J. Plasma Phys. (2026), *vol*. 92, E36 © The Author(s), 2026. Published by Cambridge University Press.
This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited. doi:10.1017/S0022377826101445

An adjoint formulation of energetic particle confinement

Christopher J. McDevitt[1] **and Jonathan S. Arnaud**[1]

[1]Nuclear Engineering Program, Department of Materials Science and Engineering, University of Florida, Gainesville, FL 32611, USA

Corresponding author: Christopher J. McDevitt, cmcdevitt@ufl.edu

(Received 17 December 2025; revision received 18 February 2026; accepted 22 February 2026)

An adjoint formulation of energetic particle confinement in axisymmetric tokamak geometry is derived and evaluated using a physics-informed neural network (PINN). The PINN estimates the mean escape time of energetic ions by solving an inhomogeneous adjoint of the drift kinetic equation with a Lorentz collision operator, yielding predictions of fast ion loss in tokamak geometry due to direct ion orbit loss and collisional transport. To our knowledge, this is the first time a PINN has been used to solve the drift kinetic equation in tokamak geometry, a challenging problem due to the large time scale separation between the rapid transit time of energetic ions and their slow collisional time scale. It is shown that a careful and intentional design of a PINN is able to learn the mean escape time across the majority of the plasma volume, suggesting a path towards constructing a rapid surrogate for use within a broader optimisation framework.

Key words: fusion plasma, plasma simulation

1. Introduction

Energetic particles are a ubiquitous aspect of magnetic fusion plasmas. Such particles emerge as a result of ion heating schemes including neutral beam injection (Hemsworth, Tanga & Antoni 2008) and ion cyclotron resonance heating (ICRH) (Adam 1987; Kazakov *et al.* 2017). For a burning deuterium–tritium plasma, the confinement of energetic alpha particles is critical for the sustainment of the fusion plasma. In addition, when such particles escape, they can cause substantial damage to plasma-facing components (Fasoli *et al.* 2007; Bonofiglo *et al.* 2024). Obtaining an understanding of the confinement time of such particles is thus vital to the operation of a magnetic fusion device.

In the present work, an initial step towards a framework through which the confinement time of an energetic ion or electron can be efficiently estimated is developed. While evaluating the confinement time of an energetic particle is often of greatest importance for non-axisymmetric magnetic fields arising from error fields (Goldston, White & Boozer 1981), magnetic islands (Carolipio *et al.* 2002), energetic

particle modes (Heidbrink & White 2020), or for a three-dimensional (3-D) stellarator equilibrium (Galeev *et al.* 1969), this work will consider the simpler problem of an axisymmetric tokamak equilibrium. Our motivation for considering this simpler problem is to investigate the potential for evaluating metrics of energetic particle transport via the solution to an inhomogeneous adjoint problem. As described later, the inhomogeneous adjoint of the drift kinetic equation provides a direct means of computing the average confinement time of an ion or electron as a function of the particle's initial phase space position (Särkimäki 2020), a quantity we will refer to as the mean escape time. While this adjoint problem can be solved by a variety of numerical methods, we will consider a physics-informed neural network (PINN) (Raissi, Perdikaris & Karniadakis 2019) in the present work. Physics-informed neural networks correspond to a machine learning approach capable of integrating both physics and data into the training of a neural network. A distinguishing feature of a PINN is that it can be extended to learn the parametric dependence of the solution of a partial differential equation (PDE; Sun *et al.* 2020). Noting that optimising the transport properties of energetic particle orbits often requires exploring a broad range of magnetic field configurations, obtaining a rapid surrogate capable of computing quantities of interest such as the mean escape time of energetic particles provides a powerful tool for assessing energetic particle transport.

Our aim in the present paper is not to provide such a parametric solution, but to instead investigate the ability of a PINN to solve the drift kinetic equation for an energetic ion population. The extension to a parametric solution will be treated in a future work. Specifically, while PINNs have previously been used to solve the relativistic (Arnaud, Mark & McDevitt 2024; McDevitt, Arnaud & Tang 2025; Arnaud, Tang & McDevitt 2025) and non-relativistic (McDevitt & Tang 2024) Fokker–Planck equations, the large time scale separation between the rapid bounce or transit times and slow collision time of energetic particles in tokamak geometry represents a far more challenging problem. Using advanced optimisation algorithms (Kiyani *et al.* 2025; Urbán *et al.* 2025; Wang *et al.* 2025), we aim to show that PINNs are able to robustly resolve several features of this challenging system, even in the absence of data. As described further later, as the energy of the ions is increased, leading to a greater time scale separation between the fast bounce/transit motion and slow collision time of energetic ions, our implementation of a PINN struggles to quantitatively describe the escape time of the best confined ion orbits, although it does resolve the solution in the outer region of the tokamak including phase space regions of prompt ion loss (Hinton & Chu 1985).

As an alternative means of evaluating the fast ion mean escape time, we introduce a new drift-kinetic solver based on the JAX framework. While in the present work, this solver is only used to verify solutions from the PINN, future work will leverage particle data generated by this drift-kinetic solver to help accurately resolve the mean escape time of well-confined ion orbits, along with facilitating the training of the PINN over a broad range of plasma parameters. This extended PINN framework, trained using both physics and particle data, would thus provide an efficient surrogate for fast ion confinement appropriate for use inside optimisation loops. The remainder of this paper is organised as follows. Section 2 derives an adjoint problem for the mean escape time of an energetic particle. Brief overviews of the physics-informed deep learning framework and JAX particle-based solver are given in § 3. The mean escape time of an energetic ion species is computed in § 4. Conclusions and a discussion of future directions are given in § 5.

https://doi.org/10.1017/S0022377826101445 Published online by Cambridge University Press

2. Adjoint of steady-state drift kinetic equation

In this section, we will define an adjoint problem based on the inhomogeneous steady-state drift kinetic equation. Our starting point will be to define the Green's function $F_s(z; z_0)$ for the steady-state drift kinetic (Helander & Sigmar 2002):

$$\nabla \cdot \left(\dot{\boldsymbol{X}} F_s\right) + \frac{\partial}{\partial v_\parallel}\left(\dot{V}_\parallel F_s\right) - C_s\left(F_s\right) = \delta\left(z - z_0\right), \tag{2.1a}$$

where $z = (\boldsymbol{x}, v_\parallel, \mu)$ is the phase space position of the particle and the characteristic equations are given by

$$\dot{\boldsymbol{X}} = v_\parallel \hat{\boldsymbol{b}} + \boldsymbol{v}_{EB} + \boldsymbol{v}_{Ds}, \tag{2.1b}$$

$$\dot{V}_\parallel = \frac{q_s}{m_s}\hat{\boldsymbol{b}} \cdot \boldsymbol{E} - \mu \hat{\boldsymbol{b}} \cdot \nabla B. \tag{2.1c}$$

Here, the subscript s is the species index, $\mu \equiv v_\perp^2/(2B)$ is the magnetic moment, $v_\parallel$ is the velocity parallel to the magnetic field, $C_s(F_s)$ is a collision operator, taken to be a Lorentz collision operator

$$C_s\left(F_s\right) = \frac{\nu_D}{2}\frac{\partial}{\partial \xi}\left[\left(1 - \xi^2\right)\frac{\partial F}{\partial \xi}\right], \tag{2.2}$$

and we have defined the $\boldsymbol{E} \times \boldsymbol{B}$ and magnetic drifts as

$$\boldsymbol{v}_{EB} = \frac{\boldsymbol{E} \times \boldsymbol{B}}{B^2}, \tag{2.3a}$$

$$\boldsymbol{v}_{Ds} = \frac{\left(v_\parallel^2 + \mu B\right)}{\omega_{cs}}\hat{\boldsymbol{b}} \times \nabla \ln B. \tag{2.3b}$$

The magnetic drifts have been simplified by the approximation $\nabla \times \hat{\boldsymbol{b}} \approx \hat{\boldsymbol{b}} \times \nabla \ln B$, appropriate in the limit of a 'vacuum plasma'. As boundary conditions, we will enforce $F_s = 0$ at the spatial boundaries whenever $U_r \equiv \boldsymbol{r} \cdot \dot{\boldsymbol{X}} < 0$, where $\hat{\boldsymbol{r}}$ is the radial unit vector and we have assumed a circular spatial boundary, but leave F_s unconstrained for $U_r \equiv \boldsymbol{r} \cdot \dot{\boldsymbol{X}} \geqslant 0$, i.e.

$$F_s\left(r = r_{wall}\right) = \begin{cases} 0, & U_r < 0, \\ \text{unconstrained}, & U_r \geqslant 0, \end{cases} \tag{2.4}$$

where r_{wall} is the location of the surrounding wall, taken as the plasma minor radius a, for simplicity. With these boundary conditions, particles are not allowed to enter the spatial domain, but particles are able to exit. The resulting Green's function can thus be physically interpreted as the steady-state particle distribution resulting from a unit source located at $z = z_0$.

To derive an adjoint problem for the mean escape time of an energetic particle, we begin by multiplying the left-hand side of (2.1*a*) by the solution to the adjoint equation T_s, which after successive integrations by parts, yields

$$\begin{aligned} &\int \mathrm{d}^3 v \mathrm{d}^3 x T_s \left[\nabla \cdot \left(\dot{\boldsymbol{X}} F_s\right) + \frac{\partial}{\partial v_\parallel}\left(\dot{V}_\parallel F_s\right) - C_s\left(F_s\right)\right] \\ &\quad = \int \mathrm{d}^3 v\, \mathrm{d}^3 x F_s \left[-\dot{\boldsymbol{X}} \cdot \nabla T_s - \dot{V}_\parallel \frac{\partial T_s}{\partial v_\parallel} - C_s^*\left(T_s\right)\right] \\ &\qquad + \int \mathrm{d}^3 v \int \mathrm{d}\Gamma\, \left[U_r T_s F_s\right]|_{r = r_{wall}}. \end{aligned} \tag{2.5}$$

https://doi.org/10.1017/S0022377826101445 Published online by Cambridge University Press

Here, the first term on the right-hand side is the Green's function times the adjoint of the steady state drift kinetic equation, and the second term on the right-hand side describes fluxes through the spatial boundary Γ. The boundary condition on the Green's function F_s (2.4) implies the spatial boundary term will vanish when $U_r < 0$, where setting the boundary conditions on T_s when $U_r \geqslant 0$ allows for distinct adjoint problems to be defined.

