

5. Algorithmic issues and numerical examples

Based on the considerations of the Sections 2.4, 3, and 4.2 we shall qualitatively study some of the introduced reduction schemes for two model systems: Ryckaert-Bellemans n -butane and the glycine dipeptide analogue with a GROMOS96 vacuum force field. The latter is a small peptide that contains a central amino acid and which is a popular benchmark system for spectroscopy and conformation dynamics [282, 283]. On the other hand Ryckaert-Bellemans' n -butane molecule is particularly convenient for our purposes, since many properties are known on analytical grounds (e.g., reaction coordinate, torsional free energy).

We have argued (and this is confirmed by numerical simulations) that hybrid Monte-Carlo (HMC) generates a diffusion-like flow, and, in point of fact, we can draw similar conclusion from HMC simulations or simulations of the corresponding Smoluchowski equation. Langevin dynamics is not used in the model studies. The reason is that (i) the dynamics of a Langevin equation can be vastly different for different friction and noise coefficients, and (ii) there is an ongoing discussion about the choice of good integrators that preserve the thermodynamical properties of the systems (e.g., temperature, equipartition of energy, invariant measure). We refer to the monograph [284] and the references therein.

5.1. The constrained hybrid Monte-Carlo algorithm

For the evaluation of constrained and conditional expectation values we confine our attention to the hybrid Monte-Carlo scheme as has been introduced in Section 4.2.2. The reason is twofold: first of all, all quantities of interest can be computed as positional averages, and, secondly, we know for sure that the constrained expectations eventually converge to the correct values (strong law of large numbers).

To this end we briefly explain how the constrained hybrid Monte-Carlo algorithm actually works. It is convenient to switch back to a representation of the equations of motion and the invariant measure in terms of the ambient space variable (q, p) . We shall also drop the mass scaling assumption. Given a symmetric, positive-definite molecular mass matrix $M \in \mathbf{R}^{n \times n}$, and an interaction potential $V : \mathbf{R}^n \rightarrow \mathbf{R}$ that is bounded from below, the unconstrained Lagrangian is defined as

$$L(q, v) = \frac{1}{2} \langle Mv, v \rangle - V(q).$$

The respective unconstrained Hamiltonian thus reads

$$H(q, p) = \frac{1}{2} \langle M^{-1}p, p \rangle + V(q).$$

Introducing the reaction coordinate constraint $\Phi(q) = \xi$, the constrained equations of motion (4.1) are then generated by the augmented Lagrangian $\hat{L} = L - \lambda^i (\Phi_i(q) - \xi_i)$. The SHAKE discretization of the equations of motion for a time step $h > 0$ is

$$\begin{aligned} q_{n+1} - 2q_n + q_{n-1} &= -h^2 M^{-1} (\nabla V(q_n) + \mathbf{D}\Phi(q_n)\lambda_n) \\ \xi &= \Phi(q_{n+1}). \end{aligned} \tag{5.1}$$

In the classical formulation of Ryckaert *et al.*, the momentum is approximated by [260]

$$p_n = M \left(\frac{q_{n+1} - q_{n-1}}{h} \right). \tag{5.2}$$

This method has two major drawbacks: First of all, the mapping $(q_n, p_n) \mapsto (q_{n+1}, p_{n+1})$ defined by (5.1)–(5.2) is not symplectic.²² Secondly, the three-term recursion in (5.1) may lead to an accumulation of round-off errors. Therefore the scheme may become unstable, as has been pointed out in [105]. A remedy of both problems is to make the SHAKE algorithm a variational integrator: following [258] we replace (5.2) by the correct discrete conjugate momentum (4.18). This amounts to formulating SHAKE as a one-step method which yields the RATTLE algorithm [261]

$$\begin{aligned}
p_{n+1/2} &= p_n - \frac{h}{2} (\nabla V(q_n) + \mathbf{D}\Phi(q_n)\lambda_n) \\
q_{n+1} &= q_n + hM^{-1}p_{n+1} \\
\xi &= \Phi(q_{n+1}) \\
p_{n+1} &= p_{n+1/2} - \frac{h}{2} (\nabla V(q_{n+1}) + \mathbf{D}\Phi(q_{n+1})\mu_n) \\
0 &= \mathbf{D}\Phi(q_{n+1})M^{-1}p_{n+1},
\end{aligned} \tag{5.3}$$

The Lagrange multipliers λ_n, μ_n are chosen, such that the two constraints are satisfied. The RATTLE integrator (or SHAKE considered as a mapping $T^*\Sigma \rightarrow T^*\Sigma$) is symplectic as following from its variational nature; cf. the related articles [246, 82].²³

Recall that hybrid Monte-Carlo (HMC) requires that we draw a initial momentum from the constrained Maxwell distribution at each Monte-Carlo step, where the constrained Maxwellian depends point-wise on the constrained position variables. This can be understood as follows: consider the unconstrained expression for the kinetic energy in terms of the velocity variables,

$$T(v) = \frac{1}{2} \langle Mv, v \rangle := \frac{1}{2} \langle v, v \rangle_M,$$

where $\langle \cdot, \cdot \rangle_M$ denotes the metric with respect to the mass matrix. We have shown in Section 4.1.3 that the constrained canonical probability distribution is simply the restriction of the unconstrained distribution. In order to restrict the Maxwell density to the constrained tangent space $T_q\Sigma$, $q \in \Sigma$, we define the M -orthogonal projection $P_{M,T} : T_q\mathbf{R}^n \rightarrow T_q\Sigma$ that is defined point-wise for each $q \in \Sigma$:

$$P_{M,T} = \mathbf{1} - M^{-1}J_\Phi(J_\Phi^T M^{-1}J_\Phi)^{-1}J_\Phi^T, \quad J_\Phi = \mathbf{D}\Phi(q).$$

Strictly speaking, $P_{M,T}$ sends vectors $v \in \mathbf{R}^n$ to vectors in $\tilde{v} \in \mathbf{R}^n$, such that \tilde{v} satisfies the hidden constraint $\mathbf{D}\Phi \cdot \tilde{v} = 0$. It can be readily checked that (i) the matrix $P_{M,T}$ meets the idempotency property $P_{M,T}^2 = P_{M,T}$, and that (ii) it is symmetric with respect to the mass-weighted scalar product $\langle \cdot, \cdot \rangle_M$. That is,

$$\langle P_{M,T}u, v \rangle_M = \langle u, P_{M,T}v \rangle_M$$

for any two vectors $u, v \in \mathbf{R}^n$. Hence $P_{M,T}$ is an orthogonal projection with respect to the metric defined by the mass matrix M . Consequently, we shall refer to $P_{M,T}$ as M -orthogonal projection. Since $P_{M,T}$ maps to the constrained velocity space, we obtain the restricted Maxwell density $\exp(-\beta T_\Sigma)$ by restricting the kinetic energy,

$$T_\Sigma(q, v) := T(P_{M,T}v) = \frac{1}{2} \langle P_{M,T}v, v \rangle_M.$$

²²The mapping preserves the wedge product though. However the thus defined flow is not a map $T^*\Sigma \rightarrow T^*\Sigma$, since the momenta do not satisfy the hidden constraint $\mathbf{D}\Phi^T M^{-1}p = 0$.

²³A convenient numerical scheme for solving the nonlinear constraint equation $\Phi(q_{n+1}) = \xi$ is provided by original SHAKE iteration; see [260]. The SHAKE iteration can be considered a nonlinear one-step Gauss-Seidel-Newton iteration as has been argued in [267]. It is incredibly stable at rather large step-size, e.g., 5fs with a torsion angle constraint and the Ryckaert-Bellemans force field.

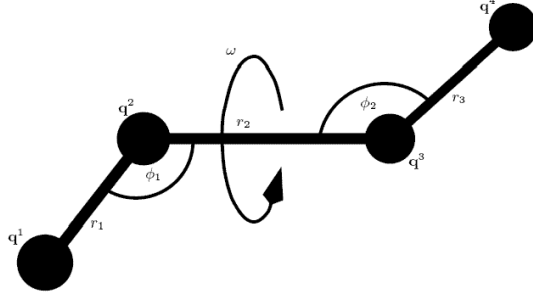


Figure 18. Ryckaert-Bellemans united-atoms butane molecule [91].

