APPLICATION OF GPU-accelerated FDTD method to electromagnetic wave propagation in plasma using MATLAB Parallel Processing Toolbox

A Preprint

Shayan Dodge
Laser and Plasma Research Institute
Shahid Beheshti University
Iran
dodgeshayan@gmail.com

Mojtaba Shafiee
Laser and Plasma Research Institute
Shahid Beheshti University
Iran
m-shafiee@sbu.ac.ir

Babak Shokri
Laser and Plasma Research Institute
Shahid Beheshti University
Iran
b-shokri@sbu.ac.ir

November 11, 2022

ABSTRACT

Since numerical computing with MATLAB offers a wide variety of advantages, such as easier developing and debugging of computational codes rather than lower-level languages, the popularity of this tool is significantly increased in the past decade. However, MATLAB is slower than other languages. Moreover, utilizing MATLAB parallel computing toolbox on the Graphics Processing Unit (GPU) face some limitations. The lack of attention to these limitations reduces the program execution speed. Even sometimes, parallel GPU codes are slower than serial. In this paper, some techniques in using MATLAB parallel computing toolbox are studied to improve the performance of solving complex electromagnetic problems by the Finite Difference Time Domain (FDTD) method. Implementing these techniques allows the GPU-Accelerated Parallel FDTD code to execute 20x faster than (basic) serial FDTD code. Eventually, GPU-Accelerated Parallel FDTD code is utilized to optimize the computational modeling of electromagnetic waves propagating in plasma. In this simulation, kinetic theory equations for plasma are used (excluding inelastic collisions), and temporal evolution is studied by the FDTD method (coupled FDTD with kinetic theory).

Keywords FDTD · GPU · Plasma · Kinetic Theory · Waves

1 Introduction

The Finite Difference Time Domain (FDTD) method among the various numerical methods is the most widely used to solve electromagnetic problems, as it has a simple structure and high accuracy. In other to solving Maxwell’s equations to obtain the distribution of electric and magnetic fields in space and time domains, the FDTD method is the best approach. Maxwell’s equations are discretized using the Yee algorithm in this method, and the distribution of electric and magnetic fields in time and space is obtained.[Elsherbeni and Demir, 2015] However, when the mesh number or the number of time steps is very large, the execution speed decreases dramatically. Therefore, execution speed is a critical factor in this condition.

*Use footnote for providing further information about author (webpage, alternative address)—not for acknowledging funding agencies.
Traditional serial FDTD codes do not have enough performance. Hence the best way to increase the FDTD code execution speed is parallelization. However, it is not an easy task because the explicit leap-frog scheme has been applied in the FDTD method.

Today, a Graphics Processing Unit (GPU) benefits from hundreds of computing cores with higher computing ability than a Central Processing Unit (CPU). Therefore, many attempts have been made to run programs on a GPU. Yu et al. [2013], Demir [2010], Jiang et al. [2011], Xiong et al. [2018], Diener and Elsherbeni [2017], Weiss et al. [2019]. In the past, parallelization of the FDTD method on a GPU using CUDA Kernels has been well developed. Besides, many works have been done in optimizing GPU FDTD programming in MATLAB. [Diener and Elsherbeni [2017], Weiss et al. [2019]. However, there are still limitations that significantly impact the code execution speed.

Since MATLAB is widely used in most engineering fields, optimized MATLAB code is required. However, despite the advantages of programming over MATLAB, such as ease of use, MATLAB is slower than lower-level languages. So increasing the execution speed in the MATLAB environment is a very important factor. FDTD method has various applications in solving electromagnetic problems, such as electromagnetic wave propagation in plasma. Since solving a system of equations is essential to simulate electromagnetic wave propagation in plasma, an optimized FDTD code is needed to execute this large number of computations.

There are several ways for plasma modeling using the FDTD method:

- The Drude model is commonly used to approximate the frequency behavior of plasma. In this case, plasma is modeled as a dispersive medium [Guo et al. 2018].

- The fluid model describes plasma using macroscopic quantities. This model solves a set of Maxwell equations and velocity moments of the Boltzmann equation simultaneously [Arcese et al. 2017].

- Finally, the last and most complete way to model plasma is the kinetic model based on solving the Boltzmann equation. In this approach, the temporal evolution of the electron distribution function is computed [Cerri et al. 2008, Russo et al. 2010, Liu et al. 2018].

