Tesi di Laurea Magistrale in
Mathematical Engineering - Ingegneria Matematica

POLITECNICO
MILANO 1863

A Scalable Solver for the Linearized Poisson-Boltzmann Equation on Cartesian Grids with Hierarchical Local Refinement

Advisor:
Prof. Carlo de Falco

Co-advisors:
Dr. Walter Rocchia, Dr. Sergio Decherchi

Candidate:
Martina Politi, 953448

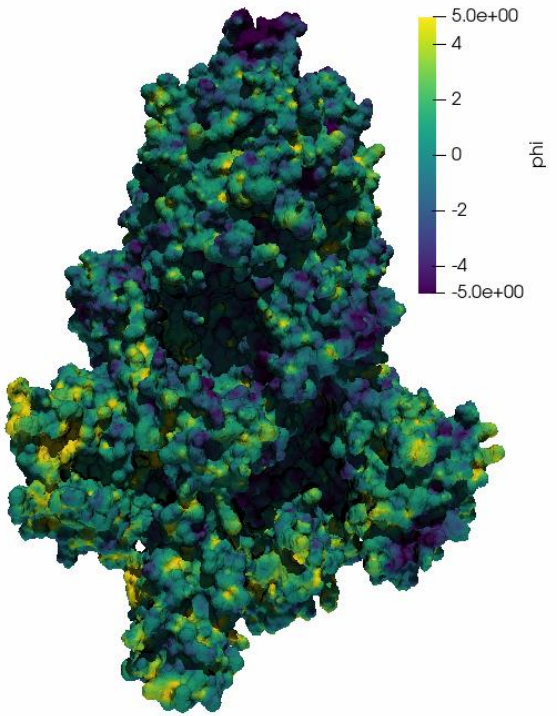
Academic year:
2020-2021

Introduction

Development of **modern calculus tools** for applications in **Molecular Biology**

Implemention of **easy_pbe solver** as an alternative to Delphi

Introduction



SARS-CoV-2, spike protein, ectodomain structure open state.

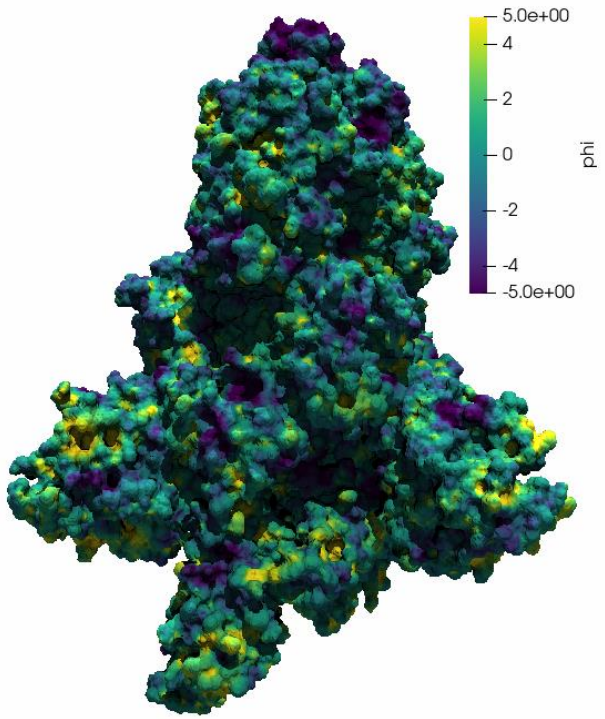
Computation of the **Electrostatic Potential**

↓

Drug discovery and **drug repurposing**

↓

Application to **COVID-19** and new diseases

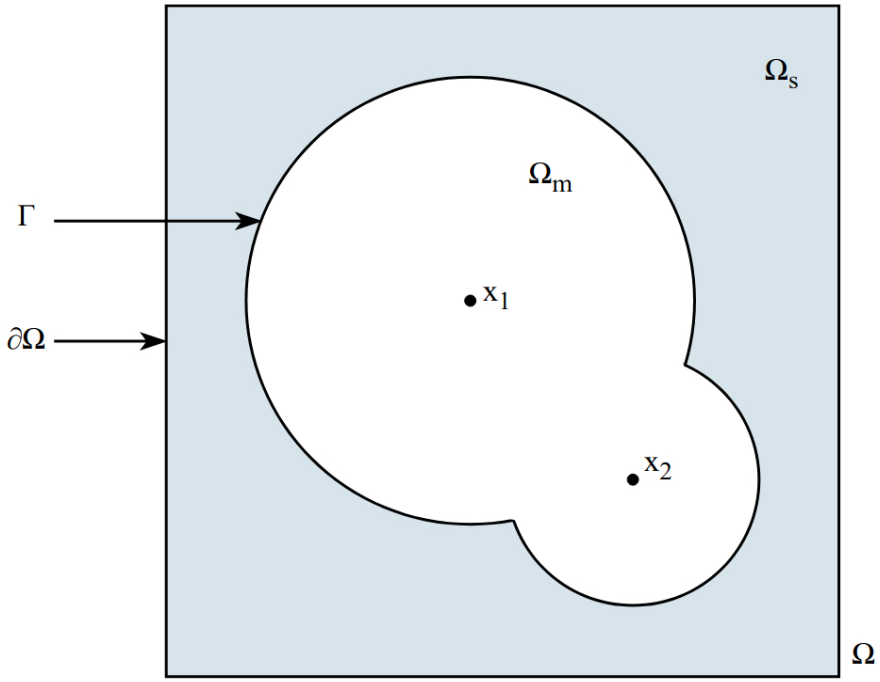


SARS-CoV-2, spike protein, Omicron variant of concern.

The Poisson-Boltzmann Equation

Nondimensional form of the PBE:

$$\begin{cases} -\nabla \cdot (\varepsilon_0 \, \varepsilon_r(x) \, \nabla\varphi(x)) = \rho^s + \rho^f, & in \; \Omega \\ \varphi(x) = g(x), & on \; \partial\Omega \end{cases}$$



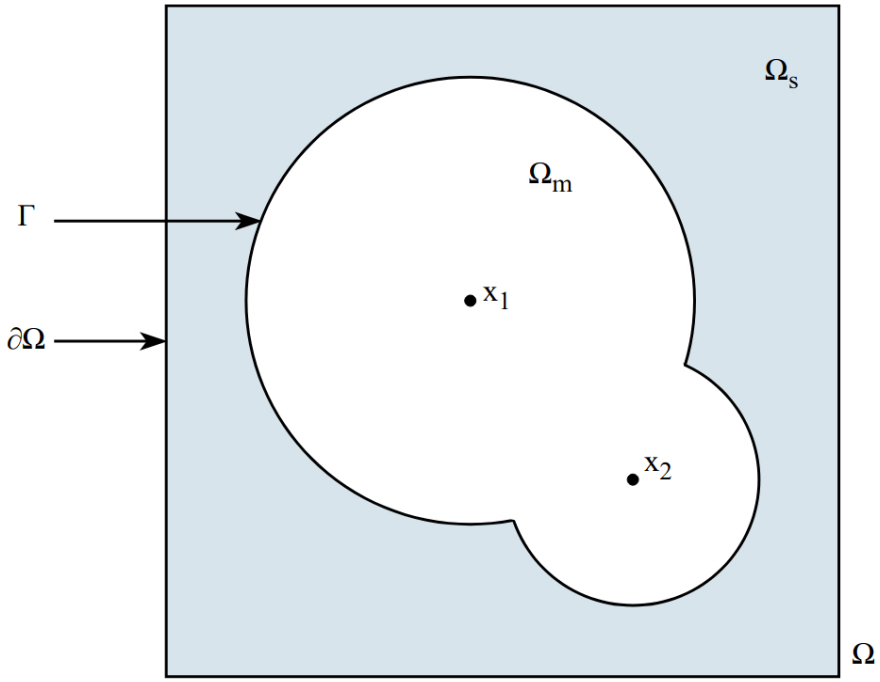
The Poisson-Boltzmann Equation

Nondimensional form of the PBE:

$$\begin{cases} -\nabla \cdot (\varepsilon_0 \, \varepsilon_r(x) \, \nabla\varphi(x)) = \rho^s + \rho^f, & in \ \Omega \\ \varphi(x) = g(x), & on \ \partial\Omega \end{cases}$$

Relative dielectric permittivity:

