

Normal Mode Partitioning of Langevin Dynamics for Biomolecules

Christopher R. Sweet^a, Paula Petrone^b, Vijay S. Pande^{c,*} and Jesús A. Izaguirre^{a†}

a) Department of Computer Science and Engineering

University of Notre Dame, Notre Dame, Indiana

b) Departments of Biophysics and c) Chemistry, and

Stanford University, Stanford, California

(Dated: Feb. 4, 2008; Received: Aug. 21, 2007; Revised: Dec. 21, 2007; Accepted: Jan. 22, 2008;

Published: Mar. 14, 2008)

Abstract

We propose a novel Normal-Mode multiple time stepping Langevin dynamics integrator called NML. The aim is to approximate the kinetics or thermodynamics of a biomolecule by a reduced model based on a normal mode decomposition of the dynamical space. Our basis set uses the eigenvectors of a mass re-weighted Hessian matrix calculated with a biomolecular force field. This particular choice has the advantage of an ordering according to the eigenvalues, which have a physical meaning being the square-root of the mode frequency. Low frequency eigenvalues correspond to more collective motions, whereas the highest frequency eigenvalues are the limiting factor for the stability of the integrator. In NML the higher frequency modes are overdamped and relaxed near to their energy minimum while respecting the subspace of low frequency dynamical modes. Our numerical results confirm that both sampling and rates are conserved for an implicitly solvated alanine dipeptide model, with only 30% of the modes propagated, when compared to the full model. For implicitly solvated systems NML gives a twofold improvement in efficiency over plain Langevin dynamics for sampling a small 22 atom (alanine dipeptide) model and in excess of an order of magnitude for sampling an 882 atom (bovine pancreatic trypsin inhibitor, or BPTI) model, with good scaling with system size subject to the number of modes propagated. NML has been implemented in the open source software PROTOMOL.

Keywords: Normal mode, Langevin dynamics, long time integrator, time coarse-graining, Brownian dynamics, slow dynamics subspace

Contents

I. Introduction	2
II. Theory of Normal Mode Partitioning	6
A. NMA and Hessian diagonalization.	7
B. Propagation using the Langevin equation.	8
C. Multiple-time-stepping discretization (NMI).	9
D. NML discretization.	11
E. NML minimization.	12
III. Computational tests	14
A. Choice of the number of modes to propagate.	16
B. Sampling using NML or NMI.	17
C. Rate calculation using NML or NMI.	19
D. Efficiency of NML.	20
E. Thermostat stability.	22
IV. Discussion	23
V. Acknowledgments	24
A. Projection matrices	24
References	25
Captions	29
Figures	31
Tables	34

I. INTRODUCTION

Molecular dynamics (MD) is a useful technique for studying the kinetics and thermodynamics of proteins and other biomolecules. A primary limitation to the application of MD to study

biologically relevant processes such as protein folding, docking, and functionally important conformational changes is the small time step size of conventional MD. Whereas the latter is typically measured in fs, dynamical processes of interest happen in μs and longer time scales. We seek to construct a reduced model of the kinetics or thermodynamics of biomolecules: that is, separate the dynamics into fast, uninteresting modes, and slow, functionally relevant modes. Once the space has been partitioned in this way, we perform MD in the reduced space to compute observables.

Among the most popular approaches to find reduced dynamical spaces for biomolecules are normal mode analysis (NMA) and principal component analysis (PCA). Normal-mode analysis has proven successful in representing collective motions of atoms, such as domain and hinge bending in proteins.¹⁻¹³ Because of this, NMA has been used to generate a generalized coordinate set in which to represent molecular motion. Principal component analysis, e.g. Ref. 14, or the essential dynamics analysis¹⁵ use several conformations or relatively long simulations to capture slow motions. In comparison, NMA requires only an initial structure and no simulation to determine the essential space.

Methods to compute observables given a reduced space model have previously been developed. Some MD propagators have attempted to take advantage of the frequency spectra provided by NMA. For instance, a scheme called LIN¹⁶ propagates the high frequency normal modes using the analytical solution of a Langevin dynamics equation. Updating these modes requires frequent re-diagonalization of the mass re-weighted Hessian (MRH), every few fs or ps. Also, the low frequency modes are updated using an implicit solver, which is generally very costly. These choices make LIN computationally expensive and limit the potential speedup of the method.

The Amplified Collective Motions (ACM) MD method¹⁷ attempts to excite motions along a few low frequency modes (typically 3-5) obtained from the Anisotropic Network Model (ANM).¹⁸ After estimating the slowest ANM modes, the velocities are split into the essential subspace and the remainder subspace. The essential subspace is weakly coupled to a higher temperature bath. The modes are updated very frequently, every few fs. Simulations of ACM on proteins show enhanced sampling. The authors of ANM do not show what the stationary distribution and partition function of the method are. Another method, LLMOD¹⁹ obtains the low frequency modes of the Hessian and does an analytical propagation along the low frequency modes. This integration step is non-physical because it ignores the coupling with high frequency modes and quickly becomes invalid. Thus, it is not possible to determine the ensemble that is sampled. An interesting feature of LLMOD is that uses an iterative Arnoldi method to compute the low frequency eigenvectors

and is thus applicable to larger molecules. Similarly, Elnémo, a web server to compute an elastic network model using a rotation-translation-block approximation is able to compute the 100 lowest-frequency modes of a protein in PDB format.²⁰

The Subspace Molecular Dynamics (SMD)²¹ is a method in which only the low frequency subspace is propagated according to projected Newtonian equations of motion, while high frequency modes remain unaltered. This elegant scheme, which works well in crystals, loses efficiency when extended to biomolecules.²² In this case, the system is not periodic and large interactions between subspaces occur, and failure to frequently re-diagonalize the MRH results in constraining the dynamics to the neighborhood of the initial equilibrium structure.

The collective Langevin dynamics²³ uses PCA to obtain a reduced space of low frequency modes. The dynamics of the remaining space is accounted for through a generalized Langevin equation with a memory kernel. This method, though elegant, requires a relatively long simulation to obtain the memory kernel for the Langevin equation.

Other work has attempted to build longer time step integrators for MD.^{24–28} One effort in this direction has been the use of multiple time-step (MTS)²⁹ integrators. Linear and nonlinear resonances limit the length of the time step to about one third of the period of the fastest vibrational mode. For MD of proteins using r-RESPA^{30,31}, a prototypical MTS integrator, the stability limit is at a time-step of 3.3 fs.³² More stable MTS integrators only offer a modest improvement of roughly a factor of 1.5.^{33,34}

The largest impediment to a time-step increase comes from the fastest internal vibrations. Algorithms to constrain bonds, such as SHAKE^{35–37} are normally implemented and popular solvent models constrain the bond lengths and angles for water. It is also a standard procedure to constrain covalent bonds to hydrogen in all-atom representations. This approach allows the time-step to be doubled over unconstrained MD.³⁸

It is unclear, however, how to use constraints on atomic degrees of freedom beyond these obvious holonomic constraints without significantly affecting the dynamics of the molecule. Several methods have been proposed that treat whole regions of a biomolecule as rigid entities, e.g. Ref. 39. Because the definition of these regions is arbitrary, the error in the kinetics or thermodynamics incurred by this procedure is difficult to estimate. In general constraints beyond just covalent bonds damage the thermodynamical and kinetic properties of the system.⁴⁰

An MTS algorithm for the diffusive Langevin equation has been devised.⁴¹ Here, bonds and angles are constrained to achieve a long outer time step, and then the unconstrained system is

propagated to bring the high frequencies to thermal equilibrium. The method relies on separation of time scales introduced by the constraints and splitting of the forces. Unfortunately, in general, the separation of frequencies between these time scales is not well defined and the coupling between high and low frequencies is not always weak.⁴²⁻⁴⁴ Also, relaxing through unconstrained steps limits the efficiency of the method, since many short time steps still need to be carried out.

We introduce a multiple time stepping propagator for Langevin dynamics of implicitly solvated biomolecules called Normal Mode Langevin (NML). We obtain the reduced model (low frequency modes) through a normal mode analysis of a mass-reweighted Hessian of a biomolecular force field (currently, CHARMM 22^{45,46}). This partitioning is used to construct a damping matrix for the Langevin equation that overdamps and propagates the high frequency modes using Brownian dynamics, whereas it propagates the relevant low frequency modes using full Langevin dynamics. In our discretization we accelerate the convergence of Brownian dynamics on high frequencies by an energy minimization.

The rationale for this treatment of the high frequencies can be understood if we consider the work of Petrone and Pande.⁴⁷ They have shown that, contrary to previous observations,⁷ only the low frequency normal modes are not enough to describe the transition between two conformations, and instead, a broad spectrum of frequencies is required. Since high frequency eigenvectors have short persistence throughout a simulation, we cannot propagate them without expensive re-diagonalization.¹⁶ As regards the treatment of high frequency modes, constraining their amplitude to zero damages the long time dynamics,²¹ whereas relaxing them to thermal equilibrium through small time step integration is expensive.⁴¹ NML minimizes the energy with respect to the high frequency modes, which in the limit of large damping coefficient is equivalent to propagating these modes using Brownian dynamics. This is a more efficient way of thermalizing these high frequency modes, as shown below.

In our NML implementation the force calculation, integration, and minimization or dampening of fast modes are performed in Cartesian coordinates. Because of this, our method combines the advantages of a generalized coordinate obtained through NMA from only one structure to select the relevant motions, and the simplicity of the equations of motion given by the Cartesian coordinates.

To validate our method we use two test cases: the alanine dipeptide and bovine pancreatic trypsin inhibitor (BPTI), for which extensive data is available in the literature. We show the range of validity of our method, calculating kinetic (for alanine dipeptide) and sampling properties of

test systems (alanine dipeptide, BPTI) as a function of the number of dynamical modes.

In Section II we describe the underlying rationale and theory for the NML method and derive the Langevin Impulse version. In Section III we apply NML to explore sampling or kinetics of blocked alanine dipeptide using a distance-dependent dielectric implicit solvent model and BPTI using the screened Coulomb potential implicit solvent model (SCPISM). In the alanine dipeptide tests we demonstrate that using NML, when only the first 22-26 modes (of 66) are used, results in both Ramachandran plots and isomerization transition rates which are in good agreement with those obtained from the Langevin Impulse propagator. This leads to only a two-fold increase in efficiency due to the small size of the system. To show the scalability of the method we also test a BPTI model, demonstrating fair to good agreement on Ramachandran plots when propagating 100-300 modes (of 2646), with efficiency increases in excess of an order of magnitude. These results are reviewed and interpreted in Section IV.

II. THEORY OF NORMAL MODE PARTITIONING

We partition the system's degrees of freedom into a subspace C and its complement C^\perp according to frequency using Normal Mode Analysis (NMA). NMA finds a harmonic approximation to the system of interest around an initial equilibrium structure, using the mass re-weighted Hessian (MRH). The diagonalization of the MRH generates the normal mode vectors, ranked in increasing order of their frequency (square root of eigenvalue). These normal modes therefore become the generalized coordinates with which to capture the dynamics of the system. In subspace C normal modes will be propagated whereas in C^\perp modes will be dampened to near their energy minimum while respecting the subspace of dynamical modes. In the case where C includes all the modes, the original Langevin dynamics are recovered. Since we wish to use standard MD force fields such as CHARMM^{45,46}, which are more readily calculated in the $3N$ Cartesian space for system size N , the system is propagated in these coordinates rather than the normal mode coordinates for greater efficiency.

We describe the method for partitioning spaces and determining the largest time step in Section II A. We derive the Langevin equations using a damping matrix to partition the system in Section II B, and do a multiple time stepping (MTS) discretization called Normal Mode Impulse (NMI, Section II C), and a more efficient method called NML (Section II D) using a minimizer (Section II E).