This work is organized into six sections. In section 2, the formulation of the FDTD method is studied. In section 3, the modified GPU FDTD code is introduced. In section 4, the validation of the modified FDTD code in the GPU is studied. In section 5, the three-dimensional propagation of electromagnetic waves in plasma is simulated using the FDTD method and Boltzmann equation. Finally, a summary and conclusion are presented in Section 6.

2 FORMULATION OF FDTD METHOD

The formulation of the FDTD method begins with Maxwell’s equations

\[
\begin{align*}
\mu_0 \frac{\partial \vec{H}}{\partial t} & = -\nabla \times \vec{E} - \vec{M}, \\
\varepsilon_0 \frac{\partial \vec{E}}{\partial t} & = \nabla \times \vec{H} - \vec{J},
\end{align*}
\]

where \( \vec{M} \) and \( \vec{J} \) are the magnetic and electric current densities, respectively. Discretized Maxwell’s equations, using the Yee algorithm, are as follows:

\[
E_x^{n+1}_{i,j,k} = C_{E_x}^{E_x} E_x^n_{i,j,k} + C_{H_z}^{E_x} H_z^n_{i,j,k} - H_z^n_{i,j,k-1} + C_{J_x}^{E_x} J_x^n_{i,j,k},
\]

where

\[
C_{E_x}^{E_x} = \frac{1}{\varepsilon_0}, \quad C_{H_z}^{E_x} = \frac{1}{\mu_0}, \quad C_{J_x}^{E_x} = \frac{\Delta t}{\Delta x},
\]

and \( \Delta t \) is the time step, and \( \Delta x \) is the spatial step.
where

\[
C_{E_x}^{i,j,k} = \frac{2\varepsilon_0 - \delta t \sigma_x^{i,j,k}}{2\varepsilon_0 + \delta t \sigma_x^{i,j,k}};
\]

\[
C_{H_x}^{i,j,k} = \frac{2\delta t}{2\varepsilon_0 \delta y + \delta t \sigma_y^{i,j,k}};
\]

\[
C_{E_y}^{i,j,k} = \frac{2\delta t}{2\varepsilon_0 \delta z + \delta t \sigma_z^{i,j,k} \delta z};
\]

\[
C_{H_y}^{i,j,k} = \frac{2\delta t}{2\varepsilon_0 + \delta t \sigma_z^{i,j,k}}.
\]

In the same way, the other components of the electric and magnetic fields are discretized.

Given that computer storage is limited, FDTD problem space should be truncated by special boundary conditions that simulate electromagnetic waves propagating continuously beyond the calculation. Convolutional Perfectly Matched Layer (CPML) is implemented among various types of absorbing boundary conditions. The equation for the CPML region is

\[
E_x^{n+1}_{i,j,k} = C_{E_x}^{i,j,k} E_x^n_{i,j,k} + \left( C_{H_x}^{i,j,k} / \kappa_{z,1,j,k} \right) \times \left[ H_z^{n+1/2}_{i,j,k} - H_z^{n+1/2}_{i,j-1,k} \right]
\]

\[
+ \left( C_{H_y}^{i,j,k} / \kappa_{y,1,j,k} \right) \times \left[ H_y^{n+1/2}_{i,j,k} - H_y^{n+1/2}_{i,j,k-1} \right]
\]

\[+ C_{E_y}^{i,j,k} E_y^{n+1/2}_{i,j,k} + C_{E_z}^{i,j,k} E_z^{n+1/2}_{i,j,k},\]

where

\[
\psi_y^{n+1/2}_{i,j,k} = b_y^e \psi_{y}^{n+1/2}_{i,j,k} + a_y^e \left[ H_z^{n+1/2}_{i,j,k} - H_z^{n+1/2}_{i,j-1,k} \right],
\]

\[
b_y^e = e^{-\left( \sigma_{\max}(\rho/\delta) n_{cpml} / \sigma_y \right)}^{\delta y},
\]

\[
a_y^e = \frac{\sigma_{max}(\rho/\delta)^{n_{cpml}}}{\delta y (\sigma_{\max}(\rho/\delta)^{n_{cpml}} \kappa_{y}^2 + \alpha_{y}^r \kappa_{y}^2)} \left[ b_y^e - 1 \right];
\]

\[
\sigma_{\max} = \sigma_{factor} \times \frac{n_{cpml} + 1}{150 \pi \delta y},
\]

\[
\alpha_{y}^r(\rho) = \alpha_{\min} + \left( \alpha_{\max} - \alpha_{\min} \right) \left( 1 - \frac{\rho}{\delta} \right).
\]

Here, \( n_{cpml} \) is the order of CPML, \( \rho \) is boundary distance to the desired point, \( \delta \) is the thickness of the CPML layer, \( \sigma_{factor} \) is about 0.7 to 1.5, \( \kappa_{max} \) is about 5 to 11, \( \alpha_{max} \) is about 0 to 0.05, \( n_{cpml} \) is about 2 to 4 and the thickness of boundary is 8 cells.

