URSI GASS 2021, Rome, Italy, 28 August - 4 September 2021

GPU-Accelerated Finite Element and Finite Difference Methods for Scattering Problems in Voxel-based Human Models

Alessandro Arduino*, Oriano Bottauscio, and Luca Zilberti
Istituto Nazionale di Ricerca Metrologica (INRIM), Torino, ITALY

Abstract

The GPU-accelerated implementations of two three-dimensional full-wave electromagnetic solvers for scattering radiofrequency problems, based on the finite element and the finite difference methods, are compared in terms of accuracy and performances, looking also at the achieved speed-up with respect to the serial CPU implementations. The finite element method shows a speed-up of about 10 for medium-size problems, whereas the acceleration rises up to 50 when the finite difference method is used. The described codes are then used to compute the specific absorption rate of a magnetic resonance imaging radiofrequency birdcage coil in a voxel-based anatomical human model.

1 Introduction

Fast and reliable computing of the distribution of radiofrequency (RF) electromagnetic (EM) fields produced within a human body by an external source is a fundamental requirement for many applications. A direct example is the dosimetric assessment of the RF radiation generated by magnetic resonance imaging (MRI) [1]. The need of shortening computational times becomes even larger when dealing with inverse problems, which repeatedly solve a direct EM problem until convergence is reached. The design of dielectric pads for shimming in high-field MRI [2] and the quantitative imaging of the electric properties from MRI measurements [3, 4] are just a couple of examples of inverse problems at RF.

Different consolidated EM models proposed in literature lead to reliable algorithms that can be accelerated to achieve the desired speed-up using parallel programming on graphics processing units (GPUs) [1]. Common frameworks for general-purpose programming of GPUs are NVIDIA's CUDA Toolkit [5] and OpenCL [6].

In this paper, the RF scattering problem is studied and a couple of methods based on the biconjugate gradient stabilized (BiCGstab) fast Fourier transform (FFT) [7] are compared. The incident EM field generated by the RF source in vacuum is assumed known and is the forcing term, whereas its alteration due to the presence of a human body (i.e., the scatterer) is computed. Both serial and parallel C++ implementations, with CUDA toolkit 9.1 [5] to run on NVIDIA's GPUs, are presented and compared in terms of performances when applied to an analytically solved model problem. Finally, the methods are tested in a more elaborated set-up featuring the voxel-based anatomical human model Duke [8] exposed to the EM field generated by a RF head coil for MRI.

2 Mathematical Formulation

Let us denote by $\{\boldsymbol{E}, \boldsymbol{H}\}$ the EM field generated by the studied RF source in the presence of the scatterer and by $\{\boldsymbol{E}^{\mathrm{i}}, \boldsymbol{H}^{\mathrm{i}}\}$ the incident EM field generated by the same source in vacuum. Denoting by $\boldsymbol{D} = \tilde{\varepsilon}\,\boldsymbol{E}$ the electric flux density, in an unbounded domain with Sommerfeld radiation conditions at infinity, it solves [7]

$$\tilde{\varepsilon}^{-1}\boldsymbol{D} - k_0^2\boldsymbol{A} - \nabla\left(\nabla\cdot\boldsymbol{A}\right) = \boldsymbol{E}^{\mathrm{i}}, \tag{1}$$

where $\tilde{\varepsilon} = \varepsilon - \mathrm{i}\sigma/\omega$ is the complex permittivity, with ε the electric permittivity, σ the electric conductivity, and ω the angular frequency of the radiation; $k_0^2 = \omega^2\,\varepsilon_0\,\mu_0$ is the propagation coefficient of the radiation, with μ_0 the magnetic permeability of vacuum; $\boldsymbol{A}$ is the scaled magnetic vector potential

$$\boldsymbol{A}(\boldsymbol{x}) = \int_{R^3} \varphi(\boldsymbol{x}-\boldsymbol{y})\left(\tilde{\varepsilon}(\boldsymbol{y})-\varepsilon_0\right)\left(\varepsilon_0\tilde{\varepsilon}(\boldsymbol{y})\right)^{-1}\boldsymbol{D}(\boldsymbol{y})\mathrm{d}\boldsymbol{y}, \tag{2}$$

with $\varphi(\boldsymbol{r}) = \exp(-\mathrm{i}\,k_0\,|\boldsymbol{r}|)/(4\,\pi\,|\boldsymbol{r}|)$ the Green's function.