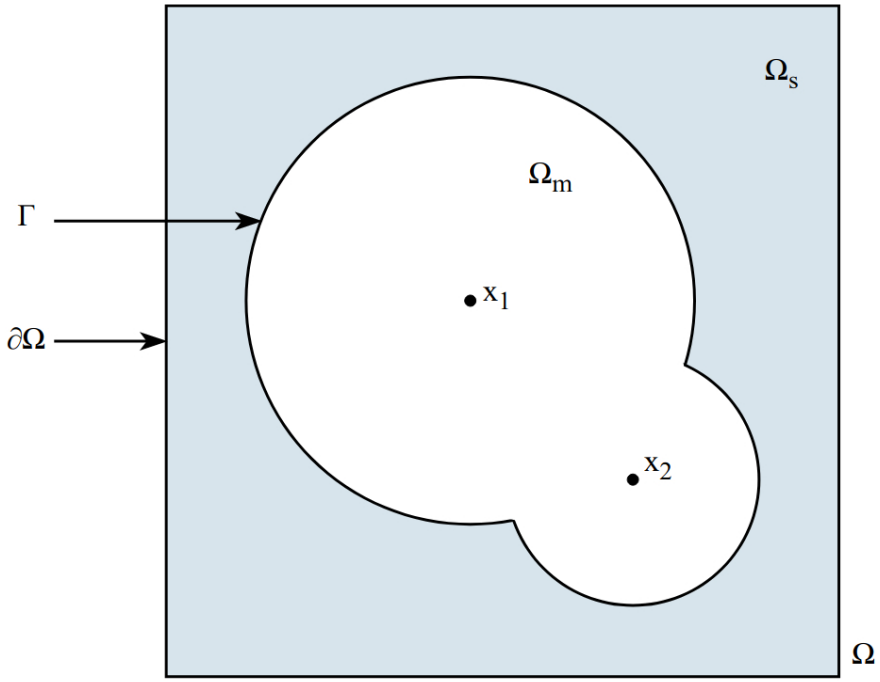
$$\begin{cases} \varepsilon_r(x) = \varepsilon_s(x), & in \ \Omega_s \\ \varepsilon_r(x) = \varepsilon_m, & in \ \Omega_m \end{cases}$$



The Poisson-Boltzmann Equation

Nondimensional form of the PBE:

$$\begin{cases} -\nabla \cdot (\varepsilon_0 \, \varepsilon_r(x) \, \nabla\varphi(x)\,) = \rho^s + \rho^f, & in\ \Omega \\ \varphi(x) = g(x), & on\ \partial\Omega \end{cases}$$



Relative dielectric permittivity:

$$\begin{cases} \varepsilon_r(x) = \varepsilon_s(x), & in\ \Omega_s \\ \varepsilon_r(x) = \varepsilon_m, & in\ \Omega_m \end{cases}$$

Fixed charge density:

$$\rho^f = \sum_i q_i \, \delta(x - x_i)$$

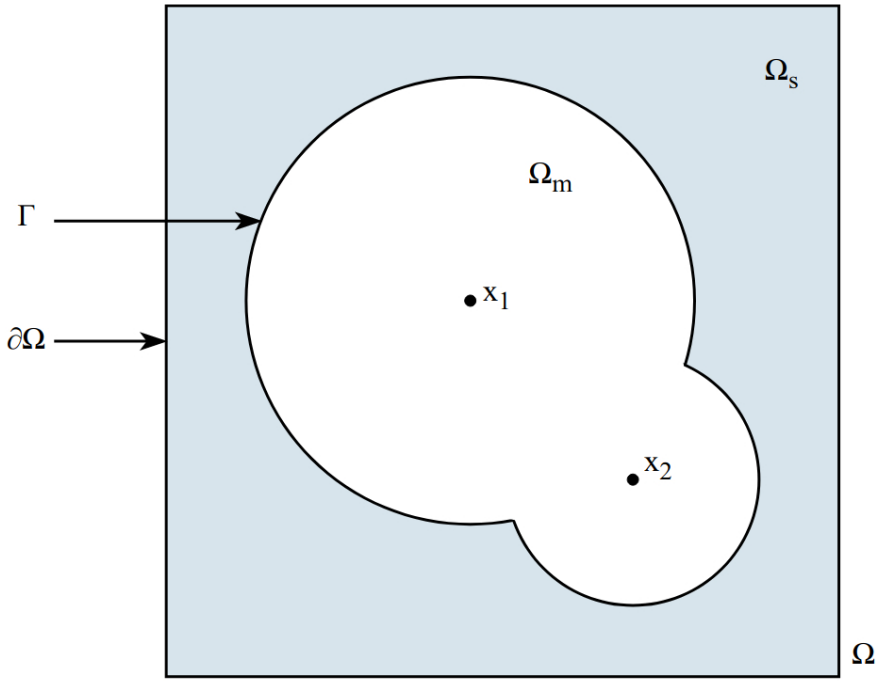
Density of solution ions:

$$\rho^s = -\varepsilon_0 \varepsilon_r(x) \, \kappa^2 \sinh\big(\, \varphi(x)\big)$$

The Poisson-Boltzmann Equation

Nondimensional form of the PBE:

$$\begin{cases} -\nabla \cdot (\varepsilon_0 \, \varepsilon_r(x) \, \nabla\varphi(x)\,) = \rho^s + \rho^f, & in\ \Omega \\ \varphi(x) = g(x), & on\ \partial\Omega \end{cases}$$



Relative dielectric permittivity:

$$\begin{cases} \varepsilon_r(x) = \varepsilon_s(x), & in\ \Omega_s \\ \varepsilon_r(x) = \varepsilon_m, & in\ \Omega_m \end{cases}$$

Fixed charge density:

$$\rho^f = \sum_i q_i \, \delta(x - x_i)$$

Density of solution ions:

$$\rho^s = -\varepsilon_0 \varepsilon_r(x) \, \kappa^2 \sinh\big(\, \varphi(x)\big)$$

Linearized PBE

$$\rho^s = -\varepsilon_0 \varepsilon_r(x) \, \kappa^2 \varphi(x)$$

Ingredients for the easy_pbe implementation

- **Finite Element discretization** on Oct-tree grids
- **Approximation of point sources by finite charge density**
- Geometric representation of the **molecular surface**
- **Grid partitioning for parallelization**

Treatment of Point Sources

Fixed charge density contains point sources:

$$\rho^f = \sum_i q_i \, \delta(x - x_i)$$

Singularities in the analytical solution of the PBE

Treatment of Point Sources

Fixed charge density contains point sources:

$$\rho^f = \sum_i q_i \, \delta(x - x_i)$$

Singularities in the analytical solution of the PBE

easy_pbe adopts **continuous first-order Lagrangian Finite Elements**

Treatment of Point Sources

Fixed charge density contains point sources:

$$\rho^f = \sum_i q_i \, \delta(x - x_i)$$

Singularities in the analytical solution of the PBE

easy_pbe adopts **continuous first-order Lagrangian Finite Elements**

Smoothed representation of the atomic charges

$$\rho_i(x) \coloneqq \sum_k N_k(x) q_{ik}\,, \qquad \int_\Omega \rho_i(x) d\omega = q_i$$

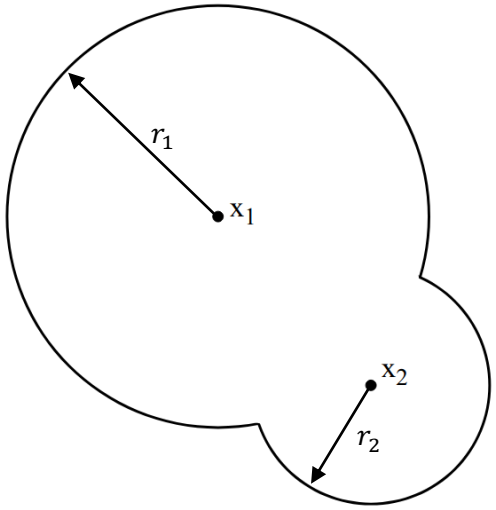
$$q_{ik} \coloneqq q_i \frac{N_k(x_i)}{\int_\Omega N_k(x) d\omega}$$

Minimizing the dependence on the adaptive grid

Geometric Representation of the Molecular Surface

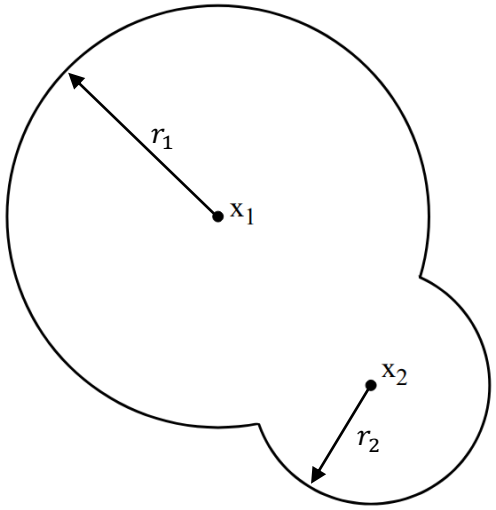
Van der Waals:

- Simple definition
- Not accurate
- Numerical problems



Geometric Representation of the Molecular Surface

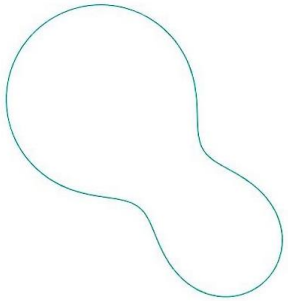
Van der Waals:
- Simple definition
- Not accurate
- Numerical problems