To measure the execution speed of the FDTD method, the number of processed cells per second is calculated [Demir 2010]:

\[
NMCPS = \frac{(n_{steps} \times n_x \times n_y \times n_z)}{t_s} \times 10^{-6},
\]

where \( NMCPS \) is the number of the cells processed per second, \( n_{steps} \) is the total number of time steps, \( t_s \) is the total simulation time, and \( n_x, n_y, \) and \( n_z \) are the number of cells in problem space in \( x, y \) and \( z \) directions, respectively.

The system specifications used in the simulation are graphics card GeForce GTX1050 and CPU Intel Core i7 – 7700HQ@2.80GHz.

3  MODIFICATION OF GPU FDTD CODE USING MATLAB

Nowadays, the GPU has played a significant role in accelerating numerical calculations. Using the MATLAB parallel processing toolbox for parallelization code has some limitations. The lack of attention to these limitations reduces the execution speed in some cases, even slower than the serial mode on the CPU unit.
Table 1: The notation of the FDTD parameters used in the code.

| Parameters | Notations |
|------------|------------|
| $x_+$, $y_+$, $z_+$ | $x_p$, $y_p$, $z_p$ |
| $x_-$, $y_-$, $z_-$ | $x_n$, $y_n$, $z_n$ |
| $C_{E_x}$, $C_{H_x}$, $C_{H_y}$ | $C_{exx}$, $C_{exy}$, $C_{exy}$ |
| $\psi_x$, $\psi_y$ | $\psi_{ex_yn}$, $\psi_{ex_yn}$ |
| $C_{\psi_x}$, $C_{\psi_y}$ | $C_{\psi_{ex_yn}}$, $C_{\psi_{ex_yn}}$ |
| $a_e$, $b_e$ | $cpml_a_yn$, $cpml_b_yn$ |

The basic FDTD source code is generated from [Elsherbeni and Demir, 2015], and in this paper mentioned source code is optimized and parallelized using MATLAB parallel processing toolbox.

In the first step, using the MATLAB parallel processing toolbox, all the matrices in the code are defined on the GPU using the `gpuArray` command.

### 3.1 Avoiding Array Indexing

The significant factor that decreases the execution speed is the array indexing, which should be avoided as much as possible. Since array indexing causes wasting time in each iteration, it should be avoided, especially when the number of iterations is large. Therefore, finding a solution to this problem is required for optimizing code. For this purpose, some corrections should be made, such as unifying the dimension of the field matrices and their coefficient matrices and defining some new auxiliary matrices before the main loop.

#### 3.1.1 Defining fields and their coefficients

Field matrices should be defined in the same size

\[
\begin{align*}
Ex &= \text{zeros}(nx+1, ny+1, nz+1, 'gpuArray'); \\
Ey &= \text{zeros}(nx+1, ny+1, nz+1, 'gpuArray'); \\
Ez &= \text{zeros}(nx+1, ny+1, nz+1, 'gpuArray');
\end{align*}
\]

After defining the coefficients of the fields, their dimensions are unified by joining a zero matrix to the coefficient matrices as follows:

\[
\begin{align*}
C_{exx} &= \text{gpuArray}(\text{cat}(1, C_{exx}, \text{zeros}(1, ny+1, nz+1))); \\
C_{exy} &= \text{gpuArray}(\text{cat}(1, C_{exy}, \text{zeros}(1, ny+1, nz+1))); \\
C_{exy} &= \text{gpuArray}(\text{cat}(1, C_{exy}, \text{zeros}(1, ny+1, nz+1)));
\end{align*}
\]

#### 3.1.2 Definition of auxiliary matrices

Since all field matrices are defined in the same size, there are additional arrays that should always be zero. Therefore, auxiliary matrices are defined so that by multiplying the auxiliary matrix by the corresponding field matrix, the additional arrays stay zero, and the main arrays do not change. The size of fields must be $(nx + 1 \times ny + 1 \times nz + 1)$ and arrays $(nx + 1, :, :)$ are additional array so in the auxiliary matrix, all the arrays corresponding to the additional arrays are zero and the others are one.

\[
\begin{align*}
Ex_A &= \text{ones}(nx+1, ny+1, nz+1, 'gpuArray'); \\
Ex_A(nx+1,:,:) &= 0;
\end{align*}
\]

