Numerical modeling of biomolecular electrostatic properties by the Element-Free Galerkin Method.

Numerical Modeling of Biomolecular Electrostatic Properties by the Element-Free Galerkin Method

Alessandra Manzin, Domenico Patrizio Ansalone, and Oriano Bottauscio
Istituto Nazionale di Ricerca Metrologica, Torino, Italy

The element-free Galerkin method is applied to the study of the electrostatic behavior of a biomolecule inside a ionic solvent. To this aim, the attention is focused on the solution of the linearized Poisson–Boltzmann equation. The numerical results put in evidence the capability of the proposed approach in approximating the steep gradients of the electrostatic potential arising in the molecular region.

***Index Terms*—Bioelectric phenomena, biological system modeling, element-free Galerkin method (EFGM), meshless methods**

I. Introduction

RECENTLY, there has been a growing interest towards the modeling of the electrostatic behavior of biomolecules (e.g., proteins, nucleic acids) in ionic solutions. The electrostatic interactions between biomolecules and solvent play a fundamental role in the definition of the structure, binding properties and kinetics of such complex systems.

In the biomolecular computing community, a widely used model to describe the electrostatic potential in proximity of a biomolecule is the nonlinear Poisson–Boltzmann (PB) equation [1]. This model represents the biomolecule as a polarized cluster of atoms and the solvent as a continuum dielectric medium, where ions satisfy the Boltzmann distribution. The accurate solution of the PB equation is a very difficult task, due to strong irregularities in the considered geometries, exponential nonlinearity and delta distribution sources. The point-charge singularities give rise to localized steep gradients in the molecular region that require a very fine mesh when the Finite Element Method (FEM) is used. To overcome these limits, we have developed a numerical model based on the element-free Galerkin method (EFGM) [2]. This technique has been proved to efficiently predict local high gradient solutions arising in magneto-hydrodynamics [3], eddy current problems with strong skin effects [4] and electromagnetic-wave scattering [5].

This paper describes the EFGM application to the linearized PB equation, considering a molecular structure whose geometry is derived from the RCSB protein data bank [6]. Particular care is devoted to the treatment of essential boundary conditions and discontinuous derivatives on molecule-solvent interface. The analysis is focused on the influence of EFGM features (node distribution, domain influence size, weight functions) on the numerical accuracy.

II. Numerical Model

For a 1:1 electrolyte, the weak formulation of the linearized PB equation, with test function w, results to be

$$\int_\Omega \varepsilon\nabla\phi\cdot\nabla w dv - \int_{\partial\Omega}\varepsilon\frac{\partial\phi}{\partial n}w ds + \int_\Omega \varepsilon\kappa^2\phi w dv = \sum_{i=1}^{N_c} q_i \int_\Omega \delta_i w dv \tag{1}$$

where ϕ is the electrostatic potential, $\delta_i = \delta(\mathbf{x}-\mathbf{x}_i)$ is the Dirac distribution at point $\mathbf{x}_i$, ε is the dielectric constant, and κ is the Debye-Hückel parameter. When neglecting the ion-exclusion layer [1], the 3-D domain Ω is decomposed into two regions: the molecule Ω_m, including N_c atomic charges q_i, with $\varepsilon_r \sim 1 \div 4$ and $\kappa = 0$, and the solvent region Ω_s, for which $\varepsilon_r \sim 78$ and κ is a function of temperature and ionic strength.

By applying the EFGM, ϕ is approximated in terms of local shape functions $\psi(\mathbf{x})$, associated with a set of N nodes distributed over Ω. It results that

$$\tilde{\phi}(\mathbf{x}) = \sum_{j=1}^{n}\psi_j(\mathbf{x})\tau_j \tag{2}$$

where $n \le N$ is the number of EFGM nodes whose influence domain contains $\mathbf{x}$ and τ_j is the unknown parameter at node j, generally non coincident with the nodal value $\phi(\mathbf{x}_j)$ [2]. Shape functions ψ_j are defined as

$$\psi_j(\mathbf{x}) = \mathbf{p}^T(\mathbf{x})\mathbf{A}^{-1}(\mathbf{x})\mathbf{B}_j(\mathbf{x}) \tag{3}$$

with

$$\begin{cases}\mathbf{A}(\mathbf{x}) = \sum\limits_{g=1}^{n} w(\mathbf{x}-\mathbf{x}_g)\mathbf{p}(\mathbf{x}_g)\mathbf{p}^T(\mathbf{x}_g)\\ \mathbf{B}_j(\mathbf{x}) = w(\mathbf{x}-\mathbf{x}_j)\mathbf{p}(\mathbf{x}_j).\end{cases}$$

In (3), $\mathbf{p}^T(\mathbf{x})$ is a linear basis function ($\mathbf{p}^T(\mathbf{x}) = [1\ x\ y\ z]$) and w is a weight function, which is different from zero only over a limited region around the associated node (influence domain). The continuity properties of w influence the regularity of shape functions, enabling highly continuous approximations. Here, exponential weight functions

$$w(r) = \begin{cases} e^{-(r/\alpha)^2}, & \text{for } r \le 1\\ 0, & \text{for } r > 1\end{cases} \tag{4}$$

are adopted and compared with the following C^2 cubic spline functions:

$$w(r) = \begin{cases} 4r^3 - 4r^2 + 2/2, & \text{for } r \le 1/2\\ -4/3r^3 + 4r^2 - 4r + 4/3, & \text{for } 1/2 < r \le 1\\ 0, & \text{for } r > 1\end{cases} \tag{5}$$

The exponential weight is actually noncontinuous, since it is different from zero at $r = 1$, but numerically it resembles a weight with C^1 continuity or higher. By assuming parameter $\alpha = 0.4$ it results that $w(1) \cong 0.002$.

Considering spherical influence domains, the argument r of the weight function is defined as

$$r = \|\mathbf{x}-\mathbf{x}_j\|/(d_{\max}\cdot c_j) \tag{6}$$

Manuscript received May 20, 2010; accepted September 18, 2010. Date of current version April 22, 2011. Corresponding author: A. Manzin (e-mail: a.manzin@inrim.it).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMAG.2010.2081353

where $d_{\max}$ is a scaling parameter and c_j is here assumed as the minimum distance between node j and its neighboring nodes.

