Normal mode splitting of Langevin dynamics

CR Sweet, P Petrone, VS Pande, JA Izaguirre
Notre Dame, Stanford

July 12, 2007

Acknowledgments

Thanks to John Chodera, Peter Minary and Bob Skeel for several suggestions.

Funding:
JAI and CRS: National Science Foundation DBI-0450067, CCF-0135195
VSP and PP: Simbios Center (NIH U54 GM072970)

JAI performed work during sabbatical visit at **Institute for Computational and Mathematical Engineering**, **Stanford**, hosted by **Eric Darve**.

Paul Brenner and Trevor Cickovski (Notre Dame) coded several aspects of the method and provided analysis scripts.

Acknowledgments

Thanks to John Chodera, Peter Minary and Bob Skeel for several suggestions.

Funding:
JAI and CRS: National Science Foundation DBI-0450067, CCF-0135195
VSP and PP: Simbios Center (NIH U54 GM072970)

JAI performed work during sabbatical visit at **Institute for Computational and Mathematical Engineering**, **Stanford**, hosted by **Eric Darve**.

Paul Brenner and Trevor Cickovski (Notre Dame) coded several aspects of the method and provided analysis scripts.

Acknowledgments

Thanks to John Chodera, Peter Minary and Bob Skeel for several suggestions.

Funding:
JAI and CRS: National Science Foundation DBI-0450067, CCF-0135195
VSP and PP: Simbios Center (NIH U54 GM072970)

JAI performed work during sabbatical visit at **Institute for Computational and Mathematical Engineering**, **Stanford**, hosted by **Eric Darve**.

Paul Brenner and Trevor Cickovski (Notre Dame) coded several aspects of the method and provided analysis scripts.

Acknowledgments

Thanks to John Chodera, Peter Minary and Bob Skeel for several suggestions.

Funding:
JAI and CRS: National Science Foundation DBI-0450067, CCF-0135195
VSP and PP: Simbios Center (NIH U54 GM072970)

JAI performed work during sabbatical visit at **Institute for Computational and Mathematical Engineering**, **Stanford**, hosted by **Eric Darve**.

Paul Brenner and Trevor Cickovski (Notre Dame) coded several aspects of the method and provided analysis scripts.

Overview

Produce reduced model (low frequency normal modes) for kinetics and sampling for implicitly solvated proteins.

Perform Langevin dynamics on reduced model to compute observables.

Simulations show that with 10%-30% of modes kinetics and sampling are recovered, with speedups of up to 30 for bovine pancreatic trypsin inhibitor (BPTI), an 882 atom protein.

Overview

Produce reduced model (low frequency normal modes) for kinetics and sampling for implicitly solvated proteins.

Perform Langevin dynamics on reduced model to compute observables.

Simulations show that with 10%-30% of modes kinetics and sampling are recovered, with speedups of up to 30 for bovine pancreatic trypsin inhibitor (BPTI), an 882 atom protein.

Overview

Produce reduced model (low frequency normal modes) for kinetics and sampling for implicitly solvated proteins.

Perform Langevin dynamics on reduced model to compute observables.

Simulations show that with 10%-30% of modes kinetics and sampling are recovered, with speedups of up to 30 for bovine pancreatic trypsin inhibitor (BPTI), an 882 atom protein.

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Reduced dynamical spaces

1. Normal modes (NMA) from diagonalization of Hessian
2. Principal components (PCA) from diagonalization of displacement covariance matrix.
3. Nonlinear coordinates

Comparison:

1. Levitt & Stern, *J. Mol. Biol.* **181:**423 (1985)
2. Amadei *et al.*, *PROTEINS* **17**:412 (1993)
3. Lange & Grübmuller, *J. Chem. Phys.* **124**:214903 (2006)

Reduced dynamical spaces

1. Normal modes (NMA) from diagonalization of Hessian
2. Principal components (PCA) from diagonalization of displacement covariance matrix.
3. Nonlinear coordinates

Comparison:

1. Levitt & Stern, *J. Mol. Biol.* **181:**423 (1985)
2. Amadei *et al.*, *PROTEINS* **17**:412 (1993)
3. Lange & Grübmuller, *J. Chem. Phys.* **124**:214903 (2006)

Reduced dynamical spaces

1. Normal modes (NMA) from diagonalization of Hessian
2. Principal components (PCA) from diagonalization of displacement covariance matrix.
3. Nonlinear coordinates

Comparison:

1. Levitt & Stern, *J. Mol. Biol.* **181:**423 (1985)
2. Amadei *et al.*, *PROTEINS* **17**:412 (1993)
3. Lange & Grübmuller, *J. Chem. Phys.* **124**:214903 (2006)

Reduced dynamical spaces

1. Normal modes (NMA) from diagonalization of Hessian
2. Principal components (PCA) from diagonalization of displacement covariance matrix.
3. Nonlinear coordinates

Comparison:

1. Levitt & Stern, *J. Mol. Biol.* **181:**423 (1985)
2. Amadei *et al.*, *PROTEINS* **17**:412 (1993)
3. Lange & Grübmuller, *J. Chem. Phys.* **124**:214903 (2006)

Reduced dynamical spaces

1. Normal modes (NMA) from diagonalization of Hessian
2. Principal components (PCA) from diagonalization of displacement covariance matrix.
3. Nonlinear coordinates

Comparison:

1. Levitt & Stern, *J. Mol. Biol.* **181:**423 (1985)
2. Amadei *et al.*, *PROTEINS* **17**:412 (1993)
3. Lange & Grübmuller, *J. Chem. Phys.* **124**:214903 (2006)

Reduced dynamical spaces

1. Normal modes (NMA) from diagonalization of Hessian
2. Principal components (PCA) from diagonalization of displacement covariance matrix.
3. Nonlinear coordinates

Comparison:

1. Levitt & Stern, *J. Mol. Biol.* **181:**423 (1985)
2. Amadei *et al.*, *PROTEINS* **17**:412 (1993)
3. Lange & Grübmuller, *J. Chem. Phys.* **124**:214903 (2006)

Langevin equation

The system will be propagated using the Langevin equation

$$\mathrm{d}\mathbf{x} = \mathbf{v}\mathrm{d}t,\ \mathbf{M}\mathrm{d}\mathbf{v} = \mathbf{f}\mathrm{d}t - \mathbf{\Gamma}\mathbf{M}\mathbf{v}\mathrm{d}t + \sqrt{2k_{\mathrm{B}}T}\,\mathbf{\Gamma}^{\frac{1}{2}}\mathbf{M}^{\frac{1}{2}}\mathrm{d}\mathbf{W}(t), \qquad (1)$$

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

$$\mathbf{\Gamma} = k_{\mathrm{B}}T\mathbf{D}^{-1}\mathbf{M}^{-1}. \qquad (2)$$

$\mathbf{D}$ 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.

Langevin equation

The system will be propagated using the Langevin equation

$$\mathrm{d}\mathbf{x} = \mathbf{v}\mathrm{d}t,\ \mathbf{M}\mathrm{d}\mathbf{v} = \mathbf{f}\mathrm{d}t - \mathbf{\Gamma}\mathbf{M}\mathbf{v}\mathrm{d}t + \sqrt{2k_{\mathrm{B}}T}\,\mathbf{\Gamma}^{\frac{1}{2}}\mathbf{M}^{\frac{1}{2}}\mathrm{d}\mathbf{W}(t), \qquad (1)$$

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

$$\mathbf{\Gamma} = k_{\mathrm{B}}T\mathbf{D}^{-1}\mathbf{M}^{-1}. \qquad (2)$$

$\mathbf{D}$ 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.

