

Global minimization on rugged energy landscapes

Patricia Amara, Jianpeng Ma and John E. Straub

ABSTRACT. Three algorithms for global energy minimization based on the simulated annealing of the classical density distribution are presented. Two algorithms are based on the approximate solution of the classical Liouville equation for the dynamics of a system coupled to a heat bath using Fokker-Planck and Smoluchowski dynamics. A third algorithm is based on annealing the classical density distribution directly in temperature and is the classical analog of imaginary time quantum dynamics. These three methods are compared with standard simulated annealing based on molecular dynamics. The results for a model potential, Lennard-Jones clusters, water clusters, a model homopolymer, and a model heteropolymer demonstrate that by annealing the continuous density distribution (representing a volume of phase or configuration space) the likelihood of finding the global minimum is dramatically enhanced.

1. Introduction

When a system must be described quantum mechanically, except for the most simple systems, approximate basis set methods and perturbation theory must be employed. However, once one assumes that a system (for example, a liquid, crystal or protein) is well described by classical mechanics, one has the luxury of calculating the classical dynamics exactly by the numerical solution of Newton's equations of motion. To some extent, this explains why relatively little work has been done to develop approximate classical dynamical simulation algorithms.

When one wishes to find the global energy minimum of a complex molecular system, or search for binding sites on a protein for a ligand or potential drug, one finds that exact molecular dynamics simulation does not provide the most effective means to sample the potential energy hypersurface [1]. As a result, approximate classical molecular dynamics methods have been developed to address these hard problems; many are based on the dynamics of a continuous classical density distribution (representing an ensemble of systems) which is defined by the Liouville equation. Due to the similarity between the Liouville equation of classical statistical mechanics and the quantum mechanical von Neumann (quantum Liouville) equation [2, 3] and Schrödinger equation [4, 5, 6], many of the methods developed

1991 *Mathematics Subject Classification.* Primary 65K10, 90C30; Secondary 92C05, 92E10.
JES gratefully acknowledges the support of the National Science Foundation (CHE-9306375).

to treat quantum mechanical dynamics have been extended as approximate, coarse grained classical dynamics methods [7, 8, 9, 10, 11, 12].

In this paper, we summarize the development of methods based on the dynamical and thermodynamical evolution of the classical density distribution. We present general derivations of moment equations for the classical density distribution. Just as methods based on Gaussian packet dynamics have been particularly effective in providing insights and useful algorithms in quantum dynamics, we have found that approximate Gaussian density distribution based methods are effective in providing both (1) effective “black box” optimization algorithms and (2) a profitable way of thinking about global optimization problems where many potentially useful areas of research remain unexplored. Results from the application of these methods to a number of optimization problems including atomic clusters, water clusters, model homopolymer collapse and model protein folding are described. These applications help to provide an understanding of the mechanism used by these optimization methods as well as their limitations and long suits.

2. Evolution of the classical density distribution in time and temperature

In this section we present simulated annealing algorithms based on the evolution of the classical density distribution defined by three different equations of motion. (1) The Fokker-Planck equation is used to derive equations of motion in time for the classical phase space density distribution $\rho(\mathbf{x}, \mathbf{p}, t)$. (2) The Smoluchowski equation is used to derive equations of motion in time for a reduced configurational space density distribution $\hat{\rho}(\mathbf{x}, t)$ where the inertial terms dependent on the momenta are removed. (3) The Bloch equation is used to derive equations of motion in *temperature* for the configurational density distribution $\hat{\rho}(\mathbf{x}, \beta)$. In each case, we first derive the general moment equations which provide an exact description of the evolution of the distribution in time or temperature. We then focus on the special case of the Gaussian density approximation.

2.1. Fokker-Planck packet dynamics of $\rho(\mathbf{x}, \mathbf{p}, t)$.

2.1.1. Time evolution of the total position and momentum moments $\mathbf{M}_{n,k}$.

The total position and momentum moment $\mathbf{M}_{n,k}$ is defined in general as a symmetrized tensor of the form

$$\mathbf{M}_{n,k} = \int d\mathbf{x}d\mathbf{p} \mathcal{P}(\mathcal{M}^{n,k}) \rho = \int d\mathbf{x}d\mathbf{p} \mathcal{P}(\Delta\mathbf{x}^n \Delta\mathbf{p}^k) \rho$$

where $\mathcal{P}(\cdot)$ stands for the average over all possible permutations of the individual tensor indices and $n+k$ is the rank of the tensor $\mathbf{M}_{n,k}$. The increment of phase space is $d\mathbf{x}d\mathbf{p} = dx_1 dx_2 dx_3 dp_1 dp_2 dp_3$ (in three dimensions), $\Delta\mathbf{x} = \mathbf{x} - \mathbf{x}_0 = \Delta x_i \hat{e}_i$ and $\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0 = \Delta p_i \hat{e}_i$ where $\mathbf{x}_0 = \langle \mathbf{x} \rangle$ and $\mathbf{p}_0 = \langle \mathbf{p} \rangle$ and $\langle \cdot \rangle$ indicates an average over the normalized density distribution.

