Physics-informed neural networks for solving the Boltzmann equation of the electron velocity distribution function in weakly ionized plasmas

Physics-informed neural networks for solving the Boltzmann equation of the electron velocity distribution function in weakly ionized plasmas

Satoru Kawaguchi[1, 2]* and Tomoyuki Murakami[1]

[1]*Department of Systems Design Engineering, Faculty of Science and Technology, Seikei University, Musashino, Tokyo, 180-8633, Japan.*

[2]*Division of Information and Electronic Engineering, Graduate School of Engineering, Muroran Institute of Technology, Muroran, Hokkaido, 050-8585, Japan*

E-mail: skawaguchi@mmm.muroran-it.ac.jp

The equilibrium electron velocity distribution function (EVDF) and electron transport coefficient in weakly ionized plasmas under crossed DC uniform electric and magnetic fields are calculated via the Boltzmann equation (BE) using physics-informed neural networks (PINNs). The latent solution of the BE is represented by an artificial neural network, and then the neural network is trained to respect the BE. By leveraging automatic differentiation, no mesh generation in velocity space is required, allowing us to calculate the three-dimensional EVDF properly with 0.01% of memory capacity required for the conventional mesh-based method. The EVDF and electron transport coefficients in SF_6 calculated from the PINNs are benchmarked by comparing with those calculated from the Monte Carlo simulation (MCS). In most cases, the relative difference between the electron transport coefficient calculated from the PINNs and MCS is found to be within 1%.

1. Introduction

The Boltzmann equation (BE) provides the foundation for calculating the electron velocity distribution function (EVDF) in weakly ionized plasmas.[1)] The electron transport coefficients and rate coefficients can be calculated from the EVDF.[1, 2)] These coefficients are required for simulating electron transport and chemistry in plasmas.[3-6)] Furthermore, electron collision cross section sets have been determined and validated by comparing calculated and measured electron transport coefficients.[7-9)]

The direct numerical solution (DNS) of the BE, which requires no expansion of the EVDF using orthogonal functions, is preferable to calculate the EVDF accurately. Drallos and Wadehra [10, 11)] applied the DNS for calculating time evolution of the EVDF under DC uniform electric fields in Ne and Ar. Maeda and Makabe [12)] developed the DNS for investigating the electron swarm transport under RF electric fields in Ar. Their method was used to calculate the electron transport coefficients in CF_4 and CF_4-Ar mixtures. [13)] Sugawara recently developed the DNS for calculating the EVDF under DC uniform electric fields crossed with DC magnetic fields at a right angle.[14, 15)] Owing to relying on discretizing velocity space, the conventional DNS requires a huge memory capacity for storing the EVDF.

Physics-informed neural networks (PINNs) provide a meshfree approach to solving the partial differential equation (PDE) [16, 17)] and have attracted tremendous attention in computational physics.[17-20)] In the PINNs approach, the latent solution of the PDE is represented by an artificial neural network (ANN), and then the ANN is trained to respect both the PDE and boundary conditions. The partial derivatives of the ANN output with respect to each ANN input can be calculated by taking advantage of automatic differentiation [21)], which makes the PINNs approach meshfree. The meshfree approach would help us to handle complex geometry and to tackle high-dimensional problems with reducing memory consumption.

The PINNs approach was recently applied to calculate the equilibrium two-dimensional EVDF under DC uniform electric fields in Reid's ramp model gas and Ar via the BE. [22)] The electron energy distribution function (EEDF) and electron transport coefficients calculated from the EVDF were found to agree well with those calculated from Monte Carlo simulation (MCS) in most cases. This indicates the applicability of the PINNs approach for calculating the EVDF in plasmas via the BE. However, the tail of the EEDF calculated from the PINNs approach slightly differs from that calculated from the MCS, and this causes a relative difference of up to 13% for the ionization collision frequency.

The aim of this paper is to describe an improved PINNs approach for calculating the EVDF via the BE and to extend it into calculating the equilibrium three-dimensional (3D) EVDF under DC uniform electric fields crossed with DC uniform magnetic fields at arbitrary angles. The present expansion would contribute to studies on electron transport in magnetized plasmas, which are employed in material processing. [23, 24)]

The rest of this paper is organized as follows: the governing equation of the equilibrium EVDF under crossed DC uniform electric and magnetic fields is deduced from the BE in Sect. 2. The procedure for solving the governing equation using the PINNs approach and its benchmark calculation are presented in Sect. 3 and Sect. 4, respectively. Electron transport coefficients and EEDF calculated from the EVDF are benchmarked by comparing those calculated from MCS. Discussion on memory capacity required for representing the EVDF and future extension of the present method to calculate time dependent EVDF is given in Sect. 5. This paper is summarised in Sect. 6.

2. Theory

2.1 Governing equation for equilibrium EVDF

The governing equation for the equilibrium EVDF is obtained from the BE for the electron distribution function $f(\boldsymbol{r}, \boldsymbol{v}, t)$

$$\left[\boldsymbol{\alpha} \cdot \frac{\partial}{\partial \boldsymbol{v}} + \boldsymbol{v} \cdot \frac{\partial}{\partial \boldsymbol{r}} + \frac{\partial}{\partial t} - J_{\mathrm{C}}\right] f(\boldsymbol{r}, \boldsymbol{v}, t) = 0, \tag{1}$$

where $\boldsymbol{\alpha} = (\alpha_x, \alpha_y, \alpha_z)$ is the electron acceleration, $\boldsymbol{v} = (v_x, v_y, v_z)$ is the velocity, $\boldsymbol{r} = (x, y, z)$ is the position, and J_{C} is a collision operator.

The DC uniform electric field $\boldsymbol{E}$ and DC uniform magnetic field $\boldsymbol{B}$ are defined in boundary-free space as

$$\boldsymbol{E} = (0,0,-E), \tag{2}$$

$$\boldsymbol{B} = (0, -B\sin\psi, -B\cos\psi), \tag{3}$$

where E is the electric field strength, B is the magnetic flux density, and ψ is the angle between $\boldsymbol{E}$ and $\boldsymbol{B}$.

The electron acceleration is given by

$$\boldsymbol{\alpha} = \frac{d\boldsymbol{v}}{dt} = -\frac{e}{m}(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}), \tag{4}$$

$$\alpha_x = \omega\left(v_y \sin\psi - v_z \cos\psi\right), \tag{5}$$

$$\alpha_y = \omega \cos\psi, \tag{6}$$

$$\alpha_z = \omega\left(\frac{E}{B} + v_x \sin\psi\right), \tag{7}$$

where e is the elementary charge, m is the electron mass, and $\omega = eB/m$ is the cyclotron angular frequency.

By integrating Eq. (1) over $\boldsymbol{r}$, the BE is transformed to

$$\left[\boldsymbol{\alpha} \cdot \frac{\partial}{\partial \boldsymbol{v}} + \frac{\partial}{\partial t} - J_{\mathrm{C}}\right] f(\boldsymbol{v}, t) = 0. \tag{8}$$

In the hydrodynamic equilibrium, the EVDF can be written by

$$f(\boldsymbol{v}, t) = f(\boldsymbol{v}) n(t), \tag{9}$$

where $n(t)$ is the electron number density. The $f(\boldsymbol{v})$ is normalized as

$$\int_{v} f(\boldsymbol{v}) d\boldsymbol{v} = 1. \tag{10}$$

We assume that the temporal evolution of $n(t)$ in the equilibrium can be written as

$$n(t) \propto \exp(\bar{R}_i t), \tag{11}$$

which is related to the pulsed Townsend experiment.[25] Here, $\bar{R}_i$ denotes the effective ionization collision frequency. By substituting Eqs. (9) and (11) into Eq. (8), we obtain the governing equation of the equilibrium EVDF

$$\left[\boldsymbol{\alpha} \cdot \frac{\partial}{\partial \boldsymbol{v}} + \bar{R}_i - J_{\mathrm{C}}\right] f(\boldsymbol{v}) = 0. \tag{12}$$