2.1 Weak Form and Finite Element Method

In order to use the finite element method (FEM), equation (1) is written in weak form by testing it against function $\boldsymbol{v}$,

$$\begin{aligned}\int_\Omega\left(\tilde{\varepsilon}^{-1}\boldsymbol{D} - k_0^2\boldsymbol{A}\right)\cdot\boldsymbol{v} + \int_\Omega\left(\nabla\cdot\boldsymbol{A}\right)\left(\nabla\cdot\boldsymbol{v}\right) + \\ -\int_{\partial\Omega}\left(\nabla\cdot\boldsymbol{A}\right)\boldsymbol{v}\cdot\boldsymbol{n} = \int_\Omega\boldsymbol{E}^{\mathrm{i}}\cdot\boldsymbol{v},\end{aligned} \tag{3}$$

where the boundary integral arises from application of the divergence theorem to the grad-div term. In the latter equation, the integrals are performed on a bounded domain Ω containing the volume occupied by the scatterer and a portion of air around it. The symbol $\boldsymbol{n}$ denotes the outward normal unit vector to the domain boundary $\partial\Omega$. Despite the latter equation does not contain explicitly any boundary condition, the relation (2) between $\boldsymbol{A}$ and $\boldsymbol{D}$ implicitly defines the Sommerfeld radiation condition ensuring for the existence and the unicity of the solution.

The electric flux density $\boldsymbol{D}$ and the related scaled magnetic vector potential $\boldsymbol{A}$, both of which belong to $\mathbf{H}(\mathrm{div}, \Omega)$ and such that (3) holds for any test function $\boldsymbol{v}$ in $\mathbf{H}(\mathrm{div}, \Omega)$, are looked for. In order to keep the div-conformity features of the functions in the discretization, finite Nédélec's facet elements [9] are used to discretize

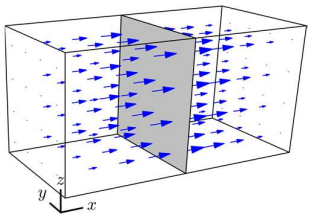

Figure 1. Representation of the vector shape function associated to an internal face of the Cartesian mesh.

(3) on a uniform Cartesian mesh, whose elements are called voxels. The vector shape function associated to a generic face of the mesh is depicted in Fig. 1 and is such that its flux through the associated face is unitary. The FEM discretization leads to the linear system

$$\mathbf{M}_{\mathrm{FEM}}\boldsymbol{d}+\mathbf{K}_{\mathrm{FEM}}\boldsymbol{a}=\boldsymbol{b}_{\mathrm{FEM}}, \tag{4}$$

where $\mathbf{M}_{\mathrm{FEM}}$ and $\mathbf{K}_{\mathrm{FEM}}$ are the sparse FEM matrices, $\boldsymbol{b}_{\mathrm{FEM}}$ is the forcing term, $\boldsymbol{d}$ and $\boldsymbol{a}$ are the vectors approximating $\boldsymbol{D}$ and $\boldsymbol{A}$, respectively, related to the value of the fields in the centers of mass of the faces in the mesh.

2.2 Finite Difference Method

The finite difference method (FDM) can be used to directly discretize (1) in a uniform Cartesian mesh. In particular, the grad-div term is discretized using the centered second order finite differences between the values at the centers of mass of the voxels in the mesh, as shown in Fig. 2. The resulting linear system can be written as follows,

$$\mathbf{M}_{\mathrm{FDM}}\boldsymbol{d}+\mathbf{K}_{\mathrm{FDM}}\boldsymbol{a}=\boldsymbol{b}_{\mathrm{FDM}}. \tag{5}$$

In this case, $\boldsymbol{d}$ approximates the values of $\boldsymbol{D}$ in the mesh voxels' barycenters. Since (2) holds everywhere, $\boldsymbol{a}$ conveniently approximates $\boldsymbol{A}$ in the voxels' barycenters of an extended mesh, bigger than the actual domain of two voxels in each direction (cf. Fig. 2). Thus, whilst $\mathbf{M}_{\mathrm{FDM}}$ is simply a diagonal matrix, $\mathbf{K}_{\mathrm{FDM}}$ is a rectangular one [4].