The Fokker-Planck equation for the time evolution of the phase space density distribution $\rho(\mathbf{x}, \mathbf{p}, t)$ is

$$(2.1) \quad \frac{\partial}{\partial t} \rho = - \left[\frac{p_i}{m} \frac{\partial}{\partial x_i} + F_i(\mathbf{x}) \frac{\partial}{\partial p_i} - \gamma \left[\frac{\partial}{\partial p_i} \left[p_i + m k_B T \frac{\partial}{\partial p_i} \right] \right] \right] \rho.$$

From this equation of motion, the time derivative of the moments $\mathbf{M}_{n,k}$ are defined

$$(2.2) \quad \frac{d}{dt} \mathbf{M}_{n,k} = \int d\mathbf{x} d\mathbf{p} \left[\mathcal{P} \left(\frac{\partial \mathcal{M}^{n,k}}{\partial t} \right) \rho + \mathcal{P} \left(\mathcal{M}^{n,k} \frac{\partial \rho}{\partial t} \right) \right].$$

The details of evaluation of this equation are given in Appendix A. The general result for the exact Fokker-Planck dynamics is

$$\frac{d}{dt} \mathbf{x}_0 = \frac{\mathbf{p}_0}{m} \quad \frac{d}{dt} \mathbf{p}_0 = \mathbf{F}_0 - \gamma \mathbf{p}_0$$

and

$$(2.3) \quad \frac{d}{dt} \mathbf{M}_{n,k} = \frac{n}{m} \mathbf{M}_{n-1,k+1} + k \mathbf{W}_{n,k-1} - \gamma [k \mathbf{M}_{n,k} - k(k-1) m k_B T \mathbf{L}_{n,k-2}]$$

where we have defined the $n+k$ rank tensors $\mathbf{W}_{n,k} = \langle \mathcal{P} (\mathcal{M}^{n,k} (F_i(\mathbf{x}) - F_{0i}) \hat{e}_i) \rangle$ and $\mathbf{L}_{n,k-2} = \langle \mathcal{P} (\mathcal{M}^{n,k-2} \hat{e}_i \hat{e}_i) \rangle$. This is our final result for the general moment equations which provide an exact description for the time evolution of the phase space density distribution in Fokker-Planck dynamics. Note that the equations of motion for the distribution center is similar to the Langevin equation for a single point in phase space. The center responds to the average force on the distribution

$$\mathbf{F}_0 = \int d\mathbf{x} d\mathbf{p} \mathbf{F}(\mathbf{x}) \rho.$$

In general, one cannot integrate the unbounded hierarchy of equations. One possible way to proceed is to define a basis for ρ and derive equations of motion for the parameters of that basis. An example is the Gaussian phase packet approximation.

2.1.2. Gaussian phase packet (GPP) Fokker-Planck dynamics.

The phase space density distribution for each degree of freedom may be approximated as a spherically symmetric GPP in the form

$$\rho(\mathbf{r}, \mathbf{p}, t) = \left(\frac{\sqrt{1-\alpha^2}}{2\pi\sigma_a\sigma_b} \right)^d \exp \left[-\frac{1}{2} \left(\frac{r-r_0}{\sigma_a} \right)^2 - \frac{1}{2} \left(\frac{p-p_0}{\sigma_b} \right)^2 - \alpha \left(\frac{r-r_0}{\sigma_a} \right) \left(\frac{p-p_0}{\sigma_b} \right) \right].$$

This approximation leads to the following time evolution for the center and second-order moments [13]

$$(2.4) \quad \begin{aligned} \dot{\mathbf{r}}_0 &= \frac{\mathbf{p}_0}{m} & \dot{\mathbf{p}}_0 &= -\nabla_{\mathbf{r}_0} \langle V \rangle - \gamma \mathbf{p}_0 \\ \dot{M}_{0,2} &= -\frac{2}{d} M_{1,1} \nabla_{\mathbf{r}_0}^2 \langle V \rangle - 2\gamma [M_{0,2} - d m k_B T] & \dot{M}_{2,0} &= \frac{2}{m} M_{1,1} \\ \dot{M}_{1,1} &= \frac{1}{m} M_{0,2} - \frac{1}{d} M_{2,0} \nabla_{\mathbf{r}_0}^2 \langle V \rangle - \gamma M_{1,1}. \end{aligned}$$

For a spherically symmetric Gaussian, the second rank tensor moments are isotropic so the scalars $M_{2,0}$, $M_{1,1}$ and $M_{0,2}$ are the trace of the second rank tensors $\mathbf{M}_{2,0}$, $\mathbf{M}_{1,1}$ and $\mathbf{M}_{0,2}$ respectively [14, 15]. These equations of motion have been used, along with the simulated annealing protocol, in the global optimization of a variety of systems discussed in the next section. In practice, one may control the external bath temperature T using a cooling schedule. The friction constant γ acts as an effective force constant to constrain the temperature of the system about the bath

temperature through the equation of motion for $M_{0,2}$. The center of the distribution follows a Langevin dynamics on a fluctuating effective potential energy surface

$$\langle V \rangle = \int d\mathbf{r} d\mathbf{p} V(\mathbf{r}) \rho.$$

This averaged potential may be thought of as a coarse grained potential function, averaged over the extent of the distribution. The effect of a broad distribution of $\langle V \rangle$ is to raise minima, lower barriers, and smooth the potential surface dramatically reducing the number of local minima. This effect of potential smoothing, which arises naturally from the dynamical equations for the density distribution, added to the simulated annealing protocol, effectively combines two of the paradigms of molecular energy minimization [16].