A. NMA and Hessian diagonalization.

Many MD systems of interest can be described by a separable Hamiltonian with $3N$ Cartesian positions \mathbf{x} and momenta \mathbf{p} , potential energy U and diagonal mass matrix \mathbf{M}

$$\mathcal{H}(\mathbf{x}, \mathbf{p}) = \frac{1}{2} \mathbf{p}^T \mathbf{M}^{-1} \mathbf{p} + U(\mathbf{x}). \quad (1)$$

For these systems we can formally expand the potential energy about an equilibrium point \mathbf{x}_0 , which we assume is a local minimum. The Hessian \mathbf{H} is a factor in the first non-constant, non-zero term of this expansion and a harmonic approximation to the original system can be found by truncating the expansion at this point. To rewrite the harmonic approximation as a set of decoupled oscillators it is insufficient to diagonalize the Hessian as the resulting oscillators would be coupled through the projected mass matrix. Instead we mass re-weight the system using $\hat{\mathbf{x}} = \mathbf{M}^{\frac{1}{2}}(\mathbf{x} - \mathbf{x}_0)$ and then diagonalize the resulting mass re-weighted Hessian,

$$\mathbf{M}^{-\frac{1}{2}} \mathbf{H} \mathbf{M}^{-\frac{1}{2}} \mathbf{Q} = \mathbf{Q} \mathbf{\Lambda}, \quad (2)$$

where $\mathbf{\Lambda}$ is the diagonal matrix of ordered eigenvalues $\lambda_1, \dots, \lambda_{3N}$ and \mathbf{Q} the matrix of column eigenvectors $\mathbf{e}_1, \dots, \mathbf{e}_{3N}$. This means that the frequency of each oscillator is uniquely defined by its associated eigenvalue. We can now project between the mode space and Cartesian space using the eigenvectors \mathbf{Q} and mass matrix \mathbf{M}

$$\mathbf{c} = \mathbf{Q}^T \mathbf{M}^{\frac{1}{2}}(\mathbf{x} - \mathbf{x}_0), \quad \mathbf{x} = \mathbf{M}^{-\frac{1}{2}} \mathbf{Q} \mathbf{c} + \mathbf{x}_0, \quad (3)$$

for mode amplitude vector \mathbf{c} . We choose a cut-off frequency $f_c = \sqrt{\lambda_i}$ to partition the normal modes such that $\mathbf{Q} = [\mathbf{Q}, \bar{\mathbf{Q}}]$, $\mathbf{Q} = [\mathbf{e}_1, \dots, \mathbf{e}_i]$ and $\bar{\mathbf{Q}} = [\mathbf{e}_{i+1}, \dots, \mathbf{e}_{3N}]$ are rectangular matrices whose columns span the subspace C and complement subspace C^\perp respectively. In the following discussion we will assume the dimensions of \mathbf{Q} to be $3N \times m$. In the linear case the time-step is bounded by the asymptotic stability of the method^{48,49} at a frequency equal to the square root of the i^{th} eigenvalue, rather than the highest frequency in the system. This can be estimated by

$$\Delta t_i \approx \Delta t_{3N} \sqrt{\frac{\lambda_{3N}}{\lambda_i}}. \quad (4)$$

where λ_i is the i^{th} diagonal element of $\mathbf{\Lambda}$, λ_{3N} is the largest eigenvalue and Δt_{3N} is the time step required to propagate the non-partitioned system. Tests in Section III confirm that this is a good heuristic for the choice of time step for the full system.

B. Propagation using the Langevin equation.

The system will be time-propagated using the Langevin equation

$$dx = vdt, \quad Mdv = fdt - \Gamma Mvdt + \sqrt{2k_B T} \Gamma^{\frac{1}{2}} M^{\frac{1}{2}} d\mathbf{W}(t), \quad (5)$$

where \mathbf{f} internal forces between particles in the system, t is time, $\mathbf{W}(t)$ is a collection of Wiener processes, k_B is the Boltzmann constant, T is the desired system temperature, \mathbf{v} are the velocities and Γ is the diagonalizable damping matrix. The system diffusion tensor \mathbf{D} gives rise to Γ :

$$\Gamma = k_B T \mathbf{D}^{-1} \mathbf{M}^{-1}. \quad (6)$$

\mathbf{D} could be chosen to model the dynamics of an implicit solvent, in which case the damping is related to solvent viscosity, or simply to sample from a canonical ensemble or for the equilibration phase of a subsequent microcanonical simulation.

We make the following physical approximations about the system:

Physical approximation 1. We use Γ to partition the system into two subspaces, C spanned by the low frequency eigenvectors \mathbf{Q} and C^\perp spanned by the high frequency eigenvectors $\bar{\mathbf{Q}}$, so that we can define independent damping heuristics for the subspaces C and C^\perp . In particular, we want to overdamp the high frequency modes.

Physical approximation 2. We assume that the solution in C^\perp , the high frequency modes, can be modeled by Brownian dynamics, the diffusive form of the Langevin equation. To this end we select a large damping coefficient $\bar{\gamma} \gg \gamma_c$, where γ_c is the critical damping coefficient, which is fixed for C^\perp . The damping matrix applied to C^\perp is then the diagonal matrix $\bar{\Gamma}_{ii} = \bar{\gamma}$ and C^\perp is ‘over-damped’.

By selecting the damping matrix applied to C to emulate system characteristics, such as water viscosity, in conjunction with the choice of $\bar{\gamma}$ we can split the system’s degrees of freedom between C and C^\perp using Γ . Here we assume that the damping matrix applied to C is a diagonal matrix $\hat{\Gamma}_{ii} = \gamma$. The resulting system damping matrix is

$$\Gamma = \mathbf{P}_f \hat{\Gamma} + \mathbf{P}_f^\perp \bar{\Gamma}, \quad (7)$$

with the projection matrices \mathbf{P}_f and \mathbf{P}_f^\perp given in Appendix A, Eqns. (A6) and (A7).

Numerically solving the Langevin equation (5) using Eqn. (7) would be computationally expensive to implement since Γ is a matrix. Instead, we split Eqn. (5) using physical approximation

2:

$$d\hat{\mathbf{x}} = \hat{\mathbf{v}}dt, \quad \mathbf{M}d\hat{\mathbf{v}} = \hat{\mathbf{f}}dt - \hat{\mathbf{\Gamma}}\mathbf{M}\hat{\mathbf{v}}dt + \sqrt{2k_B T} \hat{\mathbf{\Gamma}}^{\frac{1}{2}} \mathbf{P}_f \mathbf{M}^{\frac{1}{2}} d\mathbf{W}_1(t), \quad (8)$$

$$d\bar{\mathbf{x}} = \bar{\mathbf{\Gamma}}^{-1} \mathbf{M}^{-1} \bar{\mathbf{f}}dt + \sqrt{2k_B T} \bar{\mathbf{\Gamma}}^{-\frac{1}{2}} \mathbf{M}^{-1} \mathbf{P}_f^\perp \mathbf{M}^{\frac{1}{2}} d\mathbf{W}_2(t), \quad (9)$$

for $\hat{\mathbf{x}} = \mathbf{P}_x \mathbf{x}$, $\hat{\mathbf{v}} = \mathbf{P}_x \mathbf{v}$ and $\hat{\mathbf{f}} = \mathbf{P}_f \mathbf{f}$ in C , $\bar{\mathbf{x}} = \mathbf{P}_x^\perp \mathbf{x}$, $\bar{\mathbf{v}} = \mathbf{P}_x^\perp \mathbf{v}$ and $\bar{\mathbf{f}} = \mathbf{P}_f^\perp \mathbf{f}$ in C^\perp , and \mathbf{W}_1 , \mathbf{W}_2 are collections of independent Wiener processes. \mathbf{P}_x , \mathbf{P}_x^\perp , \mathbf{P}_f and \mathbf{P}_f^\perp are given in Appendix A, Eqns. (A3), (A4), (A6) and (A7).

C. Multiple-time-stepping discretization (NMI).

We discretize the Langevin Equations (8) and (9) using the Langevin Impulse ($\hat{\mathbf{L}}\hat{\mathbf{I}}$) method.⁵⁰ $\hat{\mathbf{L}}\hat{\mathbf{I}}$ reduces to the Euler-Maruyama method^{51,52} when applied to the Brownian dynamics Equation (9).

We apply the following numerical approximations to the system for computational efficiency.

Numerical approximation 1. We assume that the C , C^\perp splitting at eigenvector i ensures that the frequencies found in C are bounded above by $\sqrt{\lambda_i}$ and that the time step Δt given in Eqn. (4) guarantees stability of the discretization of (8).

Numerical approximation 2. We assume that the potential energy U , given $\mathbf{f} = -\nabla_x U$, can be split into U_{fast} giving rise to the fast (local) forces $\bar{\mathbf{f}}$ in C^\perp , and U_{slow} giving rise to the slow forces such that

$$U = U_{\text{fast}} + U_{\text{slow}}, \quad (10)$$

then we propagate the Brownian dynamics equation (9) using the fast forces $\bar{\mathbf{f}} \approx -\nabla_x U_{\text{fast}}$.

This discretization becomes a multiple time stepping (MTS) Langevin method with a large time step Δt associated with (8) and a smaller time step $\Delta \tau$ with (9). We call this method Normal-Mode Impulse (**NMI**):

half kick:

$$\mathbf{v}^{n+\epsilon} = \mathbf{v}^n - \frac{\Delta t}{2} \mathbf{M}^{-1} \hat{\mathbf{f}}(\mathbf{x}^n), \quad (11)$$

fluctuation 1:

$$\mathbf{v}^{n+\frac{1}{2}} = \exp(-\gamma \Delta t) \mathbf{v}^{n+\epsilon} + \sqrt{\tau_2} \sqrt{2\gamma k_B T} \mathbf{M}^{-1} \mathbf{P}_f \mathbf{M}^{\frac{1}{2}} \mathbf{z}_1^{n+1}, \quad (12)$$

$$\mathbf{x}^{n+\frac{1}{p+1}} = \mathbf{x}^n + \mathbf{P}_x \left(\tau_1 \mathbf{v}^{n+\epsilon} + \frac{1}{\gamma} \sqrt{2\gamma k_B T} \mathbf{M}^{-1} \mathbf{P}_f \mathbf{M}^{\frac{1}{2}} (\tau_3 \mathbf{z}_1^{n+1} + \tau_4 \mathbf{z}_2^{n+1}) \right), \quad (13)$$

fluctuation 2:

$$\begin{aligned} \mathbf{x}^{n+\frac{i+1}{p+1}} &= \mathbf{x}^{n+\frac{i}{p+1}} + \Delta \tau \bar{\gamma}^{-1} \mathbf{M}^{-1} \bar{\mathbf{f}}(\mathbf{x}^{n+\frac{i}{p+1}}) \\ &\quad + \sqrt{2\Delta \tau \bar{\gamma}^{-1} k_B T} \mathbf{M}^{-1} \mathbf{P}_f^\perp \mathbf{M}^{\frac{1}{2}} \mathbf{z}^{n+\frac{i+1}{p+1}}, \end{aligned} \quad (14)$$

where $i = 1, \dots, p$,

half kick:

$$\mathbf{v}^{n+1} = \mathbf{v}^{n+\frac{1}{2}} - \frac{\Delta t}{2} \mathbf{M}^{-1} \hat{\mathbf{f}}(\mathbf{x}^{n+1}), \quad (15)$$

where **fluctuation 1** updates the slow modes and **fluctuation 2** updates the high frequency modes; γ and $\bar{\gamma}$ are the scalar Langevin damping coefficients for C and C^\perp subspaces respectively; k_B is the Boltzmann constant and T is the temperature. \mathbf{z}_1 and \mathbf{z}_2 are joint, independently sampled, random variable vectors sampled from a Gaussian distribution. Constants τ_1 , τ_2 , τ_3 and τ_4 are given by

$$\tau_1 = \frac{1 - \exp(-\gamma \Delta t)}{\gamma}, \quad (16)$$

$$\tau_2 = \frac{1 - \exp(-2\gamma \Delta t)}{2\gamma}, \quad (17)$$

$$\tau_3 = \frac{\tau_1 - \tau_2}{\sqrt{\tau_2}}, \quad (18)$$

$$\tau_4 = \sqrt{\Delta t - \frac{\tau_1^2}{\tau_2}}. \quad (19)$$

These constants are a consequence of the Langevin Impulse method and their derivation can be found in.⁵⁰

We evaluate NMI in Section III B. Even though it is possible to obtain some computational gain by the slow-fast splitting of the potential energy, and ‘over-damping’ overcomes problems

associated with MTS methods for MD, the computational gain is limited by the need to do $O(\Delta t)$ Brownian dynamics steps with a small time step $\Delta\tau$, e.g., 1 or 2 fs.

