

FINDING THE NEEDLE IN THE HAYSTACK: ALGORITHMS FOR CONFORMATIONAL OPTIMIZATION

Ioan Andricioaei and
John E. Straub

Department Editors:

Harvey Gould

hgould@clarku.edu

Jan Tobochnik

jant@kzoo.edu

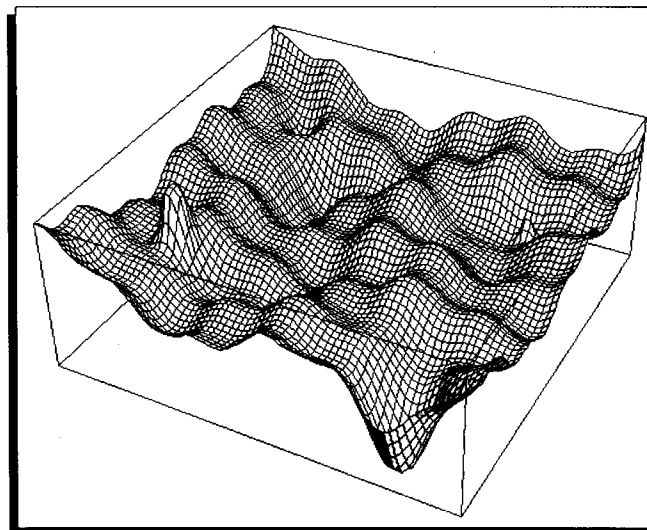


Figure 1. Schematic representation of the geographical analogy to the potential-energy surface. A downhill walk on this landscape starting anywhere in the upper left corner might lead to the erroneous conclusion that the first minimum encountered is lowest on the surface.

The search for the thermodynamic ground state of a system such as a cluster, glass, protein, or other macromolecule has many important applications.¹ An example is the protein-folding problem² from molecular biophysics. Finding the native conformation of a protein (the set of positions of the atoms of the protein in its biologically active state) has implications for drug design. It also is of much research interest because it is an example of self-organization of a complex system.

If we adopt the hypothesis that the native conformation of a protein is the one that minimizes the free energy, we can predict this conformation by finding the free energy as a function of some subset of the atomic coordinates. Assuming that the folded state of lowest free energy is the state of lowest potential energy, we see that the protein-folding problem becomes a problem of global energy minimization. However, even small proteins have hundreds of atoms, and finding the lowest-energy conformation in a many-dimensional configuration space is a computationally demanding problem. This problem is complicated by the presence of a large number of local minima separated by a multiple-scale distribution of energy barriers.

To obtain a pictorial view of the problem, let us imagine a rugged geographical landscape (an analog of the potential-energy surface) with many scattered hills and valleys having various heights (see Fig. 1).³ Our goal is to find a general method for isolating the lowest point on this hy-

persurface for systems that cannot be exhaustively searched. Starting from a random point and simply moving downhill by steepest descent according to

$$\frac{d\mathbf{r}}{ds} = \mathbf{F}(\mathbf{r}), \quad (1)$$

where the force $\mathbf{F}(\mathbf{r}) = -\nabla U(\mathbf{r})$ is derived from the potential $U(\mathbf{r})$ and s is the step size, is not a good idea in general. The problem is that on a rugged landscape the system will be trapped with overwhelming probability in a local rather than the global minimum.

Simulated annealing

A popular algorithm used to find the global minimum of a potential-energy function is called *simulated annealing*.⁴⁻⁶ The idea is analogous to the physical process of cooling a hot glass. If the cooling is very fast, the quench brings the glass into a strained state, where it can be easily broken. If the cooling schedule is sufficiently slow, an unstrained glass is obtained at the end of the annealing.