2.2. Smoluchowski packet dynamics of $\hat{\rho}(\mathbf{x}, t)$.

2.2.1. Time evolution of the total position moments \mathbf{M}_n .

In many situations, the momenta of a system relax quickly compared to the coordinates and it is possible to eliminate the momenta from the density distribution adiabatically. The resulting reduced density distribution function is $\hat{\rho}(\mathbf{x}, t)$ whose dynamics may be defined in terms of symmetrized n^{th} rank moment tensors in configuration space

$$\mathbf{M}_n = \int d\mathbf{x} \mathcal{P}(\mathcal{M}^n) \rho = \int d\mathbf{x} \mathcal{P}(\Delta \mathbf{x}^n) \rho.$$

The dynamics of the reduced configurational space distribution function $\hat{\rho}(\mathbf{x}, t)$ is defined by the Smoluchowski equation

$$(2.5) \quad \frac{\partial}{\partial t} \rho = \frac{1}{m\gamma} \left[\frac{\partial}{\partial x_i} \left[-F_i(\mathbf{x}) + k_B T \frac{\partial}{\partial x_i} \right] \rho \right].$$

The time derivative of the moments \mathbf{M}_n is given by

$$(2.6) \quad \frac{d}{dt} \mathbf{M}_n = \int d\mathbf{x} \mathcal{P} \left(\frac{\partial \mathcal{M}^n}{\partial t} \right) \rho + \mathcal{P} \left(\mathcal{M}^n \frac{\partial \rho}{\partial t} \right).$$

Evaluation of the integrals is described in Appendix B. Combined with the equation of motion for the center of the distribution, we arrive at

$$\frac{d}{dt} \mathbf{x}_0 = \frac{1}{m\gamma} \mathbf{F}_0$$

and

$$(2.7) \quad \frac{d}{dt} \mathbf{M}_n = \frac{1}{m\gamma} [n \mathbf{W}_{n-1} + k_B T n(n-1) \mathbf{L}_{n-2}]$$

where the n^{th} rank tensors are $\mathbf{M}_n = \langle \mathcal{P}(\mathcal{M}^n) \rangle$, $\mathbf{W}_n = \langle \mathcal{P}(\mathcal{M}^n (F_i(\mathbf{x}) - F_0)) \rangle$ and $\mathbf{L}_n = \langle \mathcal{P}(\mathcal{M}^{n-2} \hat{e}_i \hat{e}_i) \rangle$. These equations of motion describe the Smoluchowski dynamics of the density distribution exactly. Note that the equation of motion for the center of the distribution has the form of a steepest descent equation employing the force averaged over the density distribution.

2.2.2. GPP Smoluchowski dynamics.

The configurational distribution may be approximated as a spherical Gaussian

$$\hat{\rho}(\mathbf{r}, t) = \left(\frac{2\pi M_2}{d} \right)^{-d/2} \exp \left[-\frac{d}{2M_2} (\mathbf{r} - \mathbf{r}_0)^2 \right].$$

\mathbf{r}_0 is the center of the distribution and M_2 is the trace of the isotropic second rank tensor \mathbf{M}_2 . The resulting equations of motion for the center and second-order moment are

$$(2.8) \quad \begin{aligned} \frac{d\mathbf{r}_0}{dt} &= -\frac{1}{m\gamma} \nabla_{\mathbf{r}_0} \langle V \rangle \\ \frac{dM_2}{dt} &= \frac{1}{m\gamma} \left[2dk_B T - \frac{2}{d} M_2 \nabla_{\mathbf{r}_0}^2 \langle V \rangle \right]. \end{aligned}$$

The center of the distribution follows a steepest descent path on a dynamic effective potential. When the packets are wide and the distribution is delocalized the effective potential is smoothed. At low temperatures, the packet narrows and the exact physical potential surface is recovered. As in the case of Fokker-Planck dynamics, a simulated annealing protocol can be followed by controlling the bath temperature T . In Fokker-Planck dynamics, the system temperature fluctuates about the bath temperature T , while in Smoluchowski dynamics, the system temperature equals the bath temperature at all times.

2.3. Thermodynamical annealing in temperature of $\hat{\rho}(\mathbf{x}, \beta)$.

With all dynamical annealing algorithms, the choice of the cooling schedule is important. Typically, a realizable cooling schedule is far from optimal. Therefore, we have explored a different form of annealing where the density distribution is integrated in temperature directly and there is no real time dynamics.

2.3.1. Temperature evolution of the total position moments \mathbf{M}_n .

The temperature dependence of the reduced configurational distribution function $\hat{\rho}(\mathbf{x}, \beta)$ is described by the reduced Bloch equation

$$(2.9) \quad \frac{\partial}{\partial \beta} \hat{\rho}(\mathbf{x}, \beta) = -(V(\mathbf{x}) - \langle V(\mathbf{x}) \rangle) \hat{\rho}(\mathbf{x}, \beta).$$

The symmetrized tensor representing the position moment \mathbf{M}_n is

$$\mathbf{M}_n = \int d\mathbf{x} \mathcal{P}(\mathcal{M}^n) \rho = \int d\mathbf{x} \mathcal{P}(\Delta \mathbf{x}^n) \rho.$$

It follows that the temperature dependence of the moments \mathbf{M}_n is given by

$$(2.10) \quad \frac{d}{d\beta} \mathbf{M}_n = -\langle \mathcal{P}(V(\mathbf{x}) \mathcal{M}^n) \rangle + \langle \mathcal{P}(\mathcal{M}^n) \rangle \langle V(\mathbf{x}) \rangle.$$

These moment equations provide an exact description of the temperature evolution of the density distribution.

2.3.2. Gaussian Density Annealing (GDA).

In the GDA method we approximate the density as a Gaussian distribution which is fully defined by the center of the distribution \mathbf{r}_0 and variance M_2 or

$$\hat{\rho}(\mathbf{r}, \beta) = \left(\frac{2\pi M_2}{d} \right)^{-d/2} \exp \left[-\frac{d}{2M_2} (\mathbf{r} - \mathbf{r}_0)^2 \right]$$

as for Smoluchowski dynamics. The temperature evolution for the center and second-order moment is given by

$$(2.11) \quad \begin{aligned} \frac{d\mathbf{r}_0}{d\beta} &= -\frac{1}{d}M_2\nabla_{\mathbf{r}_0}\langle V \rangle \\ \frac{dM_2}{d\beta} &= -\frac{1}{d^2}M_2^2\nabla_{\mathbf{r}_0}^2\langle V \rangle. \end{aligned}$$

Again, the effective potential $\langle V \rangle$ appears. The center follows the gradient of the effective potential while the variance adjusts to the Laplacian of the effective potential. As with all the GPP based algorithms, the annealing dynamics is informed by higher order derivatives of the potential, rather than the force alone as is the case with classical molecular dynamics.