Defining $K(p) = T(M^{-1}p)$, the phase space analogue of T_Σ is found to be

$$K_\Sigma(q, p) := \frac{1}{2} \langle P_{M,T}^* p, p \rangle_{M^{-1}}, \quad P_{M,T}^* = M P_{M,T} M^{-1}.$$

It is easy to see that $P_{M,T}^*$ is idempotent and symmetric with respect to $\langle \cdot, \cdot \rangle_{M^{-1}}$. Hence $P_{M,T}^*$ is the M^{-1} -orthogonal projection onto the constrained momentum space $T_q^* \Sigma$. In other words, $P_{M,T}^*$ sends $p \in \mathbf{R}^n$ to $\tilde{p} \in \mathbf{R}^n$, such that \tilde{p} satisfies the hidden constraint $\mathbf{D}\Phi^T M^{-1} \tilde{p} = 0$. Omitting normalization, the constrained Maxwellian is

$$\varrho_\Sigma(q, p) \propto \exp(-\beta K_\Sigma(q, p)) \quad (5.4)$$

which is exactly the ambient space analogue of the constrained density (4.8) in local coordinates. The easiest way to draw momenta from the constrained distribution (5.4) is to generate a vector p due to the unconstrained distribution $\exp(-\beta K(p))$, and then apply the projection onto the constrained cotangent plane $T_q^* \Sigma$, $q \in \Sigma$. This then yields a vector $\tilde{p} = P_{M,T}^* p$ that is properly distributed according to ϱ_Σ . In this way the projection maintains the full dimensionality for the HMC algorithm, and we can completely work in the ambient space coordinates q and p .

The algorithm We summarize the considerations from Section 4.2.2 and the last few paragraphs. Given an initial position q_0 that satisfy the constraint $\Phi(q_0) = \xi$, the constrained hybrid Monte-Carlo algorithm proceeds as follows.

- (i) Draw a random vector due to the unconstrained momentum distribution

$$p \sim \exp(-\beta K(p)), \quad K(p) = \frac{1}{2} \langle M^{-1}p, p \rangle.$$

- (ii) Project p so as to satisfy the hidden constraint, i.e., $p_0 = P_{M,T}^* p$ with

$$P_{M,T}^* = \mathbf{1} - J_\Phi(J_\Phi^T M^{-1} J_\Phi)^{-1} J_\Phi^T M^{-1}, \quad J_\Phi = \mathbf{D}\Phi(q_0).$$

- (iii) Propagate $(\tilde{q}_1, \tilde{p}_1) = \Psi_\tau(q_0, p_0)$, where Ψ_τ is the numerical flow up to time $\tau > 0$, that is defined by the RATTLE discretization (5.3).

- (iv) Accept $q_1 = \tilde{q}_1$ with probability

$$r = \min \left(1, \frac{\exp(-\beta H(\tilde{q}_1, \tilde{p}_1))}{\exp(-\beta H(q_0, p_0))} \right),$$

or reject, i.e., set $q_1 = q_0$. (Here $H = K + V$ is the unconstrained Hamiltonian.)

- (v) Repeat.

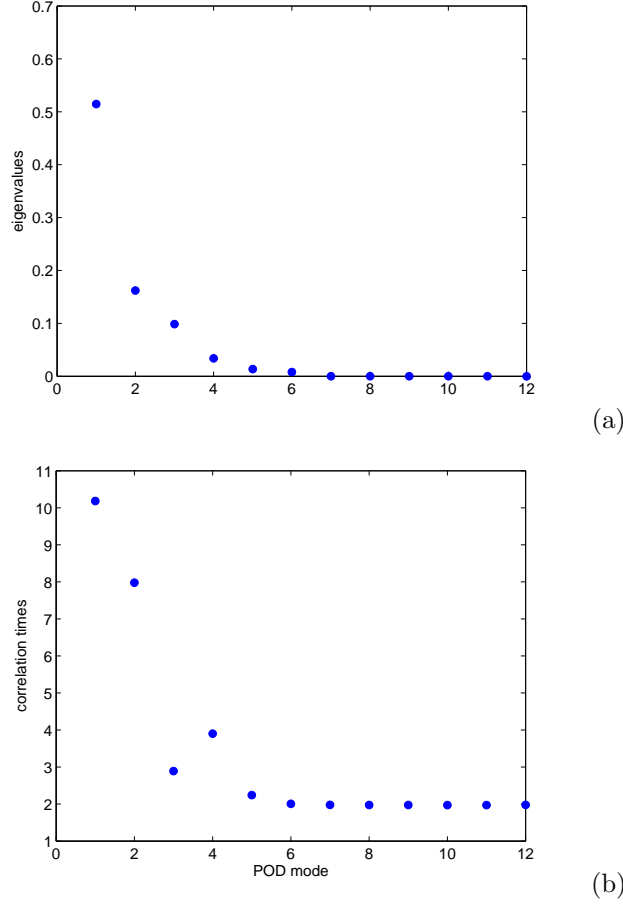


Figure 19. The figures show the POD analysis of the Cartesian configuration data of a butane molecule. The data stems from a HMC run at $T = 300\text{K}$ and $\tau = 50\text{fs}$ observation time span (step-size $h = 1\text{fs}$). The total number of steps is $N = 200\,000$ Upper panel: eigenvalues of the covariance matrix. Lower panel: characteristic time scales of the rotated modes (the scale is arbitrary)

5.2. Ryckaert-Bellemans n -butane

Proper orthogonal decomposition We study the spatio-temporal decomposition of the 12-dimensional Cartesian configuration space of a united atoms butane molecule. To this end we generate a HMC time series at $T = 300\text{K}$ with observation interval $\tau = 50\text{fs}$ between the HMC points. The integration is carried out with an ordinary Leapfrog/Verlet integrator with time step $h = 1\text{fs}$. For the chosen parameters h, τ the HMC acceptance probability is nearly one. Denoting by $\{q_1, \dots, q_N\}$ the HMC Markov chain of length $N = 200\,000$ we estimate the covariance matrix by

$$\hat{C} = \frac{1}{N-1} \sum_{k=1}^N (q_k - \bar{q}) (q_k - \bar{q})^T .$$

where \bar{q} denotes the arithmetic mean of the data

$$\bar{q} = \frac{1}{N} \sum_{k=1}^N q_k.$$

Moreover the data has been aligned in order to remove overall translations and rotations (i.e., the rigid body symmetry which lowers the rank \hat{C} by six).

Let us write the singular value decomposition of the symmetric covariance matrix as $\hat{C} = U\Lambda V^T$ with $\Lambda = \text{diag}(\lambda_k)$ with $\lambda_6, \dots, \lambda_{12} = 0$ and orthogonal matrices $U = V$. The POD modes are defined as

$$z_k = V^T (q_k - \bar{q}).$$

For the butane data we observe that the eigenvalues decay almost exponentially with one dominant eigenvalue that explains about 50% of the total variance (see Figure 19a). The first two modes cover about 81% of the total variance. To determine the number of important modes, we take a look at the corresponding decorrelation times

$$\tau_{d,i} = \frac{1}{N} \sum_{k=1}^N |c_k^i|,$$

where the discrete autocorrelation function c_k^i of the i -th POD mode at time lag k is estimated via fast Fourier transform of the data [285]. It can be seen in Figure 19b that the first two POD modes are slow in the sense that their autocorrelation functions decay slowly. The characteristic timescales of the remaining modes are comparably shorter. Note the following two features of the characteristic timescales: First of all the last six modes do not show any interesting behaviour which reflects the symmetry-reduction by means of the molecular alignment. (The same is true for the eigenvalues of the covariance matrix.) Furthermore the decay time τ_d is not a monotonic function of the number of modes, i.e., there may always be slow modes that have small variance.

Let us study the approximation capabilities of POD by means of the single torsion angle ω which is the observable of interest for the conformation dynamics. If we denote by $P_k \in \mathbf{R}^{n \times k}$ the matrix that contains the first k eigenvectors of the covariance matrix, we can define the rank- k approximant of the original data q_k as

$$\hat{q}_l = P_k P_k^T (q_l - \bar{q}) + \bar{q} = P_k (z_l^1, \dots, z_l^k)^T + \bar{q}.$$

Accordingly we obtain a reconstruction of the torsion angle by $\hat{\omega}_l = \omega(\hat{q}_l)$. We find in Figure 20a that even the single mode approximation yields the correct qualitative conformation behaviour between the three metastable sets (one *trans* and two *cis* conformations) although the variance is reduced as expected. This is contrasted with the two-mode approximation in Figure 20b which captures the main dynamical features of the conformation dynamics. Yet the invariant distribution of the torsion angle is not fully captured by the two-mode approximation, since small errors can accumulate over time and thus lead to the wrong distribution (see Figure 21 below).