D. NML discretization.

The computational efficiency limitation of NMI motivates the use of a further numerical approximation. This leads to the derivation of the NML method described below.

Numerical approximation 3. We replace Eqn. (14) in the Euler-Maruyama method with a mass-weighted minimization scheme followed by some steps of noise. This approach comes from the observation that for sufficiently large $\bar{\gamma}$ and neglecting the noise term, Eqn. (14) minimizes the potential energy w.r.t. C^\perp as shown in Proposition 1, Section II E. Furthermore, each minimization step will partially remove the previous noise term in Eqn. (14). Thus, we replace the p steps in Eqn. (14) by:

$$\tilde{\mathbf{x}}^k = [G(\mathbf{x}^{n+\epsilon})]^k, \quad (20)$$

$$\mathbf{x}^{n+1} = \left[G(\tilde{\mathbf{x}}^k) + \sqrt{2\zeta_l k_B T} \mathbf{M}^{-1} \mathbf{P}_f^\perp \mathbf{M}^{\frac{1}{2}} \mathbf{z}_2^k \right]^l, \quad (21)$$

where k is the number of ‘non-perturbed’ minimization steps, and l the number of ‘perturbed’ minimization steps and ζ_l is the minimizer ‘line-search’ solution described below. Here $[\cdot]^j$ denotes repeating the step j times using the positions from the previous step. An example of a solution method is provided in Section II E. The minimization function G , for step n , minimization iteration i and position \mathbf{x} is

$$\mathbf{x}^{n+\frac{i+1}{p+1}} = G\left(\mathbf{x}^{n+\frac{i}{p+1}}\right) = \mathbf{x}^{n+\frac{i}{p+1}} + \zeta_i \mathbf{M}^{-1} \bar{\mathbf{f}}\left(\mathbf{x}^{n+\frac{i}{p+1}}\right), \quad (22)$$

where $\mathbf{M}^{-1} \bar{\mathbf{f}}$ is the direction of the position move and scalar ζ_i the distance in that direction determined by a ‘line-search’ method. Eqn. (22) is equivalent to Eqn. (14) with $\Delta\tau \bar{\gamma}^{-1} = \zeta_i$ and without noise. Further details of the minimization scheme can be found in Section II E. Minimization allows the forces introduced by coupling into the slow subspace C to be reduced to a value which is arbitrarily close to zero. If a local minimum is attained then the forces introduced into the slow subspace by the fast subspace are due to the random forces only, justifying Physical approximations 1 and 2 in Section II B. NML is thus defined by the equations of NMI with Eqn. (14) replaced by Equations (20) - (22).

Finally, we introduce a computational optimization which is equivalent to NML but more efficient. Note that the positions in C are moved only in **fluctuation 1** and those in C^\perp are moved only

in **fluctuation 2**. From this, and since the minimizer both relaxes the system after the ‘fluctuation 1’ step and removes the previous random perturbation, defined as

$$\delta\bar{\mathbf{x}} = \sqrt{2\zeta_l k_B T} \mathbf{M}^{-1} \mathbf{P}_f^\perp \mathbf{M}^{\frac{1}{2}} \mathbf{z}_2^k, \quad (23)$$

some computational advantage can be gained by removing $\delta\mathbf{x}$ before the next ‘fluctuation 2’. Eqn. (20) then becomes

$$\tilde{\mathbf{x}}^k = [G(\mathbf{x}^{n+\epsilon} - \delta\bar{\mathbf{x}})]^k. \quad (24)$$

E. NML minimization.

We first show that a mass-weighted steepest descent minimizer can be used to replace the drift term in Eqn. (14) in the Euler-Maruyama method.

Proposition 1 *For large $\bar{\gamma}$ the drift term in Eqn. (14) minimizes the potential energy w.r.t. C^\perp .*

If we were to propagate Eqn. (14), neglecting the noise term, at a step number n and for force

$$\bar{\mathbf{f}}^n = -\nabla_{\bar{\mathbf{x}}} U(\hat{\mathbf{x}} + \bar{\mathbf{x}}^n + \mathbf{x}_0) = -\nabla_{\bar{\mathbf{x}}} U^n, \quad (25)$$

then $\bar{\mathbf{x}}^{n+1} = \bar{\mathbf{x}}^n - \Delta\tau\bar{\gamma}^{-1}\mathbf{M}^{-1}\nabla_{\bar{\mathbf{x}}}U^n$. We can expand U around $\hat{\mathbf{x}} + \mathbf{x}_0$ using a Taylor series

$$U^{n+1} = U^n - \Delta\tau\bar{\gamma}^{-1}(\nabla_{\bar{\mathbf{x}}}U^n)^\mathbf{T} \mathbf{M}^{-1}\nabla_{\bar{\mathbf{x}}}U^n + (\Delta\tau\bar{\gamma}^{-1})^2(\cdot) + \dots. \quad (26)$$

We choose $\Delta\tau\bar{\gamma}^{-1}$ to be sufficiently small, such that Eqn. (26) converges and all terms above the second term can be ignored. Then, since $(\nabla_{\bar{\mathbf{x}}}U^n)^\mathbf{T} \mathbf{M}^{-1}\nabla_{\bar{\mathbf{x}}}U^n \geq 0$, we have

$$U^{n+1} \leq U^n \quad \forall n, \quad (27)$$

with equality when $\bar{\mathbf{f}} = \mathbf{0}$. This means that for sufficiently large $\bar{\gamma}$ and number of steps p , such that $p\Delta\tau = \Delta t$ for C propagator time step Δt , the drift term in Eqn. (14) will minimize the potential energy U w.r.t. C^\perp . \square

By considering Proposition 1 and defining $\bar{\gamma}(t, \Delta\tau)$, dependent on time and step size, we can construct a method which minimizes more rapidly. The minimizer function G for mass-weighted steepest descent Eqn. (22) is equivalent to the drift term of Eqn. (14), for large $\bar{\gamma}$, when

$$\Delta\tau \bar{\gamma}(t^n + i\Delta\tau, \Delta\tau)^{-1} = \zeta_i, \quad (28)$$

where scalar ζ_i represents the distance to move in the direction of the mass-weighted force, t^n is the time corresponding to $\bar{\mathbf{x}}^n$ and $\Delta\tau$ and i are as defined for Eqn. (14). This method will require much fewer steps to achieve a given level of minimization by determining $\bar{\gamma}(t, \Delta\tau)$ as part of a minimization scheme, improving efficiency. To obtain an increase in efficiency we require that the number of minimizer iterations $k < p/n_f$ where p is the number of steps in Eqn. (14) and n_f is the number of force evaluations per minimization step. This scheme has the advantage that using the mass-weighted force as a position move direction guarantees that the solution is in C^\perp .

We must choose a threshold u_{th} as a termination condition for the minimization algorithm. If the method gives solutions U^n which are asymptotic to the actual minimum then comparing subsequent solutions $U^{n+1} - U^n$ at minimization step $n+1$ will indicate how far from the minimum the solution is. The algorithm will then terminate when

$$U^{n+1} - U^n \leq u_{th}, \quad (29)$$

is satisfied.

The mass-weighted steepest descent method uses a quadratic approximation of the potential energy $U(\bar{\mathbf{x}} + \zeta \mathbf{M}^{-1} \bar{\mathbf{f}}) \approx a\zeta^2 + b\zeta + U(\bar{\mathbf{x}})$, where ζ is the scalar distance along the direction of the mass-weighted force $\mathbf{M}^{-1} \bar{\mathbf{f}}$. This approximation is then used to estimate the ζ corresponding to minimum U . The method is defined as follows:

$$\zeta_I = \lambda_{3N}^{-1}, \quad (30)$$

$$\bar{\mathbf{f}}^{k-1} = -\mathbf{P}_f^\perp \nabla_{\mathbf{x}} U_{\text{fast}}(\bar{\mathbf{x}}^{k-1}), \quad (31)$$

$$\bar{\mathbf{f}}^\epsilon = -\mathbf{P}_f^\perp \nabla_{\mathbf{x}} U_{\text{fast}}(\bar{\mathbf{x}}^{k-1} + \zeta_I \mathbf{M}^{-1} \bar{\mathbf{f}}^{k-1}), \quad (32)$$

$$a = -\frac{1}{\zeta_I} \left(\frac{U_{\text{fast}}(\bar{\mathbf{x}}^{k-1} + \zeta_I \mathbf{M}^{-1} \bar{\mathbf{f}}^{k-1}) - U_{\text{fast}}(\bar{\mathbf{x}}^{k-1})}{\zeta_I} - \mathbf{M}^{-1} \bar{\mathbf{f}}^{k-1} \cdot \bar{\mathbf{f}}^\epsilon \right), \quad (33)$$

$$b = \mathbf{M}^{-1} \bar{\mathbf{f}}^{k-1} \cdot \bar{\mathbf{f}}^\epsilon - 2a\zeta_I, \quad (34)$$

$$\zeta_F = -\frac{b}{2a}, \quad (35)$$

$$G(\bar{\mathbf{x}}^{k-1}) = \bar{\mathbf{x}}^k = \bar{\mathbf{x}}^{k-1} + \zeta_F \mathbf{M}^{-1} \bar{\mathbf{f}}^{k-1}, \quad (36)$$

where U_{fast} is the PE that gives rise to the fast (local) forces. These were defined as bonded plus short-range non-bonded forces in our tests. The initial estimate of ζ_I is found by solving the discreet quadratic approximation, from the MRH diagonalization, for the highest frequency mode. For a system size N this is mode number $3N$, denoted c_{3N} , and the conservative forces

are $f_{3N} = -\lambda_{3N} c_{3N}$, with solution $\zeta_I = \lambda_{3N}^{-1}$ for eigenvalue λ_{3N} . The number of steps required k is found from satisfying the termination condition in Eqn. (29) for a given threshold u_{th} . The ‘simple’ mass-weighted method uses ζ_I directly to calculate $U_{fast}(\bar{\mathbf{x}}^{k-1} + \zeta_I \mathbf{M}^{-1} \bar{\mathbf{f}}^{k-1})$. If this is less than $U_{fast}(\bar{\mathbf{x}}^{k-1})$ we set $\bar{\mathbf{x}}^k = \bar{\mathbf{x}}^{k-1} + \zeta_I \mathbf{M}^{-1} \bar{\mathbf{f}}^{k-1}$, otherwise the quadratic fit is used to find ζ_F .