3. Applications

The methods described above have been applied to a series of model and molecular potential energy minimization problems. In this section, we provide a summary of the results.

3.1. One-dimensional rough potentials.

One might assume that exact classical mechanical annealing should be superior to an approximate, coarse grained annealing method. However, application of exact MD and the approximate GPP based annealing methods has shown that the GPP dynamics, while providing an approximate thermodynamics and dynamics for the system, can be superior in locating low energy regions of the potential energy hypersurface.

To understand how this is accomplished, we have studied the one-dimensional rough potential [17]

$$(3.1) \quad V(x) = \frac{1}{2}\kappa x^2 + \epsilon\cos[qx + \pi].$$

At high temperature, the density distribution is delocalized over a broad region of the potential surface. As the temperature is lowered, the distribution can collapse and become localized in a potential energy minimum. For this simple potential, it is possible to isolate two critical temperatures for the localization-delocalization transition (above which the packet is delocalized and below which the packet collapses). (1) For an *annealing* run, where the temperature is monotonically decreased, the critical temperature for the packet collapse transition is T_a . It is possible to derive an upper bound for this critical annealing temperature in terms of the potential parameters [18]

$$(3.2) \quad T_a = -\frac{4\kappa}{q^2k_B}\ln\frac{\kappa}{\epsilon q^2}.$$

(2) For a *heating* run, the initially collapsed distribution is delocalized above a critical temperature T_h . In general, for the GPP representation of the density distribution $T_a \ll T_h$. This hysteresis effectively reduces the temperature of the collapse transition for an annealing run. A similar observation has been made by Orešič and Shalloway in a study of Lennard-Jones cluster optimization [19]. By lowering the collapse transition temperature, the distribution remains delocalized over a broader region of temperature, possibly leading to further enhanced sampling.

This analysis implies that thermodynamic critical temperatures are dependent on the form of the approximation to the density distribution and the details of the annealing run. Coarse-grained dynamics might therefore be designed to move a critical temperature in such a way as to improve the probability that the algorithm will locate the global energy minimum. This is discussed below in the context of a homopolymer collapse transition.

3.2. Lennard-Jones clusters.

The benchmark for molecular optimization algorithms is the potential energy optimization of Lennard-Jones clusters. We have demonstrated that the GPP based Fokker-Planck and GDA algorithms are very effective in isolating the global energy minimum for clusters as large as $N = 55$ [11, 13]. The GPP based dynamical annealing algorithms are significantly more effective than standard molecular dynamics methods for the same cooling schedule. The GDA algorithm, which does not depend on a cooling schedule, was found to be approximately as effective as the GPP based Fokker-Planck method for the exponential cooling schedule $T(t) = 2 \exp(-t/10)$. Of course, while the GDA algorithm may be improved on only by modifying the representation of the density distribution, the GPP method may be improved by optimizing the cooling schedule. We have done some work in this direction through the use of a cooling schedule where the rate of change in temperature with time is proportional to the heat capacity $C_v(T)$ at constant volume

$$(3.3) \quad dT/dt = -\text{constant} \frac{1}{C_v(T)}.$$

Similar ideas have been explored previously [20]. Scaling the cooling rate in this way allows the system to spend approximately the same period of time dt in each energy increment dE as the temperature is being reduced. This simple technique leads to a significant improvement in the results for the same investment of computational time [13].

Ideally, one would like to understand the relative success or failure of optimization for a particular system in terms of the details of the potential energy hypersurface. We found that the probability of finding the global energy minimum for clusters of $N = 8$ to 19 was strongly correlated with the size of the energy gap between the global minimum and lowest energy local minimum. “Magic number” clusters, which have particularly low lying global energy minima and large gaps, were easier optimization problems than clusters where the energy gap was small [13]. This implies that for molecular optimization problems, the difficulty does not scale in a simple way with N but depends on the characteristics of the potential energy hypersurface.

3.3. Water clusters.

Tsoo and Brooks have applied a variant of the GDA algorithm to study the optimization of water clusters [21]. Water clusters provide notoriously difficult molecular optimization problems (empirically more difficult than Lennard-Jones clusters and small peptides). In their study, they recognized that a straightforward integration of the GDA equations of motion in reciprocal temperature β can result in a too rapid decrease in the variances (M_2) of the packets with relatively little movement of the centers. They found that it was useful to provide bounds on the variances while updating the position of the distribution center. The results of

Tsoo and Brooks show that their variation on the GDA algorithm performs slightly better than the Diffusion Equation Method (DEM) of Scheraga and coworkers [22] but is not successful in isolating the global energy minimum in certain cases.

This insight led to our development of the Adiabatic GDA method whereby one replaces the equation of motion for the center

$$(3.4) \quad d\mathbf{r}_0/d\beta = -(1/d)M_2\nabla_{\mathbf{r}_0}\langle V \rangle = 0$$

with a local energy minimization routine. The algorithm follows the protocol (1) solve Eq. (3.4) using a minimization routine, (2) integrate $dM_2/d\beta$ for a given step in β , (3) return to (1). This method allows the centers to respond instantaneously to changes in the variances. The resulting algorithm is a significant improvement over the standard GDA method [16].