2.3 Scaled Magnetic Vector Potential

In order to make equations (4) and (5) solvable, the dependence of the potential $\boldsymbol{A}$ to the flux density $\boldsymbol{D}$ must be approximated. For both FEM and FDM, $\boldsymbol{d}$ estimates each component of $\boldsymbol{D}$ in a three-dimensional Cartesian grid—parallel faces' barycenters or voxels' barycenters. This feature allows exploiting the FFT for an efficient evaluation of the convolution product in (2), which can be formally expressed as a matrix-vector product [10]

$$\boldsymbol{a}=\mathbf{G}_{\mathrm{X}}\boldsymbol{d}, \tag{6}$$

whose dense matrix $\mathbf{G}_{\mathrm{X}}$ is never actually evaluated. Finally, the linear system to be solved can be written as

$$\left(\mathbf{M}_{\mathrm{X}}+\mathbf{K}_{\mathrm{X}}\mathbf{G}_{\mathrm{X}}\right)\boldsymbol{d}=\boldsymbol{b}, \tag{7}$$

where X can be either FEM or FDM.

3 Code Implementation

The linear system (7) is solved with the matrix-free iterative solver BiCGstab [7] implemented in serial and parallel C++ codes with CUDA toolkit 9.1 [5]. In particular, the matrix-vector product with $\mathbf{G}_{\mathrm{X}}$ is evaluated using the libraries FFTW3 [11] and cuFFT [5] for FFT computation. In order to reduce the time consumed by matrix $\mathbf{G}_{\mathrm{X}}$ application, the convolution kernel FFT is computed once and stored throughout the procedure.

The sparse FEM matrices $\mathbf{M}_{\mathrm{FEM}}$ and $\mathbf{K}_{\mathrm{FEM}}$ are assembled by CPU with the library EIGEN3 [12]. Then, in the parallel code, they are transferred to GPU, stored throughout the procedure and managed with the library cuSPARSE [5]. On the other hand, the sparse FDM matrices $\mathbf{M}_{\mathrm{FDM}}$ and $\mathbf{K}_{\mathrm{FDM}}$ are not explicitly computed, but are, instead, applied as filters. In particular, the diagonal contributions are implemented as the Hadamard product of vectors, whereas the grad-div term is computed by applying the filter whose stencil is shown in Fig. 2. In the parallel code version, this is done by means of CUDA kernels implemented *ad hoc*.

The BiCGstab method provided by EIGEN3 is used for the serial code, whereas the library cuBLAS [5] is employed for implementing it to work on GPU. Neither the serial nor the parallel solvers make use of matrix preconditioning. The typical bottleneck in GPU programming—communications between host and device—are completely avoided during the execution.

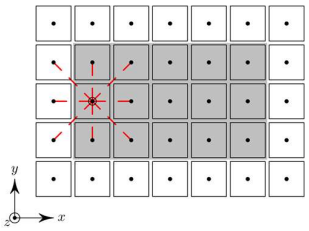

Figure 2. Representation of the FDM mesh with voxel barycenters denoted by dots and the computational domain Ω colored in grey. The stencil in the plane of the grad-div operator is pictured for the circled node by the dashed lines. $\boldsymbol{A}$ is evaluated also in white voxels.

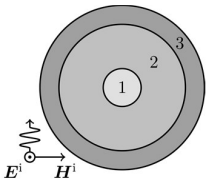

Figure 3. Schematic representation of the reference model problem.

4 Validation and Performance Evaluation

The codes are tested against an analytically solved model problem: the scattering of a layered sphere illuminated by a plane wave at 128 MHz. Three layers are considered (cf. Fig. 3), whose geometric and electric properties are collected in Table 1. From the inner to the outer layer, the properties mimic those of cerebrospinal fluid (CSF), white matter and skull, respectively [13].

Table 1. Properties of the Model Problem

Region	Radius / cm	Relative permittivity	Electric conductivity / (S/m)
1	2.0	84.0	2.14
2	8.0	52.5	3.42×10^{-1}
3	10.0	14.7	6.74×10^{-2}

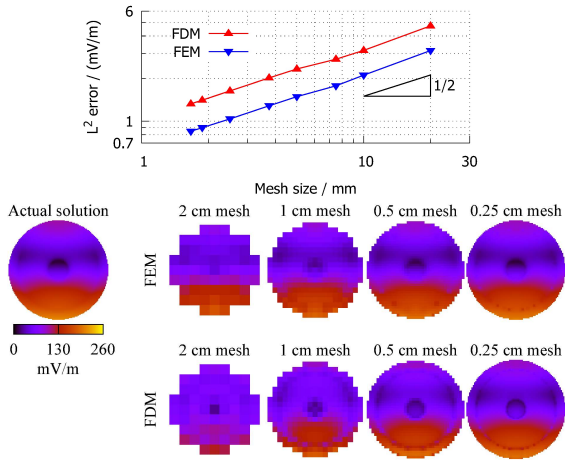

Figure 4. Top, trend of the L^2 error against different mesh sizes. Bottom, map of the actual solution (in the same plane of Fig. 3) to be compared with FEM and FDM approximations with different mesh sizes.

