J. Plasma Phys. (2024), *vol*. 90, 905900409 © The Author(s), 2024.
Published by Cambridge University Press
This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.
doi:10.1017/S0022377824000679

A physics-constrained deep learning surrogate model of the runaway electron avalanche growth rate

J.S. Arnaud [1],†, **T.B. Mark** [1] **and C.J. McDevitt** [1]

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

(Received 7 March 2024; revised 4 May 2024; accepted 8 May 2024)

A surrogate model of the runaway electron avalanche growth rate in a magnetic fusion plasma is developed. This is accomplished by employing a physics-informed neural network (PINN) to learn the parametric solution of the adjoint to the relativistic Fokker–Planck equation. The resulting PINN is able to evaluate the runaway probability function across a broad range of parameters in the absence of any synthetic or experimental data. This surrogate of the adjoint relativistic Fokker–Planck equation is then used to infer the avalanche growth rate as a function of the electric field, synchrotron radiation and effective charge. Predictions of the avalanche PINN are compared against first principle calculations of the avalanche growth rate with excellent agreement observed across a broad range of parameters.

Keywords: runaway electrons, fusion plasma

1. Introduction

The unintentional generation of a large relativistic electron population continues to pose a substantial obstacle to the success of the tokamak reactor concept. Such runaway electrons (REs) may be inadvertently generated by the strong electric fields coinciding with a tokamak disruption and obtain energies of several MeV (Hender *et al.* 2007). Due to their high energy and often localized impact, REs have the potential to induce substantial damage to plasma facing components (Matthews *et al.* 2016). Obtaining a robust description of their formation processes has thus emerged as a topic of immediate importance to tokamak devices in addition to its intrinsic interest to plasma physics.

One of the primary challenges in describing REs is the multi-physics nature of a fusion plasma. In particular, alongside a description of RE kinetics, an accurate treatment of RE formation requires the self-consistent evolution of the background magnetohydrodynamic (MHD) equilibrium, impurity transport and radiative losses. As a result, a promising approach toward developing an integrated description of RE formation and evolution involves the identification of reduced models of RE kinetics that can be coupled to

† Email address for correspondence: j.arnaud@ufl.edu

a broader plasma physics framework (Bandaru *et al.* 2021; Hoppe, Embreus & Fülöp 2021; Liu *et al.* 2021; Sainterme & Sovinec 2024). Developing such a reduced RE module has, however, posed a challenge to the plasma physics community. While several reduced models of runaway generation processes such as Dreicer generation (Dreicer 1959; Kruskal & Bernstein 1962; Connor & Hastie 1975; Hesslow *et al.* 2019*b*), hot tail generation (Helander *et al.* 2004; Smith & Verwichte 2008; Yang *et al.* 2024) or avalanche generation (Rosenbluth & Putvinski 1997; Aleynikov & Breizman 2015; Martín-Solís, Loarte & Lehnen 2015; Hesslow *et al.* 2019*a*) have been developed, these reduced models often struggle to quantitatively describe the inherently kinetic physics that characterize RE generation. Furthermore, many of these models were derived under highly simplified assumptions, where their generalization to more realistic plasma conditions is non-trivial.

The rapidly evolving field of deep learning suggests a new pathway to developing reduced RE models. While often computationally intensive to train, the online deployment of deep learning based models is typically orders of magnitude faster than traditional plasma physics codes, thus providing an efficient surrogate that may be called by a broader plasma physics framework. In the present paper, our aim will be to develop a physics-informed neural network (PINN) to provide an accurate reduced model of the RE avalanche. In contrast to purely data-driven paradigms, this deep learning approach embeds the physics model into the training of a deep neural network, enabling the PINN to make predictions in the absence of synthetic or experimental data. In so doing, the trained PINN encodes the underlying kinetic solution, allowing for greater interpretability, along with predicting quantities of interest (QoI) such as the rate of RE generation.

This approach was recently used to predict the number of hot tail seed electrons in an axisymmetric model of the thermal quench (McDevitt 2023). Our aim in the current paper is to develop a PINN to describe the avalanche amplification of REs (Sokolov 1979; Jayakumar, Fleischmann & Zweben 1993) as a function of local plasma parameters such as the electric field strength, effective charge Z_{eff} and the strength of synchrotron radiation. In contrast to the hot tail seed mechanism, the avalanche mechanism of RE formation requires a pre-existing 'seed' RE population to be present. Once established, large-angle collisions of this RE seed with the cold background plasma will result in the initially cold electrons being scattered to energies up to half of the initial energy of the seed electron. For a sufficiently large electric field, these 'secondary' electrons will be accelerated to relativistic energies, resulting in the exponential growth of the original seed. Such a process is of particular importance to reactor scale tokamak plasmas due to the ability of this mechanism to convert the ohmic plasma current into RE current, even in the limit where only a minuscule seed population of REs is present (Martín-Solís, Loarte & Lehnen 2017; Vallhagen *et al.* 2020).

The remainder of this paper is organized as follows: § 2 provides a brief overview of the physics-informed deep learning framework employed. The adjoint of the relativistic Fokker–Planck equation, together with its solution, is described in § 3. Section 4 evaluates the avalanche growth rate and threshold across a broad range of parameters and verifies the predictions against those from a traditional RE solver. Conclusions along with a brief discussion are given in § 5.

2. Physics-constrained deep learning

Physics-constrained deep learning methods have emerged as a powerful means to efficiently describe complex physical processes. In addition to exploiting available data, such methods seek to embed physical constraints into the training of a neural network (Lagaris, Likas & Fotiadis 1998; Karpatne *et al.* 2017; Lusch, Kutz & Brunton 2018; Wang *et al.* 2020; Karniadakis *et al.* 2021), thus providing a natural means of avoiding

https://doi.org/10.1017/S0022377824000679 Published online by Cambridge University Press

overfitting, along with allowing for greater generalizability to unseen parameter regimes. Physics-informed neural networks (Karniadakis *et al.* 2021) have emerged as a particularly prominent example. A PINN, in its simplest form, incorporates the partial differential equation (PDE), boundary, and initial conditions into the loss function, yielding

$$\text{Loss} = \frac{1}{N_{\text{PDE}}}\sum_{i}^{N_{\text{PDE}}} \mathcal{R}^2\left(\boldsymbol{p}_i, t_i; \boldsymbol{\lambda}_i\right) + \frac{1}{N_{\text{bdy}}}\sum_{i}^{N_{\text{bdy}}}\left[P_i - P\left(\boldsymbol{p}_i, t_i; \boldsymbol{\lambda}_i\right)\right]^2 + \frac{1}{N_{\text{init}}}\sum_{i}^{N_{\text{init}}}\left[P_i - P\left(\boldsymbol{p}_i, t=0; \boldsymbol{\lambda}_i\right)\right]^2, \tag{2.1}$$

where $P(\boldsymbol{p}, t; \boldsymbol{\lambda})$ is the dependent variable (the runaway probability function for the present work), $\mathcal{R}(\boldsymbol{p}, t; \boldsymbol{\lambda})$ is the residual of the PDE, $\boldsymbol{p}$ and t are the independent variables (momentum and time) and $\boldsymbol{\lambda}$ represents parameters of the physical system. Here, the first term represents the loss against the PDE, the second term represents the loss against the boundary conditions and the third term represents the loss against the initial conditions for time-dependent problems. Noting that derivatives of the neural network output P with respect to its inputs $(\boldsymbol{p}, t; \boldsymbol{\lambda})$ can be evaluated by automatic differentiation, a feature provided by standard machine learning libraries (Abadi *et al.* 2016; Paszke *et al.* 2017), no discretization of the PDE is required. As a result, PINNs are inherently mesh free and only require specification of a distribution of training points. After minimization of the loss function, the dependent variable P will satisfy the PDE, boundary and initial conditions up to the loss defined by (2.1). Hence, a PINN provides a means of solving PDEs, where the value of the loss achieved after training provides an estimate of the accuracy of the solution.

A powerful property of PINNs is that they may be used to learn the parametric solution to a PDE (McDevitt, Fowler & Roy 2024; Sun *et al.* 2020). In particular, since the parameters of the physical problem $\boldsymbol{\lambda}$ are inputs into the neural network, after minimization of the loss, the PINN can predict $P(\boldsymbol{p}, t; \boldsymbol{\lambda})$ across a broad range of parameters $\boldsymbol{\lambda}$. While obtaining a parametric solution of a PDE often requires extensive offline training, the online execution of a PINN is rapid, where an individual prediction typically requires a few microseconds. Physics-informed neural networks thus provide a framework for developing efficient surrogate models of a PDE. A significant limitation of the above approach, however, is that PINNs often fail to train when treating the challenging PDEs that characterize many scientific and engineering applications (Wang, Yu & Perdikaris 2022). A primary aim of this paper will therefore be to develop custom output layers to the PINN that enable it to robustly evaluate the adjoint to the relativistic Fokker–Planck equation across a broad range of plasma conditions in the absence of synthetic or experimental data.