3.4. Collapse transitions in Lennard-Jones homopolymers.

The GPP dynamical algorithm has been used to study the collapse of Lennard-Jones homopolymers [23]. The homopolymer was modeled as a freely jointed chain of Lennard-Jones atoms connected by harmonic bonds. An interesting point regarding the approximate thermodynamics generated by the GPP representation of the system emerged. In examining, for example, the radius of gyration as a function of temperature for the GPP and exact MD simulations, the homopolymer collapse transition temperature T_θ was found to occur at a significantly lower temperature in the GPP simulation than in the MD simulation. We believe that this is a manifestation of the property observed previously for the one-dimensional rough potential [18] where the critical annealing temperature T_a was suppressed. For the model one-dimensional potential and the Lennard-Jones homopolymer, as the temperature is lowered the GPP dynamics leads to a more extended or delocalized distribution at a given temperature. Therefore, above T_θ the GPP representation of the system has an effectively altered thermodynamics.

The time scale for nucleation or homopolymer collapse can depend on the difference between a thermodynamic transition temperature and a quench temperature. For example, de Gennes [24] has proposed that the characteristic time for the homopolymer collapse transition following a quench to a temperature $T_Q < T_\theta$ increases monotonically with the temperature ratio

$$(3.5) \quad \frac{T_\theta - T_Q}{T_\theta}.$$

Altering the thermodynamic transition temperature can lead to effectively faster (or slower) transitions. Therefore, if an approximate representation of the system leads to an altered value of T_θ , the dynamics may also be affected in an important way. Potentially, this could be used as a criterion to develop more effective coarse grained dynamics for global optimization.

3.5. Folding model proteins.

We have completed an extensive study of global optimization for 22-mer and 46-mer model proteins [25] based on the three letter code potential of Honeycutt and Thirumalai [26]. The three letters correspond to hydrophobic-type beads (B), hydrophilic-type beads (L), and neutral-type beads (N). The potential consists of a linear polymer chain of harmonic bonds

$$(3.6) \quad V_{\text{bond}}(r) = \frac{\kappa_b}{2} (r - \sigma)^2$$

and angles

$$(3.7) \quad V_{\text{angle}}(\theta) = \frac{\kappa_{\theta}}{2} (\theta - \theta_0)^2$$

where the equilibrium angle is $\theta_0 = 105^\circ$. The dihedral angle potential is

$$(3.8) \quad E_{\text{torsion}}(\phi) = A(1 + \cos[\phi]) + B(1 + \cos[3\phi])$$

where the dihedral constants are defined to be $A = 0$ and $B = 0.2$ if two or more of the defining atoms i, j, k and l are type N, and $A = B = 1.2$ otherwise.

The nonbonded potential is made of soft-sphere type repulsions and Lennard-Jones terms defined as

$$(3.9) \quad \begin{aligned} V_{BB}(r) &= 4\epsilon \left[(\sigma/r)^{12} - (\sigma/r)^6 \right] \\ V_{LB,LL}(r) &= (8\epsilon/3) \left[(\sigma/r)^{12} + (\sigma/r)^6 \right] \\ V_{NN,NL,NB}(r) &= 4\epsilon (\sigma/r)^{12}. \end{aligned}$$

The potential models the basic properties of attractive interactions between hydrophobic residues (B) and encourages the formation of a hydrophobic core. The repulsions between hydrophilic sites (L) and all other residues model their tendency to be found at the protein surface.

We arrived at a number of conclusions relating to the relative effectiveness of the GPP based algorithms, simulated annealing using MD, and the DEM [22] in a study of the 22-mer $(\text{LB})_5\text{N}_2(\text{LB})_5$, a β -turn, and 46-mer $\text{B}_9\text{N}_3(\text{LB})_4\text{N}_3\text{B}_9\text{N}_3(\text{LB})_5\text{L}$, a β -turn sandwich [25].

(1) Comparative analysis of the success in locating the global energy minimum for the 22-mer model showed that the GPP based methods, along with the DEM, are significantly more effective in isolating the global energy minimum than simulated annealing with MD. In particular, the GDA, DEM and GPP based Fokker-Planck dynamics algorithms were most effective.

(2) The probability of finding the global energy minimum for three sequences of the 22-mer, (a) the hetero sequence $(\text{LB})_5\text{N}_2(\text{LB})_5$, (b) the hybrid sequence $\text{B}_9\text{N}_3(\text{LB})_5$ and the (c) the homo sequence B_{22} , (which uses the torsion potential of the hybrid sequence), was studied. The results showed that the hetero sequence (a) was easiest to fold (for all algorithms studied). We found that the energy gap between the global energy minimum and the next highest energy, incorrectly folded state was well correlated with algorithmic success. This observation was made previously in the context of Lennard-Jones clusters [13].

(3) In comparing the results for folding the 22-mer and 46-mer, we found that phase space based dynamical annealing methods (such as the GPP based Fokker-Planck dynamics) had a distinct advantage over configurational space based methods (such as the GDA) in finding the global energy minimum of the 46-mer. It appears that the configurational space based methods, such as the GPP based Smoluchowski algorithms and the GDA algorithm, which are based on the evolution of $\hat{\rho}(\mathbf{r})$ in the absence of ballistic motion, are very sensitive to the initial conditions. The phase space based dynamical methods, which maintain inertial terms in their dynamics, were better able to overcome an unfortunate initial placement.

4. Conclusions

Results for a variety of systems indicate that when the simulated annealing protocol is followed using an approximate GPP based classical dynamics, it is possible to combine the best aspects of the simulated annealing and potential smoothing paradigms for molecular optimization. Regarding the relative success of these methods, it has been emphasized that the effect of potential smoothing is to reduce the number of local minima and thereby ease the search of conformational space [22, 27]. We have shown that a related change in the thermodynamic transition temperatures can be used, in a compatible way, to account for the improved ability of the GPP based methods to locate the global energy minimum for a variety of molecular optimization problems.