#### 3.1.3 Definition of boundary condition coefficients

By considering the location of each boundary, the 3D matrix of boundary condition coefficients is defined. The boundary condition coefficients related to the $E_x$ field at the $-y$ boundary are given as follows:

Since coefficients $\psi_{ex_yn}$ and $CP\psi_{ex_yn}$ are associated with the $Ex$ field, then additional arrays in these coefficients and $Ex$ are same.

\[
\begin{align*}
\psi_{ex_yn} &= \text{cat}(1, \text{zeros}(1, n_{cpml_yn}, nz+1), \psi_{ex_yn}); \\
\psi_{ex_yn} &= \text{cat}(2, \text{zeros}(nx+1, ny-n_{cpml_yn}, nz+1)); \\
\psi_{ex_yn} &= \text{gpuArray}(\text{cat}(2, \text{zeros}(nx+1, nz+1), \psi_{ex_yn}));
\end{align*}
\]
\begin{verbatim}
CPsi_ex_yn = cat(1, CPsi_ex_yn, zeros(1, n_cpml_yn, nz + 1));
CPsi_ex_yn = cat(2, CPsi_ex_yn, zeros(nx+1, ny-n_cpml_yn, nz + 1));
CPsi_ex_yn = gpuArray(cat(2, zeros(nx+1, nz+1), CPsi_ex_yn));

Since coefficients \textit{cpml\_a\_e\_yn} and \textit{cpml\_b\_e\_yn} are associated with the \textit{Hz} field, additional arrays in these coefficients and \textit{Hz} are same.

\texttt{cpml\_a\_e\_yn = cpml\_a\_e\_yn * ones(nx, n_cpml_yn, nz + 1);}
\texttt{cpml\_a\_e\_yn = cat(2, cpml\_a\_e\_yn, zeros(nx, ny-n_cpml_yn, nz + 1));}
\texttt{cpml\_a\_e\_yn = gpuArray(cat(1, cpml\_a\_e\_yn, zeros(1, nyp1, nz + 1)));}
\texttt{cpml\_b\_e\_yn = cpml\_b\_e\_yn * ones(nx, n_cpml_yn, nz + 1);}
\texttt{cpml\_b\_e\_yn = cat(2, cpml\_b\_e\_yn, zeros(nx, ny-n_cpml_yn, nz + 1));}
\texttt{cpml\_b\_e\_yn = cat(2, zeros(nx, 1, nz + 1), cpml\_b\_e\_yn);}
\texttt{cpml\_b\_e\_yn = gpuArray(cat(1, cpml\_b\_e\_yn, zeros(1, nyp1, nz + 1)));}

Since coefficients \textit{cpml\_a\_A\_e\_yn} and \textit{cpml\_b\_A\_e\_yn} are associated with the \textit{Hx} field, additional arrays in these coefficients and \textit{Hx} are same.

\texttt{cpml\_a\_A\_e\_yn = cpml\_a\_e\_yn * ones(nx+1, n_cpml_yn, nz);}
\texttt{cpml\_a\_A\_e\_yn = cat(2, zeros(nx+1, nz), cpml\_a\_x\_e\_yn);}
\texttt{cpml\_a\_A\_e\_yn = gpuArray(cat(3, cpml\_a\_x\_e\_yn, zeros(nx+1, ny+1, 1)));}
\texttt{cpml\_b\_A\_e\_yn = cpml\_b\_e\_yn * ones(nx+1, n_cpml_yn, nz);}
\texttt{cpml\_b\_A\_e\_yn = cat(2, cpml\_b\_x\_e\_yn, zeros(nx+1, ny-n_cpml_yn, nz));}
\texttt{cpml\_b\_A\_e\_yn = cat(2, zeros(nx+1, 1, nz), cpml\_b\_e\_yn);}
\texttt{cpml\_b\_A\_e\_yn = gpuArray(cat(3, cpml\_b\_x\_e\_yn, zeros(nx+1, ny+1, 1)));}
\end{verbatim}

