

[C. R. SWEET](#), [P. PETRONE](#),
[V. S. PANDE](#), [J. A. IZAGUIRRE](#)

NORMAL MODE PARTITIONING OF LANGEVIN DYNAMICS FOR BIOMOLECULES

SUPPLEMENTARY INFORMATION

[Introduction](#)

[Additional evidence related to NML](#)

[Localization of SMD](#)

[Averaged Hessian](#)

[3D sampling plots](#)

[Comparison of NMI discretization and NML](#)

[sampling](#)

[kinetics](#)

[Implicit Solvent models](#)

[Comparison of speedup for different models](#)

[Simulation and implementation details](#)

[Input files](#)

[Analysis scripts](#)

[Directions to download executable](#)

[Implementation of NML in ProtoMol](#)

[List of source code files](#)

[List of configuration file options](#)

[Protomol C2 switch](#)

INTRODUCTION

There are three main aims of this supplementary information. First, we present *additional support for some claims related to NML*, the normal mode partitioning of Langevin dynamics integrator. Second, we present a *comparison of NML and the normal mode impulse NMI method* of Eqn. (14) in the paper. The main conclusions, reported in the paper, are that NMI is able to sample correctly, but is unable to reproduce kinetics. This is evidence that the minimization of the energy due to the fast frequency modes effectively eliminates coupling between spaces, and thus points out the advantage of the

approach taken by NML as advocated in the paper. Additionally, NMI is less efficient than NML. Third, we produce *details needed for anyone to reproduce our results*. In this document we explain the testing methodology, a set of input files needed for one of the simulations, and the analysis methodology. Then we explain the structure of the software implementation of these methods. Finally a table of simulation input and outputs is given; these are available as a downloadable tar file for which a link is provided. Further information on the method is available at <http://www.normalmodes.info>.

ADDITIONAL EVIDENCE RELATED TO NML

Localization of SMD

Comparison of Ramachandran plots for alanine dipeptide propagated with the Langevin Impulse (Figure S1) and 'Subspace Molecular Dynamics' (SMD) method (Figure S2). The dynamics of the SMD method is localized by the subspace coupling, even after a 100ns trajectory.

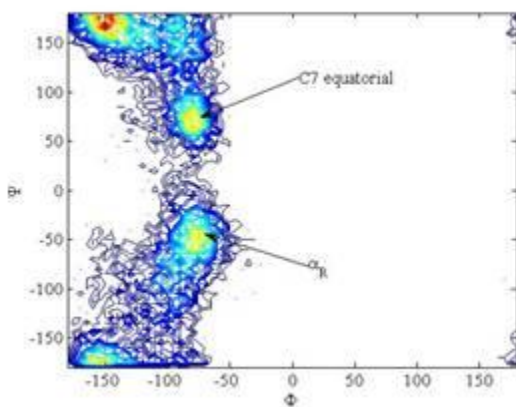


Figure S1. LI propagator.

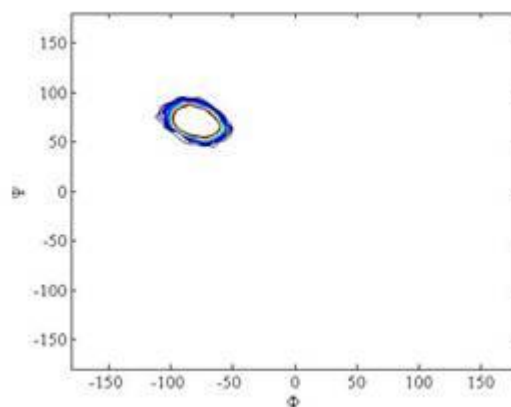


Figure S2. SMD propagator.

