

PAPER

Implicit and coupled fluid plasma solver with adaptive Cartesian mesh and its applications to non-equilibrium gas discharges

To cite this article: Robert R Arslanbekov and Vladimir I Kolobov 2021 *Plasma Sources Sci. Technol.* **30** 045013

View the [article online](#) for updates and enhancements.



IOP | ebooks™

Bringing together innovative digital publishing with leading authors from the global scientific community.

Start exploring the collection—download the first chapter of every title for free.

Implicit and coupled fluid plasma solver with adaptive Cartesian mesh and its applications to non-equilibrium gas discharges

Robert R Arslanbekov¹ and Vladimir I Kolobov^{1,2,*} 

¹ CFD Research Corporation, Huntsville, AL, United States of America

² University of Alabama in Huntsville, Huntsville, AL, United States of America

E-mail: Vladimir.Kolobov@uah.edu

Received 10 February 2021

Accepted for publication 18 March 2021

Published 26 April 2021



Abstract

We present a new fluid plasma solver with adaptive Cartesian mesh (ACM) based on a full-Newton (nonlinear, implicit) scheme for non-equilibrium gas discharge plasma. The electrons and ions are described using drift–diffusion approximation coupled to Poisson equation for the electric field. The electron-energy transport equation is solved to account for electron thermal conductivity, Joule heating, and energy loss of electrons in collisions with neutral species. The rate of electron-induced ionization is a function of electron temperature and could also depend on electron density (important for plasma stratification). The ion and gas temperature are kept constant. The transport equations are discretized using a non-isothermal Scharfetter–Gummel scheme to resolve possible large temperature gradients in the sheaths. We demonstrate the new solver for simulations of direct current (DC) and radiofrequency (RF) discharges. The implicit treatment of the coupled equations allows using large time steps. The full-Newton method (FNM) enables fast nonlinear convergence at each time step, offering significantly improved simulation efficiency. We discuss the selection of time steps for solving different plasma problems. The new solver enables solving several problems we could not solve before with existing software: two- and three-dimensional structures of the entire DC discharges including cathode and anode regions, electric field reversals and double-layer formation, the normal cathode spot and an anode ring, moving striations in diffuse and constricted DC discharges, and standing striations in RF discharges. The developed FNM-ACM technique offers many benefits for tackling the disparity of gas discharge plasma systems' time scales and nonlinearity.

Keywords: gas discharges, multi-fluid model, implicit coupled solver, low temperature plasma, adaptive Cartesian mesh, stratification

 Supplementary material for this article is available [online](#)

(Some figures may appear in colour only in the online journal)

1. Introduction

Plasmas are characterized by a disparity of time scales and nonlinear behavior. The disparity of the time scales comes

from the distinction of electron and ion mass. The nonlinearities appear from coupling charged particle transport with electric fields and from the ionization processes, which are highly sensitive to electron energy spectra. Implicit solvers can address the disparity of time scales [1]. Such solvers enable

* Author to whom any correspondence should be addressed.

using time steps exceeding the fast processes' characteristic time scales by effectively removing the time derivatives from the corresponding transport equations for electrons. Ideally, implicit solvers allow obtaining steady-state solutions in a single time step using nonlinear iterations to address coupling and nonlinearities of equations describing plasmas.

These strategies have been implemented in some form in most of the existing kinetic and fluid plasma solvers [2]. Existing multi-fluid, multi-temperature models use finite volume (FV) [3, 4] or finite element [5] space discretization and solve plasma transport equations sequentially using iterations for solving each equation. The sequential solution reduces computer memory usage but limits the time step by the convergence requirement—the 'sequentialization penalty'. For thermal plasma, a fully implicit multi-fluid reactive model with a single temperature for all species was implemented [6], assuming quasi-neutrality and a local thermodynamic equilibrium model for electrons. Fully implicit solvers require considerable memory, which limits the size of the problem they can solve.

For kinetic plasma simulations, particle-in-cell (PIC) methods and discrete velocity models have been developed. They are very computationally expensive compared to the fluid models because they typically operate on the shortest time- and length scales and calculate the velocity distribution functions of plasma species rather than their macro-parameters such as density, mean velocity, and temperature. Implicit and semi-implicit particle-based and mesh-based kinetic solvers have also been developed [7, 8]. It appears that the best procedure for addressing the disparity of time scales is to separate the fast and slow processes into separate blocks to enable individual control of the time steps and use 'recycling'. Such a 'recycling' procedure is well known for the PIC methods [9].

Recent demands for understanding and addressing the disparity of temporal, spatial, and energy scales in plasmas come from developing adaptive kinetic-fluid solvers [10]. Many plasma problems require space, time, and model adaptation for an efficient solution [11, 12]. The need for implicit solvers and adaptive time steps has been recognized recently [13] for problems that require resolving both electron and ion time scales. A typical example is a gas discharge maintained by alternating electric fields at high frequencies [from radio-frequency (RF) to microwave and optical range]. Typically, at high frequencies, ions do not respond to the time variations of the fields maintaining the plasma; they only respond to a slow-varying electric field generated by the plasma.

In contrast, electrons could respond to the high-frequency field dynamics forming sheaths and skin layers, where electrons acquire kinetic energy for gas ionization. How to solve such problems in the general case with implicit coupled solvers? How to select the appropriate time step? Can we develop optimal strategies for simulations of collisional plasma operating at different frequency ranges? The present paper attempts to address some of these challenges.

Implicit methods or, more generally, full Newton methods (FNMs) have been introduced in the late 80s and early 90s for modeling semiconductor devices [14–20]. They became a standard in most of the commercially available simulators,

such as Medici, Dessis, and others. The NanoTCAD software [21, 22] relies on FNM implementation on a binary Cartesian mesh. The FNM approach is also used now in many computational fluid dynamics (CFD) codes [23–26] to overcome the numerical stiffness intrinsically present in CFD problems. The FNM approach is typically applied to the set of Navier–Stokes equations for the density, momentum, and energy equations describing the mean flow properties while solving the often-stiff chemistry for species fractions in the same global matrix constructed from the Jacobians.

Automatic mesh generation and dynamic mesh adaptation to a solution and/or a changing geometry is a hot topic in modern computational physics. Many magnetohydrodynamics (MHD) codes for space plasma use adaptive mesh refinement (AMR) [27], which is also an essential part of today's CFD codes such as Converge [28], Simerics [29], FloEFD [30], and others. We have previously implemented an adaptive mesh and algorithm refinement (AMAR) framework [31], which allows us to adapt the computational mesh to locally required spatial resolution and select kinetic or fluid models on a cell-by-cell basis. In the AMAR framework, an open-source Gerris flow solver, GFS [32], is used for generating adaptive Cartesian mesh (ACM) for complex embedded boundaries. GFS produces a 2:1 balanced grid, which means that (a) the levels of direct neighbors cannot differ by more than one and (b) the levels of diagonal neighbors cannot differ by more than one. These additional constraints simplify the gradient and flux calculations. Fully threaded tree, the pointer-based data structure of GFS allows fast access to cell neighbors, facilitating efficient implementation of dynamic AMR and parallelization. Besides the cut-cell technique for the boundary treatment, we have also implemented the immersed boundary method [33].

The present work is devoted to combining ACM and FNM techniques for fluid plasma simulations. The authors are not aware of any publications describing the FNM implementation for ACM. Commercial software COMSOL uses FNM with the finite element method (FEM) in semiconductor and plasma modules [34]. A recently developed Zapdos code [35] also relies on FEM and FNM for solving partial differential equations (PDEs) via interfacing with PETSc library [36]. Both COMSOL and Zapdos solve the fluid plasma equations using a logarithmic transformation for the particle densities and the electron energy density to ensure the solution's positivity. The log transformation introduces additional nonlinearity and makes the implementation more involved than the natural variables used in the present work. Most of the current plasma codes for modeling gas discharges are based on a segregated solution approach. For example, our previous ACM plasma code [31] uses explicit solvers suitable for the treatment of fast (electron) time scales. The CFD-ACE+ code [50] uses implicit solvers. However, the PDEs are solved sequentially (i.e., in an uncoupled or loosely coupled manner; e.g., via additional sub-iterations), which may pose numerical issues for solving problems with strong coupling (e.g., large plasma densities).

The structure of the paper is as follows. Section 2 discusses the implementation of the FNM-ACM solver for the commonly-used drift–diffusion (DD) plasma model. Section 3 contains simulation results demonstrating the capabilities of

the method. We discuss the selection of time steps and compare the efficiency of the explicit and FNM plasma solvers. Finally, section 4 includes a conclusion and outlook.

2. Full Newton method implementation in the ACM framework

2.1. Plasma model

The plasma model equations adopted in this work include the balance equations for the electron (n_e) and ion (n_i) densities coupled to Poisson equation for the electrostatic field:

$$q_e \frac{\partial n_e}{\partial t} - \nabla \cdot \mathbf{J}_e = q_e S_e, \quad (1)$$

$$q_e \frac{\partial n_i}{\partial t} + \nabla \cdot \mathbf{J}_i = q_e S_i. \quad (2)$$

In these equations $q_e = |e|$ is the absolute value of the electron charge, S_k are the source (e.g., ionization) and sink (e.g., recombination) terms, $\mathbf{J}_e$ and $\mathbf{J}_i$ are the electron and ion current densities, respectively. The electron energy transport equation is:

$$\frac{\partial(n_e \epsilon_e)}{\partial t} + \nabla \cdot \boldsymbol{\xi}_e = \mathbf{E} \cdot \mathbf{J}_e - n_e \sum_{r=1}^{N_{\text{reactions}}} K_r \Delta \epsilon_r, \quad (3)$$

where ϵ_e is the average electron energy, K_r is the reaction rate and $\Delta \epsilon_r$ is the electron energy change per electron per collision. For inelastic collisions $\Delta \epsilon_r$ corresponds to the energy threshold of an r -type collision (e.g., excitation or ionization processes). The energy loss corresponding to elastic collisions is given by

$$\Delta \epsilon_r = \frac{3}{2} \frac{m_e}{M_r} k_B (T_e - T_g), \quad (4)$$

where m_e is the electron mass and M_r is the mass of the heavy particle (e.g., atom) and T_g is the common temperature of neutral plasma species.

The electron current density is

$$\mathbf{J}_e = q_e D_e \nabla n_e - q_e n_e \mu_e \nabla \left(\varphi - \frac{k_B T_e}{q_e} \right), \quad (5)$$

or, in a different form

$$\mathbf{J}_e = q_e D_e \nabla n_e + q_e n_e \left\{ \mu_e \nabla (-\varphi) + D_e \frac{\nabla T_e}{T_e} \right\}. \quad (6)$$

Similarly, for the ion species ($i = i_1, i_2, \dots$)

$$\mathbf{J}_i = -q_e Z_i D_i \nabla n_i + q_e Z_i n_i \left\{ \mu_i \nabla (-\varphi) - D_i \frac{\nabla T_i}{T_i} \right\}. \quad (7)$$

Here, the electron and ion mobilities are μ_e and μ_i , and the corresponding diffusion coefficients, $D_{e,i} = \frac{k_B T_{e,i}}{q_e} \mu_{e,i}$. The DD approximation for the electron and ion fluxes neglects inertial effects, which is a reasonable approximation for collision-dominated gas discharge plasmas. Furthermore, the ion motion is dominated by drift. The ion diffusion term is negligible in most computational domains and can only be comparable to the drift term at zero electric field points. It is known that the

ion inertia effects are essential in weakly collisional plasma, and they will be added to our solver in future work.

The electron energy flux $\boldsymbol{\xi}_e$ in equation (3) can then be written as:

$$\boldsymbol{\xi}_e = -\kappa_e \nabla T_e - (\epsilon_e + k_B T_e) \frac{\mathbf{J}_e}{q_e}. \quad (8)$$

The average electron energy ϵ_e consists of thermal energy, $3/2 k_B T_e$, and the electron kinetic energy $1/2 m_e v^2$. We can neglect the electron directed kinetic energy compared with the thermal energy to obtain:

$$\boldsymbol{\xi}_e = -\kappa_e \nabla T_e - \frac{5 k_B T_e}{2} \frac{\mathbf{J}_e}{q_e}. \quad (9)$$

According to (9), the electron energy flux $\boldsymbol{\xi}_e$ contains thermal conduction and convection terms. The thermal conductivity coefficient (κ_e) for electrons is given by the Wiedemann–Frantz law:

$$\kappa_e = \left(\frac{5}{2} + c_e \right) \frac{k_B^2}{q_e} T_e \mu_e n_e. \quad (10)$$

For semiconductors, the value of $c_e = -0.5$ is commonly used [18], while $c_e = 0$ is common for plasma. In most publications devoted to plasmas, the energy density flux is used in the form:

$$\boldsymbol{\xi}_e = -\frac{5}{3} D_e \nabla(n_e \epsilon_e) - \frac{5}{3} \mu_e \nabla \varphi (n_e \epsilon_e). \quad (11)$$

However, the energy flux expressed by equation (9) has a more transparent physical meaning since it separates the thermo-diffusion term explicitly. Equation (9) also indicates that in the quasi-neutral plasma, where the electron flux ($\mathbf{J}_e$) transforms into a ambipolar diffusion flux, $\mathbf{J}_a \sim \mathbf{J}_i$, the thermal diffusion (proportional to the free electron diffusion coefficient) dominates over convection. Indeed, the ratio of the free electron diffusion coefficient to the ambipolar diffusion coefficient scales as $D_e/D_a \sim \sqrt{M/m_e} \gg 1$ for typical $T_e \gg T_i$. The temperatures of ions and neutrals are kept constant, and the background gas is assumed motionless.