3.1.4 Corrections the main loop

The main loop is composed of two functions: the first function calculates the magnetic field, and the second function calculates the electric field using this magnetic field. In these two functions, two new series of matrices are defined in addition to the fields, their coefficients, and boundary conditions coefficients. The following is an example of these matrices:

\begin{verbatim}
Ez_y_Zeros = (Ez(:, 1:end-1,:), zeros(nx+1,1, nz+1, 'gpuArray'));
Ez_y_Shift = (Ez(:, 2:end,:), zeros(nx+1,1, nz+1, 'gpuArray'));
Hz_y_Zeros = (zeros(nx+1,1, nz+1, 'gpuArray'), Hz(:, 1:end-1,:), ...
             zeros(nx+1,1, nz+1, 'gpuArray'));
Hz_y_Shift = (zeros(nx+1,1, nz+1, 'gpuArray'), Hz(:, 1:end-2,:), ...
             zeros(nx+1,1, nz+1, 'gpuArray'));
\end{verbatim}

- The \textit{Update\_H} function is defined as follows:
  Evolution of magnetic fields at the boundary
  \begin{verbatim}
Psi_hx_yn = cpml\_b\_my\_yn.*Psi_hx_yn+cpml\_a\_my\_yn.*(Ez_y_Shift-Ez_y_Zeros);
  Hx = Hx+CPsi_hx_yn.*Psi_hx_yn;
Psi_hx_yp = cpml\_b\_my\_yp.*Psi_hx_yp+cpml\_a\_my\_yp.*(Ez_y_Shift-Ez_y_Zeros);
  Hx=Hx+CPsi_hx_yp.*Psi_hx_yp;
Psi_hz_zn = (cpml\_b\_mz\_zn).*Psi_hz_zn+(cpml\_a\_mz\_zn).*Ey_z_Shift-Ey_z_Zeros;
  Hz=Hz+CPsi_hz_zn.*Psi_hz_zn;
Psi_hz_zp = cpml\_b\_mz\_zp.*Psi_hz_zp+cpml\_a\_mz\_zp.*(Ey_z_Shift-Ey_z_Zeros);
  Hy=Hy+CPsi_hz_zp.*Psi_hz_zp;
\end{verbatim}

Evolution of magnetic field
\begin{verbatim}
Hx=((Chhx.*Hx)+(Chhex.*(Ey_z_Shift-Ey_z_Zeros))...+(Chhez.*(Ez_y_Shift-Ez_y_Zeros))).*Hx_A;
\end{verbatim}

- The \textit{Update\_E} function is defined as follows:
  Evolution of electric fields at the boundary
Evolution of electric fields

\[ E_x = (C_{exex} \cdot E_x + C_{exhz} \cdot (H_{z_y_Zeros} - H_{z_y_Shift}) + \ldots + C_{exhy} \cdot (H_{z_y_Zeros} - H_{z_y_Shift})) \cdot E_x_A; \]

As shown in Fig. 1, the GPU code execution speed for mesh number smaller than 1\(MC\) is slower than the CPU code. Only in larger than 2\(MC\), the GPU code execution speed is about three times faster than the basic CPU code. However, the modification can improve execution speed up to 11 times.

Table 2: The auxiliary matrices are defined in the modified FDTD code.

| Parameters             | Definition                                      |
|------------------------|-------------------------------------------------|
| \(E_x \cdot A\)        | \(E_x\) auxiliary matrix                        |
| \(cpml\_a\_e\_yn\)    | \(cpml\_a\_e\_yn\) auxiliary matrix            |
| \(cpml\_b\_e\_yn\)    | \(cpml\_b\_e\_yn\) auxiliary matrix            |
| \(E_z\_y\_Shift\)     | \(E_z\) transformation matrix in \(y\) direction|
| \(E_z\_y\_Zeros\)     | \(E_z\) dimension matrix in \(y\) direction     |

4 VALIDATION OF THE MODIFIED FDTD CODE IN THE GPU

To validate the modified code, analytically and numerically solution of a Gaussian plane wave propagation in a vacuum is compared.

The Gaussian pulse is expressed as follows:

\[ E(x, t) = \exp \left( -\left( \frac{c(t - t_0) - x}{c\tau} \right)^2 \right), \]  (14)

where \(t_0\) is the delay time and \(\tau\) is the pulse width. In this case \(\tau = 0.5ns\) and \(t_0 = 2.25ns\).

The normalized global error \(L_2\) is calculated to evaluate the discretization error. If \(u\) is the exact value of the function and \(\bar{u}\) is the calculated result, the normalized global error between these two quantities is [Li et al., 2016].
\begin{align*}
L_2 &= \sqrt{\int (\bar{u} - u)^2 \, dx} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\bar{u}_i - u_i)^2},
\end{align*}

(15)

where \( N \) is the number of mesh.

Figure 2: Validation for the Gaussian pulse (a) comparison between exact and numerical solutions calculated by the modified GPU code (b) the global error in various grid sizes.