Normal mode splitting

$$\mathcal{H}(\mathbf{x}) = \frac{1}{2}\dot{\mathbf{x}}^{\mathsf{T}}\mathbf{M}\dot{\mathbf{x}} + U(\mathbf{x}), \tag{3}$$

for positions $\mathbf{x}$, potential energy U and diagonal mass matrix $\mathbf{M}$.

$$\mathbf{M}^{-\frac{1}{2}}\mathbf{U}_{xx}(\mathbf{x}_0)\mathbf{M}^{-\frac{1}{2}}\mathcal{Q} = \mathcal{Q}\mathbf{\Lambda}, \tag{4}$$

$\mathbf{\Lambda}$ is the diagonal matrix of ordered eigenvalues $\lambda_1, \cdots, \lambda_{3N}$ and $\mathcal{Q}$ the matrix of column eigenvectors $\mathbf{e}_1, \cdots, \mathbf{e}_{3N}$

$$\mathcal{Q} = [\mathbf{Q}, \bar{\mathbf{Q}}], \mathbf{Q} = [\mathbf{e}_1, \cdots, \mathbf{e}_i] \text{ and } \bar{\mathbf{Q}} = [\mathbf{e}_{i+1}, \cdots, \mathbf{e}_{3N}], \tag{5}$$

for rectangular $\mathbf{Q}$ and $\bar{\mathbf{Q}}$ which span spaces C and $C^{\perp}$.

Normal mode splitting

$$\mathcal{H}(\mathbf{x}) = \frac{1}{2}\dot{\mathbf{x}}^{\mathsf{T}}\mathbf{M}\dot{\mathbf{x}} + U(\mathbf{x}), \tag{3}$$

for positions $\mathbf{x}$, potential energy U and diagonal mass matrix $\mathbf{M}$.

$$\mathbf{M}^{-\frac{1}{2}}\mathbf{U}_{xx}(\mathbf{x}_0)\mathbf{M}^{-\frac{1}{2}}\mathcal{Q} = \mathcal{Q}\mathbf{\Lambda}, \tag{4}$$

$\mathbf{\Lambda}$ is the diagonal matrix of ordered eigenvalues $\lambda_1, \cdots, \lambda_{3N}$ and $\mathcal{Q}$ the matrix of column eigenvectors $\mathbf{e}_1, \cdots, \mathbf{e}_{3N}$

$$\mathcal{Q} = [\mathbf{Q}, \bar{\mathbf{Q}}], \mathbf{Q} = [\mathbf{e}_1, \cdots, \mathbf{e}_i] \text{ and } \bar{\mathbf{Q}} = [\mathbf{e}_{i+1}, \cdots, \mathbf{e}_{3N}], \tag{5}$$

for rectangular $\mathbf{Q}$ and $\bar{\mathbf{Q}}$ which span spaces C and $C^{\perp}$.

Normal mode splitting

$$\mathcal{H}(\mathbf{x}) = \frac{1}{2}\dot{\mathbf{x}}^{\mathsf{T}}\mathbf{M}\dot{\mathbf{x}} + U(\mathbf{x}), \tag{3}$$

for positions $\mathbf{x}$, potential energy U and diagonal mass matrix $\mathbf{M}$.

$$\mathbf{M}^{-\frac{1}{2}}\mathbf{U}_{xx}(\mathbf{x}_0)\mathbf{M}^{-\frac{1}{2}}\mathcal{Q} = \mathcal{Q}\mathbf{\Lambda}, \tag{4}$$

$\mathbf{\Lambda}$ is the diagonal matrix of ordered eigenvalues $\lambda_1, \cdots, \lambda_{3N}$ and $\mathcal{Q}$ the matrix of column eigenvectors $\mathbf{e}_1, \cdots, \mathbf{e}_{3N}$

$$\mathcal{Q} = [\mathbf{Q}, \bar{\mathbf{Q}}], \mathbf{Q} = [\mathbf{e}_1, \cdots, \mathbf{e}_i] \text{ and } \bar{\mathbf{Q}} = [\mathbf{e}_{i+1}, \cdots, \mathbf{e}_{3N}], \tag{5}$$

for rectangular $\mathbf{Q}$ and $\bar{\mathbf{Q}}$ which span spaces C and $C^{\perp}$.

Normal mode splitting

$$\mathcal{H}(\mathbf{x}) = \frac{1}{2}\dot{\mathbf{x}}^{\mathsf{T}}\mathbf{M}\dot{\mathbf{x}} + U(\mathbf{x}), \tag{3}$$

for positions $\mathbf{x}$, potential energy U and diagonal mass matrix $\mathbf{M}$.

$$\mathbf{M}^{-\frac{1}{2}}\mathbf{U}_{xx}(\mathbf{x}_0)\mathbf{M}^{-\frac{1}{2}}\mathcal{Q} = \mathcal{Q}\mathbf{\Lambda}, \tag{4}$$

$\mathbf{\Lambda}$ is the diagonal matrix of ordered eigenvalues $\lambda_1, \cdots, \lambda_{3N}$ and $\mathcal{Q}$ the matrix of column eigenvectors $\mathbf{e}_1, \cdots, \mathbf{e}_{3N}$

$$\mathcal{Q} = [\mathbf{Q}, \bar{\mathbf{Q}}], \mathbf{Q} = [\mathbf{e}_1, \cdots, \mathbf{e}_i] \text{ and } \bar{\mathbf{Q}} = [\mathbf{e}_{i+1}, \cdots, \mathbf{e}_{3N}], \tag{5}$$

for rectangular $\mathbf{Q}$ and $\bar{\mathbf{Q}}$ which span spaces C and $C^{\perp}$.

Splitting of Langevin equation

Physical Approximation 1. We will use the NM splitting to define a diffusion tensor that overdamps the high frequency modes in $C^{\perp}$ and propagates realistically the low frequency modes in C. The diffusion tensor is chosen to satisfy the fluctuation-dissipation theorem, and hence detailed balance. (cf. C. Gardiner, *Handbook of Stochastic Methods*, Springer Verlag, 2004).

Physical Approximation 2. The overdamped high frequency modes will be approximated using Brownian dynamics.

Splitting of Langevin equation

Physical Approximation 1. We will use the NM splitting to define a diffusion tensor that overdamps the high frequency modes in $C^{\perp}$ and propagates realistically the low frequency modes in C. The diffusion tensor is chosen to satisfy the fluctuation-dissipation theorem, and hence detailed balance. (cf. C. Gardiner, *Handbook of Stochastic Methods*, Springer Verlag, 2004).

Physical Approximation 2. The overdamped high frequency modes will be approximated using Brownian dynamics.

Splitting of Langevin equation

Physical Approximation 1. We will use the NM splitting to define a diffusion tensor that overdamps the high frequency modes in $C^{\perp}$ and propagates realistically the low frequency modes in C. The diffusion tensor is chosen to satisfy the fluctuation-dissipation theorem, and hence detailed balance. (cf. C. Gardiner, *Handbook of Stochastic Methods*, Springer Verlag, 2004).

Physical Approximation 2. The overdamped high frequency modes will be approximated using Brownian dynamics.

Partitioned Langevin equations

We select different (diagonal) damping matrices $\hat{\mathbf{\Gamma}}$ and $\bar{\mathbf{\Gamma}}$ for spaces C and $C^{\perp}$, respectively.

$$\mathbf{\Gamma} = \mathbf{P_f}\,\hat{\mathbf{\Gamma}} + \mathbf{P_f^{\perp}}\,\bar{\mathbf{\Gamma}}.$$

We can split the Langevin equation (1) into

$$\mathrm{d}\hat{\mathbf{x}} = \hat{\mathbf{v}}\mathrm{d}t,\ \mathbf{M}\mathrm{d}\hat{\mathbf{v}} = \hat{\mathbf{f}}\mathrm{d}t - \hat{\mathbf{\Gamma}}\mathbf{M}\hat{\mathbf{v}}\mathrm{d}t + \sqrt{2k_{\mathrm{B}}T}\,\hat{\mathbf{\Gamma}}^{\frac{1}{2}}\mathbf{P_f}\,\mathbf{M}^{\frac{1}{2}}\mathrm{d}\mathbf{W}_1(t), \tag{6}$$

$$\mathrm{d}\bar{\mathbf{x}} = \bar{\mathbf{\Gamma}}^{-1}\mathbf{M}^{-1}\bar{\mathbf{f}}\mathrm{d}t + \sqrt{2k_{\mathrm{B}}T}\,\bar{\mathbf{\Gamma}}^{-\frac{1}{2}}\mathbf{M}^{-1}\mathbf{P_f^{\perp}}\,\mathbf{M}^{\frac{1}{2}}\mathrm{d}\mathbf{W}_2(t), \tag{7}$$

where $\hat{\mathbf{x}} = \mathbf{P_x}\mathbf{x}$, $\hat{\mathbf{v}} = \mathbf{P_x}\,\mathbf{v}$ and $\hat{\mathbf{f}} = \mathbf{P_f}\,\mathbf{f}$ in C, $\bar{\mathbf{x}} = \mathbf{P_x^{\perp}}\,\mathbf{x}$, $\bar{\mathbf{v}} = \mathbf{P_v^{\perp}}\,\mathbf{v}$ and $\bar{\mathbf{f}} = \mathbf{P_f^{\perp}}\,\mathbf{f}$ in $C^{\perp}$.

