Stochastic Methods in Electrostatics: Applications to Biological and Physical Science

Michael Mascagni ^a

Department of Computer Science and School of Computational Science Florida State University, Tallahassee, FL 32306 USA E-mail: mascagni@fsu.edu URL: http://www.cs.fsu.edu/~mascagni

Research supported by ARO, DoD, DOE, NATO, and NSF

^aWith help from Drs. James Given, Chi-Ok Hwang, and Nikolai Simonov

Outline of the Talk

First-Passage Algorithms

- Walk on Spheres (WOS)
- Greens Function First Passage (GFFP)
- Simulation-Tabulation (S-T)
- Walk on Subdomains (biochemistry)
- Walk on the Boundary

Applications

- Materials Science
- Biochemistry

Last-Passage Algorithms

Conclusions and Future Work

Stochastic Methods for Partial Differential Equations (PDEs)

Examples for Solving Elliptic PDEs (Path Integrals)

- Exterior Laplace problems and electrostatics
- Electrical capacitance
- Charge density

Advantages of Stochastic Algorithms (Curse of Dimensionality)

- Can avoid complex discrete objects
- Can deal with complicated geometries/interfaces
- Can often cope with singular solutions

Brownian Motion and the Diffusion/Laplace Equations

Cauchy problem for the diffusion equation:

$$u_t = \frac{1}{2}\Delta u \tag{1}$$

$$u(x,0) = f(x) \tag{2}$$

in 1-D:

$$u(x,t) = \int_{-\infty}^{\infty} \omega(x-y,t) f(y) dy$$
(3)

where

$$\omega(x - y, t) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x - y)^2}{2t}}$$
(4)

Brownian Motion and the Diffusion/Laplace Equations

$$u(x,t) = \mathbb{E}_x[f(X^x(t))] \tag{5}$$

• $X^{x}(t)$: a Brownian motion which has $\omega(x - y, t)$ as the transition probability of going from x to y in time t

• $\mathbb{E}_{x}[.]$: an expectation *w.r.t.* Brownian motion

$$\mathbb{E}_x[f(X^x(t))] = \int_{-\infty}^{\infty} \omega(x - y, t) f(y) dy$$
(6)

The First Passage (FP) Probability is the Green's Function

A related elliptic boundary value problem (Dirichlet problem):

$$\Delta u(\mathbf{x}) = 0, \quad \mathbf{x} \in \Omega$$
$$u(\mathbf{x}) = f(\mathbf{x}), \quad \mathbf{x} \in \partial \Omega$$
(7)

- Distribution of z is uniform on the sphere
- Mean of the values of u(z) over the sphere is u(x)
- u(x) has mean-value property and harmonic
- Also, u(x) satisfies the boundary condition

$$u(\mathbf{x}) = \mathbb{E}_x[f(X^{\mathbf{x}}(t_{\partial\Omega}))]$$

Prof. Michael Mascagni: Stochastic Electrostatics

(8)



The First Passage (FP) Probability is the Green's Function (Cont.)

Reinterpreting as an average of the boundary values

$$u(\mathbf{x}) = \int_{\partial\Omega} p(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) dy$$
(9)

Another representation in terms of an integral over the boundary

$$u(\mathbf{x}) = \int_{\partial\Omega} \frac{\partial g(\mathbf{x}, \mathbf{y})}{\partial \mathbf{n}} f(\mathbf{y}) dy$$
(10)

 $g(\mathbf{x}, \mathbf{y})$ – Green's function of the Dirichlet problem in Ω

$$\implies p(\mathbf{x}, \mathbf{y}) = \frac{\partial g(\mathbf{x}, \mathbf{y})}{\partial \mathbf{n}}$$

(11)

'Walk on Spheres' (WOS) and Green's Function First Passage (GFFP) Algorithms

• Green's function is known

 \implies direct simulation of exit points and computation of the solution through averaging boundary values

• Green's function is unknown

 \implies simulation of exit points from standard subdomains of Ω ,

e.g. spheres

 \implies Markov chain of 'Walk on Spheres' (or GFFP algorithm)

 $\{\mathbf{x_0} = \mathbf{x}, \mathbf{x_1}, \ldots\}$

 $\mathbf{x_i} \to \partial \Omega$ and hits ε -shell is $N = O(\ln(\varepsilon))$ steps