The system of the particle density and electron energy transport equations is completed by Poisson equation for the electrostatic potential φ :

$$\nabla \cdot (\epsilon \mathbf{E}) = \nabla \cdot (-\epsilon \nabla \varphi) = q_e \left(\sum Z_i n_i - n_e \right), \quad (12)$$

where $\epsilon = \epsilon_r \epsilon_0$ is the permittivity of the plasma and ϵ_0 is that of vacuum (for gas, $\epsilon_r = 1$).

The system of plasma equations is fully defined by setting appropriate boundary conditions at plasma boundaries. At metal walls, the following boundary conditions are set for the electron and ion fluxes:

$$\mathbf{n} \cdot \mathbf{J}_e = q_e \left(\frac{1}{4} v_e n_e - \sum \gamma_i (\mathbf{n} \cdot \boldsymbol{\Gamma}_i) \right), \quad (13)$$

$$\mathbf{n} \cdot \mathbf{J}_i = q_e \left(\frac{1}{4} v_i n_i + a n_i \mu_i (\mathbf{n} \cdot \mathbf{E}) \right), \quad (14)$$

where γ_i is the secondary electron emission coefficient due to i th ion bombardment, the ion flux $\Gamma_i = \mathbf{J}_i/q_e$, and

$$a = \begin{cases} 1, & (\mathbf{n} \cdot \mathbf{E}) > 0 \\ 0, & (\mathbf{n} \cdot \mathbf{E}) \leq 0 \end{cases},$$

with $\mathbf{n}$ being a unit vector normal to the wall. For the energy flux, the following boundary condition is set:

$$\mathbf{n} \cdot \boldsymbol{\xi}_e = \frac{1}{3} v_e (n_e \epsilon_e) - 2k_B T_e \sum \gamma_i (\mathbf{n} \cdot \Gamma_i). \quad (15)$$

At metal boundaries, the electrostatic potential can be set to given values (so-called Dirichlet-type boundary conditions), which are in turn explicitly prescribed or computed from external circuit solutions.

At dielectric walls, the same set of boundary conditions defined in equations (13)–(15) is used (with typically $\gamma_i = 0$). In order to compute the plasma potential at these boundaries, we introduce a surface charge, σ , for which we solve the following time dependent (local) balance equation:

$$\frac{\partial \sigma}{\partial t} = \sum (\mathbf{n} \cdot \mathbf{J}_i) + (\mathbf{n} \cdot \mathbf{J}_e) \quad (16)$$

at all boundary faces. The surface charge density is then used to compute the appropriate boundary conditions at boundary faces for the electrostatic potential through Gauss' law:

$$\varepsilon (\mathbf{n} \cdot \mathbf{E}) = \varepsilon (\mathbf{n} \cdot (-\nabla \varphi)) = \sigma. \quad (17)$$

At metal boundaries, the electrostatic potential is set to given values, which can be explicitly prescribed or computed from external circuit solutions. At dielectric walls, we calculate a surface charge, σ , by integrating charged particle fluxes at the boundary faces.

The DD model for electrons and ions described above is common for gas discharge and semiconductor device simulations [37]. Often, the rates of electron-induced reactions in the plasma model are obtained by solving the *local* Boltzmann equation for electrons [38] over a range of n_e and the reduced electric field, E/N , to obtain look-up-tables for the ionization rate as a function of n_e and electron temperature T_e [39, 50]. Instead of using this procedure, in the present work, we express the electron-induced ionization rate in the Arrhenius form:

$$R_e(T_e) = A T_e^B e^{-E_a/T_e}, \quad (18)$$

where A , B , and E_a are constants. For modeling plasma stratification, we introduce a nonlinearity arising due to Maxwellization of the electron energy distribution function (EEDF) [40, 41]:

$$R_e(n_e, T_e) = A T_e^B e^{-E_a/T_e} \begin{cases} \exp\left(\frac{n_e}{n_c}\right), & n_e < n_1 \\ \exp\left(\frac{n_1}{n_c}\right), & n_e > n_1 \end{cases}. \quad (19)$$

Here, n_c controls the rate of the nonlinear dependence, and n_1 defines the saturation value. The volume recombination is expressed in the form:

$$S_{\text{recomb}} = \beta n_i n_e. \quad (20)$$

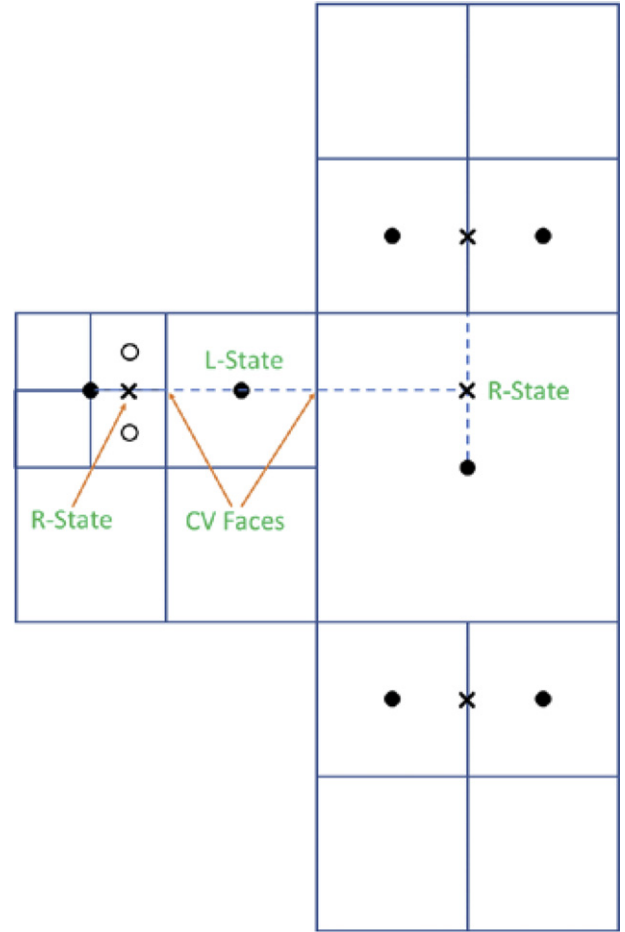


Figure 1. Schematic representation of face reconstruction on ACM with the cell to the left being one-level up and the cell to the right being one-level down (after reference [32]). Arrows show cell faces (here, right, and left). Also shown are right-state locations (crosses) involved in face reconstruction in the developed SG scheme.

The recombination rate, β , was varied to study the transition between diffuse to constricted discharges.

2.2. Discretization of the enhanced SG scheme

We have utilized the FV approach with cell-centered variables to discretize the plasma equations. Figure 1 illustrates a control-volume (CV) cell together with neighboring cells of different refinement levels. In our ACM framework, the refinement level difference of neighbor cells cannot be larger than one, so that only one hanging node per any given face is allowed. The FV spatial reconstruction scheme brings the neighbor cell values to the same level along the coordinate axes by invoking neighboring cells in the perpendicular direction to the CV cell's current face across which the fluxes are computed. Then, the right-state (whose locations are marked by cross symbols in figure 1) is reconstructed by a linear combination of all involved neighbor cell-centered values, while the left-state consists of the cell-centered values of the CV cell itself. This way, second order spatial accuracy is achieved when performing reconstruction at faces involving not axis aligned cells (so-called transversal or diagonal neighbors) [32].

The Scharfetter–Gummel (SG) scheme for computing the particle and energy fluxes at cell faces assumes that all quantities, except charged species densities, remain unchanged between the left (L) and right (R) states shown in figure 1 [14, 16, 18]. The electron and ion density fluxes at the faces are computed as

$$J_e^f = \frac{q_e \mu_e^f U_{T_e}^f}{\Delta_f} \left[n_{e,R} B \left(\frac{\varphi_R - \varphi_L}{U_{T_e}^f} \right) - n_{e,L} B \left(-\frac{\varphi_R - \varphi_L}{U_{T_e}^f} \right) \right],$$

$$J_i^f = \frac{Z_i q_e \mu_i^f U_{T_i}^f}{\Delta_f} \left[n_{i,L} B \left(\frac{\varphi_R - \varphi_L}{U_{T_i}^f} \right) - n_{i,R} B \left(-\frac{\varphi_R - \varphi_L}{U_{T_i}^f} \right) \right],$$

where the face quantities (locations indicated are arrows in figure 1) are labeled by ‘f’ superscript index, $q_e = |e|$ is the absolute value of the electron charge, $U_{T_{e,i}} = k_B T_{e,i} / q_e$ is the thermal electrical constant, $\mu_{e,i}^f$ are electron and ion mobilities at cell faces, Δ_f is the face normal distance between left and right (neighbor) cell centers or positions (depending on the neighboring cell refinement levels), and $B(x) = x / (e^x - 1)$ is the Bernoulli function.

The face values of the electron and ion mobilities are interpolated as

$$\mu_{e,i}^f = \mu_{e,i}^L + w^f (\mu_{e,i}^L - \mu_{e,i}^R), \quad (23)$$

with a face weight, w^f , determined by the face neighbor geometry (fine–fine, coarse–fine, etc; e.g., $w^f = 0.5$ on a fine–fine interface) while the right-state values are reconstructed by a linear combination of neighboring cell-center values. Similarly, $U_{T_{e,i}}^f$ and all other face properties such as electron and ion diffusion coefficients are computed.

An essential feature of the developed model is an extended, nonisothermal SG scheme. Such schemes are commonly and successfully used in semiconductor TCAD codes but are not common for plasma simulations. The extended SG scheme for the electron current continuity equation takes into account the spatial variation of the electron temperature. When the electron temperature varies across neighboring cells, the nonisothermal SG scheme for the electron density flux becomes [18, 19]:

$$J_e^f = \frac{q_e D_e^f (T_{e,R} - T_{e,L})}{\Delta_f \ln \left(\frac{T_{e,R}}{T_{e,L}} \right)} \left[B \left(\tilde{\Delta} \right) \frac{n_{e,R}}{T_{e,R}} - B \left(-\tilde{\Delta} \right) \frac{n_{e,L}}{T_{e,L}} \right], \quad (24)$$

where

$$\tilde{\Delta} = \frac{\ln(T_{e,R}/T_{e,L})}{(T_{e,R} - T_{e,L})} \left[\frac{q_e}{k_B} (\varphi_R - \varphi_L) - 2(T_{e,R} - T_{e,L}) \right]. \quad (25)$$

We have used these equations in our FNM code. The value of such a treatment could be particularly important near plasma boundaries and electrodes (cathode and anode) where the electron temperature can vary sharply over short distances.

There are several approaches to the energy flow’s discretization in the nonisothermal SG scheme [20]. We adopted the discretization scheme, in which the electron energy flow is written as [20]

$$S_e^f = -\frac{5}{2} \frac{k_B D_e^f (T_{e,R} - T_{e,L})}{\Delta_f \ln(T_{e,R}/T_{e,L})} \times \left[B \left(\tilde{\Delta} \right) n_{e,R} - B \left(-\tilde{\Delta} \right) n_{e,L} \right].$$

The Joule heating source often represents a source of numerical instabilities. An implicit treatment of the electron energy source made it possible to increase the time step by several orders of magnitude in conventional (non-FNM) plasma codes [4]. In our FNM approach, we have attempted two techniques. In the first one, the Joule heating term was rewritten as [18]

$$\mathbf{E} \cdot \mathbf{J}_e = -\nabla \cdot (\varphi \mathbf{J}_e) + q_e \varphi \left(-\frac{\partial n_e}{\partial t} + S_e \right). \quad (27)$$

In the second approach, by converting the CV cell (Ω) integration to summation over cell faces, we obtained [17, 20]

$$\int_{\Omega} \mathbf{J}_e \cdot \nabla \varphi dV = \int_{\Omega} \nabla \cdot (\varphi \mathbf{J}_e) dV - \int_{\Omega} \varphi \nabla \cdot (\mathbf{J}_e) dV$$

$$= \sum_f (\varphi_f - \varphi_L) J_e^f \Delta S^f,$$

with ΔS^f being the face surface area and f index denotes summation over cell faces. We have implemented and tested both these methods and found that the second scheme gives more robust results and better nonlinear convergence at each time step.

The implicit treatment of the boundary conditions is essential for the FNM scheme consistency to achieve better convergence properties. The boundary conditions at metal electrodes and dielectric walls are treated implicitly by adequately setting the corresponding fluxes across the boundary cell faces and computing the related Jacobian terms.