Level-set:
- *Blobby* surface

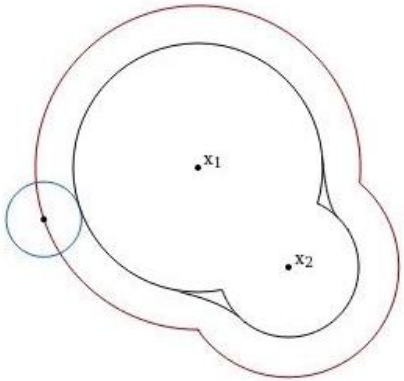
$$S(x) \coloneqq \{x \in \mathbb{R}^3 : G(x) = 1\}$$

$$G(x) = \sum_{i=1}^{n_a} e^{B\left(\frac{\|x - x_i\|^2}{r_i^2} - 1\right)}$$



Blobby surface with B = -2.5

Alpha-shapes:
- *Skin* surface
- *SES* surface



Oct-tree Grids

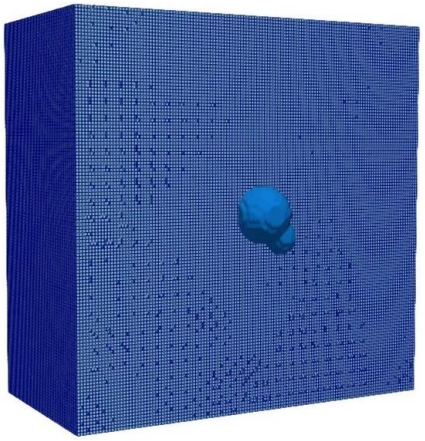
Smaller linear system dimension

Tensor product cartesian grids

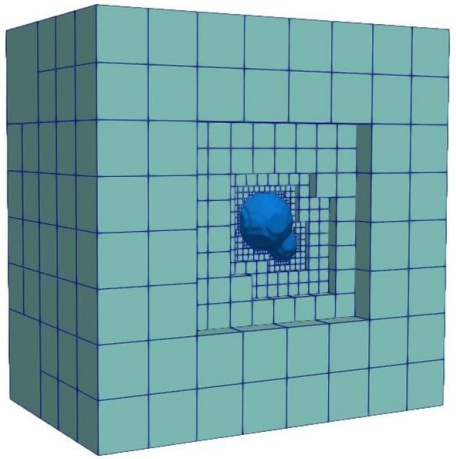
Oct-tree grids

Adaptive, simplicial, conforming meshes

Easier grid management

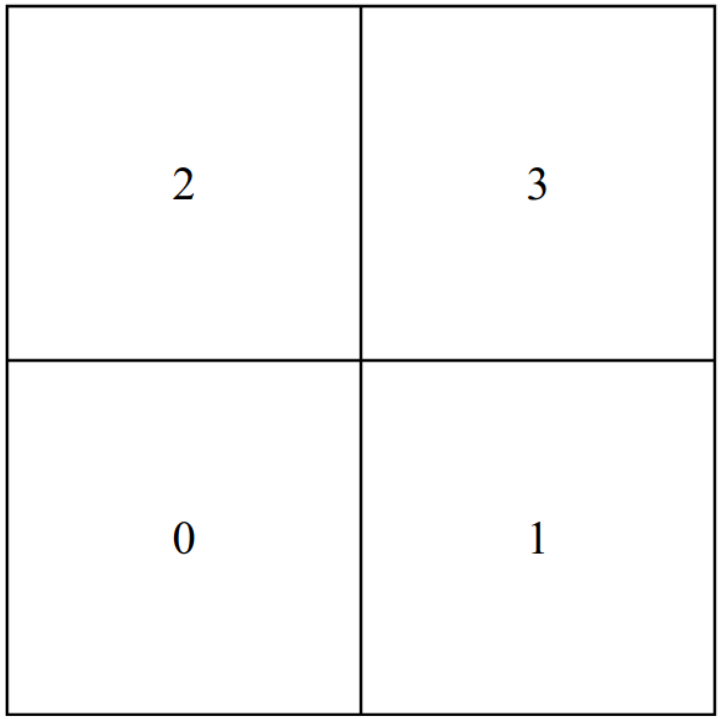


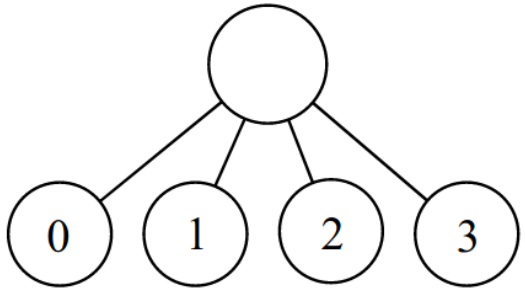
Uniformly refined mesh
Uniform level = 7
Number of cells = 2097152



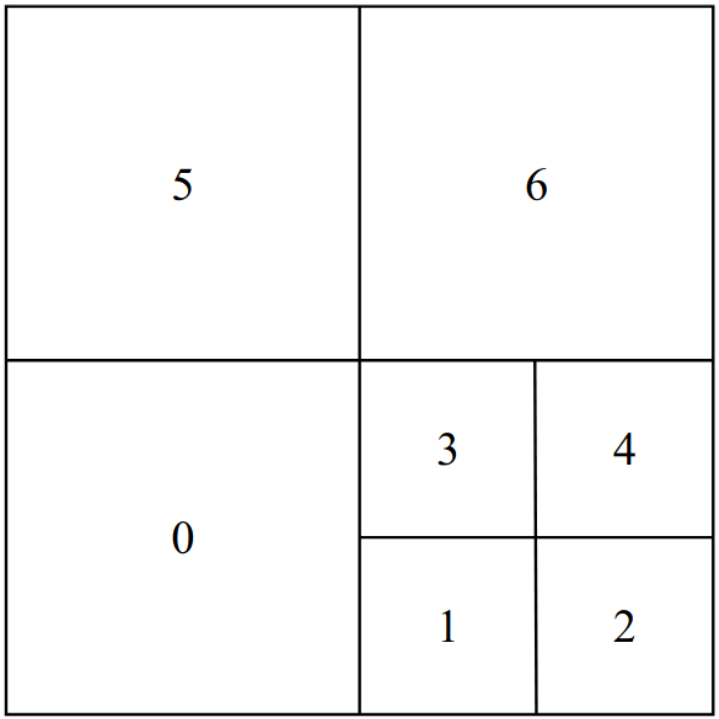
Oct-tree grid with hierarchical refinement
Min/max level = 2 - 7
Number of cells = 6392