Since EFGM shape functions are not interpolating, special techniques have to be employed to handle essential boundary conditions. The external region is truncated at a sufficiently large distance from the molecular surface, where the potential is negligible. On the boundary $\partial\Omega$, homogeneous Dirichlet conditions are imposed following the "substitution method", which introduces this constraint equation

$$\tilde{\phi}(u_g)|_{\partial\Omega} = \sum_{j=1}^{n} \psi_j(u_g)\tau_j = 0 \tag{7}$$

in correspondence of each meshless node g belonging to $\partial\Omega$.

The high-order regularity of weight and EFGM shape functions is a disadvantage when treating the discontinuity of the normal derivative of potential at the molecular surface Γ. To introduce this discontinuity, an approach based on the "visibility criterion" is here adopted [7]. When constructing weight functions, the influence domains cut by Γ are divided into two parts, belonging to Ω_m or Ω_s. Thus, nodes contained in region $\Omega_m(\Omega_s)$ are only influenced by nodes inside the same region plus the ones belonging to the interface. The nodes on Γ are duplicated, giving rise to two sets of nodes g^m and g^s. To enforce the continuity of the electrostatic potential ϕ and of the normal component of the electric displacement field, the following interface conditions are imposed:

$$\begin{aligned} \tilde{\phi}\left(u_g^m\right)\Big|_\Gamma &= \tilde{\phi}\left(u_g^s\right)\Big|_\Gamma \\ \varepsilon_m \frac{\partial\tilde{\phi}\left(u_g^m\right)}{\partial n}\Big|_\Gamma &= \varepsilon_s \frac{\partial\tilde{\phi}\left(u_g^s\right)}{\partial n}\Big|_\Gamma \end{aligned} \tag{8}$$

where n represents the direction normal to Γ.

It is important to point out that the treatment of boundary and interface conditions through the introduction of (7) and (8) reduces the sparsity of stiffness matrix.

III. Numerical Analysis

The accuracy of EFGM is initially tested computing the electrostatic potential generated by a simple biomolecule, the tyrosine, whose representation is shown in Fig. 1. The tyrosine contains 24 atomic charges and is one of the 20 amino acids used by cells to synthesize proteins. The relative dielectric constant of the molecule is assumed equal to 1 and $\kappa^{-1} \sim 9.6$ Å in the solvent region.

Domain Ω is truncated at a distance from the molecule barycentre corresponding to ∼3 times the molecule average size, where the electrostatic potential is almost negligible. To compute integrals over Ω and define the molecule surface, a mesh of tetrahedra is introduced. The integral points are located following a 4-point standard Gaussian quadrature rule.

The EFGM results are validated by comparison to a reference solution obtained by the Boundary Element Method (BEM), which has been proved to be very accurate in the numerical treatment of the linearized PB equation [8]. The reference solution (see Fig. 1) is computed considering a fine discretization of the molecular surface (∼3500 triangular elements). The analysis is

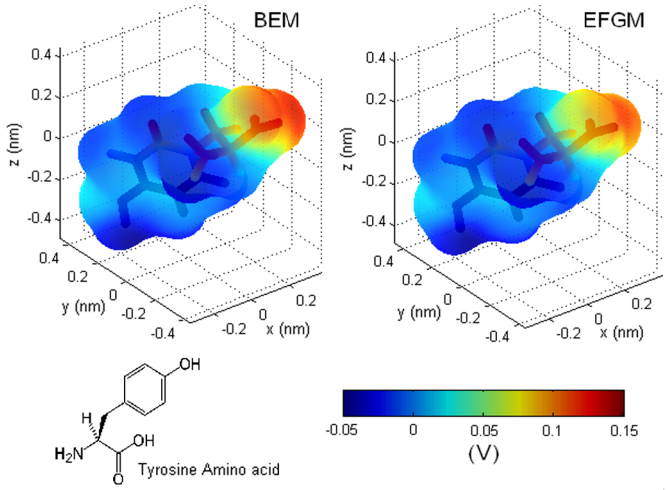

Fig. 1. Surface electrostatic potential maps of tyrosine, computed with BEM and EFGM. The solutions are obtained by considering the same surface tessellation, corresponding to ∼3500 elements. The EFGM nodes are ∼25000 in total.

focused on the role of EFGM node number and location, parameter $d_{\max}$ and weight function type.

A. Role of EFGM Node Number and Location

The EFGM nodes, which can be disjointed from the background mesh, are here placed on the tetrahedron vertices (inside Ω_m and Ω_s) and in correspondence of the charge location (inside Ω_m). It is important to point out that the singularities due to Dirac delta distribution are avoided, since the EFGM shape functions do not satisfy the Kronecker delta criterion. On the molecule surface Γ, the EFGM nodes are positioned on the barycentre of the interface tehaedron faces to allow the computation of the potential normal derivative.

The EFGM solution is computed from a distribution of ∼10400 nodes, assuming exponential weight functions with $\alpha = 0.4$ and $d_{\max} = 2.0$. The quadrature mesh and the node location are derived from a surface mesh with ∼1400 triangular elements. As shown in Figs. 2, which report the electrostatic potential along the Cartesian axes, a good agreement with the BEM reference solution is reached when the EFGM nodes are placed also in correspondence of the charge location. When meshless nodes are not arranged on charge position, the spatial behavior of the electrostatic potential inside the molecule region is not correctly reproduced.

The comparison with a FEM solution, obtained by using the quadrature mesh of EFGM and by imposing homogeneous Dirichlet boundary conditions on $\partial\Omega$, puts in evidence the better accuracy of EFGM. To obtain comparable accuracy, the FEM mesh has to be considerably increased, considering a grid with ∼25000 nodes.