 $\mathbf{x}_{\mathbf{N}}$ simulates exit point from Ω with $O(\varepsilon)$ accuracy



WOS:



Timing of the 'Walk on Spheres' Algorithm





Prof. Michael Mascagni: Stochastic Electrostatics

Slide 11 of 56



The Simulation-Tabulation (S-T) Method for Generalization

• Green's function for the non-intersected surface of a sphere located on the surface of a reflecting sphere





Another S-T Application: Mean Trapping Rate

In a domain of nonoverlapping spherical traps :



Biological Electrostatics: Motivation

Electrostatics are Extremely Important in Quantitative Biochemistry:

- Ligand binding
- Protein-protein interactions
- Protein-nucleic acid interactions
- Prediction from primary structure information
- In Vivo Electrostatics Must Include the Solvent
 - Explicit solvent model: individual water/ions computed, often with Molecular Dynamics
 - Implicit solvent models: continuum model of water and dissolved ions used, Poisson-Boltzmann equation

Molecular Electrostatics Problem Implicit solvent model

• Poisson equation for the electrostatic potential, ϕ :

$$-\nabla \epsilon(x) \nabla \phi(x) = 4\pi \rho(x) , \ x \in \mathbb{R}^3$$

dielectric permittivity, $\epsilon,$ and charge density, $\rho,$ are position-dependent

- molecule Ω a compact cavity in \mathbb{R}^3 with low $\epsilon = \epsilon_i$
- surrounded by solvent with larger $\epsilon = \epsilon_e$
- point charges, q_m , at x_m inside molecule
- Boltzmann distribution of mobile ions in solvent

Molecular Electrostatics Problem (Cont.)

Explicit geometric models for solute molecule

- van der Waals surface: union of intersecting spheres (atoms): $\Omega = \bigcup_{m=1}^{M} B(x^{(m)}, r^{(m)})$ point charges – at their centers, $x_m = x^{(m)}$
- contact and reentrant surface, Γ : $\partial\Omega$ smoothed by the probe molecule of the solute rolling on it
- ion-accessible surface $\partial \Omega'$:

 $\Omega' = \bigcup_{m=1}^{M} B(x^{(m)}, r^{(m)} + r_{ion});$ ion-exclusion layer between $\partial \Omega'$ and Γ

Molecular Electrostatics Problem (Cont.)

Mathematical model (one-surface geometry)

• Poisson equation for the electrostatic potential, ϕ , inside a molecule

$$-\epsilon_i \Delta \phi_i(x) = \sum_{m=1}^M 4\pi q_m \delta(x - x_m) , \ x \in \Omega$$

• linearized Poisson-Boltzmann equation outside, $x \in \mathbb{R}^3 \setminus \overline{\Omega}$:

$$\Delta \phi_e(x) - \kappa^2 \phi_e(x) = 0 ,$$

• Continuity condition on the boundary

$$\phi_i = \phi_e \ , \ \epsilon_i \frac{\partial \phi_i}{\partial n(y)} = \epsilon_e \frac{\partial \phi_e}{\partial n(y)} \ , \ y \in \Gamma \equiv \partial \Omega$$

Electrostatic Potential, Field and Energy (Linear Problem)

• Point values of the potential: $\phi(x) = \phi^{(0)}(x) + g(x)$ Here, singular part of ϕ :

$$g(x) = \sum_{m=1}^{M} \frac{q_m}{\epsilon_i} \frac{1}{|x - x_m|}$$

• Free electrostatic energy of a molecule = linear combination of point values of the regular part of the electrostatic potential $\phi^{(0)}$:

$$E = \frac{1}{2} \sum_{m=1}^{M} \phi^{(0)}(x_m) q_m ,$$

• Point values of the electrostatic field: $\nabla \phi(x)$

Monte Carlo Estimates for Point Potential Values

Two different approaches to constructing Monte Carlo algorithms

- **1.** Probabilistic representation for the solution
- **2.** Classical potential theory

First approach

Laplace equation for the regular part of the potential inside Ω

$$\Delta \phi^{(0)} = 0$$

Probabilistic representation

$$\phi^{(0)}(x) = \mathbb{E}_x[\phi^{(0)}(x^*)]$$

 x^* – exit point from Ω of Brownian motion starting at x

'Walk on Spheres' Algorithm

x – center of a sphere \Rightarrow exit points are distributed isotropically. Ball B(x, R) lies entirely in Ω . Strong Markov property of Brownian motion \Rightarrow probabilistic representation holds valid for exit points.