The $J_{\mathrm{C}} f(\boldsymbol{v})$ consists of the inflow of the electron at $\boldsymbol{v}$ due to elastic collision $J_{\mathrm{el}}^{\mathrm{in}} f(\boldsymbol{v})$, excitation collision $J_{\mathrm{ex}}^{\mathrm{in}} f(\boldsymbol{v})$, and ionization collision $J_{\mathrm{ion}}^{\mathrm{in}} f(\boldsymbol{v})$, and the outflow of the electron at $\boldsymbol{v}$ due to collisions, $J^{\mathrm{out}} f(\boldsymbol{v})$:

$$J_{\mathrm{C}} f(\boldsymbol{v}) = J_{\mathrm{el}}^{\mathrm{in}} f(\boldsymbol{v}) + J_{\mathrm{ex}}^{\mathrm{in}} f(\boldsymbol{v}) + J_{\mathrm{ion}}^{\mathrm{in}} f(\boldsymbol{v}) - J^{\mathrm{out}} f(\boldsymbol{v}). \tag{13}$$

The electron scattering is assumed to be isotropic, and these terms are described as [26, 27]

$$J_{\mathrm{el}}^{in} f(\boldsymbol{v}) = \frac{1}{4\pi} N q_{\mathrm{el}}(v_{\mathrm{el}}) v \left(1 + \frac{2m}{M}\right)^2 f(v_{el}), \tag{14}$$

$$J_{\mathrm{ex}}^{in} f(\boldsymbol{v}) = \frac{1}{4\pi} N q_{\mathrm{ex}}(v_{\mathrm{ex}}) \frac{v_{\mathrm{ex}}^2}{v} f(v_{ex}), \tag{15}$$

$$J_{ion}^{in} f(\boldsymbol{v}) = \frac{1}{4\pi\Delta} N q_{\mathrm{i}}(v_{\mathrm{i}1}) \frac{v_{\mathrm{i}1}^2}{v} f(v_{i1}) + \frac{1}{4\pi(1-\Delta)} N q_{\mathrm{i}}(v_{\mathrm{i}2}) \frac{v_{\mathrm{i}2}^2}{v} f(v_{i2}), \tag{16}$$

$$J^{out} f(\boldsymbol{v}) = N q_{\mathrm{T}}(v) v f(\boldsymbol{v}), \tag{17}$$

with the electron speed distribution

$$f(v) = \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} f(v \sin\theta \cos\phi, v \sin\theta \sin\phi, v \cos\theta) \sin\theta \, \mathrm{d}\theta \mathrm{d}\phi, \tag{18}$$

where θ and ϕ are the polar and azimuthal angles of $\boldsymbol{v}$ and v is the electron speed. Here, N is the number density of gas molecules, M is the molecular mass, Δ:1-Δ is the electron energy partition ratio after the ionization collision, $q_{\mathrm{el}}(v)$ is the elastic cross section, $q_{\mathrm{ex}}(v)$ is the

excitation cross section, $q_{\mathrm{i}}(v)$ is the ionization cross section, and $q_{\mathrm{T}}(v)$ is the total cross section defined as

$$q_{\mathrm{T}}(v) = q_{\mathrm{el}}(v) + q_{\mathrm{ex}}(v) + q_{\mathrm{a}}(v) + q_{\mathrm{i}}(v), \tag{19}$$

where $q_{\mathrm{a}}(v)$ is the electron attachment cross section. The value of Δ is set to 0.5. The elastic momentum transfer cross section is used instead of $q_{\mathrm{el}}(v)$. The electron speeds v_{el}, v_{ex}, v_{i1}, and v_{i2} are described as

$$v_{\mathrm{el}} = v\sqrt{1 + \frac{2m}{M}}, \tag{20}$$

$$v_{\mathrm{ex}} = \sqrt{v^2 + \frac{2\varepsilon_{\mathrm{ex}}}{m}}, \tag{21}$$

$$v_{\mathrm{i1}} = \sqrt{\frac{v^2}{\Delta} + \frac{2\varepsilon_{\mathrm{i}}}{m}}, \tag{22}$$

$$v_{\mathrm{i2}} = \sqrt{\frac{v^2}{1-\Delta} + \frac{2\varepsilon_{\mathrm{i}}}{m}}, \tag{23}$$

where $\varepsilon_{\mathrm{ex}}$ and ε_{i} are the threshold energy of $q_{\mathrm{ex}}(v)$ and $q_{\mathrm{i}}(v)$, respectively.

The mean electron energy $\langle\varepsilon\rangle$, flux electron drift velocity $\boldsymbol{V} = (V_x, V_y, V_z)$, ionization collision frequency R_{i}, and electron attachment collision frequency R_{a} are defined as

$$\langle\varepsilon\rangle = \int_{\boldsymbol{v}} \frac{1}{2} m\boldsymbol{v}^2 f(\boldsymbol{v}) \mathrm{d}\boldsymbol{v}, \tag{24}$$

$$V_k = \int_{\boldsymbol{v}} v_k f(\boldsymbol{v}) d\boldsymbol{v} \quad (k = x, y, z) \tag{25}$$

$$R_i = \int_{\boldsymbol{v}} N q_i(v) v f(\boldsymbol{v}) d\boldsymbol{v}, \tag{26}$$

$$R_a = \int_{\boldsymbol{v}} N q_a(v) v f(\boldsymbol{v}) d\boldsymbol{v}. \tag{27}$$

The effective ionization collision frequency is defined as $\overline{R_i} = R_i - R_a$.

2.2 Nondimensionalization of the governing equation

Since introducing dimensionless quantities is useful in numerical calculations [4, 28] and it is preferable for the ANN inputs to be distributed on [-1, 1] [29], the physical quantity appearing in the governing equation is converted to the dimensionless quantity. The dimension of the physical quantity considered here can be expressed as a product of time T, length L, mass M, and electric charge Q. Thus, the physical quantity x the dimension of which is $\mathrm{T}^a\mathrm{L}^b\mathrm{M}^c\mathrm{Q}^d$

is converted to the dimensionless quantity x^* as follows,

$$x^* = \frac{x}{\left(\nu_{\mathrm{T,m}}^{-1}\right)^a \left(v_{\mathrm{m}}\nu_{\mathrm{T,m}}^{-1}\right)^b m^c e^d}, \tag{28}$$

where $\nu_{\mathrm{T,m}}$ denotes the maximum of the total electron collision frequency and v_{m} is the maximum of the electron speed considered in the calculation.

The dimensionless version of the governing equation is described as

$$\hat{L} f^*(\boldsymbol{v}^*) = 0, \tag{29}$$

where $\hat{L}$ is the operator given by

$$\hat{L} = \left(v_y^* B^* \cos\psi - v_z^* B^* \sin\psi\right)\frac{\partial}{\partial v_x^*} - B^* \cos\psi \frac{\partial}{\partial v_y^*} + \left(v_x^* B^* \sin\psi + E^*\right)\frac{\partial}{\partial v_z^*} + \bar{R}_i^{\,*} - J_{\mathrm{C}}^*. \tag{30}$$

3. PINNs approach for solving the governing equation

The schematic diagram of the PINNs approach for solving the governing equation is illustrated in Fig. 1. The ANN architecture used in this work is described in Sect. 3.1. The procedure for training the ANN is given in Sect. 3.2.