The mesh refinement/coarsening criteria can be specified in the simulation scripts based on a combination of several events (see reference [32] for details). In these events, the minimum and maximum refinement levels are specified together with a set of sensitivity thresholds (or minimum variation of solution variable to be resolved across all neighboring cells of any given cell) for each particular grid adaptation event, which can be a function of cell position and solution variables. Such criteria can be based on computational cell locations (e.g., proximity to the plasma walls and electrodes), gradients or magnitudes of the solution variables, or any functions or combinations of such variables. Hence, depending on the problems being solved, one can build a rather sophisticated set of grid adaptation events to ensure proper grid resolution and solver efficiency for steady-state and transient problems. We have used simple grid adaptation events based on gradients of the electrostatic potential and electron density for plasmas controlled by ambipolar diffusion in the present work. More sophisticated grid adaptation criteria have been previously used for high-pressure gas breakdown simulations and streamer simulations

with AMR [42]. For steady-state simulations, we typically adapt the grid every 100–1000 time steps, while for transient simulations, the grid is adapted more often (e.g., each time step or every tenth step).

We paid attention to properly redistributing conserved quantities across smaller cells upon refinement. The zeroth-order algorithm would uniformly distribute the parent cell solution across all children cells of the next refinement level. We used the reconstruction approach based on the first-order spatial accuracy approximation, which employs slope-limited (e.g., Van Leer) gradients. The gradients were computed at cell centers of parent (coarse) cells, from which the values at cell centers of children (fine) cells were constructed. For some variables (such as electron density), we apply the logarithmic transformation to ensure further robustness (which preserves positiveness). During cell coarsening, the children's cells' variables are cast into their larger parent cell using a cell volume averaging technique, ensuring proper quantity conservation.

2.3. FNM implementation

In the section, we describe the FNM technique's implementation for solving the discretized system of the plasma transport equations coupled to Poisson equation. In the FNM method, the full set of plasma equations with N ion species is represented as

$$\mathbf{F}(\varphi, n_e, n_{i1}, \dots, n_{iN}, T_e) = 0. \quad (29)$$

It is also possible to use the total electron energy ($\mathcal{E}_e = n_e \bar{\epsilon}_e$) as an independent variable (instead of the electron temperature). Such an approach has been used in reference [20] to secure Jacobian matrix diagonal dominance (especially in the regions with low electron density) as well as in reference [35] in combination with log transformation of $\mathcal{E}_e$. We have implemented both approaches in our FNM scheme and have not observed significant advantages of the latter approach for the conducted test problems. However, as we continue implementing new features (such as implicit treatment of the finite-rate chemistry) and broaden the range of test problems, we keep assessing these two approaches for better performance.

Then, in the selected FNM scheme, the complete set of multi-ion plasma equations with N_{ions} ion species can be represented in a vector form as

$$\mathbf{F}(\mathbf{Q}) = 0, \quad (30)$$

where the $N_{\text{ions}} + 3$ solution vector (defined at cell centers of CV cells) is

$$\mathbf{Q} = \begin{pmatrix} \varphi \\ n_e \\ n_{i1} \\ \dots \\ n_{iN} \\ T_e \end{pmatrix}, \quad (31)$$

$$\mathbf{F} = \begin{pmatrix} F_\varphi \\ F_{n_e} \\ F_{n_{i1}} \\ \dots \\ F_{n_{iN}} \\ F_{T_e} \end{pmatrix} = \begin{pmatrix} \nabla \cdot (-\epsilon \nabla \varphi) - q_e \left(\sum_i Z_i n_i - n_e \right) \\ q_e \frac{\partial n_e}{\partial t} - \nabla \cdot \mathbf{J}_e - q_e S_e \\ q_e \frac{\partial n_{i1}}{\partial t} + \nabla \cdot \mathbf{J}_{i1} - q_e S_{i1} \\ \dots \\ q_e \frac{\partial n_{iN}}{\partial t} + \nabla \cdot \mathbf{J}_{iN} - q_e S_{iN} \\ \frac{\partial (n_e \frac{3}{2} k_B T_e)}{\partial t} + \nabla \cdot \xi_e - \mathbf{E} \cdot \mathbf{J}_e + n_e \sum_r K_r \epsilon_r \end{pmatrix}. \quad (32)$$

As one can see, we include the ion species into the same matrix together with electrons and electrostatic potential. Such a treatment resembles the techniques commonly used in semiconductor device modeling. It may be advantageous for plasma systems to separate the electron and ion transport simulations, thus considerably reducing the matrix size and the computational cost. In the present work, we explored the fully coupled approach to illustrate the method's capabilities. The fully coupled method may be advantageous for modeling ion–ion plasmas and magnetized plasmas where electron transport is strongly affected by the magnetic field.

By introducing the following update rule (for each implicit time step advance $m \leftarrow m + 1$) with k being the nonlinear (implicit) sub-iteration index, one can write down a compact form with an update vector $\Delta \mathbf{Q}$

$$\mathbf{Q}^{m+1,k+1} = \mathbf{Q}^{m+1,k} + \Delta \mathbf{Q}^{m+1,k+1}, \quad (33)$$

where the initial condition for each Newton sub-iteration k is:

$$\mathbf{Q}^{m+1,k=0} = \mathbf{Q}^m. \quad (34)$$

Following reference [18], we employ the first-order backward-differentiation-formula to the time-derivative terms. For the time discretization, composite techniques employing both the trapezoidal rule and the second-order backward-differentiation-formula were used, as described in reference [18]. A Newton iteration is then solved as

$$\frac{\partial \mathbf{F}(\mathbf{Q}^k)}{\partial \mathbf{Q}} \Delta \mathbf{Q}^{k+1} = -\mathbf{F}(\mathbf{Q}^k), \quad (35)$$

where the time marching index m is omitted for brevity.

In the FV formulation, the discretized flux vector Φ across each face of a CV cell becomes a function of the left state $\mathbf{Q}_L$ (composed of the cell center values of the CV cell, i.e., $\mathbf{Q}_L = \mathbf{Q}$) and the right state $\mathbf{Q}_R$ (involving direct as well as indirect neighboring cell values, see figure 1), i.e., $\Phi = \Phi(\mathbf{Q}_L, \mathbf{Q}_R)$. Then the discretized form of the Jacobian matrix for a Newton iteration takes two forms for the left and right (neighbor) state vectors. For plasma with one ion-species, we can write:

$$\mathbf{r}_L^k = \frac{\partial \mathbf{F}(\mathbf{Q}^k)}{\partial \mathbf{Q}_L} = \begin{pmatrix} \frac{\partial F_\varphi}{\partial \varphi_L} & \frac{\partial F_\varphi}{\partial n_{e,L}} & \frac{\partial F_\varphi}{\partial n_{i,L}} & \frac{\partial F_\varphi}{\partial T_{e,L}} \\ \frac{\partial F_{n_e}}{\partial \varphi_L} & \frac{\partial F_{n_e}}{\partial n_{e,L}} & \frac{\partial F_{n_e}}{\partial n_{i,L}} & \frac{\partial F_{n_e}}{\partial T_{e,L}} \\ \frac{\partial F_{n_i}}{\partial \varphi_L} & \frac{\partial F_{n_i}}{\partial n_{e,L}} & \frac{\partial F_{n_i}}{\partial n_{i,L}} & \frac{\partial F_{n_i}}{\partial T_{e,L}} \\ \frac{\partial F_{T_e}}{\partial \varphi_L} & \frac{\partial F_{T_e}}{\partial n_{e,L}} & \frac{\partial F_{T_e}}{\partial n_{i,L}} & \frac{\partial F_{T_e}}{\partial T_{e,L}} \end{pmatrix}, \quad (36)$$

where some of the Jacobian matrix elements are zero (such as, $\frac{\partial F_{\varphi}}{\partial T_{e,L}}$, and $\frac{\partial F_{Te}}{\partial n_{i,L}}$). Similarly, the Jacobian matrix Υ_R^k corresponding the right state vector $\mathbf{Q}_R$ can be written down. In the implemented numerical scheme, all Jacobian matrix entries were computed analytically based on the non-thermal SG scheme density and energy flux definitions given above. Analytic computation of the partial derivatives makes the code run faster and more robust, as no numerical differentiation needs to be performed.

The Jacobian matrix was assembled as a sparse matrix using the compressed row storage format. This matrix was allocated at the beginning of each time step based on the current grid state together with the cell-connectivity matrix. The cell-connectivity matrix was computed only once at the beginning of each time step. The elements of the Jacobian matrix were updated at each nonlinear Newton sub-iteration. At the end of the time step both matrices are deallocated. Although these were not very time-consuming operations for the problems solved in the present paper, the matrices' allocation and the computation of cell-connectivity need to be done only after the computational grid has undergone adaption. This approach will be implemented in future work to increase the code's efficiency for larger problems. Such an implementation is straightforward in the current framework.

2.4. Linear matrix solver

At the core of the implemented FNM is a linear matrix solver that solves the resulting system of linear equations:

$$\Upsilon_{L,R}^k \Delta \mathbf{Q}_{L,R}^{k+1} = -\mathbf{F}(\mathbf{Q}^k). \quad (37)$$

The corresponding matrix is of general unsymmetrical, sparse type. As pointed out in reference [43], linear systems arising from the discretization of the semiconductor and plasma transport models can be highly ill-conditioned and, therefore, quite challenging for direct and preconditioned iterative solvers. That work then discusses recent advances in the development of robust direct and iterative sparse linear solvers. It was obtained in reference [43] that for the preconditioned iterative Krylov subspace solvers, nonsymmetric permutations combined with scaling unsymmetrical reorderings gave the best results in terms of the number of required iterations and the time to compute the solution. As it is further discussed in detail in reference [20], there are three main ways to solve this equation: the direct method (LU lower–upper decomposition), fixed-point iteration algorithms (such as Gauss–Seidel and super relax iteration schemes), and Krylov subspace iteration methods (conjugate gradients class, minimum residual class, and others). With the direct methods being computationally costly (matrix bandwidths of ~ 15 – 30 are obtained in our typical TCAD and plasma simulations), one tends to rely on the iteration methods, which in turn must tackle two issues: large matrix condition numbers (due to which the LU algorithm can fail and the number of iterations in Krylov subspace iteration methods drastically increase) and the round-off errors (important to minority carrier densities in semiconductors and sheath regions in plasmas). Algorithms relying on

transport free quasi-minimal residual (TFQMR) and generalized minimal residual methods, as well conjugate gradient squared method (CGS), bi-conjugate gradient, bi-conjugate gradient stabilized were observed to give good results, but they must be combined with ways of minimizing the residual, see reference [20] for details. The current (at the time of writing reference [20]) commercial TCAD software packages, such as Medici integrated the LU and CGS methods, with Dessis adopting the LU and TFQMR methods, while the TCAD code in reference [20] builds upon the LU method (for smaller problems) as well as on the Krylov subspace methods available in PETSc [44]. The modern commercial TCAD software packages have been reported to successfully make use of such matrix solvers as PARDISO [45].

In the present work, we used a high-performance iterative solver [46] available in our NanoTCAD framework. It uses high order preconditioning by incomplete decomposition to ensure good accuracy, reliable stability, and fast convergence. The resulting linear algebraic system is solved using a CGS-type iterative method with preconditioning by incomplete decomposition. To avoid diagonal pivot degeneration, the Kershaw diagonal modification is used, and the apparent computational complexity of the solver estimates as $\sim O(N^{5/4})$, which allows efficient computations for larger matrices arising in multi-dimensional settings. Alternatively, to facilitate the usage of the developed techniques by other research groups, we are working on interfacing and adapting the AMR FNM framework to open-source linear matrix solvers available in the PETSc and Trilinos [47] suites. This will also allow us to have more comprehensive customization over the broad range of available solver types and their tuning parameters to achieve better convergence for each problem of interest, as well as to tackle code's efficient parallelization. The results of these developments will be reported elsewhere.

2.5. Nonlinear convergence and time stepping

In the developed FNM approach, the system of governing equations is advanced from one time to the next, $t \rightarrow t + \Delta t$ (time marching index m in equation (33)), while performing nonlinear sub-iterations for each given time step (index k in equations (16) and (18)). Because of the nature of the transient governing equations, nonlinear convergence within each time step is largely controlled by the selected time step (Δt), which defines the diagonal dominance of the Jacobian matrix. Also, because of the AMR capabilities, the convergence (both local and global) can be efficiently controlled by dynamically adapting the computational grid to resolve the critical plasma features, e.g., large space charge and strong electric field. The time local (within a given time step, index k) convergence of the FNM technique (as any Newton iteration technique) is also determined strongly by initial conditions (set at $k = 0$ in equation (34)). If these initial conditions (at time t) are close to the solution at the end of the sub-iteration cycle (at time $t + \Delta t$), when large time steps can be used. This is typically the case for steady-state [e.g., direct current (DC)] discharges. For transient simulations of RF discharges or gas breakdown simulations, smaller time steps are required. Since the

system of governing equations is strongly nonlinear and depends greatly on the initial conditions, it is not possible to provide a general recipe for selecting the maximum allowed time step Δt *a priori*, contrary to the explicit solvers where the time step can be most often estimated accurately from the courant–friedrichs–lewy (CFL) condition and from the dielectric relaxation time controlling the particle–field coupling. In our simulations with the FNM scheme, the time step was determined by experimenting with the nonlinear convergence rates over each time step (typically, convergence over 4–5 orders of magnitude suffices). Since an implicit time step costs more than an explicit time advance, one can expect that the FNM scheme will be most efficient for steady-state or slowly-varying problems by allowing large time steps. At the same time, transient problems evolving on the (fast) electron-time scale can be more efficiently solved by explicit techniques.