Averaged Hessian

Diagonalization at a potential energy (PE) minimum should give all positive eigenvalues as the PE is convex at that point. In practice negative eigenvalues will exist, even after careful minimization. We use an averaging technique, where the mass re-weighted Hessian is averaged over an NVE trajectory. In Figure S3 we compare the number of negative eigenvalues with the number of steps over which the Hessian is averaged. The initial velocities are drawn from a Boltzmann distribution at two different temperatures for comparison.

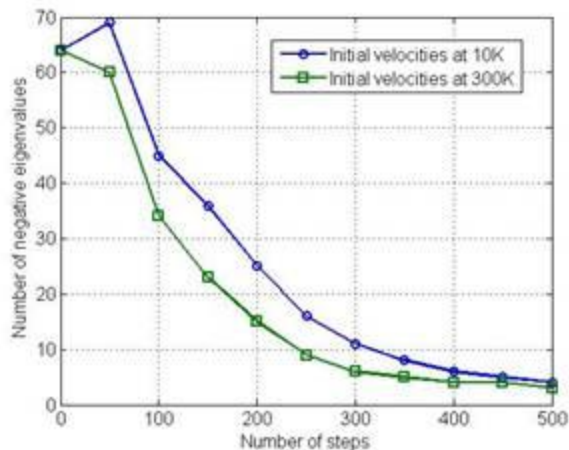


Figure S3 Number of remaining negative eigenvalues with number of steps the Hessian is averaged over.

3D sampling plots

Probability plots for alanine dipeptide, 100ns trajectory and BPTI, 10ns trajectory, for different numbers of propagated modes. The BPTI Ramachandran plot is constructed using a subset of the 57 backbone residue dihedrals which excludes the residue set {1 4 5 6 12 28 29 36 37 47 49 50 56 57}.

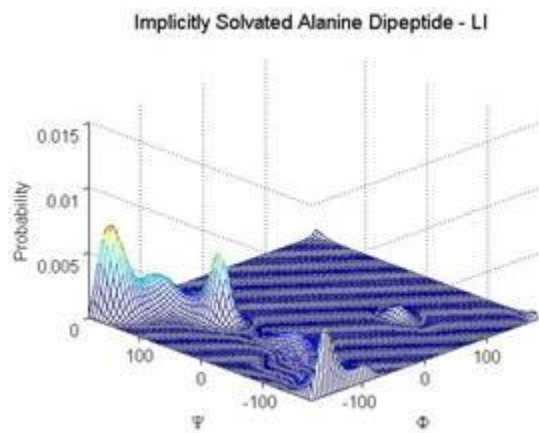


Figure S4 (a)

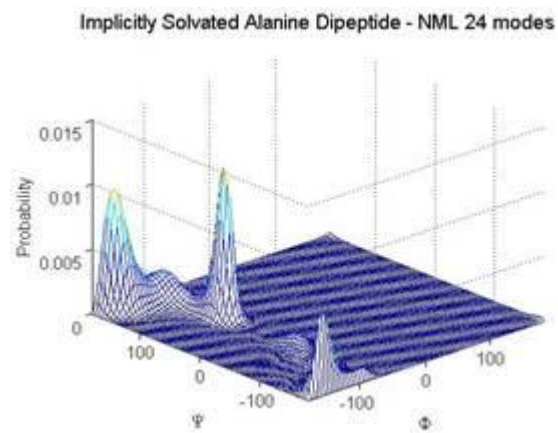


Figure S4 (b)

Imp. Solv. BPTI- 43 of 57 Backbone Dihedrals - LI

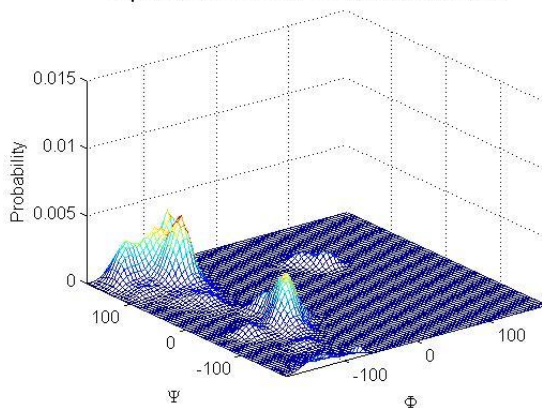


Figure S5 (a)