Numerical approximation 1. The C, $C^{\perp}$ splitting at eigenvector i ensures that the frequencies found in C are bounded above by $\sqrt{\lambda_i}$.

Numerical approximation 2. The potential energy 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

$$U = U_{\text{fast}} + U_{\text{slow}},$$

then we propagate the Brownian dynamics equation (7) using the fast forces $\bar{\mathbf{f}} \approx -\nabla_{\bar{\mathbf{x}}} U_{\text{fast}}$.

Numerical approximation 3. Use a MTS splitting of the Euler-Maruyama method for Eqn. (7) where, from Physical Approximation 2, $\bar{\mathbf{\Gamma}}$ is a matrix with diagonal values $\bar{\gamma}$

$$\bar{\mathbf{x}}^{n+\frac{i}{p+1}} = \bar{\mathbf{x}}^{n+\frac{i-1}{p+1}} + \Delta\tau\,\bar{\gamma}^{-1}\mathbf{M}^{-1}\bar{\mathbf{f}}^{n+\frac{i-1}{p+1}}, \quad i = 1, \cdots, p, \tag{8}$$

$$\bar{\mathbf{x}}^{n+1} = \bar{\mathbf{x}}^{n+\frac{p}{p+1}} + \sqrt{2\Delta t\,\bar{\gamma}^{-1}k_{\text{B}}T}\,\mathbf{M}^{-1}\mathbf{P}_{\mathbf{f}}^{\perp}\,\mathbf{M}^{\frac{1}{2}}\mathbf{z}^{n+1}, \tag{9}$$

for Gaussian random variable $\mathbf{z} = \{z_1, \cdots, z_{3N}\}$, $z_i \sim N(0,1)$. Eqn. (8) is repeated p times such that $p\Delta\tau = \Delta t$.

Numerical approximation 1. The C, $C^{\perp}$ splitting at eigenvector i ensures that the frequencies found in C are bounded above by $\sqrt{\lambda_i}$.

Numerical approximation 2. The potential energy 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

$$U = U_{\text{fast}} + U_{\text{slow}},$$

then we propagate the Brownian dynamics equation (7) using the fast forces $\bar{\mathbf{f}} \approx -\nabla_{\bar{\mathbf{x}}} U_{\text{fast}}$.

Numerical approximation 3. Use a MTS splitting of the Euler-Maruyama method for Eqn. (7) where, from Physical Approximation 2, $\bar{\mathbf{\Gamma}}$ is a matrix with diagonal values $\bar{\gamma}$

$$\bar{\mathbf{x}}^{n+\frac{i}{p+1}} = \bar{\mathbf{x}}^{n+\frac{i-1}{p+1}} + \Delta\tau\,\bar{\gamma}^{-1}\mathbf{M}^{-1}\bar{\mathbf{f}}^{n+\frac{i-1}{p+1}}, \quad i = 1, \cdots, p, \tag{8}$$

$$\bar{\mathbf{x}}^{n+1} = \bar{\mathbf{x}}^{n+\frac{p}{p+1}} + \sqrt{2\Delta t\,\bar{\gamma}^{-1}k_{\text{B}}T}\,\mathbf{M}^{-1}\mathbf{P}_{\mathbf{f}}^{\perp}\,\mathbf{M}^{\frac{1}{2}}\mathbf{z}^{n+1}, \tag{9}$$

for Gaussian random variable $\mathbf{z} = \{z_1, \cdots, z_{3N}\}$, $z_i \sim N(0,1)$. Eqn. (8) is repeated p times such that $p\Delta\tau = \Delta t$.

Numerical approximation 1. The C, $C^{\perp}$ splitting at eigenvector i ensures that the frequencies found in C are bounded above by $\sqrt{\lambda_i}$.

Numerical approximation 2. The potential energy 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

$$U = U_{\text{fast}} + U_{\text{slow}},$$

then we propagate the Brownian dynamics equation (7) using the fast forces $\bar{\mathbf{f}} \approx -\nabla_{\bar{\mathbf{x}}} U_{\text{fast}}$.

Numerical approximation 3. Use a MTS splitting of the Euler-Maruyama method for Eqn. (7) where, from Physical Approximation 2, $\bar{\boldsymbol{\Gamma}}$ is a matrix with diagonal values $\bar{\gamma}$

$$\bar{\mathbf{x}}^{n+\frac{i}{p+1}} = \bar{\mathbf{x}}^{n+\frac{i-1}{p+1}} + \Delta\tau\,\bar{\gamma}^{-1}\mathbf{M}^{-1}\bar{\mathbf{f}}^{n+\frac{i-1}{p+1}}, \quad i = 1, \cdots, p, \tag{8}$$

$$\bar{\mathbf{x}}^{n+1} = \bar{\mathbf{x}}^{n+\frac{p}{p+1}} + \sqrt{2\Delta t\,\bar{\gamma}^{-1}k_{\text{B}}T}\,\mathbf{M}^{-1}\mathbf{P}_{\mathbf{f}}^{\perp}\,\mathbf{M}^{\frac{1}{2}}\mathbf{z}^{n+1}, \tag{9}$$

for Gaussian random variable $\mathbf{z} = \{z_1, \cdots, z_{3N}\}$, $z_i \sim N(0,1)$. Eqn. (8) is repeated p times such that $p\Delta\tau = \Delta t$.

Acceleration of BD using Minimization

Acceleration of BD using Minimization

Acceleration of BD using Minimization

Langevin Impulse (LÎ) discretization

A half step using LÎ, cf. Skeel & Izaguirre, *Mol. Phys.* **100:**3885, 2002

half kick:

$$\mathbf{v}^{n+\epsilon} = \mathbf{v}^{n} - \frac{\Delta t}{2}\mathbf{M}^{-1}\mathbf{P}_{\mathbf{f}}\nabla_{\mathbf{x}}U(\mathbf{x}^{n}),$$

fluctuation-slow:

$$\mathbf{v}^{n+\frac{1}{2}} = \exp(-\gamma\Delta t)\mathbf{v}^{n+\epsilon} + \sqrt{\tau_2}\sqrt{2\gamma k_{\mathrm{B}}T}\mathbf{M}^{-1}\mathbf{P}_{\mathbf{f}}\,\mathbf{M}^{\frac{1}{2}}\mathbf{z}_1^{n+1},$$

$$\mathbf{x}^{n+\epsilon} = \mathbf{x}^{n} + \mathbf{P}_{\mathbf{x}}\left(\tau_1\mathbf{v}^{n+\epsilon} + \frac{1}{\gamma}\sqrt{2\gamma k_{\mathrm{B}}T}\mathbf{M}^{-1}\mathbf{P}_{\mathbf{f}}\,\mathbf{M}^{\frac{1}{2}}\boldsymbol{\zeta}^{n+1}\right),$$

fluctuation-fast:

$$\tilde{\mathbf{x}}^{k} = \left[G\left(\tilde{\mathbf{x}}^{k-1}\right)\right]^{k}, \text{ with } \tilde{\mathbf{x}}^{0} = \mathbf{x}^{n+\epsilon},$$

$$\mathbf{x}^{n+1} = \tilde{\mathbf{x}}^{k} + \sqrt{2\Delta t\bar{\gamma}^{-1}k_{\mathrm{B}}T}\,\mathbf{M}^{-1}\mathbf{P}_{\mathbf{f}}^{\perp}\,\mathbf{M}^{\frac{1}{2}}\mathbf{z}_2^{n+1}$$

Discretization of split Langevin equations (cont.)