For this minimization step, the damping coefficient required to obtain the equivalent damping using Eqn. (14) can be calculated from Eqn. (28). Using this $\bar{\gamma}$ in the noise term of Eqn. (21) satisfies the fluctuation-dissipation theorem.⁵³

We recognize that for much larger systems a more sophisticated approach may be necessary. Note that when PE minimization techniques which use a position move in the force direction are to be used, the position change must be restricted to C^\perp . This holds for the mass-weighted force but for other methods an additional projection is required to provide a new C^\perp force

$$\bar{\mathbf{f}}^* = \mathbf{P}_x^\perp \bar{\mathbf{f}}, \quad (37)$$

for projection matrix \mathbf{P}_x^\perp given by Eqn. (A4), such that the new position $(\mathbf{x} + \zeta_F \bar{\mathbf{f}}^*)$ is invariant in C .

III. COMPUTATIONAL TESTS

We apply NML to study the kinetics and sampling of implicit solvent models of blocked alanine dipeptide (ACE ALA NME) and BPTI using the CHARMM 22 force field. NML is implemented in the open source software PROTOMOL.⁵⁴

Supplementary information (SI) to this paper is available.⁷¹ This contains a detailed comparison of NML and NMI. In particular, it is shown that using the optimal damping coefficient, NMI is able to recover good sampling compared to $\hat{\mathbf{L}}\hat{\mathbf{I}}$. The SI contains all input files and analysis scripts to reproduce results in this section and instructions to download the executable. Figures S1 and S2 in the SI show that the dynamics using Subspace Molecular Dynamics are localized around the equilibrium structure.

The implicit solvent model used for the alanine dipeptide tests is a sigmoidal distance-dependent dielectric⁵⁵ to account for screening of electrostatic interactions due to solvent. The equation for the dielectric $\varepsilon(r)$ as a function of distance r , and parameters ϵ_s (the bulk dielectric

constant), and S (screening factor), is:

$$\varepsilon(r) = \epsilon_s \frac{\epsilon_s - 1}{2} (S^2 r^2 + 2Sr + 2) \exp(-Sr). \quad (38)$$

For these simulations we used parameter values $\epsilon_s = 78$ and $S = 0.3$.

For BPTI the distance-dependent dielectric gave unsatisfactory results for both $\hat{L}\hat{I}$ and NML: for instance, the Ramachadran plots were severely distorted compared to explicit solvent, the molecule kept expanding in volume, and these results were sensitive to the value of S . Thus, we used a more accurate implicit solvent model, the screened Coulomb Potential implicit solvent model (SCPISM), which overcame these problems. SCPISM uses the relation between the physically measurable dielectric function $\varepsilon(r)$ and the screening function $D(r)$, which for a single spherical particle in a pure solvent is given by:

$$\varepsilon(r) = D(r) \left[1 + \frac{r}{D(r)} \frac{d}{dr} D(r) \right]. \quad (39)$$

Experiments and classical electrostatic theory of polar liquids (Debye-Lorentz-Sack theory) suggest that $\varepsilon(r)$ has sigmoidal distance dependent form, a fact exploited by SCPISM and Ref. 55. SCPISM uses $\varepsilon(r)$ obtained from experiments or theory as input to the model, and then performs a numerical integration to obtain $D(r)$. The SCPISM uses atom type-dependent sigmoidal functions parameterized for the CHARMM 22 force field. Importantly, the $D(r)$ approaches bulk medium value of the screening far from the protein, and thus SCPISM does not introduce a boundary that separates the protein from the solvent.

The electrostatic component of the energy of a protein in a polar solvent is given by:

$$U(r) = \frac{1}{2} \sum_{i \neq j}^N \frac{q_i q_j}{D_s(r_{ij}) r_{ij}} + \frac{1}{2} \sum_{i=1}^N \frac{q_i^2}{D_s(R_{i,B})} \left[\frac{1}{D_s(R_{i,B})} - 1 \right], \quad (40)$$

where the first term corresponds to the interaction energy and the second to the self energy. $D_s(r)$ is a nonlinear screening term that accounts for all screening mechanisms, and $R_{i,B}$ is the Born radius of atom i , which accounts for the local environment around the atom and thus changes with protein conformation. The Born radii are estimated using a contact map method, which uses a cutoff computation around each atom, and are also modified to account for individual hydrogen bonding interactions in the protein. The latter improves the accuracy of the method. The form of the screening function is

$$D_s(r_{ij}) = \frac{1 + \epsilon_s}{1 + (\epsilon_s - 1)/2 \exp(-\alpha_{ij}r_{ij})}, \quad (41)$$

where ϵ_s is the bulk dielectric as before, and α_{ij} is a parameter that depends on the interacting atoms. More implementation details can be found in the references, including the forces and parameters for CHARMM 22.⁵⁶⁻⁶² For use of SCPISM in NMI and NML, we derived and implemented the Hessian for SCPISM (derivation found in the SI). Since we desire to reproduce kinetics for the implicit solvent biomolecules simulated, the equations of motion in C are described by a scalar Langevin equation with damping coefficient $\gamma = 91 \text{ ps}^{-1}$ to account for water viscosity.

For all of the following tests the NMA was carried out using a Hessian which had been averaged over a 500 fs MD constant energy trajectory using Leapfrog. The mass-weighted steepest descent minimizer was sufficient, and extremely efficient, for AD and BPTI. It generally reduced the PE to thresholds of $u_{th} = 0.1 \text{ kcal mol}^{-1}$, cf. Eqn. (29), within 2-5 iterations for alanine dipeptide and 2-8 iterations for BPTI.

Section III A provides discussion on the choice of the number of modes to propagate. Sampling results are presented in Section III B, rates for alanine dipeptide are shown in in Section III C. In Section III D we address the efficiency of NML. The stability of the thermostat is determined in Section III E. In all cases the results are compared to the Langevin Impulse (LÎ) method.

It is important to notice that the diagonalization of MRH often incurs negative eigenvalues for the lowest frequencies. Numerical experiments showed that NML’s stability improves if these negative eigenvalues are not present. We eliminate this artifact by averaging the Hessian in the neighborhood of the minimized structure \mathbf{x}_0 . A comparison showing that the number of negative eigenvalues decreases with averaging trajectory length can be found in Figure S3 of the SI.

A. Choice of the number of modes to propagate.

The number of modes propagated will determine the dynamics which can be recovered from the simulations. To estimate the dimension of the propagated subspace we consider the dihedral space of the molecule. Based on some numerical experiments, there seems to be a correlation between the number of backbone plus side chain dihedrals and the number of modes that should be propagated in the slow subspace. BPTI has 57 residues, and thus 171 backbone dihedrals. Of these 57 residues, 42 have side chain dihedrals χ_1 thru possibly χ_4 . GLY and ALA have no side chain dihedrals, and PRO has only two conformations. We use the following guesses for the

number of modes. Note that 6 modes to account for translational and rotational invariance of the potential energy should be added to these numbers:

1. Number of backbone ϕ, ψ dihedrals: 114 for BPTI.
2. Number of backbone ϕ, ψ dihedrals plus side chain dihedrals: BPTI has 109 side chain dihedrals for a total of 223.
3. Number of backbone ϕ, ψ, ω dihedrals plus side chain dihedrals: for BPTI this number is 280.

We used for simplicity 100, 200, and 300 modes. The results below show that choice (3) above was a good guess on the number of modes needed to obtain very good agreement with the Ramachandran plot of the $\hat{L}\hat{I}$ simulation. Choices (1) and (2) give rougher agreement: the location of the minima are preserved but their probability gets accentuated, which is consistent with the notion that the sampling is more localized to the low frequency events (see Figures S5(a)-(d) in the SI, which show 3D sampling plots). Blocked alanine dipeptide has backbone ϕ, ψ , and ω but no side chain dihedrals. The analysis above is not helpful in the case of this small molecule.

Alternatively, if one selects a target time step it is possible to find an approximate number of modes that can be propagated within the stability limit of the method. For example, we are given a model which can be propagated with the $\hat{L}\hat{I}$ method using a time step of Δt_{LI} and have a target time step Δt_T . If the eigenvalues are ordered from smallest to largest and the largest has value λ_{3N} then we find the i^{th} eigenvalue such that

$$\lambda_i \leq \frac{\lambda_{3N} \Delta t_{LI}^2}{\Delta t_T^2}, \quad (42)$$

is satisfied. The required set of ordered eigenvectors is then $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_i\}$. Examples of potential speedups for BPTI, PIN1 WW domain, and Calmodulin are in Figures S9 and Table S1 in the SI.

B. Sampling using NML or NMI.

Tests were conducted using NML to determine the minimum number of propagated modes which was consistent with good sampling. The alanine dipeptide tests use the implicit solvent model of (38), and the BPTI tests use the Screened Coulomb Potential Implicit Solvent Model

(SCPISM). The BPTI structure has PDB id 6PTI. After short minimization and equilibration for 50 ps using CHARMM⁶³ version 33, the structure was run for 10 ns using PROTOMOL. The switching function used to partition the forces at a cutoff point is given in the SI. Both programs used the CHARMM 22 force field.

For alanine dipeptide good sampling was observed when 22 (of 66) or more modes were propagated, and the Ramachandran free energy plot (in units of kcal mol⁻¹) for NML is comparable to that of $\hat{L}\hat{I}$ in Figure 1 at this point, both simulated for 500 ns. We use the metric proposed in Ref. 41 to compare the free energy plots. Over the regions which are well sampled on both simulations, that is free energy $< 3 k_B T$, the RMSD between the $\hat{L}\hat{I}$ and NML free energy plots is 0.12 kcal mol⁻¹. This compares favorably with the results of the Multiple Time Step Diffusive Langevin Method where an RMSD of 0.17 kcal mol⁻¹ was obtained for alanine dipeptide.⁴¹ Sampling of AD with NMI is reported in Figures S4(a)-(b) in the SI. An optimum choice of damping coefficient for NMI allows recovery of sampling, as expected from the theory.

The probability of being in the C7 equatorial, C5 axial and α_R , as defined in Figure 1, are given in Table I for both methods.

This level of coarse-graining allows a 5.5 fs time step with an efficiency increase of 2.1 for a minimization termination threshold (29) of 1.0 kcal mol⁻¹. It was found empirically that to obtain sampling comparable to that produced by the $\hat{L}\hat{I}$ propagator required that the minimization in C^\perp be resolved to 1 kcal mol⁻¹ or better for alanine dipeptide.

The BPTI Ramachandran free energy comparisons are shown in Figure 2 for 300 modes propagated, of a total of 2646, for a 10ns trajectory. The plots were constructed by including 43 of the 57 backbone residues dihedral $\phi - \psi$ angle pairs and were chosen to remove the residue set {1 4 5 6 12 28 29 36 37 47 49 50 56 57}. These dihedrals pairs are ones for which the minimized structure give points in the disallowed region of the Ramachadran plot, and includes all GLY. Our conclusions are not modified if we include all dihedrals, but we do this for greater clarity. The results compare well with the $\hat{L}\hat{I}$ results. Propagating 300 modes allows a 20 fs time step with an efficiency increase of 6-11 fold for minimization termination thresholds (29) of 0.5-1.0 kcal mol⁻¹. The RMSD between the $\hat{L}\hat{I}$ and NML free energy plots, over the regions which are well sampled on both simulations, that is free energy $< 3 k_B T$, is 0.184 kcal mol⁻¹.

For Langevin propagators stability or accuracy issues can arise if the condition $\gamma\Delta t \leq 1$ is not met. Since the viscosity of water requires a coefficient of around $\gamma = 91 \text{ ps}^{-1}$ then for a 1 fs time step we have $\gamma\Delta t = 0.091$, however for a 20 fs time step we have $\gamma\Delta t = 1.82$. In our tests we limit γ such that the condition is met for large time steps, for a time step of 20 fs we use $\gamma = 30 \text{ ps}^{-1}$. Since there is experimental evidence that slow modes tend to be over-damped we do not envisage this to be a problem until extremely large time steps are used for sampling. This might be an issue for kinetics for even smaller time steps.