An adjoint problem for the mean escape time (Grasman *et al.* 2013; Liu *et al.* 2017) of an energetic particle can be derived by considering the inhomogeneous adjoint equation

$$\dot{\boldsymbol{X}} \cdot \boldsymbol{\nabla} T_s + \dot{V}_{\parallel} \frac{\partial T_s}{\partial v_{\parallel}} + C_s^* (T_s) = -1. \tag{2.6}$$

Substituting (2.6) along with the steady-state Green's function defined by (2.1*a*) into the right- and left-hand sides, respectively, of (2.5), yields

$$T_s (\boldsymbol{x}_0, \boldsymbol{v}_0) = N_s + \int \mathrm{d}\Gamma \int \mathrm{d}^3 v \, [U_r T_s F_s]|_{r_{wall}}, \tag{2.7}$$

where Γ is the surface of the surrounding wall and $N_s \equiv \int \mathrm{d}^3 v \int \mathrm{d}^3 x F_s$ is the number of particles remaining in the domain due to the unit source on the right-hand side of (2.6). At the spatial boundary we will take

$$T_s (r = r_{wall}) = \begin{cases} \text{unconstrained}, & U_r \leqslant 0, \\ 0, & U_r > 0. \end{cases} \tag{2.8}$$

This boundary condition follows since a particle born on the spatial boundary with an outward directed velocity would be lost immediately, such that $T_s = 0$, however, for an inward directed velocity, this particle would not be immediately lost. For the boundary condition defined by (2.8) and noting that $F_s = 0$ for $U_r < 0$, (2.4) yields

$$T (\boldsymbol{x}_0, \boldsymbol{v}_0) = N_s. \tag{2.9}$$

Further noting that N_s will scale directly with the confinement time of the particles, $T_s(\boldsymbol{x}_0, \boldsymbol{v}_0)$ will act as a measure of how well particles injected at a given phase space location are confined on average, a quantity often referred to as the mean escape time or expected exit time (Grasman *et al.* 2013). We note that for the present example, all ions will escape for $t \to \infty$, such that the mean escape time will be finite throughout the domain.

3. Physics-constrained deep learning framework

The primary focus of the remainder of this work will be the solution of the adjoint problem derived in § 2. A brief overview of our implementation of a PINN is given in § 3.1, where the reader is referred to Cuomo *et al.* (2022), Wang *et al.* (2023), Toscano *et al.* (2025) for additional information on this rapidly expanding area. In addition, a newly developed GPU accelerated particle-based solver for the drift kinetic equation is described in § 3.2. For the present work, this particle-based solver will be used to validate the PINN's solutions to the drift kinetic equation.

3.1. *Physics-informed neural networks*

Physics-constrained machine learning methods (Karniadakis *et al.* 2021) seek to incorporate physical constraints into the training of a neural network. Physics-informed neural networks correspond to a prominent example of this area. An

https://doi.org/10.1017/S0022377826101445 Published online by Cambridge University Press

attractive feature of PINNs is that they can be used to solve PDEs and ordinary differential equations (ODEs) directly in the absence of data (van Milligen, Tribaldos & Jiménez 1995; Raissi *et al.* 2019). They also provide a mechanism through which physical constraints can be used to enrich sparse data sets (Cai *et al.* 2021; Mathews *et al.* 2022; McDevitt, Fowler & Roy 2024). In its simplest form, the loss function of a PINN can be written as

$$\text{Loss} = \frac{1}{N_{data}} \sum_{i=1}^{N_{data}} \left[T_i - T\left(\boldsymbol{z}_i; \boldsymbol{\lambda}_i\right)\right]^2 + \frac{w_{PDE}}{N_{PDE}} \sum_{i=1}^{N_{PDE}} \mathcal{R}^2\left(\boldsymbol{z}_i; \boldsymbol{\lambda}_i\right), \tag{3.1}$$

where T is the quantity we seek to predict, the mean escape time for the current work; $\mathcal{R}$ is the residual of the governing PDE; w_{PDE} is a scalar weight applied to the PDE bias of the loss; $\boldsymbol{z} = (\boldsymbol{x}, \boldsymbol{v})$ are the independent variables; and $\boldsymbol{\lambda}$ describes physical parameters, such as collisionality or inverse aspect ratio. The first term represents a bias due to data, where the data may be boundary conditions, or synthetic or experimental data. The second term is a bias due to the governing equations, which, for the present work, will be the inhomogeneous adjoint of the steady-state drift kinetic equation. In the absence of synthetic or experimental data, the first term will only contain the boundary conditions. Successfully minimising the loss will thus correspond to identifying a solution to the PDE while satisfying the boundary conditions. Unlike traditional PDE solvers, a parametric solution to a PDE can be straightforwardly obtained by training across a broad range of physics parameters $\boldsymbol{\lambda}$. Thus, while the training time of such a model may be substantially longer than a traditional PDE solver, once trained, the fast inference time of a PINN allows for fast online prediction times across a broad range of plasma conditions. This property can be used for the development of fast surrogate models, where some recent examples include runaway electron generation (McDevitt 2023; Arnaud *et al.* 2025), MHD equilibrium (Jang *et al.* 2024), plasma thrusters (Luo *et al.* 2025) or non-thermal ion distributions (McDevitt & Tang 2024).

Key elements in obtaining accurate results from PINNs correspond to (i) the identification of an appropriate neural network architecture, (ii) the phase space weighting of the residual and (iii) the selection of an efficient optimisation algorithm for minimising the loss. Regarding item (i), we will employ a fully connected feedforward neural network, but with an 'output layer' that enforces positivity as a hard constraint. In so doing, this both prevents certain unphysical solutions from being predicted by the PINN and, perhaps more importantly, narrows the space of solutions that the optimiser searches across. Defining T_{NN} to be the output of the hidden layers of the neural network, a simple transformation that enforces positivity is given by

$$T_s = \exp\left(T_{NN}\right). \tag{3.2}$$

In addition to enforcing positivity, the exponential dependence of T_s on the output of the hidden layers T_{NN} provides a natural means of capturing the large disparity in escape times for energetic ions. Specifically, ions born near the plasma core are expected to be well confined, where they will only be lost by the relatively slow process of collisional transport, in contrast to those born near the edge, which may be lost on a very short time scale by direct orbit loss.

An additional challenge that arises is ensuring that the PINN is able to robustly connect the boundary condition (2.8) applied at the edge of the tokamak to the solution throughout the core plasma. Noting that we expect the escape time T_s to be

https://doi.org/10.1017/S0022377826101445 Published online by Cambridge University Press

small near the edge, training points in this edge region can be given greater emphasis by weighting the residual of the PDE by

$$\mathcal{F}(z;\boldsymbol{\lambda}) = \frac{A}{A + T_s(z;\boldsymbol{\lambda})}, \tag{3.3}$$

where A is a constant, such that the loss can be written as

$$\text{Loss} = \frac{1}{N_{bdy}} \sum_{i=1}^{N_{bdy}} \left[T_{s,i} - T_s(z_i;\boldsymbol{\lambda}_i)\right]^2 + \frac{w_{PDE}}{N_{PDE}} \sum_{i=1}^{N_{PDE}} \left[\mathcal{F}(z_i;\boldsymbol{\lambda}_i)\,\mathcal{R}(z_i;\boldsymbol{\lambda}_i)\right]^2. \tag{3.4}$$

Here, the boundary conditions defined by (2.8) are enforced by the first term, $\mathcal{R}$ is the residual of the inhomogeneous adjoint defined by (2.6) and the choice of A sets how much emphasis is given to the edge region. In particular, noting that near the edge, T_s will be of the order of a transit time, since ions on loss orbits will directly escape from the plasma, if A is chosen to satisfy $A \ll 1$, (3.3) will scale as $\propto 1/T_s$ such that the edge region will be heavily weighted, but the inner region of the plasma with $T_s \gg 1$ will receive minimal weighting. We thus expect this choice to allow for the outer region of the plasma to be accurately described at the expense of poor accuracy in the inner region. In contrast, by choosing a larger value of A, improved accuracy for the inner region can be achieved at the expense of having a looser connection to the edge boundary condition, which is vital for achieving a physically meaningful solution. We have found that values of A between 10 and 100 allow for robust convergence of the PINN, where all results in the present paper are for $A = 100$. In addition, the relative weight of the PDE residual compared with the boundary condition is set by the scalar w_{PDE} weighing the first term in (3.4). We have found $w_{PDE} = 100$ provides relatively rapid convergence of the loss.

The last key component of our PINN implementation will be the choice of an effective optimisation algorithm. While the ADAM optimiser (Kingma & Ba 2014) is used ubiquitously across a range of machine learning applications, this first-order optimisation algorithm often is not able to achieve low losses for challenging PDEs. A common strategy is thus to use ADAM for the first phase of training, after which limited memory BFGS (L-BFGS) (Liu & Nocedal 1989), a second-order optimisation algorithm, is used to more tightly converge the loss. Motivated by Kiyani *et al.* (2025), Wang *et al.* (2025), and Urbán *et al.* (2025) and, we have adopted a different optimisation strategy that we have found to yield substantially improved accuracy. The first optimiser that we apply corresponds to the quasi-second-order optimiser SOAP (Vyas *et al.* 2024). This scalable optimiser exhibits similar computational performance as ADAM, but can yield losses up to an order of magnitude smaller than those achievable by ADAM across a broad range of challenging applications of PINNs (Wang *et al.* 2025). For the second stage of optimisation, we use the self-scaled Broyden (SSBroyden) method (Al-Baali, Spedicato & Maggioni 2014). Like L-BFGS, SSBroyden approximates the Hessian, but introduces additional degrees of freedom into the optimisation algorithm that have yielded substantial improvements in training across a range of complex PDEs (Kiyani *et al.* 2025; Urbán *et al.* 2025). While each training epoch is substantially more computationally demanding compared with ADAM or SOAP for the SciPy implementation we are currently employing, we have found it results in substantially faster convergence of the loss for the adjoint problem treated in this work. The primary limitations of SSBroyden as an optimisation routine is that the SciPy implementation used in the present work

https://doi.org/10.1017/S0022377826101445 Published online by Cambridge University Press

does not efficiently use GPUs or support mini-batching. These limitations substantially increase the training time of the SSBroyden phase of training and limit the number of training points employed.

3.2. *JONTA: a particle-based drift kinetic solver*

Particle-based drift kinetic simulations of fast ion orbits are computationally intensive (White & Chance 1984; Hirvijoki *et al.* 2014) requiring substantial computational resources and efficient software. While this has predominately been done with Fortran or C++ libraries, with GPU acceleration facilitated by libraries such as Kokkos Trott *et al.* (2021), maintaining the support of these solvers can be cumbersome, due to its dependency on external backend libraries to effectively use high-performance computing (HPC) hardware. Perhaps most importantly, given that PINNs are trained on GPUs and libraries with a large user base, this motivates the development of a traditional drift kinetic solver that runs on GPUs and uses the same libraries that are used by PINNs, facilitating the operation of the broader framework that contains an interplay between the PINN and a drift kinetic solver.

With the aforementioned discussion, we present Just anOther fuNcTionAl pusher (JONTA), which is built on the PyTorch (Paszke *et al.* 2019) and JAX (Bradbury *et al.* 2018) libraries that run on GPU accelerated hardware. For the present work, we will use the JAX backend and employ a four stage Runge–Kutta (RK4) integration scheme, which is part of a broader suite containing differentiable integrator methods. Guiding-centre equations are evolved with RK4 and collisions are implemented with a Monte-Carlo operator described later. The version of JONTA used in this manuscript together with the PINN training script are available on github.com/cmcdevitt2/DeepPlasma.