Hence follows 'random walk on spheres' algorithm for general domains with regular boundary:

$$x^{k} = x^{k-1} + d(x^{k-1}) \times \omega^{k}$$
, $k = 1, 2, \dots$

Here

 $d(x^{k-1})$ – distance from x^{k-1} to the boundary $\{\omega^k\}$ – sequence of independent unit isotropic vectors x^k is exit point from the ball, $B(x^{k-1}, d(x^{k-1}))$, for Brownian motion starting at x^{k-1}

Green's Function First Passage Simulation

Other domains with known Green's function $(G) \iff$ one-step simulation of exit points distributed on the boundary in accordance with $\partial G/\partial n$

For general domains:

Efficient way to simulate x^* – combination of 'walk in subdomains' approach and 'walk on spheres' algorithm

The whole domain, Ω , is represented as a union of intersecting subdomains:

$$\Omega = \bigcup_{m=1}^{M} \Omega_m$$

Simulate exit point separately in every Ω_m

Green's Function First Passage Simulation (cont.)

 $x^0 = x, x^1, \dots, x^N$ – Markov chain, every x^{i+1} is exit point from the corresponding subdomain for Brownian motion starting at x^i

For spherical subdomains, $B(x_m^i, R_m^i)$, exit points are distributed in accordance with the Poisson's kernel

$$\frac{|x^i - x^i_m|}{4\pi R^i_m} \frac{|x^i - x^i_m|^2 - R^i_m}{|x^i - y|^3}$$

 $x^* = x^N$ is exit point of Brownian motion from Ω Schwartz lemma \Rightarrow Markov chain $\{x^i\}$ converges to x^* geometrically

'Walk on Spheres' and 'Walk in Subdomains' Algorithms



Figure 1: Walk in subdomains example.

Monte Carlo Estimate for Point Potential Value

On every step

$$\phi^{(0)}(x^i) = \mathbb{E}[\phi^{(0)}(x^{i+1})|x^i]$$

Hence

$$\phi^{(0)}(x) = \mathbb{E}\phi^{(0)}(x^*) \equiv \mathbb{E}[\phi(x^*) - g(x^*)]$$

Values of the electrostatic potential on the boundary, $\phi(x^*)$, are not known. We can use their Monte Carlo estimates instead

• First approach

Discretization and randomization of the boundary condition $(y \in \Gamma, n = n(y) - \text{normal vector});$

$$\phi(y) = p_i \phi(y - hn) + p_e \phi(y + hn) + O(h^2)$$

$$\phi(x_1^*) = \mathbb{E}(\phi(x_2^0 | x_1^*) + O(h^2))$$

 $\begin{aligned} x_2^0 &= x_1^* - hn \text{ with probability } p_i \text{ (reenter molecule)} \\ x_2^0 &= x_1^* + hn \text{ with probability } p_e = 1 - p_i \text{ (exit to solvent)} \\ p_i &= \frac{\epsilon_i}{\epsilon_i + \epsilon_e} \end{aligned}$

• Second approach

Exact treatment of boundary conditions (mean-value theorem for boundary point, y, in the ball B(y, a) with surface S(y, a)):

$$\phi(y) = \frac{\epsilon_e}{\epsilon_e + \epsilon_i} \int_{S_e(y,a)} \frac{1}{2\pi a^2} \frac{\kappa a}{\sinh(\kappa a)} \phi_e + \frac{\epsilon_i}{\epsilon_e + \epsilon_i} \int_{S_i(y,a)} \frac{1}{2\pi a^2} \frac{\kappa a}{\sinh(\kappa a)} \phi_i$$
(12)
$$- \frac{(\epsilon_e - \epsilon_i)}{\epsilon_e + \epsilon_i} \int_{\Gamma \bigcap B(y,a) \setminus \{y\}} \frac{\cos \varphi_{yx}}{2\pi |y - x|^2} Q_{\kappa,a} \phi + \frac{\epsilon_i}{\epsilon_e + \epsilon_i} \int_{B_i(y,a)} [-2\kappa^2 \Phi_\kappa] \phi_i$$

