ISTITUTO NAZIONALE DI RICERCA METROLOGICA
Repository Istituzionale

H-matrix Sparsification Applied to Bioelectromagnetic Analysis of Large Scale Human Models

(Article begins on next page)

22 November 2024

H-matrix Sparsification Applied to Bioelectromagnetic Analysis of Large Scale Human Models

Piergiorgio Alotto[1], Paolo Bettini[1,4], Oriano Bottauscio[2], *IEEE Senior Member*, Mario Chiampi[3], and Luca Zilberti[2]

[1] Dipartimento Ingegneria Industriale – Università degli studi di Padova – Via Gradenigo, 6/A – 35131 Padova, Italia
[2] Istituto Nazionale di Ricerca Metrologica (INRIM) – Strada delle Cacce, 91 – 10135 Torino, Italia
[3] Dipartimento Energia – Politecnico di Torino – Corso Duca degli Abruzzi, 24 – 10129 Torino, Italia
[4] Consorzio RFX – Corso Stati Uniti, 4 – 35127 Padova, Italia

This work proposes the use of H-matrices with Adaptive Cross Approximation for the solution of the very large sparse/dense block linear systems typically arising in hybrid differential/integral formulations applied to bioelectromagnetic problems which make use of large scale human models. The code implementing this approach is described in detail and the usefulness of the approach is first demonstrated on a benchmark provided with an analytical solution and then applied to a magnetic resonance imaging problem.

Index Terms— Biomedical computing, Hybrid differential/integral formulations, Sparse matrices.

I. INTRODUCTION

THE INCREASING application of electromagnetic fields in medicine, in particular for Magnetic Resonance Imaging (MRI), has produced significant progress both in diagnostics and therapeutic treatments. However, the interaction of the electromagnetic fields with biological tissues has raised the need of verifying patient's safety issues [1]. Excluding the possibility of in-vivo measurements, the problem has to be solved by computational means, making use of highly accurate anatomical models, like the one shown in Fig. 1, discretized into millions of volume elements (voxels), in order to correctly reproduce fine geometrical details.

Hybrid differential/integral formulations, which are found to be particularly convenient for the solution of such problems [2], give rise to large sparse/dense block matrices, whose size may easily lead to impractical memory and computational time requirements. Therefore, a compression of the dense blocks, e.g. with libraries implementing Adaptive Cross Approximation (ACA) coupled with hierarchical matrix (H-matrix) arithmetics [3]-[5], is required in order to solve realistic problems. While such approaches have been used successfully for dense systems, their application to sparse/dense block matrices is rare due to the difficulty of generating suitable preconditioners for the iterative solvers. Furthermore, their application to bioelectromagnetic problems, as the one presented in this paper, is novel.

II. FORMULATION

The electromagnetic analysis is performed within an unbounded, magnetically homogeneous domain Ω_{ext}, with vacuum magnetic permeability μ_0, containing a conductive subdomain Ω (e.g. a geometrical solid or a human body). The conductive subdomain Ω is discretized with cubic voxels, directly exploiting the features of the anatomical datasets (e.g. [8]), while the surrounding space Ω_{ext}, which contains also the field sources, is unmeshed. The problem is formulated in the frequency domain (angular frequency ω), using phasors to represent the field quantities.

Adopting a $\mathbf{T}$-φ formulation, the magnetic field is decomposed into the unperturbed source field ($\mathbf{H}_s$) and a reaction field, due to the total (i.e. ohmic and dielectric) current density distribution $\mathbf{J}$ induced in Ω. Expressing the reaction field in Ω through an electric vector potential $\mathbf{T}$ ($\mathbf{J} = \nabla \times \mathbf{T}$) and the gradient of a magnetic scalar potential φ, the total field $\mathbf{H}$ becomes: $\mathbf{H} = \mathbf{H}_s + \mathbf{T} + \nabla \varphi$.