Where γ and $\bar{\gamma}$ are the scalar Langevin damping coefficients for C and $C^{\perp}$ subspaces respectively. Constants are given by

$$\tau_1 = \frac{1-\exp(-\gamma\Delta t)}{\gamma}, \qquad \tau_2 = \frac{1-\exp(-2\gamma\Delta t)}{2\gamma}$$

$$\tau_3 = \frac{\tau_1-\tau_2}{\sqrt{\tau_2}}, \qquad \tau_4 = \sqrt{\Delta t - \frac{\tau_1^2}{\tau_2}}$$

$$\boldsymbol{\zeta}^{n+1} = \tau_3\,\mathbf{z}_1^{n+1} + \tau_4\,\mathbf{z}_2^{n+1}$$

The minimizer G is equivalent to the overdamped MTS integrator for

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

Simulation protocols

Test systems
Blocked alanine dipeptide (AD, 22 atoms)
Bovine pancreatic trypsin inhibitor (BPTI, 882 atoms)
Both using CHARMM 22 force field. Implemented using ProtoMol

Full Langevin simulations
Langevin Impulse, $\gamma = 91$ ps^{-1}, Time step: 1 fs
C^2 cutoffs for nonbonded forces:
AD (on at 9 Å, off at 12 Å; minimizer on at 4.5 Å, off at 5.5 Å
BPTI (on at 15 Å, off at 20 Å; minimizer on at 5.5 Å, off at 6.5 Å)

Simulation protocols

Test systems
Blocked alanine dipeptide (AD, 22 atoms)
Bovine pancreatic trypsin inhibitor (BPTI, 882 atoms)
Both using CHARMM 22 force field. Implemented using ProtoMol

Full Langevin simulations
Langevin Impulse, $\gamma = 91$ ps^{-1}, Time step: 1 fs
C^2 cutoffs for nonbonded forces:
AD (on at 9 Å, off at 12 Å; minimizer on at 4.5 Å, off at 5.5 Å
BPTI (on at 15 Å, off at 20 Å; minimizer on at 5.5 Å, off at 6.5 Å)

Simulation protocol continued

Normal-mode Langevin simulations
Hessian averaged over 500 fs NVE MD using leapfrog. Diagonalized using Lapack
Damping stage: potential energy differences of 0.1 kcal mol^{-1} within 2-5 iterations for AD and 2-8 iterations for BPTI.

Implicit Solvent Model
Screened Coulomb potential of Shen & Freed, *Biophys. J.* **82:**1791 (2002)

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

For these simulations we used bulk dielectric $D = 78$ and $S = 0.3$.
More sophisticated implicit solvent models can be used (e.g., GB, or Poisson-Boltzmann, etc.).

Simulation protocol continued

Normal-mode Langevin simulations
Hessian averaged over 500 fs NVE MD using leapfrog. Diagonalized using Lapack
Damping stage: potential energy differences of 0.1 kcal mol^{-1} within 2-5 iterations for AD and 2-8 iterations for BPTI.

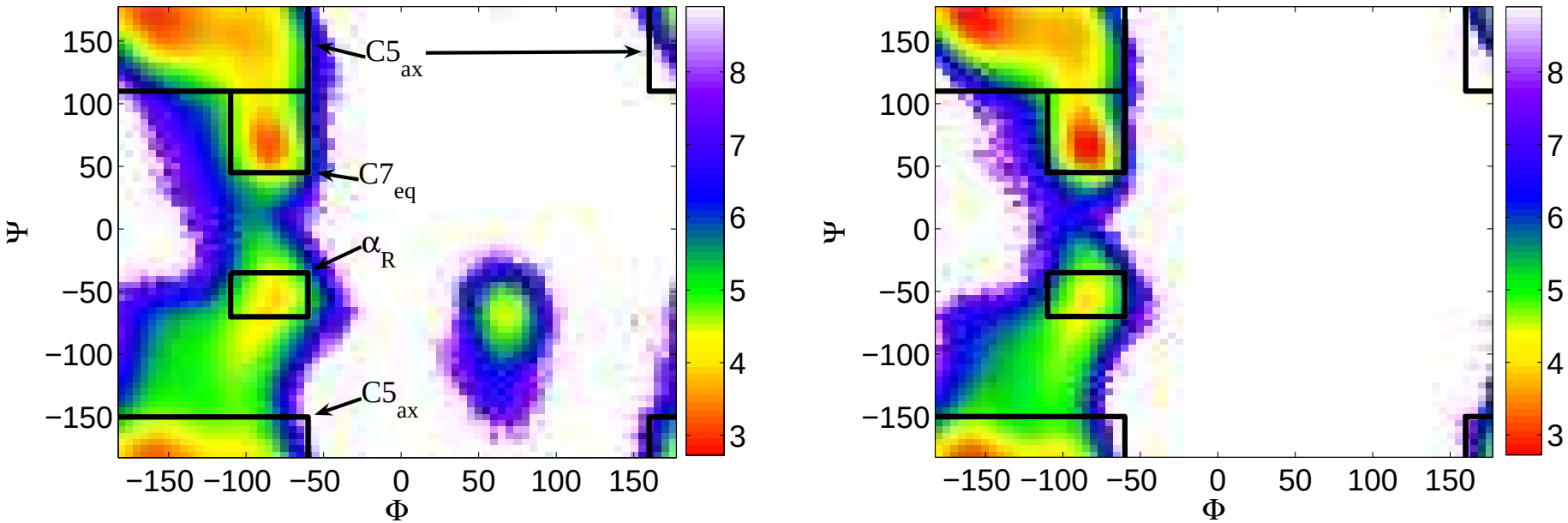
Implicit Solvent Model
Screened Coulomb potential of Shen & Freed, *Biophys. J.* **82:**1791 (2002)

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

For these simulations we used bulk dielectric $D = 78$ and $S = 0.3$.
More sophisticated implicit solvent models can be used (e.g., GB, or Poisson-Boltzmann, etc.).