When carrying out the training of the PINN, a single Nvidia A100 GPU will be used. A fully connected feedforward neural network is employed, containing six hidden layers, each having a width of 64 neurons. Roughly a million training points are used and distributed across the input space $(\boldsymbol{p}, t, \boldsymbol{\lambda})$ according to a Hammersley distribution. The first 15,000 steps of training are done with the adaptive moment estimation (ADAM) optimizer (Kingma & Ba 2014), whereas the rest of the training is done with the limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) optimizer (Nocedal 1980). During training a small number of training points are added to regions where the residual is maximal. An independent set of test points are also applied to verify accuracy of predictions away from training points. The python script used for training the PINN is

https://doi.org/10.1017/S0022377824000679 Published online by Cambridge University Press

written using the DeepXDE library (Lu *et al.* 2021) with TensorFlow (Abadi *et al.* 2016) as the backend and can be found in the RunAwayPINNs github repository.

3. Steady-state runaway probability function

3.1. *Adjoint relativistic Fokker–Planck equation*

This section describes the implementation of the adjoint to the relativistic Fokker–Planck equation. While a direct evaluation of the particle distribution via a PINN is possible (see McDevitt & Tang (2024) for example), obtaining an accurate description of the runaway electron population is particularly challenging due to the very small number of electrons expected to run away. This electron population, typically orders of magnitude less than the bulk electron population, will provide a negligible contribution to the loss function defined by (2.1) unless special care is taken. An alternative means of describing REs is via the solution of the adjoint to the relativistic Fokker–Planck equation. This approach has been pursued in a variety of contexts, including wave-driven currents in magnetized fusion plasmas (Antonsen & Chu 1982; Taguchi 1983; Fisch 1986; Karney & Fisch 1986) and RE formation (Karney & Fisch 1986; Liu *et al.* 2016, 2017; Zhang & del Castillo-Negrete 2017; McDevitt 2023). As will be discussed below, the adjoint problem treated in this analysis describes the probability P of an electron at an initial momentum space location running away at a later time. Thus, its range will be $P \in (0, 1)$ and provides a quantity well suited for the optimization problem posed in (2.1). In the absence of synchrotron radiation, the adjoint to the steady-state relativistic Fokker–Planck equation can be written as

$$\left[-E_\parallel \xi - C_F\right] \frac{\partial P}{\partial p} - \left(1-\xi^2\right) \frac{E_\parallel}{p} \frac{\partial P}{\partial \xi} = -\frac{\nu_D}{2} \frac{\partial}{\partial \xi} \left[\left(1-\xi^2\right) \frac{\partial P}{\partial \xi}\right], \tag{3.1a}$$

where the collisional coefficients are taken to be

$$\nu_D = (1+Z_{\text{eff}}) \frac{\gamma}{p^3}, \quad C_F = \frac{1+p^2}{p^2}. \tag{3.1b, c}$$

Here, the relativistic momentum p is normalized to the electron mass m_e and speed of light $c\,p \to p/(m_e c)$, the electron's pitch is defined by $\xi \equiv p_\parallel/p$, time is normalized as $t \to t/\tau_c$, where $\tau_c \equiv 4\pi\epsilon_0^2 m_e^2 c^3/(e^4 n_e \ln\Lambda)$ is the collision time of a relativistic electron for a given free electron density n_e and Coulomb logarithm $\ln\Lambda$ with the physical constants of the permittivity of free space ϵ_0 and elementary charge e, the collisional coefficients ν_D and C_F are normalized to τ_c and the parallel electric field is normalized to the Connor–Hastie electric field $E_\parallel \to E_\parallel/E_c$, where $E_c \equiv m_e c/(e\tau_c)$ (Connor & Hastie 1975). Energy diffusion has been neglected due to this term being exceptionally small for conditions typical of a tokamak disruption. In particular, noting that the energy diffusivity scales with the electron temperature $T_e/(m_e c^2)$ for the high energies characteristic of REs, we expect this approximation to be well satisfied for a low temperature post-thermal quench plasma. Furthermore, the collision frequencies used in (3.1) assume the limit $v > v_{\text{Te}}$, where v_{Te} is the electron thermal velocity. For the parameters of interest, the critical speed for an electron to run away will be much larger than v_{Te}, hence we anticipate that this approximation will not substantially impact our results.

Concerning the boundary conditions, we will enforce $P = 0$ at $p = p_{\min}$ and $P = 1$ along the upper boundary, where the energy flux $U_p \equiv -E_\parallel \xi - C_F$ is positive. Specifically, the high energy boundary condition enforces that P is unity for values of the pitch ξ where electric field acceleration exceeds collisional drag such that the electron will be accelerated out of the simulation domain. Further noting that for values of the pitch near $|\xi| \approx 1$, (3.1*a*) is nearly hyperbolic, we anticipate that electrons accelerated through the

https://doi.org/10.1017/S0022377824000679 Published online by Cambridge University Press

high energy boundary will be accelerated to arbitrarily high energies, and thus have a zero probability of returning to the simulation domain. Hence, any electron located at the high energy boundary $p_{\max}$ with $U_p > 0$, is treated as a RE. With these boundary conditions, the quantity $P(p, \xi)$ indicates the probability that an electron initially located at (p, ξ) will run away at a later time, and is often referred to as the runaway probability function (RPF) (Karney & Fisch 1986).

When near the threshold electric field for avalanche generation, synchrotron radiation substantially impacts the RPF. While adding synchrotron radiation leads to a modest modification of (3.1), it does complicate the physical interpretation of the RPF. In particular, with the inclusion of synchrotron radiation, the adjoint equation takes the form

$$\left[-E_\parallel \xi - C_F - \alpha\gamma p\left(1-\xi^2\right)\right]\frac{\partial P}{\partial p} + \left(1-\xi^2\right)\left[-\frac{E_\parallel}{p} + \alpha\frac{\xi}{\gamma}\right]\frac{\partial P}{\partial \xi} = -\frac{\nu_D}{2}\frac{\partial}{\partial \xi}\left[\left(1-\xi^2\right)\frac{\partial P}{\partial \xi}\right], \tag{3.2}$$

where the strength of synchrotron radiation is set by the parameter $\alpha \equiv \tau_c/\tau_s$, where $\tau_s \equiv 6\pi\epsilon_0 m_e^3 c^3/(e^4 B^2)$. Here, (3.2) is solved similarly as before with a boundary condition of $P = 0$ at $p = p_{\min}$ along with the condition that $P = 1$ on the high energy boundary $p = p_{\max}$ when the energy flux, now defined by $U_p^{(\alpha)} \equiv -E_\parallel \xi - C_F - \alpha\gamma p(1-\xi^2)$, is positive. While this problem formulation is directly analogous to the case that neglects synchrotron radiation, the physical interpretation of the RPF is slightly modified. Specifically, as shown in Andersson, Helander & Eriksson (2001), Decker *et al.* (2016), Guo, McDevitt & Tang (2017) and McDevitt, Guo & Tang (2018), synchrotron radiation damping and pitch-angle scattering results in electrons obtaining a saturated energy, achieved via the formation of a circulation pattern in momentum space centred about an O-point. Thus, an electron accelerated through the high energy boundary will have a non-zero probability of returning to the simulation domain after a finite time, rather than being accelerated to arbitrarily high energy, as was the case when synchrotron radiation was neglected. For the case with synchrotron radiation, the RPF should thus be given the narrower interpretation as the probability that an electron reaches the high energy boundary of the simulation domain before slowing down to the low energy bulk, rather than the probability of an electron being accelerated to arbitrarily high energy. For this analysis the high energy boundary was chosen to be 5 MeV, which is a sufficiently high energy to capture the critical energy for an electron to run away in all cases except when asymptotically close to threshold.