To investigate the effect of a further reduced set of modes the BPTI Ramachandran free energy for both 100 and 200 modes propagated, of a total of 2646, for 10ns trajectories are shown in Figure 3. Again the plots include 43 of the 57 residues backbone dihedral $\Phi - \Psi$ angle pairs. Depending on the minimization level (0.1 to 1.0 kcal mol⁻¹) propagating only 200 modes allows a 34 fs time step for real speedups between 12 to 16, whereas propagating 100 modes allows a 70 fs time step for real speedups between 21 and 30. The RMSD between the L \hat{I} and NML free energy plots, over the regions which are well sampled on both simulations, that is free energy $3 k_B T$, is 0.234 kcal mol⁻¹ for Figure 3(a) and 0.213 kcal mol⁻¹ for Figure 3(b), which also compares well with the results of Ref. 41.

C. Rate calculation using NML or NMI.

With a small molecule like alanine dipeptide it is possible to sample for a sufficient length of time to measure the rates of transition between two states: in this case we measure the isomerization rate between the C7 equatorial and α_R PE wells, cf. Ref. 64. For a model with states A and B the rate from A to B , denoted k_{AB} , can be calculated using the approximation proposed by Best and Hummer⁶⁵ from the probability of transition P_{TP} and the average transition time $\langle t_{TP} \rangle$

$$2c_A k_{AB} = \frac{P_{TP}}{\langle t_{TP} \rangle}, \quad (43)$$

where c_A is the equilibrium mole fraction of conformation A . We consider conformation A to be C7 equatorial and C5 axial combined, giving $c_A = 0.83$ from Table 1. Since the C7 axial conformation is also present for this model we ignore trajectories which include it.

Rates for NMI were calculated for a model with 30 modes propagated and the remaining 36 modes over-damped. A time step of $\Delta t = 3 \text{ fs}$ was used to solve Eqn. (8) and the over damped Eqns. (14) were propagated with time steps $\Delta\tau$ of 0.5 and 1 fs. The behavior of the method with

varying $\bar{\gamma}$ was assessed by determining the rates, and the results are shown in Figure 4. Although the results compare poorly with the $\hat{L}\hat{I}$ result of 2.75 ns^{-1} , the optimum $\Delta\tau/\bar{\gamma}$ lies in the region between λ_{3N}^{-1} , the discrete quadratic solution for a mass-weighted minimizer, and the stability limit for small $\bar{\gamma}$.

The NML method results are compared to the $\hat{L}\hat{I}$ dynamics as shown in Figures 5(a) and 5(b) for 100 ns trajectories with 1 fs time step and minimization threshold of 0.1 and 0.5 kcal mol⁻¹, respectively. We denote the rates obtained by propagating the first n modes using NML as k_n , and the rate obtained from $\hat{L}\hat{I}$ as $k_{LI} \approx 2.75 \text{ ns}^{-1}$. A similar rate calculation for alanine dipeptide on explicit solvent has been reported as 10 ns^{-1} .⁶⁶ The crude screened Coulomb implicit solvent model used probably accounts for the difference in rate with the $\hat{L}\hat{I}$ simulation. The results show good correlation between k_n and k_{LI} in the region above $n = 24$ for minimizations to 0.1 and 0.5 kcal mol⁻¹, with slightly better results for the lower threshold below $n = 36$.

We determined empirically that the transition time T fits a Gamma distribution with shape parameter $k = 2.7$ and scale parameter $\theta = 2,549$. That means that the standard deviation $\sigma_T = \sqrt{k}\theta$ is close to its expected value $E(T) = k\theta$. Thus, in Figure 5 we do not use $\sigma(T)$ to quantify error bars. Rather, we compute the standard deviation resulting from running four simulations for each data point, with two different initial random velocity seeds and two different initial configurations.

For NML to provide rates comparable to those produced by the $\hat{L}\hat{I}$ propagator it can be seen from Figure 5 that the minimization threshold (27) must be resolved to 0.5 kcal mol⁻¹ or lower. This is a more stringent requirement than was found for sampling in Section III B, where 1 kcal mol⁻¹ was sufficient.

From our results, we observe that in the case of alanine dipeptide, NML (a) permits the simulation of the transition between two energy minima (states α_R and $C7$ equatorial), (b) provides rates that agree with $\hat{L}\hat{I}$. Despite the harmonic approximation of the Hessian that produced the original eigenvectors in state $C7$ equatorial, the low frequency eigenvectors are long-lived enough to be able to allow a transition between $C7$ equatorial and α_R .

D. Efficiency of NML.

The increased efficiency of the NML method comes primarily from the increased time steps available from reducing the number of modes to be propagated. The efficiency is then determined

from the reduced number of propagation steps, inversely proportional to the time step, and the cost of minimization and diagonalization. Diagonalization can be amortized over the simulation length without significant cost. For alanine dipeptide only one diagonalization was necessary to sample for up to 500 ns, whereas for BPTI we sampled for up to 10 ns with one diagonalization.

As explained in Section II A, the increase in time step is dependent on the frequency f_c at which the modes are partitioned and the relationship is shown for both the alanine dipeptide and BPTI models in Figures 6 and 7. Diagrams (a) show the maximum time step assuming a 1 fs time step for the $\hat{L}\hat{I}$ method. To assess the stability margin of the method the time steps were increased until the expected value, and variance, of the kinetic energy deviated from the statistically correct values. A threshold of 5% error was used for trajectories of 1ns, for alanine dipeptide, and 0.1ns, for BPTI. On average this occurred when the actual time step is between 2 and 3 times the projected time step, which compares well with a similar test for the $\hat{L}\hat{I}$ method where just under a 3 fs time step was possible for both models.⁶⁷ For both alanine dipeptide and BPTI we observe a jump in the projected maximum time step in Figures 6(a) and 7(a), blue plot. This shows the separation of scales between the H-X (X some heavy atom) bonds and confirms the twofold increase in time step observed empirically when using constraint methods such as SHAKE.³⁵⁻³⁷

The efficiency calculations are based on 1 ns trajectories for the alanine dipeptide and 0.05 ns trajectories for BPTI using the PROTOMOL⁵⁴ package on a 3 GHz Dual core Intel computer. The time steps used for different number of modes correspond with the eigenvector predicted values (blue line), rather than the maximum attainable, for good stability and so that a comparison can be made with $\hat{L}\hat{I}$ at 1 fs time step. For alanine dipeptide a 1ns trajectory took an average (over 5 runs) of 175.3 seconds and for the BPTI a 0.05 ns trajectory took an average (again over 5 runs) of 6820.8 seconds. The minimizer used was the ‘simple’ steepest descent using mass-weighting of the forces and the minimization was carried out to different accuracies. The range of the Lennard Jones and Coulombic forces limited by using a C^2 continuity switch, details of which are given in the SI. For alanine dipeptide the switch distances were: switch on at 4.5Å off at 5.5Å for the minimizer, on at 9Å off at 12Å for the propagator. For BPTI the switch distances were: switch on at 5.5Å off at 6.5Å for the minimizer, on at 15Å off at 20Å for the propagator. The cost of diagonalization was amortized over the 1 ns and 0.05 ns trajectories respectively. For alanine dipeptide the diagonalization took less than 0.2 seconds, for BPTI all of the modes were found in 57 seconds and 300 of 2646 could be found in 31 seconds.

The results are shown in Figures 6b and 7b. It is clear that the small size of the alanine dipeptide

limits the efficiency gains possible, as the calculation of both minimization and propagator force fields are of comparable cost. Propagating 22 modes using a 5.5 fs time step (the point beyond which sampling and rates begin to deviate from the $\hat{L}\hat{I}$ results) give an efficiency increase of 2.1 times. For BPTI, minimizing to $1.0 \text{ kcal mol}^{-1}$, which is appropriate for sampling gives the following efficiency gains: 11 for propagating 300 modes using a time step of 20 fs; 16 for propagating 200 modes using a time step of 34 fs; and 30 for propagating 100 modes using a time step of 70 fs.

The cost of the method can be separated into the normal CHARMM force field calculations e.g. $\mathcal{O}(N)$ for cutoff computations or multigrid methods or $\mathcal{O}(N \ln N)$ for particle mesh Ewald; the cost for diagonalization $\mathcal{O}(N^3)$; and the NM projection cost $\mathcal{O}(mN)$, where m is number of conserved modes.

For the simple alanine dipeptide model only one diagonalization was required, but we envisage that this will need to be carried out periodically for larger systems. If we diagonalize as often as once every N steps the cost reduces to $\mathcal{O}(N^2)$.

The projection cost will depend on the number of modes propagated, if this is proportional to the size of the system this will have a cost of $\mathcal{O}(N^2)$. However, larger systems will have a richer set of low frequency modes so it may be desirable to increase the time step, reducing the cost to the point that, at fixed m , we have $\mathcal{O}(N)$. This trend can be observed even in the small models discussed above: to obtain sampling comparable to the $\hat{L}\hat{I}$ propagator for the 22 atom model required 30% of the modes whereas the 882 atom model required only 11%. Note that the latter value is not a lower bound; that is, we did not perform an exhaustive analysis to determine the minimum number of modes that need to be kept for BPTI.

E. Thermostat stability.

We investigated the stability of our thermostat by studying the effect of increasing the time step on the mean and variance of the kinetic energy of the system. In Figure 8 results for alanine dipeptide for time steps from 1 fs to 16 fs are presented when propagating 22 modes. The results show that the thermostat works well beyond the predicted time step of 5.5 fs, and is consistent with stability results for Langevin dynamics integrators.^{67,68}

IV. DISCUSSION

We have developed a method that allows a long-time step for Langevin dynamics simulations. The degrees of freedom of the biomolecule are partitioned into its normal modes ranked by frequency. The low frequency modes are propagated according to a Langevin integrator and the high frequency modes are dampened near to their energy minimum and propagated using an accelerated Brownian dynamics propagator. This scheme allows us to speed up the run time step 2 fold for a small system such as alanine dipeptide. We also show that it scales extremely well with system size, subject to the number of modes propagated: the time-step is increased 11 fold for a molecule of 882 atoms (BPTI). The two Langevin integrators presented here satisfy fluctuation-dissipation.

Interest has grown in the use of internal-coordinates as an alternative to Cartesian coordinates for the calculation of normal modes. Since they describe the positions of the atoms in terms of distances, angles and dihedral angles, with respect to some origin atom they are seen as a more natural method of describing the geometry of biomolecules. However, the equations of motion in nonlinear coordinates become very difficult to solve. In contrast, NML can be used with well established force fields such as CHARMM 22, which are readily calculated in Cartesian coordinates.

In our tests, sampling and kinetics of small biomolecules are preserved when even only 10%-30% of modes are propagated, as indicated by the transition rates and Ramachandran plots studied. NML permits an increase in runtime efficiency which is close to the theoretical limit of being proportional to inverse time step for few propagated modes. The increase in time-step is essential to simulate time-scales that are currently out of the reach of conventional dynamics. NML can be readily applied with better implicit solvent models, such as Poisson-Boltzmann or Generalized-Born (GB), which could improve agreement with explicitly solvated simulations and experiment. All that is needed is to compute Hessians. For instance, the package NAB implements GB with analytical Hessians.⁶⁹

We would like to apply NML to study proteins undergoing conformational changes. We need to understand when the propagated space C no longer adequately describes the real dynamical space of the protein. In this case, NML needs to perform a re-diagonalization to obtain an updated propagated space. Work is underway to develop a metric to measure the validity of C and trigger re-diagonalization.