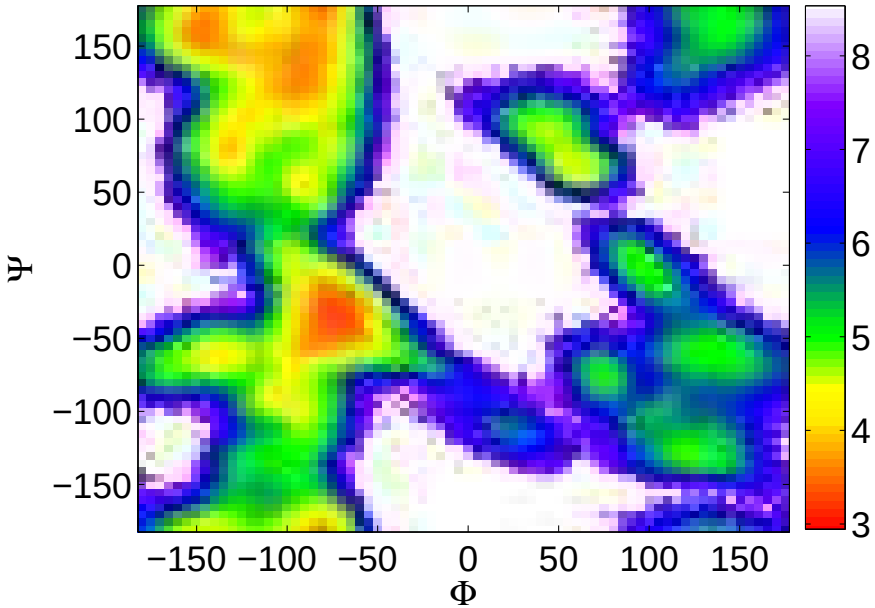
Sampling results for AD

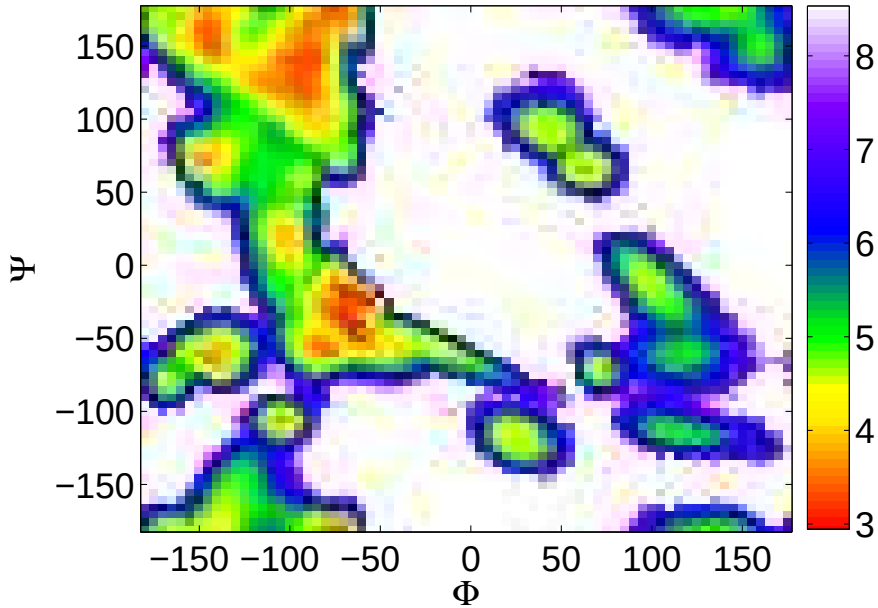
Comparison of alanine dipeptide Ramachandran free energy plots for (a) Langevin Impulse $\Delta t = 1\,\text{fs}$ and (b) NML with 22 modes (of 66) propagated $\Delta t = 5.5\,\text{fs}$. Data obtained from 500 ns trajectories with $\gamma = 91\,\text{ps}^{-1}$. The color-bar represents the free energy 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 Method RMSD of 0.17 kcal mol^{-1}



Sampling results for BPTI

Comparison of BPTI Ramachandran free energy plots for (a) Langevin Impulse $\Delta t = 1\,\text{fs}$, $\gamma = 91\,\text{ps}^{-1}$ and (b) NML with 300 modes propagated (of 2646) with $\Delta t = 20\,\text{fs}$ and $\gamma = 30\,\text{ps}^{-1}$. Data for all 55 backbone dihedral $\Phi - \Psi$ pairs from 3 ns trajectories. The color-bar represents the free energy in kcal mol^{-1} and the RMSD free energy difference between the Figures is 0.2 kcal mol^{-1}





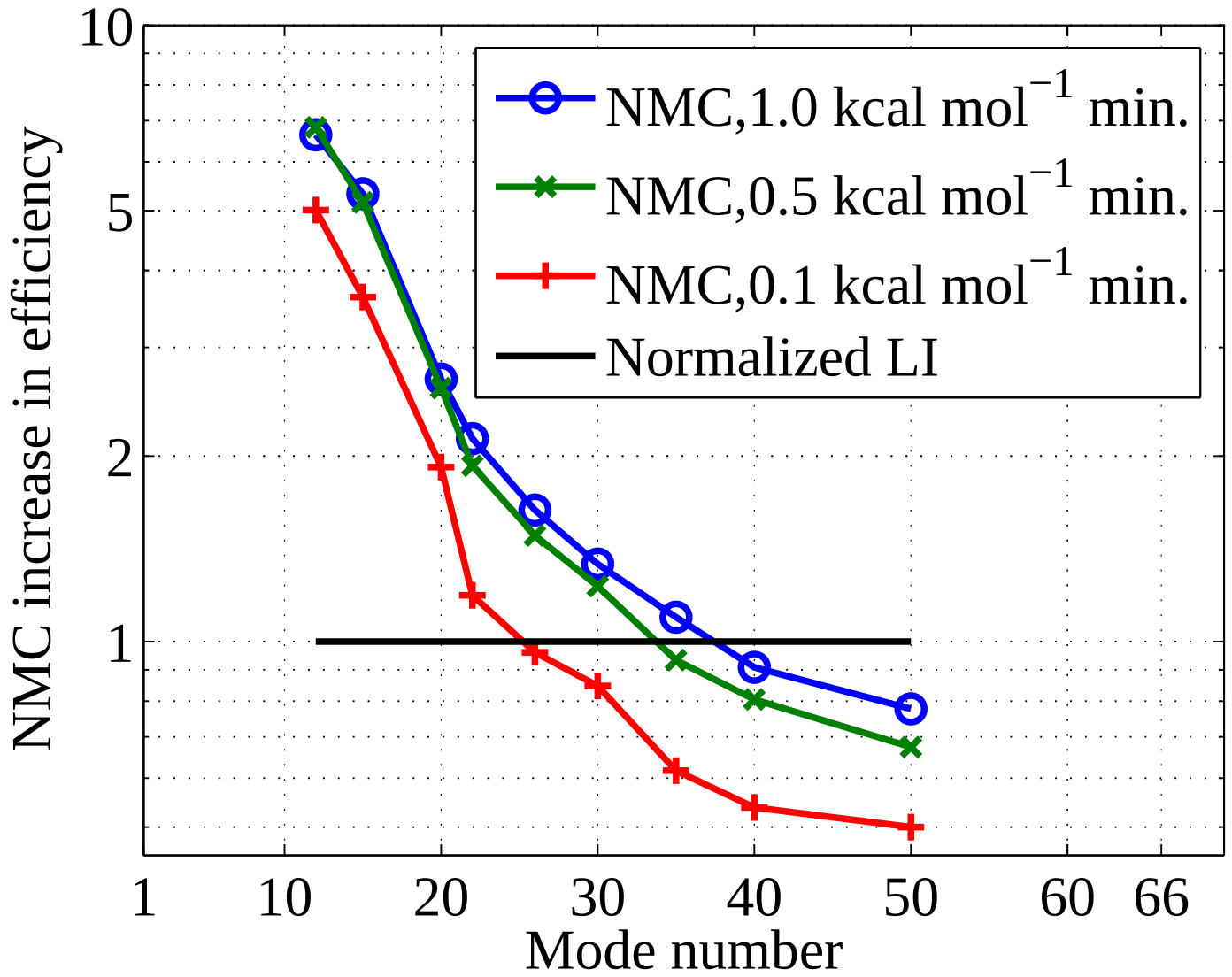
Kinetic results for AD

We compute isomerization rates between the $C7$ equatorial and α_R wells, from a 100 ns simulation from