3.2. *Embedding physical constraints into the PINN*

Our aim in this section will be to develop a PINN framework capable of robustly representing solutions to (3.2) across the three-dimensional parameter space $(E_\parallel, Z_{\text{eff}}, \alpha)$. Here, the inputs of the PINN will be $(p, \xi, E_\parallel, Z_{\text{eff}}, \alpha)$ and the output will be the RPF P. A key component of our description is enforcing a subset of physical properties of the RPF as hard constraints. In particular, we will (i) enforce the low energy boundary condition $P = 0$ at $p = p_{\min}$, (ii) constrain the RPF to have a range between zero and one and (iii) ensure the RPF vanishes when $|E_\parallel| < 1$. These three constraints are enforced by introducing a customized output layer to the neural network of the form

$$P'(p,\xi) \equiv \frac{1}{2}\left[1 + \tanh\left(\frac{E_\parallel - 1}{\Delta E}\right)\right]\left(\frac{p - p_{\min}}{p_{\max} - p_{\min}}\right) P_{\text{NN}}(p,\xi), \tag{3.3a}$$

$$P(p,\xi) \equiv \tanh\left(P'^2(p,\xi)\right). \tag{3.3b}$$

https://doi.org/10.1017/S0022377824000679 Published online by Cambridge University Press

Here, P_{NN} is the output of the hidden layers of the neural network, and ΔE is a hyperparameter whose value should satisfy $\Delta E < 1$, where for all cases in this paper it is taken to be $\Delta E = 0.1$. From (3.3) it can be verified that, regardless of the value of P_{NN}, the predicted RPF $P(p, \xi)$: (i) vanishes at the low energy boundary, (ii) has a range of zero to one and (iii) tends to zero for $E_\parallel < 1$ and $\Delta E \ll 1$ (only positive electric fields are considered when training the RPF PINN). We note in passing that the output layer defined by (3.3) results in both the value and the first derivative of P vanishing at $p = p_{\mathrm{min}}$. While this latter condition is not required when defining the RPF, it will nevertheless be satisfied as long as the value of p_{min} chosen is well below the momentum $p_{\mathrm{RE}} \equiv 1/\sqrt{E_\parallel - 1}$ where collisional drag exceeds electric field acceleration. Specifically, since energy diffusion is neglected in the present analysis, electrons located below p_{RE} will have zero probability of running away. Thus, for $p < p_{\mathrm{RE}}$ the solution of the RPF will be a constant with a magnitude of zero, and thus is consistent with pure Dirichlet and Neumann boundary conditions at $p = p_{\mathrm{min}}$.

The loss function employed is taken to have the form

$$\mathrm{Loss} = \frac{1}{N_{\mathrm{PDE}}} \sum_i^{N_{\mathrm{PDE}}} \left[\left(\frac{p_i^2}{1 + p_i^2} \right) \mathcal{R}\left(p_i, \xi_i; \boldsymbol{\lambda}_i\right) \right]^2 + \frac{1}{N_{\mathrm{bdy}}} \sum_i^{N_{\mathrm{bdy}}} \left[1 - P\left(p_i, \xi_i; \boldsymbol{\lambda}_i\right)\right]^2, \tag{3.4}$$

where $\mathcal{R}(p_i, \xi_i; \boldsymbol{\lambda}_i)$ is the residual of (3.2) and $\boldsymbol{\lambda}$ represents the physics parameters $(E_\parallel, Z_{\mathrm{eff}}, \alpha)$. Here, the first term in (3.4) penalizes deviations from the governing PDE given by (3.2), where the prefactor $p_i^2/(1 + p_i^2)$ removes the low energy divergence of the pitch-angle scattering operator. The second term in the loss function defined by (3.4) enforces the high energy boundary condition, i.e. $P = 1$ at p_{max} when $U_p^{(\alpha)} > 0$. In particular, the boundary points N_{bdy}, will only be applied at locations that satisfy both $U_p^{(\alpha)} > 0$ and $p = p_{\mathrm{max}}$.

3.3. *Parametric dependence of the runaway probability function*

In this section, we will seek to obtain solutions to the PINN in the five-dimensional space defined by the two independent coordinates (p, ξ) and the three physics parameters $(E_\parallel, Z_{\mathrm{eff}}, \alpha)$. The loss history of the PINN is shown in figure 1. Here, after roughly 100 000 epochs, the loss associated with the PDE drops below 10^{-7}, along with the boundary loss dropping to $\approx 10^{-9}$, with the loss dropping more slowly for the remaining $\sim$400 000 epochs. The test loss of the PDE reaches $\approx 10^{-9}$ by the end of the training, indicating that an accurate solution was found. The periodic spikes in the training data are due to additional training points sampled after every 100 000 epochs at locations where the residual is maximal. The test loss is only updated after every 100 000 epochs when using the L-BFGS optimizer, leading to the sharp variations evident in the dashed curves.

Four example predictions of the RPF are shown in figure 2. Here, the RPF vanishes at low energies, where drag exceeds electric field acceleration, but increases at higher energy due to the $(1 + p^2)/p^2$ drop in the collisional drag. In particular, for the parameters indicated in figure 2(*a*), the $P = 0.5$ contour is located at an approximate energy of 200 keV for $\xi = -1$. Considering a case with a large Z_{eff} (see figure 2(*b*) with $Z_{\mathrm{eff}} = 10$), the $P = 0.5$ contour shifts to higher energy, with a more gradual transition between the $P \approx 0$ and $P \approx 1$ regions. The impact of synchrotron radiation on the RPF is shown in figure 2(*c*), where the magnitude of synchrotron radiation was taken to be $\alpha = 0.2$. Compared with an otherwise identical case, but without synchrotron radiation (figure 2*a*), it is evident that the location of the $P = 0.5$ contour at $\xi = -1$ has only shifted slightly. This is due to synchrotron radiation vanishing for $\xi = -1$ and having a modest magnitude

https://doi.org/10.1017/S0022377824000679 Published online by Cambridge University Press

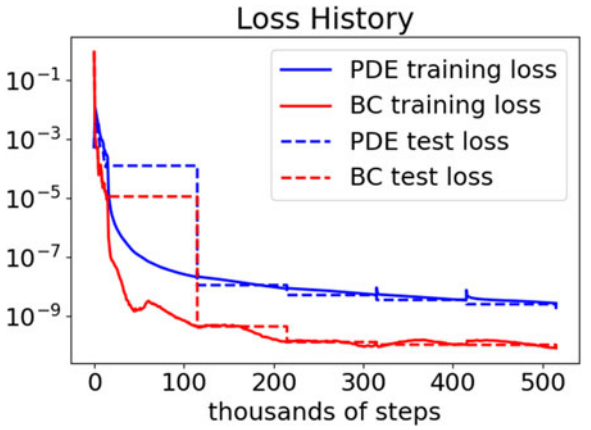


FIGURE 1. Loss history (blue lines for the PDE and red lines for the boundary condition (BC)) for a feedforward neural network with six hidden layers each with a width of 64 neurons, along with roughly 1 000 000 training points. Here, 15 000 epochs were performed with the ADAM optimizer, with the remaining steps performed using L-BFGS. The model was trained across $E_\parallel \in (1, 10)$, $Z_{\text{eff}} \in (1, 10)$ and $\alpha \in (0, 0.2)$. The range of p was chosen such that the low energy boundary was 10 keV and the high energy boundary was 5 MeV.

at low energies. In contrast, the RPF at high energies is more substantially impacted, where the region with $P \approx 1$ is now largely localized to negative values of pitch for the energy range considered. Finally, increasing the electric field to $E_\parallel = 10$, with $Z_{\text{eff}} = 1$ and $\alpha = 0$, results in a drop in the location of the $P = 0.5$ contour (compare figure 2*a*,*d*), due to the electric field being able to overcome collisional drag for a larger range of energies. The residual of (3.2), multiplied by the prefactor $p^2/(1 + p^2)$ contained in the loss defined in (3.4), is shown in figure 3 for the four cases described above. Here, it is evident that the residuals have been reduced to a small value with a maximum of $\approx$0.03 for the largest electric field case (see figure 3*d*). A direct assessment of the accuracy of the PINN's prediction of the avalanche growth rate will be made in § 4.3, where the predictions of the PINN are compared with first principle Monte Carlo simulations.