The analytical solution, evaluated according to [14], is reported in Fig. 4 (in the same plane of Fig. 3) and is compared to the approximations provided by FEM and FDM with different mesh sizes (the two applied on the same hexahedral meshes). The FEM approximation is more accurate than the FDM one, in particular at the boundary between contiguous materials. This observation is reflected in the global error measured in L^2 norm, whose trend is reported in Fig. 4. On the other hand, FDM appears to be much faster than FEM and to achieve better speed-up with respect to the computational time of CPU code, when executed on GPU. The performances of serial and parallel implementations, reported in Fig. 5, have been evaluated by running the codes on a core of an Intel Xeon E5-2699 v4 and on an NVIDIA Quadro K6000.

For FEM solver, the performances of the parallel code become favorable by increasing the problem size, up to a computational time about 10 times smaller than the serial one for the largest considered problem. Being matrix-free, FDM is more naturally parallelized than FEM and its parallel version is advantageous already for a small size problem. Its efficiency increases until a speed-up higher than 50 is reached for the largest problem. The last histogram of Fig. 5 compares the computational times of the parallel GPU implementations. The higher speed-up achieved by FDM translates in a faster execution of the solver when FDM is used instead of FEM.

It is worth noting that the speed-up results reported in Fig. 5 are independent of the simulated frequency, since

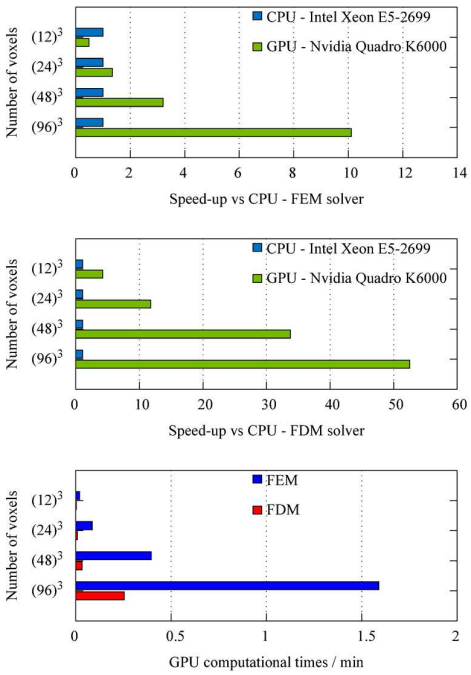

Figure 5. Top, speed-up of FEM implementation with respect to the computational time of the FEM CPU code. Middle, speed-up of the FDM implementation with respect to the computational time of the FDM CPU code. Bottom, computational times of the GPU codes.

they are due only to the matrix-vector products performed at each step of the iterative BiCGstab solver.

5 Biomedical Application

The EM field generated by a 16-leg birdcage head coil of a 3 T MRI scanner operating at the frequency of 128 MHz has been computed in order to check the codes in a more realistic case. The head coil radius is 14 cm and its height is 30 cm. The RF coil irradiates the head of the anatomical human model Duke [8], discretized with a 2 mm mesh of 124×99×121 voxels in x, y, and z direction, respectively. The reciprocal positioning of the human model and the birdcage coil is depicted in Fig. 6 and it is realistic for an MRI session of brain imaging. The electric properties are assigned to the biological tissues according to the IT'IS Foundation database [13].

A quadrature supply of the RF coil is simulated by imposing a unitary-peak current in each mesh of the coil, each one dephased of the geometrical angle. The incident field is then computed semi-analytically by integration.

Since the dosimetric quantity for safety assessment at RF is the specific absorption rate (SAR) of the EM radiation, the computed electric field is further elaborated to obtain

$$\mathrm{SAR}=\frac{\sigma|\boldsymbol{E}|^2}{2\rho}, \qquad (8)$$

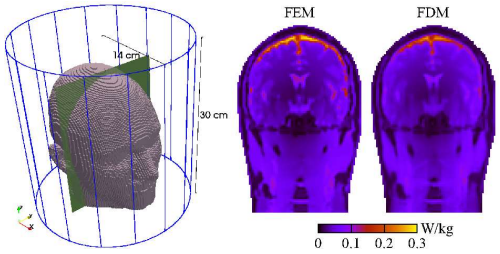

Figure 6. Left, schematic representation of the biomedical problem. Around the head model, the birdcage conductors are drawn in blue. Right, SAR distributions on the coronal plane pictured in the model scheme. The reported SAR is computed by the implemented FEM and FDM codes.