Here

$$\varphi_{yx} - \text{angle between the normal } n(y) \text{ and } y - x$$

$$\Phi_{\kappa}(x-y) = -\frac{1}{4\pi} \frac{\sinh(\kappa(a-|x-y|))}{|x-y|\sinh(\kappa a)}$$

$$- \text{ Green's function for the Poisson-Boltzmann equation in } B(y,a)$$

$$Q_{\kappa,a}(|x-y|) = \frac{\sinh(\kappa(a-|x-y|)) + \kappa|x-y|\cosh(\kappa(a-|x-y|))}{\sinh(\kappa a)}$$

Next

Randomization of approximation to (12), $y = x_1^*$, $x = x_2^0$:

 $\phi(y) = \mathbb{E}\phi(x) + O(a/2R)^3$

Here

• with probability p_e exit to solvent:

x is chosen isotropically on the surface of auxiliary sphere, $S_+(y,a)$, that lies above tangent plane; random walk survives with probability $\frac{\kappa a}{\sinh(\kappa a)}$

• with probability p_i

x is chosen isotropically in the solid angle below tangent plane; with probability $-2\kappa^2 \Phi_{\kappa}$ it is sampled in $B_i(y, a)$ (reenter molecule);

with the complementary probability x is sampled on the surface of auxiliary sphere, $S_{-}(y, a)$, that lies below tangent plane; xreenters molecule with conditional probability 1 - a/2R and xexits to solvent with conditional probability a/2R

Higher order of approximation!

Monte Carlo Algorithm (cont.)

• x_2^0 inside

Return to the boundary at x_2^* , the exit point of Brownian motion (Markov chain) starting at x_2^0 , set

$$\phi(x_2^0) = \mathbb{E}(\phi(x_2^*) - g(x_2^*) + g(x_2^0) | x_2^0)$$
(13)

Repeat the randomized treatment of the boundary condition at the point x_2^*

Monte Carlo Algorithm (cont.)

• x_2^0 outside

'Walk on spheres' algorithm $x_2^{i+1} = x_2^i + \omega \times d_i, \qquad d_i = \text{distance from } x_2^i \text{ to } \partial\Omega$ Terminates with probability $1 - \frac{\kappa d_i}{\sinh(\kappa d_i)}$ on every step, or when $d_{N_2} < \varepsilon$. x_2^* – the nearest to $x_2^{N_2}$ on the boundary $\phi(x_2^0) = \mathbb{E}(\phi(x_2^*)|x_2^0) + O(\varepsilon)$ (14)

Repeat the randomized treatment of the boundary condition at the point
$$x_2^*$$

Molecular Electrostatics Problem (cont.)

In the exterior probability of terminating Markov chain depends linearly on the initial distance to the boundary, $d_0 \Rightarrow$ Mean number of returns to the boundary is $O(d_0)^{-1}$

- Finite-difference approximation of boundary conditions, ε = h² Mean number of steps in the algorithm is O(h⁻¹ log(h) f(κ)), f is a decreasing function (f(κ) = O(log(κ)) for small κ). Estimates for point values of the potential and free energy are O(h)-biased
- New treatment of boundary conditions provides O(ā)²-biased and more efficient Monte Carlo algorithm. Mean number of steps is O((ā)⁻¹ log(ā) f(κ)), ā = a/2R.

The same simulations give point values of the gradient and free electrostatic energy

Molecular Electrostatics Problem (cont.) Mathematical model (with ion-exclusion layer)

- Poisson equation inside a molecule
- Laplace equation in the ion-exclusion layer:

$$-\Delta\phi_{lay}(x) = 0$$

- linearized Poisson-Boltzmann equation outside, $x \in \mathbb{R}^3 \setminus \overline{\Omega'}$:
- Continuity condition on the intermediate boundary, $\partial \Omega$:

$$\phi_i = \phi_{lay} \ , \ \epsilon_i \frac{\partial \phi_i}{\partial n(y)} = \epsilon_e \frac{\partial \phi_{lay}}{\partial n(y)}$$

• Continuity condition on the external boundary, $\partial \Omega'$:

$$\phi_{lay} = \phi_e \ , \ \frac{\partial \phi_{lay}}{\partial n(y)} = \frac{\partial \phi_e}{\partial n(y)}$$

Molecular Electrostatics Problem (Cont.)