In terms of a topographical analogy, the simulated-annealing process is similar to the effects of an earthquake. Imagine a marble rolling on the landscape. The amplitude of the earthquake, analogous to the thermal energy kT , can make the marble jump out of its local minimum and move

Ioan Andricioaei is a graduate student in theoretical chemical physics, and John E. Straub is an associate professor in the Department of Chemistry, Boston University, 590 Commonwealth Avenue, Boston, MA 02215. E-mail: andricio@chem.bu.edu, straub@chem.bu.edu

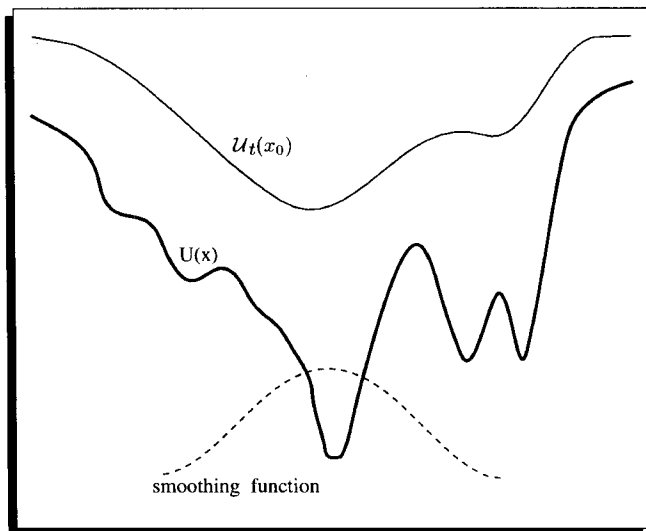


Figure 2. The effect of the coarse-graining average by a convolution of the original rugged potential with a smoothing function. The minima are raised while the maxima are somewhat lowered.

it to lower ground. By slowly decreasing the amplitude of the earthquake, the marble will decrease its tendency to jump out of minima. If this decrease is well timed, the marble will finally settle in the global minimum.

More formally, after having chosen a cooling schedule (defining how the temperature monotonically decreases as time progresses), we need to generate representative configurations of the allowed microstates using either molecular dynamics or Monte Carlo-type algorithms. Using molecular dynamics, we can solve Newton's equations of motion for each atom in the system by numerical integration. The temperature is controlled through coupling to a heat bath.⁷ In the standard Monte Carlo approach, a trial move is accepted or rejected according to a temperature-dependent update probability of the Metropolis type⁷

$$p = \min[1, e^{-\beta\Delta U}], \quad (2)$$

where $\beta = 1/kT$, and ΔU is the change in the potential energy. This acceptance probability satisfies the principle of detailed balance, that is, microscopic reversibility, assuring that a stationary probability will be reached eventually. At low T , only moves that lead to equal or lower energy ($\Delta U \leq 0$) are accepted with significant probability. In this limit, the algorithm reduces to a steepest-descent minimizer.

It can be shown that the duration of a simulation (measured in molecular-dynamics time steps or Metropolis moves) required to isolate the global energy minimum scales as⁸

$$T_{\text{sim}} \sim e^{E_{\text{max}}/E_{\text{min}}}, \quad (3)$$

where E_{max} is the maximum value in the set of all minimum energy barriers separating basins on the energy landscape, and E_{min} is the difference between the global energy minimum and the lowest local energy minimum. It turns out that

many systems of interest have a broad distribution of energy scales so that $E_{\text{max}}/E_{\text{min}}$ is large. As we discuss in the following several methods have been developed to address this problem by performing a smoothing transformation on the energy landscape.

Potential smoothing

One approach to the problem of finding the global minimum is to smooth the potential surface. The smooth potential can be obtained by any average that has a coarse-graining effect (see Fig. 2). This method stems from the observation that the efficiency of the search for the minimum of a multiextremal function is increased by replacing the function with an integral transform having an appropriate kernel and performing the search on this smoothed function.⁹

A helpful analogy is furnished by a heavy snowfall on a rugged landscape. Rolling downhill is easy on the snowy surface and is equivalent to a steep descent. After having reached the lowest surface on the snow-covered landscape, one has to wait for the inverse transformation of snowing, that is, melting, and then gradually adjust the position by local minimization. After the melting is completed, the hope is that one ends in the lowest valley.