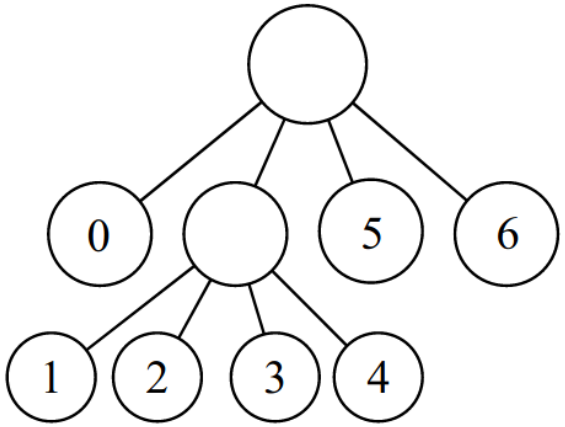
Oct-tree Grids



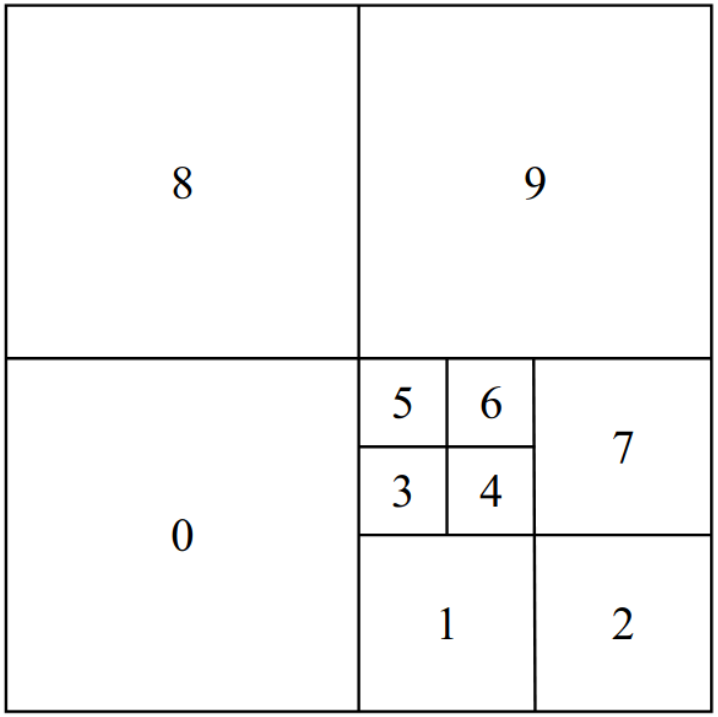


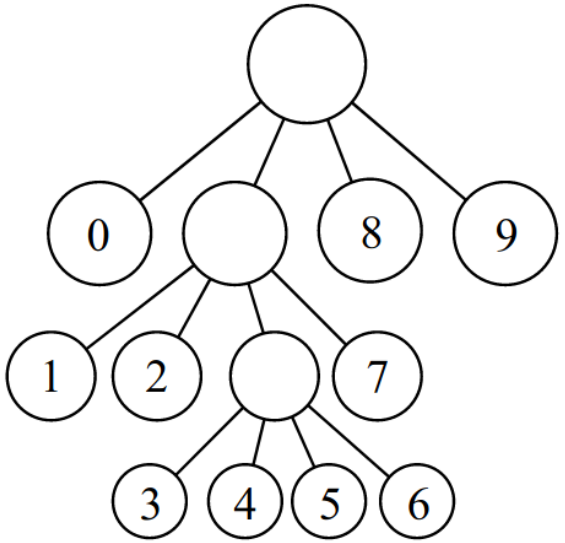
Oct-tree Grids



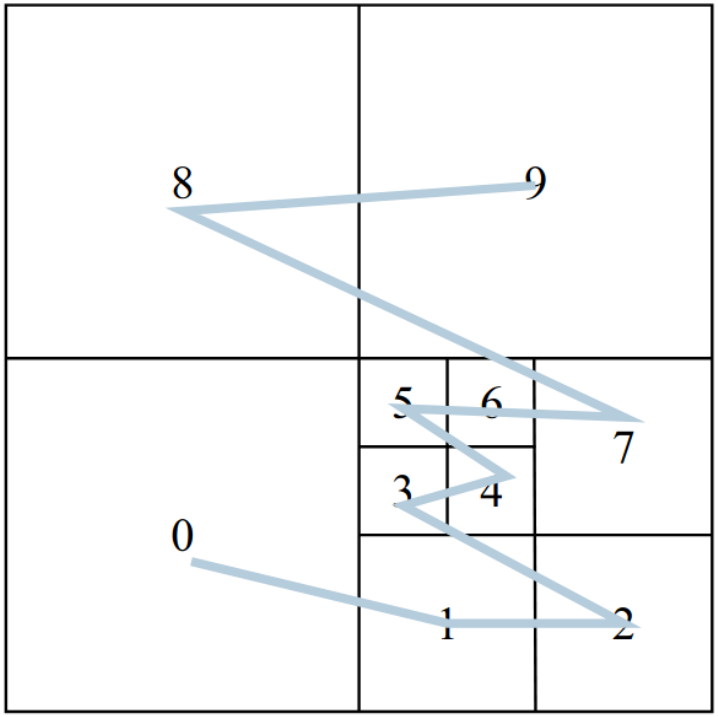


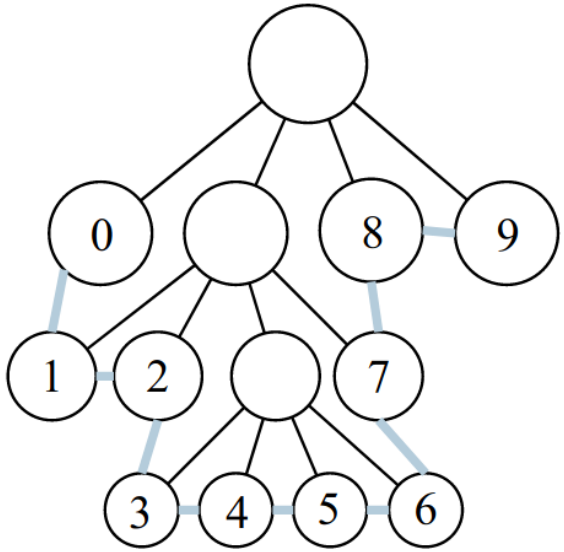
Oct-tree Grids





Oct-tree Grids

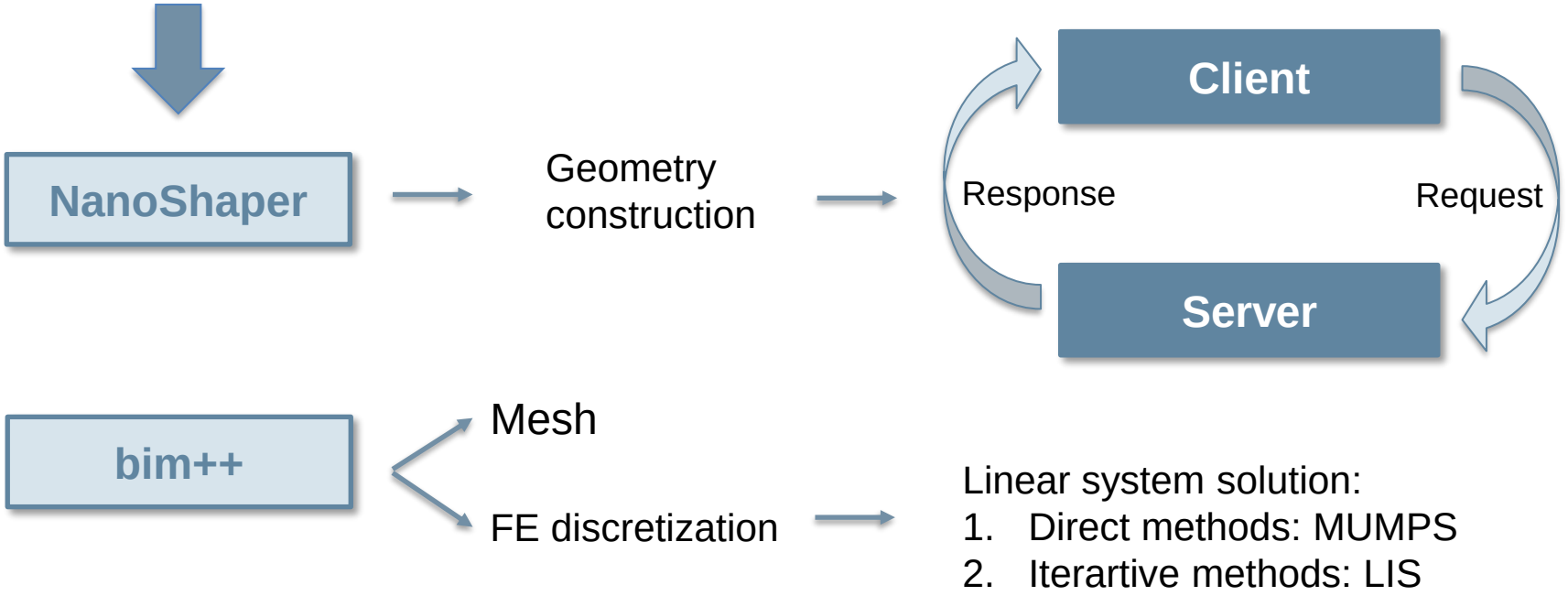




easy_pbe

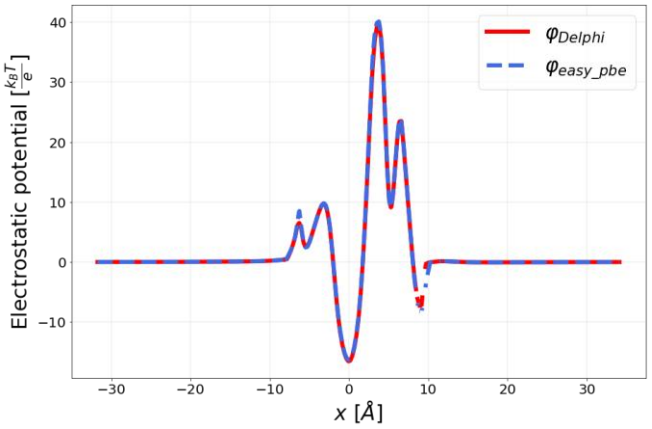
Development of a **c++ code:**

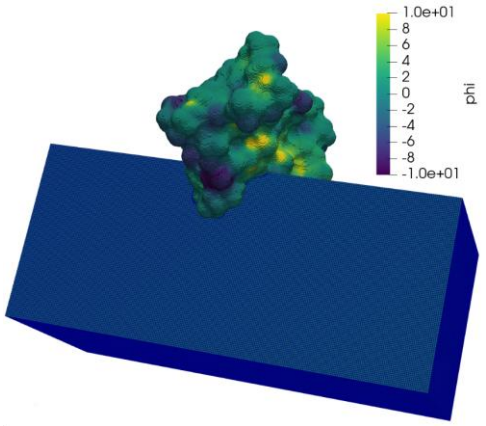
- **Scalable**
- **Efficient** in solving PBE
- Able to handle the **MS geometry**