The exact and numerical results are significantly matched in Fig. 2(a). As shown in Fig. 2(b), the order of accuracy for the modified GPU code is 2.0, which is sufficient.

5 PLASMA SIMULATION CODE IN THE GPU

5.1 Physics of the Problem

The propagation of electromagnetic waves in plasma is studied by solving the Boltzmann and Maxwell equations system. The Boltzmann equation is

\begin{equation}
\frac{\partial F}{\partial t} + \vec{v} \cdot \nabla \vec{r} F + \frac{e}{m} (\vec{E} + \mu \vec{v} \times \vec{H}) \cdot \nabla \vec{v} F = \left[ \frac{\partial F}{\partial t} \right]_{\text{coll}},
\end{equation}

(16)

where \( F(\vec{v}, \vec{r}, t) \) is the electron distribution function.

The usual method to solve the Boltzmann equation is based on the expansion of the electron distribution function using spherical harmonics \( \text{Cerri et al., [2008]} \).

\begin{equation}
\vec{F}(v) = F_0(v) + \frac{v}{u} \cdot \vec{F}_1(v),
\end{equation}

(17)

which \( F_0(v) \) describes isotropic and \( \vec{F}_1(v) = F_{1x}(v)\hat{x} + F_{1y}(v)\hat{y} + F_{1z}(v)\hat{z} \) describes non-isotropic motion of particles.

By substituting equation (17) into equation (16) and after several manipulations based on the orthogonality of spherical harmonics, the following two equations for the isotropic and non-isotropic expressions are obtained:

\begin{equation}
\frac{\partial F_0}{\partial t} + \frac{v}{3} \nabla \vec{r} \cdot \vec{F}_1 - \frac{e}{3m v^2} \frac{\partial}{\partial v} \left( \frac{\vec{E} \cdot \vec{F}_1}{v^2} \right) = C_0(F),
\end{equation}

(18)

\begin{equation}
\frac{\partial \vec{F}_1}{\partial t} + \nu \nabla \vec{r} F_0 - \frac{e}{m} \frac{\partial F_0}{\partial v} = C_1(F),
\end{equation}

(19)

where \( C_0(F) \) and \( C_1(F) \) are collision expressions. It is worth mentioning that in the present paper, only elastic collisions are considered.
5.2 Numerical Model

In this model, according to the experimental results, the isotropic electron distribution function is considered the Druyvesteyn function [van der Gaag et al., 2021, Li et al., 2018].

The anisotropic electron distribution function is also considered zero at the initial moment, and its evolution is studied. Therefore, the three-dimensional model is described by the following system of equations:

\[
\begin{align*}
\mu_0 \frac{\partial \vec{H}}{\partial t} &= -\nabla \times \vec{E} - \vec{M}, \\
\varepsilon_0 \frac{\partial \vec{E}}{\partial t} &= \nabla \times \vec{H} - \vec{J}, \\
\frac{\partial F_1}{\partial t} - e m \vec{E} \frac{\partial F_0}{\partial v} &= -\nu(v) F_1,
\end{align*}
\]

where the elastic collision frequency \( \nu(v) \) and the current density \( \vec{J} \) are as follows:

\[
\begin{align*}
\nu(v) &= N s v, \\
\vec{J} &= -e \int_{v} \bar{v} F(\bar{v}, \vec{r}, t) d^4v = -e \frac{4\pi}{3} \int_{0}^{\infty} v^3 F_1 dv,
\end{align*}
\]

where \( s \) is the cross section for elastic collision between electrons and neutral atoms.

The overall flowchart used to simulate the propagation of electromagnetic waves in plasma is shown in Fig. 3.

The anisotropic function is a four-dimensional parameter (three-dimensional space and one-dimensional velocity). Undoubtedly, calculating four-dimensional plasma parameters need sufficient execution speed, so optimizing the FDTD code’s execution speed is vital. Therefore, the following modifications are applied.

![Figure 3: The flowchart used to simulate the propagation of electromagnetic waves in plasma.](image)
The anisotropic distribution function and the current density before the main loop should be defined as follows:

\[
f_0 = \text{zeros}(nx + 1, ny + 1, nz + 1, v_{\text{dim}}, 'gpuArray'); \tag{25}
\]

\[
f_1x = \text{zeros}(nx + 1, ny + 1, nz + 1, v_{\text{dim}}, 'gpuArray'); \tag{26}
\]

\[
Jx = \text{zeros}(nx + 1, ny + 1, nz + 1, 'gpuArray'); \tag{27}
\]

where \(v_{\text{dim}}\) is the number of velocities considered. In this case \(v_{\text{dim}} = 31\).