In the field of conformational optimization, potential-smoothing methods have been applied with great success when the kernel is Gaussian.⁸ The objective is to isolate the global minimum of $U(\mathbf{r})$, the function describing the potential-energy hypersurface.¹⁰ The transform is most easily performed when the original hypersurface to be smoothed is fitted by a sum of Gaussians. For a d -dimensional Gaussian potential

$$U(\mathbf{r}) = e^{-\sum_{i=1}^d a_i x_i^2}, \quad (4)$$

the smoothed potential is the Fourier-Poisson integral transform of $U(\mathbf{r})$:¹¹

$$\mathcal{U}_t(\mathbf{r}_0) = (4\pi t)^{-d/2} \int d\mathbf{r} U(\mathbf{r}) e^{-(\mathbf{r}-\mathbf{r}_0)^2/4t}. \quad (5)$$

The transform is

$$\mathcal{U}_t(\mathbf{r}_0) = \prod_{i=1}^d (1 + 4a_i t)^{-1/2} e^{-a_i x_{0i}^2 / (1 + 4a_i t)}, \quad (6)$$

where x_{0i} is the i th component of \mathbf{r}_0 and the smoothing increases with the parameter t . In the study of atomic clusters or proteins, it often is desirable to restrict the search to a limited region of configuration space (such as the set of compact states) using a "boundary potential." For a proper choice of boundary conditions, the smoothed potential $\mathcal{U}_t(\mathbf{r}_0)$ will be a convex function when t is sufficiently large. Finding the minimum of this function is now trivial (by using a steepest-descent method).

(1) Use the force derived from the smoothed potential to perform a steepest descent

$$\frac{d\mathbf{r}_0}{ds} = -\nabla \mathcal{U}_t(\mathbf{r}_0). \quad (7)$$

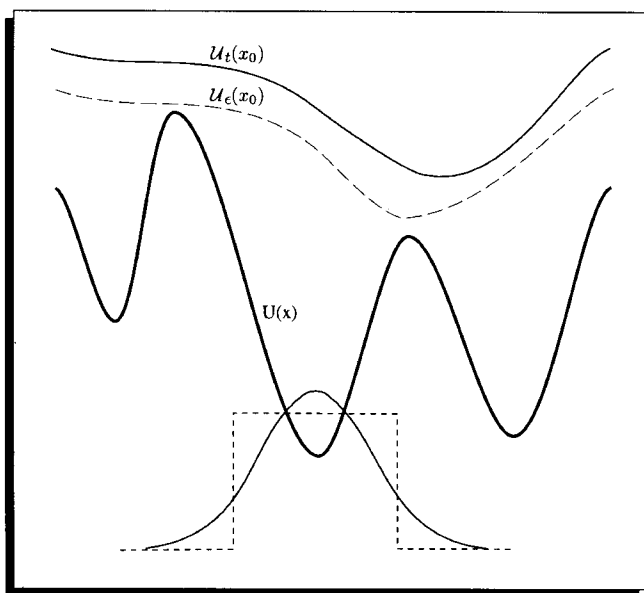


Figure 3. Comparison of the smoothing properties of the Gaussian and the impulse functions, respectively. Although their effect is similar, the impulse function-smoothing transform of the potential functions requires the evaluation of only the potential function itself.

Initially set t to a relatively large value.

(2) Reduce t by an amount Δt .

(3) As long as t is significantly different from zero, go to step 1; otherwise, stop.

This step-wise process is repeated until the value of t is brought close to zero, thereby reversing the transformation and retrieving the global minimum of the original function. A shortcoming of the Gaussian smoothing is that, for most potentials of interest, the integral transform for determining $\mathcal{U}_t(\mathbf{r}_0)$ using Eq. (5) is possible but cumbersome.

Singularities in the potential energy $U(\mathbf{r})$, which arise from the core repulsion of overlapping atoms, lead to serious problems in optimization methods based on potential-energy smoothing methods and local energy minimization. Typically, singular potentials of the type $1/r$ (Ref. 12) are replaced by exponential or Gaussian core repulsions. It is recognized that a more appealing solution is to coarse grain the probability density $\rho(\mathbf{r}) = \exp[-\beta U(\mathbf{r})]/Z(\beta)$, where $Z(\beta) = \int d\mathbf{r} \exp[-\beta U(\mathbf{r})]$ is the configurational integral.¹² Singular core repulsions in $U(\mathbf{r})$ are replaced by regions of low probability density in $\rho(\mathbf{r})$ which is a well-behaved, normalized function. However, a Gaussian integral transform of the probability density is difficult. For all except the most simple potentials, one must resort to numerical quadrature to perform the transform of $\rho(\mathbf{r})$. In the following, we describe a promising alternative to Gaussian smoothing, which allows for coarse graining of $\rho(\mathbf{r})$ for a large class of problems.