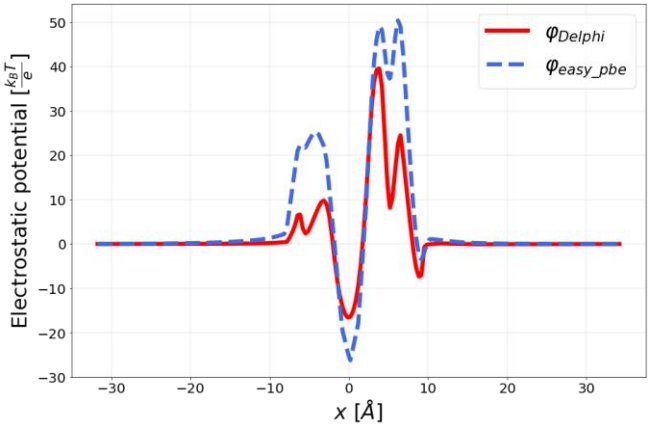
Numerical Results on Real Molecules
Validation tests: Crambin molecule

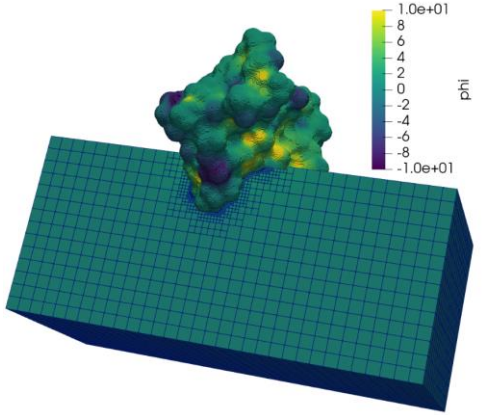
Refinement:
Minlevel = 8
Maxlevel = 9





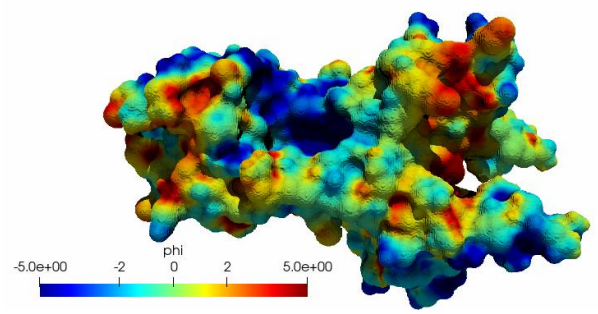
Refinement:
Minlevel = 5
Maxlevel = 9

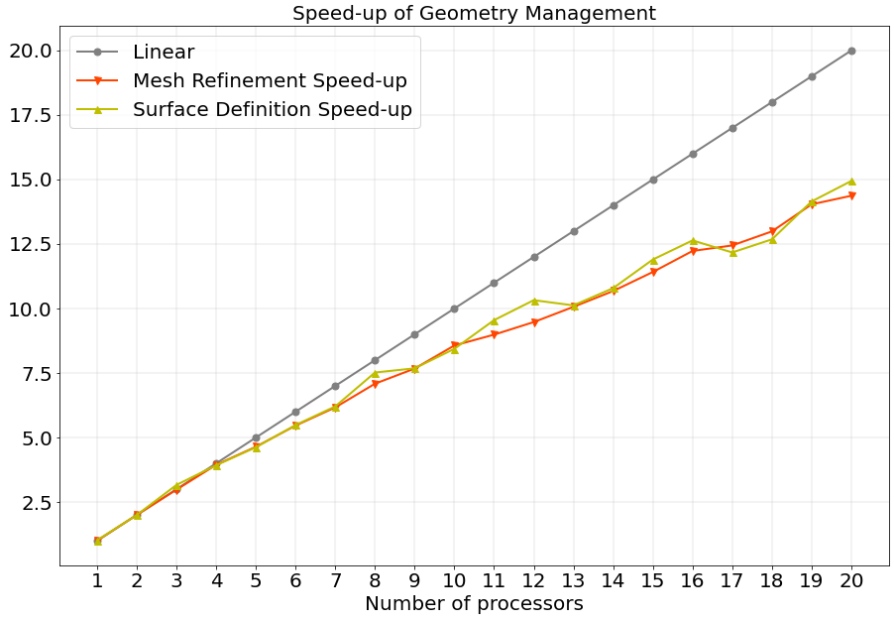


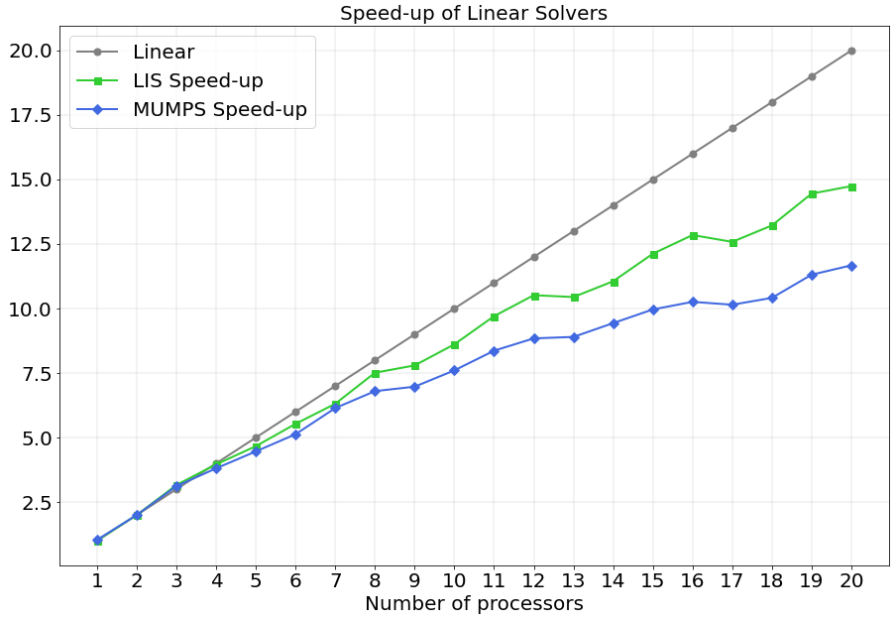


Numerical Results on Real Molecules
Parallel scaling test

Nerve growth factor protein, involved in many cell physiological processes.

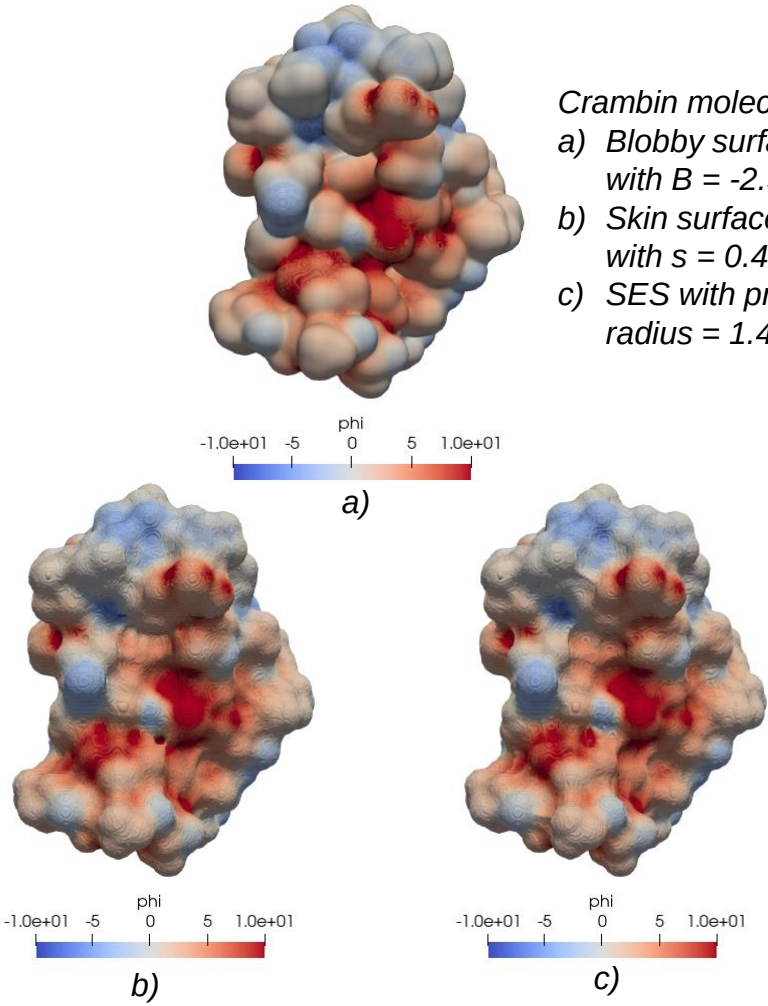






Conclusions

- **easy_pbe** is a c++ **scalable solver**
- **Oct-tree grids**
- **Different molecular surface definitions**
- **Validation** with test cases and real molecules



Crambin molecule:

a) Blobby surface with B = -2.5

b) Skin surface with s = 0.45

c) SES with probe radius = 1.4

THANK YOU FOR THE ATTENTION!