B. Role of Influence Domain Size

The scaling parameter $d_{\max}$ defines the extension of node influence domains and has a strong impact on accuracy [4]. With the node distribution previously considered, if $d_{\max} \leq 1.7$, matrix $\mathbf{A}$, which contributes to shape function construction, can be singular, i.e. non invertible, since a node can have only one neighbor in its influence domain. When $d_{\max} = 1.8 \div 2.2$ and

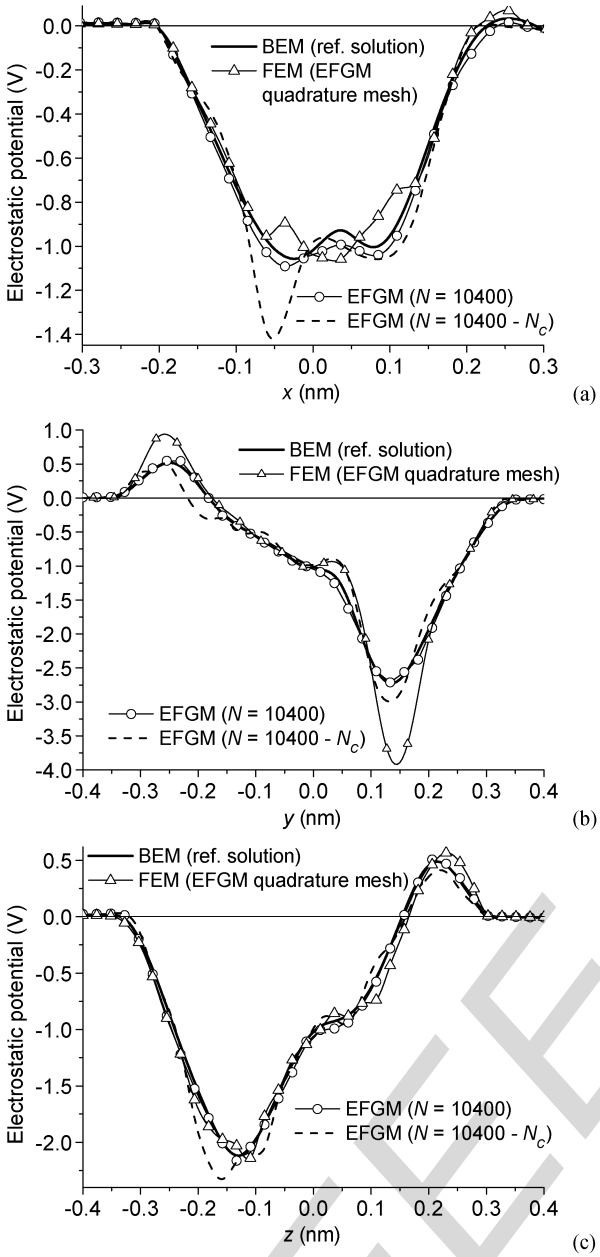

Fig. 2. Electrostatic potential along Cartesian axes, computed by BEM (reference solution), EFGM, and FEM. The EFGM solutions are obtained both including and removing meshless nodes on charge location (N_c). The FEM solution is calculated by employing the EFGM quadrature mesh.

with exponential weight functions, there is a good agreement with the BEM reference solution, mainly inside the molecule, where the potential reaches its highest values.

Larger discrepancies can arise in the solvent region, as put in evidence in Fig. 3, which reports the spatial distribution of the electrostatic potential computed along the y-axis for different values of $d_{\max}$. In the range $1.8 \div 2.2$, this parameter has a weak influence, leading to a quite accurate reconstruction of the solution. When $d_{\max} \geq 2.5$, near the molecule-solvent interface the solution is affected by spatial oscillations, whose amplitude rises with $d_{\max}$. The solution quality can be improved by increasing the node number in the solvent region near the molecule surface.

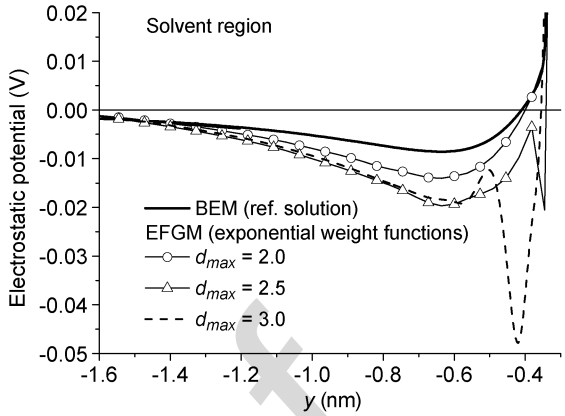

Fig. 3. Electrostatic potential in the solvent region computed along y-axis. The EFGM solution is obtained for different values of parameter $d_{\max}$ considering exponential weight functions with α equal to 0.4.

Parameter $d_{\max}$ has also a strong effect on computational burden and memory requirements. This is put in evidence in Table I, which reports the CPU time required to assemble stiffness matrix and known term, compute optimized LU factorization following the Tinney ordering scheme [9] and solve the system of equations. The most influenced contribution is the first one, which increases of about 80 times by varying $d_{\max}$ from 1.8 to 3.0. The cost of computing integrals in (1) is strongly affected by the number n of EFGM nodes involved in the calculation of shape functions at a quadrature node, which rises with $d_{\max}$. The CPU time required to construct the system matrix is proportional to n^2.

The number of matrix elements augments with $d_{\max}$, due to the overlapping of influence domains and the consequent increase of the interconnections between EFGM nodes. This produces an increment of the CPU time required to compute LU factorization and solution.

It is interesting to note that the BEM solution computed from a surface mesh with ~1400 elements requires a memory allocation comparable to the one obtained with the EFGM when $d_{\max}$ is equal to 3.0. Also with BEM the system assembling can be very time consuming due to the high number of involved boundary integral operations.

C. Role of Weight Function

Weight functions play an important role in the performance of EFGM. When adopting exponential weight functions, accurate results in the molecule region can be obtained by varying α between 0.3 and 0.5. For higher values of α, the strong discontinuities at the boundary of weight function support (at $r = 1$) introduce abrupt variations in the potential, giving rise to inacceptable results.

In the solvent region, near the molecule surface, the most accurate prediction is obtained by imposing $\alpha = 0.5$ [10], even if this value leads to $w(1) \cong 0.02$ (Fig. 4).