The method of bad derivatives

The above smoothing procedure can be simplified by replacing the Gaussian kernel by an impulse function (see

Fig. 3). We give a one-dimensional example to illustrate the method. The kernel $I_\epsilon(x, x_0)$ is the impulse function

$$I_\epsilon(x, x_0) = \frac{1}{2\epsilon} \{ \Theta[x - (x_0 - \epsilon)] - \Theta[x - (x_0 + \epsilon)] \}, \quad (8)$$

where the Heaviside function $\Theta(x)$ is zero for $x < 0$ and unity for $x \geq 0$. The integral transform of the potential can be written as

$$\mathcal{U}_\epsilon(x_0) = \int_{-\infty}^{\infty} U(x) I_\epsilon(x, x_0) dx = \frac{1}{2\epsilon} \int_{x_0 - \epsilon}^{x_0 + \epsilon} U(x) dx, \quad (9)$$

where \mathcal{U}_ϵ is a coarse-grained potential. The force derived from the gradient of this coarse-grained potential has the remarkably simple form of

$$\mathcal{F}_\epsilon(x_0) \equiv -\nabla_{x_0} \mathcal{U}_\epsilon(x_0) = -\frac{1}{2\epsilon} [U(x_0 + \epsilon) - U(x_0 - \epsilon)]. \quad (10)$$

Equation (10) is a finite-difference approximation to the gradient of the actual potential. We need not keep ϵ small (as we would to obtain a good approximation to the gradient) because the smoothing is exact for all ϵ . However, in terms of a finite-difference interpretation, the derivative is “bad” for large enough ϵ . This “method of bad derivatives” may be applied as follows using a series of local energy minimizations on a decreasingly coarse-grained potential-energy function.

(1) Use the force in (10) to perform a steepest descent

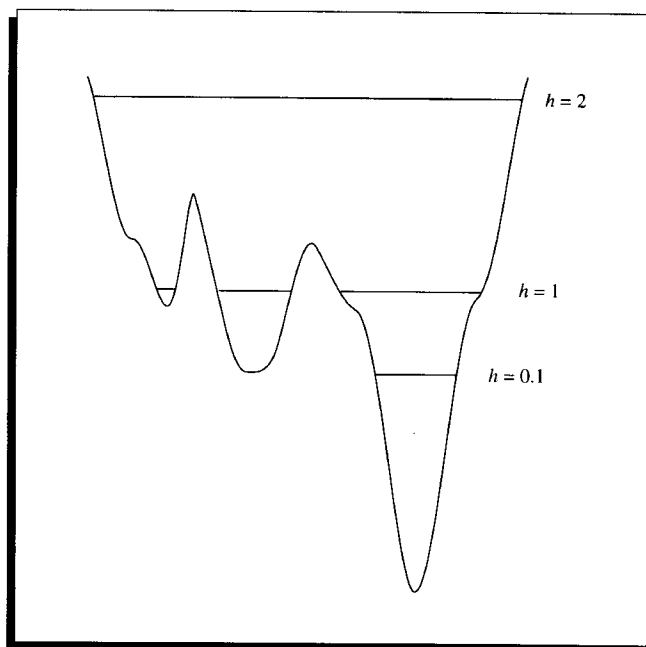


Figure 4. Schematic representation of the quantum-annealing method. The approximate ground state is shown for several values of the control parameter h . For large values of h , the packet wave function is delocalized. Decreasing h tends to localize the wave packet in the global minimum.

$$\frac{dx_0}{ds} = \mathcal{F}_\epsilon(x_0). \tag{11}$$

Initially set ϵ to a relatively large value.

(2) Reduced ϵ by an amount $\Delta\epsilon$.

(3) As long as ϵ is significantly different from zero, go to step 1; otherwise, stop.

The last solution gives the best guess for the position of the global energy minimum. Generalization to multidimensional energy surfaces requires approximation but a rather straight forward generalization of the finite difference derivative suggests itself. Alternatively, the steepest-descent algorithm may be replaced by a simulated-annealing protocol employing molecular dynamics or Monte Carlo. In contrast to Gaussian transforms of the potential, there are no approximations, transforms, or derivatives involved in the computation of the force $\mathcal{F}_\epsilon(x_0)$. Because the method is “derivative free,” it is of interest for problems where one seeks the extremum of a nondifferentiable function (such as the Weierstrass function or a stochastic signal).