The solution of the eddy-current problem is obtained by means of a hybrid technique, which uses the Finite Element Method (FEM) inside the subdomain Ω and the Boundary
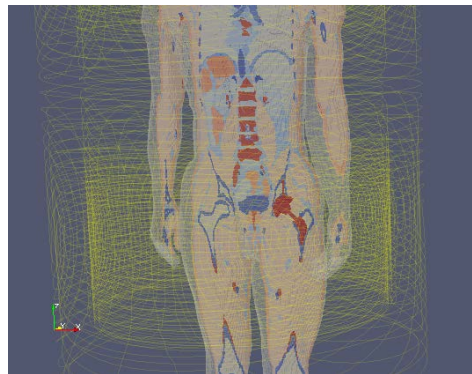

Fig. 1. Anatomical human model 'Duke', belonging to the IT'IS Virtual Family [8], modified with the insertion of a realistic unilateral metallic hip prosthesis. The model includes more than 80 tissues (shown with different colors in the figure) and is discretized in more than $8 \cdot 10^6$ cubic voxels.

Element Method (BEM) for Ω_{ext}. In the FEM region, the unknowns are the nodal values of φ and the projections of $\mathbf{T}$ on the mesh edges. In the BEM external space, $\mathbf{J}$ is assumed to be null and $\mathbf{T}$ is not defined ($\mathbf{H} = \mathbf{H}_s + \nabla\varphi$). The BE unknowns are the normal gradient of φ on each surface element, into which the bounding surface $\partial\Omega$ is discretized by the FE mesh. In order to bound $\mathbf{J}$ inside Ω, the tangential components of $\mathbf{T}$ on the FE-BE interface ($\partial\Omega$) and the related edge unknowns are imposed to zero.

Inside the subdomain Ω the FE field equations, written in the weak form, for the scalar and vector potentials are respectively:

$$\int_\Omega \nabla\varphi \cdot \nabla\upsilon \, dv - \int_{\partial\Omega} \upsilon(\nabla\varphi \cdot \mathbf{n})^{(i)} ds + \int_\Omega \mathbf{T} \cdot \nabla\upsilon \, dv = 0 \quad (1)$$

$$\int_\Omega \frac{1}{\tilde{\sigma}} \nabla \times \mathbf{T} \cdot \nabla \times \mathbf{w} \, dv + j\omega\mu_0 \int_\Omega (\mathbf{T} + \nabla\varphi) \cdot \mathbf{w} \, dv$$

$$= -j\omega\mu_0 \int_\Omega \mathbf{H}_s \cdot \mathbf{w} \, dv \quad (2)$$

where υ and $\mathbf{w}$ are respectively the nodal and the edge shape functions, $\tilde{\sigma} = \sigma + j\omega\varepsilon$ is the complex conductivity (σ: conductivity, ε: permittivity, j: imaginary unit), $\mathbf{n}$ is the local normal unit vector on the body surface $\partial\Omega$ and the superscript i indicates the internal side of $\partial\Omega$. In the external region Ω_{ext}, the following BEM field equation holds:

$$\xi\varphi + \int_{\partial\Omega} (\nabla\varphi \cdot \mathbf{n})^{(e)}\Psi ds - \int_{\partial\Omega} \varphi(\nabla\Psi \cdot \mathbf{n}) ds = 0 \quad (3)$$

where Ψ is the Green's function for quasi-stationary problems, $\xi = 0.5$ a BEM coefficient and e indicates the external side of $\partial\Omega$. Since relation $\mathbf{n} \times \mathbf{T} = 0$ on $\partial\Omega$ is the only constraint applied to $\mathbf{T}$, the formulation is ungauged and leads to indeterminate systems, which can be solved through a GMRES algorithm, anyway. The formulation based on equations (1) – (3) gives rise to a matrix structure composed of nine blocks, according to the following scheme:
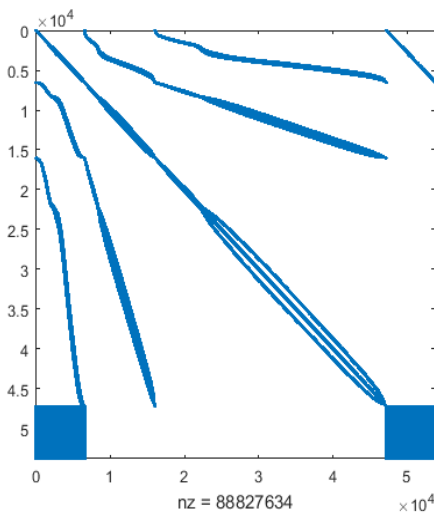