Appendix A. Derivation of $dM_{n,k}/dt$ for Fokker-Planck dynamics

The right hand side of Eq. (2.2) may be divided into two integrals

$$I_1 = \int d\mathbf{x}d\mathbf{p} \mathcal{P} \left(\frac{\partial \mathcal{M}^{n,k}}{\partial t} \right) \rho \quad I_2 = \int d\mathbf{x}d\mathbf{p} \mathcal{P} \left(\mathcal{M}^{n,k} \frac{\partial \rho}{\partial t} \right).$$

Given that

$$(A.1) \quad \frac{\partial}{\partial t} \Delta x_i \hat{e}_i = -\dot{x}_{0i} \hat{e}_i \quad \frac{\partial}{\partial t} \Delta p_i \hat{e}_i = -\dot{p}_{0i} \hat{e}_i$$

and that for the Fokker-Planck dynamics

$$\dot{x}_{0i} \hat{e}_i = \frac{p_{0i}}{m} \hat{e}_i \quad \dot{p}_{0i} \hat{e}_i = (F_{0i} - \gamma p_{0i}) \hat{e}_i$$

we find

$$(A.2) \quad \frac{\partial}{\partial t} \mathcal{P} (\mathcal{M}^{n,k}) = \mathcal{P} \left(-\frac{n}{m} (p_{0i} \hat{e}_i) \mathcal{M}^{n-1,k} - k (F_{0i} - \gamma p_{0i}) \hat{e}_i \mathcal{M}^{n,k-1} \right).$$

We advise that one performs the time derivatives using the Einstein summation convention. The permutation operator insures the simple final result. Therefore, I_1 can be written as $I_1 = I_{11} + I_{12} + I_{13}$

$$(A.3) \quad I_1 = -\frac{n}{m} \int d\mathbf{x}d\mathbf{p} \mathcal{P} (\mathcal{M}^{n-1,k} (p_{0i} \hat{e}_i)) \rho - k \int d\mathbf{x}d\mathbf{p} \mathcal{P} (\mathcal{M}^{n,k-1} (F_{0i} \hat{e}_i)) \rho \\ + \gamma k \int d\mathbf{x}d\mathbf{p} \mathcal{P} (\mathcal{M}^{n,k-1} (p_{0i} \hat{e}_i)) \rho.$$

Substituting $\partial \rho / \partial t$ by its expression given in Eq. (2.1), I_2 can be written as $I_2 = I_{21} + I_{22} + I_{23} + I_{24} + I_{25}$

$$(A.4) \quad I_2 = -\frac{1}{m} \int d\mathbf{x}d\mathbf{p} \mathcal{P} \left(\mathcal{M}^{n,k} p_i \frac{\partial \rho}{\partial x_i} \right) - \int d\mathbf{x}d\mathbf{p} \mathcal{P} \left(\mathcal{M}^{n,k} F_i(\mathbf{x}) \frac{\partial \rho}{\partial p_i} \right) \\ + \gamma \int d\mathbf{x}d\mathbf{p} \mathcal{P} \left(\mathcal{M}^{n,k} p_i \frac{\partial \rho}{\partial p_i} \right) + \gamma \int d\mathbf{x}d\mathbf{p} \mathcal{P} (\mathcal{M}^{n,k}) \rho \\ + \gamma m k_B T \int d\mathbf{x}d\mathbf{p} \mathcal{P} \left(\mathcal{M}^{n,k} \frac{\partial^2 \rho}{\partial p_i^2} \right).$$

Holding the variable \mathbf{p} constant, I_{21} may be integrated by parts with respect to \mathbf{x} . For all following integrations by parts we divide the integrand into two parts;

the integrated term is always a differential of ρ . We assume that the surface terms vanish since the normalized density distribution goes to 0 at ∞ . This leads to

$$I_{21} = \frac{n}{m} \int d\mathbf{x} d\mathbf{p} \mathcal{P} (\mathcal{M}^{n-1, k} (p_i \hat{e}_i)) \rho.$$

In the same way, holding \mathbf{x} constant, I_{22} and I_{23} may be integrated by parts with respect to \mathbf{p} leading to

$$I_{22} = k \int d\mathbf{x} d\mathbf{p} \mathcal{P} (\mathcal{M}^{n, k-1} F_i(\mathbf{x}) \hat{e}_i) \rho$$

and

$$I_{23} = -\gamma k \int d\mathbf{x} d\mathbf{p} \mathcal{P} (\mathcal{M}^{n, k-1} p_i \hat{e}_i) \rho - \gamma \int d\mathbf{x} d\mathbf{p} \mathcal{P} (\mathcal{M}^{n, k}) \rho.$$

I_{25} may be integrated by parts twice with respect to \mathbf{p} leading to

$$(A.5) \quad I_{25} = \gamma m k_B T k (k-1) \int d\mathbf{x} d\mathbf{p} \mathcal{P} (\mathcal{M}^{n, k-2} (\hat{e}_i \hat{e}_i)) \rho.$$

Grouping all terms from Eq. (A.3) and Eq. (A.4) we find a final expression for the total time derivative of the moment $d\mathbf{M}_{n, k}/dt = (I_{11} + I_{21}) + (I_{12} + I_{22}) + (I_{13} + I_{23} + I_{24}) + I_{25}$ which is

$$(A.6) \quad \begin{aligned} \frac{d}{dt} \mathbf{M}_{n, k} &= \frac{n}{m} \int d\mathbf{x} d\mathbf{p} \mathcal{P} (\mathcal{M}^{n-1, k+1}) \rho + k \int d\mathbf{x} d\mathbf{p} \mathcal{P} (\mathcal{M}^{n, k-1} (F_i(\mathbf{x}) - F_{0i}) \hat{e}_i) \rho \\ &\quad - \gamma k \int d\mathbf{x} d\mathbf{p} \mathcal{P} (\mathcal{M}^{n, k}) \rho + \gamma k (k-1) m k_B T \int d\mathbf{x} d\mathbf{p} \mathcal{P} (\mathcal{M}^{n, k-2} \hat{e}_i \hat{e}_i) \rho. \end{aligned}$$