Imp. Solv. BPTI- 43 of 57 Backbone Dihedrals - NML 300 modes

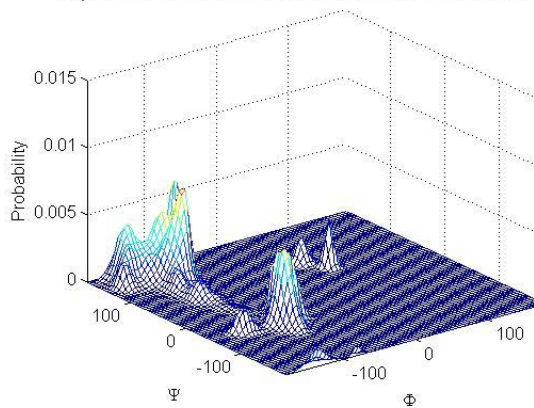


Figure S5 (b)

Imp. Solv. BPTI- 43 of 57 Backbone Dihedrals - NML 200 modes

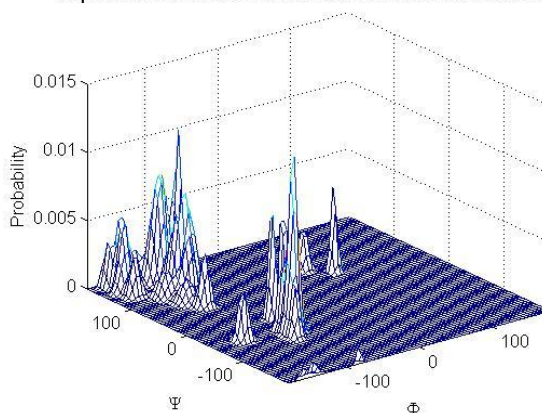


Figure S5 (c)

Imp. Solv. BPTI- 43 of 57 Backbone Dihedrals - NML 100 modes

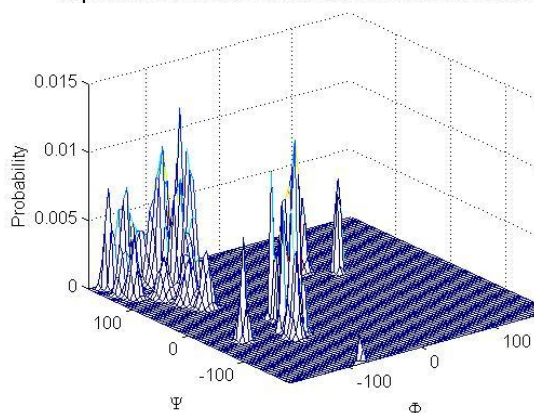


Figure S5 (d)

COMPARISON OF NMI DISCRETIZATION AND NML

SAMPLING

The NMI discretization is able to sample correctly, similarly to NML, as expected.

The over-damped Normal Mode Impulse method (NMI), upon which NML is based, was tested to study the relationship between the damping coefficient, sampling and stability. 30 modes of 66 were propagated. For very large damping coefficients the sampling was poor, but there exists a region close to the stability limit where sampling is much improved. In Figures S6(a)-(c) the Ramachandran plots for alanine dipeptide are shown for the full system with Langevin impulse, the very large damping coefficient and the optimal damping coefficient. The graph shows the rate from conformation C7equatorial to α_R with different damping coefficients.