3.4. *Critical energy to run away*

The location of the $P = 0.5$ contour $(p_{\text{crit}}, \xi_{\text{crit}})$ provides a useful reference point for identifying the critical energy and pitch above which electrons are likely to run away. The parametric solution of the PDE given by the PINN thus provides $(p_{\text{crit}}, \xi_{\text{crit}})$ across the entire domain by simply extracting the $P = 0.5$ surface. To show the dependence of $(E_\parallel, Z_{\text{eff}}, \alpha)$ on the critical energy to run away, we evaluate the energy at which $P = 0.5$ and $\xi = -1$ across the entire training region, which is shown in figure 4. Here, the $\log_{10}$ of the critical energy in units of eV is plotted, where the empty regions represent scenarios where $E_\parallel$ is below the avalanche threshold E_{av} (see § 4.1 below). The dependence of $E_\parallel$ and Z_{eff} on the critical energy can be seen in figure 4(*a*) for $\alpha = 0$, where increasing the electric field from $E_\parallel = 2$ to $E_\parallel = 10$ decreases the critical energy from $\sim$1 MeV to $\sim$36 keV. At a given electric field (e.g. $E_\parallel = 4$), the critical energy increases by almost an order of magnitude as Z_{eff} is increased from one to ten, indicating that higher Z elements in the plasma increase the critical energy for REs. Moreover, the addition of synchrotron radiation (compare figure 4*b*,*c*), shows that the region where $E_\parallel < E_{\text{av}}$ is increased (larger white region). As α is increased further (compare figure 4*a*,*c*), however, we also see that

https://doi.org/10.1017/S0022377824000679 Published online by Cambridge University Press

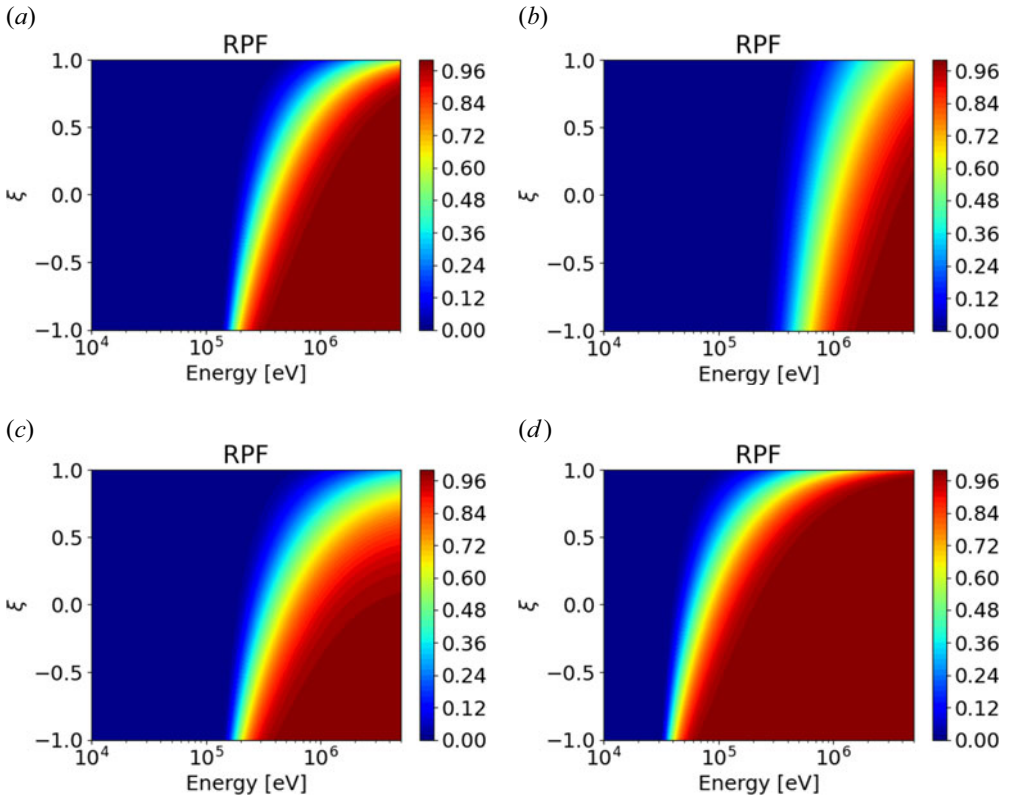


FIGURE 2. Runaway probability functions for different values of the physics parameters $(E_\parallel, Z_{\text{eff}}, \alpha)$. Panel (a) is for $(E_\parallel = 3, Z_{\text{eff}} = 1, \alpha = 0)$, panel (b) is for $(E_\parallel = 3, Z_{\text{eff}} = 10, \alpha = 0)$, panel (c) is for $(E_\parallel = 3, Z_{\text{eff}} = 1, \alpha = 0.2)$ and panel (d) is for $(E_\parallel = 10, Z_{\text{eff}} = 1, \alpha = 0)$.

the shape and range of the critical energy varies modestly, indicating the critical energy is modestly impacted by α when $E_\parallel \gg 1$.

4. Surrogate model of the avalanche growth rate

4.1. *Secondary source of runaway electrons*

Our aim in this section will be to utilize the RPF PINN described in § 3.3 to estimate the avalanche growth rate. Denoting the source of secondary electrons as $S(p, \xi)$, the rate that REs form due to the avalanche mechanism can be expressed as (Liu *et al.* 2017)

$$\left.\frac{\mathrm{d}n_{RE}}{\mathrm{d}t}\right|_{\text{av}} = \int \mathrm{d}^3 p S(p, \xi) P(p, \xi), \tag{4.1}$$

where n_{RE} is the density of REs. Here, $S(p, \xi)$ indicates the rate and momentum space distribution of REs, whereas $P(p, \xi)$ indicates the probability that an electron at a given momentum space location (p, ξ) runs away. By integrating over momentum space, $S(p, \xi)P(p, \xi)$ will thus indicate the expected rate that REs form due to the avalanche mechanism. The primary challenge with evaluating (4.1) is due to $S(p, \xi)$ depending on the electron distribution f_e, i.e.

$$S(p, \xi) = \int \mathrm{d}^3 p' S_0\left(p', \xi', p, \xi\right) f_e\left(p', \xi'\right), \tag{4.2}$$

https://doi.org/10.1017/S0022377824000679 Published online by Cambridge University Press

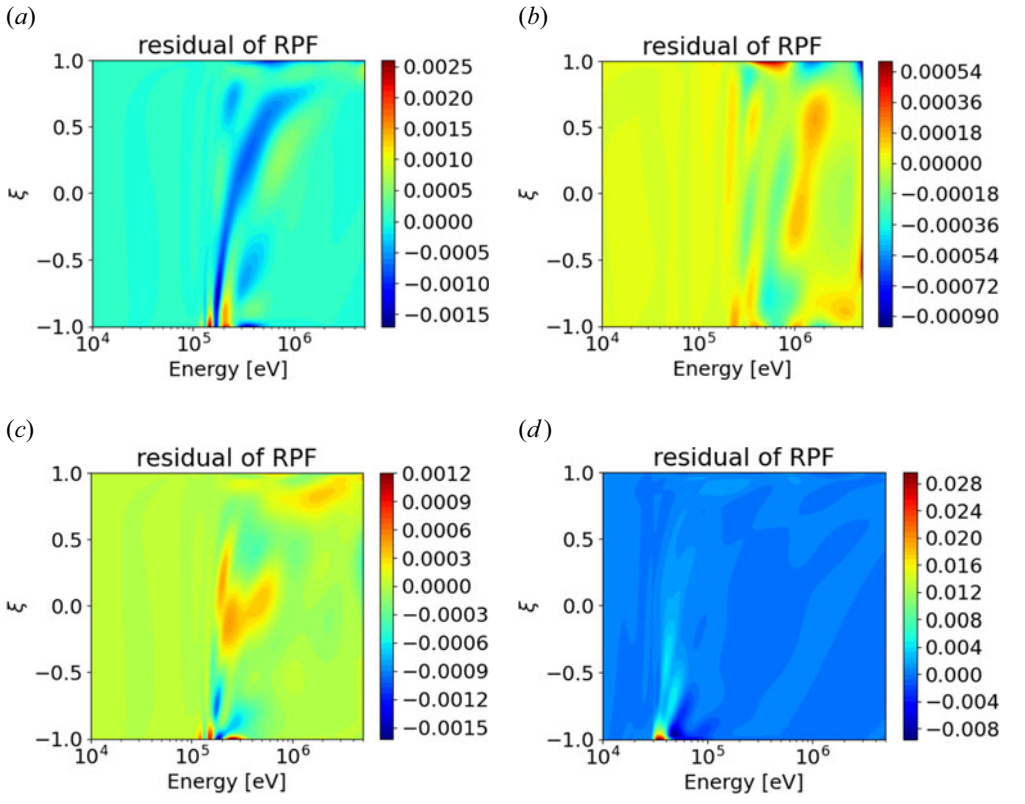


FIGURE 3. Residual of (3.2) (multiplied by the prefactor $p^2/(1+p^2)$) for different values of the physics parameters $(E_\parallel, Z_{\text{eff}}, \alpha)$. Panel (a) is for $(E_\parallel = 3, Z_{\text{eff}} = 1, \alpha = 0)$, panel (b) is for $(E_\parallel = 3, Z_{\text{eff}} = 10, \alpha = 0)$, panel (c) is for $(E_\parallel = 3, Z_{\text{eff}} = 1, \alpha = 0.2)$ and panel (d) is for $(E_\parallel = 10, Z_{\text{eff}} = 1, \alpha = 0)$.