Bibliografy

1. https://github.com/carlodefalco/easy_pbe.git.
2. http://p4est.github.io.
3. http://optimad.github.io/bitpit/modules/index.html#PABLO.
4. https://www.ssisc.org/lis/lis-ug-en.pdf.
5. N. Baker, M. Holst, and F. Wang. Adaptive multilevel finite element solution of the poisson-boltzmann equation ii. refinement at solvent-accessible surfaces in biomolecular systems. J. Comput. Chem., 21:1343– 1352.
6. M. Born. Volumen und hydratationswärme der ionen. Z. Phys., 1(45), 1920.
7. A. H. Boschitsch and M. O. Fenley. A fast and robust poisson–boltzmann solver based on adaptive cartesian grids. Journal of Chemical Theory and Computation, 7(5):524–1540, 05 2011.
8. S. Decherchi, C. E. Catalano J. Colmenare and, M. Spagnuolo, E. Alexov, and W. Rocchia. Between algorithm and model: different molecular surface definitions for the poisson-boltzmann based electrostatic characterization of biomolecules in solution. Commun. Comput. Phys., 13:61–89, January 2013. 26.
9. M. Holst, N. Baker, and F. Wang. Adaptive multilevel finite element solution of the poisson–boltzmann equation i. algorithms and examples. J. Comput. Chem., 21:1319–1342.
10. B. Honig and A. Nicholls. Classical electrostatics in biology and chemistry. SCIENCE, 268, MAY 1995.
11. A. Kucherova, S. Strango, S. Sukenik, and M. Theillard. Computational modeling of protein conformational changes - application to the opening sars-cov-2 spike. Journal of Computational Physics, 444:110591, 2021.
12. A. Lee, W. Geng, and S. Zhao. Regularization methods for the poisson-boltzmann equation: Comparison and accuracy recovery. Journal of Computational Physics, 426:109958, 2021.
13. C. Li, Z. Jia, A. Chakravorty, S. Pahari, Y. Peng, S. Basu, M. Koirala, S.K. Panday, M. Petukh, L. Li, and E. Alexov. Delphi suite: New developments and review of functionalities. J Comput Chem., 40(28):2502– 2508, Oct 2019.
14. L. Li, C. Li, S. Sarkar, J. Zhang, S. Witham, Z. Zhang, L. Wang, N. Smith, M. Petukh, and E. Alexov. Delphi: a comprehensive suite for delphi software and associated resources. BMC Biophys, 4(1), May 2012.
15. J. A. Wilson, A. Bender, T. Kaya, and P. A. Clemons. Alpha shapes applied to molecular shape characterization exhibit novel properties compared to established shape descriptors. J. Chem. Inf. Model., 49(10):2231–2241, 2009.

Introduction

Understanding of physical/chemical properties of molecules in acqueous solutions

FOR

Describing the strength and the nature of the **ELECTROSTATIC INTERACTIONS**

1920

Max Born, Volumen und hydratationswärme der ionen [10].

1995

Barry Honig and Anthony Nicholls, Classical electrostatics in biology and chemistry [17].

2021

Application to the opening **Sars-Cov-2 spike protein** [19].

Treatment of Point Sources

Presence of Point Sources induces **SINGULARITIES**

Smoothed representation of the atomic charges

$$\rho_i(x) \coloneqq \sum_k N_k(x) q_{ik}\,, \qquad \int_\Omega \rho_i(x)\,d\omega = q_i$$

$$q_{ik} \coloneqq q_i \frac{N_k(x_i)}{\int_\Omega N_k(x)\,d\omega}$$

Regularized reformulation

$$\varphi(x) = \varphi_m(x) + \varphi_s(x)$$

Transfers the effect of the point charges to:

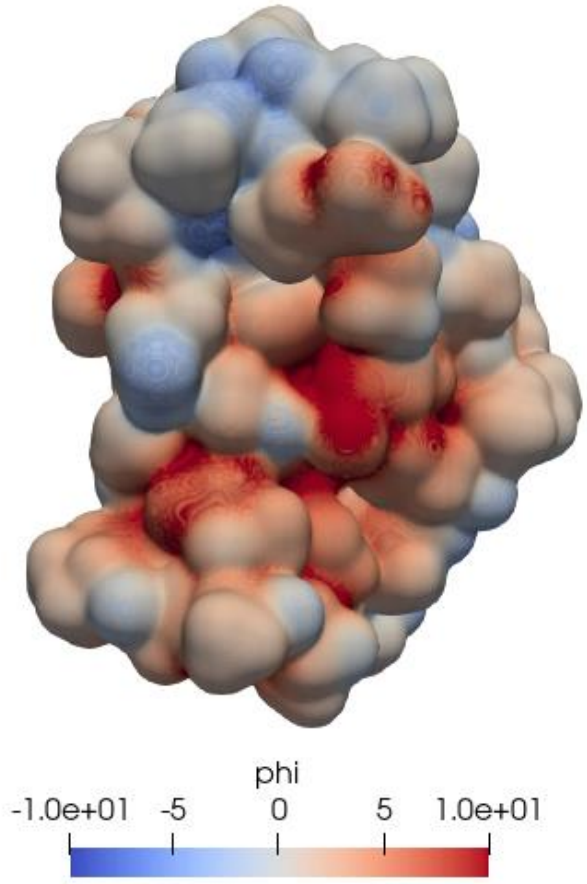
- The boundary conditions **on $\partial\Omega$**;
- The interface condition **on Γ**.

Geometric Representation of the Molecular Surface: *Blobby Surface*

$$S(x) \coloneqq \{x \in \mathbb{R}^3 : G(x) = 1\}$$

$$G(x) = \sum_{i=0}^{n_a} exp\left(B\left(\frac{\|x - x_i\|^2}{r_i^2} - 1\right)\right)$$

PROS	CONS
Simple and intuitive definition	Cannot be partitioned in analytical patches
Easy to implement	No easy projection methods
Smooth surface	Difficult B tuning



Crambin molecule, Blobby surface with B = -2.5.

Geometric Representation of the Molecular Surface: *Skin Surface*

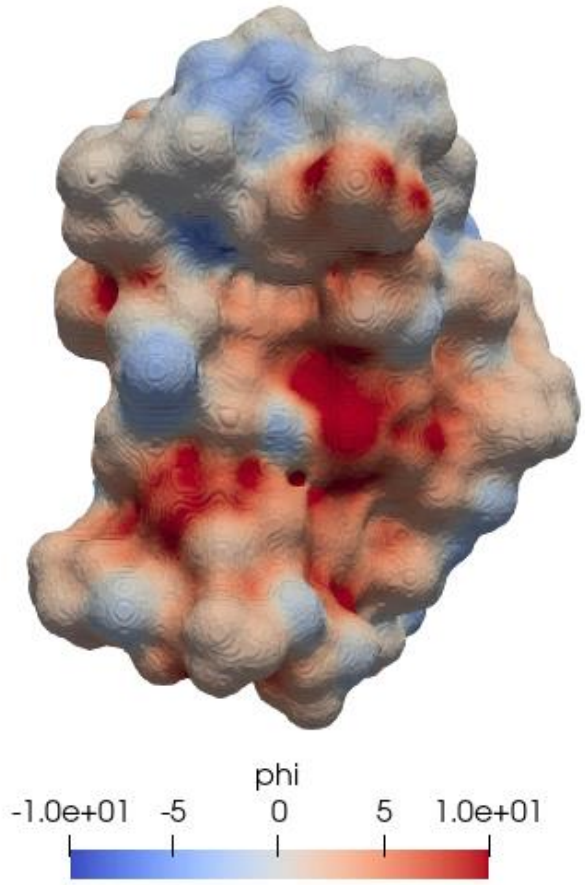
$$S = \{p_i = (x_i, w_i), x_i \in \mathbb{R}^3, i = 1, \ldots, n_a\}$$

$$w_i = \frac{r_i^2}{s}$$

Mixed cells:

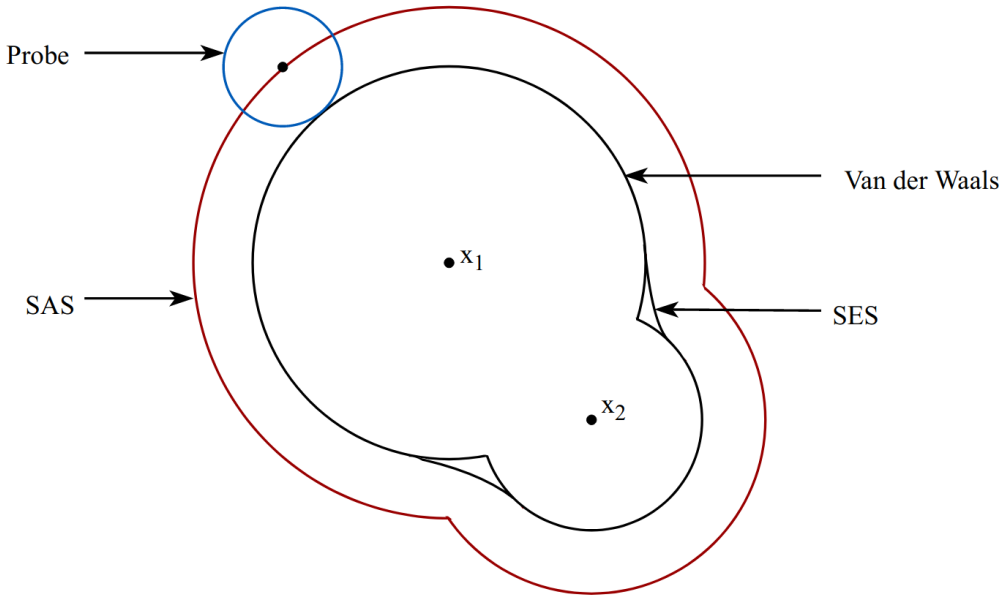
$$\mu_X^s = \{s \cdot v_X \oplus (1 - s) \cdot \delta_X\}$$