By comparing exponential to cubic spline weight functions, the quality of the solution in the solvent region is better when using the former ones, as shown by the graphs of Fig. 5

TABLE I
ROLE OF INFLUENCE DOMAIN SIZE ON CPU TIME AND MEMORY REQUIREMENTS. PERFORMANCES ARE REFERRED TO THE CASE WITH $d_{\max} = 1.8$

d_{max}	CPU time (p.u.)			N. of matrix elements	
	Assembling	LU Factoriz.	Solution	Before LU	After LU
1.8	1	1	1	N_e ~$1{\cdot}10^6$	~29 N_e
2.0	1.7	1.4	1.3	~1.4 N_e	~34 N_e
2.2	3.6	2.0	1.6	~1.9 N_e	~40 N_e
2.5	12	2.4	2.0	~2.7 N_e	~43 N_e
3.0	78	3.6	3.0	~4.7 N_e	~52 N_e

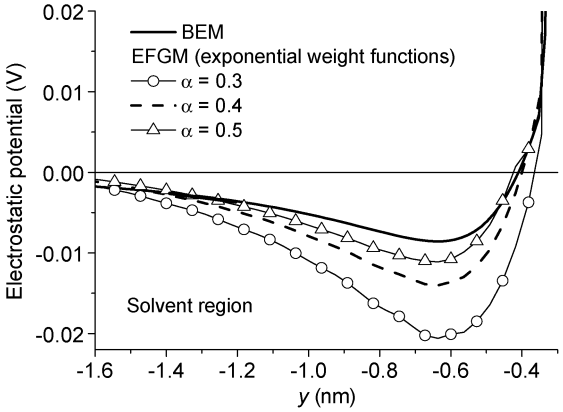

Fig. 4. Electrostatic potential in the solvent region computed along y-axis. The EFGM solution is obtained considering exponential weight functions and $d_{\max} = 2.0$ for different values of parameter α.

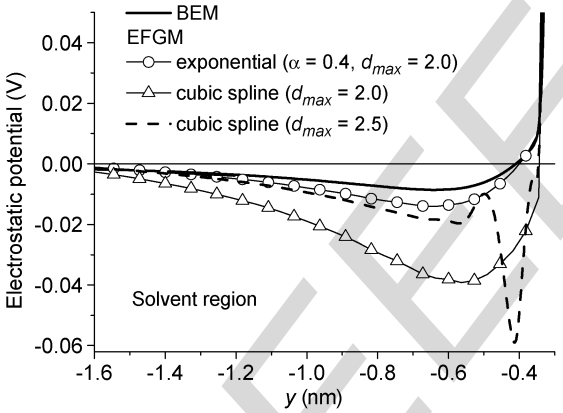

Fig. 5. Electrostatic potential in the solvent region computed along y-axis. The EFGM solution is also obtained considering cubic spline weight functions and parameter $d_{\max}$ equal to 2.0 and 2.5.

IV. CONCLUSIONS

The EFGM has been applied for the first time to the solution of the linearized PB equation in 3-D domains. Like BEM and differently from FEM, the EFGM is able to accurately reconstruct the spatial distribution of electrostatic potential inside molecule region also with a low number of unknowns. To reach this goal, the EFGM nodes have to be placed on charge location too.

Differently from BEM, the open-boundary conditions are not exactly treated and ad-hoc strategies have to be used to handle discontinuities in the normal derivative of the potential at the molecular surface. Concerning memory allocation, the EFGM is more competitive than BEM and it could be simply adapted for the treatment of the non linear PB equation.

The EFGM is less advantageous than FEM in terms of CPU usage and memory requirements. First, the calculation of integrals in the weak formulation involves a matrix inversion at each quadrature point, leading to high computational cost in the system assembling. Second, the introduction of additional equations to handle interface and boundary conditions reduces the sparsity of stiffness matrix. The performances can be improved by using parallel computing techniques.

The study has also put in evidence the importance of a suitable choice of weight function and influence domain size in the computation of the electrostatic potential in the solvent region. The accurate prediction of the potential spatial distribution around the molecule surface is fundamental in the determination of the solvation energy and in the analysis of binding processes between complex molecular structures.

ACKNOWLEDGMENT

The work was supported by the Regione Piemonte (Italy) under Project "Metrology on a cellular and macromolecular scale for regenerative medicine (METREGEN)", Bando Converging technologies.

REFERENCES

[1] B. K. Lu *et al.*, "Review article. Recent progress in numerical methods for the Poisson–Boltzmann equation in biophysical applications," *Commun. Comput. Phys.*, vol. 3, no. 5, pp. 973–1009, May 2008.
[2] T. Belytschko and J. Dolbow, "An introduction to programming the meshless element free Galerkin method," *Arch. Comput. Meth. Eng.*, vol. 5, pp. 207–241, 1998.
[3] S. L. L. Verardi, J. M. Machado, and Y. Shiyou, "The application of interpolating MLS approximations to the analysis of MHD flows," *Finite Elem. Anal. Des.*, vol. 39, pp. 1173–1187, 2003.
[4] O. Bottauscio, M. Chiampi, and A. Manzin, "Element-free Galerkin method in eddy-current problems with ferromagnetic media," *IEEE Trans. Magn.*, vol. 42, no. 5, pp. 1577–1584, May 2006.
[5] A. Manzin and O. Bottauscio[illegible]ment-free Galerkin method for the analysis of electromagnetic-wave scattering," *IEEE Trans. Magn.*, vol. 44, no. 6, pp. 1366–1369, Jun. 2008.
[6] [Online]. Available: http://www.pdb.org/pdb/home/home.do
[7] Y. C. Cai and H. H. Zhu[illegible]ct imposition of essential boundary conditions and treatment of material discontinuities in the EFG method," *Comput. Mech.*, vol. 34, pp. 330–338, 2004.
[8] B. Lu, D. Zhang, and J. A. McCammon, "Computation of electrostatic forces between solvated molecules determined by the Poisson–Boltzmann equation using a boundary element method," *J. Chem. Phys.*, vol. 122, p. 214102, 2005.
[9] E. C. Ogbuobiri, W. F. Tinney, and J. W. Walker, "Sparsity-directed decomposition for Gaussian elimination on matrices," *IEEE Trans. Power App. Syst.*, vol. PAS-89, no. 1, pp. 141–150, 1970.
[10] L. Xuan, Z. Zeng, B. Shanker, and L. Udpa, "Element-free Galerkin method for static and quasi-static electromagnetic field computation," *IEEE Trans. Magn.*, vol. 40, no. 1, pp. 12–20, Jan. 2004.