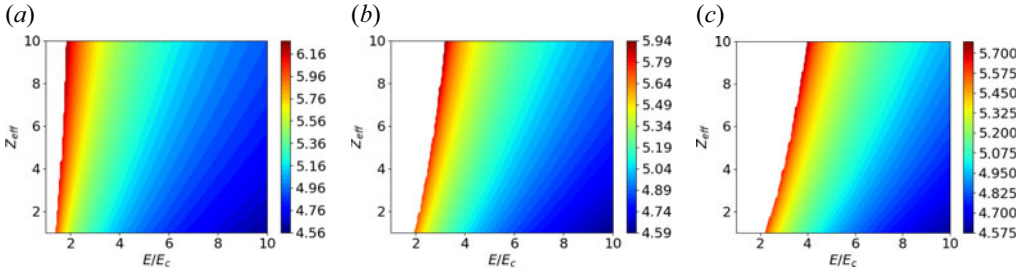


FIGURE 4. Value of $\log_{10} E_{\text{crit}}$, where E_{crit} is the critical energy to run away in eV. Panel (a) is for $\alpha = 0$, panel (b) is for $\alpha = 0.1$ and panel (c) is for $\alpha = 0.2$.

where $S_0(p', \xi', p, \xi)$ is defined by

$$S_0\left(p', \xi', p, \xi\right) = n_e c r_e^2 \frac{v'}{2\pi p^2} \frac{\mathrm{d}\sigma_M\left(p', p\right)}{\mathrm{d}p} \Pi\left(p', \xi', p; \xi\right), \tag{4.3}$$

where $r_e = e^2/(4\pi\epsilon_0 m_e c^2)$ is the classical electron radius, $\mathrm{d}\sigma_M/\mathrm{d}p$ is the Møller cross-section (Møller 1932; Ashkin, Page & Woodward 1954), $\Pi(p', \xi', p; \xi)$ describes

https://doi.org/10.1017/S0022377824000679 Published online by Cambridge University Press

the pitch-angle dependence of secondary electron generation (see Boozer (2015) for an explicit expression), $\mathrm{d}^3p = 2\pi p^2\,\mathrm{d}p\,\mathrm{d}\xi$ and all variables have been non-dimensionalized $p \to p/m_e c$, $v' \to v'/c$, $\sigma_M \to \sigma_M/r_e^2$. Noting that the solution of the adjoint relativistic Fokker–Planck equation does not directly yield the RE distribution $f_e(p', \xi')$, a closure relation will need to be introduced to evaluate the rate of RE generation via avalanching. The simplest closure, introduced in Rosenbluth & Putvinski (1997), involves taking the limit where REs are assumed to have asymptotically high energies and a pitch of $\xi = -1$. While idealized, this closure has been shown to provide a good approximation to the full Møller source evaluated using a self-consistently computed RE distribution (McDevitt *et al.* 2018). In the limit of $p' \to \infty$ and $\xi' = -1$, (4.2) asymptotes to

$$S\left(p,\xi\right) = n_e n_{\mathrm{RE}} c r_e^2 \frac{v}{\gamma^2 - 1}\frac{1}{\left(\gamma - 1\right)^2}\delta\left(\xi - \xi_1\right), \tag{4.4}$$

where we have introduced the Lorentz factor $\gamma \equiv \sqrt{1 + p^2}$ and ξ_1 is defined by

$$\xi_1 = -\sqrt{\frac{\gamma - 1}{\gamma + 1}}. \tag{4.5}$$

Using (4.4), (4.1) reduces to

$$\left.\frac{\mathrm{d}n_{RE}}{\mathrm{d}t}\right|_{av} = 2\pi n_e n_{\mathrm{RE}} c r_e^2 \int \mathrm{d}p\, p^2 \frac{v}{\left(\gamma^2 - 1\right)\left(\gamma - 1\right)^2} P\left(p, \xi_1\right). \tag{4.6}$$

Noting that the right-hand side of (4.6) is directly proportional to the RE density n_{RE}, this implies an exponentially growing solution with a growth rate given by

$$\tau_c \gamma_{\mathrm{av}} = \frac{1}{2\ln\Lambda}\int \mathrm{d}p\, \frac{v}{\left(\gamma - 1\right)^2} P\left(p, \xi_1\right), \tag{4.7}$$

where ξ_1 is defined by (4.5). Thus, once $P(p, \xi)$ has been evaluated, the avalanche growth rate can be directly inferred from (4.7).

A caveat when evaluating (4.7) is that, while this integral formally extends to $p \to \infty$, the RPF is evaluated assuming a finite p_{max}. The error induced by this approximation can be estimated by considering the magnitude of the integrand in (4.7) across the range of momenta used in this paper (see figure 5, where $p_{\mathrm{min}} \approx 0.2$ and $p_{\mathrm{max}} \approx 10.74$). Here, we have taken $P = 1$ inside the integrand of (4.7), such that figure 5 provides an upper bound on the value of the integrand. Noting that the integrand has decayed to a value of $\approx 3.5 \times 10^{-4}$ at the upper boundary, this implies a small contribution to the avalanche growth rate for secondary electrons born with $p > p_{\mathrm{max}}$, particularly for parameters where the system is well above marginality.

4.2. *Parametric dependence of RE avalanche*

Using the RPF evaluated in § 3, (4.7) can be used to infer the avalanche growth rate across the parameter space $(E_\parallel, Z_{\mathrm{eff}}, \alpha)$ (see figure 6). Here, the avalanche growth rate increases approximately linearly with the electric field when $E_\parallel \gg 1$, with the slope and threshold sensitive to Z_{eff} and α. Specifically, the value of α significantly impacts the avalanche threshold (i.e. where $\gamma_{\mathrm{av}} \approx 0$), but has a negligible impact at large electric fields. In contrast, Z_{eff} strongly impacts the avalanche growth rate for all values of the electric field. A feature of (4.7) is that since it indicates the number of REs generated

https://doi.org/10.1017/S0022377824000679 Published online by Cambridge University Press

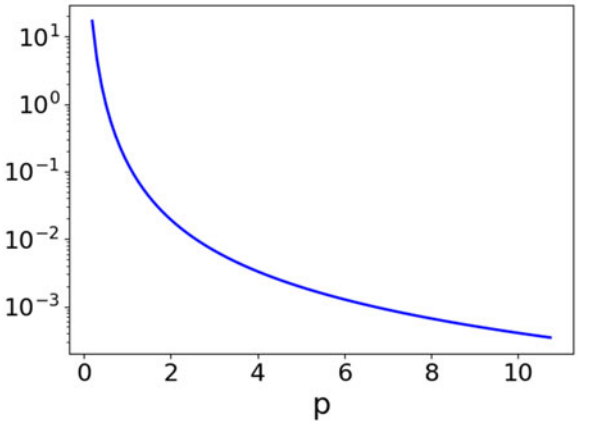


FIGURE 5. The integrand $v/(\gamma-1)^2/(2\ln\Lambda)$ of (4.7) when $P=1$. The low and high energy bounds are $p_{\min}\approx 0.2$ and $p_{\max}\approx 10.74$, respectively. The Coulomb logarithm was taken to be $\ln\Lambda=15$.

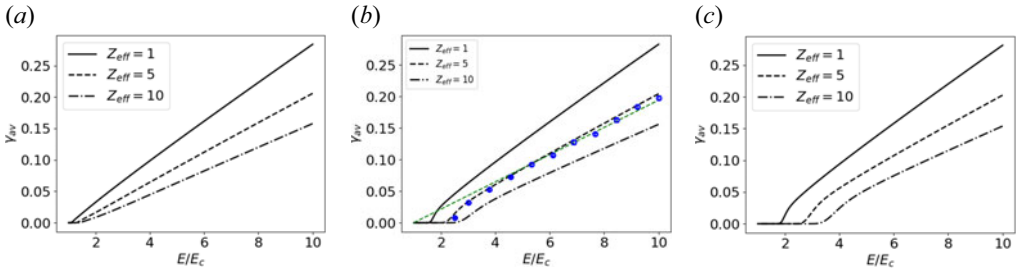


FIGURE 6. Avalanche growth rate vs electric field for different values of $Z_{\rm eff}$ and α. Panel (a) is for $\alpha=0$, panel (b) is for $\alpha=0.1$ and panel (c) is for $\alpha=0.2$. The blue points in panel (b) represent Monte Carlo simulations with $n_e=5\times10^{21}$ m^{-3}, and (4.8) is shown as the dashed green curve. The Coulomb logarithm was taken to be $\ln\Lambda=15$.

via large-angle collisions, it is thus positive definite. As a result, avalanche growth rates predicted by (4.7) will not account for the decay of the RE population when $E_\parallel<E_{\rm av}$ and will instead asymptote to zero. While such behaviour is strictly correct when focusing solely on the avalanche growth mechanism, when developing a RE model appropriate for coupling with a self-consistent MHD solver it will be necessary to account for the decay of the RE distribution when $E_\parallel<E_{\rm av}$. The inclusion of this additional physics will be the subject of future work.