In our typical FNM plasma simulations, low initial plasma density was set up. Once a voltage was applied to one of the electrodes (at time $t = 0$), electron motion and multiplication first developed at the fast electron time scale. The electron motion produced space charge and high electric fields, enhancing further electron multiplication and the development of electron avalanche. During this initial highly transient phase, the time steps Δt were typically set to 0.05–0.1 ns for the problems described below. After completing this stage, the plasma dynamics occur on the slow ion time scale, and the FNM scheme allows increasing the time steps. In our simulations, we have found that it is advantageous to gradually ramp up Δt during the following few 1000's time steps with the final allowed time step being set to the values which ensure good nonlinear convergence (typically, 4–5 orders of magnitude). Typical achieved Δt 's ranged between a few ns for RF plasma problems and 200–500 ns for DC problems. A typical example of a 2D DC plasma in argon gas at 400 mTorr with an inter-electrode gap of 2 cm is shown in figure 2. The cathode on the left is grounded ($\gamma = 0.1$), the voltage of 200 V is applied to the anode (on the right), and the top and bottom boundaries are dielectric walls. AMR was based on magnitudes of space charge and electric field. A fixed number of nonlinear sub-iterations (10) were performed with the total number of time steps of 15 000. During the first 5000 steps, Δt was increased to 200 ns. The time histories of residuals plotted in figure 2 show that for the three main solution variables (electrostatic potential, electron density, and temperature), the local residuals (index k in equations (16) and (18)) drop by a least four orders of magnitude (by almost ten orders of magnitude at the beginning of simulation when the time step is still small), while the global residuals (index m in equation (33)) dropped by almost 12 orders of magnitude thus indicating complete convergence. Such convergence rates (to almost machine precision) may not be required for most of the typical plasma simulations, but they demonstrate superior capabilities of the developed FNM code. Our preliminary results indicate that DC plasma simulations with the developed FNM code are up to a factor of 100 more efficient than those obtained with our explicit code [31]. More detailed comparison studies are in order when fully implicit treatment of finite-rate chemistry

is implemented, which will allow increasing further the efficiency of the new FNM framework for a broad range of plasma problems.

3. Results of simulations

We have applied the new solver to simulations of various DC and RF discharges in 2D and 3D settings. The code is capable of handling 2D-axisymmetric geometries, which are typical for these systems, by adequately setting the cell and face metrics in the electron and ion particle and the energy fluxes across faces of CV cells.

3.1. Direct current glow discharges

Figure 3 shows the results of 2D simulations of a DC discharge in a long cylindrical dielectric tube of radius $R = 1$ cm and the inter-electrode length $d = 7$ cm. The cathode is grounded, and a voltage of 235 V is applied to the anode. The initial plasma density was set to $\sim 10^8 \text{ cm}^{-3}$. The time step for this case was 50 ns.

The results shown in figure 3 illustrate the typical structure of a DC glow discharge containing a cathode region, an axially uniform positive column, and an anode region. The plasma density (a) has a large peak in a negative glow near the cathode, passes through a minimum in the Faraday dark space before reaching an axially constant value in positive column plasma, and decays near the anode. The electric potential distribution (b) connects the equipotential electrodes with a positive column plasma. The radial (ambipolar) electric fields are established to equalize the electron and ion fluxes to the wall. A complicated redistribution of the electric potential takes place in the cathode region, forming a collisional double layer. The ionization rate (c) has a sharp peak in the negative glow and decreases sharply in the Faraday dark space before increasing again in the positive column. The ionization rate decreases near the anode, forming an anode dark space and produces an off-axis peak on the anode surface.

Figure 4 shows distributions of plasma parameters along the discharge axis. The ionization rate has a sharp peak near the boundary of the cathode sheath with plasma; it vanishes in the Faraday dark space, increases again in the positive column plasma, and decreases near the anode forming an anode dark space. The axial electric field changes its sign at two points a and b in the cathode region. It remains constant in the positive column, where the plasma density is axially uniform. The axial electric field's value in the positive column is controlled by plasma to balance the ionization rate and the particle loss to the wall by diffusion and surface recombination. The electric field decreases near the anode and even reverses the sign at the point c near the anode. The electron temperature is high in the cathode sheath, passes through a minimum in the cathode region before reaching a constant value in the positive column, and decreases near the anode. The DC discharge's calculated axial structure has all the features observed in experiments and described in textbooks on gas discharge physics [48].

Figure 5 shows the electron density's radial distributions and the ionization rate in the positive column and at the anode

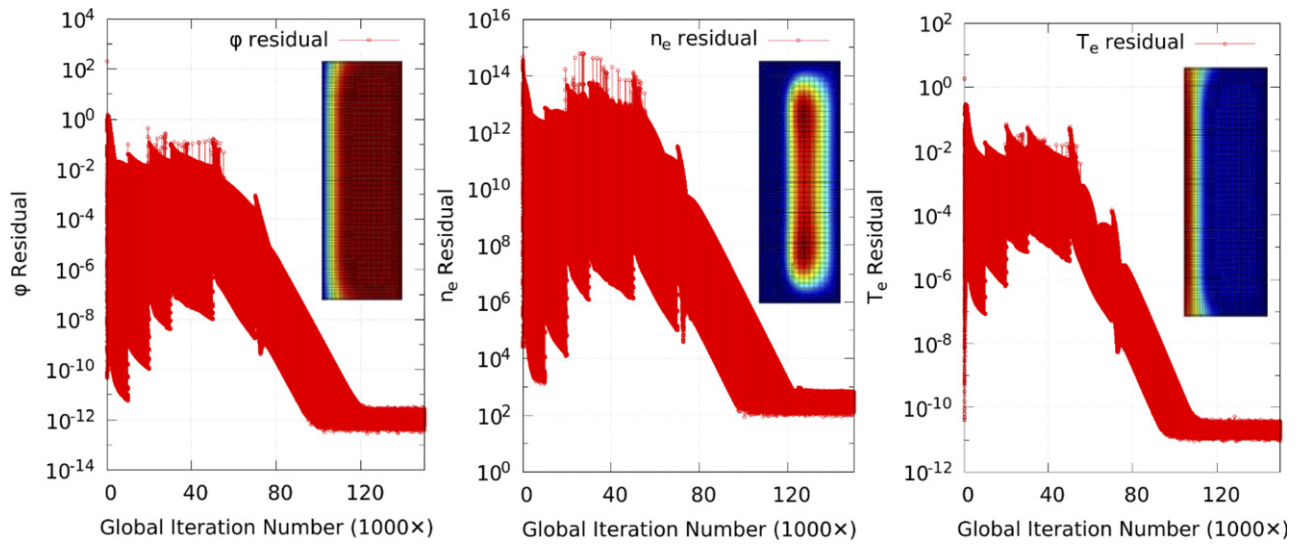


Figure 2. Example of residual evolution obtained with the implemented FNM for short 2D DC discharge cell (Ar, pressure 400 mTorr, gap 2 cm, applied DC voltage 200 V) with AMR and final time step, Δt , of 200 ns. Shown are residuals for electrostatic potential, electron density, and temperature vs global iteration number). The inserts show corresponding plasma profiles with overlaid AMR grid.

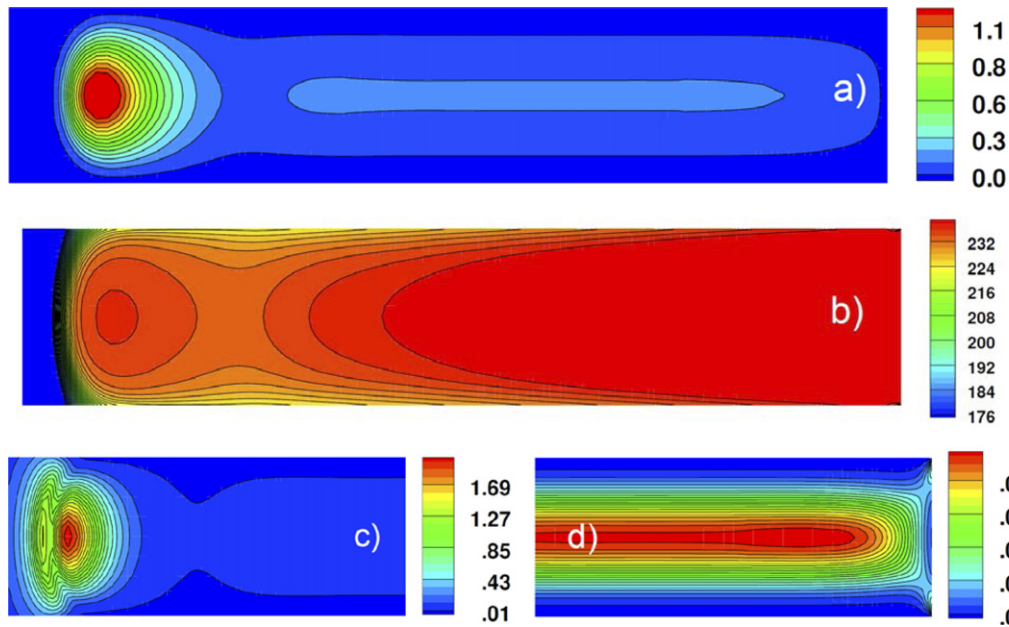


Figure 3. Spatial distributions of electron density (a) in 10^{10} cm^{-3} , electrostatic potential (b) in V, and ionization rate (in $10^{15} \text{ cm s}^{-1}$) near cathode (c) and near anode (d).

surface. In the positive column (a), the radial distributions of the ionization rate and plasma density are similar. At the anode, the radial distribution of the ionization rate forms a ring with a minimum on the axis and sharp peaks near the wall (b). Simultaneously, the radial distribution of the plasma density remains monotonic, as in the positive column. The formation of the anode ring that can break into different spots along the azimuthal direction has been observed in experiments. It is of interest to understand the self-organization of the anode region in DC discharges.

Figure 6 illustrates the structure of the cathode region based on 3D simulations of a DC discharge in a rectangular chamber with dimensions $1 \times 1 \times 3 \text{ cm}$ for gas pressure 0.4 Torr and voltage 300 V. Iso-surfaces of the ion density as well as three slices with contour lines are shown in figure 6. The conditions correspond to a normal discharge with the cathode sheath length of the order of the transverse chamber size R . The phenomenon of the normal current density is one of the key concepts in gas discharge physics [48]. The cathode spot covers only a part of the cathode surface. The current density in the

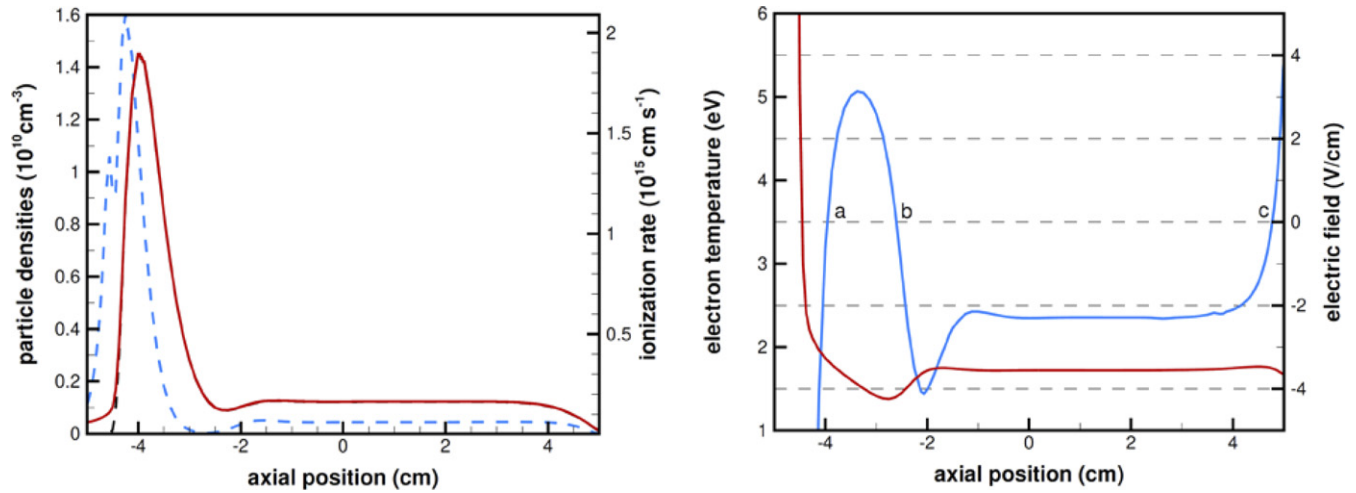


Figure 4. Axial distributions of the electron (dashed) and ion (red line) densities and the ionization rate (dashed blue line) (a), electron temperature (red), and the axial electric field (b) on the axis.