Quantum-mechanical annealing

Potential smoothing algorithms provide a means of lowering the energy barriers to allow for more rapid thermally activated barrier crossing. Another way of enhancing barrier crossing is to enhance the probability of tunneling through energy barriers using quantum-mechanical annealing. The motivations for exploring quantum-mechanical annealing are twofold. The most obvious is that, although classical trajectories are confined in potential-energy wells higher than the energy of the system, quantum systems may tunnel through energy barriers. Second, while a classical particle is localized and sees the potential only at a single point in space, the quantum system can be delocalized and can use nonlocal information to find low energy wells. In terms of the landscape analogue of the energy surface, instead of climbing over the hills, the walker is provided with the means to bore a hole through the hillside into the adjacent valley. The landscape would be transparent for the walker, so that, even on a “golf course,” the holes would be visible from a distance.¹³

The quantum-mechanical annealing protocol consists of finding an estimate of the ground-state wave function for large \hbar and taking the classical limit $\hbar \rightarrow 0$ (see Fig. 4). In this way, the final probability density associated with the wave function will be localized in the region of the classical ground state configuration.

We focus on two algorithms based on the approximate solution of the imaginary-time Schrödinger equation. For the stationary-state wave functions $\Psi_n(\mathbf{r}, \tau)$, the time-dependent wave equation is

$$\begin{aligned} \frac{\partial \Psi_n(\mathbf{r}, \tau)}{\partial \tau} &= \frac{\hbar^2}{2m} \nabla^2 \Psi_n(\mathbf{r}, \tau) - U(\mathbf{r}) \Psi_n(\mathbf{r}, \tau) \\ &= -E_n \Psi_n(\mathbf{r}, \tau), \end{aligned} \tag{12}$$

where $\tau = it/\hbar$ is the imaginary time, and E_n are the eigenvalues of the energy corresponding to the eigenfunctions

$\psi_n(\mathbf{r})$. By integrating (12) once with respect to τ , we find

$$\Psi_n(\mathbf{r}, \tau) = e^{-E_n \tau} \psi_n(\mathbf{r}), \tag{13}$$

where $\psi_n(\mathbf{r})$ is an eigenfunction of the coordinates only [the solution of $H\psi_n(\mathbf{r}) = E_n\psi_n(\mathbf{r})$]. We can expand an arbitrary wave function Ψ in terms of the eigenfunctions of the stationary states:

$$\Psi(\mathbf{r}, \tau) = \sum_n a_n e^{-E_n \tau} \psi_n(\mathbf{r}). \tag{14}$$

In the limit of large τ , we can approximate (14) as

$$\Psi(\mathbf{r}, \tau) \sim e^{-E_0 \tau} \psi_0(\mathbf{r}). \tag{15}$$

That is, the ground-state wave function of energy E_0 makes an exponentially larger contribution to Ψ than any excited state with higher energy.

Note that the Schrödinger equation, written as

$$\frac{\partial \Psi}{\partial \tau} = \frac{\hbar^2}{2m} \nabla^2 \Psi - [U(\mathbf{r}) - E_0] \Psi, \tag{16}$$

has the same form as a diffusion equation for the concentration $\Psi(\mathbf{r})$. Additionally, there is a birth–death term which increases (decreases) the concentration if the ground state E_0 (initially chosen as a trial energy) is greater than (less than) the potential energy $U(\mathbf{r})$. The simulation of such a process consists of evolving a number of “random walkers” that simulate the diffusion and the birth–death processes. E_0 is iteratively adjusted to create a steady-state population of walkers. This method is known as diffusion quantum Monte Carlo.¹⁴ It may be used to find the ground-state wave function and the ground-state energy at a given value of $\hbar^2/2m$. By taking the classical limit ($\hbar \rightarrow 0$), we are able to find the global potential energy minimum of the system.¹⁵