It will be of interest to compare the predictions of the PINN with available analytic theories. Considering the avalanche growth rate given by (Rosenbluth & Putvinski 1997)

$$\tau_c\gamma_{\rm RP}=\frac{1}{\ln\Lambda}\sqrt{\frac{\pi}{3\,(Z_{\rm eff}+5)}}\left(E_\parallel-1\right), \tag{4.8}$$

a comparison between the PINN predictions and (4.8) is shown in figure 6(b). For reference, a small number of avalanche growth rates computed by a Monte Carlo code (McDevitt, Guo & Tang 2019) using the complete Møller source are also shown (the blue points on figure 6b). Details on how this Monte Carlo data set was generated are given in § 4.3 below. It is apparent that both the PINN and (4.8) yield results in reasonable

https://doi.org/10.1017/S0022377824000679 Published online by Cambridge University Press

agreement with the Monte Carlo points, where the PINN yields more accurate results near threshold. In particular, while (4.8) implies an avalanche threshold field of $E_\parallel = 1$ regardless of the values of (Z_{eff}, α), the PINN is able to account for the variation of the avalanche threshold for non-zero values of α and different Z_{eff}. In addition, at larger electric fields a modest difference in the slope of the avalanche growth rate predicted by (4.8) compared with the Monte Carlo data is evident. The predictions of the PINN, in contrast, are able to recover the correct slope of the avalanche growth rate, albeit with a modest offset in the magnitude as will be discussed further in § 4.3 below.

When well above marginality the avalanche growth rate is most conveniently described by evaluating the amount of poloidal flux needed to increase the amplitude of a RE seed by one order of magnitude (Rosenbluth & Putvinski 1997; Boozer 2018). This quantity can be evaluated by noting that, for $E_\parallel \gg 1$, the avalanche growth rate scales linearly with $E_\parallel$, i.e.

$$\gamma_{\text{av}} = \gamma_0 \left(\frac{E_\parallel}{E_c} - 1 \right) \approx \gamma_0 \frac{E_\parallel}{E_c} \approx \frac{\gamma_0}{E_c} \frac{1}{R_0} \frac{\partial \psi}{\partial t}, \tag{4.9}$$

where γ_0 is a constant that depends on $(Z_{\text{eff}}, \alpha, \ln \Lambda)$, R_0 is the major radius and ψ is the poloidal flux function. After integrating (4.9) over the time interval $t_f - t_i$, the number of exponentials of the RE population can be written in terms of the change of poloidal flux, and a constant ψ_{exp} that defines the efficiency of the avalanche. In particular, integrating (4.9) over the time interval $t_f - t_i$, yields

$$N_{\text{exp}} = \int_{t_i}^{t_f} \mathrm{d}t \, \gamma_{\text{av}} \approx \frac{\gamma_0}{E_c} \frac{1}{R_0} \int_{t_i}^{t_f} \mathrm{d}t \, \frac{\partial \psi}{\partial t} = \frac{\Delta \psi}{\psi_{\text{exp}}}, \tag{4.10}$$

where $\Delta \psi \equiv \psi(t_f) - \psi(t_i)$ and we have defined $\psi_{\text{exp}} \equiv R_0 E_c / \gamma_0$, which is related to the amount of poloidal flux required to effect one exponential amplification of the RE population. Thus, once γ_0 is inferred, the efficiency through which the decay of the poloidal flux leads to an amplification of the seed RE population can be evaluated. Further defining the quantity $\psi_{10} \equiv \ln 10 \psi_{\text{exp}}$ (which delineates base ten amplifications), the efficiency of the avalanche growth rate for a broad range of parameters is shown in figure 7(*a*). Here, the avalanche growth rate is most efficient for low values of Z_{eff} and increases by nearly a factor of two for $Z_{\text{eff}} = 10$, agreeing with previous results (McDevitt *et al.* 2019). We have not indicated the dependence of ψ_{10} on α since, for large electric fields, the avalanche growth rate will be independent of α.

In addition, the avalanche threshold is strongly impacted by the parameters (Z_{eff}, α). In particular, as Z_{eff} or α are increased, the threshold electric field E_{av}, where the avalanche growth rate is zero, is increased. The dependence of the avalanche threshold on (Z_{eff}, α) is shown in figure 7(*b*), where the PINN predictions are the solid lines. The Monte Carlo results are the 'o' markers, and the dashed lines represent empirical fit formula given by (B 15) of McDevitt *et al.* (2019). Here, since the avalanche growth rates predicted by the PINN are positive definite, we define the avalanche threshold as the value of the electric field where $\tau_c \gamma_{\text{av}} = 2 \times 10^{-3}$. From figure 7(*b*) it is apparent that the PINN predicts E_{av} particularly well across α for $Z_{\text{eff}} = 1$, with somewhat larger deviations evident for $Z_{\text{eff}} = 5$. The systematic under-prediction of the avalanche threshold is due to the assumption in (4.7) that primary electrons have infinite energy. Such an assumption overestimates the number of secondary electrons generated and, perhaps more importantly, neglects that the primary electron population itself will slowly decay to the thermal bulk when $E \approx E_{\text{av}}$.

https://doi.org/10.1017/S0022377824000679 Published online by Cambridge University Press

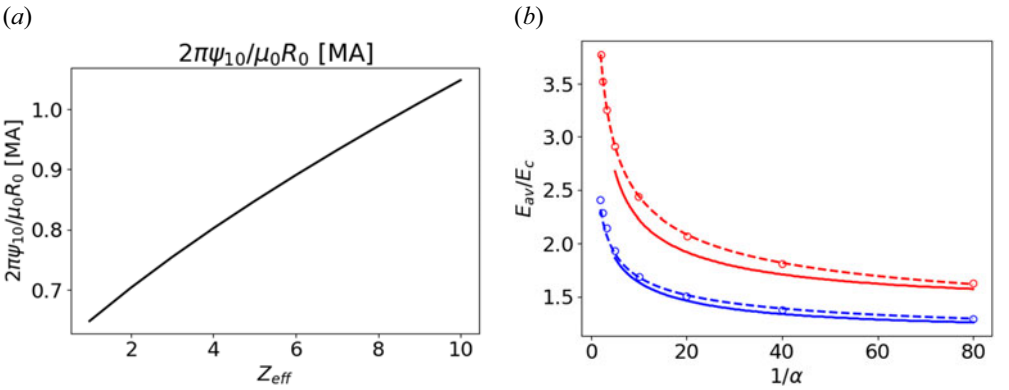


FIGURE 7. (*a*) The value of $2\pi\psi_{10}/\mu_0 R_0$ as a function of Z_{eff}, with a Coulomb logarithm of $\ln \Lambda = 15$. (*b*) Avalanche threshold E_{av} as a function of the synchrotron radiation strength α. The solid lines represent the PINN predictions, the dashed lines and 'o' markers represent (B 15) and Monte Carlo results from McDevitt *et al.* (2019), respectively, and WebPlotDigitizer was used to extract the Monte Carlo values. The blue values represent $Z_{\text{eff}} = 1$ and the red values represent $Z_{\text{eff}} = 5$. The Coulomb logarithm was taken to be $\ln \Lambda = 20$.

4.3. *Verification of the surrogate model*

In this section we will verify the PINN's predictions of the avalanche growth rate against first principle Monte Carlo simulations across the training region $(E_\parallel, Z_{\text{eff}}, \alpha)$. The RunAway Monte Carlo (RAMc) code is employed (see McDevitt *et al.* (2019) for a detailed description), which evolves the guiding centre motion of relativistic electrons and includes effects from small-angle collisions, large-angle collisions and synchrotron radiation. Large-angle collisions are evaluated using the full Møller source (Møller 1932). In order to avoid toroidal corrections to the avalanche growth rate (McDevitt & Tang 2019; Arnaud & McDevitt 2024), all REs are initialized near $r = 0$, and a large tokamak device was chosen with a minor radius of $a = 2$ m and major radius $R_0 = 6$ m, in order to render spatial transport negligible. A geometry with circular flux surfaces was selected, with a safety factor profile taken to be $q(r) = 2.1 + 2(r/a)^2$.