JONTA will be used to solve the guiding centre equations defined by (2.1) together with a Monte Carlo collision operator describing pitch-angle scattering. The guiding centre equations can be written as

$$\dot{X} = \hat{\boldsymbol{x}} \cdot \dot{\boldsymbol{X}}, \tag{3.5a}$$

$$\dot{Z} = \hat{\boldsymbol{z}} \cdot \dot{\boldsymbol{X}}, \tag{3.5b}$$

$$\dot{v} = \frac{\partial v}{\partial v_\parallel}\frac{\mathrm{d}v_\parallel}{\mathrm{d}t} + \frac{\partial v}{\partial \mu}\frac{\mathrm{d}\mu}{\mathrm{d}t} + \frac{\partial v}{\partial \boldsymbol{X}} \cdot \frac{\mathrm{d}\boldsymbol{X}}{\mathrm{d}t}, \tag{3.5c}$$

$$\dot{\xi} = \frac{\partial \xi}{\partial v_\parallel}\frac{\mathrm{d}v_\parallel}{\mathrm{d}t} + \frac{\partial \xi}{\partial \mu}\frac{\mathrm{d}\mu}{\mathrm{d}t} + \frac{\partial \xi}{\partial \boldsymbol{X}} \cdot \frac{\mathrm{d}\boldsymbol{X}}{\mathrm{d}t}, \tag{3.5d}$$

where we have chosen cylindrical coordinates (R, Z), defined $X \equiv R - R_0$ and made the transformation from $(\boldsymbol{x}, v_\parallel, \mu)$ to $(\boldsymbol{x}, \xi, v)$. Assuming an axisymmetric tokamak geometry with circular flux surfaces, and noting the relations $\xi^2 = v_\parallel^2/(v_\parallel^2 + 2\mu B)$ and $v^2 = v_\parallel^2 + 2\mu B$, (3.5) can be written as

$$\dot{X} = \hat{\boldsymbol{x}} \cdot v_\parallel \hat{\boldsymbol{b}} = -\xi v \frac{Z}{q R_0}, \tag{3.6a}$$

$$\dot{Z} = \hat{\boldsymbol{z}} \cdot v_\parallel \hat{\boldsymbol{b}} + \hat{\boldsymbol{z}} \cdot \boldsymbol{v}_{Ds} = \xi v \frac{X}{q R_0} - \frac{1}{2}\rho_i^* \left(1 + \xi^2\right) v^2 \frac{a}{R_0}, \tag{3.6b}$$

$$\dot{v} = 0, \tag{3.6c}$$

$$\dot{\xi} = \frac{\partial \xi}{\partial v_\parallel}\frac{\mathrm{d}v_\parallel}{\mathrm{d}t} + \frac{\partial \xi}{\partial \boldsymbol{X}} \cdot \frac{\mathrm{d}\boldsymbol{X}}{\mathrm{d}t} = -\frac{1}{2}\left(1 - \xi^2\right) v \frac{Z}{q R_0}\frac{a}{R_0}\frac{B}{B_0}. \tag{3.6d}$$

https://doi.org/10.1017/S0022377826101445 Published online by Cambridge University Press

Here, we normalised time to a/v_{Ti}, where a is the plasma minor radius taken to be equal to r_{wall}, speed v has been normalised to v_{Ti}, space to a, and we have defined $\rho_i^* = \rho_i/a$, $\rho_i = v_{Ti}/\omega_c^{(0)}$, $\omega_c^{(0)} = ZeB_0/m_i$, where B_0 is the magnetic field on axis, and q is the safety factor. For convenience, we have neglected the electric field, and assumed the limit of a small inverse aspect ratio and low-β plasma. Specifically, the magnetic field was taken to have the form:

$$\boldsymbol{B} = I\,(\psi)\,\nabla\varphi + \nabla\varphi \times \nabla\psi, \tag{3.7}$$

$$B_\varphi\,(r,\theta) = \frac{B_0}{1+\epsilon\cos\theta}, \tag{3.8}$$

$$B_\theta\,(r,\theta) = \frac{B_\theta\,(r)}{1+\epsilon\cos\theta}, \tag{3.9}$$

with

$$R = R_0 + r\cos\theta, \tag{3.10}$$

$$q = \frac{r\,B_0}{R_0 B_\theta\,(r)}, \tag{3.11}$$

and the poloidal flux function is related to the poloidal magnetic field by $\mathrm{d}\psi/\mathrm{d}r = R_0 B_\theta(r)$. The magnitude of the total magnetic field can be conveniently written as

$$\frac{B\,(r,\theta)}{B_0} = \frac{R_0}{R}\sqrt{1+\left(\frac{r}{q\,R_0}\right)^2} \approx \frac{R_0}{R}, \tag{3.12}$$

where in the last line, we take $(a/R_0)^2 \ll 1$.

Equations (3.6*a*) and (3.6*b*) describe the spatial evolution of the energetic ions due to parallel streaming and a vertical drift arising from the grad-B and curvature drifts. Our motivation for neglecting the electric field is that in this initial study, we will assume stationary and homogeneous density, temperature and current profiles, such that no electric field is expected to arise. In the absence of the electric field, the collisionless velocity space dynamics are set by the mirror force, which only impacts the pitch of the ions such that the speed v will be constant. While a more complete set of guiding centre equations can be easily adopted, for this initial study, (3.6) will prove sufficient to capture several non-trivial aspects of the collisionless orbit of energetic ions.

Aside from the speed v, this reduced set of guiding centre equations conserves toroidal canonical momentum in the absence of collisions, which can be written as

$$p_\varphi = m_i R\frac{B_\varphi}{B}v_\parallel - Ze\psi. \tag{3.13}$$

For circular flux surfaces and making analogous approximations as the guiding centrer equations (3.6) allows the normalised toroidal canonical momentum to be expressed as

$$\frac{p_\varphi}{R_0 m_i v_{Ti}} = \frac{R}{R_0}\xi v - \frac{1}{2}\frac{1}{\rho_i^*}\frac{a}{R_0}\frac{r^2}{q}, \tag{3.14}$$

where we have taken the safety factor to be a constant for simplicity. Figure 1 provides a numerical demonstration of the conservation of toroidal canonical

https://doi.org/10.1017/S0022377826101445 Published online by Cambridge University Press

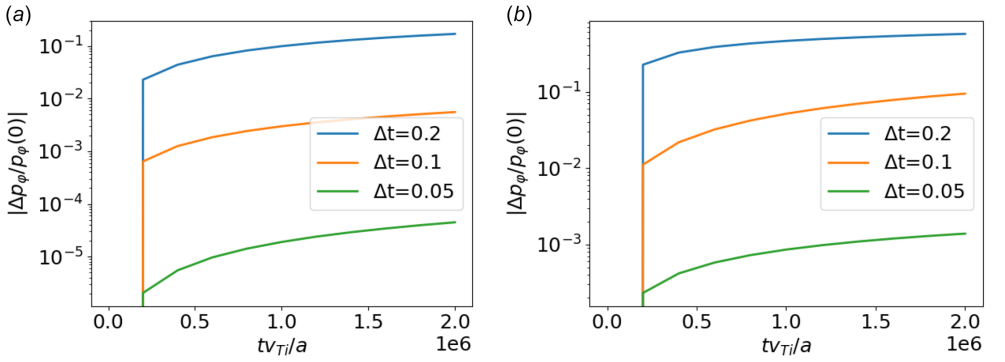


FIGURE 1. Change of toroidal canonical momentum for (*a*) 20 keV ions and (*b*) 50 keV ions. Ten million deuterium ions were initialised randomly across the spatial and pitch domains. The time step was varied from $\Delta t = 0.2$ (blue curve) to $\Delta t = 0.1$ (orange curve) and $\Delta t = 0.05$ (green curve). The tokamak was assumed to have a minor radius of $a = 0.5$ [m], an inverse aspect ratio of $a/R_0 = 1/3$, a magnetic field of $B_0 = 2$ [T] and a constant safety factor $q = 2$.

momentum conservation for ions with 20 and 50 keV. Here, while the error in toroidal canonical momentum conservation grows with time, after a time $2 \times 10^6 a/v_{Ti}$, the error remains below 10^{-2} for a time step of $\Delta t = 0.1a/v_{Ti}$ for 20 keV ions and $\Delta t = 0.05a/v_{Ti}$ for 50 keV ions. An error of less than 1 % will be sufficiently accurate for the present study. JONTA simulations described in the remainder of this work will use a time step of $\Delta t = 0.1a/v_{Ti}$ when studying 20 keV ions, and $\Delta t = 0.05a/v_{Ti}$ for 50 keV ions.

The collision operator is taken to be a Lorentz operator (2.2), where the deflection frequency in the limit of $v \gg v_{Ti}$ can be expressed as

$$\frac{a\nu_D}{v_{Ti}} = \left(a\frac{\hat{\nu}_{si}}{v_{Ti}}\right)\frac{v_{Ti}^3}{v^3} = \pi\left(n_i r_e^2 a\right) Z_s^2 Z_i^2 \left(\frac{m_e c^2}{T_i}\right)^2 \ln\Lambda \frac{v_{Ti}^3}{v^3}, \tag{3.15}$$

where $r_e = 2.8179 \times 10^{-15}$ [m] is the classical electron radius, Z_i is the charge state of the background ions and Z_s is the charge state of the energetic ion. The Lorentz collision operator is implemented by the Monte Carlo equivalent (Boozer & Kuo-Petravic 1981)

$$\xi_{n+1} = \xi_n\left(1 - \nu_D \Delta t\right) \pm \sqrt{\left(1 - \xi_n^2\right)\nu_D \Delta t}, \tag{3.16}$$

where Δt is the collisional time step, which is taken to be the same as the collisionless time step. While the Lorentz collision operator does not account for the slowing down of energetic ions, a critical aspect of describing energetic particle evolution, pitch-angle scattering, will result in collisional cross field transport. The addition of a more complete set of physical processes including 3-D magnetic fields, energetic particle modes and collisional slowing down will be pursued in future work.