Another possibility is to solve the imaginary-time Schrödinger equation approximately. One may use a Gaussian wave-packet representation of the wave function. If \mathbf{r} is a d -dimensional vector of Gaussian random variables, the corresponding multivariate probability-density function in the spherically symmetric case has the form¹⁶

$$\Psi(\mathbf{r}, \tau) = (2\pi\sigma^2)^{-d/4} e^{-(\mathbf{r}-\mathbf{r}_0)^2/4\sigma^2}, \tag{17}$$

where the center of the packet is at \mathbf{r}_0 , and the second moment is $M_2 = d\sigma^2$. An approximate wave function for a many-body system consists of a Hartree product of single-particle, spherically symmetric Gaussian wave packets. The wave function evolves in imaginary time according to the wave equation (12). In general, for an expectation value $\langle A \rangle$ the equation of motion is¹⁷

$$\frac{d\langle A \rangle}{d\tau} = -\langle [A, H] \rangle + \left\langle \frac{\partial A}{\partial \tau} \right\rangle. \tag{18}$$

Using $\langle \mathbf{r} \rangle = \mathbf{r}_0$, the position of the center, and $\langle (\mathbf{r} - \mathbf{r}_0)^2 \rangle = M_2$, the variance of the Gaussian wave packet, we may write the imaginary-time equations of motion in the form¹⁸

$$\frac{d\mathbf{r}_0}{d\tau} = -\frac{2}{d} M_2 \nabla_{\mathbf{r}_0} \langle U \rangle, \quad (19)$$

$$\frac{dM_2}{d\tau} = \frac{d\hbar^2}{2m} - \frac{2}{d^2} M_2^2 \nabla_{\mathbf{r}_0}^2 \langle U \rangle. \quad (20)$$

$\langle U \rangle$ is the potential energy averaged by weighting with the Gaussian wave-packet probability density

$$\langle U \rangle = \int d\mathbf{r} \Psi^2(\mathbf{r}, \tau) U(\mathbf{r}). \quad (21)$$

This averaged potential has the form of the Gaussian transform discussed earlier [see Eq. (5)].

The algorithm consists of the three steps.

(1) Solve the equations of motion (19) and (20), initially setting h to a relatively large value.

(2) Reduce h by an amount Δh .

(3) As long as h is significantly different from zero, go to step 1; otherwise, stop.

Note that the center \mathbf{r}_0 moves in a very steep descent according to the effective potential $\langle U \rangle$. The second moment of the free particle wave packet spreads at a constant rate of $d\hbar^2/2m$; the term proportional to the curvature of the effective potential acts to decrease (increase) the rate of expansion when the curvature is positive (negative).

The ability of the wave packets to expand and tunnel is enhanced as h is increased. In the classical limit ($\hbar=0$), the global minimum of the energy surface can be found after a quantum-mechanical evolution.¹⁸ This method has been successfully applied to Lennard-Jones clusters and a related algorithm has been employed to fold model proteins.¹⁹

In general, conformational optimization of many-body systems in molecular physics and biophysics remains an unsolved problem. In recent years, various algorithms that use potential smoothing to simplify the problem of isolating the global energy minimum have been devised. Although Gaussian transforms have led to important results, the method of bad derivatives, which requires function evaluations only and is derivative free, appears to be promising. Quantum-mechanical annealing is an optimization paradigm that should lead to a variety of optimization algorithms based on alternative representations of the system, such as path integrals, which are well suited for condensed-phase simulation.

Suggestions for further study

(1) Find the minimum-energy conformation of a cluster of 13 Lennard-Jones particles using a Monte Carlo simulated-annealing protocol. The Lennard-Jones potential for the interaction of two rare-gas atoms is

$$U(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right], \quad (22)$$

where σ is the atomic diameter, and ϵ is the depth of the

energy minimum located at $r = 2^{1/6}\sigma$. The global energy minimum of the 13-atom cluster is an icosahedron (12 atoms at the vertices and one in the center) with an energy of -44.327ϵ . Use a boundary potential to confine the particles (and thus prevent dissociation) of the form²⁰

$$V_{\text{conf}} = \frac{k}{2} \sum_{i>j=1}^{13} r_{ij}^2, \quad (23)$$

with $k=0.0001$. The initial positions of the atoms can be placed randomly in a cube.