In the main loop, the anisotropic distribution function and the currents density are computed as follows

\[
f_1x = (1/(2 + dt \times nu_m)) \times ((2 - dt \times nu_m) \times \ldots \tag{28}
\]

\[
f_1x + (dt \times e./m_e) \times (def_f_0) \times (2 \times E_x); \\
Jx = (-4 \times pi \times e./3) \times gam \times sum(v^3 \times f_1x, 4); \tag{29}
\]

which \(def_f_0\) is a derivative of the isotropic distribution function relative to velocity, and \(gam\) is the velocity step.

Table 3: The utilized parameters for numerical study.

| Parameters                                      | Amount           |
|------------------------------------------------|------------------|
| Center frequency of the Gaussian-modulated pulse \(f_c\) | 20 GHz           |
| Bandwidth of the Gaussian-modulated pulse \(\Delta f\) | 20 GHz           |
| Time constant of the Gaussian-modulated pulse \(\tau\) | 48 ps            |
| Time shift of the Gaussian-modulated pulse \(t_0\)  | 217 ps           |
| Time Bandwidth of the Gaussian-modulated pulse \(\Delta t\) | 146 ps           |
| Electron density \(n\)                           | \(10^{19}\) m\(^3\) |
| Neutral density \(N\)                            | \(10^{23}\) m\(^3\) |
| Electron temperature \(T_e\)                     | 1 ev             |

5.3 Result

The electromagnetic wave propagation in Argon plasma is considered to evaluate the performance of the modified plasma code. In this problem, the z-axis is divided into three regions. The two regions are air in \(z < 0 \text{ mm}\) and \(z > 4 \text{ mm}\), and the plasma region is located in \(0 < z < 4 \text{ mm}\). The incident plane wave is a Gaussian modulated, irradiated from air into plasma.

Plasma and electromagnetic waves properties are presented in Table 3.

Figure 4: Temporal evolution of (a) the Gaussian-modulated plane wave radiation into plasma and (b) the plasma current density at the center of air-plasma interface along the z-axis.
Fig. 4(a) shows two-dimensional views of the electric field evolution along the z-axis at different times. As shown in Fig. 4(a), the Gaussian-modulated plane wave, which contains a frequency range, is irradiated to the plasma-air interface. Plasma behaves like a filter, so the waves with a frequency higher than the plasma frequency can be propagated into plasma, and the waves with a frequency lower than the plasma frequency are reflected from its interface. In this work, the electron-neutral cross-section described in reference [Alves, 2014] for cold Argon plasma is used.

According to the Figs. 4(a) and (b) in the initial time steps (t = 0 to t = \(\Delta t/2\)), there is no interaction between the electric field and plasma, so the current density and the electric field inside plasma are zero. Since the electric field is irradiated into the air-plasma interface from \(t = \Delta t/2\), the plasma current density corresponding to the electric field on the interface continues to rise until it reaches a peak of \(-30 \text{ mA/cm}^2\) at the \(t = \Delta t\) and then falls gradually until the electric field gets out of plasma (t = 3\(\Delta t/2\)).

In Fig. 5, the three-dimensional profile of the electric field is shown, (a) before the interface, (b) in the absorption region, (c) after the interaction (the high-frequency part is propagated, and the low-frequency part is reflected).

Some scattering parameters, such as \(S_{11}\) and \(S_{21}\), were studied to investigate the reflection, absorption, and transmission of electromagnetic waves in plasma. These parameters show the reflection and transmission from plasma, respectively. Therefore, two ports are defined at the beginning (z = 0mm) and end (z = 4mm) of plasma to measure these two scattering parameters.

As shown in Fig. 6 plasma completely reflects the electromagnetic waves at the frequencies below the plasma frequency (28.5GHz). So in these frequencies, \(s_{11}\) is zero. Moreover, plasma behaves like a lossy medium at the frequencies above the plasma, and a fraction of the wave is reflected, transmitted, and absorbed by the plasma.