3.1 Artificial neural network

The dimensionless EVDF $f^*(\boldsymbol{v}^*)$ is represented by the multi-layer feedforward neural network improved by Wang *et al.* [30] The neural network architecture is illustrated in Fig. 2. The coordinate of sampling points $(v_{x,i}^*, v_{y,i}^*, v_{z,i}^*)$ (i = 1, 2, ..., N_d) in velocity space is stored in the (N_d×3) matrix X. The matrices U, V, and $H^{(k)}$ are calculated as follows:

$$U = \sigma(XW_1 + b_1), \tag{31}$$

$$V = \sigma(XW_2 + b_2), \tag{32}$$

$$H^{(1)} = Z^{(1)} = \sigma\left(XW_{z,1} + b_{z,1}\right), \tag{33}$$

$$Z^{(k)} = \sigma\left(H^{(k)}W_{z,k} + b_{z,k}\right) \quad (k = 2,3,\ldots,N_L), \tag{34}$$

$$H^{(k+1)} = \left(1 - Z^{(k)}\right) \odot U + Z^{(k)} \odot V \ (k = 1,2,\ldots,N_L - 1), \tag{35}$$

$$f^*(X) = \exp\left[-H^{(L)}W\right], \tag{36}$$

where W_k, $W_{z,k}$, and W are weight matrices, b_k and $b_{z,k}$ are bias matrices, $\sigma(A)$ is the nonlinear activation function applied to each element of matrix A, $\odot$ denotes Hadamard product (element-wise multiplication), and N_L is the number of the hidden layers. The size of the matrix shown above is summarized in Table I. The size of the weight and bias matrices increases with the number of units in the layer N_{u}. The ability of the ANN for representing a function, therefore, increases with N_{L} and N_{u}. The Mish function [31],

$$\mathrm{Mish}(x) = x \tanh(\log(1 + e^x)), \tag{37}$$

is used as an activation function. Here, the addition of a matrix A and a vector b yields a matrix C as follows: [32)]

$$C_{i,j} = A_{i,j} + b_j. \tag{38}$$

The parameter to be optimized in the ANN is summarized as

$$\boldsymbol{\theta} = \{W_1, b_1, W_2, b_2, W_{z,1}, \dots W_{z,N_L}, b_{z,1}, \dots, b_{z,N_L}, W\}. \tag{39}$$

The bias matrix is initialized to zero, and the weight matrix is initialized by using random numbers described by He *et al.* [33)] unless otherwise mentioned.

3.2 Training of the artificial neural network

How well the ANN respects the governing equation is measured by loss function defined as

$$\mathcal{L} = \frac{1}{N_d}\sum_{i=1}^{N_d}\left|\hat{L}f^*\left(v_{x,i}^*, v_{y,i}^*, v_{z,i}^*\right)\right|. \tag{40}$$

When we focus on a small volume centered on $(v_{x,i}^*, v_{y,i}^*, v_{z,i}^*)$ in velocity space, electrons flow in and out due to acceleration and collision with molecules. In the hydrodynamic equilibrium regime, the inflow and the outflow should be balanced, namely, $\hat{L}f^*\left(v_{x,i}^*, v_{y,i}^*, v_{z,i}^*\right) = 0$. The loss function is the mean absolute error of $\hat{L}f^*\left(v_{x,i}^*, v_{y,i}^*, v_{z,i}^*\right)$. The loss function would be defined as the mean square error (MSE) of $\hat{L}f^*\left(v_{x,i}^*, v_{y,i}^*, v_{z,i}^*\right)$; however, when we use the MSE, the tail of the EEDF calculated from the ANN tends to slightly differ from the EEDF calculated from MCS. The points $(v_{x,i}^*, v_{y,i}^*, v_{z,i}^*)$ are sampled as follows: the energy $\varepsilon_i \in [0, \varepsilon_\mathrm{m}]$, the cosine of the polar angle $\cos\theta_i \in [-1,1]$, and the azimuthal angle $\phi_i \in [0,2\pi]$ are sampled by Latin hypercube sampling.[34)] Here, $\varepsilon_\mathrm{m} = (1/2)mv_\mathrm{m}^2$. Then, $v_{x,i}^*$, $v_{y,i}^*$, and $v_{z,i}^*$ are obtained as

$$v_{x.i}^* = \sqrt{\frac{2\varepsilon_i}{m}}\sin\theta_i\cos\phi_i \Big/ v_m\,, \tag{41}$$

$$v_{y.i}^* = \sqrt{\frac{2\varepsilon_i}{m}}\sin\theta_i\sin\phi_i \Big/ v_m\,, \tag{42}$$

$$v_{z.i}^* = \sqrt{\frac{2\varepsilon_i}{m}}\cos\theta_i \Big/ v_m\,. \tag{43}$$

The distribution of the sampled ε_i, $\cos\theta_i$, and ϕ_i is uniform. Therefore, the distribution of the electron speed $v_i = \sqrt{(2\varepsilon_i)/m}$ is not uniform. The points are updated at each iteration, avoiding the overfitting of the ANN.

The parameter $\boldsymbol{\theta}$ is optimized to minimize the value of $\mathcal{L}$ by using AMSGrad [35)] with

step size $\alpha = 10^{-2}$, exponential decay rates $\beta_1 = 0.9$ and $\beta_2 = 0.999$, and small parameter $\epsilon = 10^{-7}$.

In order to obtain the unique solution of the governing equation, the normalization constraint of f^* is necessary. Instead of adding the term describing such constraint into $\mathcal{L}$, the collision term $J_C^* f^*(v_{x,i}^*, v_{y,i}^*, v_{z,i}^*)$ in $\hat{L} f^*\left(v_{x,i}^*, v_{y,i}^*, v_{z,i}^*\right)$ is calculated by using the normalized electron speed distribution function $f^*(v^*)$ defined as

$$f^*(v^*) = \frac{\int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi} f^*(v^*\sin\theta\cos\phi\,, v^*\sin\theta\sin\phi\,, v^*\cos\theta)\sin\theta\, d\theta d\phi}{C}, \tag{44}$$

with the normalization constant

$$C = \int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi}\int_{v^*=0}^{1} f^*(v^*\sin\theta\cos\phi\,, v^*\sin\theta\sin\phi\,, v^*\cos\theta)\, v^{*2}\sin\theta\, dv^* d\theta d\phi\,. \tag{45}$$

Here, $f^*(v^*)$ is calculated at fixed points v_j^* at each iteration, and the speed distribution function at k iteration $f^{*(k)}\left(v_j^*\right)$ is updated as

$$f^{*(k)}\left(v_j^*\right) = \frac{1}{2} f^{*(k-1)}\left(v_j^*\right) + \frac{1}{2}\frac{f^*\left(v_j^*\right)}{C}, \tag{46}$$

where $f^{*(0)}\left(v_j^*\right) = f^*\left(v_j^*\right)/C$. Then, $f^*(v_i^*)$ is calculated by interpolating $f^{*(k)}\left(v_j^*\right)$ using cubic spline interpolation. Here, $v_i^* = \sqrt{{v_{x,i}^*}^2 + {v_{y,i}^*}^2 + {v_{z,i}^*}^2}$

The dimensionless effective ionization collision frequency $\bar{R}_i^*$ is also calculated every iteration, and $\bar{R}_i^*$ in k iteration $\bar{R}_i^{*(k)}$ is updated as

$$\bar{R}_i^{*(k)} = \frac{1}{2} R_i^{*(k-1)} + \frac{1}{2}\bar{R}_i^*, \tag{47}$$

where $R_i^{*(0)} = \bar{R}_i^*$.

The procedure of the PINNs approach for calculating the EVDF is summarized as follows.

(1) Initialize the parameter of the ANN $\boldsymbol{\theta}$ at the first iteration (k = 1)

(2) Sample points ($v_{x,i}^*, v_{y,i}^*, v_{z,i}^*$) ($i$ = 1, 2, ..., N_d) in velocity space.

(3) Calculate $f^*\left(v_j^*\right)/C$ and $\bar{R}_i^*$, and update $f^{*(k)}\left(v_j^*\right)$ and $\bar{R}_i^{*(k)}$.

(4) Calculate $f^*(v_i^*)$ by interpolating $f^{*(k)}\left(v_j^*\right)$.

(5) Calculate $\mathcal{L}$ and update $\boldsymbol{\theta}$ to minimize $\mathcal{L}$ using AMSGrad.

(6) $k \rightarrow k + 1$

(7) Repeat steps (2) to (6) until the value of $\mathcal{L}$ becomes a minimum.