Numerical Modeling of Biomolecular Electrostatic Properties by the Element-Free Galerkin Method

Alessandra Manzin, Domenico Patrizio Ansalone, and Oriano Bottauscio
Istituto Nazionale di Ricerca Metrologica, Torino, Italy

The element-free Galerkin method is applied to the study of the electrostatic behavior of a biomolecule inside a ionic solvent. To this aim, the attention is focused on the solution of the linearized Poisson–Boltzmann equation. The numerical results put in evidence the capability of the proposed approach in approximating the steep gradients of the electrostatic potential arising in the molecular region.

***Index Terms*—Bioelectric phenomena, biological system modeling, element-free Galerkin method (EFGM), meshless methods**

I. Introduction

RECENTLY, there has been a growing interest towards the modeling of the electrostatic behavior of biomolecules (e.g., proteins, nucleic acids) in ionic solutions. The electrostatic interactions between biomolecules and solvent play a fundamental role in the definition of the structure, binding properties and kinetics of such complex systems.

In the biomolecular computing community, a widely used model to describe the electrostatic potential in proximity of a biomolecule is the nonlinear Poisson–Boltzmann (PB) equation [1]. This model represents the biomolecule as a polarized cluster of atoms and the solvent as a continuum dielectric medium, where ions satisfy the Boltzmann distribution. The accurate solution of the PB equation is a very difficult task, due to strong irregularities in the considered geometries, exponential nonlinearity and delta distribution sources. The point-charge singularities give rise to localized steep gradients in the molecular region that require a very fine mesh when the Finite Element Method (FEM) is used. To overcome these limits, we have developed a numerical model based on the element-free Galerkin method (EFGM) [2]. This technique has been proved to efficiently predict local high gradient solutions arising in magneto-hydrodynamics [3], eddy current problems with strong skin effects [4] and electromagnetic-wave scattering [5].

This paper describes the EFGM application to the linearized PB equation, considering a molecular structure whose geometry is derived from the RCSB protein data bank [6]. Particular care is devoted to the treatment of essential boundary conditions and discontinuous derivatives on molecule-solvent interface. The analysis is focused on the influence of EFGM features (node distribution, domain influence size, weight functions) on the numerical accuracy.

II. Numerical Model

For a 1:1 electrolyte, the weak formulation of the linearized PB equation, with test function w, results to be

$$\int_\Omega \varepsilon\nabla\phi\cdot\nabla w dv - \int_{\partial\Omega}\varepsilon\frac{\partial\phi}{\partial n}w ds + \int_\Omega \varepsilon\kappa^2\phi w dv = \sum_{i=1}^{N_c} q_i \int_\Omega \delta_i w dv \tag{1}$$

where ϕ is the electrostatic potential, $\delta_i = \delta(\mathbf{x}-\mathbf{x}_i)$ is the Dirac distribution at point $\mathbf{x}_i$, ε is the dielectric constant, and κ is the Debye-Hückel parameter. When neglecting the ion-exclusion layer [1], the 3-D domain Ω is decomposed into two regions: the molecule Ω_m, including N_c atomic charges q_i, with $\varepsilon_r \sim 1 \div 4$ and $\kappa = 0$, and the solvent region Ω_s, for which $\varepsilon_r \sim 78$ and κ is a function of temperature and ionic strength.

By applying the EFGM, ϕ is approximated in terms of local shape functions $\psi(\mathbf{x})$, associated with a set of N nodes distributed over Ω. It results that

$$\tilde{\phi}(\mathbf{x}) = \sum_{j=1}^{n}\psi_j(\mathbf{x})\tau_j \tag{2}$$

where $n \le N$ is the number of EFGM nodes whose influence domain contains $\mathbf{x}$ and τ_j is the unknown parameter at node j, generally non coincident with the nodal value $\phi(\mathbf{x}_j)$ [2]. Shape functions ψ_j are defined as

$$\psi_j(\mathbf{x}) = \mathbf{p}^T(\mathbf{x})\mathbf{A}^{-1}(\mathbf{x})\mathbf{B}_j(\mathbf{x}) \tag{3}$$

with

$$\begin{cases}\mathbf{A}(\mathbf{x}) = \sum\limits_{g=1}^{n} w(\mathbf{x}-\mathbf{x}_g)\mathbf{p}(\mathbf{x}_g)\mathbf{p}^T(\mathbf{x}_g)\\ \mathbf{B}_j(\mathbf{x}) = w(\mathbf{x}-\mathbf{x}_j)\mathbf{p}(\mathbf{x}_j).\end{cases}$$

In (3), $\mathbf{p}^T(\mathbf{x})$ is a linear basis function ($\mathbf{p}^T(\mathbf{x}) = [1\ x\ y\ z]$) and w is a weight function, which is different from zero only over a limited region around the associated node (influence domain). The continuity properties of w influence the regularity of shape functions, enabling highly continuous approximations. Here, exponential weight functions

$$w(r) = \begin{cases} e^{-(r/\alpha)^2}, & \text{for } r \le 1\\ 0, & \text{for } r > 1\end{cases} \tag{4}$$

are adopted and compared with the following C^2 cubic spline functions:

$$w(r) = \begin{cases} 4r^3 - 4r^2 + 2/2, & \text{for } r \le 1/2\\ -4/3r^3 + 4r^2 - 4r + 4/3, & \text{for } 1/2 < r \le 1\\ 0, & \text{for } r > 1\end{cases} \tag{5}$$

The exponential weight is actually noncontinuous, since it is different from zero at $r = 1$, but numerically it resembles a weight with C^1 continuity or higher. By assuming parameter $\alpha = 0.4$ it results that $w(1) \cong 0.002$.

Considering spherical influence domains, the argument r of the weight function is defined as

$$r = \|\mathbf{x}-\mathbf{x}_j\|/(d_{\max}\cdot c_j) \tag{6}$$

Manuscript received May 20, 2010; accepted September 18, 2010. Date of current version April 22, 2011. Corresponding author: A. Manzin (e-mail: a.manzin@inrim.it).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMAG.2010.2081353

where $d_{\max}$ is a scaling parameter and c_j is here assumed as the minimum distance between node j and its neighboring nodes.