Free energy calculation and optimal prediction We shall restrict our attention to the single torsion angle of the butane molecule; it will serve as the reaction coordinate from now on. Let us start with the free energy along the torsion angle which equals the torsion potential independently of the temperature [286]. (Note that this is not true in general, but only for the n -butane potential that does not involve Lennard-Jones interactions.) We compare the two quantities: Helmholtz free energy

$$F(\omega) = -\beta^{-1} \ln \int_{\Sigma} \exp(-\beta V) \|\nabla \Phi\|^{-1} d\sigma_{\omega},$$

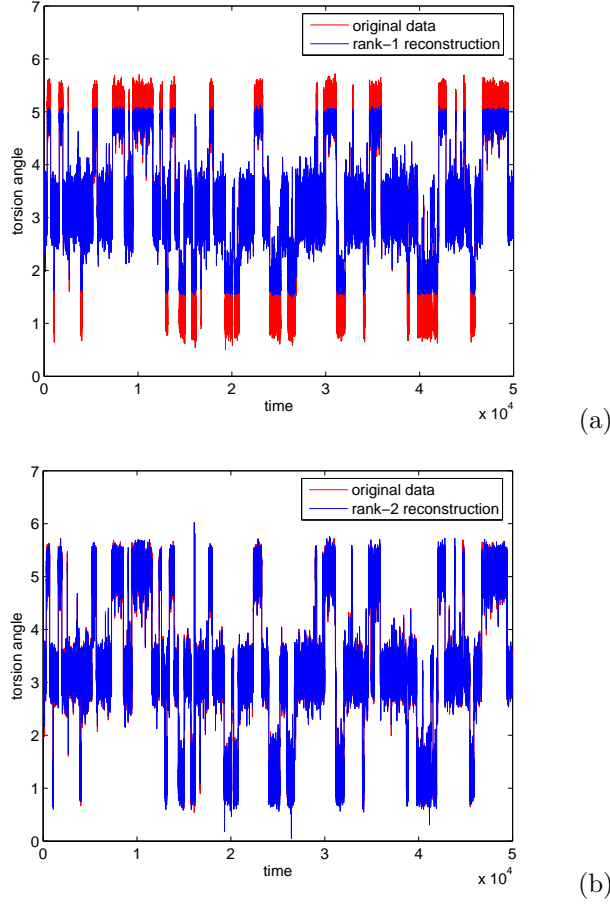


Figure 20. The plot compares the single mode approximation of the central torsion angle with the two-mode approximation for the first 50 000 HMC steps.

and geometric free energy

$$G(\omega) = -\beta^{-1} \ln \int_{\Sigma} \exp(-\beta V) d\sigma_{\omega},$$

where $\Sigma = \Phi^{-1}(\omega)$ in either case. Both energies are computed by Thermodynamic Integration using the constrained hybrid Monte-Carlo algorithm at temperature $T = 300\text{K}$ with $\tau = 50\text{fs}$ between the HMC points and 5 000 sample points per constrained run. The interval $[-\pi, \pi[$ was subdivided using a uniform grid of 60 values between 0.14 and 6.14 (see Section 5.1 for the simulation details).

The standard and the geometric free energy are shown in Figure 22a. One observes that the geometric free energy barrier from the *cis* to the *trans* conformation is about 1.0kJ/mol higher than the standard free energy barrier, which confirms that free energy barriers give a only lower bound for the (reversible) work that is needed to switch between different conformations. Simultaneously we have sampled the effective inverse mass, that appears in the optimal prediction Hamiltonian (3.63) and that is depicted in Figure 22b. The plot clearly indicated that the inverse mass or

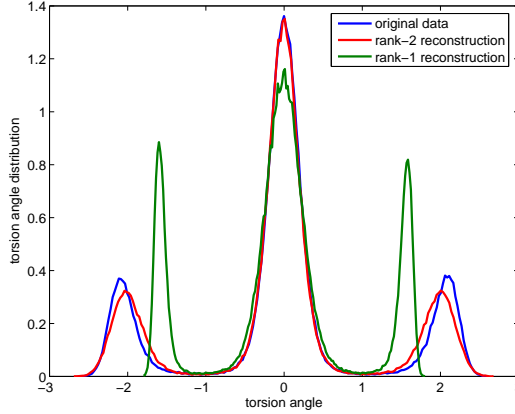


Figure 21. The plot shows histograms the torsion angle distribution at $T = 300\text{K}$ that were computed over the full time series and approximations thereof. The variance in both reconstructions is lowered as compared to the original data. The red histogram moreover illustrates that the correct distribution may not be reached although the dynamics seems well-captured. The reason is that small errors can accumulate over time and thus lead to the wrong distribution.

inverse metric, respectively, depends on the reaction coordinate, even in this simple case of a single torsion angle living on the unit circle S^1 . We observe that the kinetic energy favours the *trans* conformation, which is characterized by a rather slim shape with respect to the principal axis of inertia and which should be contrasted with to the more clustered *cis* conformations. This seems somehow counter-intuitive since one could expect that the mass distribution of a rotating molecule tends to spread out due to centrifugal forces. However here the situation is different since the rotation we are dealing with is an internal motion of the molecule. Likewise the respective kinetic energy is *internal* and should not be confused with the rotational energy of a rigid-body. Physically speaking, the effective mass indicates the redistribution of atomic masses for different conformations. Since the kinetic energy tends to stabilize the more compact *trans* conformation by slightly increasing the total energy of the *cis* conformations, we shall speak of an *internal centripetal force*.²⁴

In order to study the dynamical properties of the reduced model we compare the torsion dynamics of the full HMC simulation to a HMC simulation of the optimal prediction Hamiltonian as has been defined in Section 3.3.1:

$$E(\omega, \eta) = \frac{1}{2g(\omega)}\eta^2 + G(\omega).$$

Here

$$\frac{1}{g(\omega)} = \mathbf{E}_{\Sigma} \|\nabla \Phi\|_{M^{-1}}^2 \quad \text{with} \quad \|x\|_{M^{-1}}^2 = \langle M^{-1}x, x \rangle$$

is the mass-weighted effective inverse mass. A phase plot of the effective Hamiltonian is presented in Figure 23. For the sake of comparison we also do a HMC simulation in

²⁴Although the conjugate momentum to the torsion angle is certainly angular momentum, the reader should not be tempted to interpret the effective mass as moment of inertia. Both *cis* and *trans* have the same moments of inertia with respect to the rotation axis. In fact it is constant.

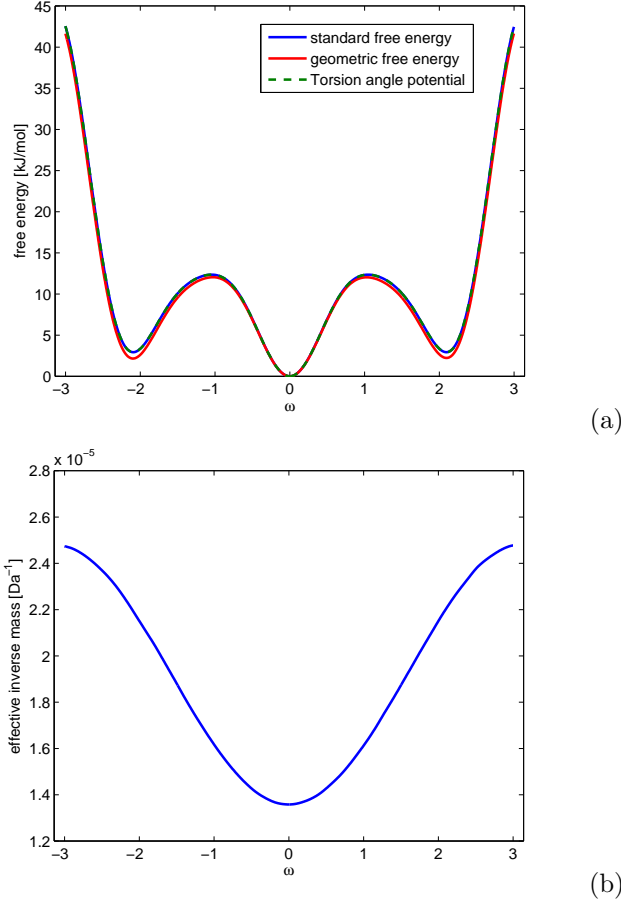


Figure 22. Comparison of free energy $F(\omega)$ and geometric free energy $G(\omega)$ at $T = 300\text{K}$ along the single torsion angle. For the Ryckaert-Bellemans *n*-butane the free energy does not depend on T . The energy barrier in the geometric free energy between *cis* and *trans* conformation is about 1.0kJ/mol higher than for the standard free energy. The lower panel shows the angle-dependent effective inverse mass that appears in the optimal prediction Hamiltonian.

the free energy landscape. The corresponding free energy Hamiltonian is defined by

$$H_{\text{free}}(\omega, \eta) = \frac{1}{2\mu}\eta^2 + F(\omega),$$

where μ is the effective mass

$$\mu = \beta^{-1} \text{cov}(\dot{\Phi})^{-1}.$$

Here $\dot{\Phi}$ means the torsion angle velocity. The definition of an effective mass by means of the covariance is based on the assumption that the kinetic energy in the simulated ensemble is equally partitioned among all degrees of freedom (equipartition of energy).²⁵ We run HMC simulations for E , H_{free} and for the full system, each at

²⁵Keep in mind that the hybrid Monte-Carlo trajectory does not contain momenta or velocities. Therefore we have computed the torsion angle velocity taking finite differences along the trajectory.