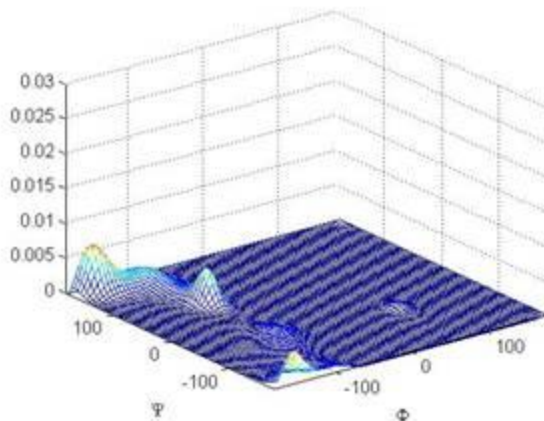


Figure S6(a). LI sampling.

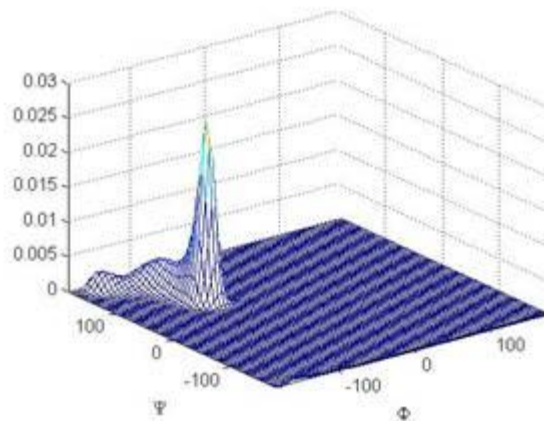


Figure S6(b). NMI sampling large γ .

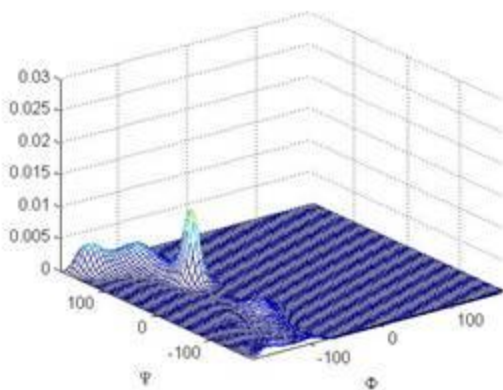


Figure S6(c). NMI sampling, optimum γ .

KINETICS

The NMI discretization is unable to recover kinetics, unlike NML.

For Normal Mode Impulse (NMI) we see in Figure S7 that the optimal damping coefficient is close to the reciprocal of the largest eigenvalue $\lambda_{3N} \approx 950$, which would be the line-search solution for a minimizer for the discrete quadratic approximation. The method was unstable for $\Delta\tau/\gamma$ of greater than 2.16×10^{-3} , indicating that the propagator for high frequency degrees of freedom must be in the minimization region for stability. For comparison the Langevin Impulse rate (assuming a 0.83 population) was 2.75 ns^{-1} . The results for NML are presented in Figure S8, showing excellent rates when compared to LI.

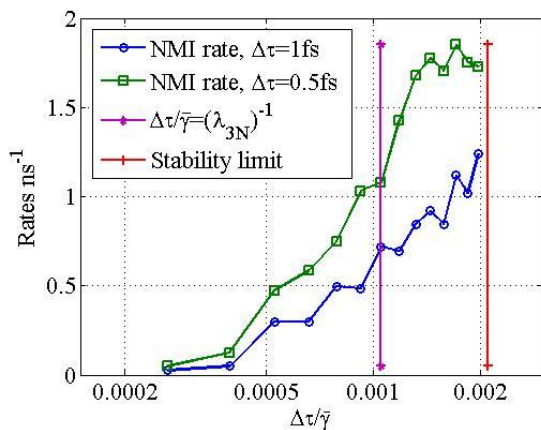


Figure S7. NMI rates.