4. Benchmark calculations

4.1 Configurations

SF_6 is chosen as an ambient gas in the benchmark calculation of the PINNs approach. The set of electron collision cross sections of SF_6 is taken from Ref. 36. Elastic collision and inelastic collisions for excitation, electron attachment, and ionization are taken into account. Here, N is assumed to be 3.535×10^{22} m^{-3}. The value of the reduced electric field E/N and that of the reduced magnetic field B/N range from 250 Td to 2000 Td (1 Td = 10^{-21} Vm^2) and from 0 Hx to 2000 Hx (1 Hx = 10^{-27} Tm^3), respectively. The value of ψ ranges from 0° to 90°.

The ability of the ANN for representing a function depends on N_L and N_u and affects the accuracy of the EVDF represented by the trained ANN [22]. In the present benchmark, the ANN is constructed with $N_L = 4$ and $N_u = 100$.

4.2 Preparation of the reference data for the benchmark calculation

The EVDF, EEDF, and electron transport coefficients in SF_6 under electric and magnetic fields are calculated by MCS to prepare reference data for the benchmark. Details of the MCS are given in Refs 7 and 37, where electron transport under DC uniform electric fields was considered. Due to the presence of magnetic fields, the calculation of the electron position and velocity is modified according to Eq. (4). The flight time of electrons between the collisions τ is calculated by using the null collision method.[38] Given an initial position $\boldsymbol{r}_0 = (x_0, y_0, z_0)$ and a velocity $\boldsymbol{v_0} = (v_{x0}, v_{y0}, v_{z0})$, the electron position $\boldsymbol{r}_1 = (x_1, y_1, z_1)$ and velocity $\boldsymbol{v}_1 = (v_{x1}, v_{y1}, v_{z1})$ after τ are calculated by

$$x_1 = x_0 + \frac{C_0}{\omega}(1-\cos(\omega\tau)) + \frac{C_1}{\omega}\sin(\omega\tau) - \frac{E}{B}\tau\sin\psi, \tag{48}$$

$$y_1 = y_0 + \frac{C_0}{\omega}\sin(\omega\tau)\cos\psi - \frac{C_1}{\omega}(1-\cos(\omega\tau))\cos\psi + \frac{\omega}{2}\frac{E}{B}\tau^2\sin\psi + C_2\tau\sin\psi, \tag{49}$$

$$z_1 = z_0 - \frac{C_0}{\omega}\sin(\omega\tau)\sin\psi + \frac{C_1}{\omega}(1-\cos(\omega\tau))\sin\psi + \frac{\omega}{2}\frac{E}{B}\tau^2\cos^2\psi + C_2\tau\cos\psi, \tag{50}$$

$$v_{x1} = C_0\sin(\omega\tau) + C_1\cos(\omega\tau) - \frac{E}{B}\sin\psi, \tag{51}$$

$$v_{y1} = C_0\cos(\omega\tau)\cos\psi - C_1\sin(\omega\tau)\cos\psi + \omega\frac{E}{B}\tau\sin\psi\cos\psi + C_2\sin\psi, \tag{52}$$

$$v_{z1} = -C_0\cos(\omega\tau)\sin\psi + C_1\sin(\omega\tau)\sin\psi + \omega\frac{E}{B}\tau\cos^2\psi + C_2\cos\psi, \tag{53}$$

where C_0, C_1, and C_2 are given by

$$C_0 = v_{y0}\cos\psi - v_{z0}\sin\psi, \tag{54}$$

$$C_1 = v_{x0} + \frac{E}{B}\sin\psi\,, \tag{55}$$

$$C_2 = v_{y0}\sin\psi + v_{z0}\cos\psi\,. \tag{56}$$

4.3 Results

The EVDF at *E/N* = 2000 Td, *B/N* = 2000 Hx, and $\psi = 90$ is firstly calculated. Figure 3 shows the value of $\mathcal{L}$ at different N_d as a function of the number of iterations. The loss value tends to decrease with increasing the number of iterations and then converges. The converged value tends to decrease with increasing N_d, indicating that the accuracy of the EVDF represented by the ANN increases with N_d.

Figure 4 shows the value of the electron attachment rate coefficient R_a/N calculated from the trained ANN as a function of the number of iterations. The value of R_a/N calculated from the MCS is also shown in Fig. 4. The N_d affects the convergence of the R_a/N, and we need to set the N_d greater than 512 to obtain R_a/N properly. The effect of N_d on other electron transport coefficients is found to be small. Hereafter, we set N_d = 2048 with a margin.

The EVDF at *B/N* = 2000 Hx and $\psi = 90°$ is calculated by varying the value of *E/N* to 2000, 1750, 1500, 1250, 1000, 750, 500, and 250 Td in order. Figure 5 shows the value of $\mathcal{L}$ as functions of the number of iterations. In the calculation below 2000 Td, retraining of the ANN trained in the previous condition is carried out; that is, the weight and bias matrices are initialized by using those matrices of the trained ANN. The value of $\mathcal{L}$ below *E/N* = 2000 Td immediately converges between 2.0×10^{-3} and 5.0×10^{-3} while 4,000 iterations. This indicates that retraining of the ANN helps us to speed up and improve the training of the ANN for the EVDF in another condition.

The computation time is about 40 seconds per 1,000 iterations by using Nvidia Quadro RTX 8000, the theoretical computing performance of which is 16.31 TFLOPS in single precision.

Figure 6 shows the electron velocity distribution $f(v_x, v_z)$ calculated from the PINNs and MCS. The $f(v_x, v_z)$ is calculated from the EVDF as

$$f(v_x, v_z) = \int_{v_y=-\infty}^{\infty} f\big(v_x, v_y, v_z\big)dv_y \tag{57}$$

and is normalized to satisfy

$$\int_{v_z=-\infty}^{\infty}\int_{v_x=-\infty}^{\infty} f(v_x, v_z)dv_x dv_z = 1\,. \tag{58}$$

The present meshfree method allows us to obtain a smooth distribution. Anisotropy of the

distribution is observed and the peak of the distribution deviates from the origin to (v_x, v_z) = (-0.714$\times10^6$ m/s, 0.306$\times10^6$ m/s). The $f(v_x, v_z)$ calculated from the PINNs agrees well with that calculated from the MCS.

Figure 7 shows the EEDF $f(\varepsilon)$ calculated from the PINNs and MCS at various reduced electric fields at *B/N* = 2000 Hx and $\psi = 90°$. The EEDF is calculated from

$$4\pi v^2 f(v)\mathrm{d}v = f(\varepsilon)\mathrm{d}\varepsilon. \tag{59}$$

The EEDF spreads over higher energies with increasing *E/N* values due to electron acceleration by an electric field, and accordingly, the peak position of the EEDF shifts to higher energy. The PINNs excellently reproduce the EEDF calculated from the MCS. It is worth noting that the good agreement is also observed in the tail of the EEDF, which was difficult to reproduce using the PINNs approach in the previous work. [22)]

The largest difference in the EEDF was found between *E/N* = 250 Td and 2000 Td. We found that the EVDF at *E/N* = 250 Td can be obtained while 4000 iterations by retraining the ANN trained at *E/N* = 2000 Td. It was also found that the EVDF at *E/N* = 2000 Td can be obtained during the same iterations by retraining the ANN trained at *E/N* = 250 Td.

Figure 8 shows the mean electron energy, electron drift velocity, ionization rate coefficient R_i/N, and electron attachment rate coefficient R_a/N at *B/N* = 2000 Hx and $\psi = 90°$ calculated from the PINNs and MCS as functions of *E/N*. The electron drift velocity in the z-direction V_z and that in the opposite to the *x*-direction $-V_x$ increase with *E/N* values. As shown in Fig.7, with increasing *E/N* values the EEDF spreads to high energy region, where ionization cross section becomes competitive with cross sections for other collisions. Accordingly, the value of mean electron energy and that of R_i/N increase with *E/N* values. Due to the large electron attachment cross section of SF_6 at low electron energy, the value of R_a/N increases with decreasing *E/N* values. The loss of electrons due to the electron attachment is therefore superior to the growth of electrons due to ionization below 500 Td.