Mathematical model (nonlinear)

• nonlinear Poisson-Boltzmann equation outside (1-to-1 electrolyte):

$$\Delta\phi_e(x) - \kappa^2 \sinh\phi_e(x) = 0$$

Second order approximation to the non-linear term:

$$\Delta\phi_e(x) - \kappa^2\phi_e(x) = \frac{\kappa^2}{6}\phi_e^3(x)$$





Capacitance of a Conductor G

$$C = -\frac{1}{4\pi} \int_{\Gamma} \frac{\partial u}{\partial n} \, ds \; ,$$

u – solution of the external Dirichlet problem for the Laplace equation

$$\Delta u(x) = 0 , \ x \in G_1 = \mathbb{R}^3 \setminus \overline{G} ,$$
$$u(y) = 1 , \ y \in \Gamma ,$$
$$\lim_{|x| \to \infty} u(x) = 0 .$$

By Green's formula

$$C = 4\pi R \overline{u}(R) = \mathbb{E} \left(4\pi u(R\omega) \right) = \mathbb{E} \left(4\pi \xi(R\omega) \right),$$

 ξ – Monte Carlo estimate for u, ω – unit isotropic vector, S(0, R) – sphere containing G.

Monte Carlo Estimates Based on Integral Equations

Consider u(x) to be an integral functional of an integral equation solution:

$$u(x) = \int_Y h_x(y)\mu(y)d\sigma(y)$$

$$\mu(y) = \int_Y k(y, y')\mu(y')d\sigma(y') + f(y) \equiv K\mu(y) + f(y)$$

Example:

In the energy calculation, assume there are no charges outside the 'molecule' G. The non-singular part of the solution can be represented as a single-layer potential

$$u^{(0)}(x) = \int_{\Gamma} \frac{1}{2\pi} \frac{1}{|x-y|} \mu(y) d\sigma(y)$$

^

Integral Equations (Cont.)

Potential's density satisfies the integral equation

$$\mu(y) = -\lambda_0 \int_{\Gamma} \frac{1}{2\pi} \frac{\cos \varphi_{yy'}}{|y - y'|^2} \mu(y') d\sigma(y') + f(y)$$

Here $\lambda_0 = \frac{\epsilon_e - \epsilon_i}{\epsilon_e + \epsilon_i}$. The Neumann series

$$\sum_{i=0}^{\infty} (-\lambda_0 K)^i f$$

for this equation converges, but slowly.

Substitution of spectral parameter to speed up the convergence:

00

$$\mu = \sum_{i=0}^{n} l_i^{(n)} (-\lambda_0 K)^i f + O(q^{n+1})$$

Integral Equations (Cont.)

Monte Carlo Estimate

Markov chain of *random walk on the boundary*: $p_0(y)$ – initial distribution density

$$p(y_i \to y_{i+1}) = \frac{1}{2\pi} \frac{\cos \varphi_{y_{i+1}y_i}}{|y_{i+1} - y_i|^2}$$

- transition density (uniform in the solid angle)

The estimate (biased, for a convex Γ)

$$u(x) = \mathbb{E}\left[\sum_{i=0}^{n} l_i^{(n)} (-\lambda_0)^i \frac{f(y_0)}{p_0(y_0)} h_x(y_i)\right]$$

Capacitance: Random Walk on the Boundary

Capacitance

$$C = \int_{\Gamma} \mu(y) \ d\sigma(y)$$

Charge distribution

$$\mu(y) = -\frac{1}{4\pi} \frac{\partial u}{\partial n}(y)$$

is the eigenfunction of the integral operator K:

$$\mu(y) = \int_{\Gamma} \frac{\cos \varphi_{yy'}}{2\pi |y - y'|^2} \mu(y') d\sigma(y')$$

Capacitance: Random Walk on the Boundary (Cont.)