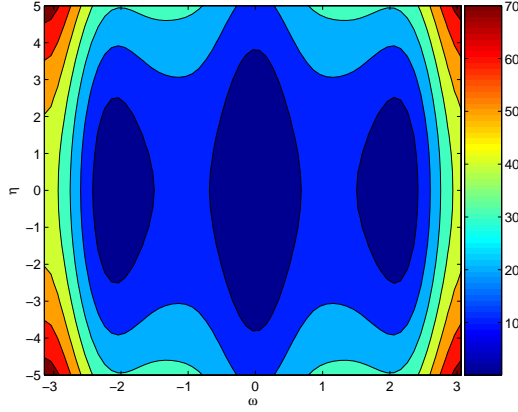


Figure 23. Optimal prediction Hamiltonian E at $T = 300\text{K}$ (arbitrary units).

room temperature $T = 300\text{K}$ with $\tau = 50\text{fs}$ and total length $N = 500\,000$ steps. The results in terms of torsion angle distribution and decay of correlations is presented in Figure 24 below. By definition the separable free energy Hamiltonian reproduces the correct marginal distribution of the torsion angle. However this is not true for the optimal prediction Hamiltonian: the marginal distribution here is²⁶

$$\int_{\mathbf{R}} \exp(-\beta E(\omega, \eta)) d\eta \propto \exp(-\beta G(\omega)) \left(\sqrt{\mathbf{E}_{\Sigma} \|\nabla \Phi\|_{M^{-1}}^2} \right)^{-1},$$

which should not be confused with the respective Fixman relation (3.26),

$$\exp(-\beta F(\omega)) = \exp(-\beta G(\omega)) \mathbf{E}_{\Sigma} \|\nabla \Phi\|_{M^{-1}}^{-1}.$$

Nevertheless it seems that the deviation is not too severe, since the optimal prediction simulation almost reaches the correct torsion angle distributions (see Figure 24a). As a dynamical observable we have computed the autocorrelation functions for each torsion angles. It can be seen in Figure 24b that the decay of correlations of the optimal prediction simulation and the free energy simulation is close to the behaviour of the full HMC Markov chain. Unfortunately the correct decay of correlations is no indicator for the correct transition behaviour between the conformations. Indeed if we define metastable sets by subdividing the interval $[-\pi, \pi[$ into three boxes $M_1 = [-\pi, -1[$, $M_2 = [-1, 1[$, and $M_3 = [1, \pi[$ and compute the transition matrix of the respective Markov chains by simply counting transitions within the HMC time step τ , we find the following three transition matrices:

$$P_{\text{full}} = \begin{pmatrix} 0.9927 & 0.0073 & 0 \\ 0.0022 & 0.9956 & 0.0022 \\ 0 & 0.0075 & 0.9925 \end{pmatrix} \quad (5.5)$$

$$P_{\text{op}} = \begin{pmatrix} 0.9927 & 0.0073 & 0 \\ 0.0022 & 0.9956 & 0.0022 \\ 0 & 0.0075 & 0.9925 \end{pmatrix} \quad (5.6)$$

²⁶The occurrence of mass-weighted metric is owed to the fact that the constrained expectation is understood as the configuration space marginal of a constrained Hamiltonian system. Therefore the respective Gibbs density involves the molecular masses by virtue of the Riemannian surface element.

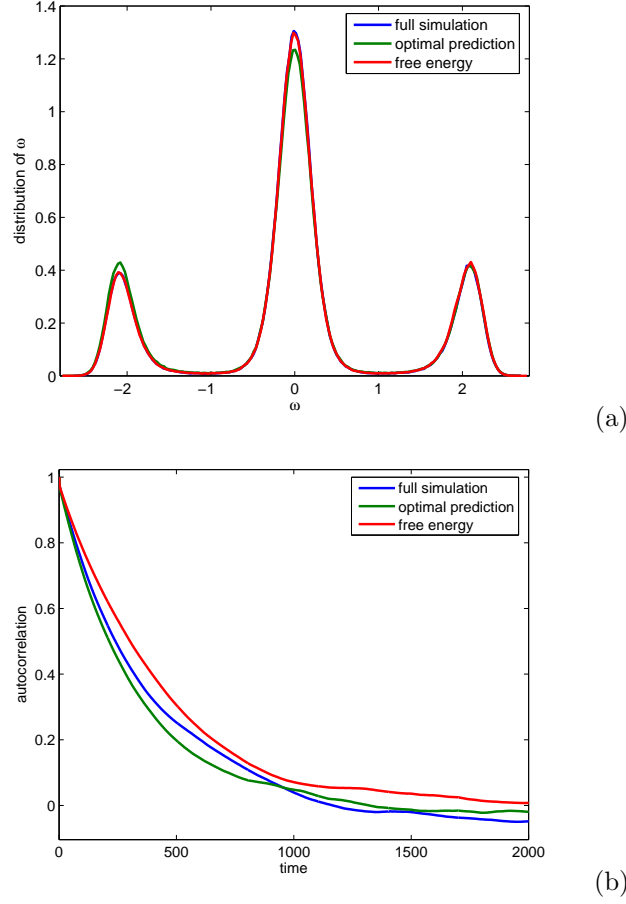


Figure 24. Comparison of reduced models for conformation dynamics. Upper panel: distribution of the torsion angles at $T = 300\text{K}$ over a 25ns HMC simulation. Lower panel: autocorrelation functions of the torsion angle (arbitrary time units).

$$P_{\text{free}} = \begin{pmatrix} 0.9941 & 0.0059 & 0 \\ 0.0016 & 0.9968 & 0.0016 \\ 0 & 0.0056 & 0.9944 \end{pmatrix}. \quad (5.7)$$

The three states M_1, M_2, M_3 are chosen such that the respective conformations are separated by the saddle points of the free energy landscape. Notice that $P_{\text{op}} = P_{\text{full}}$. In particular the probability to stay inside the *cis* transformations is lower as compared to the free energy system. Provided the coarse-grained dynamics is still Markovian, then the three-state Markov chains are completely characterized by the respective transition matrices (e.g., exit times, transition rates). Of course the deviations between the transition matrices are not very drastic, but we have observed that the identity $P_{\text{op}} = P_{\text{full}}$ is quite robust with regard to different discretizations. Moreover we should keep in mind that the matrices we all computed from finite samples.

Recapitulating, the optimal prediction Hamiltonian provides a reasonable reduced model for studying the dynamics between conformations on short time intervals (the

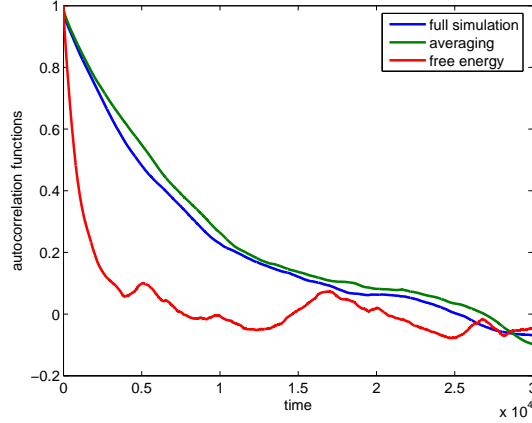


Figure 25. Decay of correlations for the torsion angle using Brownian dynamics at temperature $T = 300\text{K}$. (The time axis is arbitrarily scaled.)

correct invariant distribution is clearly given by the free energy). The main feature of optimal prediction however is that it allows for a physical interpretation of the different dynamical contributions in terms of inertial forces and forces coming from the reduced potential (i.e., the geometric free energy).

Brownian motion Just for illustration we repeat the simulation using the reduced diffusion model (3.46) from Section 3.2. The equation of motion is

$$\dot{\omega} = -\frac{1}{m(\omega)}G'(\omega) - \frac{1}{2\beta} \frac{m'(\omega)}{m(\omega)^2} + \sqrt{\frac{2}{\beta} \frac{1}{m(\omega)}} \dot{W}$$

where

$$m(\omega) = \mathbf{E}_{\Sigma} \|\nabla \Phi\|^{-2}$$

should not be confused with the effective mass g in the optimal prediction Hamiltonian. In particular the constrained expectation is understood with respect to the mass-free constrained Brownian dynamics. We have run a full simulation at $T = 300\text{K}$ and step-size $h = 1\text{fs}$, and also simulated diffusion in the respective free energy landscape using the Euler-Maruyama scheme. Then we compared both realizations to the reduced diffusion model. The full length of the simulation was $N = 1\,000\,000$ steps. (Note that the total length of the HMC trajectory was 25ns as compared to only 1ns here.)