Figure 9 shows the relative difference in the electron transport and rate coefficients between the PINNs and MCS defined as

$$\Delta = \frac{\zeta_{\mathrm{PINNs}} - \zeta_{\mathrm{MCS}}}{\zeta_{\mathrm{MCS}}}, \tag{60}$$

where ζ_{PINNs} and ζ_{MCS} represent a quantity calculated from the PINNs and MCS, respectively. The relative difference is found to be within 1%.

Retraining of the ANN is found to be also effective to calculate the EVDF by varying ψ and *B/N* at fixed *E/N* values. Figure 10 shows the relative difference between the electron transport coefficients calculated from the PINNs and MCS at *E/N* = 2000 Td as functions of

ψ and *B/N*. Here, the training of the ANN was stopped at 4000 iterations. Agreement on the electron transport coefficients between the two methods is within 1% in most cases. The large discrepancy is found in V_x at $\psi = 0°$, V_y at $\psi = 0°$ and 90°, and V_x at *B/N* = 0 Hx. In such conditions, these values are zero in theory, and those calculated from the PINNs and MCS are found to fluctuate around zero.

The collision processes considered in this work are essential for calculating the EVDF and EEDF in various gases. The benchmark results indicate the applicability of the present method for calculating the EVDF in middle and high *E/N* regions. In a low *E/N* region, rotational excitation and superelastic collisions not considered here are sometimes important for determining the EVDF in molecule gases, and consideration of these collisions would be required.

5. Discussion

5.1 Memory capacity required for representing the EVDF

In the ANN used in this work, the layers U, L, and $H^{(1)}$ have $d \times N_u + N_u = 3 \times 100 + 100$ parameters, and the layer $H^{(k+1)}$ (k = 1, 2, ..., N_L - 1) has $N_u \times (N_u + 1) = 100 \times (100 + 1)$ parameters. The 3D EVDF is therefore approximated by ANN having 41,700 parameters. Sugawara [15] calculated the equilibrium 3D EVDF in SF_6 at *E/N* = 100 – 2000 Td, *B/N* = 100 – 2000 Hx, and $\psi = 90°$ by DNS. The EVDF was stored in 3D arrays the size of which is 10000×45×720. The PINNs approach allows us to represent 3D EVDF properly with approximately 0.01% of the memory capacity used in the mesh-based DNS.

5.2 Extension of the PINNs approach for calculating time dependent EVDF

The time dependent calculation of the EVDF is needed for considering AC electric fields and for examining the relaxation process of the EVDF. Since the PINNs approach provides a general framework for solving the PDE, the extension of the present method for calculating time dependent EVDF would be possible.

The number of units in the input layer of the ANN is modified according to the number of variables for the EVDF. When we consider boundary conditions, the loss function is modified as

$$\mathcal{L} = \mathcal{L}_{\mathrm{PDE}} + \lambda \mathcal{L}_{\mathrm{B}}, \tag{61}$$

where $\mathcal{L}_{\mathrm{PDE}}$ and $\mathcal{L}_{\mathrm{B}}$ denote the part of the loss function regarding the governing equation and the boundary condition, respectively, and λ is the parameter controlling $\mathcal{L}_{\mathrm{B}}$.

When we calculate the temporal evolution of the EVDF $f(v_x, v_y, v_z, t)$ under DC electric

and magnetic fields, we put 4 units in the input layer of the ANN. Equation (8) is the governing equation for the EVDF, and the operator $\hat{L}$ is modified. When the Dirichlet boundary condition is applied, the loss function is written as:

$$\mathcal{L} = \frac{1}{N_{d1}} \sum_{i=1}^{N_{d1}} \left| \hat{L} f^* \left(v_{x,i}^*, v_{y,i}^*, v_{z,i}^*, t_i^* \right) \right| + \frac{\lambda}{N_{d2}} \sum_{j=1}^{N_{d2}} \left| f^* \left(v_{x,j}^*, v_{y,j}^*, v_{z,j}^*, t_j^* \right) - y_j^* \right|, \qquad (62)$$

where N_{d1} and N_{d2} denote the number of sampling points, and y_j* denotes the given value of f* at $\left(v_{x,j}^*, v_{y,j}^*, v_{z,j}^*, t_j^* \right)$. We sample points ($v_{x,i}$*, $v_{y,i}$*, $v_{z,i}$*, t_i*) and $\left(v_{x,j}^*, v_{y,j}^*, v_{z,j}^*, t_j^* \right)$ on the domain of the EVDF and on the boundary, respectively. The ANN representing the EVDF is obtained by training the ANN to minimize the loss function.

Compared to the mesh-based method, the theoretical foundation of the PINNs has not yet fully established[17]. So far, the ANN architecture is empirically designed in accordance with the problem. The size of the ANN required for representing the EVDF properly would become large with increasing the dimension of the EVDF. Accordingly, the training of the ANN might become difficult. When we extend the PINNs for calculating time dependent EVDF, modification of the ANN architecture and training procedure of the ANN might be required.

6. Conclusions

The equilibrium EVDF under DC uniform electric fields crossed with DC magnetic fields at arbitrary angles was calculated via the BE by applying the PINNs approach. The present method does not require any expansion of the EVDF using orthogonal functions and discretizing velocity space.

The EVDF in SF_6 is calculated in a wide range of *E/N* by varying with the value of *B/N* and that of ψ. The present mesh-free approach enables us to obtain the smooth EVDF and reproduce the anisotropy of the EVDF properly.

The mean electron energy, electron drift velocity, ionization rate coefficient, and electron attachment rate coefficient were calculated from the EVDF and compared with those calculated from the MCS. It was found that the relative difference in those electron transport coefficients between the PINNs and MCS was within 1% in most cases. Furthermore, good agreement in the EEDF between the two methods was also found. This indicates the validity of the present method.

The PINNs approach allows us to represent 3D EVDF properly with approximately 0.01% of memory capacity required for the mesh-based DNS. This feature would enable us

to calculate high-dimensional EVDF with saving computational costs.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Number JP20KK0089.

References

(1) K. Kumar, H. R. Skullerud, and R. E. Robson, Aust. J. Phys. **33**, 343 (1980).

(2) R. Robson, R. White, and M. Hildebrandt, *Fundamentals of Charged Particle Transport in Gases and Condensed Matter* (CRC Press, Boca Raton, 2017), Chap. 2.

(3) Z. Lj. Petrović, S. Dujko, D. Marić, G. Malović, Ž. Nikitović, O. Šašić, J. Jovanović, V. Stojanović, and M. Radmilović-Rađenović, J. Phys. D: Appl. Phys. **42**, 194002 (2009).

(4) D. A. Konovalov, D. G. Cocks, and R. D. White, Eur. Phys. J. D **71**, 258 (2017).

(5) F. Tochikubo and A. Komuro, Jpn. J. Appl. Phys. **60**, 040501 (2021).

(6) M. J. Kushner, J. Phys. D: Appl. Phys. **42**, 194013 (2009).

(7) S. Kawaguchi, K. Takahashi, and K. Satoh, Plasma Sources Sci. Technol. **30**, 035010 (2021).

(8) O. Šašić, J. de Urquijo, A. M. Juárez, S. Dupljanin, J. Jovanović, J. L. Hernández-Ávila, E. Basurto, and Z. Lj. Petrović, Plasma Sources Sci. Technol. **19**, 034003 (2010).

(9) P. W. Stokes, S. P. Foster, M. J. E. Casey, D. G. Cocks, O. González-Magaña, J. de Urqijo, G. García, M. J. Brunger, and R. D. White, J. Chem. Phys. **154**, 084306 (2021).

(10) P. J. Drallos and J. M. Wadehra, J. Appl. Phys. **63**, 5601 (1988).

(11) P. J. Drallos and J. M. Wadehra, Phys. Rev. A **40**, 1967 (1989).