To verify our implementation of the Monte Carlo collision operator, we compared JONTA predictions of neoclassical transport with an analytic expression for the neoclassical diffusivity. Using the radial coordinate $r = \sqrt{X^2 + Z^2}$, rather than the toroidal flux $\psi_t = r^2 B_0/2$, the expression for the neoclassical diffusivity derived by Rosenbluth, Hazeltine & Hinton (1972), Boozer & Kuo-Petravic (1981) can be

https://doi.org/10.1017/S0022377826101445 Published online by Cambridge University Press

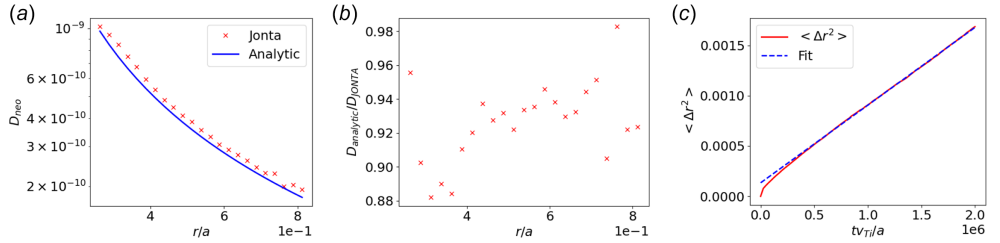


FIGURE 2. (*a*) Comparison of neoclassical diffusivity computed from JONTA (red 'x' markers) and the analytic expression given by (3.17) (solid blue curve). (*b*) Ratio of the analytic expression for the neoclassical diffusivity to JONTA's prediction. (*c*) Example fit of squared radial ion displacement versus time (dashed blue curve) with the numerically computed evolution (solid red curve) for particles with an initial radius of ∼0.51. One million total marker particles were used, randomly distributed over pitch and radius. Spatially resolved estimates of transport were made by binning marker particles based on their initial location. The parameters were chosen to be $a = 2$ [m], $a/R_0 = 1/6$, $B_0 = 10$ [T], $n_i = 10^{18}$ [m^{-3}], $q = 2$, $\ln \Lambda = 15$, and the ions were assumed to have an energy of 20 [keV].

written as

$$\frac{D_{neo}}{a v_{Ti}} = 0.689 \sqrt{2\varepsilon} q^2 \left(\frac{R_0}{r}\right)^2 \left(\frac{m_s v^2}{2T_i}\right) \left(\frac{a \nu_D}{v_{Ti}}\right), \tag{3.17}$$

where $\varepsilon = r/R_0$ and ν_D is the pitch-angle scattering frequency defined in (3.15). To estimate the neoclassical transport predicted by JONTA, we will consider a large tokamak device at low density. Specifically, we will take the density to be $n_i = 10^{18}$ [m^{-3}], the minor radius $a = 2$ [m], a small inverse aspect ratio of $a/R_0 = 1/6$ and a large toroidal magnetic field of $B_0 = 10$ [T]. By considering a low density plasma, this ensures the system is in the banana transport regime throughout the majority of the device, whereas the large device size and strong magnetic field results in the banana orbits of the fast ions being much smaller than the system size, a necessary condition for estimating an average diffusivity. A comparison of the neoclassical diffusivity estimated by JONTA with (3.17) is shown as a function of minor radius in figure 2. Here, good agreement is evident between the computed neoclassical diffusivity and the approximate analytic expression. In particular, the maximum deviation is roughly 12 %, with a typical deviation less than 10 %. Note that we have removed the inner 20 % of the plasma radius from the comparison, since the de-trapping time scales with ν_D/ε and is thus very short near the magnetic origin, such that this region is not in the banana transport regime.

3.3. *Example ion orbits and coordinate conventions*

Before discussing solutions to the adjoint drift kinetic equation, it will be useful to describe the magnetic geometry employed. A plot of the circular flux surface geometry assumed, together with four example ion orbits, are shown in figure 3. Here, the blue contours indicate magnetic flux surfaces, and the red and green curves indicate example passing and trapped ion orbits, respectively. The toroidal magnetic field and current are assumed to come out of the page, with the poloidal magnetic field in the counter-clockwise direction. Passing ions with a positive pitch (co-current) will thus rotate in the counter-clockwise direction when their motion is projected onto the poloidal plane as they propagate along a magnetic field line,

https://doi.org/10.1017/S0022377826101445 Published online by Cambridge University Press

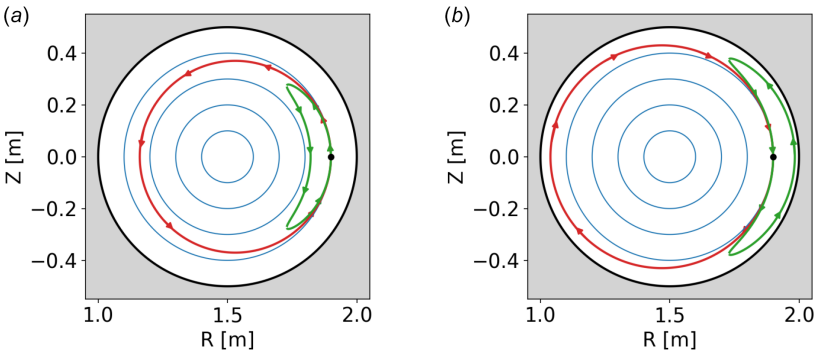


FIGURE 3. Circular flux surface geometry used in this work and example collisionless ion orbits. (*a*) Co-current ions with a passing ion with initial pitch $\xi = 0.8$ (red curve) and a trapped ion with initial pitch $\xi = 0.3$ (green curve). (*b*) Counter-current ions with a passing ion with initial pitch $\xi = -0.8$ (red curve) and a trapped ion with initial pitch $\xi = -0.3$ (green curve). Flux surface contours are shown in blue. The black marker indicates the initial position of the ions and arrows indicate their direction. The grad-B drift points downward in our example geometry. We assumed a deuterium ion with 20 keV, a tokamak with a minor radius of 0.5 m, inverse aspect ratio $a/R_0 = 1/3$, constant q-profile of $q = 2$ and an on-axis magnetic field strength of $B_0 = 2$ T.

with a grad-B drift vertically downward. Ions with a negative pitch (counter-current) will move in a clockwise sense when moving along field lines, but will still have a vertically downward grad-B drift. From figure 3, it is apparent that co-current ions (figure 3*a*) born on the weak field side will tend to be better confined compared with counter-current ions (figure 3*b*) born at the same location, due to the magnetic drifts taking the counter-current ions closer to the spatial boundary. Furthermore, a counter-current passing ion when scattered into the trapped region will move onto an orbit that takes it closer to the plasma edge, as evident from figure 3(*b*), such that these ions will be more susceptible to being lost.

4. Mean escape time

4.1. *Training details of the PINN*

This section will describe solutions for the mean escape time of a fast ion population at 20 and 50 keV using the PINN implementation described in § 3.1, along with a comparison to solutions evaluated with JONTA. Loss histories for ions with 20 and 50 keV are shown in figure 4. For these cases, we assumed a midsize tokamak with a minor radius of $a = 0.5$ m, aspect ratio $R_0/a = 3$ and deuterium fast ions. The background deuterium plasma is assumed to have a temperature of 1 keV, a density of $n_i = 10^{20}$ m^{-3} and the safety factor is taken to be $q = 2$. A representative Coulomb logarithm of $\ln \Lambda = 15$ was used in the collision operator. Each of these quantities were taken to be constant across the plasma radius, though non-trivial spatial profiles could easily be introduced. A fully connected feedforward network with seven hidden layers each with a width of fifty neurons was used. A network of this size, together with six million training points, used the majority of the 180 GB of VRAM present on the NVIDIA Blackwell 200 GPU used to train the network. While we have observed improvements in predictions of the PINN by increasing the size of the network, we found the present network size and number of training points offers a good balance in ensuring dense coverage over the ion phase space,

https://doi.org/10.1017/S0022377826101445 Published online by Cambridge University Press

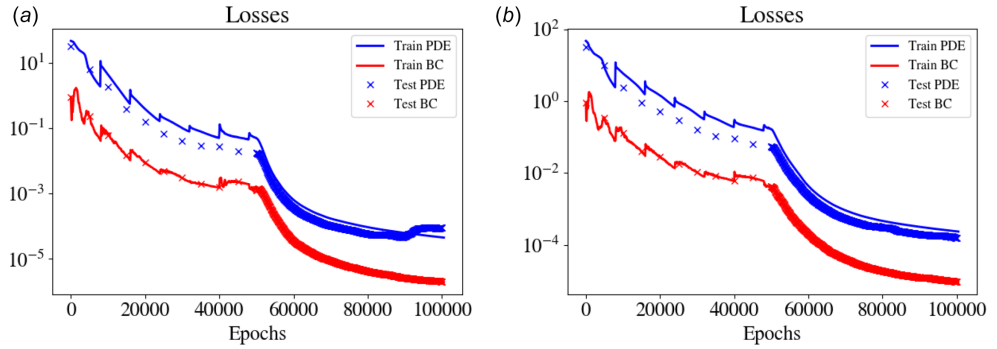


FIGURE 4. Loss histories for (*a*) 20 keV and (*b*) 50 keV ions. Solid lines indicate the training loss whereas 'x' markers indicate the test loss. Six million training points were used and two million test points. The same seed for the pseudorandom number generator was used in both cases.

while providing a sufficiently expressive network to approximate the solution. For a network of this size, the inference time on an M3 MacBook Pro was measured to be two microseconds per prediction.

The SOAP optimiser was used for the first 50 000 epochs, with SSBroyden used for the remaining 50 000 epochs. Six million training points were used with two million uniformly distributed test points applied. An adaptive residual based sampling method was employed during the SOAP phase of training (Wu *et al.* 2023), such that training points are redistributed near regions with large residuals. From figure 4, it is evident that SOAP drives a slow decay of the loss, with SSBroyden driving a sharper decline. The periodic spikes in the training loss during the SOAP phase are due to the training points being resampled every 8000 epochs. As a result of training points being concentrated in regions with large residuals, the test loss is consistently lower than the training loss, except for the last 10 000 epochs of training for the case of 20 keV ions. This crossing of the test and training loss is due to a large residual forming at some isolated regions immediately adjacent to $r = a$ and will be discussed further later. We note that the SOAP optimiser being employed in this paper is able to efficiently train on a GPU, whereas the SciPy implementation of SSBroyden that we have employed primarily uses a CPU. As a result, the vast majority of the training time is spent during the SSBroyden phase of optimisation.

Considering the evolution of the predicted solution during training, figure 5(*a*) shows the predicted mean escape time for a 20 keV ion after the SOAP optimisation phase, 10 000 epochs into the SSBroyden phase (figure 5*b*) and the final predicted mean escape time (figure 5*c*). A reference solution from JONTA is shown in figure 10(*d*) for comparison. It is evident that SOAP is able to approximately capture the structure of the solution near the tokamak edge, but fails to recover the long mean escape time of ions initially located at small minor radii. This is due to the large discrepancy in time scales between ions that begin on loss orbits, and are thus lost on a transit time scale, and those initially located deep in the plasma interior. These latter ions must collisionally diffuse out of the plasma and are thus lost on a much longer time scale, roughly 10^5 slower for the present example. After 10 000 epochs with SSBroyden, the training and test loss of the residual and boundary terms drop further, and it is evident from figure 5 that the relatively long confinement time of ions initially located in the interior is better captured, though its

https://doi.org/10.1017/S0022377826101445 Published online by Cambridge University Press