Since EFGM shape functions are not interpolating, special techniques have to be employed to handle essential boundary conditions. The external region is truncated at a sufficiently large distance from the molecular surface, where the potential is negligible. On the boundary $\partial\Omega$, homogeneous Dirichlet conditions are imposed following the "substitution method", which introduces this constraint equation

$$\tilde{\phi}(u_g)|_{\partial\Omega} = \sum_{j=1}^{n} \psi_j(u_g)\tau_j = 0 \tag{7}$$

in correspondence of each meshless node g belonging to $\partial\Omega$.

The high-order regularity of weight and EFGM shape functions is a disadvantage when treating the discontinuity of the normal derivative of potential at the molecular surface Γ. To introduce this discontinuity, an approach based on the "visibility criterion" is here adopted [7]. When constructing weight functions, the influence domains cut by Γ are divided into two parts, belonging to Ω_m or Ω_s. Thus, nodes contained in region $\Omega_m(\Omega_s)$ are only influenced by nodes inside the same region plus the ones belonging to the interface. The nodes on Γ are duplicated, giving rise to two sets of nodes g^m and g^s. To enforce the continuity of the electrostatic potential ϕ and of the normal component of the electric displacement field, the following interface conditions are imposed:

$$\begin{aligned} \tilde{\phi}\left(u_g^m\right)\Big|_\Gamma &= \tilde{\phi}\left(u_g^s\right)\Big|_\Gamma \\ \varepsilon_m \frac{\partial\tilde{\phi}\left(u_g^m\right)}{\partial n}\Big|_\Gamma &= \varepsilon_s \frac{\partial\tilde{\phi}\left(u_g^s\right)}{\partial n}\Big|_\Gamma \end{aligned} \tag{8}$$

where n represents the direction normal to Γ.

It is important to point out that the treatment of boundary and interface conditions through the introduction of (7) and (8) reduces the sparsity of stiffness matrix.

III. Numerical Analysis

The accuracy of EFGM is initially tested computing the electrostatic potential generated by a simple biomolecule, the tyrosine, whose representation is shown in Fig. 1. The tyrosine contains 24 atomic charges and is one of the 20 amino acids used by cells to synthesize proteins. The relative dielectric constant of the molecule is assumed equal to 1 and $\kappa^{-1} \sim 9.6$ Å in the solvent region.

Domain Ω is truncated at a distance from the molecule barycentre corresponding to $\sim$3 times the molecule average size, where the electrostatic potential is almost negligible. To compute integrals over Ω and define the molecule surface, a mesh of tetrahedra is introduced. The integral points are located following a 4-point standard Gaussian quadrature rule.

The EFGM results are validated by comparison to a reference solution obtained by the Boundary Element Method (BEM), which has been proved to be very accurate in the numerical treatment of the linearized PB equation [8]. The reference solution (see Fig. 1) is computed considering a fine discretization of the molecular surface ($\sim$3500 triangular elements). The analysis is focused on the role of EFGM node number and location, parameter $d_{\max}$ and weight function type.

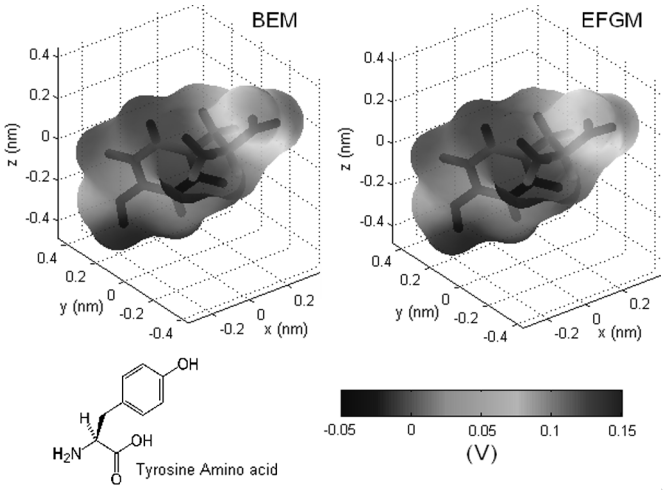

Fig. 1. Surface electrostatic potential maps of tyrosine, computed with BEM and EFGM. The solutions are obtained by considering the same surface tessellation, corresponding to $\sim$3500 elements. The EFGM nodes are $\sim$25000 in total.

A. Role of EFGM Node Number and Location

The EFGM nodes, which can be disjointed from the background mesh, are here placed on the tetrahedron vertices (inside Ω_m and Ω_s) and in correspondence of the charge location (inside Ω_m). It is important to point out that the singularities due to Dirac delta distribution are avoided, since the EFGM shape functions do not satisfy the Kronecker delta criterion. On the molecule surface Γ, the EFGM nodes are positioned on the barycentre of the interface tehaedron faces to allow the computation of the potential normal derivative.

The EFGM solution is computed from a distribution of $\sim$10400 nodes, assuming exponential weight functions with $\alpha = 0.4$ and $d_{\max} = 2.0$. The quadrature mesh and the node location are derived from a surface mesh with $\sim$1400 triangular elements. As shown in Figs. 2, which report the electrostatic potential along the Cartesian axes, a good agreement with the BEM reference solution is reached when the EFGM nodes are placed also in correspondence of the charge location. When meshless nodes are not arranged on charge position, the spatial behavior of the electrostatic potential inside the molecule region is not correctly reproduced.

The comparison with a FEM solution, obtained by using the quadrature mesh of EFGM and by imposing homogeneous Dirichlet boundary conditions on $\partial\Omega$, puts in evidence the better accuracy of EFGM. To obtain comparable accuracy, the FEM mesh has to be considerably increased, considering a grid with $\sim$25000 nodes.