where ρ is the tissue density. In Fig. 6, the results obtained in the highlighted coronal section by FEM and FDM implementations are reported. Despite the global distribution of SAR is almost the same for the two methods, a slightly higher maximum value and more distributed hotspots are estimated by FEM, which has proved to be more accurate than FDM in the model problem. The execution of the parallel GPU code for FEM requires about 1 minute with a speed-up of about 10 with respect to the corresponding serial code. On the other hand, FDM provides the solution in less than 15 seconds—about 50 times faster than the corresponding CPU code.

6 Conclusion

Although FEM leads to more accurate approximations of the scattered EM fields, FDM can be efficiently parallelized to work on GPUs and achieves a speed-up of about 50 on a medium-size problem. Given the small computational times, in the solution of a single direct problem, the higher accuracy of FEM makes it the preferable choice; but when the direct solution must be repeated multiple times to solve an inverse problem, the velocity of FDM becomes relevant. The fast solution achieved by FDM could then be used as starting guess for a final refinement using FEM.

7 Acknowledgements

The results here presented have been developed in the framework of the 18HLT05 QUIERO project. This project has received funding from the EMPIR programme co-financed by the Participating States and from the European Union's Horizon 2020 research and innovation programme.

8 References

1. O. Bottauscio, M. Chiampi, J. Hand, and L. Zilberti, "A GPU computational code for eddy-current problems in voxel-based anatomy," *IEEE Trans. Magn.*, **51**, March 2015, art. num. 5100904.

2. W. M. Brink, R. F. Remis, and A. G. Webb, "A theoretical approach based on electromagnetic scattering for analyzing dielectric shimming in high-field MRI," *Magn. Reson. Med.*, **75**, May 2016, pp. 2185-2194.

3. R. Hong, S. Li, J. Zhang, Y. Zhang, N. Liu, Z. Yu, and Q. H. Liu, "3-D MRI-based electrical properties tomography using the volume integral equation method," *IEEE Trans. Microw. Theory*, **65**, December 2017, pp. 4802-4811.

4. R. L. Leijsen, W. M. Brink, C. A. T. van den Berg, A. G. Webb, and R. F. Remis, "3-D Contrast Source Inversion-Electrical Properties Tomography," *IEEE Trans. Med. Imag.*, **37**, September 2018, pp. 2080-2089.

5. (2019). *CUDA Toolkit*, accessed on 2019. [Online]. Available: https://developer.nvidia.com/cuda-toolkit

6. (2019). *OpenCL*, accessed on 2019. [Online]. Available: https://www.khronos.org/opencl/

7. P. Zwamborn, and P. M. van den Berg, "The three-dimensional weak form of the conjugate gradient FFT method for solving scattering problems," *IEEE Trans. Microw. Theory*, **40**, September 1992, pp. 1757-1766.

8.M.-C. Gosselin, *et al.*, "Development of a new generation of high-resolution anatomical models for medical device evaluation: the virtual population 3.0," *Phys. Med. Biol.*, **59**, August 2014, pp. 5287-5303.

9. V. Le-Van, G. Meunier, O. Chadebec, and J.-M. Guichon, "A volume integral formulation based on facet elements for nonlinear magnetostatic problems," *IEEE Trans. Magn.*, **51**, July 2015, art. num. 7002206.

10. R. D. Ryne (Nov. 2011). "On FFT-based convolutions and correlations, with application to solving Poisson's equation in an open rectangular pipe." [Online]. Available: https://arxiv.org/abs/1111.4971

11. M. Frigo, and S. G. Johnson, "The design and implementation of FFTW3," *Proc. IEEE,* **93**, February 2005, pp. 216-231.

12. G. Guennebaud, *et al.* (2010). *Eigen V3*, accessed on 2018. [Online]. Available: http://eigen.tuxfamily.org

13. P.A. Hasgall, *et al.*, "IT'IS Database for thermal and electromagnetic parameters of biological tissues," Version 4.0, May 2018. [Online]. Available: itis.swiss/database.

14. J.-P. Schäfer, "Implementierung und Anwendung analytischer und nume-rischer Verfahren zur Lösung der Maxwellgleichungen für die Untersu-chung der Lichtausbreitung in biologischem Gewebe," PhD thesis, Uni-versität Ulm, 2011. [Online]. Available: http://vts.uni-ulm.de/doc.asp?id=7663