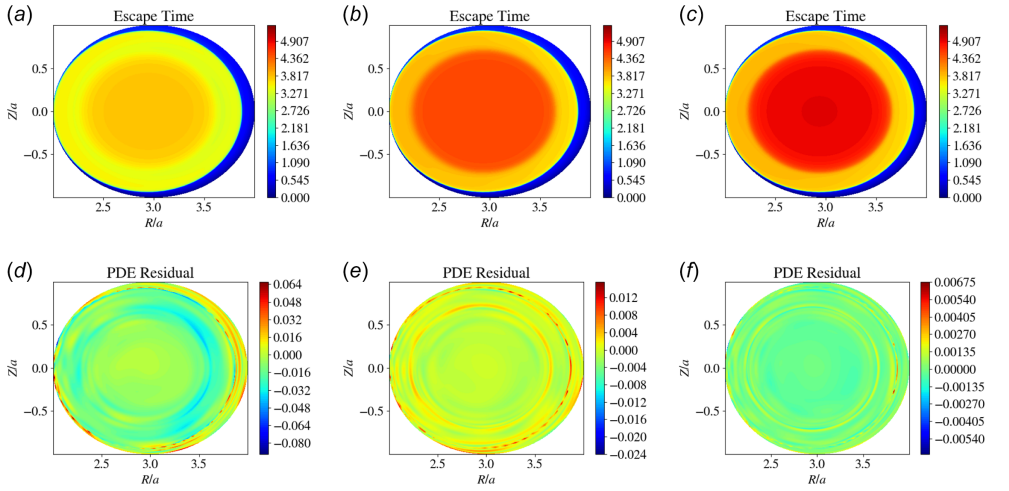


FIGURE 5. (*a*) Mean escape time for an ion initialised in the counter-current direction with a pitch of $\xi = -0.8$, after different stages of training. The profile of the ion's escape time is shown in panels (*a–c*), with the residual indicated in panels (*d–f*). The first column is after the SOAP phase of training (50 000 epochs), the second column is 10 000 into the SSBroyden phase (60 000 net epochs) and the final column is after 50 000 epochs of SSBroyden (100 000 epochs total). The quantity plotted is $\log_{10}(1 + T)$, yielding a $\log_{10}$ scale for large values of T, but vanishing when $T = 0$. Deuterium ions with 20 keV were assumed.

value is still substantially under predicted. Finally, after 50 000 epochs of SSBroyden, the predicted escape time of ions initially located in the interior slowly increases, though the final value is still below that predicted by the particle-based code JONTA. It has been verified that additional training epochs with SSBroyden leads to a more accurate recovery of the mean escape time of the best confined ion orbits, though as described later, begins to overfit the solution near the plasma edge at isolated locations.

Turning to the distribution of training losses, two-dimensional projections of the resulting test error are shown in figure 6, with the magnitude of the markers set to be proportional to the magnitude of the residual. Hence, darker regions represent regions with the largest residuals. From figure 6, the outer 30 % of the tokamak contained the largest losses. This follows due to this region containing direct loss orbits, whose detailed form can be difficult to resolve. It is also evident from figure 6(*d*), where the magnitude of the residuals are plotted as a function of the minor radius, that for $r \approx a$, a small number of very large residuals are present, with the largest having a magnitude of 0.589 located at $r/a \approx 0.998$ and $\theta \approx 3.041$. These large residuals at the very edge of the plasma are responsible for the increase in the test loss during the last 10 000 epochs of training, but as discussed further later, do not substantially impact the accuracy of the solution throughout the majority of the plasma volume. Finally, regions with $r/a > 0.99$ and poloidal angles on either the outboard ($-\Delta\theta < \theta < \Delta\theta$) or inboard ($\theta > -\pi + \Delta\theta$ and $\theta < \pi - \Delta\theta$) midplanes were set to zero, leading to empty regions in figure 6(*a*). For this example, we set $\Delta\theta = 0.1$. As described further later, these regions contain discontinuities in the solution arising from changes in boundary conditions between the lower and upper halves of the tokamak, which cannot be precisely resolved. By removing training points from this

https://doi.org/10.1017/S0022377826101445 Published online by Cambridge University Press

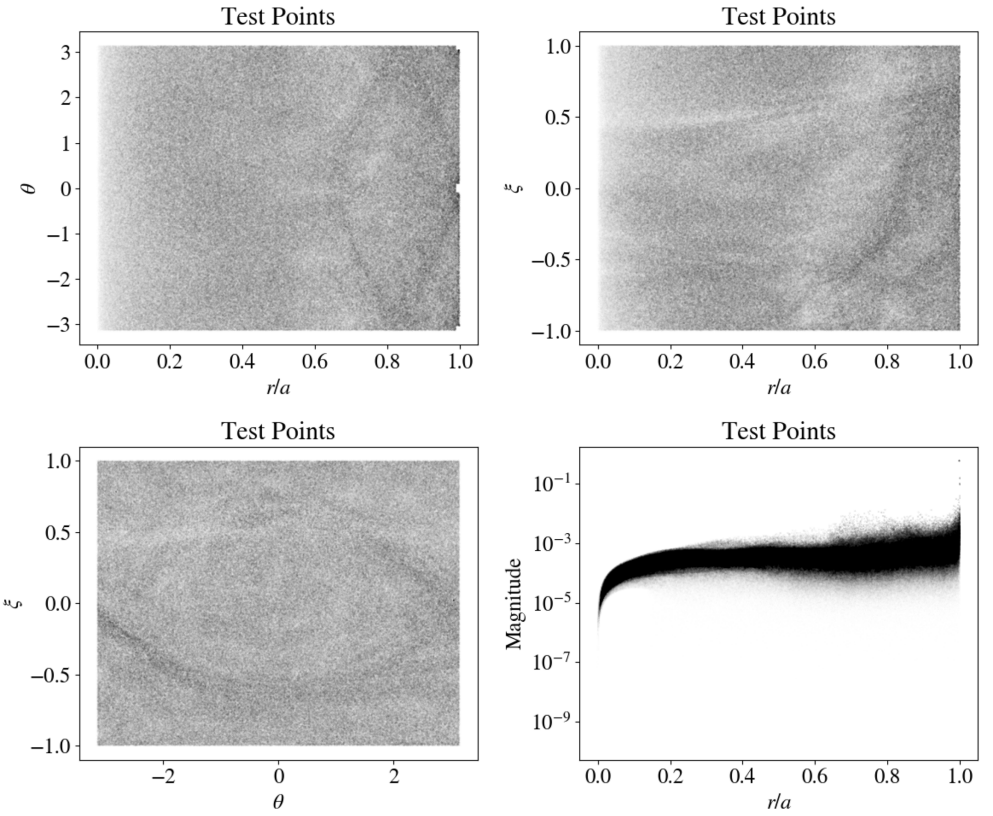


FIGURE 6. Test loss distribution for 20 keV ions. A uniform random distribution with one million points was used. The size of the markers is proportional to the magnitude of the residual.

region, this allows the PINN to find a solution that smoothly transitions across these narrow regions. We note that the largest residuals in the test loss are located at poloidal angles just outside of these regions, and are thus due to the sharp variation of the solution between the upper and lower regions of the tokamak near the plasma edge (see figure 9*a*).

4.2. *Phase space structure of the mean escape time*

The phase space distribution of the fast ion mean escape time can be conveniently illustrated by taking slices in the three-dimensional (R, Z, ξ) space. The mean escape time of ions born with $\theta = 0$ (weak field side) is shown in figure 7 on a $\log_{10}$ scale. Here, counter-current ions ($\xi < 0$) born in the edge of the plasma are lost quickly due to direct orbit loss, with both co- and counter-current ions initially located deeper in the plasma having much longer mean escape times. Further noting that the white contour indicates the trapped-passing boundary, it is evident that co-current passing ions (positive pitches above the white contour) have far better confinement in comparison to counter-current ions. Similarly, in the trapped region, ions that begin with $\xi > 0$ are much better confined compared with ions with $\xi < 0$. Comparing the left and right panels of figure 3, this difference in confinement results from trapped ions born with $\xi < 0$ at $\theta = 0$ having orbits that extend out to larger radii (figure 3*b*), in contrast to trapped ions born with $\xi > 0$ (figure 3*a*). The ions with the worst

https://doi.org/10.1017/S0022377826101445 Published online by Cambridge University Press

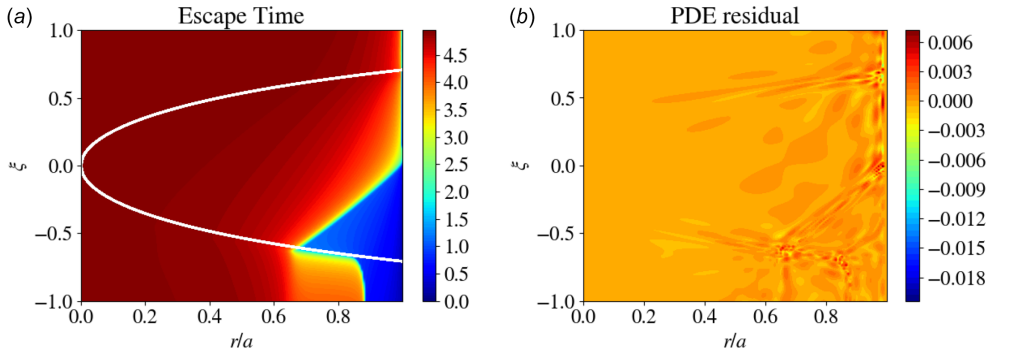


FIGURE 7. (*a*) Escape time for an ion initially located at $\theta = 0$. The quantity plotted is $\log_{10}(1 + T)$, yielding a $\log_{10}$ scale for large values of T, but vanishing when $T = 0$. The white curve indicates the trapped-passing boundary. (*b*) Residual of (2.6). A deuterium ion with 20 keV was assumed.

confinement are those just inside the trapped-passing boundary and $\xi < 0$, which are expected to have the largest banana width and thus be most susceptible to orbit loss.

Figure 8 shows slices of major radius R and vertical height Z for different pitches. These slices allow for the difference in confinement of co- and counter-current ions to be investigated more clearly. Considering passing co-current ions ($\xi = 0.8$), these ions will propagate in the counter-clockwise direction in figure 8(*a*). Noting that ion orbits drift downward, this will result in ions that are initially on the outboard side not being on direct loss orbits, in contrast to those born on the inboard side. This leads to a notable in–out asymmetry of the fast ion population, as evident in figure 8(*a*). This trend is reversed for passing ions that begin in the counter-current direction (figure 8*d*). A more subtle feature is that counter-current ions that are not on direct loss orbits are more likely to be lost in comparison to co-current ions. This is evident by the orange region in figure 8(*d*). The relatively poor confinement of these counter-current ions is due to the collisional trapping of these ions resulting in the orbit exploring larger minor radii (compare the red and green curves in figure 3*b*). Hence, passing counter-current ions that are within a banana width of the plasma edge are far more susceptible to being lost to the plasma compared with passing co-current ions.