The speed of the method can also be optimized. In particular, the diagonalization of the aver-

aged Hessian with few propagated modes can be improved through a Krylov subspace approximation, and it might be possible to construct more efficient minimizers. Since NML is linear-algebra intensive, it would map well to modern specialized architectures such as GPUs.⁷⁰

V. ACKNOWLEDGMENTS

JAI acknowledges partial funding from NSF grants CCF-0135195, DBI-0450067, and CCF-0622940. VSP and PP acknowledge support from the Simbios Center (NIH U54 GM072970). JAI performed this work while on sabbatical at the Institute for Computational and Mathematical Engineering at Stanford University. Santanu Chatterjee, Paul Brenner and Trevor Cickovski from Notre Dame coded several aspects of this method and provided many of the analysis scripts. Thanks to John Chodera and Peter Minary from Stanford University and Bob Skeel from Purdue University for critical reading of earlier versions of the manuscript and several suggestions. The authors acknowledge helpful discussions with Eric Darve from Stanford. Sergio Hassan from NIH greatly assisted us in correctly implementing SCPISM in our code and comparing it to the CHARMM implementation.

APPENDIX A: PROJECTION MATRICES

The NMA partitions the system into two subspaces, C spanned by the matrix of low frequency eigenvectors \mathbf{Q} and C^\perp spanned by the high frequency eigenvectors $\bar{\mathbf{Q}}$. To project the system positions \mathbf{x} , velocities \mathbf{v} and forces \mathbf{f} onto these subspaces requires a set of projection matrices which we derive below. The position and velocity projection matrices are equivalent and we denote them as \mathbf{P}_x and \mathbf{P}_x^\perp for C and C^\perp respectively. The force projection matrices we denote as \mathbf{P}_f for C and \mathbf{P}_f^\perp for C^\perp .

The choice of system partition separates the positions around the equilibrium point \mathbf{x}_0 into $\hat{\mathbf{x}}$ in C and $\bar{\mathbf{x}}$ in C^\perp such that $\mathbf{x} = \hat{\mathbf{x}} + \bar{\mathbf{x}} + \mathbf{x}_0$. Differentiating this equation w.r.t. time gives the associated velocity equation $\mathbf{v} = \hat{\mathbf{v}} + \bar{\mathbf{v}}$. The projections are

$$\hat{\mathbf{x}} = \mathbf{P}_x(\mathbf{x} - \mathbf{x}_0), \quad \bar{\mathbf{x}} = \mathbf{P}_x^\perp(\mathbf{x} - \mathbf{x}_0), \quad (\text{A1})$$

$$\hat{\mathbf{v}} = \mathbf{P}_x \mathbf{v}, \quad \bar{\mathbf{v}} = \mathbf{P}_x^\perp \mathbf{v}, \quad (\text{A2})$$

where

$$\mathbf{P}_x = \mathbf{M}^{-\frac{1}{2}} \mathbf{Q} \mathbf{Q}^T \mathbf{M}^{\frac{1}{2}}, \quad (\text{A3})$$

and

$$\mathbf{P}_x^\perp = \mathbf{M}^{-\frac{1}{2}} (\mathbf{I} - \mathbf{Q}\mathbf{Q}^T) \mathbf{M}^{\frac{1}{2}}, \quad (\text{A4})$$

from Eqn. (3), where the positions are mapped from the Cartesian space to mode space and back to Cartesian space.

We also need projection matrices applicable to the forces. Given the system forces \mathbf{f} , we can determine $\hat{\mathbf{f}}$ in C by considering Newton's second law $\mathbf{M}\ddot{\mathbf{x}} = \mathbf{f}$, and Eqns. (3) to derive a projection matrix. We can map Newton's second law into \hat{c} , the modes in the C subspace

$$\ddot{\hat{c}} = \mathbf{Q}^T \mathbf{M}^{-\frac{1}{2}} \mathbf{f}. \quad (\text{A5})$$

To evolve the system in Cartesian space we can map the equation back to the Cartesian space and multiply through by \mathbf{M} . The normal mode subspace force projected onto the the real space is

$$\hat{\mathbf{f}} = \mathbf{P}_f \mathbf{f} = \left(\mathbf{M}^{\frac{1}{2}} \mathbf{Q}\mathbf{Q}^T \mathbf{M}^{-\frac{1}{2}} \right) \mathbf{f}, \quad (\text{A6})$$

for C force projection matrix \mathbf{P}_f .

$\bar{\mathbf{f}}$ can be found in a similar manner to $\hat{\mathbf{f}}$ but using the complement of the projection matrix

$$\bar{\mathbf{f}} = \mathbf{P}_f^\perp \mathbf{f} = \left(\mathbf{M}^{\frac{1}{2}} (\mathbf{I} - \mathbf{Q}\mathbf{Q}^T) \mathbf{M}^{-\frac{1}{2}} \right) \mathbf{f}, \quad (\text{A7})$$

for C^\perp force projection matrix \mathbf{P}_f^\perp .

* Electronic address: pande@stanford.edu

† Electronic address: izaguirr@nd.edu

¹ C. S. M. Levitt and P. Stern, *J. Mol. Biol.* **181**, 423 (1985).

² O. Marques and Y. Sanejouand, *Proteins* **23**, 557 (1995).

³ F. Tama and Y. Sanejouand, *Protein Eng.* **14**, 1 (2001).

⁴ G. Li and Q. Cui, *Biophys. J.* **86**, 743 (2004).

⁵ J. Ma, *Structure (Camb)* **13**, 373 (2005).

⁶ O. M. F. Tama and C. B. 3rd, *J. Struct Biol.* **147**, 315 (2004).

⁷ O. M. F. Tama, F.X. Gadea and Y. Sanejouand, *Proteins* **41**, 1 (2000).

⁸ M. Delarue and Y. Sanejouand, *J. Mol. Biol.* **320**, 1011 (2002).

⁹ G. T. P. Durand and Y. Sanejouand, *Biopolymers* **34**, 759 (1994).

- 10 B. Brooks and M. Karplus, Proc. Natl. Acad. Sci. USA **82**, 4995 (1985).
- 11 G. Li and Q. Cui, Biophys J. **83**, 2457 (2002).
- 12 A. A. I. Bahar and B. Erman, Fold. Des. **2**, 173 (1997).
- 13 R. J. P. Doruker and I. Bahar, J. Comput. Chem. **23**, 119 (2002).
- 14 M. A. Balsera, W. Wriggers, Y. Oono, and K. Schulten, J. Phys. Chem. **100**, 2567 (1996).
- 15 A. Amadei, A. B. M. Linnsen, and H. J. C. Berendsen, PROTEINS: Struct., Func., and Genetics **17**, 412 (1993).
- 16 G. Zhang and T. Schlick, J. Comp. Chem. **14**, 1212 (1993).
- 17 Z. Zhang, Y. Shi, and H. Liu, Biophys. J. **84**, 3583 (2003).
- 18 A. R. Atilgan, S. R. Durell, R. L. Jernigan, M. C. Demirel, O. Keskin, and I. Bahar, Biophys. J. **80**, 505 (2001).
- 19 I. Kolossvary and G. Keser³; J. Comp. Chem. **22**, 21 (2001).
- 20 K. Suhre and Y.-H. Sanejouand, Nucl. Acids Res. **32**, W610 (2004).
- 21 B. Space and Herschel, J. Chem. Phys. **99**, 9070 (1993).
- 22 B. S. A. Askar and A. Rabitz, J. Phys. Chem. **99**, 7330 (1995).
- 23 O. Lange and H. Grubmuller, J. Chem. Phys. **124**, 2149 (2006).
- 24 E. B. T. Schlick and M. Mandziuk, Annu Rev Biophys Biomol Struct **26**, 181 (1997).
- 25 A. Nyberg and T. Schlick, Chem. Phys. Lett. **198**, 538 (1992).
- 26 K. Watanabe, J. Phys. Chem. **99**, 5680 (1995).
- 27 S. Gibson, J. Comput. Chem. **11**, 468 (1990).
- 28 H. Scully, Mol. Simul. **11**, 67 (1993).
- 29 D. T. W.B. Streett and G. Saville, Mol. Phys. **35**, 639 (1978).
- 30 H. Grubmüller, H. Heller, A. Windemuth, and K. Schulten, Molecular Simulation **6**, 121 (1991).
- 31 M. Tuckerman, B. J. Berne, and G. J. Martyna, J. Chem. Phys. **97**, 1990 (1992).
- 32 Q. Ma, J. A. Izaguirre, and R. D. Skeel, SIAM J. Sci. Comput. **24**, 1951 (2003).
- 33 B. García-Archilla, J. M. Sanz-Serna, and R. D. Skeel, SIAM J. Sci. Comput. **20**, 930 (1998).
- 34 J. A. Izaguirre, S. Reich, and R. D. Skeel, J. Chem. Phys. **110**, 9853 (1999).
- 35 G. C. J.P. Ryckaert and H. J. Berendsen, J. Comput. Phys. **23**, 327 (1977).
- 36 W. F. van Gunsteren and H. J. C. Berendsen, Mol. Phys. **34**, 1311 (1977).
- 37 H. Andersen, J. Comp. Phys. **52**, 24 (1983).
- 38 K. D. Hammonds and J.-P. Ryckaert, Comput. Phys. Commun. **62**, 336 (1991).

- ³⁹ G. Ciccotti, M. Ferrario, and J.-P. Ryckaert, *Mol. Phys.* **47**, 1253 (1982).
- ⁴⁰ W. F. van Gunsteren and M. Karplus, *Macromolecules* **15**, 1528 (1982).
- ⁴¹ P. Eastman and S. Doniach, *Proteins* **30**, 215 (1998).
- ⁴² W. F. van Gunsteren, *Mol. Phys.* **40**, 1015 (1980).
- ⁴³ J. Kottalam and D. Case, *Biopolymers* **29**, 1409 (1990).
- ⁴⁴ J. Smith, *Q Rev Biophys* **24**, 227 (1991).
- ⁴⁵ A. D. MacKerell Jr., D. Bashford, M. Bellott, R. L. Dunbrack Jr., J. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, et al., *FASEB J.* **A143**, 6 (1992).
- ⁴⁶ A. D. MacKerell Jr., D. Bashford, M. Bellott, R. L. Dunbrack Jr., J. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, et al., *J. Phys. Chem. B* **102**, 3586 (1998).
- ⁴⁷ P. Petrone and V. S. Pande, *Biophys. J.* **90**, 1583 (2006).
- ⁴⁸ B. Leimkuhler and S. Reich, *Simulating Hamiltonian Dynamics* (Cambridge University Press, 2004).
- ⁴⁹ P. F. Batcho and T. Schlick, *J. Chem. Phys.* **115**, 4019 (2001).
- ⁵⁰ R. D. Skeel and J. A. Izaguirre, *Mol. Phys.* **100**, 3885 (2002).
- ⁵¹ D. L. Ermak and J. A. McCammon, *J. Chem. Phys.* **69**, 1352 (1987).
- ⁵² P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, vol. 23 of *Applications of Mathematics: Stochastic Modelling and Applied Probability* (Springer-Verlag, New York, 1992), second corrected printing 1995.
- ⁵³ C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences* (Springer-Verlag, New York, 1985).
- ⁵⁴ T. Matthey, T. Cickovski, S. S. Hampton, A. Ko, Q. Ma, M. Nyerges, T. Raeder, T. Slabach, and J. A. Izaguirre, *ACM Trans. Math. Softw.* **30**, 237 (2004).
- ⁵⁵ K. F. F. Min-Yi Shen, *J. Comp. Chem.* **26**, 691 (2005).
- ⁵⁶ S. Hassan, E. Mehler, D. Zhang, and H. Weinstein, *PROTEINS: Struc., Func., and Genetics* **51**, 109 (2003).
- ⁵⁷ S. Hassan, F. Guarneri, and E. Mehler, *J. Phys. Chem. B* **104**, 6478 (2000).
- ⁵⁸ S. A. Hassan, F. Guarneri, and E. Mehler, *J. Phys. Chem. B* **104**, 6490 (2000).
- ⁵⁹ S. Hassan and E. Mehler, *PROTEINS: Struc., Func., and Genetics* **47**, 45 (2002).
- ⁶⁰ S. Hassan, *J. Phys. Chem. B* **50**, 19501 (2004).
- ⁶¹ S. Hassan and E. Mehler, *Int. J. Quant. Chem.* **102**, 986 (2005).
- ⁶² X. Li, S. Hassan, and E. Mehler, *PROTEINS: Structure, Function, and Bioinformatics* **60**, 464 (2005).