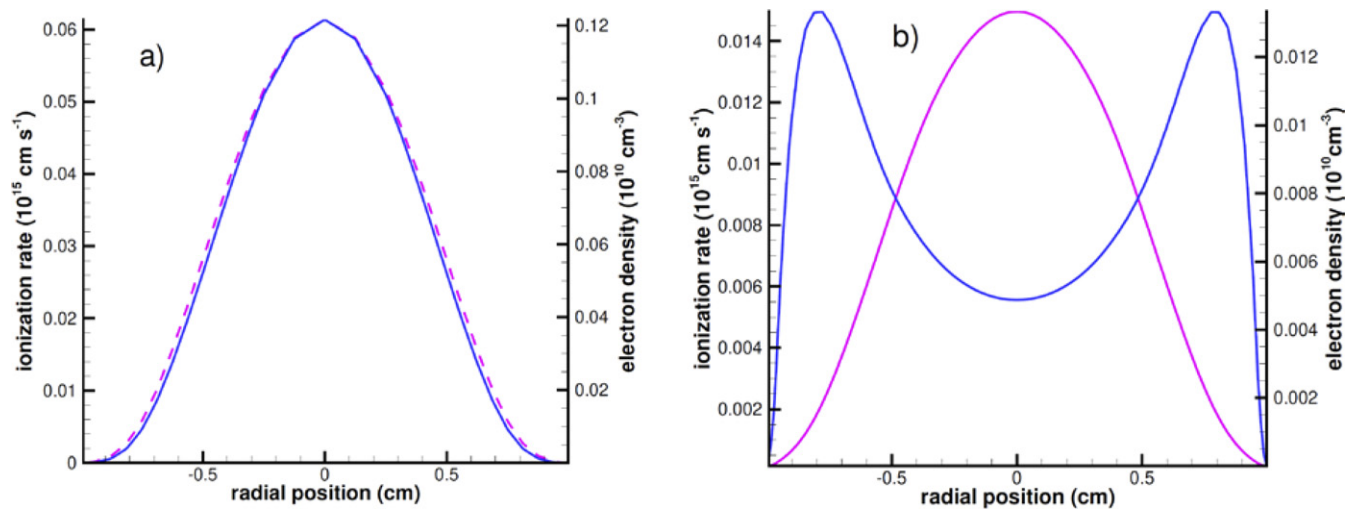


Figure 5. Radial distributions of electron density (dashed pink line) and ionization rate (solid blue line) in the positive column (a), and at the anode (b).

spot remains constant, and the size of the spot increases with the increasing current until the spot covers the entire surface of the cathode. These processes have been previously described in numerous publications [49] and simulated with different codes [37, 50, 51]. It is seen in figure 6 that the transverse distributions of the ion density gradually change from axially symmetric distributions on the cathode surface to a rectangular profile in quasi-neutral plasma.

Figure 7 shows the axial distributions of the electron and ion densities and the electric field along the axis, similar to the 2D simulations in the cylindrical tube discussed above. The axial distribution of plasma density has two maxima: the first at the plasma-sheath boundary and the second close to anode. The electric field reverses sign three points marked *a*, *b*, and *c* in figure 7. The first field reversal (point *a*) occurs close to the plasma-sheath boundary, the second reversal (point *b*) occurs near the minimum of plasma density, and the third field

reversal (point *c*) is close to the point of maximum plasma density. The key difference with the previous 2D simulations is the lower current density in the 3D case and the short inter-electrode gap, which corresponds to the absence of a positive column.

Our fluid model describes the cathode region's typical structure and captures qualitatively all the essential features observed in experiments and previous simulations. The fluid model's success is ensured by incorporating the energy balance of electrons. It is known that the local field approximation (when the ionization rate is assumed to be a function of the local electric field instead of the electron mean energy) gives no electric field reversals in the cathode region [52, 53]. The lack of quantitative description of gas discharges by fluid models is also known and well described in several publications [54, 55].

Experiments and kinetic theory reveal the presence of three distinct electron groups in the cathode region [56]. The fast

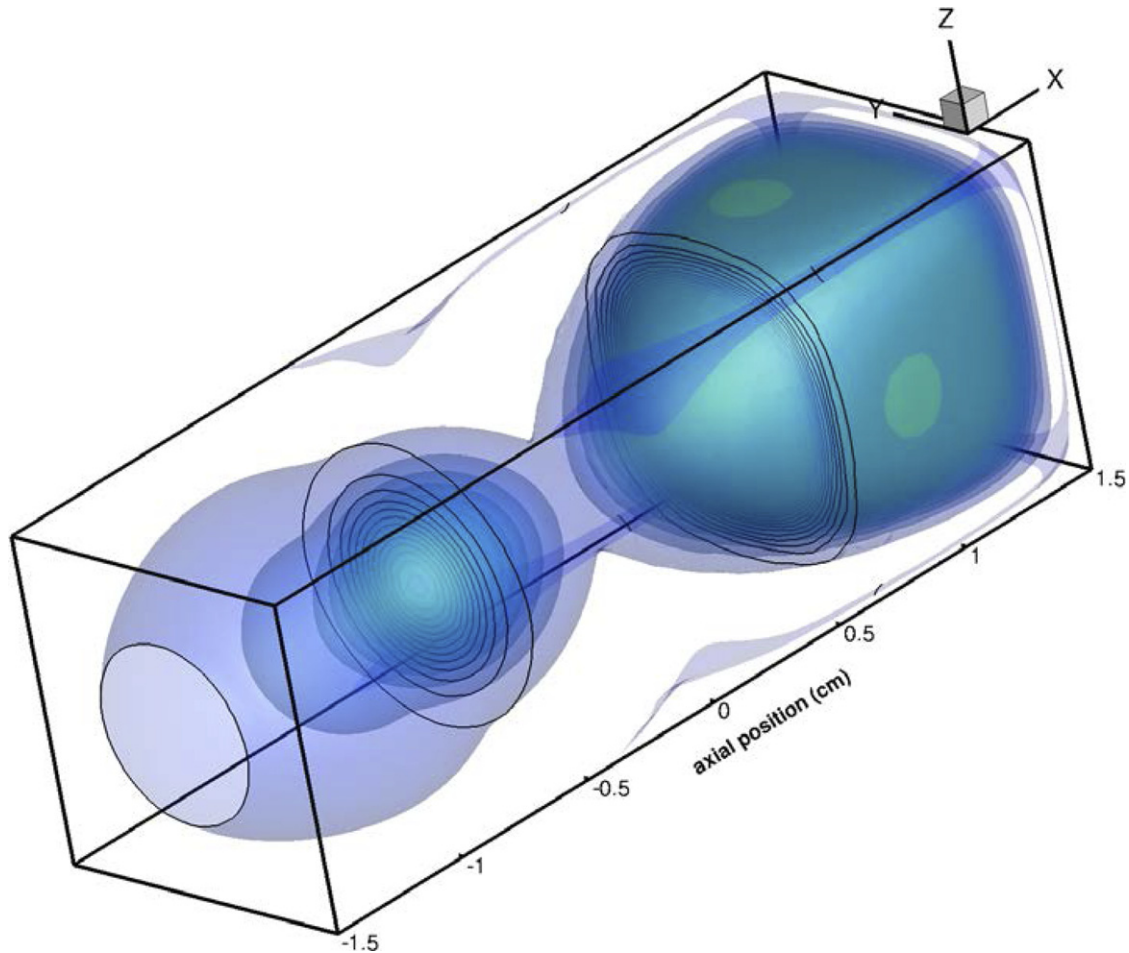


Figure 6. Iso-surfaces of ion density with three slices with contour lines for a DC discharge in a rectangular chamber with dimensions $2 \times 2 \times 6$ cm: Ar, 0.4 Torr, voltage 350 V.

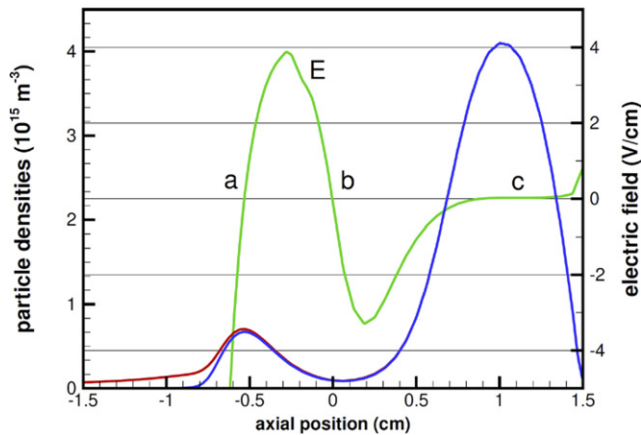


Figure 7. Axial distributions of the electron (blue) and ion (red) densities and the electric field (green) on the axis for the conditions shown in figure 6. The two electric field reversals at points *a* and *b* correspond to the formation of a collisional double layer in the cathode region.

electrons with kinetic energies exceeding the first excitation potential of atoms produce non-local ionization in the negative glow. The fast electrons, which consist of the primary electrons injected from the cathode and those generated in the

cathode sheath, may have an anisotropic velocity distribution function (electron beam). The range of their penetration into plasma corresponds to the length of the negative glow. The second electron group includes slow electrons trapped in the potential well formed in the cathode region. These electrons have near-Maxwellian energy distribution with a temperature that can be close to the gas temperature. They are responsible for the large peak of the plasma density in the cathode region but give zero contribution to the current. The current is transported by intermediate electrons, which occupy the energy range between the slow and fast electrons. The formation of distinct electron groups associated with the double-layer formation in collisional and near-collisionless plasma is typical for gas discharges and space plasmas [57].

Obviously, one cannot expect a quantitative description of physics from a fluid model that describes all the complexity of electron kinetics in terms of an average electron. A recent paper [55] illustrated how the one-dimensional fluid model could be improved to better describe the cathode region's longitudinal structure by incorporating non-local ionization induced by the fast electrons. The addition of the non-local ionization decreases the electron temperature in the negative glow and increases the length of the Faraday dark space between the negative glow and the positive column. In future work, we

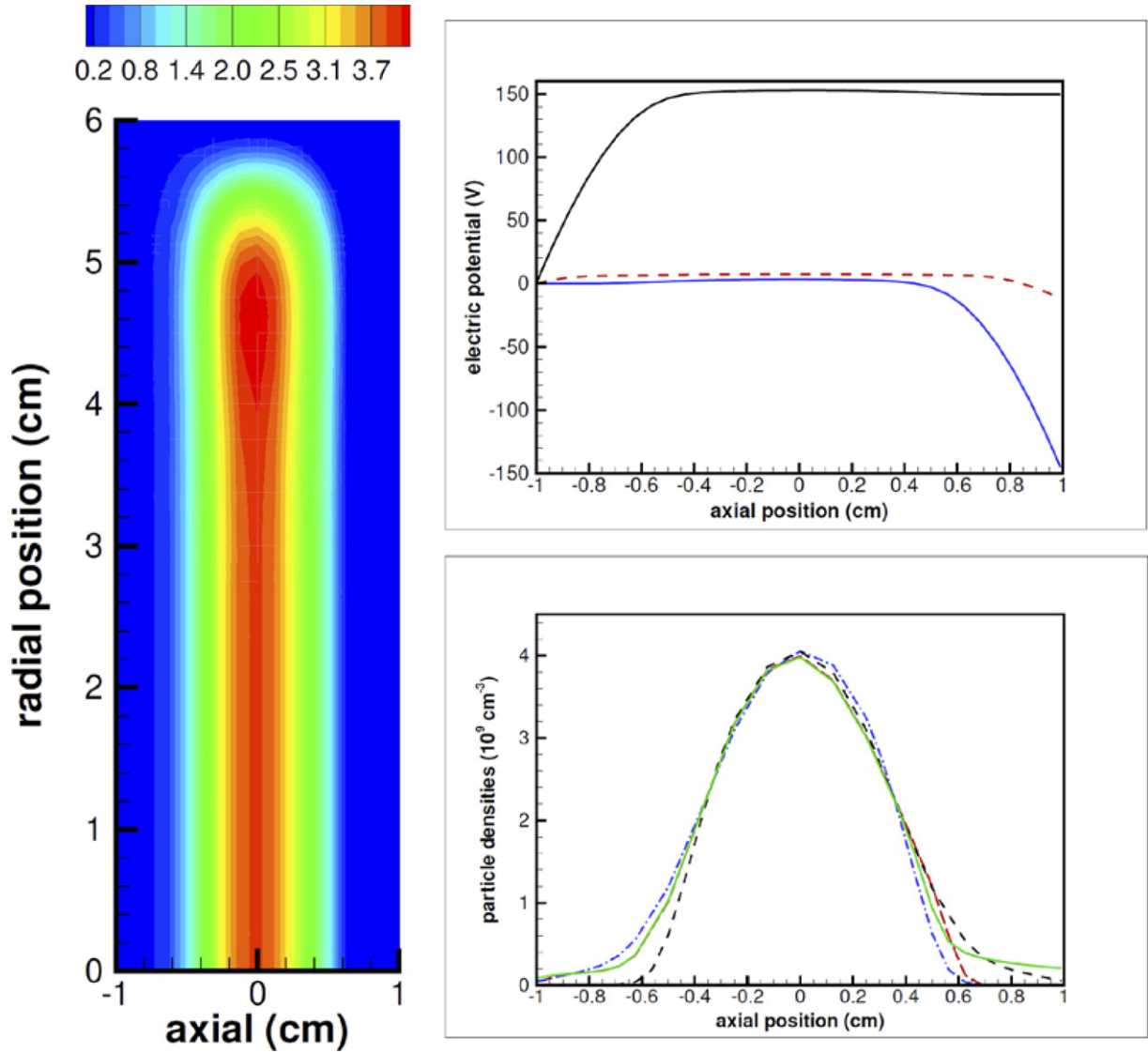


Figure 8. Instantaneous spatial distribution of electron density (left) and axial distributions of the electric potential (top) and electron and ion densities (right) at three times during the RF period.

plan to replace the fluid model for electrons with a grid-based kinetic solver.

