortium and has partially received funding from the Euratom research and training programme 2014–2018 and 2019–2020 under grant agreement No. 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

[1] M. Garcia-Munoz, S. Äkäslompolo, P. de Marne, M. G. Dunne, R. Dux, T. E. Evans, N. M. Ferraro, S. Fietz, C. Fuchs, B. Geiger, A. Herrmann, M. Hoelzl, B. Kurzan, N. Lazanyi, R. M. McDermott, M. Nocente, D. C. Pace, M. Rodriguez-Ramos, K. Shinohara, E. Strumberger, W. Suttrop, M. A. Van Zeeland, E. Viezzer, M. Willensdorfer, and E. Wolfrum, "Fast-ion losses induced by ELMs and externally applied magnetic perturbations in the ASDEX Upgrade tokamak," Plasma Physics and Controlled Fusion **55**, 124014 (2013).

[2] M. A. Van Zeeland, N. M. Ferraro, B. A. Grierson, W. W. Heidbrink, G. J. Kramer, C. J. Lasnier, D. C. Pace, S. L. Allen, X. Chen, T. E. Evans, M. García-Muñoz, J. M. Hanson, M. J. Lanctot, L. L. Lao, W. H. Meyer, R. A. Moyer, R. Nazikian, D. M. Orlov, C. Paz-Soldan, and A. Wingen, "Fast ion transport during applied 3D magnetic perturbations on DIII-D," Nuclear Fusion **55**, 073028 (2015).

[3] T. Kurki-Suonio, K. Särkimäki, S. Äkäslompolo, J. Varje, Y. Liu, S. Sipilä, O. Asunta, E. Hirvijoki, A. Snicker, J. Terävä, M. Cavinato, M. Gagliardi, V. Parail, and G. Saibene, "Protecting ITER walls: fast ion power loads in 3D magnetic field," Plasma Physics and Controlled Fusion **59**, 014013 (2017).

[4] G. Bormetti, G. Callegaro, G. Livieri, and A. Pallavicini, "A backward Monte Carlo approach to exotic option pricing," European Journal of Applied Mathematics **29**, 146–187 (2018), arXiv:1511.00848.

[5] G. Zhang and D. del-Castillo-Negrete, "A backward Monte-Carlo method for time-dependent runaway electron simulations," Physics of Plasmas **24**, 092511 (2017), arXiv:1708.00947.

[6] Charles F. F. Karney and Nathaniel J. Fisch, "Current in wave-driven plasmas," The Physics of Fluids **29**, 180–192 (1986).

[7] E. Hirvijoki, "Eliminating poor statistics in Monte-Carlo simulations of fast-ion losses to plasma-facing components and detectors," (2019), arXiv:1905.04952.

[8] E. Hirvijoki, O. Asunta, T. Koskela, T. Kurki-Suonio, J. Miettunen, S. Sipilä, A. Snicker, and S. Äkäslompolo, "ASCOT: Solving the kinetic equation of minority particle species in tokamak plasmas," Computer Physics Communications **185**, 1310–1321 (2014), arXiv:1308.1904.

[9] K. Särkimäki, E. Hirvijoki, and J. Terävä, "Adaptive time-stepping Monte Carlo integration of Coulomb collisions," Computer Physics Communications **222**, 374–383 (2018), arXiv:1701.05043.

[10] E. Hirvijoki, A. Brizard, A. Snicker, and T. Kurki-Suonio, "Monte Carlo implementation of a guiding-center Fokker-Planck kinetic equation," Physics of Plasmas **20**, 092505 (2013), arXiv:1305.2013.

[11] O. Asunta, J. Govenius, R. Budny, M. Gorelenkova, G. Tardini, T. Kurki-Suonio, A. Salmi, and S. Sipilä, "Modelling neutral beams in fusion devices: Beamlet-based model for fast particle simulations," Computer Physics Communications **188**, 33–46 (2015).