Fig. 2. Typical sparsity pattern of the linear system (4)

$$\begin{bmatrix} S_{nbn,nbn} & S_{nbn,nint} & S_{nbn,nf} \\ S_{nint,nbn} & S_{nint,nint} & 0 \\ F_{nf,nbn} & 0 & F_{nf,nf} \end{bmatrix} \quad (4)$$

where S or F indicate that a block is sparse or full, respectively, *nbn* is the number of boundary nodes (on $\partial\Omega$), *nint* is the total number of internal nodes and edges, while *nf* is the number of boundary faces. Fig. 2 shows a typical sparsity pattern of the linear system (4) in which the blocks are of the real relative sizes.

III. H-Matrix Low Rank Approximation

An original way of handling the linear system (4) was proposed in [6], adopting a massively parallelized solver. Anyway, in order to allow the solution of realistic large problems, it is mandatory to compress the linear system (4), and in particular its dense matrix blocks, with suitable techniques. Among others, the fast multiple method (FMM) is maybe the most popular for both low and high frequency problems. One more recent approach for integral operators with asymptotically smooth kernels is based on the Adaptive Cross Approximation (ACA) coupled with hierarchical matrix (H-matrix) representations [3],[5]. In contrast with the FMM, where the kernel is approximated by a sum of spherical multipole functions, ACA generates low-rank approximations of far-field blocks from the entries of the original matrix. From an implementation viewpoint, ACA can directly use the computational routines of the existing code without any major change. The specific implementation within our code has been achieved through the HLIBpro library [5]. A useful feature of this library is that it natively supports the assembly of a block H−matrix where the single blocks are H−matrices computed by different routines like in the present case.

The method works as follows. In a first step the degrees of freedom are partitioned and clustered according to a geometrical criterion. Then each cluster pair (σ,τ), corresponding to the sub matrix $A_{(\sigma,\tau)}$ is tested against the admissibility criterion $\min\{\text{diam}(\sigma),\text{diam}(\tau)\} \leq \eta \ \text{dist}(\sigma,\tau)$ where $\text{diam}(\sigma)$ is the cluster diameter, $\text{dist}(\sigma,\tau)$ is the distance between the clusters and η is the admissibility parameter. When the cluster pair satisfies the criterion, the corresponding matrix block is classified as belonging to the far-field, otherwise the clusters are halved and the procedure is applied recursively until the number of elements is larger than a specified threshold. The matrix blocks are then stored with a hierarchical H-matrix structure. The near-field sub-matrices are calculated exactly, whereas the far-field interactions are approximated with the ACA technique. Consider a matrix block $\mathbf{M} \in \mathbb{C}^{m \times n}$ that represents a far-field interaction; in principle, if the singular value decomposition (SVD) is applied to $\mathbf{M}$, only a few singular values are needed to represent the matrix, obtaining the low rank approximation $\widetilde{\mathbf{M}}_k$ such that

$$\left\|\mathbf{M} - \widetilde{\mathbf{M}}_k\right\|_F \leq \alpha\|\mathbf{M}\|_F \quad (5)$$

where $k < m,n$ is the number of singular values used to represent $\mathbf{M}$, α is a specified accuracy and $\| \bullet \|_F$ is the Frobenius norm. The low rank approximation can be obtained in a smarter way without the construction of the SVD by choosing a subset of rows and columns, forming a cross, of the matrix such that

$$\widetilde{\mathbf{M}}_k = \mathbf{U}\mathbf{V}^T, \ \mathbf{U} \in \mathbb{C}^{m \times k}, \mathbf{V} \in \mathbb{C}^{n \times k} \qquad (6)$$

where $\mathbf{V}^T$ is the transpose of vector $\mathbf{V}$.