It turns out that the naïve free energy model gives a remarkably wrong decay of correlations, whereas the averaged model does a fairly good job (see Figure 25). Apparently, it is necessary to introduce additional friction and noise parameters $\gamma, \sigma > 0$, similar to the effective mass μ in the free energy Hamiltonian, that controls the decay of correlations. Following [287, 288, 289] we can estimate the diffusion coefficient σ from the decay of the velocity autocorrelation function (Einstein-Green-Kubo formula). That is, we consider the reduced equation

$$\gamma \dot{\bar{\omega}} = -F'(\bar{\omega}) + \sigma \dot{W},$$

where

$$2\gamma = \beta\sigma^2 \quad \text{and} \quad \sigma^2 = \int_0^\infty c_\omega(t) dt$$

with $c_\omega(t)$ denoting the velocity autocorrelation function of the torsion angle. The result is shown in Figure 26, where it turns out that the parametrized free energy diffusion model now reproduces the correct decay of correlations. (It better be, for we have chosen friction and noise coefficients in such a way as to reflect the correct correlations.) Moreover the model captures the correct marginal distribution of the torsion angle by definition of the standard free energy — at least for a sufficiently long simulation. As we have argued, the decay of correlations does not tell us anything about other dynamical observables such as transition rates between metastable conformations. For further comparison we compute the Markov transition matrices between the sets $M_1 = [-\pi, -1[$, $M_2 = [-1, 1[$, and $M_3 = [1, \pi[$:

$$P_{\text{full}} = \begin{pmatrix} 0.9988 & 0.0012 & 0 \\ 0.0005 & 0.9990 & 0.0005 \\ 0 & 0.0011 & 0.9989 \end{pmatrix} \quad (5.8)$$

$$P_{\text{avg}} = \begin{pmatrix} 0.9988 & 0.0012 & 0 \\ 0.0005 & 0.9990 & 0.0005 \\ 0 & 0.0011 & 0.9989 \end{pmatrix} \quad (5.9)$$

$$P_{\text{free}} = \begin{pmatrix} 0.9982 & 0.0018 & 0 \\ 0.0004 & 0.9991 & 0.0005 \\ 0 & 0.0012 & 0.9988 \end{pmatrix}. \quad (5.10)$$

As before in case of optimal prediction, we observe that the transfer operator of the averaged model coincides exactly with the full propagator, $P_{\text{full}} = P_{\text{avg}}$. Conclusively, and in accordance with the considerations of the Hamiltonian system, we expect that the diffusion model with the geometric free energy is able to capture the correct transition rates between the *cis* and the *trans* conformations. The standard free energy model gives the correct statistics of the conformations though, and the deviation between the free energy propagator and the full one is not very drastic anyway.

Fixman potential: replacing fast degrees of freedom In Section 3.4 we have introduced an *ad-hoc* method to derive a semi-analytic reduced model. Since the calculation of effective models or free energy profiles, may be numerically expensive, we have proposed to replace the energy of the non-approximating modes in the system by an appropriately parametrized harmonic modelling potential. Averaging over the harmonic degrees of freedom then leads to a reduced model in terms of the reaction coordinate and an additional Fixman potential, the parametrization of which is open to choice. Moreover there is some freedom in choosing an appropriate metric on the approximant. For the butane molecule we suggest the following: Choosing the torsion angle as reaction coordinate the configuration space is essentially $S^1 \subset \mathbf{R}^{12}$. Given a metric $h(\omega)$ on S^1 , we consider the modelling Hamiltonian

$$H_{\text{fix}}(\omega, \eta) = \frac{1}{2h(\omega)} + V_{\text{tor}}(\omega) + \frac{1}{2\beta} \ln \det K(\omega).$$

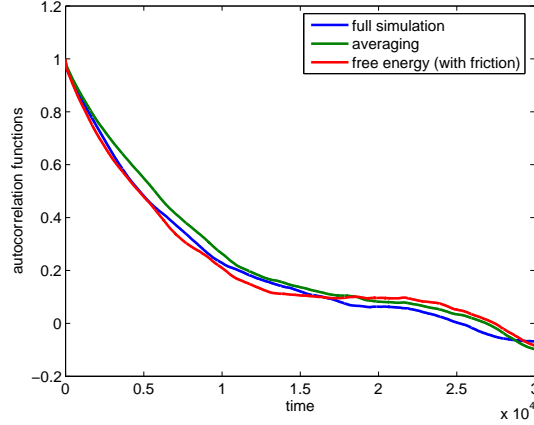


Figure 26. Decay of correlations for the torsion angle dynamics for Brownian dynamics with friction, where noise and friction coefficients σ, γ are computed from the Einstein-Green-Kubo relations (cf. Figure 25 above).

Here $K \in \mathbf{R}^{5 \times 5}$ is a yet unspecified, symmetric, positive-definite matrix.²⁷ As a potential V_{tor} , we choose the torsion potential; for the metric we choose either the constant effective mass $h(\omega) = \mu$ or $h(\omega) = g(\omega)$ from the last subsection. We decide to employ the covariance matrix of the symmetry-reduced data to feed the Fixman potential. There are essentially two possibilities

$$K_1(\omega) = \mathbf{E}_\xi(q_r(t) - \bar{q}_r)(q_r(t) - \bar{q}_r)^T,$$

or

$$K_2(\omega) = \mathbf{E}_\Sigma(q_r(t) - \bar{q}_r)(q_r(t) - \bar{q}_r)^T,$$

where \mathbf{E}_ξ and \mathbf{E}_Σ mean conditional or constrained averages, respectively, and the subscript r indicates that the data is aligned in order to account for the overall rotations and translations. Strictly speaking, the K_i denote only the irreducibly part of the covariance matrices. Running a sufficiently long unconstrained simulation the conditional covariance matrix can be computed from sorting the data according to the different values of the torsion angle. Alternatively, one could run short constrained simulations to get a local estimate of K_1 or K_2 . Here we have run constrained simulations on a 30 points grid each with about 5 000 HMC sampling points in order to obtain rather accurate estimates. In fact, it turns out that $\det K_1$ and $\det K_2$ are quite different (see Figure 27): the modelling potential that involves the constrained covariance matrix K_2 is right between the free energy and the geometric free energy $G(\omega)$, whereas the one involving K_1 is significantly different.

Hence we shall only consider two instances of the latter: a separable system with constant effective mass $h(\omega) = \mu$ and a non-separable system with $h(\omega) = g(\omega)$. For each system we run a HMC simulation at $T = 300\text{K}$ with $\tau = 50\text{fs}$. The total length is $T = 25\text{ns}$. The results are shown in Figure 28 below: it seems that the

²⁷Clearly, the normal space over S^1 is 11-dimensional. But if we further take into account the $SE(3)$ -symmetry of rigid motions, then the quotient space is effectively only 5-dimensional. Therefore the matrix in the Fixman potential has rank $s = 5$, and it suffices to consider its irreducible part.

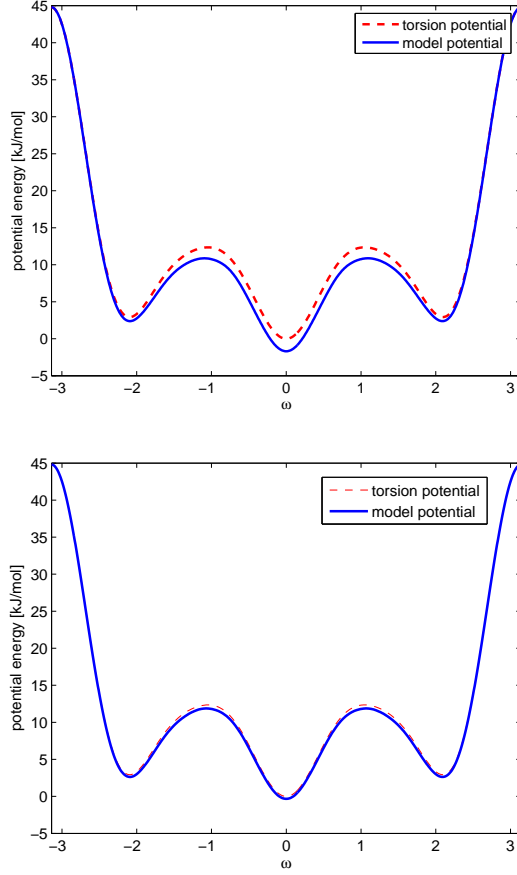


Figure 27. Different modelling potentials at $T = 300\text{K}$. Upper panel: potential energy for a Fixman potential computed from a conditional covariance matrix. Lower panel: modelling potential that has been computed from the constrained covariance matrix.

separable system with the constant effective mass performs slightly better in terms of torsion angle distribution and autocorrelation function, whereas the torsion-dependent mass overemphasizes the influence of the *cis* conformations' inertia which is already accounted for by the Fixman potential.