3.2. Capacitively coupled plasmas

We have performed 2D axisymmetric simulations of capacitively coupled plasma (CCP) in a dielectric tube of radius $R = 6$ cm, inter-electrode gap $d = 2$ cm, driven by RF voltages at frequencies of 100, 10, and 1 kHz. The secondary emission coefficient was $\gamma = 0.1$. Figure 8 shows spatial distributions of plasma parameters in argon at a pressure of 1 Torr, frequency of 100 kHz, and voltage 150 V. The time step in these simulations was between 2 and 10 ns.

Figure 9 compares the time modulation of electron density and temperature on the axis in the center of discharge (at $x = 0$) for different driving frequencies. At 100 kHz, the plasma density modulation in plasma is negligible, which is consistent with figure 8. With decreasing driving frequency,

the time modulations of electron density in plasma become substantial. The electron temperature is strongly modulated at all these frequencies.

Figure 10 compares the time-variations of electron and ion current densities at the center of the electrode at 100 and 10 kHz. Both electron and ion currents are strongly modulated during the RF period. The electron current's negative values are due to the secondary electron emission from the cathode, which reaches maximal value during the peaks of the ion current. The time modulation of the ion current is specific to low-frequency CCP operating in the so-called dynamics regime [58]. In this regime, the ion transit time through the sheath is comparable to the RF period, and the ion current at the electrode is absent during a substantial part of the RF period. At higher frequencies, the ions respond to the electric field's time-averaged value in the sheath. The ion current at the electrode becomes nearly constant during the RF period, and the electron current has sharp peaks when the sheath voltage has minimal values. The ion current is constant during the RF period [59].

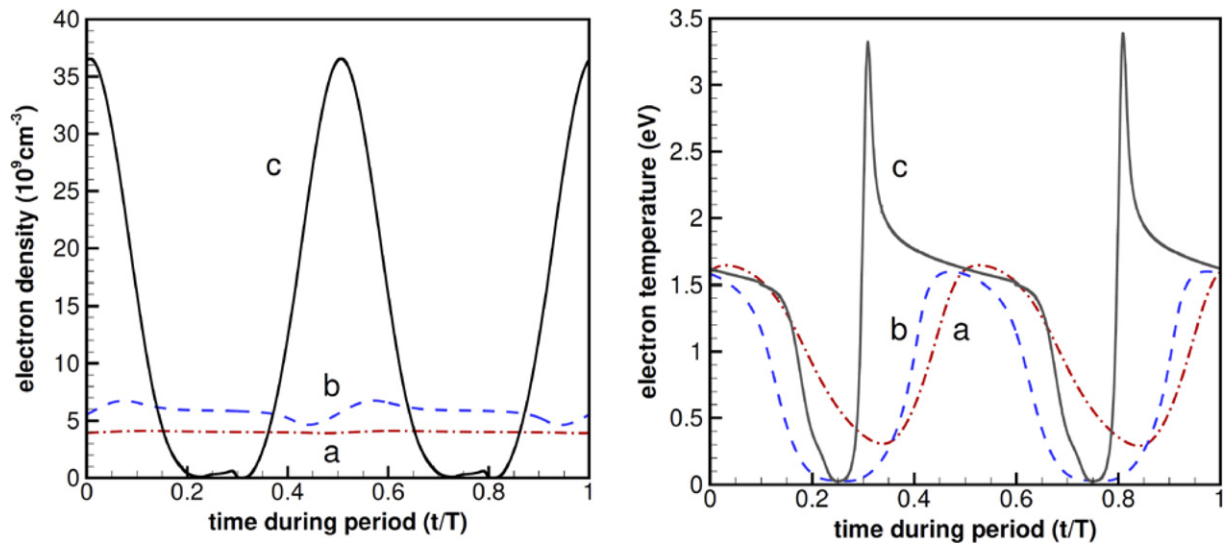


Figure 9. Time modulation of electron density (left) and temperature (right) in plasma at frequencies 100 (a), 10 (b) and 1 (c) kHz.

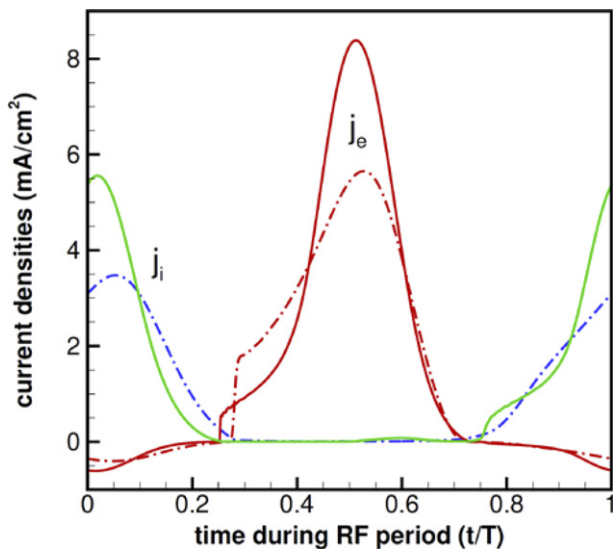


Figure 10. Time modulation of the electron (red lines) and ion current densities (blue and green lines) at electrode for driving frequencies of 100 (dash-dot lines) and 10 (solid lines) kHz.

3.3. Plasma stratification in argon discharges

Plasma stratification (i.e., pattern formation along the direction of discharge current) often occurs in DC and RF discharges over a wide range of gas pressures and discharge currents [60, 61]. In DC discharges of noble gases, striations (ionization waves) usually move along the electric field direction, while standing striations are typical for molecular gases. We have applied the FNM plasma solver for simulations of striations in argon discharges and discovered the common mechanism of stratification in DC and RF discharges at relatively high currents (plasma densities). The nonlinear dependence of the ionization rate on electron density is the leading cause of plasma stratification. The present paper provides details of simulations and describes some new results compared to those briefly reported in two letters [40, 41].

3.3.1. Moving striations in DC discharges. Simulations were performed for a cylindrical dielectric tube of radius $R = 1$ cm and length $L = 14$ cm. The anode was grounded, and a voltage was applied to the cathode. The electric potential at the dielectric walls was calculated from the local surface charge, which evolved in time based on electron and ion fluxes. During our simulations of the discharge development, striations first originate in the cathode region and gradually propagate towards the anode. However, the striations themselves move towards the cathode that corresponds to the backward waves (with the group and phase velocities in opposite directions), which are well known from the analytical theory of small-amplitude striations under these conditions. The corresponding DC intermediate movie (<https://stacks.iop.org/PSST/30/045013/mmedia>) illustrating the dynamics of discharge stratification and mesh adaptation is available on the journal web site.

Figure 11 shows instantaneous spatial distributions of electron density in diffuse and constricted discharges at two gas pressures (1 Torr and 400 Torr) in argon. The radius of the plasma column in the diffuse discharge is controlled by surface recombination, and the striation wavelength is about the radius of the discharge tube. In the constricted discharge, the radius of the plasma column is controlled by volume recombination and is smaller than R . The striation wavelength is smaller, and the plasma radius changes over striation wavelength during the wave propagation. This agrees well with the reported experimental observations of the striations in constricted discharges. The bottom part of figure 11 shows a dynamically adapted mesh for the constricted discharge. The mesh adaptation criterion is based on the gradients of the electrostatic potential and electron density. A DC constricted AMR movie available on the journal web site illustrates the dynamics of plasma stratification in DC discharges.

Figure 12 shows distributions of the electric field, electron density, and temperature on the axis of the diffuse (left) and constricted (right) discharges. In the diffuse discharge, the

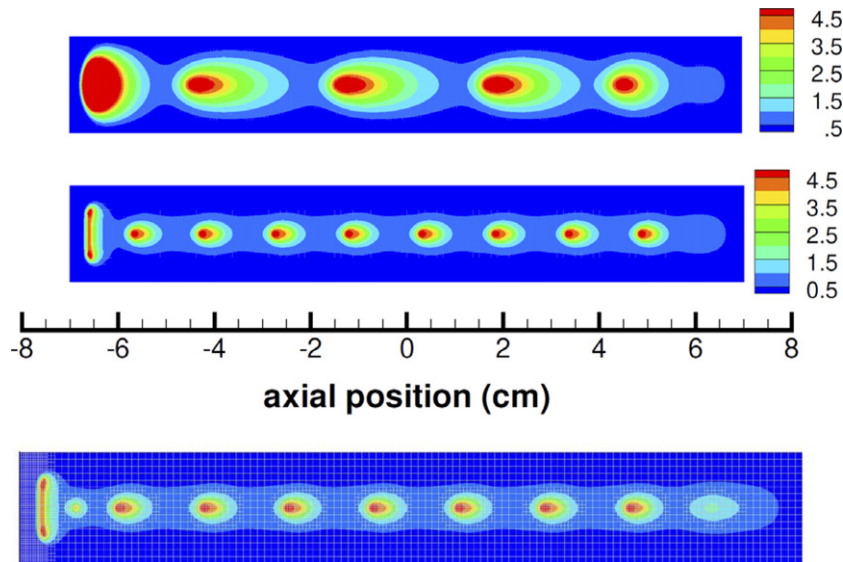


Figure 11. Instantaneous spatial distributions of electron density (10^{10} cm^{-3}) for the diffuse discharge (top) and constricted discharge (middle). The bottom part shows the dynamically adapted mesh for the constricted discharge (argon 400 Torr).

maximal value of the electric field is observed near the maximal gradient of plasma density, which corresponds to the dominance of the ambipolar component of the electric field over the conduction component. The maximal value of the electron temperature is shifted towards the cathode compared to plasma density, which corresponds to the propagation of striations towards the cathode. Electric field reversals take place in the constricted discharge, which indicates that striations are substantially nonlinear in the constricted positive column. In diffuse discharge, electric field reversals are negligible for these conditions.

In the constricted discharge, the maximal value of the electric field is observed again near the point of maximal gradient of plasma density. This corresponds to the dominance of the ambipolar electric field over the conduction field. The field is close to zero at the points of maximal plasma density on the axis. That is in good agreement with the two-dimensional theory of striations in constricted discharges. The maximal values of the electron temperature are shifted towards the cathode compared to the maximums of plasma density, which corresponds to the propagation of striations towards the cathode. The electric field changes sign between the maximum plasma density and the maximum electron temperature. This corresponds to highly nonlinear waves under these conditions.

Figure 13 shows the radial distributions of electron densities in two phases of waves, which correspond to the maximal and minimal density on the axis. In the diffuse discharge (left), the radial distribution of plasma densities changes weakly. In the constricted discharge (right), the radial distribution of plasma density changes substantially over the striation wavelength. The radius of plasma has a minimum at the point of maximal plasma density on the axis, which is in good agreement with the two-dimensional theory of striations in constricted discharges.

3.3.2. Standing striations in RF discharges. The key difference between moving striations in DC discharges and standing striations in RF discharges is the absence of the RF electric field's time-average component in the striation-free positive column. In our experiments and simulations, we have observed that any discharge asymmetry and the appearance of even a small DC component of the electric field in CCP results in a movement of striations along the axis. Figure 14 shows an example of 2D axisymmetric simulations of CCP for tube radius $R = 1 \text{ cm}$, inter-electrode lengths $L = 14 \text{ cm}$ (a) and 20 cm (b), gas pressure 0.5 Torr, driving frequency of 20 MHz, and $n_e = 4 \times 10^{15} \text{ m}^{-3}$. The striation wavelength varies discontinuously with changing the inter-electrode distance so that an integer number of standing waves always forms between the electrodes. Six striations are observed in our simulations for $L = 14 \text{ cm}$, and eight striations are formed for $L = 20 \text{ cm}$. It is seen that the two central striations at $L = 20 \text{ cm}$ are slightly longer than the rest of the waves. They will break into an additional pair with a further increase of the inter-electrode distance L .

Figure 15 shows the time variations of electron density in the middle of the gap (at $x = 0$) and close to the electrode (at $x = L/2$) during the plasma stratification process. It is seen that standing striations are formed during a few thousands of RF cycles. The time of striation formation is also large compared to the particle ambipolar diffusion time to the wall. A CCP-movie available on the journal web site illustrates the transient process of striation formation. They propagate from electrodes towards the discharge center. In fully striated plasma, despite substantial time variations of the power deposition, the electron density, and the electron temperature do not vary in time under these conditions, except within the oscillating sheaths near the electrodes. The developed code helped clarify the plasma stratification mechanism and study the nonlinear stage often realized in the experiments [41].