B. Role of Influence Domain Size

The scaling parameter $d_{\max}$ defines the extension of node influence domains and has a strong impact on accuracy [4]. With the node distribution previously considered, if $d_{\max} \leq 1.7$, matrix $\mathbf{A}$, which contributes to shape function construction, can be singular, i.e. non invertible, since a node can have only one neighbor in its influence domain. When $d_{\max} = 1.8 \div 2.2$ and

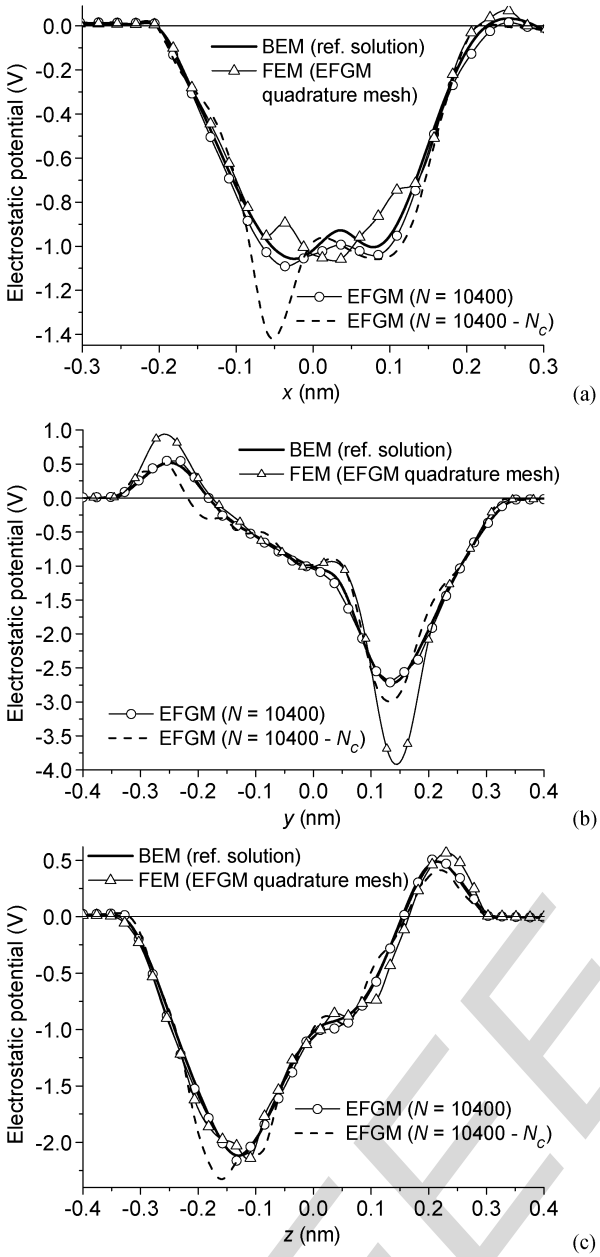

Fig. 2. Electrostatic potential along Cartesian axes, computed by BEM (reference solution), EFGM, and FEM. The EFGM solutions are obtained both including and removing meshless nodes on charge location (N_c). The FEM solution is calculated by employing the EFGM quadrature mesh.

with exponential weight functions, there is a good agreement with the BEM reference solution, mainly inside the molecule, where the potential reaches its highest values.

Larger discrepancies can arise in the solvent region, as put in evidence in Fig. 3, which reports the spatial distribution of the electrostatic potential computed along the y-axis for different values of $d_{\max}$. In the range $1.8 \div 2.2$, this parameter has a weak influence, leading to a quite accurate reconstruction of the solution. When $d_{\max} \geq 2.5$, near the molecule-solvent interface the solution is affected by spatial oscillations, whose amplitude rises with $d_{\max}$. The solution quality can be improved by increasing the node number in the solvent region near the molecule surface.

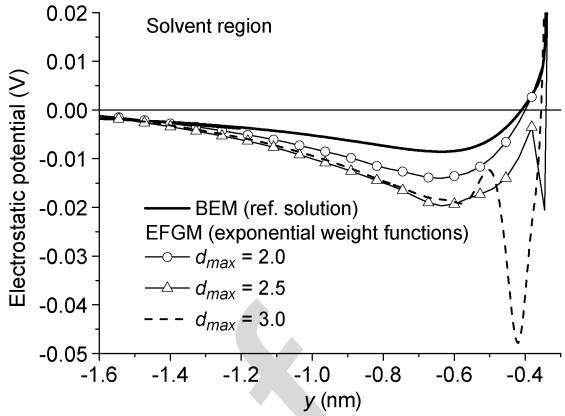

Fig. 3. Electrostatic potential in the solvent region computed along y-axis. The EFGM solution is obtained for different values of parameter $d_{\max}$ considering exponential weight functions with α equal to 0.4.

Parameter $d_{\max}$ has also a strong effect on computational burden and memory requirements. This is put in evidence in Table I, which reports the CPU time required to assemble stiffness matrix and known term, compute optimized LU factorization following the Tinney ordering scheme [9] and solve the system of equations. The most influenced contribution is the first one, which increases of about 80 times by varying $d_{\max}$ from 1.8 to 3.0. The cost of computing integrals in (1) is strongly affected by the number n of EFGM nodes involved in the calculation of shape functions at a quadrature node, which rises with $d_{\max}$. The CPU time required to construct the system matrix is proportional to n^2.

The number of matrix elements augments with $d_{\max}$, due to the overlapping of influence domains and the consequent increase of the interconnections between EFGM nodes. This produces an increment of the CPU time required to compute LU factorization and solution.

It is interesting to note that the BEM solution computed from a surface mesh with $\sim$1400 elements requires a memory allocation comparable to the one obtained with the EFGM when $d_{\max}$ is equal to 3.0. Also with BEM the system assembling can be very time consuming due to the high number of involved boundary integral operations.

C. Role of Weight Function

Weight functions play an important role in the performance of EFGM. When adopting exponential weight functions, accurate results in the molecule region can be obtained by varying α between 0.3 and 0.5. For higher values of α, the strong discontinuities at the boundary of weight function support (at $r = 1$) introduce abrupt variations in the potential, giving rise to inacceptable results.

In the solvent region, near the molecule surface, the most accurate prediction is obtained by imposing $\alpha = 0.5$ [10], even if this value leads to $w(1) \cong 0.02$ (Fig. 4).