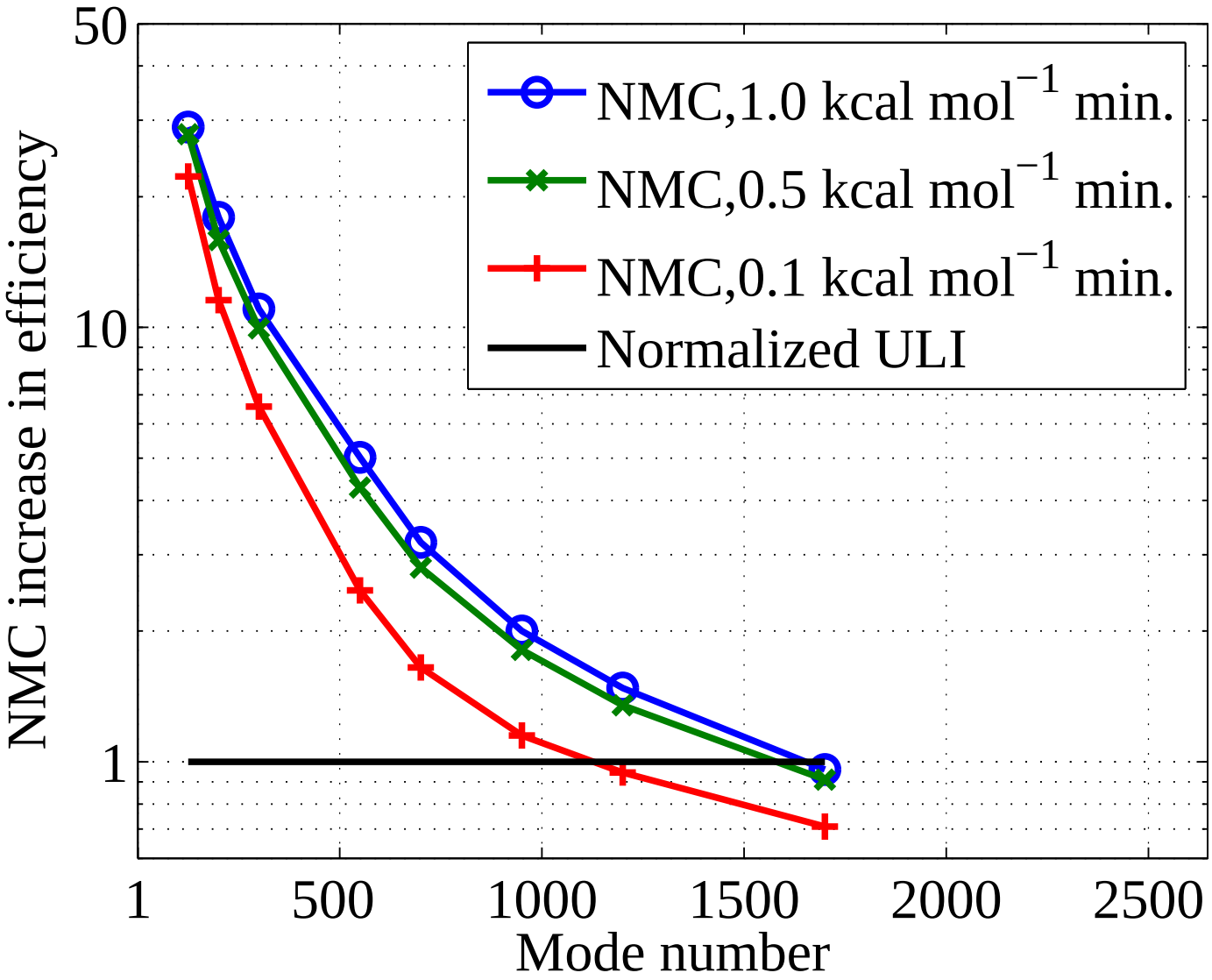
$$2c_A k_{AB} = \frac{P_{TP}}{\langle t_{TP} \rangle}, \tag{11}$$

where c_A is mole fraction of A (0.5). Formula: Best & Hummer, *PNAS* **102**:6732 (2005).

Efficiency results for AD



Efficiency results for BPTI



Discussion

Our splitting of the Langevin equation for biomolecules produces a method that overdamps high frequency modes and propagates correctly low frequency modes.

CPU time speedups in excess of an order of magnitude for 882 atom protein for sampling; speedup of 2 for 22 atom AD for kinetics and sampling.

Rediagonalization is our next major area of study. Improved over other methods though.

Possible optimizations include Krylov subspace diagonalization, improved minimizers, and enhanced integrators for Langevin dynamics with non-diagonal tensors.

Discussion

Our splitting of the Langevin equation for biomolecules produces a method that overdamps high frequency modes and propagates correctly low frequency modes.

CPU time speedups in excess of an order of magnitude for 882 atom protein for sampling; speedup of 2 for 22 atom AD for kinetics and sampling.

Rediagonalization is our next major area of study. Improved over other methods though.

Possible optimizations include Krylov subspace diagonalization, improved minimizers, and enhanced integrators for Langevin dynamics with non-diagonal tensors.

Discussion

Our splitting of the Langevin equation for biomolecules produces a method that overdamps high frequency modes and propagates correctly low frequency modes.

CPU time speedups in excess of an order of magnitude for 882 atom protein for sampling; speedup of 2 for 22 atom AD for kinetics and sampling.

Rediagonalization is our next major area of study. Improved over other methods though.

Possible optimizations include Krylov subspace diagonalization, improved minimizers, and enhanced integrators for Langevin dynamics with non-diagonal tensors.

Discussion

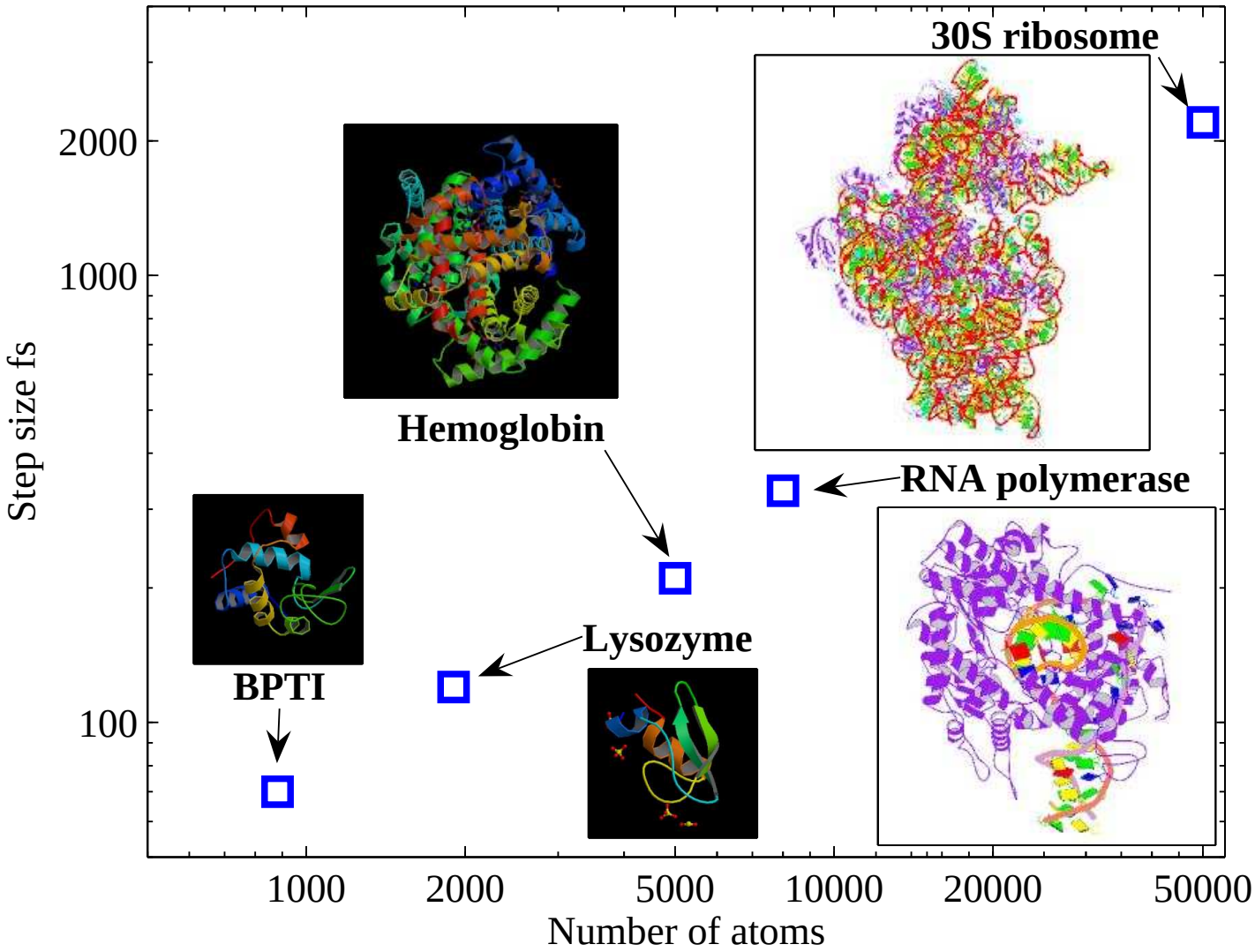
Our splitting of the Langevin equation for biomolecules produces a method that overdamps high frequency modes and propagates correctly low frequency modes.