(2) Scheraga and co-workers¹¹ have examined the one-dimensional conformational model of a substituted ethane molecule ($\text{H}_3\text{C}-\text{CH}_3$) for which the only variable is the torsion angle θ (as you look down the carbon-carbon bond). The form of the potential as a function of θ is

$$U(\theta) = \sum_{i=1}^3 \sum_{j=1}^3 [b_{ij} r_{ij}^{-12}(\theta) - a_{ij} r_{ij}^{-6}(\theta)], \quad (24)$$

where

$$r_{ij}^2(\theta) = \alpha_e + \beta_e \cos(\theta + \Gamma_{ij}).$$

The parameters α_e and β_e are given by $\alpha_e = L^2 + 2l^2$ and $\beta_e = -2l^2$, where $L = R(1 - 2\cos\theta)$ and $l = R\sin\phi$ with $R = 1.54 \text{ \AA}$ and $\phi = 109.5^\circ$. Γ_{ij} are the elements of the matrix

$$\Gamma = \frac{2\pi}{3} \begin{pmatrix} 0 & -1 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix}. \quad (26)$$

The values of b_{ij} (in units of kcal $\text{\AA}^{12}/\text{mol}$) and a_{ij} (in units of 10^4 kcal $\text{\AA}^6/\text{mol}$) are $b_{11}=372.5$, $b_{12}=348.7$, $b_{13}=367.2$, $b_{22}=344.7$, $b_{23}=349.5$, $b_{33}=367.2$, and $a_{11}=28.58$, $a_{12}=16.82$, $a_{13}=20.52$, $a_{22}=17.82$, $a_{23}=15.75$, $a_{33}=14.49$. Find the global minimum of this function using both the Gaussian and the impulse function as smoothing kernels.

(3) Write a program that uses the method of bad derivatives to calculate the global maximum of the function $\exp[-\beta U(\mathbf{r})]$, where $U(\mathbf{r})$ is a nonconvex potential-energy function. A one-dimensional example of U is²¹

$$U(x) = -\log(5e^{-10(x+1)^2} + 50e^{-10(x-2)^2} + 10e^{-(x-3)^2} + 30e^{-10(x-4)^2}). \quad (27)$$

Comment on the relation between the maximum value of ϵ (the smoothing width of the impulse function) and the characteristic length scales in the problem. Also comment on the relationship between this algorithm and the search for the free-energy minimum, rather than the potential-energy minimum.

(4) Apply the quantum annealing method employing Gaussian wave packets to an asymmetric double-well potential of the form

$$U(r) = \frac{7}{8}r^4 - \frac{1}{4}r^2 - \frac{1}{8}r. \quad (28)$$

Plot the approximate ground-state wave function as the parameter $D = \hbar^2/2m$ is decreased toward zero and notice the localization of the wave function. For nonzero values of D , you should see some delocalization because of the contribution from the second minimum to the approximate ground state.

(5) Show that the $\hbar=0$ limit of the imaginary-time quantum annealing method corresponds to a classical simulated-annealing algorithm based on direct evolution in reciprocal temperature (no real time) of the equilibrium density distribution according to the classical Bloch equation²²

$$\frac{\partial \rho(\mathbf{r}, \beta)}{\partial \beta} = -(H - \langle H \rangle) \rho(\mathbf{r}, \beta), \quad (29)$$

where H is the Hamiltonian. Demonstrate that even in the $\hbar=0$ limit, the classical density distribution can tunnel through energy barriers and enhance the sampling of the potential-energy surface.²⁰

Acknowledgments

One of the authors (J.E.S.) received support from the Alfred P. Sloan Foundation and recognizes the generous support of the National Science Foundation (Grant No. CHE-9306375). The authors thank the department editors for their helpful comments.

From the editors. Please send us your comments and suggestions for future columns.