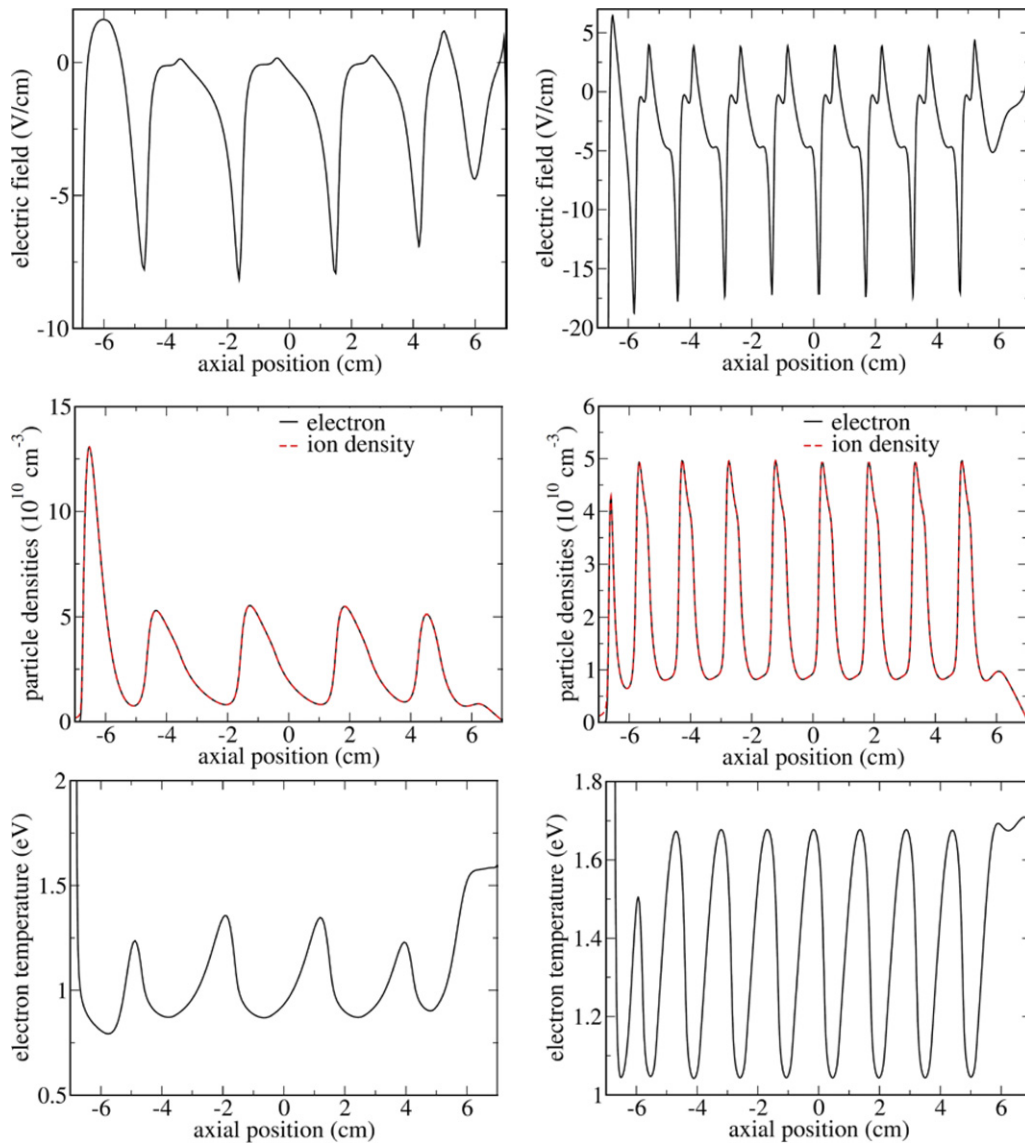


Figure 12. The axial distributions of the normalized electric field, electron density, and temperature in the diffuse (left) and constricted (right) discharges.

3.4. Comparison of explicit and FNM solvers

Detailed benchmarking of the new FNM solver with previously developed codes is not the goal of the present work and could be a subject of future studies. In this section, we will only provide some comments and estimates to compare the accuracy and efficiency of the previously developed explicit and the new FNM code. The programming language for both codes is C with inheritance [32]. Both the explicit and FNM codes were run on Intel Xeon CPU-E5-2680 v4 2.40 GHz processors in serial mode. However, the numerical schemes in the explicit code [31] and the new FNM code are quite different. In particular, the explicit code used the semi-implicit Poisson solver, fully explicit isothermal SG scheme for the particle transport, and a scheme with no special treatment of the Joule heating term and boundary conditions. Hence the results of the two codes can be expected to be somewhat different even on the same computation grids. Direct benchmarking of the explicit and FNM codes would require the implementation of

the same spatial and temporal discretization schemes, with the same treatment of different terms of the governing equations.

However, some estimates can still be made using the two codes in terms of the memory footprint and timings for comparable accuracy of the solution. We have selected two 2D axisymmetric, steady-state problems for comparative studies. One problem is a short DC plasma case (400 mTorr of Ar gas, gap 2 cm, applied voltage 200 V) and the other problem is a long DC plasma case (1 Torr of Ar gas, inter-electrode distance 5 cm, tube radius 1 cm, applied voltage 100 V). The number of solved species was 5 in both cases: electrons, ground-state Ar species (allocated but not solved for), and (for generality) three types of ions, of which only the Ar^+ ion played a role under simulated discharge conditions. The short DC plasma case has a distinctive feature of stronger electric fields between the electrodes, and thus shorter CFL time steps allowed in the explicit code. The second case is characterized by larger plasma formation times due to the presence of a long positive column region controlled by the ambipolar diffusion.

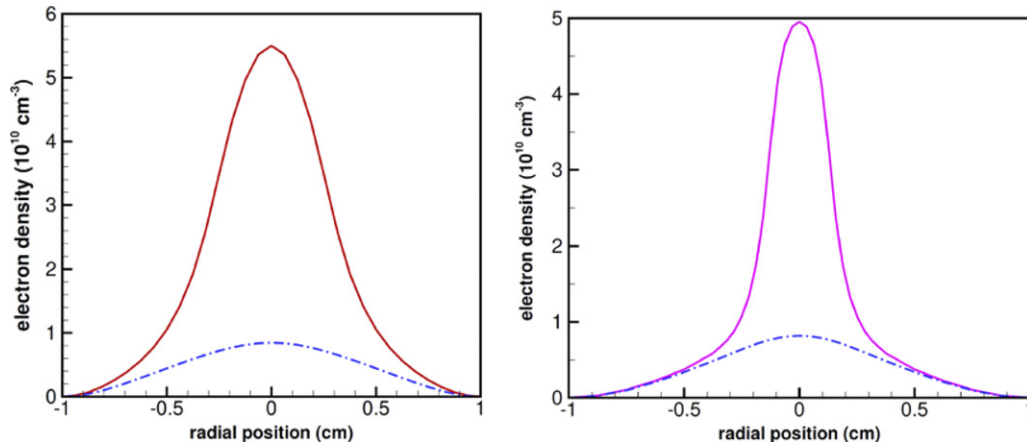


Figure 13. Radial distributions of electron densities in two phases of striation corresponding to the maximal and minimal density on the axis in the diffuse (left) and constricted (right) plasma.

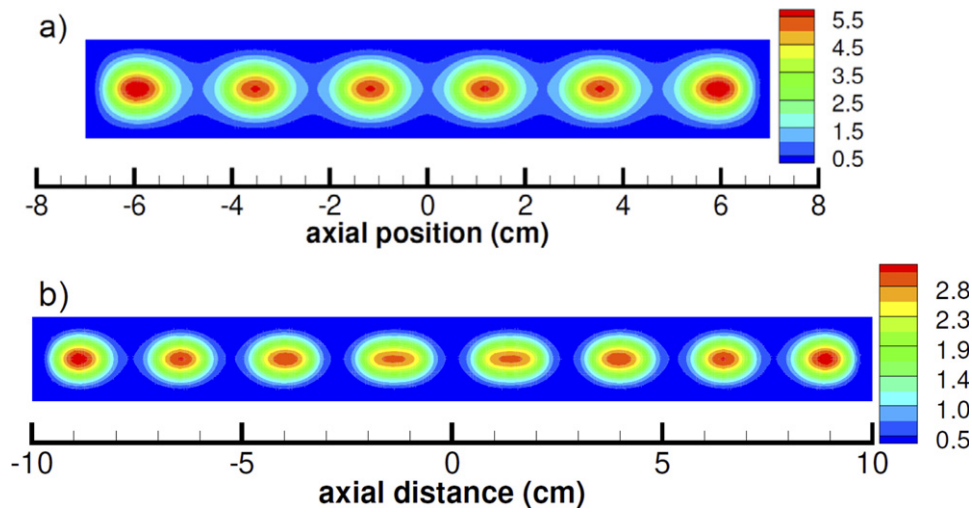


Figure 14. Spatial distributions of the electron density (10^{10} cm^{-3}) obtained in 2D-axisymmetric simulations of CCP at 20 MHz for $L = 14 \text{ cm}$ (a) and $L = 20 \text{ cm}$ (b).

For each setup, we have selected the same AMR grids (with cell clustering around the cathode, anode, and dielectric walls regions) to be used in the explicit and FNM codes. The number of cells was close to 2 K in both cases. RAM required to run the explicit code was $\sim 350 \text{ MB}$ while the FNM code required $\sim 400 \text{ MB}$ ($\sim 0.3\text{--}0.4 \text{ MB}$ per cell). The extra RAM of $\sim 50 \text{ MB}$ required to run the FNM code could be attributed to the Jacobian matrix storage, as well to the linear matrix solver storage (which was found to scale exceptionally well with the number of degrees of freedom, see references in the paper). The explicit code relied on the semi-implicit implementation of the Poisson solver [31] while Poisson equation itself was solved by the multigrid technique [32] with the convergence of 3–4 orders of magnitude (typically, in 4–5 iterations) thus ensuring implicitness of this solver. The results showed that 1000 time steps for the explicit code took $\sim 25 \text{ s}$ of CPU time, while 1000 steps of the FNM code took $\sim 500 \text{ s}$, thus the explicit code being ~ 20 faster per time step than the

FNM code. The physical time steps for the explicit code were about 10–50 ps at the initial stage of plasma evolution and later dropped to several ps after the narrow plasma sheaths were formed.

On the contrary, the physical time step in the FNM code started with $\sim 100 \text{ ps}$ and could be gradually ramped to $0.1 \mu\text{s}$ (sometimes even larger, up to $1 \mu\text{s}$) during the first several 1000's steps, which then remained at this level until the end of the simulation. We have run these simulations to a physical time of 1 ms, which is a typical time scale for the plasma to establish due to chemical reactions and ambipolar diffusion processes. For the explicit code with a time step of $\sim 10 \text{ ps}$ to reach out to such times, it requires $\sim 100 \text{ M}$ ($1 \text{ ms}/10 \text{ ps}$) time steps. For the FNM code with the time step of $0.1 \mu\text{s}$, it requires only $\sim 10 \text{ K}$ time steps. While the FNM code was found to be ~ 20 times more expensive, because of the drastic difference ($\sim 10\,000\times$) of the time steps required, the final simulation could be obtained ~ 500 times faster in terms of the

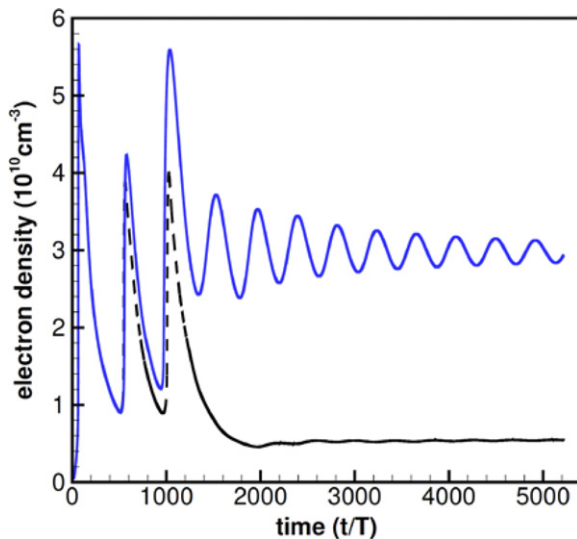


Figure 15. Oscillations of plasma density on the axis at $x = 0$ (dashed black line) and $x = L/4$ (solid blue line) obtained in 2D axisymmetric simulations of a 20 MHz CCP with $L = 14$ cm.

CPU time. Since the plasma can be established faster than 1 ms (say, 0.2 ms), the resulting (conservative) speed up estimation is 100X.

It is worth noting that our explicit code has the semi-implicit Poisson solver thereby avoiding the strict dielectric relaxation time step limitation (which can be a fraction of ps at the studied plasma densities); hence, the explicit time step was only limited by the CFL condition. The FNM solver relied on the explicit chemistry implementation, which was likely to be the main cause of its observed time step limitation. More accurate benchmarking of the explicit and FNM codes could involve a larger set of cases (e.g., with a larger number of cells and plasma species, as well as time transient cases, such as RF plasmas). Such benchmarking could be a subject of separate studies. However, even these preliminary results have already shown the great advantages of developed FNM technology. This technology is currently being further advanced via fast–slow modular splitting and chemistry implicitness for efficient plasma simulations.