Cross-cuts of the pitch and poloidal angle distribution of loss times at different radii are shown in figure 9. Here, when approaching $r/a = 1$, a discontinuity appears in the solution. This is evident in figure 9(*a*), where for $\theta \approx 0$ and $\theta \approx \pm\pi$, the solution changes rapidly when traversing these discontinuous regions. The origin of this discontinuity is that for ions in the lower half of the tokamak (i.e. $-\pi < \theta < 0$) and $r/a = 1$, they are lost immediately. In contrast, ions located just above $\theta = 0$ or $\theta = \pm\pi$ are often not on direct loss orbits and, thus, can be confined for a long time. For example, co-current ions just above $\theta = 0$ will drift inward and, thus, have a finite escape time, whereas those just below $\theta = 0$ are lost immediately to the boundary. An analogous discontinuity in the solution is present for $\theta = \pm\pi$ for counter-current ions. To relax these discontinuities, we have removed all training and test points from these regions, such that the PINN is able to approximate these discontinuities with sharply varying, but continuous solutions, as evident in figure 9(*a*). Moving radially away from the $r/a = 1$ boundary, figure 9(*b–d*) shows the solution at radial locations between $r/a = 0.95$ and $r/a = 0.8$, indicating vastly different mean escape

https://doi.org/10.1017/S0022377826101445 Published online by Cambridge University Press

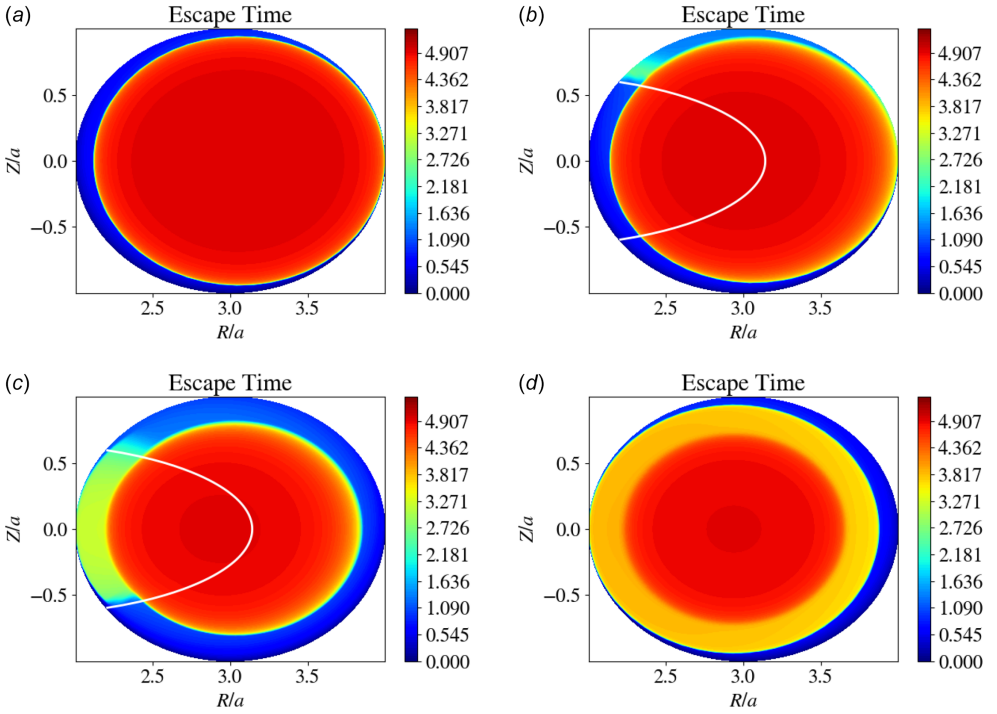


FIGURE 8. Slices of escape time in units of $\log_{10}(1+T)$ for different initial pitch values: (*a*) for $\xi = 0.8$; (*b*) for $\xi = 0.3$; (*c*) for $\xi = -0.3$; and (*d*) for $\xi = -0.8$. The solid white curve indicates the location of the trapped-passing boundary, where ions born with major radii R to the right of the white curve are trapped.

times of ions initially located at different pitch and poloidal angles. In particular, the phase space region of prompt ion loss, which initially occupies roughly half of the (ξ, θ) domain for $r/a \approx 1$, shrinks to a narrow sliver just inside the trapped-passing boundary for $r/a = 0.8$. This is indicated in figure 9(*d*), where the solid white curve indicates the trapped-passing region.

4.3. *Comparison with Monte Carlo solutions*

The mean escape time can also be computed by the particle-based solver JONTA described in § 3.2. Here, marker particles are initialised randomly across the spatial domain at a given pitch. Once a particle's minor radius exceeds r_{wall}, its escape time and location are saved. By pushing a large number of marker particles, the mean escape time can be computed as a function of the ion's initial R and Z. Noting the conceptual simplicity of this approach, we will use it to validate the predictions of the PINN. Figure 10 shows four slices in the (R, Z) plane at constant pitch computed from JONTA. Here, ten million marker particles are randomly distributed at a given pitch, with their escape times recorded after they escape from the domain. These escape times are then binned, allowing for the mean escape time to be computed over the (R, Z) plane for a given pitch. The particles were evolved until $t_{final} = 2 \times 10^6$, a time at which nearly all of the original particles remained in the domain. Four cross-cuts are shown in figure 10. These four cross-cuts are for the same physics parameters and pitches as figure 8. Good agreement between the mean escape time

https://doi.org/10.1017/S0022377826101445 Published online by Cambridge University Press

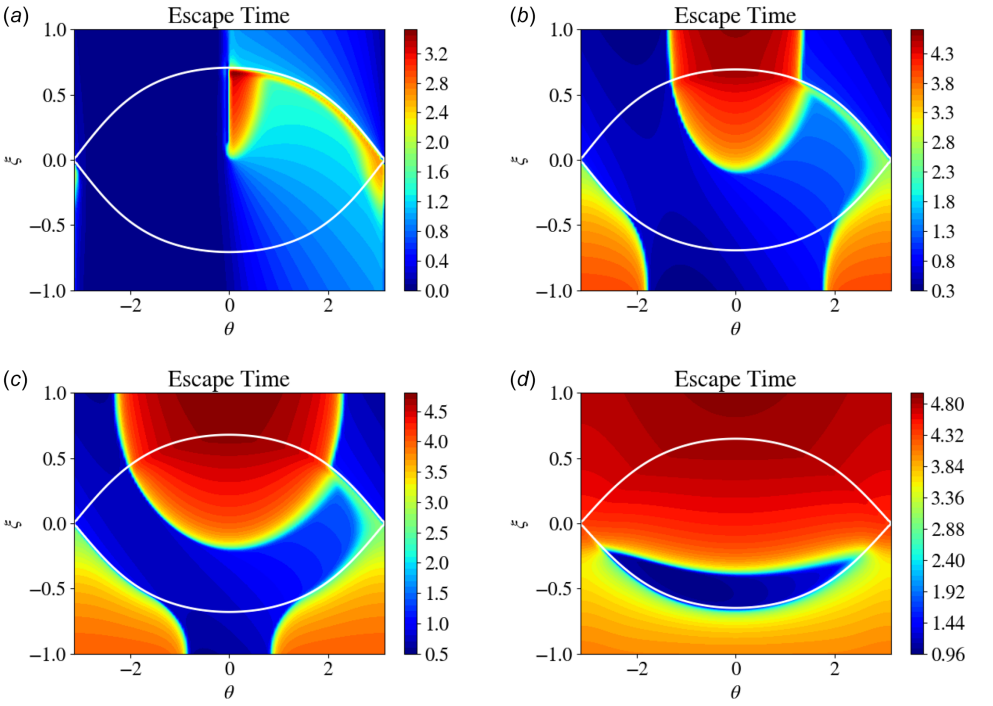


FIGURE 9. Slices of escape time in units of $\log_{10}(1+T)$ for different initial radii: (*a*) for $r/a = 0.999$; (*b*) for $r/a = 0.95$; (*c*) for $r/a = 0.9$; and (*d*) for $r/a = 0.8$. The while curves are the trapped-passing boundary.

predicted by JONTA and the PINN are evident. In particular, both approaches are able to recover the rapid escape of ions that are initiated on direct loss orbits (blue regions), ions lost due to collisional trapping/detrapping directly onto a loss orbit (yellow and green regions) and ions initially located near the magnetic origin that must slowly diffuse out (red regions). The JONTA simulations also reveal a limitation of our implementation of a PINN. In particular, while the PINN is able to distinguish the complex edge structure of confined versus unconfined orbits, with excellent agreement with predictions of JONTA, it struggles to accurately capture the very long mean escape time for ions initially located deep inside the plasma. Specifically, the dark red region evident in figure 10 of the JONTA data is not present in the PINN's solution (see figure 8), indicating roughly a factor of two and a half discrepancy in predicted confinement for these 20 keV well-confined ions. The reason for this quantitative inaccuracy is that the PINN struggles to simultaneously resolve the fast direct orbit dynamics together with the slow collisional transport of the core ions, though as indicated in figure 5, this discrepancy slowly shrinks as the model is trained. The resolution of this disparate physics will require (i) the inclusion of data in training the PINN, e.g. using data taken directly from JONTA for example, (ii) allowing the PINN to train for more epochs which has been verified to slowly reduce the discrepancy, or (iii) modifications to the training of the PINN. These extensions will be discussed further later.

Now, considering an energetic ion population with 50 keV, similar physical trends are evident, but the larger ion orbits lead to a larger number of ions beginning on direct loss orbits (see figure 11). In addition, the higher energy of the particles

https://doi.org/10.1017/S0022377826101445 Published online by Cambridge University Press

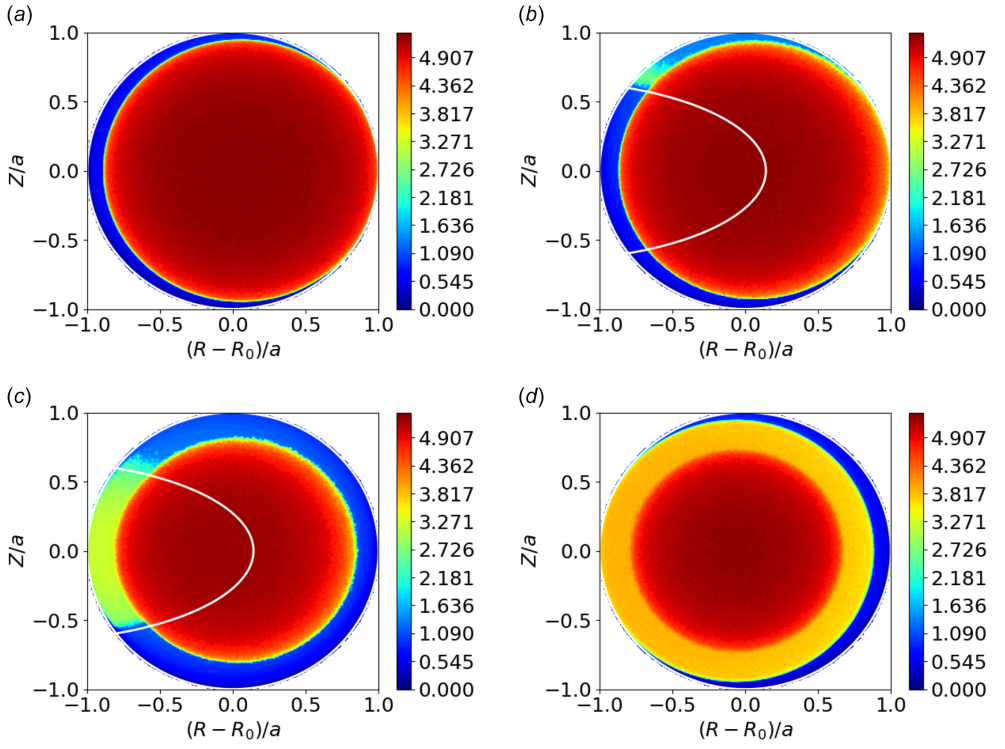


FIGURE 10. Slices of escape time in units of $\log_{10}(1+T)$ for different initial pitch values and an energy of 20 keV computed from the JONTA code: (*a*) for $\xi = 0.8$; (*b*) for $\xi = 0.3$; (*c*) for $\xi = -0.3$; and (*d*) for $\xi = -0.8$. Ten million markers were used, where the particles were integrated for $t_{final} = 2 \times 10^6$.