- ⁶³ B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, *J. Comp. Chem.* **4**, 187 (1983).
- ⁶⁴ J. D. Chodera, W. C. Swope, J. W. Pitera, and K. A. Dill, *Multiscale Modeling & Simulation* **5**, 1214 (2006).
- ⁶⁵ R. B. Best and G. Hummer, *PNAS* **102**, 6732 (2005).
- ⁶⁶ P. G. Bolhuis, C. Dellago, and D. Chandler, *Proc. natl. Acad. Sci. USA* **97**, 5877 (2000).
- ⁶⁷ B. Mishra and T. Schlick, *J.Chem. Phys.* **105**, 299 (1996).
- ⁶⁸ D. A. Beard and T. Schlick, *J. Chem. Phys.* **112**, 7313 (2000).
- ⁶⁹ R. Brown and D. Case, *J. Comp. Chem.* **27**, 1662 (2006).
- ⁷⁰ E. Elsen, M. Houston, V. Vishal, E. Darve, P. Hanrahan, and V. Pande, *Proc. Supercomputing '06* (2006).
- ⁷¹ See EPAPS Document No. XXX for Supplementary Information (SI) to this paper. This document can be reached through a direct link in the online article's reference section or via the EPAPS homepage (<http://www.aip.org/pubservs/epaps.html>)

CAPTIONS

Figure 1

Comparison of alanine dipeptide Ramachandran free energy plots for (a) Langevin Impulse $\Delta t = 1$ fs and (b) NML with 22 modes (of 66) propagated $\Delta t = 5.5$ fs. Data obtained from 500 ns trajectories with $\gamma = 91$ ps⁻¹. The color-bar represents the free energy in kcal mol⁻¹ and the RMSD free energy difference between the Figures is 0.119 kcal mol⁻¹, which compares well with the Multiple Time Step Diffusive Langevin Method⁴¹ RMSD of 0.17 kcal mol⁻¹.

Figure 2

Comparison of BPTI Ramachandran free energy plots for (a) Langevin Impulse $\Delta t = 1$ fs, $\gamma = 91$ ps⁻¹ and (b) NML with 300 modes propagated (of 2646) with $\Delta t = 20$ fs and $\gamma = 30$ ps⁻¹. Data for 43 of the 57 residues backbone dihedral $\Phi - \Psi$ pairs from 10 ns trajectories. The color-bar represents the free energy in kcal mol⁻¹ and the RMSD free energy difference between the Figures is 0.184 kcal mol⁻¹.

Figure 3

BPTI Ramachandran free energy plots for NML with (a) 200 modes propagated (of 2646) with $\Delta t = 34$ fs, $\gamma = 28$ ps⁻¹ and (b) 100 (of 2646) with $\Delta t = 70$ fs, $\gamma = 12$ ps⁻¹. Data for 43 of the 57 residues backbone dihedral $\Phi - \Psi$ pairs from 10 ns trajectories. The color-bar represents the free energy in kcal mol⁻¹, the RMSD free energy difference between (a) and Figure 2(a) is 0.213 kcal mol⁻¹ and (b) and Figure 2(a) is 0.234 kcal mol⁻¹.

Figure 4

Calculated transition rates for the over-damped NMI method for different $\bar{\gamma}$, where 30 modes were propagated and the remaining 36 modes over-damped with time steps of $\Delta\tau = 0.5$ and $\Delta\tau = 1$ fs. The comparable rate for $L\hat{I}$ is 2.75 ns⁻¹. Results show that the optimum $\Delta\tau/\bar{\gamma}$ lies in the region between λ_{3N}^{-1} , the discrete quadratic solution for a mass-weighted minimizer, and the stability limit for small $\bar{\gamma}$.

Figure 5

Calculated transition rates for NML for different minimization accuracies compared to Langevin Impulse (red line) for alanine dipeptide, as a function of modes propagated. Mode number indicates the last propagated mode, after which all higher frequency mode forces are relaxed. The error bars are determined from running simulations with different initial random velocity seeds and initial conditions, and represent one standard deviation. As expected, for mode 66, all modes

are propagated and the rate is close to the Langevin Impulse rate.

Figure 6

(a) Alanine dipeptide maximum expected time step (blue), and maximum obtained time step for 1 ns trajectories (green). (b) NML efficiency increase by reducing the number of propagated modes. Mode number indicates the last propagated mode. Efficiency also depends on the minimization step, which has a computational cost depending on the threshold. We show curves for 1.0, 0.5 and 0.1 kcal mol⁻¹ thresholds compared to Langevin Impulse ($\hat{L}\hat{I}$).

Figure 7

BPTI maximum expected time step (blue), and maximum obtained time step for 0.1 ns trajectories (green). (b) NML efficiency increase by reducing the number of propagated modes. Mode number indicates the last propagated mode. Efficiency also depends on the minimization step and we show curves for 1.0, 0.5 and 0.1 kcal mol⁻¹ thresholds compared to Langevin Impulse ($\hat{L}\hat{I}$).

Figure 8

Kinetic energy stability for alanine dipeptide, propagating 22 modes, with increasing time step. The time step predicted from the eigenvalues is 5.5 fs.

FIGURES

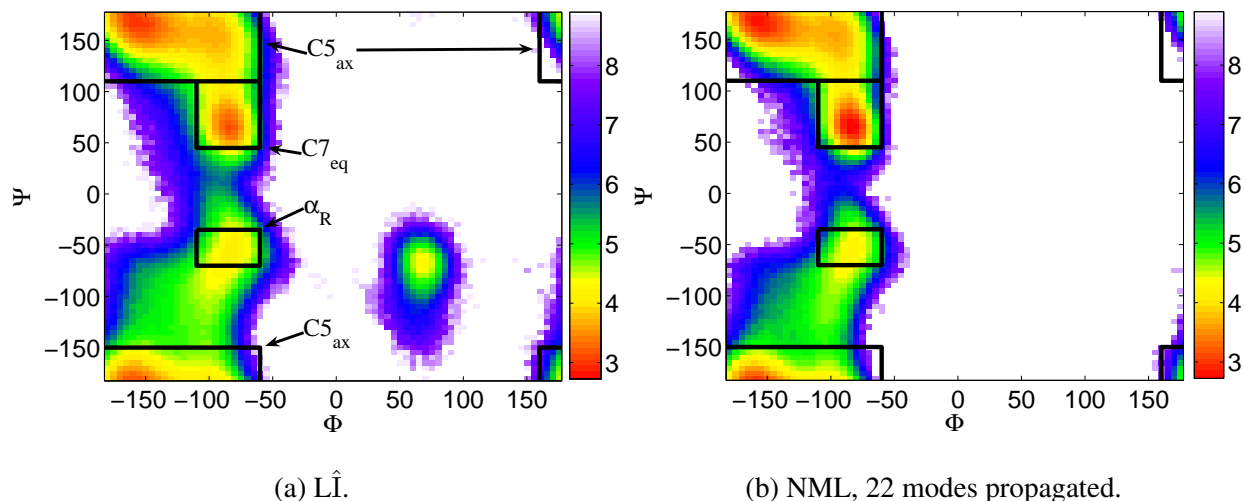


FIG. 1: Comparison of alanine dipeptide Ramachandran free energy plots for (a) Langevin Impulse $\Delta t = 1$ fs and (b) NML with 22 modes (of 66) propagated $\Delta t = 5.5$ fs. Data obtained from 500 ns trajectories with $\gamma = 91$ ps $^{-1}$. The color-bar represents the free energy in kcal mol $^{-1}$ and the RMSD free energy difference between the Figures is 0.119 kcal mol $^{-1}$, which compares well with the Multiple Time Step Diffusive Langevin Method⁴¹ RMSD of 0.17 kcal mol $^{-1}$

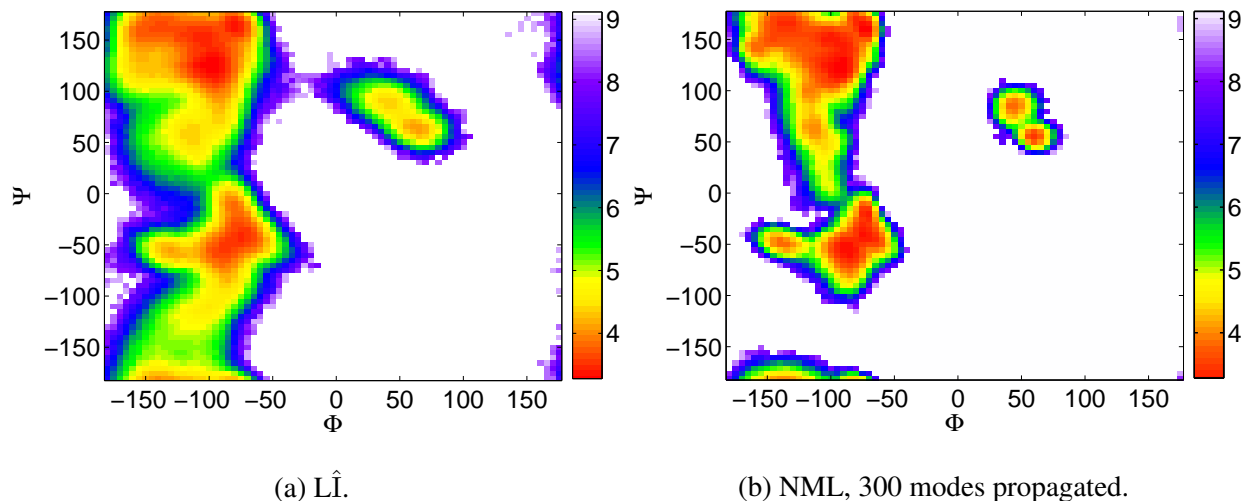
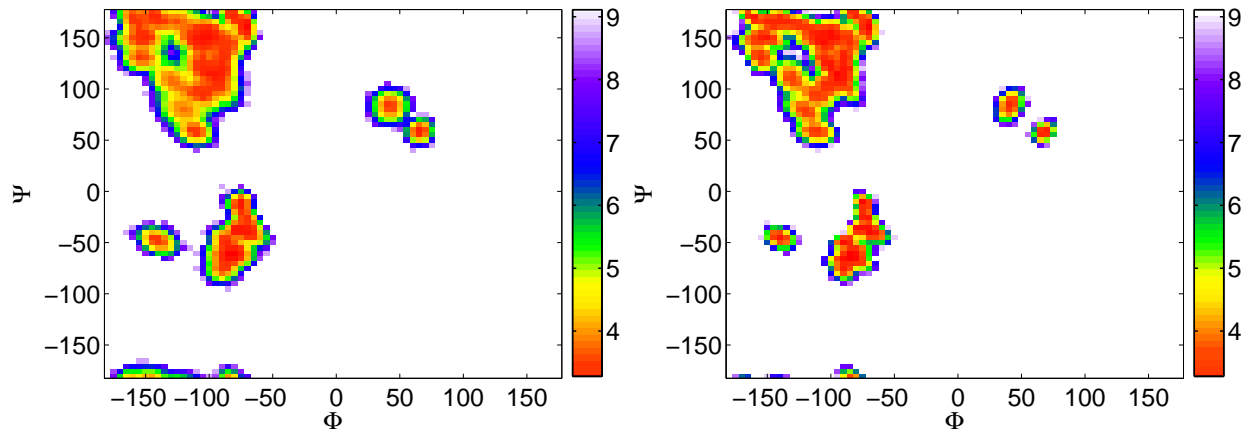


FIG. 2: Comparison of BPTI Ramachandran free energy plots for (a) Langevin Impulse $\Delta t = 1$ fs, $\gamma = 91$ ps $^{-1}$ and (b) NML with 300 modes propagated (of 2646) with $\Delta t = 20$ fs and $\gamma = 30$ ps $^{-1}$. Data for 43 of the 55 residues backbone dihedral $\Phi - \Psi$ pairs from 10 ns trajectories. The color-bar represents the free energy in kcal mol $^{-1}$ and the RMSD free energy difference between the Figures is 0.184 kcal mol $^{-1}$



(a) NML, 200 modes propagated.