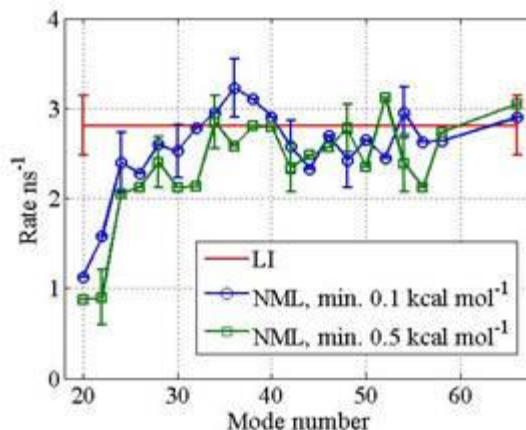


Figure S8. NML rates.

IMPLICIT SOLVENT MODELS

The implicit solvent model used for the alanine dipeptide tests is a sigmoidal distance-dependent dielectric to account for screening of electrostatic interactions due to solvent. For BPTI the distance dependent dielectric gave unsatisfactory results: for instance, the Ramachandran 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). SCPISM uses the relation between the physically measurable dielectric function $\epsilon(r)$ and the screening function $D(r)$.

The derivation of the SCPISM Hessian, required for correct diagonalization, can be found in the [zipped archive](#) discussed in Section SIMULATION AND IMPLEMENTATION DETAILS.

COMPARISON OF SPEEDUP FOR DIFFERENT MODELS

The following figure and table compares the speedup that can be achieved for different protein models.

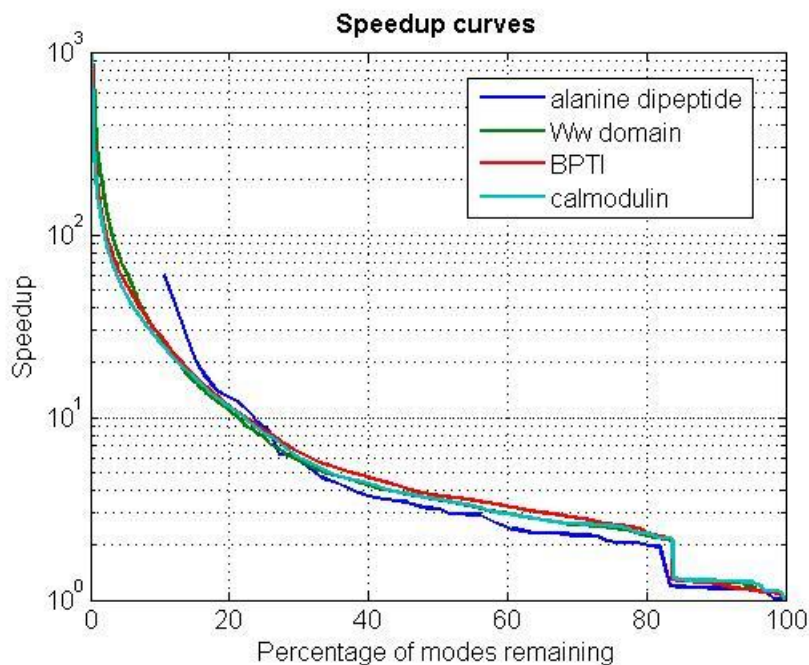


Figure S9.

Molecule	Total No. of modes	Residues*2/ Speedup	Speedup for % modes				
			10%	20%	40%	60%	80%
Alanine dipeptide	66	1/-	67.8	12.9	3.8	2.6	2.0
Ww domain	1653	66/78.3	27.5	10.8	4.2	3.0	2.3
BPTI	2646	114/61.4	28.3	11.5	4.7	3.3	2.3
calmodulin	6603	288/53.7	25.7	11.3	4.3	3.0	2.3

Table S1

SIMULATION AND IMPLEMENTATION DETAILS

The following input files are available in a zipped archive which can be found here:
[Normal Mode Langevin supplemental input files.](#)

Input files

Alanine dipeptide input files for SAMPLING, RATES, EFFICIENCY and KINETIC ENERGY.

alan.nm.conf	NML config file for sampling.
alan.nm2.conf	NML config file for rates.
alan.nm3.conf	NML config file for efficiency and kinetic energy.
alan.lang.conf	Langevin Impulse config file.

par_all27_prot_lipid.inp	par input file.
alan_mineq.psf	psf input file.
minC7eq.pdb	pdb input file.
eigVmC7eq	eigenvector file.