References

1. The problem belongs to the general class of different search problems that have the universal property called nondeterministic polynomial (NP) time completeness. These NP-complete search problems require a computational time that increases faster than a polynomial-time (P) function of the size of the problem. The question, "does an efficient solution for a NP-complete problem exist," known as the P=NP problem, is a major focus of work in optimization theory. This problem is universal in the sense that if one member of the class of problems can be solved, the others can be too.
2. H. A. Scheraga, *Rev. Comp. Chem.* **III**, 73 (1992); *Theor. Biochem. Mol. Biophys.* **2**, 231 (1991); *Chem. Scr. A* **29**, 3 (1989).
3. Of course, the real energy surface is a many-dimensional one, and there are qualitatively different aspects as the number of dimensions increases. If the potential is one-dimensional, it is necessary to climb the barrier between the two minima to go from one local minimum to the next. However, if the potential is two-dimensional, the next minima is also reachable by going around the hill. In fact, "relaxing the dimensionality" has been used effectively in the field of optimization. See, for example, G. M. Crippen, *J. Comp. Chem.* **3**, 471 (1982); *ibid.* **5**, 548 (1984); E. O. Purisma and H. A. Scheraga, *Proc. Natl. Acad. Sci.* **83**, 2782 (1986); *J. Mol. Biol.* **196**, 697 (1987).
4. S. Kirkpatrick, C. D. Gelatt, Jr., and M. P. Vecchi, *Science* **220**, 671 (1983).
5. E. Aarts and J. Korst, *Simulated Annealing and Boltzmann Machines* (Wiley, New York, 1990).
6. A. Silverman and J. Adler, *Comput. Phys.* **6**, 277 (1992).
7. M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford University Press, Bristol, 1990); H. Gould and J. Tobochnik, *An Introduction to Computer Simulation Methods: Applications to Physical Systems*, 2nd ed. (Addison-Wesley, Reading, MA, 1996).
8. P. Amara, J. Ma, and J. E. Straub, *DIMACS Series in Discrete Mathematics and Theoretical Computer Science* (American Mathematical Society, Washington DC, 1996), Vol. 23; J. E. Straub, in *New Developments in Theoretical Studies in Proteins*, edited by R. Elber (World Scientific, Singapore, 1996).
9. V. V. Zakharov, *Eng. Cybernet.* **4**, 637 (1970).
10. It was shown that this transformation is equivalent to a "diffusion" process where the rugged potential surface is treated as an initially nonuniform concentration profile (Ref. 11). The smoothing is similar to the uniformization of the concentration profile, modeled by the diffusion equation, $\partial \mathcal{C} / \partial t = \nabla^2 \mathcal{C}$, with the initial condition $\mathcal{C}(\mathbf{r}, 0) = U(\mathbf{r})$. It is well known that the solution of the diffusion equation has a uniform concentration distribution as an attractor. The hope is that, starting with a rugged multiextremal distribution, a convex hypersurface is obtained at some point towards the convergence towards the attractor. In one dimension it can be shown that the last surviving minimum corresponds to the basin of greatest area.
11. L. Piela, J. Kostrowicki, and H. A. Scheraga, *J. Phys. Chem.* **93**, 3339 (1989); J. Kostrowicki, L. Piela, B. J. Cherayil, and H. Scheraga, *ibid.* **95**, 4113 (1991).
12. D. Shalloway, in *Recent Advances in Global Optimization*, edited by C. A. Floudas and P. M. Pardalos (Princeton University Press, Princeton, 1992), p. 433.
13. E. B. Baum, *Phys. Rev. Lett.* **57**, 2764 (1986).
14. D. Ceperley and B. Alder, *Science* **231**, 555 (1986).
15. A. B. Finnila, M. A. Gomez, C. Sebenik, C. Stenson, and J. D. Doll, *Chem. Phys. Lett.* **219**, 343 (1994).
16. C. W. Gardiner, *Handbook of Stochastic Methods* (Springer, Berlin, 1983).
17. A. Messiah, *Quantum Mechanics* (North-Holland, Amsterdam, 1961).
18. P. Amara, D. Hsu, and J. E. Straub, *J. Phys. Chem.* **97**, 6715 (1993).
19. P. Amara and J. E. Straub, *J. Phys. Chem.* **99**, 14840 (1995).
20. J. Ma and J. E. Straub, *J. Chem. Phys.* **101**, 533 (1994); *ibid.* **103**, 9113 (1995).
21. M. Orešič and D. Shalloway, *J. Chem. Phys.* **101**, 9844 (1994).
22. R. W. Zwanzig, in *Lectures in Theoretical Physics*, edited by W. E. Britton, B. W. Downs, and J. Downs (Wiley Interscience, New York, 1961), Vol. III, p. 106.