(b) NML, 100 modes propagated.

FIG. 3: BPTI Ramachandran free energy plots for NML with (a) 200 modes propagated (of 2646) with $\Delta t = 34$ fs, $\gamma = 28$ ps $^{-1}$ and (b) 100 (of 2646) with $\Delta t = 70$ fs, $\gamma = 12$ ps $^{-1}$. Data for 43 of the 55 residues backbone dihedral $\Phi - \Psi$ pairs from 10 ns trajectories. The color-bar represents the free energy in kcal mol $^{-1}$, the RMSD free energy difference between (a) and Figure 2(a) is 0.213 kcal mol $^{-1}$ and (b) and Figure 2(a) is 0.234 kcal mol $^{-1}$.

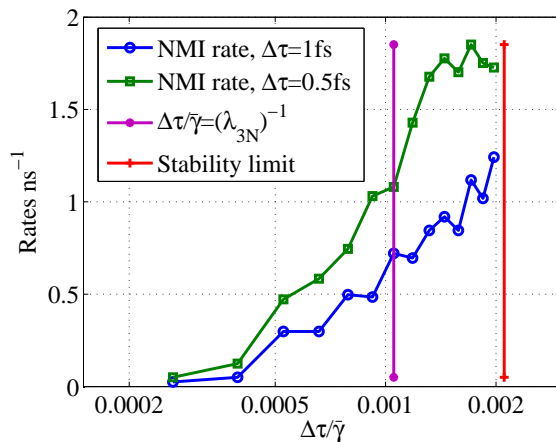


FIG. 4: Calculated transition rates for the over-damped NMI method for different $\bar{\gamma}$, where 30 modes were propagated and the remaining 36 modes over-damped with time steps of $\Delta\tau = 0.5$ and $\Delta\tau = 1$ fs. The comparable rate for $\hat{L}\hat{I}$ is 2.75 ns $^{-1}$. Results show that the optimum $\Delta\tau/\bar{\gamma}$ lies in the region between λ_{3N}^{-1} , the discrete quadratic solution for a mass-weighted minimizer, and the stability limit for small $\bar{\gamma}$.

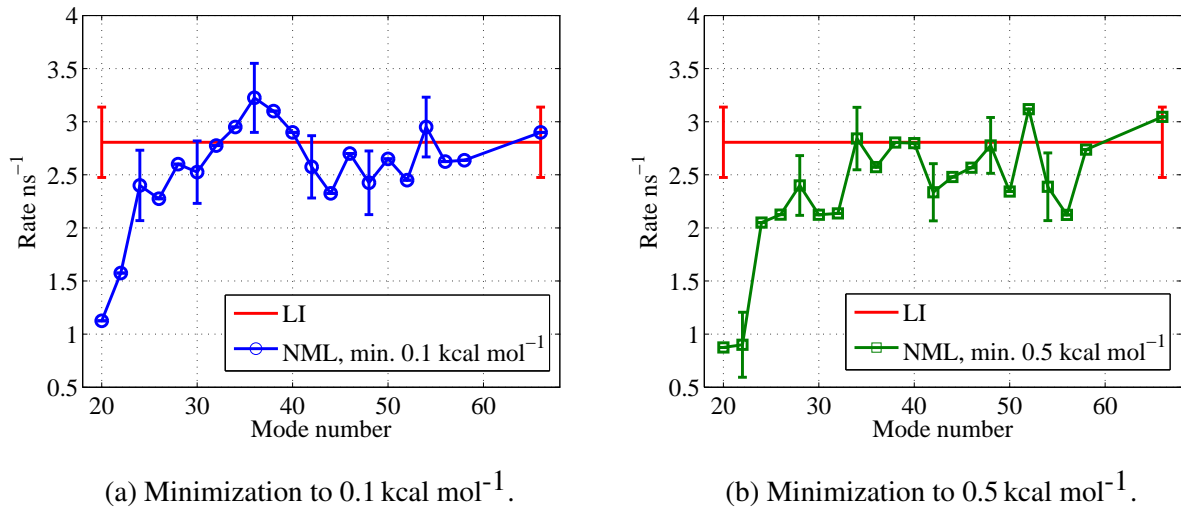


FIG. 5: Calculated transition rates for NML for different minimization accuracies compared to Langevin Impulse (red line) for alanine dipeptide, as a function of modes propagated. Mode number indicates the last propagated mode, after which all higher frequency mode forces are relaxed. The error bars are determined from running simulations with different initial random velocity seeds and initial conditions, and represent one standard deviation. As expected, for mode 66, all modes are propagated and the rate is close to the Langevin Impulse rate.

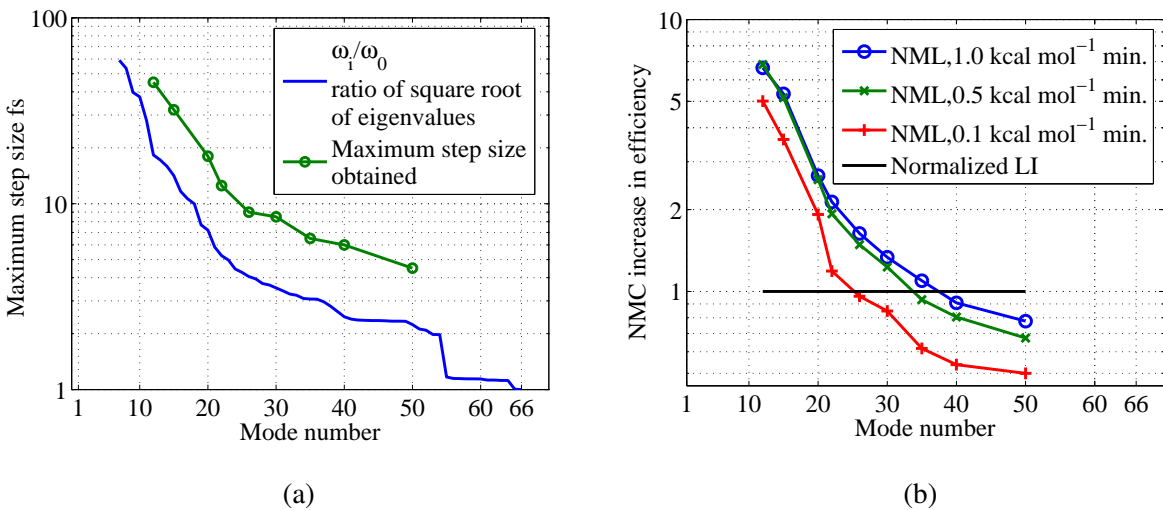


FIG. 6: (a) Alanine dipeptide maximum expected step size (blue), and maximum obtained step size for 1 ns trajectories (green). (b) NML efficiency increase by reducing the number of propagated modes. Mode number indicates the last propagated mode. Efficiency also depends on the minimization step, which has a computational cost depending on the threshold. We show curves for 1.0, 0.5 and 0.1 kcal mol⁻¹ thresholds compared to Langevin Impulse ($\hat{L}\hat{I}$).

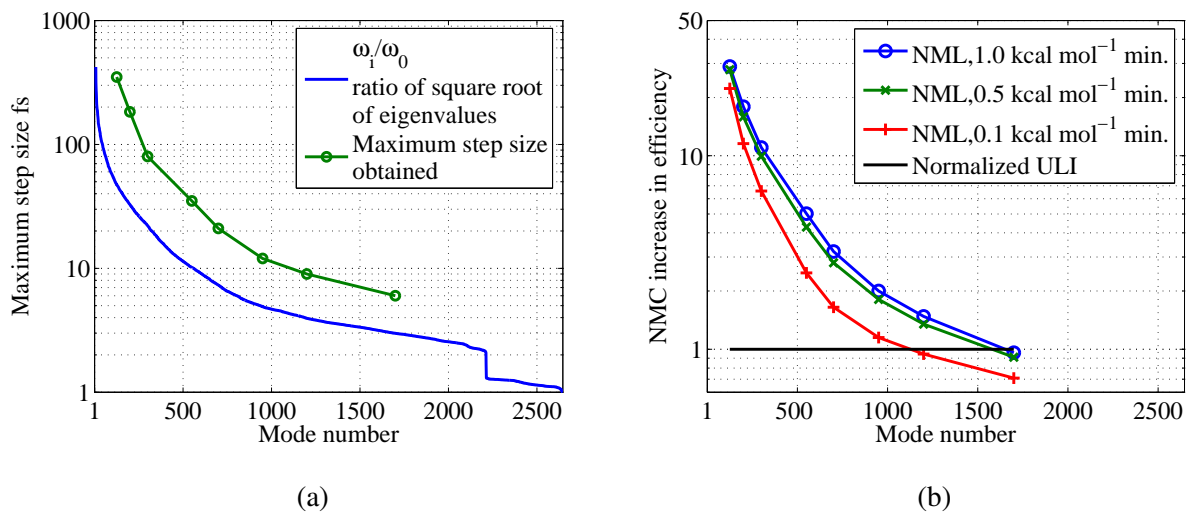


FIG. 7: BPTI maximum expected step size (blue), and maximum obtained step size for 0.1 ns trajectories (green). (b) NML efficiency increase by reducing the number of propagated modes. Mode number indicates the last propagated mode. Efficiency also depends on the minimization step and we show curves for 1.0, 0.5 and 0.1 kcal mol⁻¹ thresholds compared to Langevin Impulse (\hat{L}).

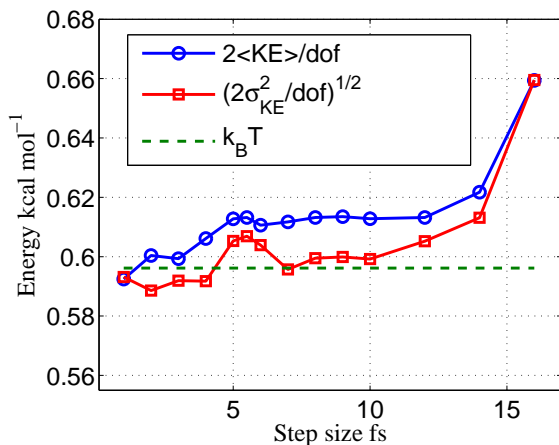


FIG. 8: Kinetic energy stability for alanine dipeptide, propagating 22 modes, with increasing time step. The time step predicted from the eigenvalues is 5.5 fs.

TABLES

Method	C7 equ.	C5 ax.	α_R
L \hat{I}	0.183	0.637	0.049
NML	0.271	0.632	0.039

TABLE I: Probabilities of the alanine dipeptide model occupying C7 equatorial, C5 axial and α_R positions in the $\Phi - \Psi$ plane, as defined in Figure 1, for NML and L \hat{I} propagators.