(12) K. Maeda and T. Makabe, Jpn. J. Appl. Phys. **33** 4173 (1994).

(13) M. Kurihara, Z. Lj. Petrović, and T. Makabe, J. Phys. D: Appl. Phys. **33**, 2146 (2000).

(14) H. Sugawara, IEEE Trans. Plasma Sci. **47**, 1071 (2019).

(15) H. Sugawara, Plasma Sci. Technol. **21**, 094001 (2019).

(16) M. Raissi, P. Perdikaris, G. E. Karniadakis, J. Compt. Phys. **378**, 686 (2019).

(17) G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, Nat. Rev. Phys. **3**, 422 (2021).

(18) C. Rao, H. Sun, and Y. Liu, Theor. App. Mech. Lett. **10**, 207 (2020).

(19) L. H. Queiroz, F. P. Santos, J. P. Oliveira, and M. B. Suzan Jr., Digit. Chem. Eng. **1**, 100002 (2021).

(20) N. Zobeiry and K. D. Humfeld, Eng. Appl. Artif. Intell. **101**, 104232 (2021).

(21) A. G. Baydin, B. A. Pearlmutter, A. A. Radul, and J. M. Siskind, J. Mach. Learn. Res. **18**, 1 (2018).

(22) S. Kawaguchi, K. Takahashi, H. Ohkama, and K. Satoh, Plasma Sources Sci. Technol. **29**, 025021 (2020).

(23) T. Uchida and S. Hamaguchi, J. Phys. D: Appl. Phys. **41**, 083001 (2008).

(24) J. T. Gudmundsson, Plasma Sources Sci. Technol. **29**, 113001 (2020).

(25) H. Tagashira, Y. Sakai, and S. Sakamoto, J. Phys. D: Appl. Phys. **10**, 1051 (1977).
(26) H. Itoh, Y. Miura, N. Ikuta, Y. Nakao, and H. Tagashira, J. Phys. D: Appl. Phys. **21**, 922 (1988).
(27) T. Makabe and Z. Lj. Petrovic, Plasma Electronics (CRC Press, Boca Raton, 2014) 2nd ed., Chap.5.
(28) D. A. Konovalov, D. G. Cocks, and R. D. White, arXiv:1612.09008 (2016).
(29) F. Chollet, *Deep Learning with Python* (Manning, Shelter Island, 2018), Chap. 4.
(30) S. Wang, Y. Teng, and P. Perdikaris, SIAM J. Sci. Compt. **43**, A3055 (2021).
(31) D. Misra, arXiv:1908.08681.
(32) I. Goodfellow, Y. Bengo, and A. Courville, *Deep learning* (MIT Press, 2016), Chap. 2.
(33) K. He, X. Zhang, S. Ren, and J. Sun, 2015 IEEE Int. Conf. Comput. Vis. 1026 (2015).
(34) M. D. McKay, R. J. Beckman, and W. J. Conover, Technometrics, **21**, 239 (1979).
(35) S. J. Reddi, S. Kale, and S. Kumar, arXiv:1904.09237.
(36) H. Itoh, T. Matsumura, K. Satoh, H. Date, Y. Nakao, and H. Tagashira, J. Phys. D: Appl. Phys. **26**, 1975 (1993).
(37) S. Kawaguchi, K. Takahashi, K. Satoh, and H. Itoh, Jpn. J. Appl. Phys. **55**, 07LD03 (2016).
(38) H. R. Skullerud, J. Phys. D: Appl. Phys. **1**, 1567 (1968).

Figure Captions

Fig. 1. Schematic diagram of the PINNs approach for solving the governing equation.

Fig. 2. Schematic diagram of the artificial neural network used in this work.

Fig. 3. The values of the loss function $\mathcal{L}$ in SF_6 at *E/N* = 2000 Td, *B/N* = 2000 Hx, and $\psi = 90°$ at different N_d as a function of the number of iterations.

Fig. 4. The value of the electron attachment rate coefficient in SF_6 at *E/N* = 2000 Td, *B/N* = 2000 Hx, and $\psi = 90°$ at different N_d as a function of the number of iterations.

Fig. 5. The values of the loss function $\mathcal{L}$ in SF_6 at *B/N* = 2000 Hx, and $\psi = 90°$ as a function of the number of iterations. The value of *E/N* varies with 2000, 17500, 1500, 1250, 1000, 750, 500, and 250 Td in order.

Fig. 6. Electron velocity distribution $f(v_x, v_z)$ in SF_6 at *E/N* = 2000 Td, *B/N* = 2000 Hx, and $\psi = 90°$. The contours of $f(v_x, v_z)$ are plotted from 0.008 to 0.024 in steps of 0.002.

Fig. 7. Electron energy distribution function in SF_6 at *E/N* = 250, 750, 1250, and 2000 Td. The values of *B/N* and ψ are fixed at 2000 Hx and 90°, respectively.

Fig. 8. Electron transport coefficients in SF_6 at *B/N* = 2000 Hx and $\psi = 90°$ as functions of *E/N*: (a) mean electron energy, (b) electron drift velocity, and (c) ionization and electron attachment rate coefficients. Those coefficients calculated from the PINNs and the MCS are shown in filled and open circles, respectively.

Fig. 9. Relative difference in the electron transport coefficient calculated from the PINNs and MCS in SF_6 at *B/N* = 2000 Hx and $\psi = 90°$ as functions of *E/N*

Fig. 10. Relative difference in the electron transport coefficient between the PINNs and MCS

in SF_6 (a) at *E/N* = 2000 Td and *B/N* = 2000 Hx and (b) at *E/N* = 2000 Td and $\psi = 90°$.

Table I. The size of the matrix used in the artificial neural network. The number of units in the layer is denoted by N_u.

Symbol	Size
X	$N_d \times 3$
W_1, W_2, $W_{z,1}$	$d \times N_u$
b_1, b_2, $b_{z,k}$ $(k = 2, 3, \ldots, N_L)$	$1 \times N_u$
$W_{z,k}$ $(k = 2, 3, \ldots, N_L)$	$N_u \times N_u$
W	$N_u \times 1$
$f^*(X)$	$N_d \times 1$

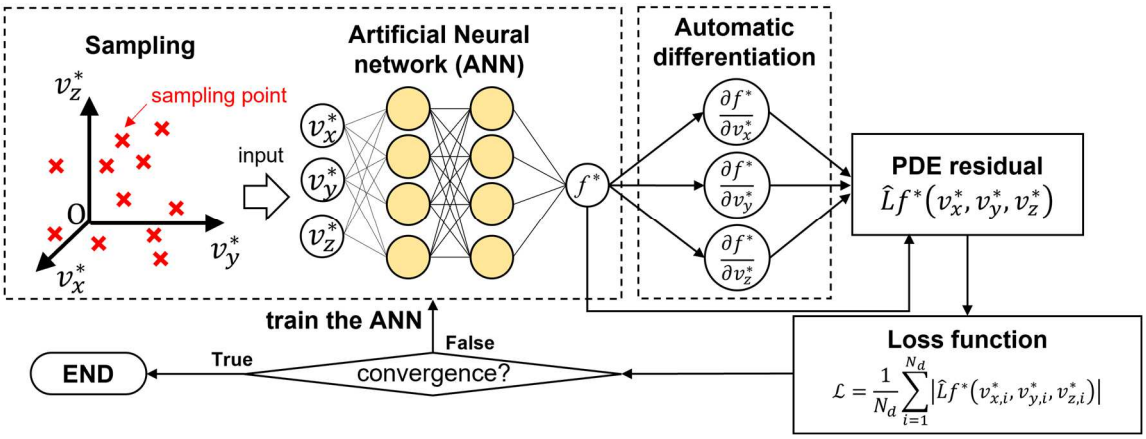


Fig.1.