For a convex G, stationary distribution of isotropic random walk on boundary:

$$\pi_{\infty} = \frac{1}{C} \mu$$

By the ergodic theorem

$$C = \left(\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} v(y_n)\right)^{-1}$$

for $v(y) = \frac{1}{|x-y|}$ (arbitrary $x \in G$), since inside G the potential $\int_{\Gamma} \frac{1}{|x-y'|} \mu(y') d\sigma(y') = 1.$





First-Passage Results: Cumulative Charge Distribution

Charge Density on a Circular Disk via Last-Passage

Approach from the Outside

• P(x): prob. of diffusing from ϵ above lower FP surface to ∞

$$P(x) = \int_{\partial \Omega_y} g(x, y, \epsilon) p(y, \infty) dS$$
(15)

$$\sigma(x) = -\frac{1}{4\pi} \frac{d}{d\epsilon} \bigg|_{\epsilon=0} \phi(x) = \frac{1}{4\pi} \frac{d}{d\epsilon} \bigg|_{\epsilon=0} P(x) \quad (16)$$

$$\sigma(x) = \frac{1}{4\pi} \int G(x, y) p(y, \infty) dS \quad (17)$$

$$F(x) = \frac{1}{4\pi} \int_{\partial \Omega_y} G(x, y) p(y, \infty) dS$$

where

$$G(x,y) = \frac{d}{d\epsilon} \bigg|_{\epsilon=0} g(x,y,\epsilon)$$
(18)

• G(x, y) satisfies a point dipole problem

Unit Cube Edge Distribution (Cont.)

$$\sigma(x,\delta_e) = \delta_e^{\pi/\alpha - 1} \sigma_e(x) \tag{19}$$

• $\sigma(x, \delta_e)$: charge on a curve parallel to the edge separated by δ_e

- $\sigma_e(x)$: edge distribution
- α : angle between the two intersecting surfaces, here $\alpha = 3\pi/2$

$$\sigma_e(x) = \frac{1}{4\pi} \lim_{\delta_e \to 0} \delta_e^{1-\pi/\alpha} \int_{\partial \Omega_e} G(x, y) p(y, \infty) dS$$
(20)

• $\partial \Omega_e$: cylindrical surface that intersects the pair of absorbing surfaces meeting at angle α

near the corner is approximately -0.20, that is, $\sigma_e \sim \delta_c^{-1/5}$

Conclusions and Future Work

Conclusions

- Stochastic algorithms are very effective in a wide range of partial differential equation and integral equation settings
- Efficiency comes from choosing among the appropriate variant: WOS, GFFP, S-T, 'Walk on the Boundary,' or 'Walk on Subdomains'
- Many applications can be addresses, here the examples are related through electrostatics

Conclusions and Future Work (Cont.)

Future Work

- Molecular Electrostatics
 - More complicated functionals of the solution
 - Derivatives (forces)
 - Nonlinear problem via branching processes and expansions
- Multiscale Monte Carlo

Local Bibliography

- M. Mascagni and N. A. Simonov (2004), "Monte Carlo Methods for Calculating Some Physical Properties of Large Molecules," *SIAM Journal on Scientific Computing*, **26(1)**: 339–357.
- N. A. Simonov and M. Mascagni (2004), "Random Walk Algorithms for Estimating Effective Properties of Digitized Porous Media," *Monte Carlo Methods and Applications*, **10**: 599–608.
- M. Mascagni and N. A. Simonov (2004), "The Random Walk on the Boundary Method for Calculating Capacitance," *Journal of Computational Physics*, **195(2)**: 465–473.
- C.-O. Hwang and M. Mascagni (2003), "Analysis and Comparison of Green's Function First-Passage Algorithms with "Walk on Spheres" Algorithms," *Mathematics and Computers in* Simulation, **63**: 605-613.
- J. A. Given, C.-O. Hwang and M. Mascagni (2002), "First- and last-passage Monte Carlo algorithms for the charge density distribution on a conducting surface," *Physical Review E*, 66, 056704, 8 pages.
- (2001), "The Simulation-Tabulation Method for Classical Diffusion Monte Carlo," Journal of Computational Physics, 174: 925–946.
- C.-O. Hwang, J. A. Given, and M. Mascagni (2000), "On the Rapid Calculation of Permeability for Porous Media Using Brownian Motion Paths," *Physics of Fluids*, **12**: 1699–1709.