The mentioned results described the physical behavior of plasma exactly. Furthermore, the execution speed of the modified GPU code is presented in Fig. 7. This figure shows that the plasma simulation execution speed significantly increased (up to 20 times) using GPU and attention to the mentioned modifications.
The modified codes’ significant advantages are summarized in Table 4.

| Modification                      | Advantages                         |
|-----------------------------------|------------------------------------|
| Modified FDTD GPU code            | 11x faster than basic GPU code      |
| Modified plasma code in GPU       | 20x faster than basic plasma code in CPU |

6 Conclusion

Several programming techniques were presented to reduce the execution time of the FDTD simulation in MATLAB software. Using MATLAB parallel computing toolbox and mentioned techniques, FDTD code was optimized on GPU. Validation of the modified code was examined, and the desired order of accuracy for the modified code was shown. Eventually, the interaction of electromagnetic waves with plasma (modeled by two-term Boltzmann approximation) was simulated on the GPU by applying the proposed techniques, and its significant effect on the execution speed of the plasma code was proved.

References

Atef Z Elsherbeni and Veysel Demir. *The finite-difference time-domain method for electromagnetics with MATLAB® simulations*, volume 2. IET, 2015.

Wenhua Yu, Xiaoling Yang, and Wenxing Li. *VALU, AVX and GPU acceleration techniques for parallel FDTD methods*. Scitech Pub Inc, 2013.

Veysel Demir. A stacking scheme to improve the efficiency of finite-difference time-domain solutions on graphics processing units. *The Applied Computational Electromagnetics Society Journal (ACES)*, pages 323–330, 2010.

Wang-Qiang Jiang, Min Zhang, Hui Chen, and Yong-Ge Lu. Cuda-based four-path method with application to the electromagnetic scattering of a two-layer canopy. *Waves in Random and Complex Media*, 21(3):529–541, 2011.

Lang-lang Xiong, Xi-min Wang, Song Liu, Zhi-yun Peng, and Shuang-ying Zhong. The electromagnetic waves propagation in unmagnetized plasma media using parallelized finite-difference time-domain method. *Optik*, 166: 8–14, 2018.

Joseph E Diener and Atef Z Elsherbeni. Fdtd acceleration using matlab parallel computing toolbox and gpu. *The Applied Computational Electromagnetics Society Journal (ACES)*, pages 283–288, 2017.

Alec J Weiss, Atef Z Elsherbeni, Veysel Demir, and Mohammed F Hadi. Using matlab’s parallel processing toolbox for multi-cpu and multi-gpu accelerated fdtd simulations. *The Applied Computational Electromagnetics Society Journal (ACES)*, pages 724–730, 2019.

Li-xin Guo, Wei Chen, Jiang-ting Li, Yi Ren, and Song-hua Liu. Scattering characteristics of electromagnetic waves in time and space inhomogeneous weakly ionized dusty plasma sheath. *Physics of Plasmas*, 25(5):053707, 2018.

Emanuele Arcese, François Rogier, and Jean-Pierre Boeuf. Plasma fluid modeling of microwave streamers: Approximations and accuracy. *Physics of Plasmas*, 24(11):113517, 2017.
Graziano Cerri, Franco Moglie, Ruggero Montesi, Paola Russo, and Eleonora Vecchioni. Fdtd solution of the maxwell–boltzmann system for electromagnetic wave propagation in a plasma. *IEEE transactions on antennas and propagation*, 56(8):2584–2588, 2008.

P Russo, G Cerri, and E Vecchioni. Self-consistent model for the characterisation of plasma ignition by propagation of an electromagnetic wave to be used for plasma antennas design. *IET microwaves, antennas & propagation*, 4(12):2256–2264, 2010.

Jian-Xiao Liu, Ze-Kun Yang, Lu Ju, Lei-Qing Pan, Zhi-Gang Xu, and Hong-Wei Yang. Boltzmann finite-difference time-domain method research electromagnetic wave oblique incidence into plasma. *Plasmonics*, 13(5):1699–1704, 2018.

Yong Li, Haiyan Xie, Yonghui Guo, and Chun Xuan. Verification of fdtd code using the method of exact solutions. In *2016 Progress in Electromagnetic Research Symposium (PIERS)*, pages 679–685. IEEE, 2016.

Thijs van der Gaag, Hiroshi Onishi, and Hiroshi Akatsuka. Arbitrary eedf determination of atmospheric-pressure plasma by applying machine learning to oes measurement. *Physics of Plasmas*, 28(3):033511, 2021.

Jinming Li, Chengxun Yuan, Zhongxiang Zhou, AA Kudryavtsev, Xiaoou Wang, and Xiaodong Chen. Effects of druvesteyn distribution to transmission coefficients in plasma. In *2018 12th International Symposium on Antennas, Propagation and EM Theory (ISAPE)*, pages 1–3. IEEE, 2018.

LL Alves. The ist-lisbon database on lxcat. In *Journal of Physics: Conference Series*, volume 565, page 012007. IOP Publishing, 2014.