For further comparison we compute again the Markov transition matrices between the sets $M_1 = [-\pi, -1[$, $M_2 = [-1, 1[$, and $M_3 = [1, \pi[$. We find

$$P_{\text{full}} = \begin{pmatrix} 0.9936 & 0.0064 & 0 \\ 0.0017 & 0.9965 & 0.0018 \\ 0 & 0.0063 & 0.9937 \end{pmatrix} \quad (5.11)$$

$$P_{\text{non}} = \begin{pmatrix} 0.9915 & 0.0085 & 0 \\ 0.0012 & 0.9976 & 0.0012 \\ 0 & 0.0090 & 0.9910 \end{pmatrix} \quad (5.12)$$

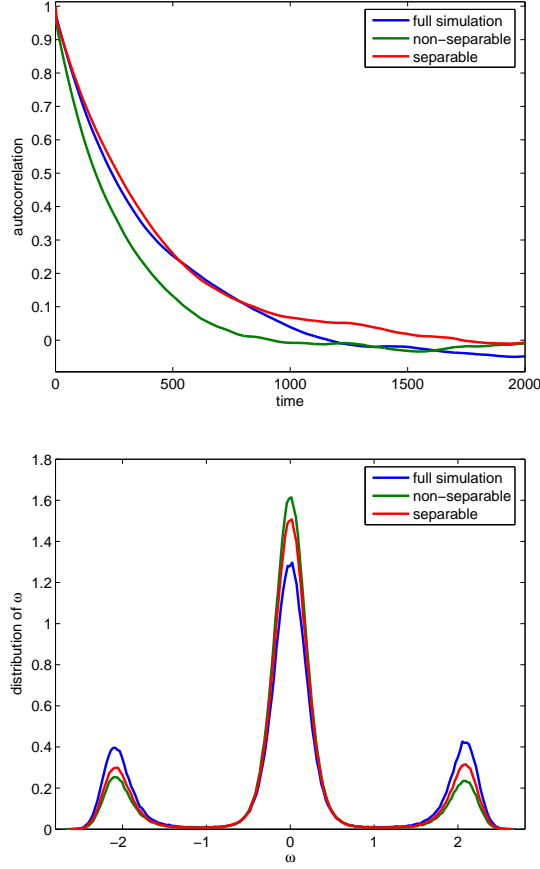


Figure 28. Simulation results for the modelling Hamiltonians involving the conditionally averaged covariance matrix at $T = 300\text{K}$. Upper panel: autocorrelation functions. Lower panel: distributions of the central torsion angles

$$P_{\text{sep}} = \begin{pmatrix} 0.9935 & 0.0065 & 0 \\ 0.0012 & 0.9976 & 0.0012 \\ 0 & 0.0066 & 0.9934 \end{pmatrix}. \quad (5.13)$$

Notice that the two (1,1) and (3,3) entries in the transition matrix P_{non} reflect the effect of the rotational energy of the *cis* transformation. To confirm that the separable model scores well, we compute the matrix norm of the differences $\Delta P_{\text{non}} = P_{\text{full}} - P_{\text{non}}$ and $\Delta P_{\text{sep}} = P_{\text{full}} - P_{\text{sep}}$. Indeed, we have

$$\|\Delta P_{\text{non}}\|_2 = 0.0046 \quad \text{and} \quad \|\Delta P_{\text{sep}}\|_2 = 0.0014,$$

where $\|\cdot\|_2$ denotes the spectral norm that is induced by the Euclidean vector norm. One should be careful with drawing conclusions from this simple example, as there are many different routes leading towards reduced models, and there is no *a priori* justification for preferring the separable model to the non-separable one. Setting up a good parametrization already requires some physical insight into the system; cf. [41, 38]. In fact it is not even clear whether an accurate sampling of the Fixman

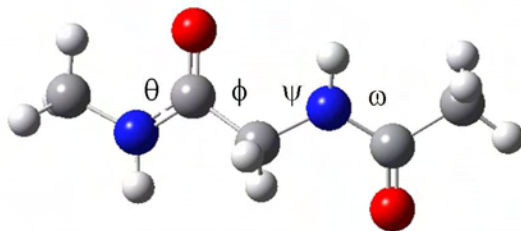


Figure 29. Glycine dipeptide in its extended C5 conformation.

potential improves the result, as we have carried out a substitution of the unresolved modes only in a small tubular neighbourhood of the approximant (here: S^1). However the simulations seem encouraging to promote further systematic studies.

5.3. Glycine dipeptide in vacuum

We consider the glycine dipeptide analogue as a paradigm for small biomolecules that exhibit interesting dynamical features, e.g., macroscopically distinct conformations and transitions between these conformations. The data is obtained from a 100ns hybrid Monte-Carlo simulation with a 100fs integration between the Monte-Carlo points. For the simulation we employ the GROMOS96 vacuum force field of GROMACS [290, 291] together with the native Java interface METAMACS [292]. It is important to note that we do not want to discuss issues whether the peptide model is physically meaningful or even realistic. It literally serves as a paradigm by means of which we can illustrate certain effects or compare different reduction schemes.

Proper orthogonal decomposition of torsion space We apply proper orthogonal decomposition (POD) to the torsion space of the glycine dipeptide analogue (GLDA). Regarding the macroscopic conformations we may confine our attention to the four central dihedral angles (see Figure 29); the leftmost and the rightmost torsions only rotate the methyl (CH_3) endgroups, which typically does not give rise to interesting physical effects. The four angles $(\theta, \phi, \psi, \omega) \in S^1 \times S^1 \times S^1 \times S^1$ span the 4-torus \mathbf{T}^4 which we shall consider as the essential configuration space. (Doing a decomposition of the Cartesian configuration space does not seem to make sense here, since the conformation dynamics of GLDA in vacuum can be completely described in terms of the torsion angles.)

In principle, we could embed the torus into a linear space, e.g., $\mathbf{T}^4 \subset \mathbf{R}^5$. But then we have to take into account that the embedding induces a nontrivial metric on \mathbf{R}^5 , which complicates both the POD and the following reduction steps. Instead we favour the idea of considering a flat torus, i.e., we regard \mathbf{T}^4 as the rectangle $[-\pi, \pi]^4$ (or $[-180, 180]^4$, respectively) upon identifying opposite edges. In this case it may happen that edges separate points lying in a single conformation which then show up as several clusters. However we are free to place the edges wherever we like. In particular we can cut the torus in a way that no clusters get separated. This amounts to a shift of the data as in Figure 30 which is optimal in the sense that the number of transitions over the edges within one Monte-Carlo step is minimized. As POD is affine

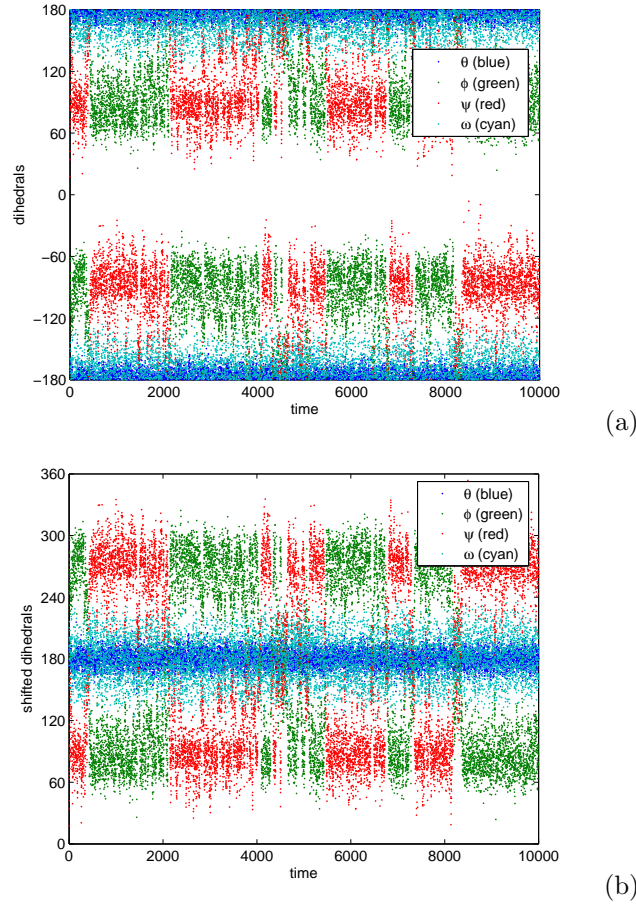


Figure 30. Optimal cut of the torus \mathbf{T}^4 . Notice that unlike before in Section 5.2, all angles are measured in degrees.

invariant it is clearly not affected by the shifting, and we can apply standard POD.