CPU time speedups in excess of an order of magnitude for 882 atom protein for sampling; speedup of 2 for 22 atom AD for kinetics and sampling.

Rediagonalization is our next major area of study. Improved over other methods though.

Possible optimizations include Krylov subspace diagonalization, improved minimizers, and enhanced integrators for Langevin dynamics with non-diagonal tensors.

Maximum theoretical speedups with system size



How others deal with fast motions

1. Langevin/Implicit/Normal-mode (LIN): Analytical propagation of high freq. NM
2. Amplified Collective Motions (ACM): Excite low ANM frequencies.
3. Subspace Molecular Dynamics (SMD): Constrain high NM frequencies.
4. Multiple time step Diffusive Langevin: Constrain fast d.o.f. - then relax them.
5. Collective Langevin Motions: Construct generalized Langevin equation from PCA or nonlinear coordinate.

1. Zhang & Schlick, *J. Comp. Chem.* **14:**1212 (1993)
2. Zhang *et al.*, *Biophys. J.* **84**:3583 (2003)
3. Space *et al.*, *J. Chem. Phys.* **99:**9070 (1993)
4. Eastman & Doniach, *Proteins* **30**:215 (1998)
5. Lange & Grübmuller, *J. Chem. Phys.* **124:**2149 (2006)

How others deal with fast motions

1. Langevin/Implicit/Normal-mode (LIN): Analytical propagation of high freq. NM
2. Amplified Collective Motions (ACM): Excite low ANM frequencies.
3. Subspace Molecular Dynamics (SMD): Constrain high NM frequencies.
4. Multiple time step Diffusive Langevin: Constrain fast d.o.f. - then relax them.
5. Collective Langevin Motions: Construct generalized Langevin equation from PCA or nonlinear coordinate.

1. Zhang & Schlick, *J. Comp. Chem.* **14:**1212 (1993)
2. Zhang *et al.*, *Biophys. J.* **84**:3583 (2003)
3. Space *et al.*, *J. Chem. Phys.* **99:**9070 (1993)
4. Eastman & Doniach, *Proteins* **30**:215 (1998)
5. Lange & Grübmuller, *J. Chem. Phys.* **124:**2149 (2006)

How others deal with fast motions

1. Langevin/Implicit/Normal-mode (LIN): Analytical propagation of high freq. NM
2. Amplified Collective Motions (ACM): Excite low ANM frequencies.
3. Subspace Molecular Dynamics (SMD): Constrain high NM frequencies.
4. Multiple time step Diffusive Langevin: Constrain fast d.o.f. - then relax them.
5. Collective Langevin Motions: Construct generalized Langevin equation from PCA or nonlinear coordinate.

1. Zhang & Schlick, *J. Comp. Chem.* **14:**1212 (1993)
2. Zhang *et al.*, *Biophys. J.* **84**:3583 (2003)
3. Space *et al.*, *J. Chem. Phys.* **99:**9070 (1993)
4. Eastman & Doniach, *Proteins* **30**:215 (1998)
5. Lange & Grübmuller, *J. Chem. Phys.* **124:**2149 (2006)

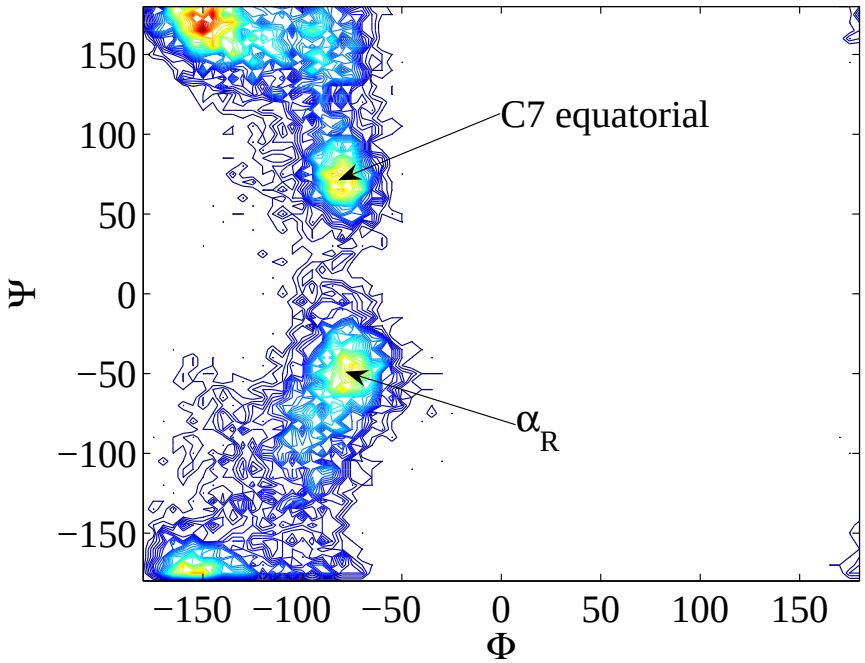
How others deal with fast motions

1. Langevin/Implicit/Normal-mode (LIN): Analytical propagation of high freq. NM
2. Amplified Collective Motions (ACM): Excite low ANM frequencies.
3. Subspace Molecular Dynamics (SMD): Constrain high NM frequencies.
4. Multiple time step Diffusive Langevin: Constrain fast d.o.f. - then relax them.
5. Collective Langevin Motions: Construct generalized Langevin equation from PCA or nonlinear coordinate.

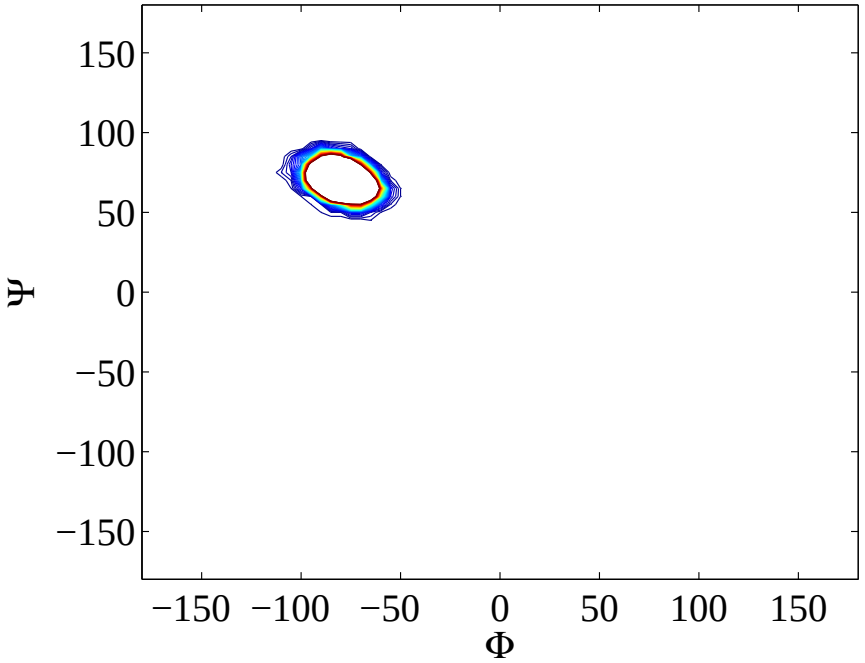
1. Zhang & Schlick, *J. Comp. Chem.* **14:**1212 (1993)
2. Zhang *et al.*, *Biophys. J.* **84**:3583 (2003)
3. Space *et al.*, *J. Chem. Phys.* **99:**9070 (1993)
4. Eastman & Doniach, *Proteins* **30**:215 (1998)
5. Lange & Grübmuller, *J. Chem. Phys.* **124:**2149 (2006)