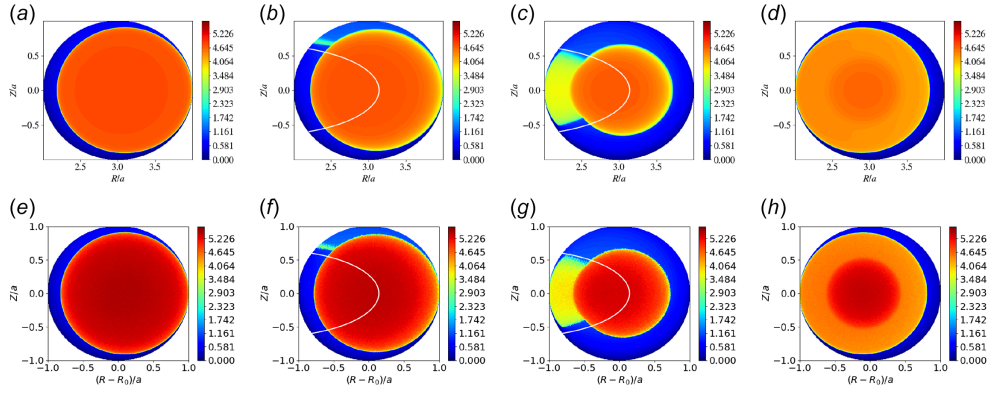


FIGURE 11. Comparison of escape time in units of $\log_{10}(1+T)$ predictions from the PINN (top row) and JONTA (bottom row). Ions were assumed to have an initial energy of 50 keV: (*a*, *e*) for $\xi = 0.8$; (*b*, *f*) for $\xi = 0.3$; (*c*, *g*) for $\xi = -0.3$; and (*d*, *h*) for $\xi = -0.8$. Ten million markers were used in the JONTA simulations, where the particles were integrated for $t_{final} = 10^7$.

https://doi.org/10.1017/S0022377826101445 Published online by Cambridge University Press

leads to a lower collision time and thus energetic ions initially located in the core have a longer escape time. This latter property results in the problem exhibiting greater stiffness, with a faster bounce time but slower collision time, resulting in the PINN struggling to accurately resolve the escape time of the best confined orbits. This is evident in figure 11, where cross-sections in the (R, Z) plane computed both with JONTA and the PINN are compared. While the PINN fails to capture the precise mean escape time of the best confined particles, it does succeed in accurately capturing the phase space structure of the mean escape time across the edge region, where ions have the shortest mean escape times. It thus acts as an effective tool for delineating well-confined regions of phase space from those where ions are directly lost or lost after a small number of collisional scattering events. We note that while for the case of 20 keV ions a factor of roughly two and a half difference in the mean escape time was present for the best confined ions, for 50 keV ions, roughly a factor of five difference in the prediction of the best confined orbits is evident. This larger deviation in the prediction of the best confined ions is due to the increased time scale separation between the collisionless ion orbits and the slow collision time. As discussed further in § 5, while further training or hyperparameter tuning can likely reduce these discrepancies, a robust means of improving the predictions of the PINN will be the incorporation of small quantities of data.

5. Conclusions

The inhomogeneous adjoint of the drift kinetic equation was used to evaluate the mean escape time of energetic ion populations. This quantity, computed in the limit of an axisymmetric tokamak plasma, captures fast ion losses due to direct orbit loss and collisional transport, thus providing a metric for energetic particle confinement. A PINN was used to evaluate the phase space dependence of the mean escape time, a challenging task due to the strong time scale separation between the fast bounce time of energetic ions and their slow collision rate. While the offline training time of the PINN is substantial, the online inference time is rapid, typically microseconds per prediction. Although the PINN failed to quantitatively capture the mean escape time of the best confined ion orbits, it is able to robustly recover the phase space structure of the mean escape time, and thus acts as a means of distinguishing well-confined orbits and phase space regions of prompt ion loss. We also note that the present analysis did not include fast ion slowing down, where the slowing down time for the 20 and 50 keV ions considered in this work would be far shorter than the mean escape times computed in this work for the best confined orbits.

For this proof-of-principle demonstration of the proposed framework, we considered several simplifications to the complete system. Perhaps the primary simplifications were the assumptions of (i) an axisymmetric magnetic geometry, (ii) a simplified collision operator, and (iii) idealised plasma profiles and shaping. Regarding item (i), since PINNs to date have not been used to solve the drift kinetic equation, beginning with an axisymmetry geometry was a logical starting point for assessing this approach's strengths and weaknesses. We anticipate extending the present approach in future work to 3-D magnetic geometry (Law, Cerfon & Peherstorfer 2022), though such an extension will likely require incorporating particle data into the training of the neural network. The particle-based solver JONTA will provide a convenient tool through which such data can be generated, where by providing an accurate calculation of the mean escape time at specific phase space locations, these may be included in the first term of the loss function defined by

https://doi.org/10.1017/S0022377826101445 Published online by Cambridge University Press

(3.1). By providing high fidelity data points at internal locations, for example, in regions with large residuals or for the best confined ion orbits where the present implementation of a PINN has struggled, this will likely expedite the training of the PINN while providing a systematic means of improving its accuracy. Our aim will be to leverage data as a means of both treating more challenging scenarios, such as non-axisymmetric geometry, along with facilitating the learning of a parametric solution to the drift kinetic equation, enabling the mean escape time to be evaluated for a broad range of plasma conditions.

An additional simplification was the use of a Lorentz collision operator, which does not account for the slowing down of energetic particles, a quantity of critical interest when assessing the amount of energy deposited by energetic ions into the bulk plasma. Despite this limitation, including Lorentz collisions allows for the collisional transport of fast ions to be assessed. Thus, while not providing an accurate indication of the mean escape time for well-confined energetic particles, since such particles will likely slow down substantially during their long confinement time, it does provide a means of rapidly assessing particles on direct loss orbits, or those that are directly lost after being collisionally scattered into or out of the trapped particle region. The predictions of the present work will thus be most relevant to the outer region of the tokamak, where energetic particles are lost on relatively short time scales. Finally, homogeneous plasma profiles and a circularly shaped plasma were assumed in the present work. This was done for convenience, where we anticipate including more realistic plasma profiles to not pose a substantial challenge. Furthermore, since PINNs are mesh free and are thus well suited to treating shaped magnetic geometry (van Milligen *et al.* 1995; Kaltsas & Throumoulopoulos 2022; Jang *et al.* 2024), we do not expect the treatment of shaped axisymmetric geometry to pose a major hurdle. Similarly, strong radial electric fields that emerge in the pedestal region (Brzozowski *et al.* 2019) along with realistic equilibrium density and temperature profiles could easily be incorporated.

Due to the strong time scale separation, evaluating fast ion transport is computationally intensive, creating a bottleneck in many optimisation workflows. For the JONTA simulations presented in this paper, the shortest simulation required one and a half hours on a single NVIDIA Blackwell 200 GPU, with the longest requiring slightly over a day. Noting that each JONTA simulation was for a fixed value of pitch, computing four distinct pitches (figures 10 and 11) thus required several hours to days of computing time. A PINN, which encompasses all values of pitch, and thus only a single PINN must be trained, required approximately three days to train, where approximate results obtained after the SOAP phase of training typically take a few hours. Both approaches are thus computationally intensive, where while the PINN is able to recover smooth predictions of the mean escape time, it is unable to accurately compute the best confined ions. JONTA, in contrast, requires roughly ten million markers to obtain a relatively smooth, well-resolved solution. We note that while the PINN also used a single NVIDIA Blackwell 200 GPU, the implementation of the SSBroyden optimiser does not make efficient use of the GPU, and thus was mostly trained using a CPU, which was the primary bottleneck in the training of the PINN.

We also note that while the present paper has focused on the mean escape time of energetic ions, several straightforward extensions are possible. One such extension would be to treat a broader range of fast ion metrics. Metrics of interest include the probability that an ion impacts a specific region of the wall from a given initial phase space location (Salewski *et al.* 2011; Särkimäki 2020) (see also Yang *et al.* (2026) for

https://doi.org/10.1017/S0022377826101445 Published online by Cambridge University Press

a runaway electron example), or the slowing down probability, which corresponds to the probability that an ion slows down to the thermal energy. Both of these fast ion metrics can be straightforwardly constructed from the steady-state homogeneous adjoint equation by applying appropriate boundary conditions. The evaluation of this broader class of fast ion metrics will be left for future work. Finally, while we have considered energetic ions in the present formulation, extensions to the relativistic drift kinetic equation describing runaway electron (RE) evolution and transport can be developed following an analogous approach. Such extensions will allow for 'neoclassical' corrections to RE generation rates to be inferred (Rosenbluth & Putvinski 1997; McDevitt & Tang 2019; Arnaud & McDevitt 2024), or account for spatial transport due to processes such as the Ware pinch (Nilsson *et al.* 2015; McDevitt, Guo & Tang 2019). While we anticipate that the extremely short bounce time of REs compared with their long collision times will pose a substantial challenge compared with the purely momentum space treatment employed by Arnaud *et al.* (2024), McDevitt *et al.* (2025), we note that the impact of 'neoclassical' effects on RE generation rates are most important near the critical energy for electrons to run away (McDevitt & Tang 2019; Arnaud & McDevitt 2024), an energy that often ranges between several keV to tens of keV during a tokamak disruption. At these modest energies, the time scale separation between the transit or bounce time of an RE and the collision time will be drastically reduced. The extension of the present framework to REs will be pursued in future work.

Acknowledgements

The authors acknowledge the University of Florida Research Computing for providing computational resources that have contributed to the research results reported in this publication.

Editor Tünde Fülöp thanks the referees for their advice in evaluating this article.

Funding

This work was supported by the Department of Energy, Office of Fusion Energy Sciences under awards DE-SC0024634 and DE-SC0024649, and the University of Florida Division of Sponsored Research Projects.

Data availability statement

The version of JONTA used in this manuscript, together with the PINN training script are available on github.com/cmcdevitt2/DeepPlasma.

Declaration of interests

The authors report no conflict of interest.

REFERENCES

ADAM, J. 1987 Review of tokamak plasma heating by wave damping in the ion cyclotron range of frequency. *Plasma Phys. Control. Fusion* **29**, 443.

AL-BAALI, M., SPEDICATO, E. & MAGGIONI, F. 2014 Broyden's quasi-Newton methods for a nonlinear system of equations and unconstrained optimization: a review and open problems. *Optim. Methods Softw.* **29**, 937–954.

https://doi.org/10.1017/S0022377826101445 Published online by Cambridge University Press

ARNAUD, J.S., MARK, T. & MCDEVITT, C.J. 2024 A physics-constrained deep learning surrogate model of the runaway electron avalanche growth rate. *J. Plasma Phys.* **90**, 905900409.