Diagonalizing the covariance matrix of the centered torsion data, it turns out that there is a single POD mode that is separated from the remaining modes by a considerable eigenvalue gap. This explains why the torsion data can be reconstructed from a single POD mode as can be seen from Figure 31. Moreover the dominant mode is much slower mixing than the other modes as is indicated by the decay of autocorrelations, which qualifies it for the systems reaction coordinate. (The eigenvalues of the covariance matrix and the characteristic time scales can be found in the caption of Figure 31.) Intriguingly the first mode depends almost completely on the two central backbone angles (ϕ, ψ) , viz.,

$$z_1 \approx \frac{1}{\sqrt{2}} (\psi - \phi) .$$

Furthermore we see from the time series that $\psi(t)$ equals $-\phi(t)$ (to all appearances). This behaviour reflects the fact that the glycine dipeptide analogue is symmetric with respect to the plane that is spanned by the peptide bond in the extended C5

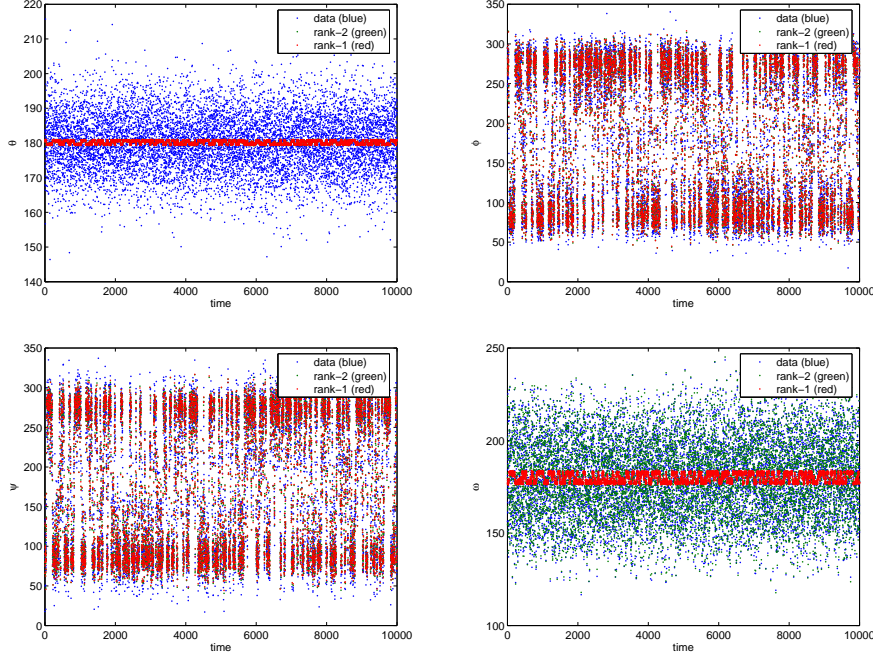


Figure 31. POD of the shifted dihedrals. The eigenvalues of the data covariance matrix are $\sigma = \{1.4448, 0.0416, 0.0371, 0.0066\}$. The corresponding autocorrelation times are $\tau = \{238.2034, 2.1211, 3.4585, 2.4682\}$.

conformation. Accordingly we have $z_1 \approx \sqrt{2}\psi$. The agreement is indeed surprising, as Figure 32 indicates. However it should be kept in mind that the approximation is likely not to capture the correct long-term behaviour (e.g., invariant distribution), since small deviations can accumulate over time.

Free energy and the optimal prediction Hamiltonian We should not be tempted to think that ψ is the slowest mode in the system, for equal reasoning applies for the other backbone angle ϕ . If we assume that dihedral angles are typically the slowest degrees of freedom in the system as compared to bond and bond angle vibrations, then it follows that only ϕ and ψ together can be considered as slow variables. Nevertheless it is instructive to first take a look at standard and geometric free energy in terms of a single backbone angle, e.g., the ψ angle.

To this end we constrain the reaction coordinate ψ , and do Thermodynamic Integration at $T = 300\text{K}$ along 36 different values ψ between -180° and $+180^\circ$. The simulation is carried out using the constrained hybrid Monte-Carlo scheme from Section 5.1, where each constrained run involves 50 000 sample points with $\tau = 100\text{fs}$ spacing between the Monte-Carlo points and internal step-size $h = 1\text{fs}$, such that the acceptance rate varies between 98% and 99%. Each integration starts from a configuration for the respective angle after minimizing the potential energy. Again we compare standard free energy

$$F(\psi) = -\beta^{-1} \ln \int_{\Sigma} \exp(-\beta V) \|\nabla \Phi\|^{-1} d\sigma_{\psi},$$

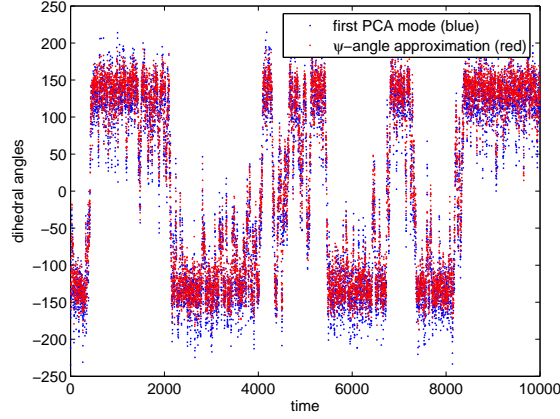


Figure 32. Approximation of the dominant POD mode by the angle ψ .

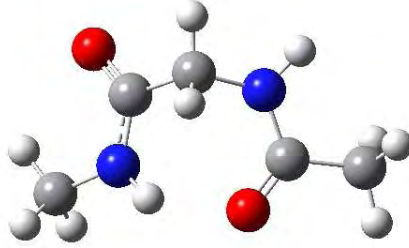


Figure 33. Glycine dipeptide in its C7 conformation.

and geometric free energy

$$G(\psi) = -\beta^{-1} \ln \int_{\Sigma} \exp(-\beta V) d\sigma_{\psi},$$

where $\Sigma = \Phi^{-1}(\psi)$ is of codimension one in the configuration space \mathbf{R}^{57} . Both standard and geometric free energy are shown in Figure 34a below. The C7 conformations at $\psi \approx \pm 85^\circ$ and the stretched C5 conformations at $\psi \approx \pm 150^\circ$ are clearly distinguished (see Figure 33). Recalling the considerations from the butane example we face the same effect, namely, that the geometric free energy favours the cluster-like C7 conformation as compared to the profile of the standard free energy. This becomes more lucid if we look at the corresponding effective inverse mass $1/g$ as it appears in the optimal prediction Hamiltonian, and which is defined by the constrained expectation

$$g(\psi) = (\mathbf{E}_{\Sigma} \|\nabla \Phi\|_{M^{-1}}^2)^{-1}.$$

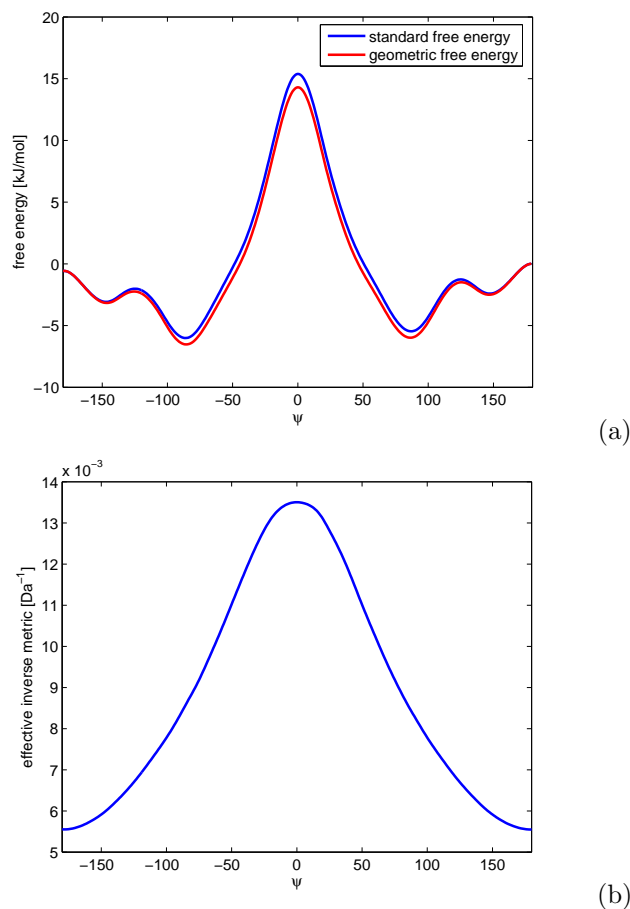


Figure 34. Free energy and effective inverse mass of the ψ angle at $T = 300\text{K}$.

The effective inverse mass is depicted in Figure 34b. It again reflects the tendency of the kinetic energy to stabilize those conformations which are slim with respect to the principal axis of inertia. (The destabilizing effect is even more evident for the knobby 0° conformation which, however, is extremely unfavourable anyway.)

Note that the effective inverse mass has almost the same shape as in the butane example before (up to a 180° shift). This suggests that the inverse effective mass is a genuine property of the chosen reaction coordinate (here: a torsion angle). This is quite remarkable, since the quantity g is computed as a configurational average which depends very well on the molecular potential and thus on the specific molecule.