The Monte Carlo avalanche simulations are set up by initializing a small population of electrons (8 for this analysis) at high momentum ($p \in (10, 20)$) and strongly aligned with the magnetic field ($\xi \in (-0.9, -1.0)$). They are then allowed to exponentially grow in time until a saturated growth rate can be identified, generally resulting in the initial seed RE population growing by several orders of magnitude. The plasma parameters chosen were a density of $n_e = 10^{21}$ m^{-3}, a temperature of $T_e = 10$ eV and a toroidal magnetic field strength that was varied to give the appropriate value of α. One caveat is that at a fixed density and temperature $\alpha \propto B^2$, which in turn can lead to orbit drifts at the low boundary of synchrotron radiation due to the weak magnetic field. These orbit drifts are thus ensured to be negligible by enforcing a minimum toroidal magnetic field of $B = 2$ T, which corresponds to a lower bound ($\alpha \approx 2.8 \times 10^{-3}$) in the training region for synchrotron radiation. Fifty randomly selected samples of $(E_\parallel, Z_{\text{eff}}, \alpha)$ are chosen as input parameters for the Monte Carlo solver and the avalanche PINN, where cases below the avalanche threshold are discarded, leaving forty-one samples. A comparison between the predictions of the avalanche PINN and Monte Carlo solver is shown in figure 8(*a*). Here, the grey dashed line represents equality (*y*-axis equal to *x*-axis) between the PINN and Monte Carlo predictions. The coefficient of determination (Guilford & Guilford

https://doi.org/10.1017/S0022377824000679 Published online by Cambridge University Press

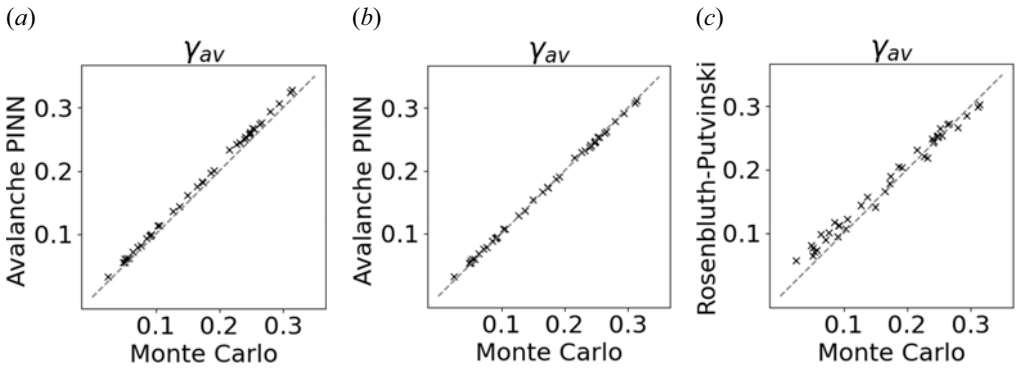


FIGURE 8. (*a*) Avalanche growth rate comparison between the PINN and Monte Carlo solver. The grey dashed line represents the coefficient of determination of $R^2 = 1$. (*b*) Same as panel (*a*), but with the predictions of the avalanche PINN multiplied by the factor 0.94886. (*c*) Avalanche growth rate comparison between (4.8) and the Monte Carlo solver. The avalanche growth rates were evaluated across $E_\parallel \in (1, 10)$, $Z_{\text{eff}} \in (1, 10)$ and $\alpha \in (2.8 \times 10^{-3}, 0.2)$. The other parameters were chosen to be $T_e = 10$ eV and $n_e = 10^{21}$ m^{-3}.

1936) between the PINN predictions and the Monte Carlo avalanche growth rate was evaluated to be $R^2 \approx 0.9849$, indicating that the PINN was able to accurately predict the avalanche growth rate across a broad range of parameters. One noticeable feature present in figure 8(*a*) is the systematic over-prediction of the avalanche growth rate by the PINN. This feature is expected, and is due to the use of the simplified Rosenbluth–Putvinski secondary source term described by (4.4), which assumes all primary electrons to have $p \to \infty$ and $\xi = -1$. We note that this feature can be mitigated by introducing a simple discrepancy model, where for the parameter regime considered in this study, multiplying the predictions of the PINN by a factor of 0.94886 significantly reduced the discrepancy between the PINN and Monte Carlo predictions. The correlation between the avalanche PINN predictions including the discrepancy model and the Monte Carlo predictions are shown in figure 8(*b*), where the coefficient of determination increased to $R^2 = 0.9988$. Improving the RE closure used when approximating the Møller source will be the subject of future work.

The PINN is also shown to robustly provide better avalanche growth rate predictions than the analytical expression provided by (4.8) (compare figure 8*a*,*c*). Here, the avalanche growth rate provided by (4.8) compared with the Monte Carlo solver has a weaker correlation than that between the PINN and Monte Carlo solver, where the coefficient of determination for figure 8(*c*) was $R^2 \approx 0.9644$. The avalanche PINN thus improves on the accuracy of the avalanche growth rate predictions, even in the absence of the 0.94886 factor.

5. Summary and conclusions

This work utilized a PINN to evaluate the steady-state solution of the adjoint to the relativistic Fokker–Planck equation. Noting that the PINN takes the physical parameters of the problem as inputs, once trained, the PINN provides an efficient surrogate model of the RPF. In addition, while a comprehensive description of the avalanche growth rate requires evaluating the primary electron distribution, a quantity not evaluated in the present approach, by invoking the often employed simplification that primary electrons have infinite energy and a pitch $\xi = -1$ (Rosenbluth & Putvinski 1997) the RE avalanche

https://doi.org/10.1017/S0022377824000679 Published online by Cambridge University Press

growth rate was shown to be directly linked to an integral of the RPF (4.7). While this approximation to the secondary source of REs is known to lead to a modest overestimate of the avalanche growth rate at intermediate values of the electric field (McDevitt *et al.* 2018), predictions from the PINN were shown to agree with direct Monte Carlo simulations that utilized a complete Møller source across a broad range of parameters with an offset of roughly 5 % (figure 8*a*).

An additional aim of this paper was to provide a proof-of-principle demonstration that physics constrained deep learning methods offer an attractive avenue through which RE surrogate models can be developed. In contrast to data-driven deep learning approaches, the present method encodes the underlying kinetic solution into the neural network, rather than just the QoI (avalanche growth rate in this case), and thus allows for greater interpretability and hence greater confidence in the accuracy of the prediction. While the present paper has focused on an idealized description of REs, generalization to more complete models of RE formation, incorporating partial screening effects (Hesslow *et al.* 2017) for example, can be accomplished by modifying the collision coefficients in the adjoint equation described by (3.2) and retraining the PINN. This extension, along with a generalization to include temporal and spatial evolution of the RPF, will be the subject of future work. We thus anticipate that the present approach provides a flexible means through which RE surrogate models can be developed across a broad range of plasma conditions.

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 U.S. Department of Energy Office of Fusion Energy Sciences under award nos. DE-SC0024634 and DE-SC0024649.

Declaration of interests

The authors report no conflict of interest.

REFERENCES

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., *et al.* 2016 {TensorFlow}: a system for {Large-Scale} machine learning. In *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, pp. 265–283. USENIX Association.

Aleynikov, P. & Breizman, B.N. 2015 Theory of two threshold fields for relativistic runaway electrons. *Phys. Rev. Lett.* **114** (15), 155001.

Andersson, F., Helander, P. & Eriksson, L.-G. 2001 Damping of relativistic electron beams by synchrotron radiation. *Phys. Plasmas* **8** (12), 5221–5229.

Antonsen, T. Jr. & Chu, K. 1982 Radio frequency current generation by waves in toroidal geometry. *Phys. Fluids* **25** (8), 1295–1296.

Arnaud, J. & McDevitt, C. 2024 The impact of collisionality on the runaway electron avalanche during a tokamak disruption. *Phys. Plasmas* **31** (6), 062509.

Ashkin, A., Page, L.A. & Woodward, W. 1954 Electron-electron and positron-electron scattering measurements. *Phys. Rev.* **94** (2), 357.

https://doi.org/10.1017/S0022377824000679 Published online by Cambridge University Press

Bandaru, V., Hoelzl, M., Reux, C., Ficker, O., Silburn, S., Lehnen, M., Eidietis, N., JOREK Team & JET Contributors 2021 Magnetohydrodynamic simulations of runaway electron beam termination in JET. *Plasma Phys. Control. Fusion* **63** (3), 035024.

Boozer, A.H. 2015 Theory of runaway electrons in ITER: equations, important parameters, and implications for mitigation. *Phys. Plasmas* **22** (3), 032504.

Boozer, A.H. 2018 Pivotal issues on relativistic electrons in ITER. *Nucl. Fusion* **58** (3), 036006.

Connor, J. & Hastie, R. 1975 Relativistic limitations on runaway electrons. *Nucl. Fusion* **15** (3), 415.

Decker, J., Hirvijoki, E., Embreus, O., Peysson, Y., Stahl, A., Pusztai, I. & Fülöp, T. 2016 Numerical characterization of bump formation in the runaway electron tail. *Plasma Phys. Control. Fusion* **58** (2), 025016.

Dreicer, H. 1959 Electron and ion runaway in a fully ionized gas. I. *Phys. Rev*. **115** (2), 238.

Fisch, N. 1986 Transport in driven plasmas. *Phys. Fluids* **29** (1), 172–179.