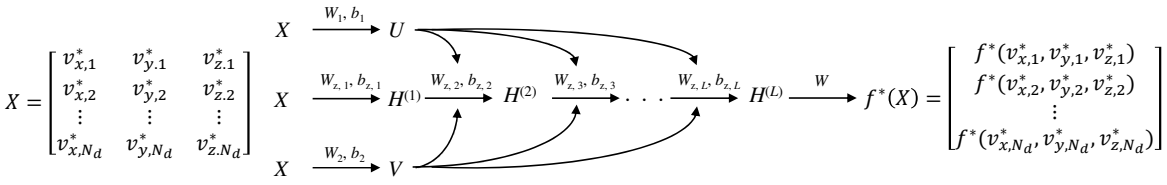

$X = \begin{bmatrix} v^*_{x,1} & v^*_{y,1} & v^*_{z,1} \\ v^*_{x,2} & v^*_{y,2} & v^*_{z,2} \\ \vdots & \vdots & \vdots \\ v^*_{x,N_d} & v^*_{y,N_d} & v^*_{z,N_d} \end{bmatrix}$
X
W_1, b_1
U
X
$W_{z,1}, b_{z,1}$
$H^{(1)}$
$W_{z,2}, b_{z,2}$
$H^{(2)}$
$W_{z,3}, b_{z,3}$
$W_{z,L}, b_{z,L}$
$H^{(L)}$
W
$f^*(X) = \begin{bmatrix} f^*(v^*_{x,1}, v^*_{y,1}, v^*_{z,1}) \\ f^*(v^*_{x,2}, v^*_{y,2}, v^*_{z,2}) \\ \vdots \\ f^*(v^*_{x,N_d}, v^*_{y,N_d}, v^*_{z,N_d}) \end{bmatrix}$
X
W_2, b_2
V

Fig.2.

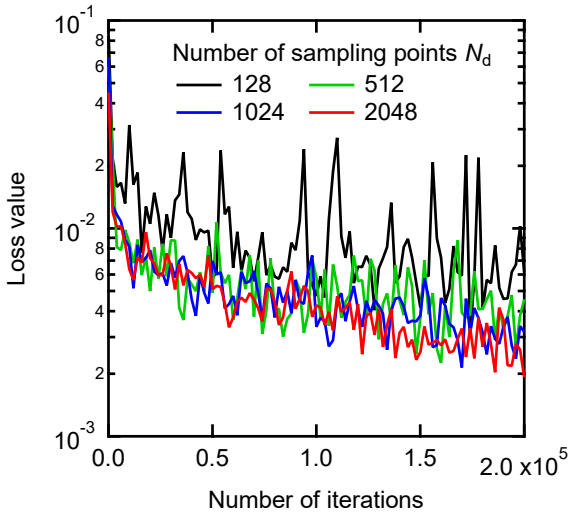

Number of sampling points N_d
128
512
1024
2048
Loss value
Number of iterations
10^{-1}
10^{-2}
10^{-3}
0.0
0.5
1.0
1.5
2.0 x10^5

Fig.3.

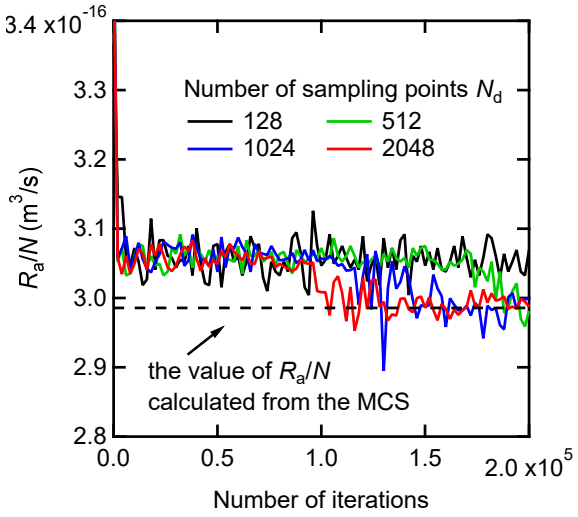

3.4 x10^-16
3.3
3.2
3.1
3.0
2.9
2.8
Number of sampling points N_d
128
512
1024
2048
R_a/N (m^3/s)
the value of R_a/N
calculated from the MCS
0.0
0.5
1.0
1.5
2.0 x10^5
Number of iterations

Fig.4

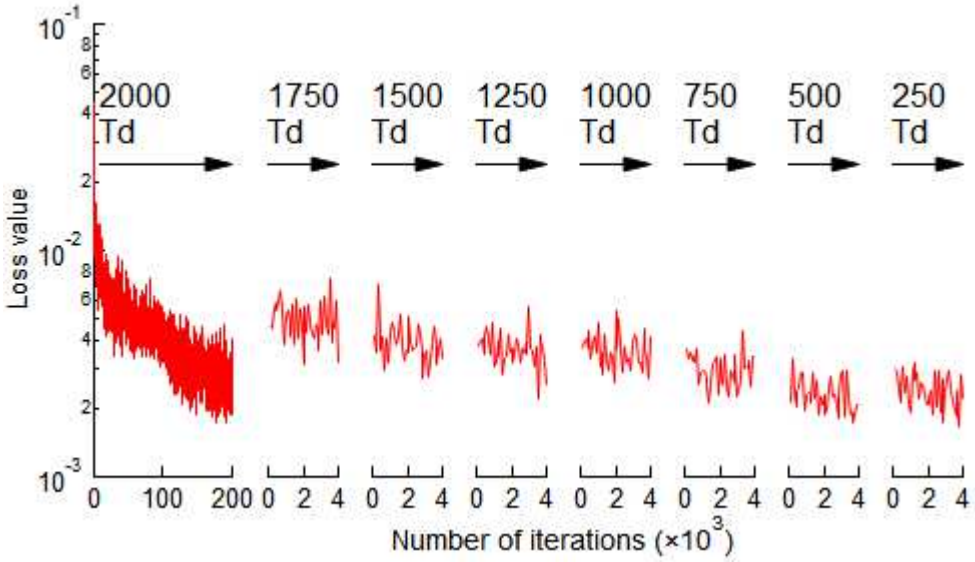

10^-1
10^-2
10^-3
Loss value
2000 Td
1750 Td
1500 Td
1250 Td
1000 Td
750 Td
500 Td
250 Td
0 100 200 0 2 4 0 2 4 0 2 4 0 2 4 0 2 4 0 2 4 0 2 4
Number of iterations (×10^3)

Fig.5.

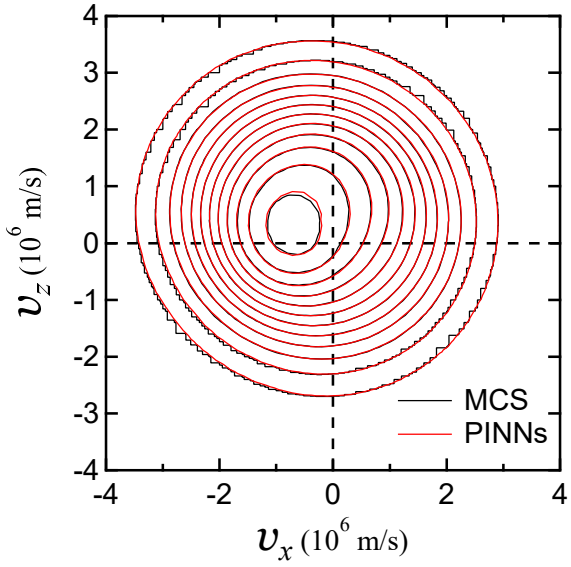

MCS
PINNs
v_z (10^6 m/s)
v_x (10^6 m/s)

Fig.6.