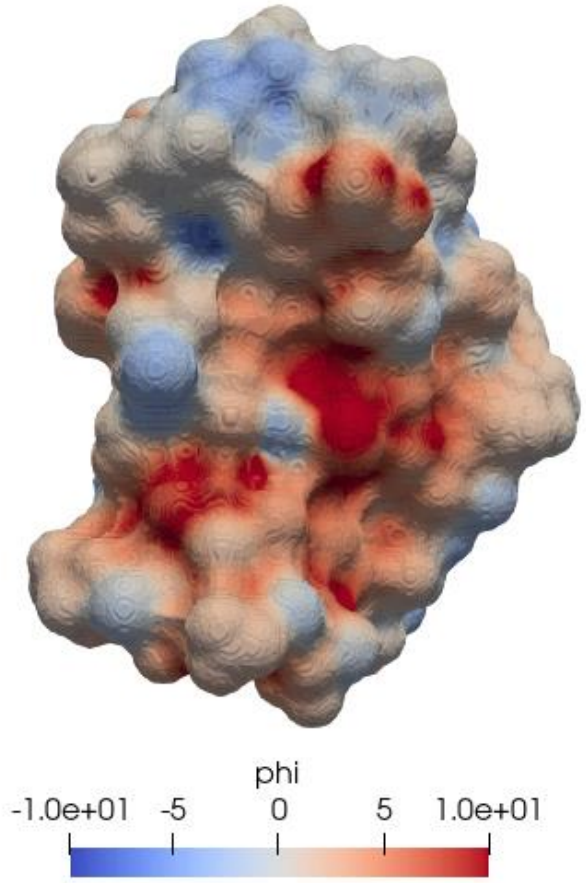
PROS	CONS
Use of fast algorithms, *e.g.* Voronoi diagram and Delaunay tetrahedrization	Higher computational costs *w.r.t.* SES
Limited pathological configurations	Unphysical high dielectric cavities



Crambin molecule, Skin surface with s = 0.45.

Geometric Representation of the Molecular Surface: *SES Surface*





Crambin molecule, SES surface with probe = 1.4.

easy_pbe Description and Implementation

Development of a **c++ code** using**:**

- **bim++** library
- **NanoShaper** library

Algorithm steps:

1. Data Input
2. Surface Creation
3. Mesh Creation and Refinement
4. Ray Caching
5. Matrix and Vectors Assembling
6. Linear System Solution

2. Surface Creation

```
ray_cache.init_analytical_surf (pb.atoms, pb.surf_type, pb.surf_param,
                                pb.stern_layer, pb.num_threads);
```

3. Mesh Creation and Refinement

Cell marked for **refinement** if

- it contains:
  - at least 1 node in Ω_m
  - at least 1 node in Ω_s
- it contains a point charge

Cell marked for **coarsening** if

- it doesn't contain:
  - the molecular surface Γ
  - a point charge

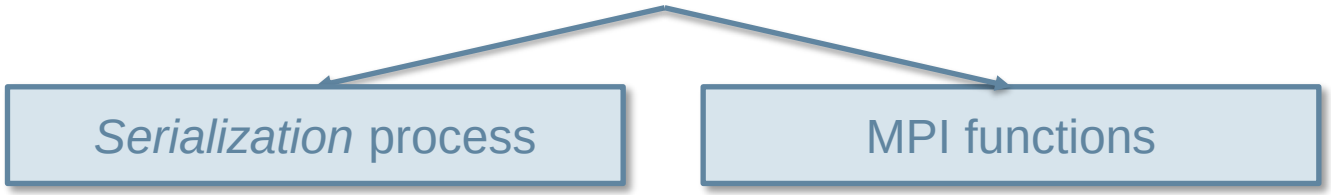
4. **Ray Caching**

NanoShaper: executed either in serial or using **multiple threads**.
easy_pbe: parallelization rely on **MPI multi-core**.

```
std::map<std::array<double, 2>, crossings_t, map_compare> rays;
```

ray_cache_t class functions involved in the **caching** operation:

- crossings_t & operator() (double x0, double x1)
- void fill_cache ()



5. Matrix and Vectors Assembling

Use of **bim++ functions**:

- bim3a_advection_diffusion
- bim3a_reaction
- bim3a_rhs

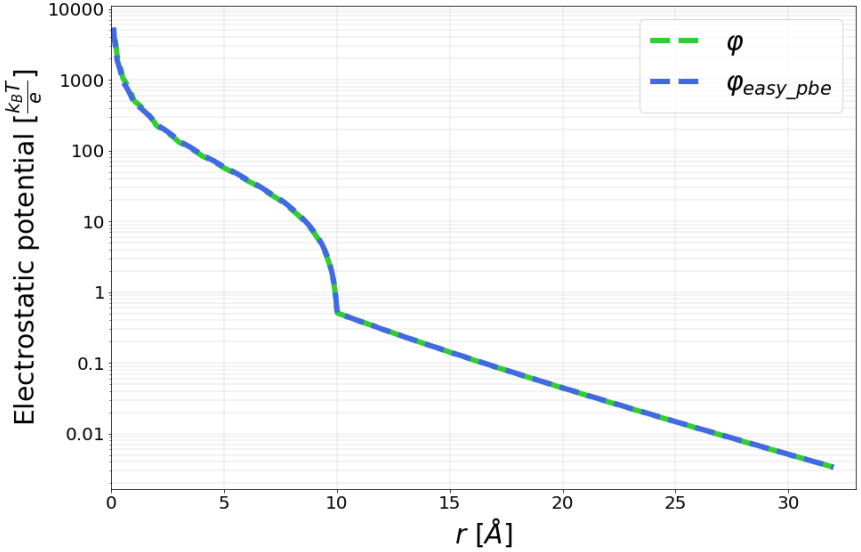
Distributed vectors and matrices

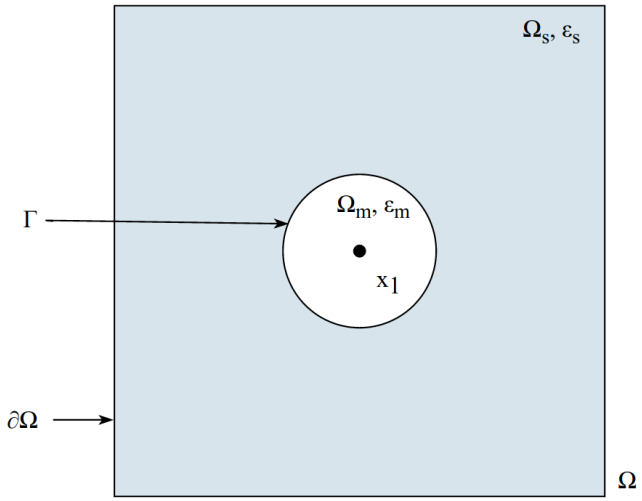
6. Linear System Solution

MUMPS

LIS

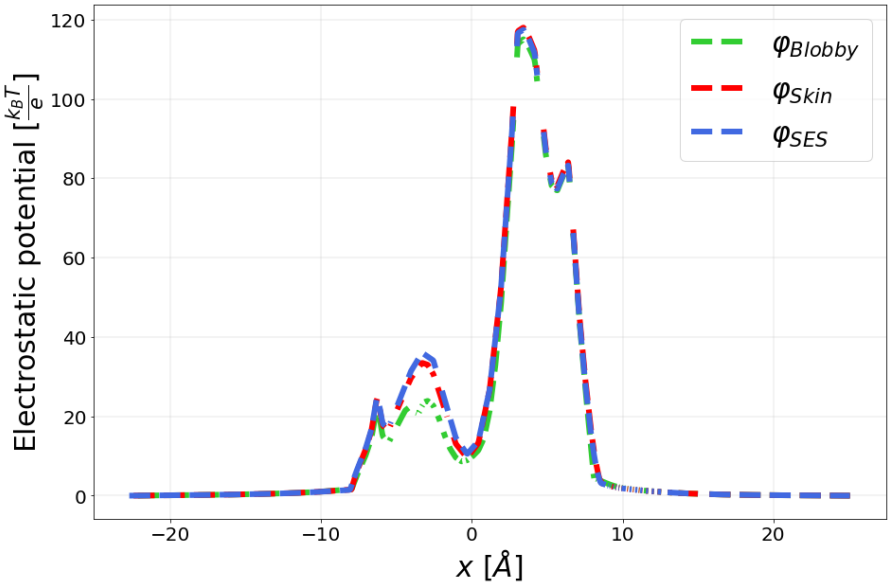
Validation Test Cases





Numerical Results on Real Molecules
Crambin test case

Surface type	Ref-Level	Times	Number of mesh cells
Blobby	2-5-9	P: 47.3448 s M: 64.404 s R: 154.745 s	1334572
Skin	2-5-9	P: 44.529 s M: 5.750 s R: 62.073 s	1218722
SES	2-5-9	P: 45.913 s M: 5.668 s R: 61.610 s	1206458



Bibliografy

[1] https://github.com/carlodefalco/easy_pbe.git.