ARNAUD, J.S. & MCDEVITT, C.J. 2024 The impact of collisionality on the runaway electron avalanche during a tokamak disruption. *Phys. Plasmas* **31**, 062509.

ARNAUD, J.S., TANG, X.-Z. & MCDEVITT, C.J. 2025 A runaway electron avalanche surrogate for partially ionized plasmas. *Nucl. Fusion* **65**, 106013.

BONOFIGLO, P. *et al.* 2024 Alpha particle loss measurements and analysis in jet dt plasmas. *Nucl. Fusion* **64**, 096038.

BOOZER, A.H. & KUO-PETRAVIC, G. 1981 Monte carlo evaluation of transport coefficients. *Phys. Fluids* **24**, 851–859.

BRADBURY, J. *et al.* 2018 Jax: composable transformations of python+ numpy programs.

BRZOZOWSKI, R.W. *et al.* 2019 A geometric model of ion orbit loss under the influence of a radial electric field. *Phys. Plasmas* **26**, 042511.

CAI, S., MAO, Z., WANG, Z., YIN, M. & KARNIADAKIS, G.E. 2021 Physics-informed neural networks (pinns) for fluid mechanics: a review. *Acta Mechanica Sin.* **37**, 1727–1738.

CAROLIPIO, E., HEIDBRINK, W., FOREST, C. & WHITE, R. 2002 Simulations of beam ion transport during tearing modesin the diii-d tokamak. *Nucl. Fusion* **42**, 853.

CUOMO, S., DI COLA, V.S., GIAMPAOLO, F., ROZZA, G., RAISSI, M. & PICCIALLI, F. 2022 Scientific machine learning through physics–informed neural networks: Where we are and what's next. *J. Sci. Comput.* **92**, 88.

FASOLI, A. *et al.* 2007 Physics of energetic ions. *Nucl. Fusion* **47**, S264.

GALEEV, A.A., SAGDEEV, R., FURTH, H. & ROSENBLUTH, M. 1969 Plasma diffusion in a toroidal stellarator. *Phys. Rev. Lett.* **22**, 511.

GOLDSTON, R.J., WHITE, R. & BOOZER, A.H. 1981 Confinement of high-energy trapped particles in tokamaks. *Phys. Rev. Lett.* **47**, 647.

GRASMAN, J. *et al.* 2013 *Asymptotic Methods for the Fokker–Planck Equation and the Exit Problem in Applications*. Springer Science & Business Media.

HEIDBRINK, W. & WHITE, R. 2020 Mechanisms of energetic-particle transport in magnetically confined plasmas. *Phys. Plasmas* **27**, 030901.

HELANDER, P. & SIGMAR, D.J. 2002 *Collisional Transport in Magnetized Plasmas*. Cambridge University Press.

HEMSWORTH, R., TANGA, A. & ANTONI, V. 2008 Status of the iter neutral beam injection system. *Rev. Sci. Instrum.* **79**, 02C109.

HINTON, F. & CHU, M. 1985 Neoclassical ion transport through the separatrix in divertor tokamaks. *Nucl. Fusion* **25**, 345.

HIRVIJOKI, E., ASUNTA, O., KOSKELA, T., KURKI-SUONIO, T., MIETTUNEN, J., SIPILÄ, S., SNICKER, A. & ÄKÄSLOMPOLO, S. 2014 Ascot: solving the kinetic equation of minority particle species in tokamak plasmas. *Comput. Phys. Commun.* **185**, 1310–1321.

JANG, B., KAPTANOGLU, A.A., GAUR, R., PAN, S., LANDREMAN, M. & DORLAND, W. 2024 Grad–shafranov equilibria via data-free physics informed neural networks. *Phys. Plasmas* **31**, 032510.

KALTSAS, D. & THROUMOULOPOULOS, G. 2022 Neural network tokamak equilibria with incompressible flows. *Phys. Plasmas* **29**, 022506.

KARNIADAKIS, G.E., KEVREKIDIS, I.G., LU, L., PERDIKARIS, P., WANG, S. & YANG, L. 2021 Physics-informed machine learning. *Nat. Rev. Phys.* **3**, 422–440.

KAZAKOV, Y.O. *et al.* 2017 Efficient generation of energetic ions in multi-ion plasmas by radio-frequency heating. *Nat. Phys.* **13**, 973–978.

KINGMA, D.P. & BA, J. 2014 Adam: a method for stochastic optimization. arXiv preprint arXiv: 1412.6980.

KIYANI, E., SHUKLA, K., URBÁN, J.F., DARBON, J. & KARNIADAKIS, G.E. 2025 Optimizing the optimizer for physics-informed neural networks and Kolmogorov–Arnold networks. *Comput. Meth. Appl. Mech. Engng* **446**, 118308.

LAW, F., CERFON, A. & PEHERSTORFER, B. 2022 Accelerating the estimation of collisionless energetic particle confinement statistics in stellarators using multifidelity monte carlo. *Nucl. Fusion* **62**, 076019.

https://doi.org/10.1017/S0022377826101445 Published online by Cambridge University Press

LIU, C., BRENNAN, D.P., BOOZER, A.H. & BHATTACHARJEE, A. 2017 Adjoint method and runaway electron avalanche. *Plasma Phys. Control. Fusion* **59**, 024003.

LIU, D.C. & NOCEDAL, J. 1989 On the limited memory bfgs method for large scale optimization. *Math. Program* **45**, 503–528.

LUO, W., REN, J.-X., CHEN, Z., MAO, R., ZHANG, G., WANG, Y. & TANG, H. 2025 Parametric study of plasma swirling flow based on magnetohydrodynamics informed deep learning. *Phys. Fluids* **37**, 077106.

MATHEWS, A., HUGHES, J.W., TERRY, J.L. & BAEK, S.-G. 2022 Deep electric field predictions by drift-reduced braginskii theory with plasma-neutral interactions based on experimental images of boundary turbulence. *Phys. Rev. Lett.* **129**, 235002.

MCDEVITT, C.J. 2023 A physics-informed deep learning model of the hot tail runaway electron seed. *Phys. Plasmas* **30**, 092501.

MCDEVITT, C.J., ARNAUD, J.S. & TANG, X.-Z. 2025 A physics-constrained deep learning treatment of runaway electron dynamics. *Phys. Plasmas* **32**, 042503.

MCDEVITT, C., FOWLER, E. & ROY, S. 2024 Physics-constrained deep learning of incompressible cavity flows. In *AIAA SCITECH, Forum*, p. 1692.

MCDEVITT, C.J., GUO, Z. & TANG, X. 2019 Spatial transport of runaway electrons in axisymmetric tokamak plasmas. *Plasma Phys. Control. Fusion* **61**, 024004.

MCDEVITT, C. & TANG, X.-Z. 2019 Runaway electron generation in axisymmetric tokamak geometry. *EPL (Europhys. Lett.)* **127**, 45001.

MCDEVITT, C.J. & TANG, X.-Z. 2024 A physics-informed deep learning description of knudsen layer reactivity reduction. *Phys. Plasmas* **31**, 062701.

NILSSON, E., DECKER, J., FISCH, N.J. & PEYSSON, Y. 2015 Trapped-electron runaway effect. *J. Plasma Phys.* **81**, 475810403.

PASZKE, A. *et al.* 2019 Pytorch: an imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems* 32.

RAISSI, M., PERDIKARIS, P. & KARNIADAKIS, G.E. 2019 Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **378**, 686–707.

ROSENBLUTH, M., HAZELTINE, R. & HINTON, F.L. 1972 Plasma transport in toroidal confinement systems. *Phys. Fluids* **15**, 116–140.

ROSENBLUTH, M. & PUTVINSKI, S. 1997 Theory for avalanche of runaway electrons in tokamaks. *Nucl. Fusion* **37**, 1355.

SALEWSKI, M. *et al.* 2011 On velocity space interrogation regions of fast-ion collective thomson scattering at iter. *Nucl. Fusion* **51**, 083014.

SUN, L., GAO, H., PAN, S. & WANG, J.-X. 2020 Surrogate modeling for fluid flows based on physics-constrained deep learning without simulation data. *Comput. Method. Appl. Mech. Engng* **361**, 112732.

SÄRKIMÄKI, K. 2020 Efficient and rigorous evaluation of fast particle losses in non-axisymmetric tokamak plasmas. *Nucl. Fusion* **60**, 036002.

TOSCANO, J.D., OOMMEN, V., VARGHESE, A.J., ZOU, Z., AHMADI DARYAKENARI, N., WU, C. & KARNIADAKIS, G.E. 2025 From pinns to pikans: recent advances in physics-informed machine learning. *Mach. Learn. Comput. Sci. Engng* **1**, 1–43.

TROTT, C.R. *et al.* 2021 Kokkos 3: Programming model extensions for the exascale era. *IEEE Trans. Parallel Distrib. Syst.* **33**, 805–817.

URBÁN, J.F., STEFANOU, P. & PONS, J.A. 2025 Unveiling the optimization process of physics informed neural networks: how accurate and competitive can pinns be? *J. Comput. Phys.* **523**, 113656.

VAN MILLIGEN, B.P., TRIBALDOS, V. & JIMÉNEZ, J. 1995 Neural network differential equation and plasma equilibrium solver. *Phys. Rev. Lett.* **75**, 3594.

VYAS, N., MORWANI, D., ZHAO, R., KWUN, M., SHAPIRA, I., BRANDFONBRENER, D., JANSON, L. & KAKADE, S. 2024 Soap: improving and stabilizing shampoo using adam. arXiv preprint arXiv: 2409.11321.

WANG, S., BHARTARI, A.K., LI, B. & PERDIKARIS, P. 2025 Gradient alignment in physics-informed neural networks: a second-order optimization perspective. arXiv preprint arXiv: 2502.00604.

https://doi.org/10.1017/S0022377826101445 Published online by Cambridge University Press

WANG, S., SANKARAN, S., WANG, H. & PERDIKARIS, P. 2023 An expert's guide to training physics-informed neural networks. arXiv preprint arXiv: 2308.08468.

WHITE, R.T & CHANCE, M. 1984 Hamiltonian guiding center drift orbit calculation for toroidal plasmas of arbitrary cross section. Tech. Rep. Princeton Plasma Physics Lab. (PPPL).

WU, C., ZHU, M., TAN, Q., KARTHA, Y. & LU, L. 2023 A comprehensive study of non-adaptive and residual-based adaptive sampling for physics-informed neural networks. *Comput. Meth. Appl. Mech. Engng* **403**, 115671.

YANG, M., LIU, Y., DEL-CASTILLO-NEGRETE, D., CAO, Y. & ZHANG, G. 2026 Generative ai models for learning flow maps of stochastic dynamical systems in bounded domains. *J. Comput. Phys.* **544**, 114434.

https://doi.org/10.1017/S0022377826101445 Published online by Cambridge University Press