Since only a reduced number of entries of the original matrix must typically be computed, it can be proved that the
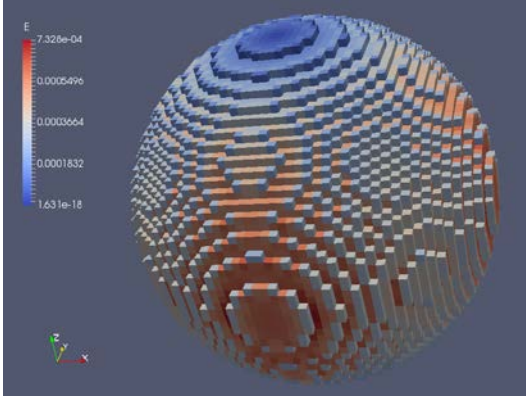

Fig. 5. Example of discretization of the sphere (50 elements along the diameter) used as validation benchmark and map of the surface induced electric field.
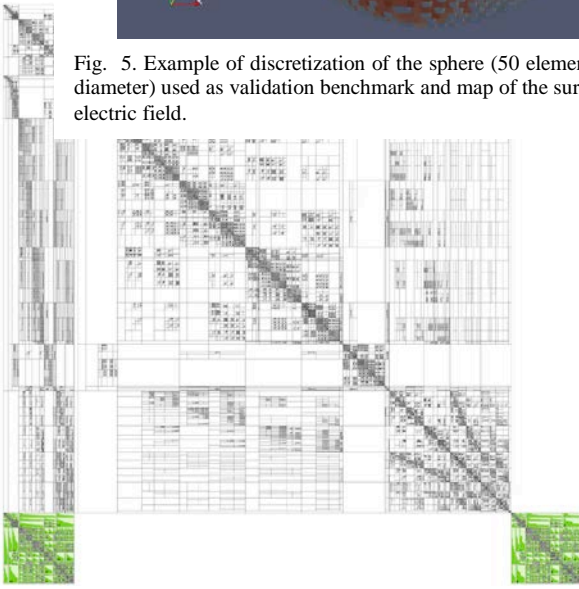

Fig. 3. Low-rank approximation of the linear system (4).

computational cost, as well as the memory consumption, of the matrix partitioning and the ACA approximation have linear-logarithmic complexity.

In contrast to other compression methods, like the FMM, the H-matrix is indeed assembled. Thus, in addition to the matrix-vector multiplication, a special version (called formatted) of H-matrix addition and H-multiplication can be defined [3]. This H-matrix arithmetic can be used to define the inversion operator or the LU decomposition of the H−matrix. When Krylov-type solvers like GMRES are used, it is possible to choose a suitable $\alpha' < \alpha$ and perform an approximate LU decomposition to be used as preconditioner.

In principle, two possibilities for clustering exist: the geometric one used for BEM dense blocks and an algebraic one which could in principle be applied for sparse FEM blocks. However, the admissibility of clusters can be checked only if row and column clusters are all geometric or all sparse, and therefore, by a cyclical application of this argument, all blocks have to be treated with geometric clustering. A typical result of the application of the H-matrix/ACA low-rank approximation technique is shown in Fig. 3.

IV. CODE STRUCTURE

The computer code implementing the proposed formulation with H-matrix/ACA low-rank approximation deserves some explanation due to its particular structure, as schematically
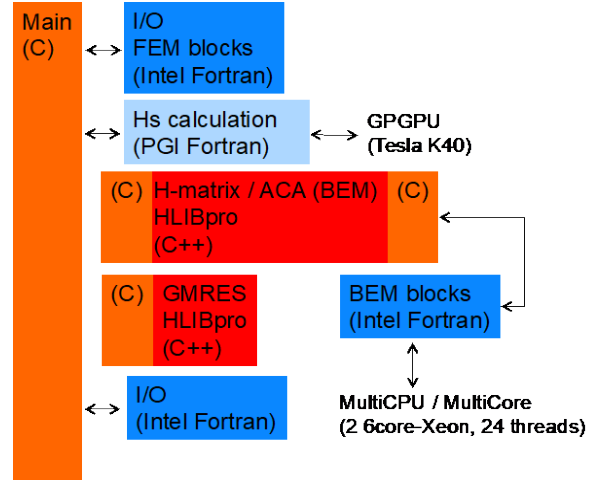

Fig. 4. Code structure

shown in Fig. 4.