BPTI input files for SAMPLING and EFFICIENCY.

Files to generate the BPTI results in the paper.

bpti.nm.conf	NML config file for sampling.
bpti.nm1.conf	NML config file for efficiency.
bpti.lang.conf	Langevin Impulse config file.
par_all27_prot_lipid.inp	par input file.
cbpti.psf	psf input file.
cbpti.min.pdb	pdb input file.

Additional files to create eigVmBPTISCP eigenvector file.

bpti.hess.conf	Diagonalization config file.
bpti.finmin1.pdb	Minimized position file.

Analysis scripts

VMD and Matlab analysis scripts.

proc_dihedrals_alan.sh	Script for extracting dihedral data for alanine dipeptide.
dihedrals_alan.tcl	Dihedral definition file for proc_dihedrals_alan.sh.
proc_dihedrals_bpti.sh	Script for extracting dihedral data for BPTI.
dihedrals_BPTI.tcl	Dihedral definition file for proc_dihedrals_bpti.sh.
RamachandranFe.m	Ramachandran Matlab script for alanine dipeptide.
RamachandranAll.m	Ramachandran Matlab script for BPTI.
RamachandranAll43.m	Ramachandran Matlab script for BPTI, 43 of 57 residues considered.
MyColormaps.mat	Colormap for Ramachandran Matlab scripts.
GetTPRates.m	Rate calculation for alanine dipeptide.
TransitionPaths.m	Matlab script required by GetTPRates.m, finds all transition paths.
TheState.m	Matlab script required by GetTPRates.m, finds current molecule state

Linux X86 Protomol executable

protomol	Light version of Protomol 2.1.
----------	--------------------------------

Implementation of NML in ProtoMol

NML is implemented in the open source molecular dynamics application Protomol, which can be found at:

<http://protomol.sourceforge.net/>

List of source code files

NML propagator:	framework/integrators/NormModeInt.cpp/h
NML minimizer:	framework/integrators/NormModeMin.cpp/h
NML simple minimizer:	framework/integrators/NormModeSmplMin.cpp/h

List of configuration file options

NML propagator:	
cyclength	1 # Legacy MTS parameter, always 1

fixmodes	44	# Number of high frequency modes constrained
gamma	91	# Langevin Gamma
seed	1234	# Langevin random seed
temperature	300	# Langevin temperature
nve	0	# NVE simulation if not 0
Berendsen	0	# Berendsen tau in fs
fdof	0	# Fixed degrees of freedom

NML minimizer

timestep	16	# timestep for propagates modes (legacy MTS position)
minimlim	0.1	# Minimizer target PE difference kcal mole ⁻¹
forcePEcheck	true	# Force PE/calcForces check at end of loop, always set true
massweight	true	# mass weighted minimization, always se true.
randforce	true	# Add random force, always set true

NML simple minimizer

As NML minimizer options.

Protomol C2 switch

The 'C2' switch implemented in Protomol is a C^1 switch defined as

$$C^1(a_{ij}) = \begin{cases} 1 & a_{ij} \leq r_0, \\ \frac{(a_{ij}^2 - r_c^2)^2 (r_c^2 + 2a_{ij}^2 - 3r_0^2)}{(r_c^2 - r_0^2)^2} & r_0 \leq a_{ij} < r_c, \\ 0 & a_{ij} > r_c, \end{cases}$$

where $a_{ij} = \|x_j - x_i\|$, r_0 is the switch-on value and r_c is the cutoff value.