Appendix B. Derivation of $d\mathbf{M}_n/dt$ for Smoluchowski dynamics

The right hand side of Eq. (2.6) may be divided into two integrals

$$I_1 = \int d\mathbf{x} \mathcal{P} \left(\frac{\partial \mathcal{M}^n}{\partial t} \right) \rho \quad I_2 = \int d\mathbf{x} \mathcal{P} \left(\mathcal{M}^n \frac{\partial \rho}{\partial t} \right).$$

Using the result of Eq. (A.1) and the fact that for the Smoluchowski dynamics

$$\dot{x}_{0i} \hat{e}_i = \frac{1}{m\gamma} F_{0i} \hat{e}_i$$

I_1 can be written as

$$(B.1) \quad I_1 = -\frac{n}{m\gamma} \int d\mathbf{x} \mathcal{P} (\mathcal{M}^{n-1} F_{0i} \hat{e}_i) \rho.$$

Substituting the expression of $\partial\rho/\partial t$ defined in Eq. (2.5) we find

$$I_2 = \frac{1}{m\gamma} \int d\mathbf{x} \mathcal{P} \left(\mathcal{M}^n \left[\frac{\partial}{\partial x_i} \left[-F_i(\mathbf{x}) + k_B T \frac{\partial}{\partial x_i} \right] \right] \right) \rho.$$

Recognizing that $-F_i(\mathbf{x}) = \partial V(x_i)/\partial x_i$

$$\begin{aligned} I_2 &= \frac{1}{m\gamma} \int d\mathbf{x} \mathcal{P} \left(\mathcal{M}^n \frac{\partial^2 V(x_i)}{\partial x_i^2} \right) \rho + \frac{1}{m\gamma} \int d\mathbf{x} \mathcal{P} \left(\mathcal{M}^n \frac{\partial V(x_i)}{\partial x_i} \frac{\partial \rho}{\partial x_i} \right) \\ &\quad + \frac{k_B T}{m\gamma} \int d\mathbf{x} \mathcal{P} \left(\mathcal{M}^n \frac{\partial^2 \rho}{\partial x_i^2} \right) \end{aligned}$$

and integrating the second term in I_2 by parts we find

$$(B.2) \quad I_2 = -\frac{n}{m\gamma} \int d\mathbf{x} \mathcal{P} \left(\mathcal{M}^{n-1} \frac{\partial V(x_i)}{\partial x_i} \hat{e}_i \right) \rho + \frac{k_B T}{m\gamma} \int d\mathbf{x} \mathcal{P} \left(\mathcal{M}^n \frac{\partial^2 \rho}{\partial x_i^2} \right).$$

Integrating the second term of Eq. (B.2) by parts twice we find

$$(B.3) \quad I_2 = \frac{n}{m\gamma} \int d\mathbf{x} \mathcal{P} \left(\mathcal{M}^{n-1} \left(-\frac{\partial V(x_i)}{\partial x_i} \right) \hat{e}_i \right) \rho \\ + \frac{n(n-1)k_B T}{m\gamma} \int d\mathbf{x} \mathcal{P} \left(\mathcal{M}^{n-2} \hat{e}_i \hat{e}_i \right) \rho.$$

Combining Eq. (B.1) and Eq. (B.3) to form $d\mathbf{M}_n/dt = I_1 + I_2$ we arrive at

$$\frac{d}{dt} \mathbf{M}_n = \int d\mathbf{x} \frac{n}{m\gamma} \mathcal{P} \left(\mathcal{M}^{n-1} (F_i(\mathbf{x}) \hat{e}_i - F_{0i} \hat{e}_i) \right) \rho \\ + \frac{k_B T n(n-1)}{m\gamma} \mathcal{P} \left(\mathcal{M}^{n-2} \hat{e}_i \hat{e}_i \right) \rho$$

leading to the final expression for $d\mathbf{M}_n/dt$ given in Eq. (2.7).