Effect of constraining in Subspace MD

Implicitly solvated blocked alanine dipeptide



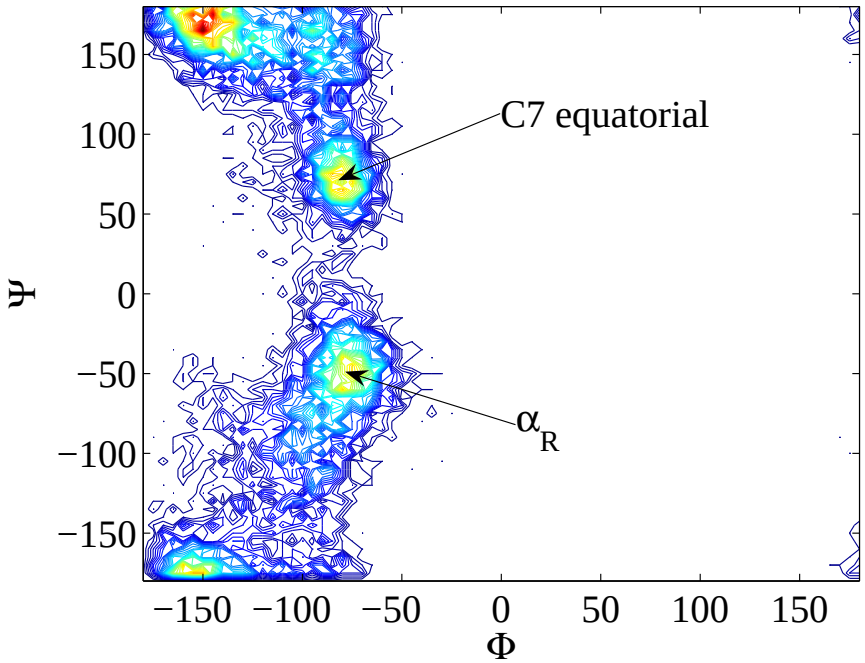
Unconstrained Langevin dynamics



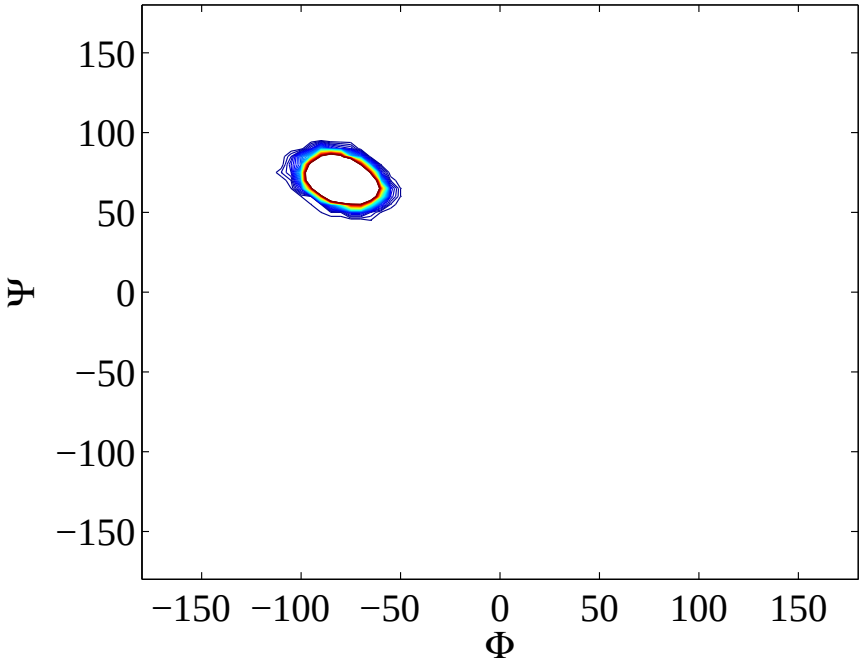
SMD. 20/66 modes constrained

Effect of constraining in Subspace MD

Implicitly solvated blocked alanine dipeptide



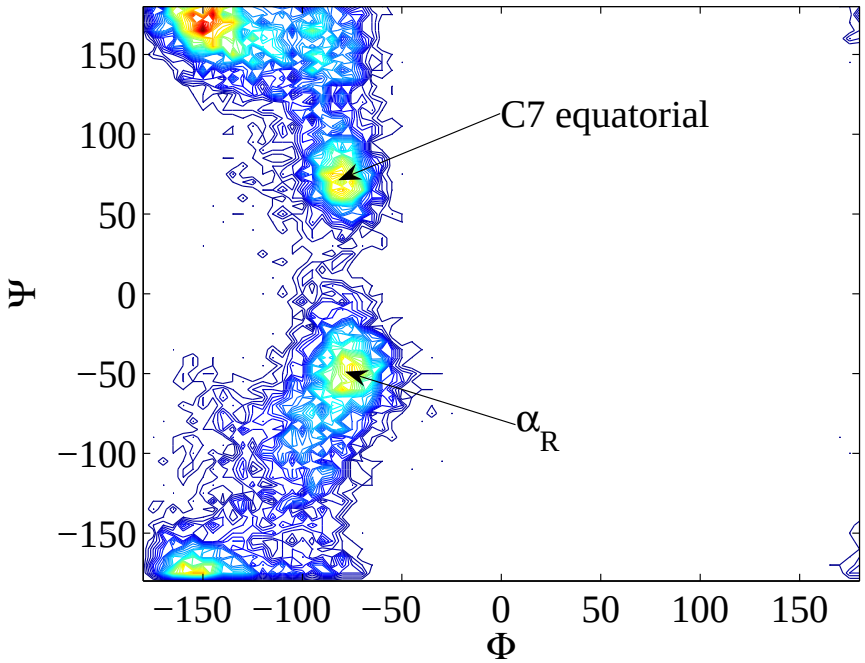
Unconstrained Langevin dynamics



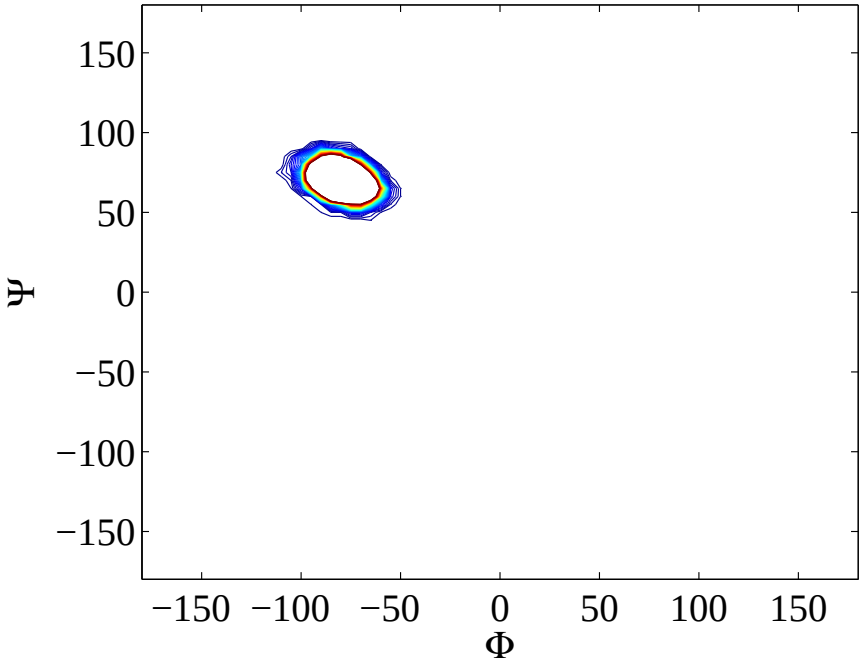
SMD. 20/66 modes constrained

Effect of constraining in Subspace MD

Implicitly solvated blocked alanine dipeptide



Unconstrained Langevin dynamics



SMD. 20/66 modes constrained