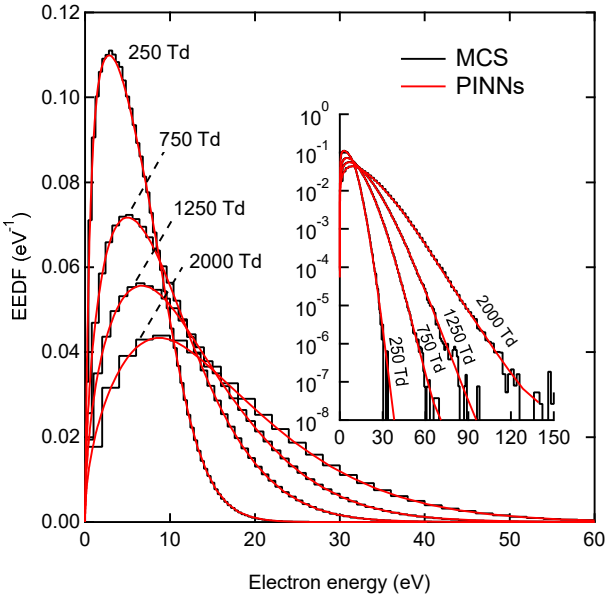
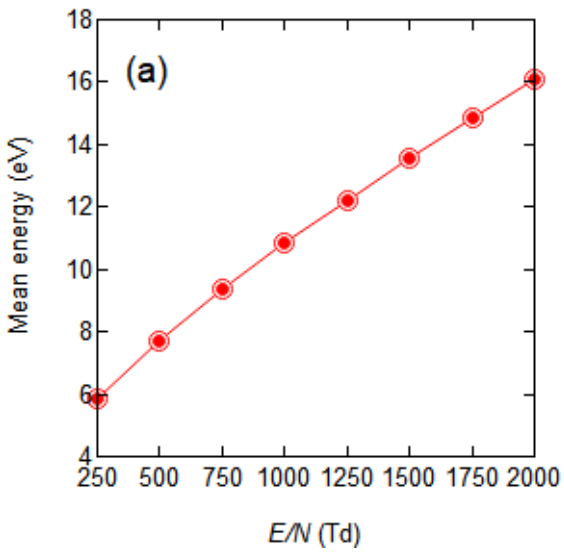
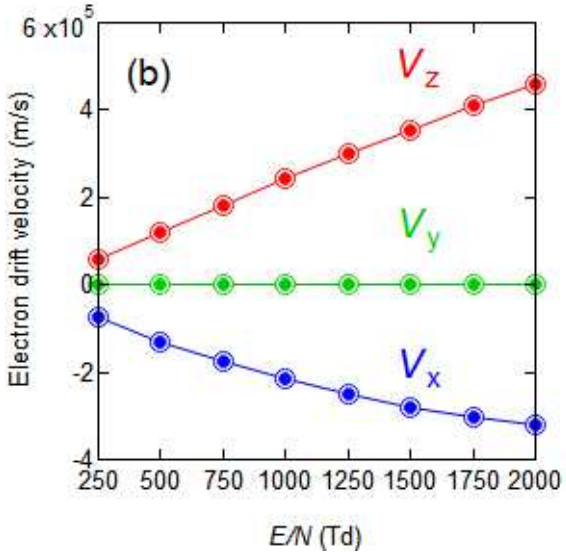

250 Td
750 Td
1250 Td
2000 Td
MCS
PINNs
EEDF (eV^{-1})
Electron energy (eV)

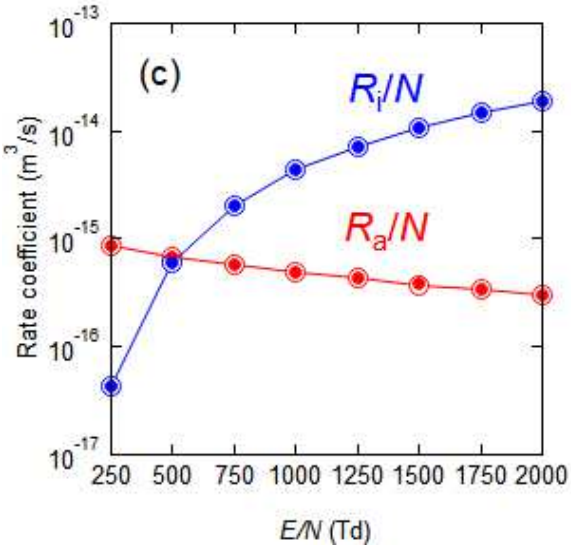
Fig.7.


(a)
Mean energy (eV)
E/N (Td)

(a)


(b)
6 x10^5
Electron drift velocity (m/s)
V_z
V_y
V_x
E/N (Td)

(b)


(c)
Rate coefficient (m^3/s)
R_i/N
R_a/N
E/N (Td)

(c)

Fig. 8.

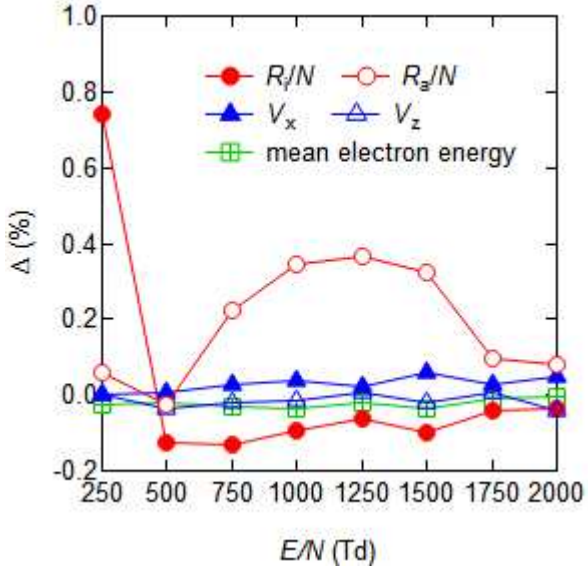

1.0
0.8
0.6
0.4
0.2
0.0
-0.2
Δ (%)
250 500 750 1000 1250 1500 1750 2000
E/N (Td)
R_i/N
R_a/N
V_x
V_z
mean electron energy

Fig. 9.

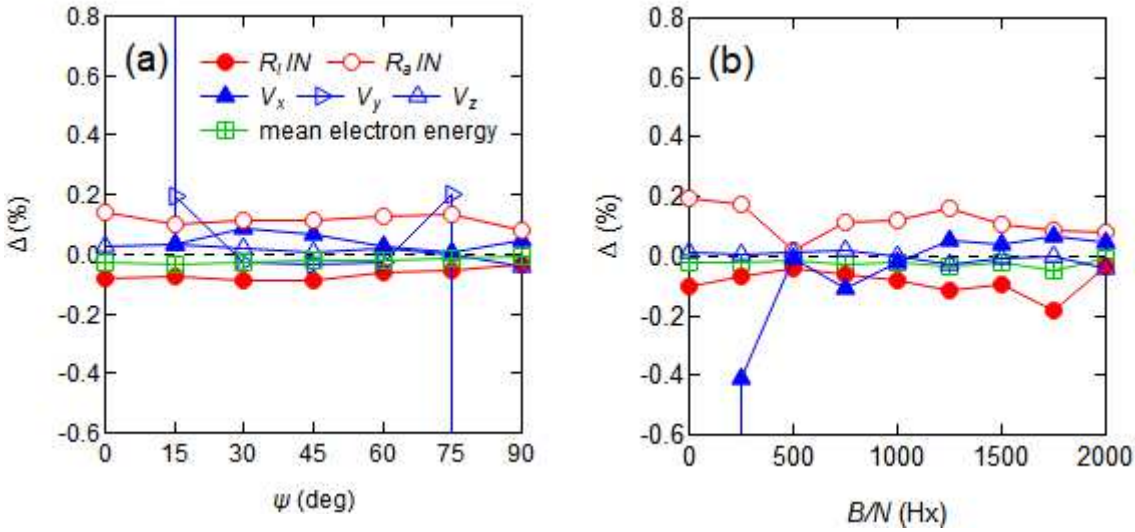

(a)
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
Δ (%)
0 15 30 45 60 75 90
ψ (deg)
R_i /N
R_a /N
V_x
V_y
V_z
mean electron energy
(b)
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
Δ (%)
0 500 1000 1500 2000
B/N (Hx)

Fig.10.