References

- [1] J. E. Straub and D. Thirumalai. *Exploring the energy landscape in proteins*. Proc. Natl. Acad. Sci. USA **90**, 809–813 (1993). J. E. Straub and D. Thirumalai. *Theoretical Probes of conformational fluctuations in S-Peptide and RNase A/3'-UMP enzyme product complex*. Proteins **15**, 360-373 (1993). J. E. Straub, A. Rashkin and D. Thirumalai. *Dynamics in rugged energy landscapes with applications to the S-Peptide and Ribonuclease A*. J. Am. Chem. Soc. **116**, 2049-2063 (1994).
- [2] J. Grad, Y. J. Yan, and S. Mukamel. *Time-dependent self-consistent field approximation for semiclassical dynamics using gaussian wavepackets in phase space*. Chem. Phys. Lett. **134**, 291-295(1987); J. Grad, Y. J. Yan, A. Haque, and S. Mukamel. *Reduced equations of motion for semiclassical dynamics in phase space*. J. Chem. Phys. **86**, 3441-3454 (1987).
- [3] R. D. Coalson and M. Karplus. *Generalized quantum Liouville equation: Its solution by wave packet dynamics*. J. Chem. Phys. **79**, 6150–6161 (1983).
- [4] E. J. Heller. *Time-Dependent approach to semiclassical dynamics*. J. Chem. Phys. **62**, 1544-1555 (1975); E. J. Heller. *The semiclassical way to molecular spectroscopy*. Acc. Chem. Res. **14**, 368-375 (1981).
- [5] N. Corbin and K. Singer. *Semiclassical molecular dynamics of wave packets*. Mol. Phys. **46**, 671-677 (1982); K. Singer and W. Smith. *Semiclassical many-particle dynamics with gaussian wave packets*. Mol. Phys. **57**, 761-775 (1986).
- [6] R. B. Gerber and M. A. Ratner. *Self-consistent-field methods for vibrational excitations in polyatomic systems*. Adv. Chem. Phys. **70**, 97–132 (1988).
- [7] R. Elber and M. Karplus. *Enhanced sampling in molecular dynamics: Use of the time-dependent Hartree approximation for a simulation of carbon monoxide diffusion through myoglobin*. J. Am. Chem. Soc **112**, 9161–9175 (1990).
- [8] R. Czerminski and R. Elber. *Computational studies of ligand diffusion in globins: I. Leghemoglobin*. Proteins **10**, 70–80 (1991).
- [9] P. Amara, D. Hsu, and J. E. Straub. *Global energy minimum searches using an approximate solution of the imaginary time Schrödinger equation*. J. Phys. Chem. **97**, 6715–6721 (1993).
- [10] A. B. Finnila, M.A. Gomez, C. Sebenik, C. Stenson, and J.D. Doll. *Quantum Annealing: A new method for minimizing multidimensional functions*. Chem. Phys. Lett **219**, 343–348 (1994).
- [11] J. Ma, D. Hsu, and J. E. Straub. *Approximate solution of the classical Liouville equation using Gaussian phase packet dynamics: Application to enhanced equilibrium averaging and global optimization*. J. Chem. Phys. **99**, 4024–4035 (1993).

- [12] R. E. Walkup. *A local-Gaussian approximation for the propagation of a classical distribution function*. *J. Chem. Phys.* **95**, 6440–6448 (1991).
- [13] J. Ma and John E. Straub. *Simulated annealing using the classical density distribution*. *J. Chem. Phys.* **101**, 533–541 (1994).
- [14] H. Jeffreys. *Cartesian Tensors*. Cambridge University Press, London, (1974).
- [15] These equations are identical to those derived from the Dirac-Frenkel variational principle. J. Frenkel, *Wave Mechanics: Advanced General Theory* (Clarendon Press, Oxford, 1934). In the Dover edition (1950), see p. 253. Frenkel cites an appendix in the Russian edition of Dirac’s *Principles of Quantum Mechanics* (1930).
- [16] J. E. Straub. In R. Elber, editor, *New Developments in Theoretical Studies of Proteins*. World Scientific, Singapore, (in press).
- [17] R. Zwanzig. *Diffusion in a rough potential*. *Proc. Natl Acad. Sci.* **85**, 2029–2030 (1988).
- [18] J. E. Straub J. Ma and P. Amara. *Simulated annealing using coarse grained classical dynamics: Fokker-Planck and Smoluchowski dynamics in the Gaussian density approximation*. submitted to *J. Chem. Phys.*
- [19] M. Orešič and D. Shalloway. *Hierarchical characterization of energy landscapes using Gaussian packet states*. *J. Chem. Phys.* **101**, 9844–9857 (1994).
- [20] P. Salamon, J. D. Nulton, J. R. Harland, J. Pedersen, G. Ruppeiner, and L. Liao. *Simulated annealing with constant thermodynamic speed*. *Comp. Phys. Comm.* **49**, 423–428 (1988).
- [21] C. Tsao and C. L. Brooks III. *Cluster structure determination using Gaussian density distribution global minimization methods*. *J. Chem. Phys.* **101**, 6405–6411 (1994).
- [22] L. Piela, J. Kostrowicki, and H. A. Scheraga. *The multiple-minima problem in the conformational analysis of molecules. Deformation of the potential energy hypersurface by the Diffusion Equation Method*. *J. Phys. Chem.* **93**, 3339–3346 (1989). J. Kostrowicki and L. Piela and B. J. Cherayil and H. A. Scheraga. *Performance of the Diffusion Equation Method in searches for optimum structures of clusters of Lennard-Jones atoms*. *ibid.*, **95** 4113–4119 (1991); R. J. Wawak and M. M. Wimmer and H. A. Scheraga. *Application of the Diffusion Equation Method of Global Optimization to Water Clusters*. *ibid.*, **96**, 5138–5145 (1992); J. Kostrowicki and H. A. Scheraga. *Application of the Diffusion Equation Method for Global Optimization to Oligopeptides* *ibid.* **96**, 7442–7449 (1992).
- [23] J. Ma, J. E. Straub, and E. I. Shakhnovich. *Simulation study of the collapse of Lennard-Jones homopolymers*. submitted to *J. Chem. Phys.*
- [24] P. G. de Gennes. *Kinetics of collapse for a flexible coil*. *J. Phys. Lett.* **46**, L-639-L-642 (1985).
- [25] P. Amara and J. E. Straub. *Folding model proteins using kinetic and thermodynamic annealing of the classical density distribution*. submitted to *J. Phys. Chem.*
- [26] J. D. Honeycutt and D. Thirumalai. *The Nature of Folded States of Globular Proteins*. *Biopolymers* **32**, 695–709 (1992).
- [27] D. Shalloway, in: *Recent advances in global optimization*, eds. A. Floudas and P. M. Pardalos (Princeton University Press, Princeton, 1992) p. 433.

DEPARTMENT OF CHEMISTRY, BOSTON UNIVERSITY, BOSTON, MASSACHUSETTS 02215

E-mail address: amara@chem.bu.edu

E-mail address: jma@chem.bu.edu

E-mail address: straub@chem.bu.edu