By comparing exponential to cubic spline weight functions, the quality of the solution in the solvent region is better when using the former ones, as shown by the graphs of Fig. 5

TABLE I
ROLE OF INFLUENCE DOMAIN SIZE ON CPU TIME AND MEMORY REQUIREMENTS. PERFORMANCES ARE REFERRED TO THE CASE WITH $d_{\max} = 1.8$

d_{max}	CPU time (p.u.)			N. of matrix elements	
	Assembling	LU Factoriz.	Solution	Before LU	After LU
1.8	1	1	1	N_e ~$1{\cdot}10^6$	~29 N_e
2.0	1.7	1.4	1.3	~1.4 N_e	~34 N_e
2.2	3.6	2.0	1.6	~1.9 N_e	~40 N_e
2.5	12	2.4	2.0	~2.7 N_e	~43 N_e
3.0	78	3.6	3.0	~4.7 N_e	~52 N_e

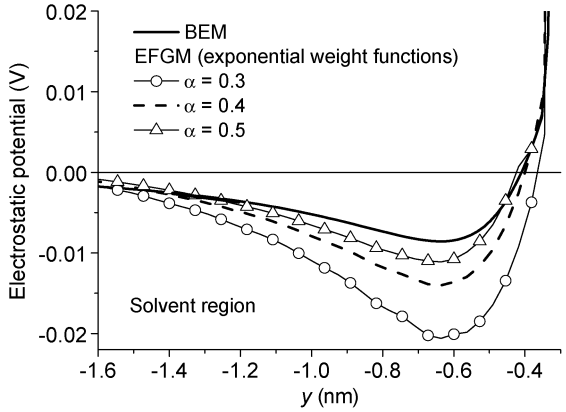

Fig. 4. Electrostatic potential in the solvent region computed along y-axis. The EFGM solution is obtained considering exponential weight functions and $d_{\max} = 2.0$ for different values of parameter α.

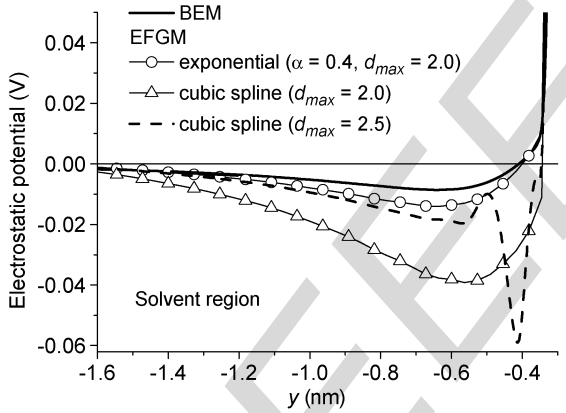

Fig. 5. Electrostatic potential in the solvent region computed along y-axis. The EFGM solution is also obtained considering cubic spline weight functions and parameter $d_{\max}$ equal to 2.0 and 2.5.

IV. CONCLUSIONS

The EFGM has been applied for the first time to the solution of the linearized PB equation in 3-D domains. Like BEM and differently from FEM, the EFGM is able to accurately reconstruct the spatial distribution of electrostatic potential inside molecule region also with a low number of unknowns. To reach this goal, the EFGM nodes have to be placed on charge location too.

Differently from BEM, the open-boundary conditions are not exactly treated and ad-hoc strategies have to be used to handle discontinuities in the normal derivative of the potential at the molecular surface. Concerning memory allocation, the EFGM is more competitive than BEM and it could be simply adapted for the treatment of the non linear PB equation.

The EFGM is less advantageous than FEM in terms of CPU usage and memory requirements. First, the calculation of integrals in the weak formulation involves a matrix inversion at each quadrature point, leading to high computational cost in the system assembling. Second, the introduction of additional equations to handle interface and boundary conditions reduces the sparsity of stiffness matrix. The performances can be improved by using parallel computing techniques.

The study has also put in evidence the importance of a suitable choice of weight function and influence domain size in the computation of the electrostatic potential in the solvent region. The accurate prediction of the potential spatial distribution around the molecule surface is fundamental in the determination of the solvation energy and in the analysis of binding processes between complex molecular structures.

ACKNOWLEDGMENT

The work was supported by the Regione Piemonte (Italy) under Project "Metrology on a cellular and macromolecular scale for regenerative medicine (METREGEN)", Bando Converging technologies.

REFERENCES

[1] B. K. Lu *et al.*, "Review article. Recent progress in numerical methods for the Poisson–Boltzmann equation in biophysical applications," *Commun. Comput. Phys.*, vol. 3, no. 5, pp. 973–1009, May 2008.
[2] T. Belytschko and J. Dolbow, "An introduction to programming the meshless element free Galerkin method," *Arch. Comput. Meth. Eng.*, vol. 5, pp. 207–241, 1998.
[3] S. L. L. Verardi, J. M. Machado, and Y. Shiyou, "The application of interpolating MLS approximations to the analysis of MHD flows," *Finite Elem. Anal. Des.*, vol. 39, pp. 1173–1187, 2003.
[4] O. Bottauscio, M. Chiampi, and A. Manzin, "Element-free Galerkin method in eddy-current problems with ferromagnetic media," *IEEE Trans. Magn.*, vol. 42, no. 5, pp. 1577–1584, May 2006.
[5] A. Manzin and O. Bottauscioc, "Element-free Galerkin method for the analysis of electromagnetic-wave scattering," *IEEE Trans. Magn.*, vol. 44, no. 6, pp. 1366–1369, Jun. 2008.
[6] [Online]. Available: http://www.pdb.org/pdb/home/home.do
[7] Y. C. Cai and H. H. Zhuc, "Direct imposition of essential boundary conditions and treatment of material discontinuities in the EFG method," *Comput. Mech.*, vol. 34, pp. 330–338, 2004.
[8] B. Lu, D. Zhang, and J. A. McCammon, "Computation of electrostatic forces between solvated molecules determined by the Poisson–Boltzmann equation using a boundary element method," *J. Chem. Phys.*, vol. 122, p. 214102, 2005.
[9] E. C. Ogbuobiri, W. F. Tinney, and J. W. Walker, "Sparsity-directed decomposition for Gaussian elimination on matrices," *IEEE Trans. Power App. Syst.*, vol. PAS-89, no. 1, pp. 141–150, 1970.
[10] L. Xuan, Z. Zeng, B. Shanker, and L. Udpa, "Element-free Galerkin method for static and quasi-static electromagnetic field computation," *IEEE Trans. Magn.*, vol. 40, no. 1, pp. 12–20, Jan. 2004.