First of all, it is noteworthy that the computation of the source terms needs to be parallelized with the PGI Fortran compiler, in order to make use of NVIDIA GPGPU cards. Indeed, the generation of the volume integrals of the source fields $\mathbf{H}_s$ is extremely expensive for the very large meshes typical of this kind of problems and the use of massively parallel computation becomes mandatory. As an example, the time needed to compute the contribution to field $\mathbf{H}_s$ by one source volume element (hexaedron) reduces from 1.1 s (CPU) to 1.7 ms (GPU), with an Intel XEON E5-2650 processor and a GPU NVIDIA Tesla K40.

Furthermore, the generation of the H-matrix/ACA low-rank approximation of the dense blocks is also parallelized thanks to HLIBpro which adopts OpenMP to distribute the computation of the blocks on the individual threads of multicore processors (in our case 24 parallel threads).

Profiling tools have showed that in both parallelized phases the GPGPU and the two Xeon E5645 6-core processors, each capable of performing 2 parallel threads per core, were fully loaded, while during the GMRES iterations the total load was 50% on average.

V. APPLICATION TO BENCHMARKS

As a validation benchmark, the proposed technique is applied to the computation of the electric field and the related specific absorption rate (SAR) produced in a homogeneous sphere (diameter $d = 100$ mm; electrical conductivity $\sigma = 1$ MS/m) subjected to a uniform sinusoidal magnetic field (frequency $f = 1$ kHz). This benchmark was chosen for several reasons: it emulates the exposure of a metallic implant to the MRI gradient coil fields, an analytical solution is available for comparison and validation [7], and finally the number of voxels can be regularly changed enabling controlled testing. The sphere is discretized into voxels (similarly to the anatomical models). In the example of Fig. 5, the diameters along the Cartesian axes

are uniformily subdivided into 50 elements (cubic voxel size: 2 mm). For this benchmark the dense blocks of (4), if fully assembled, would occupy 173 GB of memory each. Table I shows the threshold parameter α, the size and compression ratios (compr.) for the (3,1) and (3,3) dense matrix blocks and the time and iterations required by GMRES. A diagonal preconditioner was used in all these cases. To generate the table, the threshold parameter was reduced until the solver no longer converged. In all converged cases the solution was almost

TABLE I
SPHERE BENCHMARK (50 ELEMENTS ALONG THE DIAMETER)

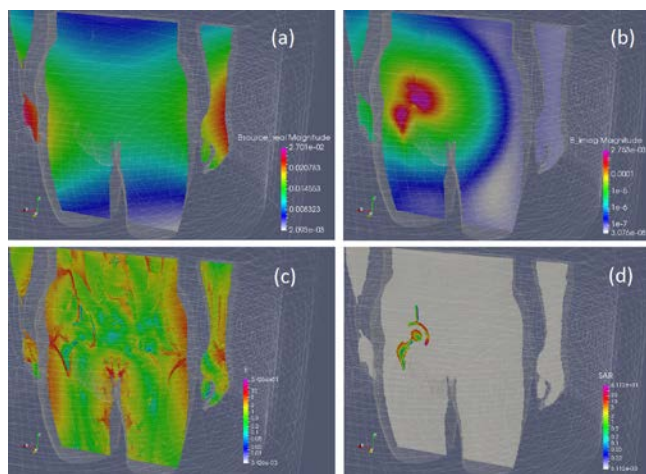| α	Block (3,1)		Block (3,3)		Solver	
	[size, GB]	[compr.]	[size, GB]	[compr.]	[s]	[iter.]
1E-8	7.17	4.13%	7.80	4.49%	12324	1851
1E-6	4.08	2.35%	4.88	2.81%	11562	1851
1E-4	2.55	1.47%	2.53	1.46%	29870	n.c.

Fig. 6. Results for the 'Duke' benchmark problem. All maps are generated in a coronal plane including the hip prosthesis: (a) unperturbed B field (T) as generated by the MRI gradient coils, (b) reaction B field (T) due to the eddy currents in the implant, (c) induced E field (V/m), (d) power density (W/kg) in the metallic implant.

identical with the analytical one, so that results of comparison are not reported for brevity. Almost similar results were obtained by increasing the number of unknowns (100 and 150 elements along the sphere diameter).