Guilford, J.P. & Guilford, R.B. 1936 Personality factors S, E, and M, and their measurement. *J. Psychol*. **2** (1), 109–127.

Guo, Z., McDevitt, C.J. & Tang, X.-Z. 2017 Phase-space dynamics of runaway electrons in magnetic fields. *Plasma Phys. Control. Fusion* **59** (4), 044003.

Helander, P., Smith, H., Fülöp, T. & Eriksson, L.-G. 2004 Electron kinetics in a cooling plasma. *Phys. Plasmas* **11** (12), 5704–5709.

Hender, T., Wesley, J., Bialek, J., Bondeson, A., Boozer, A., Buttery, R., Garofalo, A., Goodman, T., Granetz, R., Gribov, Y., *et al.* 2007 MHD stability, operational limits and disruptions. *Nucl. Fusion* **47** (6), S128.

Hesslow, L., Embréus, O., Stahl, A., DuBois, T.C., Papp, G., Newton, S.L. & Fülöp, T. 2017 Effect of partially screened nuclei on fast-electron dynamics. *Phys. Rev. Lett*. **118**, 255001.

Hesslow, L., Embréus, O., Vallhagen, O. & Fülöp, T. 2019*a* Influence of massive material injection on avalanche runaway generation during tokamak disruptions. *Nucl. Fusion* **59** (8), 084004.

Hesslow, L., Unnerfelt, L., Vallhagen, O., Embréus, O., Hoppe, M., Papp, G. & Fülöp, T. 2019*b* Evaluation of the dreicer runaway generation rate in the presence of high-impurities using a neural network. *J. Plasma Phys*. **85** (6), 475850601.

Hoppe, M., Embreus, O. & Fülöp, T. 2021 Dream: a fluid-kinetic framework for tokamak disruption runaway electron simulations. *Comput. Phys. Commun*. **268**, 108098.

Jayakumar, R., Fleischmann, H. & Zweben, S. 1993 Collisional avalanche exponentiation of runaway electrons in electrified plasmas. *Phys. Lett.* A **172** (6), 447–451.

Karney, C.F. & Fisch, N.J. 1986 Current in wave-driven plasmas. *Phys. Fluids* **29** (1), 180–192.

Karniadakis, G.E., Kevrekidis, I.G., Lu, L., Perdikaris, P., Wang, S. & Yang, L. 2021 Physics-informed machine learning. *Nat. Rev. Phys*. **3** (6), 422–440.

Karpatne, A., Atluri, G., Faghmous, J.H., Steinbach, M., Banerjee, A., Ganguly, A., Shekhar, S., Samatova, N. & Kumar, V. 2017 Theory-guided data science: a new paradigm for scientific discovery from data. *IEEE Trans. Knowl. Data Engng* **29** (10), 2318–2331.

Kingma, D.P. & Ba, J. 2014 Adam: a method for stochastic optimization. arXiv:1412.6980.

Kruskal, M. & Bernstein, I. 1962 Princeton Plasma Physics Lab. Rep. No. Matt-Q-20.

Lagaris, I.E., Likas, A. & Fotiadis, D.I. 1998 Artificial neural networks for solving ordinary and partial differential equations. *IEEE Trans. Neural Netw*. **9** (5), 987–1000.

Liu, C., Brennan, D.P., Bhattacharjee, A. & Boozer, A.H. 2016 Adjoint Fokker–Planck equation and runaway electron dynamics. *Phys. Plasmas* **23** (1), 010702.

Liu, C., Brennan, D.P., Boozer, A.H. & Bhattacharjee, A. 2017 Adjoint method and runaway electron avalanche. *Plasma Phys. Control. Fusion* **59** (2), 024003.

Liu, C., Zhao, C., Jardin, S.C., Ferraro, N.M., Paz-Soldan, C., Liu, Y. & Lyons, B.C. 2021 Self-consistent simulation of resistive kink instabilities with runaway electrons. *Plasma Phys. Control. Fusion* **63** (12), 125031.

Lu, L., Meng, X., Mao, Z. & Karniadakis, G.E. 2021 DeepXDE: a deep learning library for solving differential equations. *SIAM Rev*. **63** (1), 208–228.

Lusch, B., Kutz, J.N. & Brunton, S.L. 2018 Deep learning for universal linear embeddings of nonlinear dynamics. *Nat. Commun*. **9** (1), 1–10.

https://doi.org/10.1017/S0022377824000679 Published online by Cambridge University Press

MARTÍN-SOLÍS, J., LOARTE, A. & LEHNEN, M. 2015 Runaway electron dynamics in tokamak plasmas with high impurity content. *Phys. Plasmas* **22** (9), 092512.

MARTÍN-SOLÍS, J., LOARTE, A. & LEHNEN, M. 2017 Formation and termination of runaway beams in ITER disruptions. *Nucl. Fusion* **57** (6), 066025.

MATTHEWS, G., BAZYLEV, B., BARON-WIECHEC, A., COENEN, J., HEINOLA, K., KIPTILY, V., MAIER, H., REUX, C., RICCARDO, V., RIMINI, F., *et al.* 2016 Melt damage to the jet ITER-like wall and divertor. *Phys. Scr*. **2016** (T167), 014070.

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

MCDEVITT, C.J. 2023 A physics-informed deep learning model of the hot tail runaway electron seed. *Phys. Plasmas* **30** (9), 092501.

MCDEVITT, C.J., GUO, Z. & TANG, X.-Z. 2018 Relation of the runaway avalanche threshold to momentum space topology. *Plasma Phys. Control. Fusion* **60** (2), 024004.

MCDEVITT, C.J., GUO, Z. & TANG, X. 2019 Avalanche mechanism for runaway electron amplification in a tokamak plasma. *Plasma Phys. Control. Fusion* **61**, 054008.

MCDEVITT, C.J. & TANG, X.-Z. 2024 A physics-informed deep learning description of Knudsen layer reactivity reduction. *Phys. Plasmas* **31** (6), 062701.

MCDEVITT, C. & TANG, X.-Z. 2019 Runaway electron generation in axisymmetric tokamak geometry. *Europhys. Lett*. **127** (4), 45001.

MØLLER, C. 1932 Zur theorie des durchgangs schneller elektronen durch materie. *Ann. Phys*. **406** (5), 531–585.

NOCEDAL, J. 1980 Updating quasi-newton matrices with limited storage. *Maths Comput*. **35** (151), 773–782.

PASZKE, A., GROSS, S., CHINTALA, S., CHANAN, G., YANG, E., DEVITO, Z., LIN, Z., DESMAISON, A., ANTIGA, L. & LERER, A. 2017 Automatic differentiation in PyTorch. In *NIPS 2017 Workshop*.

ROSENBLUTH, M. & PUTVINSKI, S. 1997 Theory for avalanche of runaway electrons in tokamaks. *Nucl. Fusion* **37** (10), 1355.

SAINTERME, A. & SOVINEC, C. 2024 Resistive hose modes in tokamak runaway electron beams. *Phys. Plasmas* **31** (1), 010701.

SMITH, H.M. & VERWICHTE, E. 2008 Hot tail runaway electron generation in tokamak disruptions. *Phys. Plasmas* **15** (7), 072502.

SOKOLOV, I. 1979 Multiplication' of accelerated electrons in a tokamak. *JETP Lett*. **29**, 218–221.

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. Meth. Appl. Mech. Engng* **361**, 112732.

TAGUCHI, M. 1983 The effect of trapped electrons on the wave-induced current. *J. Phys. Soc. Japan* **52** (6), 2035–2040.

VALLHAGEN, O., EMBREUS, O., PUSZTAI, I., HESSLOW, L. & FÜLÖP, T. 2020 Runaway dynamics in the DT phase of iter operations in the presence of massive material injection. *J. Plasma Phys*. **86** (4), 475860401.

WANG, R., KASHINATH, K., MUSTAFA, M., ALBERT, A. & YU, R. 2020 Towards physics-informed deep learning for turbulent flow prediction. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 1457–1466.

WANG, S., YU, X. & PERDIKARIS, P. 2022 When and why PINNs fail to train: a neural tangent kernel perspective. *J. Comput. Phys*. **449**, 110768.

YANG, M., WANG, P., DEL CASTILLO-NEGRETE, D., CAO, Y. & ZHANG, G. 2024 A Pseudoreversible Normalizing Flow for Stochastic Dynamical Systems with Various Initial Distributions. *SIAM J. Sci. Compu.* **46** (4), C508–C533.

ZHANG, G. & DEL CASTILLO-NEGRETE, D. 2017 A backward Monte–Carlo method for time-dependent runaway electron simulations. *Phys. Plasmas* **24** (9), 092511.

https://doi.org/10.1017/S0022377824000679 Published online by Cambridge University Press