[2] http://p4est.github.io.

[3] http://optimad.github.io/bitpit/modules/index.html#PABLO.

[4] https://www.ssisc.org/lis/lis-ug-en.pdf.

[5] N. Baker, M. Holst, and F. Wang. Adaptive multilevel finite element solution of the poisson-boltzmann equation ii. refinement at solvent-accessible surfaces in biomolecular systems. J. Comput. Chem., 21:1343– 1352.

[6] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, and P.E. Bourne. The protein data bank. Nucleic Acids Research, 28:235–242, 2000. http://www.rcsb.org/.

[7] J.F. Blinn. A generalization of algebraic surface drawing. ACM T. Graph., 1(3):235–256, 1982.

[8] A.M. Bonvin, J.A. Rullmann, R.M. Lamerichs, R. Boelens, and R. Kaptein. Direct noe refinement of crambin from 2d nmr data using a slow-cooling annealing protocol, 1993. PDB ID: 1CCM, PDB doi: http://doi.org/10.2210/pdb1CCM/pdb.

[9] A.M. Bonvin, J.A. Rullmann, R.M. Lamerichs, R. Boelens, and R. Kaptein. Ensemble iterative relaxation matrix approach: a new nmr refinement protocol applied to the solution structure of crambin. Proteins, 15:385–400, 1993. doi: http://dx.doi.org/10.1002/prot.340150406.

[10] M. Born. Volumen und hydratationswärme der ionen. Z. Phys., 1(45), 1920.

[11] A. H. Boschitsch and M. O. Fenley. A fast and robust poisson–boltzmann solver based on adaptive cartesian grids. Journal of Chemical Theory and Computation, 7(5):524–1540, 05 2011.

[12] S. Decherchi, José C., C. E. Catalano, M. Spagnuolo, E. Alexov, and W. Rocchia. Between algorithm and model: different molecular surface definitions for the poisson-boltzmann based electrostatic characterization of biomolecules in solution. Commun Comput Phys., 13:61–89, January 2013.

[13] S. Decherchi, C. E. Catalano J. Colmenare and, M. Spagnuolo, E. Alexov, and W. Rocchia. Between algorithm and model: different molecular surface definitions for the poisson-boltzmann based electrostatic characterization of biomolecules in solution. Commun. Comput. Phys., 13:61–89, January 2013. 26

[14] S. Decherchi and W. Rocchia. A general and robust ray-casting-based algorithm for triangulating surfaces at the nanoscale. PLoS One, 8(4):e59744, 2013. http://dx.doi.org/10.1371/journal.pone.0059744.

[15] H. Edelsbrunner. Deformable smooth surface design. Disrete Comput. Geom., 21:87–115, 1999.

[16] E. Jurrus et al. Improvements to the APBS biomolecular solvation software suite. Protein Science, 27:112– 128, 2018.

[17] M. Holst, N. Baker, and F. Wang. Adaptive multilevel finite element solution of the poisson–boltzmann equation i. algorithms and examples. J. Comput. Chem., 21:1319–1342.

[18] B. Honig and A. Nicholls. Classical electrostatics in biology and chemistry. SCIENCE, 268, MAY 1995.

[19] M. Kosloff, A. Travis, and D. Bosch et al. Integrating energy calculations with functional assays to decipher the specificity of g protein–rgs protein interactions. Nat Struct Mol Biol, 18:846—-853, 2011.

[20] A. Kucherova, S. Strango, S. Sukenik, and M. Theillard. Computational modeling of protein conformational changes - application to the opening sars-cov-2 spike. Journal of Computational Physics, 444:110591, 2021.

Bibliografy

[21] A. Lee, W. Geng, and S. Zhao. Regularization methods for the poisson-boltzmann equation: Comparison and accuracy recovery. Journal of Computational Physics, 426:109958, 2021.

[22] B. Lee and F. M. Richards. The interpretation of protein structures: Estimation of static accessibility. J. Mol. Biol., 55:379–400, 1971.

[23] C. Li, Z. Jia, A. Chakravorty, S. Pahari, Y. Peng, S. Basu, M. Koirala, S.K. Panday, M. Petukh, L. Li, and E. Alexov. Delphi suite: New developments and review of functionalities. J Comput Chem., 40(28):2502– 2508, Oct 2019.

[24] L. Li, C. Li, S. Sarkar, J. Zhang, S. Witham, Z. Zhang, L. Wang, N. Smith, M. Petukh, and E. Alexov. Delphi: a comprehensive suite for delphi software and associated resources. BMC Biophys, 4(1), May 2012.

[25] M. Mirzadeh, M. Theillard, and F. Gibou. A second-order discretization of the nonlinear poisson–boltzmann equation over irregular geometries using non-graded adaptive cartesian grids. Journal of Computational Physics, 230(5):2125–2140, 2011.

[26] A. Nicholls and B. Honig. A rapid finite difference algorithm, utilizing successive over-relaxation to solve the poisson-boltzmann equation. J. Comput. Chem., 12:435–445., 1991.

[27] W. Rocchia. Poisson-boltzmann equation boundary conditions for biological applications. Mathematical and Computer Modelling, 41(10):1109–1118, 2005. Modelling Complex Systems in Molecular Biology and Tumor Dynamics and Control.

[28] R. Rohs, S. M. West, A. Sosinsky, P. Liu, R. S. Mann, and B. Honig. The role of dna shape in protein – dna recognition. nature, 461:1248–1253, 2009.

[29] F. B. Sheinerman, R. Norel, and B. Honig. Electrostatic aspects of protein–protein interactions. Current Opinion in Structural Biology, 10(2):153–159, 2000. [30] H.L. Sun and T. Jiang. Crystal structure of nerve growth factor in complex with lysophosphatidylinositol, 2015. PDB ID: 4XPJ, PDB doi: http://doi.org/10.2210/pdb4XPJ/pdb.

[31] H.L. Sun and T. Jiang. The structure of nerve growth factor in complex with lysophosphatidylinositol. Acta Crystallogr F Struct Biol Commun, 71:906–912, 2015. doi: http://dx.doi.org/10.1107/ S2053230X15008870.

[32] E. Alexov W. Rocchia and B. Honig. Extending the applicability of the nonlinear poisson-boltzmann equation: Multiple dielectric constants and multivalent ions. J. Phys. Chem. B, 105(28):6507–6514.

[33] A.C. Walls, Y.J. Park, M.A. Tortorici, A. Wall, Seattle Structural Genomics Center for Infectious Disease (SSGCID), A.T. McGuire, and D. Veesler. Sars-cov-2 spike ectodomain structure (open state), 2020. PDB ID: 6VYB, PDB doi: http://doi.org/10.2210/pd6VYB/pdb.

[34] A.C. Walls, Y.J. Park, M.A. Tortorici, A. Wall, A.T. McGuire, and D. Veesler. The structure of nerve growth factor in complex with lysophosphatidylinositol. Cell, 181:281, 2020. doi: http://dx.doi.org/ 10.1016/j.cell.2020.02.058. 27

[35] J. A. Wilson, A. Bender, T. Kaya, and P. A. Clemons. Alpha shapes applied to molecular shape characterization exhibit novel properties compared to established shape descriptors. J. Chem. Inf. Model., 49(10):2231–2241, 2009.

[36] Y. Zhang, G. Xu, and C. Bajaj. Quality meshing of implicit solvation models of biomolecular structures. Comput. Aided Geom. Des., 23:510–530, 2006.

[37] T. Zhou, T. Tsybovsky, and P.D. Kwong et al. Antibodies with potent and broad neutralizing activity against antigenically diverse and highly transmissible sars-cov-2 variants. Biorxiv, 2021. doi: http: //dx.doi.org/10.1101/2021.02.25.432969.

[38] T. Zhou, T. Tsybovsky, and P.D. Kwong. Cryo-em structure of the spike of sars-cov-2 omicron variant of concern, 2021. PDB ID: 7TB4, PDB doi: http://doi.org/10.2210/pdb7TB4/pdb.