TABLE II
"DUKE" BENCHMARK

α	Block (3,1)		Block (3,3)		Solver	
	[size, GB]	[compr.]	[size, GB]	[compr.]	[s]	[iter.]
1E-6	7.77	0.85%	5.08	0.55%	12784	664

VI. APPLICATION TO A MRI DOSIMETRY PROBLEM

The considered computational method was then used to estimate the spatial distributions of the electric and magnetic fields and power deposition in a human body model radiated by the gradient coils of a MRI scanner. In particular, the analysis is performed on the waist of the anatomical human model "Duke" provided by the IT'IS Foundation [8], modified to include a metallic hip prosthesis and discretized into 3,189,606 cubic voxels (size: 2 mm × 2 mm × 2 mm). The electric properties of the tissues are obtained from the database made

available by the same IT'IS Foundation [9]. The problem involves in total 12,884,049 unknowns, divided into 3,313,672 nodal values of the magnetic scalar potential, 9,325,117 edge values of the vector electric potential and 245,260 value of the normal components of scalar potential gradient on the boundary faces. For this second benchmark the dense blocks of (4), if fully assembled, would occupy 914 GB and 922 GB of memory, respectively. They largely dominate the memory occupation of the system. Table II shows the threshold parameter α, the size and compression ratios for the (3,1) and (3,3) dense matrix blocks and the time and iterations required by GMRES. To generate the table, the threshold parameter was set to the minimum allowable value identified with the sphere benchmark. A diagonal preconditioner was used also in this case. The accuracy of the obtained solution was checked against a reference one obtained on a coarser discretization which allowed the solution of the full problem without compression. In this case the compression is even higher than in the previous case. This behaviour is typical of the H-matrix/ACA low-rank approximation technique, which produces increasing benefits as the problem size grows. The obtained spatial distributions of B and E fields in a coronal section of the body model, as well as the power density in the metallic implant, are shown in Fig. 6.

VII. CONCLUSION

This paper has demonstrated the applicability and usefulness of H-matrix techniques for the bioelectromagnetic analysis of large scale human models. The code presented here is currently being applied also to non-MRI dosimetric evaluation cases (e.g. wireless power transfer for automotive applications) and for the simulation of Transcranial Magnetic Stimulation (TMS) for both diagnosis and treatment. On the algorithmic side, work is in progress to implement alternative formulations and preconditioners.

REFERENCES

[1] F. G. Shellock and J. V. Crues, 'MR procedures: Biologic effects, safety and patient care,' *Radiology*, vol. 232, pp. 635–652, Jul. 2004.
[2] O. Bottauscio at al., 'Assessment of computational tools for MRI RF dosimetry by comparison with measurements on a laboratory phantom,' *Phys. Med. Biol.*, vol. 60, pp. 5655-80, July 2015.
[3] W. Hackbusch, 'A sparse matrix arithmetic based on H-matrices. Part I : Introduction to H-matrices,' *Computing*, vol. 62, pp. 89–108, 1999.
[4] S. Kurz, O. Rain, and S. Rjasanow, 'The adaptive cross-approximation technique for the 3-D boundary-element method,' *IEEE Trans. Magn.*, vol. 38, pp. 421–424, Mar. 2002.
[5] R. Kriemann. *HLIBpro*. [Online]. Available: http://www.hlibpro.com.
[6] O. Bottauscio, M. Chiampi, and L. Zilberti, "Massively Parallelized Boundary Element Simulation of Voxel-Based Human Models Exposed to MRI Fields," *IEEE Trans. Magn.*, vol. 50, no. 2, art. 7025504, Feb. 2014.
[7] L. Zilberti, A. Arduino, O. Bottauscio, M. Chiampi, 'The Underestimated Role of Gradient Coils in MRI Safety,' to appear on *Magn. Res. Med.*, DOI 10.1002/mrm.26544.

[8]   A. Christ et al. "The Virtual Family - Development of Surface-based Anatomical Models of Two Adults and Two Children for Dosimetric Simulations," *Phys. Med. Biol.*, vol. 55, pp. N23-38, Jan. 2010.

[9]   IT'IS Foundation (2016). Database for thermal and electromagnetic parameters of biological tissues. [Online]. Available: http://www.itis.ethz.ch/database.