4. Conclusions and outlook

We have described an initial implementation of a multi-fluid, multi-temperature plasma solver with ACM using a full-Newton (coupled, implicit) scheme. The new solver provided breakthrough capabilities for solving several gas discharge problems that we could not solve before with existing software. With the new solver, we could calculate the two-dimensional structure of the entire DC discharges, including cathode and anode regions, electric field reversals and the formation of a double layer, a normal cathode spot, and an anode ring, for realistic values of the secondary electron emission coefficient at the cathode and high discharge currents. We could also simulate moving striations in diffuse and constricted DC discharges, and standing striations in RF discharges in argon gas. We have observed good convergence rates and

proper dynamic grid adaptation for resolving time-variations of space-charge sheaths in the CCP at the electron scale and the slow plasma dynamics at the ion time scale. The fully implicit treatment of the coupled plasma equations allowed using large time steps (speedup factors up to 100 compared to explicit solvers). The full-Newton treatment enabled fast nonlinear convergence at each time step, offering significantly improved fluid plasma simulations.

In the present paper, we considered the plasma boundaries aligned with the grid lines. In future work, we will treat non-aligned boundaries using the cut-cell technique, as previously done in our explicit plasma solver. Simulating free surfaces of the complex shape appears very attractive for modeling discharges with liquid electrodes [11] using the VoF method with AMR. The developed FNM-ACM technique can efficiently treat such problems. The FNM-ACM solver's parallelization can be done efficiently using the Forest-of-Trees or space-filling curve algorithms. The work in these directions is currently underway and will be reported elsewhere.

Adding a general-purpose chemistry solver is another subject of our future work. The developed solver is ready to be linked with finite-rate chemistry solvers using multi-temperature reaction rates (for electrons, gas species, and vibrational states of molecules). The nonlinear source and loss terms for charged and neutral species can be treated using several techniques, such as a point-implicit approach [62], high-order Runge–Kutta techniques with adaptive time-stepping, as well as explicit techniques. The plan is to treat the species transport and finite-rate chemistry in a fully coupled way, as commonly used in CFD for modeling combustion and flames with complex and often stiff chemistries. The coupling of the charged particle motion and the electron energy transport with self-consistent electric fields makes plasma processes highly nonlinear and multi-scale problems. The developed FNM-ACM technique offers many benefits for tackling plasma systems' time scales' disparity and nonlinearity.

We wish to point out that the implemented code has not been optimized for maximal performance. However, the code runs fast and allows large time steps for stability, with good convergence observed for a broad range of grid resolution variation. The time steps employed in the present paper can be further increased by implicit (or semi-implicit) treatment of the finite-rate chemistry, as was discussed earlier. Among different optimization techniques in the future is an automatic time-step-control algorithm [18], which ensures maximum allowable time steps during runs. Another method is matrix bandwidth reduction, which can further speed up the code. Such a technique is easily realizable with ACM by reordering the cells for decreasing the adjacency matrix. Another advantage of the FNM is that it efficiently treats the cut cells that often invoke the small-cell problem, limiting the allowed time step in explicit algorithms and requiring rather complicated (in terms of bookkeeping) cell-merging techniques [32]. The FNM technique avoids this problem altogether by including the cut cells into a single computational matrix. The use of implicit (thus, no CFL limit) but uncoupled schemes can result in severe convergence problems, especially in small cut cells with volumes much smaller than off-body cells. Thus, the full

advantages of the developed FNM-ACM techniques will be realized for complex plasma interface simulations.

For our fluid modeling of collisional gas discharge plasmas, we have adapted the numerical techniques commonly used in semiconductor modeling with the fully coupled solution for electrons and holes in a single (global) matrix. For plasma systems, in general, this approach may not be necessary, as the electron and ion mobilities in plasmas differ by factors of 100–1000. In contrast, the charge carrier mobilities in semiconductor systems may be of the same order of magnitude. Thus, ion transport can be solved separately for the plasma systems, which can significantly reduce the computational cost, especially for modeling of more exotic plasma systems such as ion–ion plasmas, strongly magnetized or ‘liquid’ plasmas. We plan to combine the developed FNM approach with the time-scale separation methodology to treat electrons and heavy (ion and neutral) species by loosely dependent modules. The loosely coupled approach would also allow using hybrid kinetic-fluid models for electrons while fluid models treat ions. Such a system would facilitate simulations of plasmas with large numbers of heavy species and complicated chemical mechanisms.

Acknowledgments

This work was partially supported by the NSF project OIA-1655280 and DOE project DE-SC0021391.

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

ORCID iDs

Vladimir I Kolobov  <https://orcid.org/0000-0002-6197-3258>

References

- [1] Rambo P W and Denavit J 1992 *J. Comput. Phys.* **98** 317
- [2] Alves L L, Bogaerts A, Guerra V and Turner M M 2018 *Plasma Sources Sci. Technol.* **27** 023002
- [3] Kolobov V I 2003 *Comput. Mater. Sci.* **28** 302
- [4] Hagelaar G J M and Kroesen G M W 2000 *J. Comput. Phys.* **159** 1
- [5] Trelles J P 2018 *Plasma Sources Sci. Technol.* **27** 093001
- [6] Alvarez Laguna A, Lani A, Deconinck H, Mansour N N and Poedts S 2016 *J. Comput. Phys.* **318** 252
- [7] Markidis S, Lapenta G and Rizwan-uddin 2010 *Math. Comput. Simul.* **80** 1509
- [8] Palmroth M *et al* 2018 *Living Rev. Comput. Astrophys.* **4** 1
- [9] Kawamura E, Birdsall C K and Vahedi V 2000 *Plasma Sources Sci. Technol.* **9** 413
- [10] Kolobov V and Deluzet F 2019 *Front. Phys.* **7**
- [11] Levko D, Arslanbekov R R and Kolobov V I 2020 *J. Appl. Phys.* **127** 043301
- [12] Arslanbekov R and Kolobov V 2018 *J. Phys.: Conf. Ser.* **1031** 012018
- [13] Fořt J, Karel J, Trdlička D, Benkhaldoun F, Kissami I, Montavon J-B, Hassouni K and Mezei J Z 2019 *Adv. Comput. Math.* **45** 2163
- [14] Tang T W 1984 *IEEE Trans. Electron Devices* **31** 1912
- [15] McAndrew C C, Singhal K and Heasell E L 1985 *IEEE Electron Device Lett.* **6** 446
- [16] Forghieri A, Guerrieri R, Ciampolini P, Gnudi A, Rudan M and Baccarani G 1988 *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.* **7** 231
- [17] Kato A, Katada M, Kamiya T, Ito T and Hattori T 1994 *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.* **13** 1425
- [18] Choi W-S, Ahn J-G and Park Y-J 1994 *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.* **13** 899
- [19] Simlinger T 1996 Simulation von Heterostruktur-Feldeffekttransistoren *PhD Thesis* Technischen Universität Wien <http://iue.tuwien.ac.at/phd/simlinger/diss.html>
- [20] Ding G 2008 Open source two-dimensional general-purpose semiconductor simulator GSS user's guide http://gss-tcad.sourceforge.net/gss Ug_en.pdf
- [21] Turowski M, Raman A and Schrimpf R D 2004 *IEEE Trans. Nucl. Sci.* **51** 3166
- [22] Raman A and Turowski M 2021 *Mixed-Mode TCAD Tools* Eds John D Cressler and H Alan Mantooth (Boca Raton, FL: CRC Press) Ch 31 pp 345–58
- [23] Wright M J, Candler G V and Bose D 1998 *AIAA J.* **36** 1603
- [24] Gnoffo P A 1990 An upwind-biased, point-implicit relaxation algorithm for viscous, compressible perfect-gas flows *NASA Technical Report*
- [25] Kirk B, Stogner R, Oliver T A and Bauman P T 2013 Recent advancements in fully implicit numerical methods for hypersonic reacting flows *21st AIAA Computational Fluid Dynamics Conf.* (American Institute of Aeronautics and Astronautics)
- [26] Luke E A, Tong X-L, Wu J, Tang L and Cinnella P 2003 A chemically reacting flow solver for generalized grids *AIAA Conf. Proc.*
- [27] Stone J M, Tomida K, White C J and Felker K G 2020 *Astrophys. J. Suppl. Ser.* **249** 4
- [28] Converge Software <https://convergecd.com/benefits/autonomous-meshing>
- [29] Simerics Software <https://simerics.com/mp-meshing/>
- [30] FloEFD Products <https://mentor.com/products/mechanical/floefd/>
- [31] Kolobov V I and Arslanbekov R R 2012 *J. Comput. Phys.* **231** 839
- [32] Popinet S 2003 *J. Comput. Phys.* **190** 572
- [33] Arslanbekov R R, Kolobov V I and Frolova A A Analysis of compressible viscous flow solvers with adaptive Cartesian mesh *AIAA-2011-3381* <https://doi.org/10.2514/6.2011-3381>
- [34] COMSOL Multiphysics <https://doc.comsol.com/5.6/docserver/#!/com.comsol.help.com>
- [35] Lindsay A D, Graves D B and Shannon S C 2016 *J. Phys. D: Appl. Phys.* **49** 235204
- [36] Balay S *et al* 2019 *PETSc* <https://mcs.anl.gov/petsc>
- [37] Rafatov I, Bogdanov E A and Kudryavtsev A A 2012 *Phys. Plasmas* **19** 033502
- [38] Tejero-del-Caz A, Guerra V, Gonçalves D, da Silva M L, Marques L, Pinhão N, Pintassilgo C D and Alves L L 2019 *Plasma Sources Sci. Technol.* **28** 043001
- [39] Arslanbekov R R and Kolobov V I 2005 *IEEE Trans. Plasma Sci.* **33** 354–5
- [40] Arslanbekov R R and Kolobov V I 2019 *Phys. Plasmas* **26** 104501

- [41] Kolobov V I, Arslanbekov R R, Levko D and Godyak V A 2020 *J. Phys. D: Appl. Phys.* **53** 25LT01
- [42] Nikandrov D S, Arslanbekov R R and Kolobov V I 2008 *IEEE Trans. Plasma Sci.* **36** 932
- [43] Schenk O, Hagemann M and Rollin S 2003 Recent advances in sparse linear solver technology for semiconductor device simulation matrices *Int. Conf. on Simulation of Semiconductor Processes and Devices, 2003. SISPAD 2003* (Boston, MA, USA) 103–8
- [44] PETSc scalable nonlinear equations solvers (SNES) component <https://mcs.anl.gov/petsc/petsc-current/docs/manualpages/SNES/index.html>
- [45] Schenk O 2000 *Scalable Parallel Sparse LU Factorization Methods on Shared Memory Multiprocessors (Series in Microelectronics vol 89)* (Konstanz: Hartung Gorre)
- [46] Fedoseyev A I, Turowski M, Raman A, Alles M L and Weller R A 2009 *Int. J. Multiscale Comput. Eng.* **7** 9
- [47] Trilinos https://trilinos.github.io/linear_solver.html
- [48] Raizer Y P 1991 *Gas Discharge Physics* (Berlin: Springer)
- [49] Kolobov V I and Fiala A 1994 *Phys. Rev. E* **50** 3018
- [50] Arslanbekov R R and Kolobov V I 2003 *J. Phys. D: Appl. Phys.* **36** 2986
- [51] Surzhikov S T, Kozlov P V, Kotov M A, Ruleva L B and Solodovnikov S I 2019 *Dokl. Phys.* **64** 154
- [52] Kolobov V I and Tsendin L D 1989 *Sov. Phys. Tech. Phys.* **34** 1239
- [53] Grubert G, Becker M and Loffhagen D 2009 *Phys. Rev. E* **80** 036405
- [54] Derzsi A, Hartmann P, Korolov I, Karácsny J, Bánó G and Donkó Z 2009 *J. Phys. D: Appl. Phys.* **42** 225204
- [55] Yao J, Yuan C, Eliseev S, Kudryavtsev A and Zhou Z 2020 *Plasma Sources Sci. Technol.* **29** 075003
- [56] Kolobov V I and Tsendin L D 1992 *Phys. Rev. A* **46** 7837
- [57] Kolobov V, Arslanbekov R and Levko D 2020 *J. Phys.: Conf. Ser.* **1623** 012006
- [58] Nikandrov D S and Tsendin L D 2006 *Tech. Phys. Lett.* **32** 719
- [59] Raizer Y P, Shneider M N and Yatsenko N A 1995 *Radio-Frequency Capacitive Discharges* (Boca Raton, FL: CRC Press)
- [60] Kolobov V I 2017 Glow discharges: stratification *Encyclopedia of Plasma Technology* (London: Taylor and Francis)
- [61] Nakata J, Takenaka E and Masutani T 1965 *J. Phys. Soc. Japan* **20** 1698
- [62] Sekhar S K, Zaki M, Ruffin S, Kolobov V and Arslanbekov R 2011 Evaluation of viscous flow solvers with adaptive Cartesian meshes for hypersonic flows *42nd AIAA Thermophysics Conf.* <https://doi.org/10.2514/6.2011-4035>