Effective models in the Ramachandran plane The one-dimensional free energy profiles and the effective inverse mass of the optimal prediction Hamiltonian could already provide some insight into the conformation dynamics of GLDA. As in the butane example it has turned out that the extrinsic geometry of the reaction coordinate can have significant dynamical effects on the conformation dynamics that compete with the effects induced by the potential energy. We shall complete the picture by

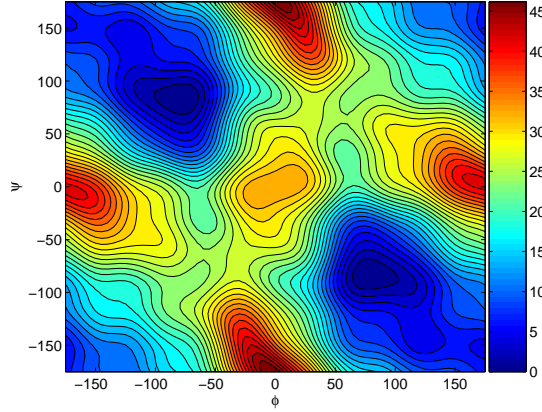


Figure 35. Standard free energy $F(\phi, \psi)$ at $T = 300\text{K}$.

looking at free energy landscapes and effective kinetic energy along the two central backbone angles ϕ and ψ .

In order to compute free energy and effective inverse mass, we perform Thermodynamic Integration in the Ramachandran plane. Such calculations are rare (e.g., [11]), although easy-to-use Thermodynamic Integration formulae in more than one dimension have been put forward during the last few years (see also [293], using a simplified force expression). We cover the Ramachandran plane with a two-dimensional, uniform 36×36 grid, and run constrained hybrid Monte-Carlo simulations at $T = 300\text{K}$ on each grid point (ϕ_i, ψ_j) . The step-size was chosen to be $h = 1\text{fs}$ with 100 integration steps between between the Monte-Carlo points. Starting from a energy-minimized configuration, each simulation involves $N = 10\,000$ sample points, hence equivalently 1ns of total integration time for each ϕ, ψ combination. Using the expressions from the Corollaries 4.15 and 4.16, we compare both standard free energy

$$F(\phi, \psi) = -\beta^{-1} \ln \int_{\Sigma} \exp(-\beta V) (\text{vol} J_{\Phi})^{-1} d\sigma_{\phi, \psi},$$

and geometric free energy

$$G(\phi, \psi) = -\beta^{-1} \ln \int_{\Sigma} \exp(-\beta V) d\sigma_{\phi, \psi}$$

with $\Sigma = \Phi^{-1}(\phi, \psi)$. Taking advantage of the reaction coordinate's periodicity, we reconstruct the smooth free energy surfaces from the mean forces by expanding F and G into truncated, two-dimensional Fourier series [294]. The respective Fourier coefficients are then determined from the sampled derivatives, ∇G and ∇F , in a least square sense which amounts to solving an underdetermined linear system of equations.

The results are shown in the Figures 35 and 36. Both plots clearly reveal the C7 conformation at about $(\phi, \psi) = (\pm 80^\circ, \mp 80^\circ)$. Moreover, but less clearly, we can see the extended C5 conformation around $(\phi, \psi) = (\pm 180^\circ, \pm 150^\circ)$ which is about 5 – 10kJ/mol higher than the C7 conformation. Again the C5 energy in the geometric free energy landscape is raised as compared to the standard free energy (circa 3.5kJ/mol). For illustration the minimized potential energy function projected

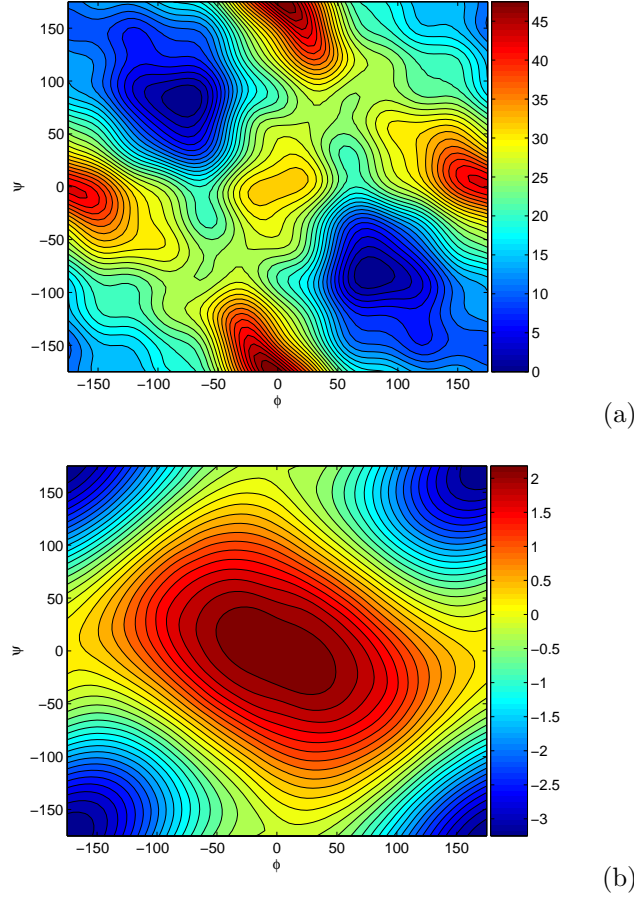


Figure 36. Geometric free energy $G(\phi, \psi)$ at $T = 300\text{K}$. The upper panel shows the geometric free energy, whereas the lower panel depicts the difference to the standard free energy. See Figure 35 and the Fixman potential in Figure 38.

onto the Ramachandran plane is shown in Figure 37 below. The most noticeable difference is that the energy barriers of the strongly repulsive O-O ring-like state at $\phi = 0^\circ$ and the H-H ring-like state at $\psi = 0^\circ$ are far more pronounced than in the free energy landscapes, which, however, does not belong to any admissible transition path anyway (cf. Figure 30). The plots indicate that both geometric and free energy are mainly affected by the potential energy landscape of the system.

For further comparison we have plotted the difference $F - G$ in Figure 36b. Letting Φ denote the reaction coordinate, the following is known on analytical grounds

$$F - G = -\beta^{-1} \ln \mathbf{E}_\Sigma (\text{vol} J_\Phi)^{-1},$$

where the mass-weighted matrix volume is defined as

$$\text{vol} J_\Phi = \sqrt{\det(\mathbf{D}\Phi \mathbf{M}^{-1} \mathbf{D}\Phi)},$$

This should be contrasted with the Fixman potential $A = -\beta^{-1} \ln \sqrt{\det g}$ that is

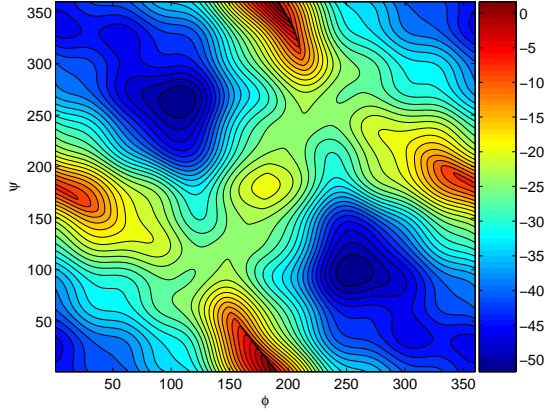


Figure 37. Minimized potential energy landscape of the backbone angles.

defined by the relation

$$\int \exp(-\beta E) d\eta \propto \exp(-\beta A) \exp(-\beta G),$$

where E is the optimal prediction Hamiltonian

$$E(\phi, \psi, \eta_1, \eta_2) = \frac{1}{2} g^{ij}(\phi, \psi) \eta_i \eta_j + G(\phi, \psi)$$

with the effective inverse mass

$$g^{-1} = \mathbf{E}_\Sigma \mathbf{D} \Phi^T M^{-1} \mathbf{D} \Phi \in \mathbf{R}^{2 \times 2}.$$

The Fixman potential A is shown in Figure 38 below. Note that clearly $A \neq F - G$, since the constrained expectation does not commute with the operations performed on the mass-weighted Gramian. (They are quite similar though.) We abstain from doing numerical simulations of the reduced system, and focus on qualitative features instead. To this end we shall study the specific form of the effective kinetic energy in more detail. Roughly, Figure 38 reflects the familiar property of the free energy to favour the extended C5 conformation: the Fixman potential achieves its global minimum in the $\pm 180^\circ$ corners of the Ramachandran plane. Moreover we see that the effective inverse mass is not diagonal, i.e., the two angles are coupled by the effective kinetic energy, where the off-diagonal terms' order of magnitude in Figure 38 is about one tenth of that of the diagonal entries. The reader should compare this to the calculations of the kinetic energy for a system with constrained bonds in [295] which reveal quite similar features. Additionally we see that the reduced model inherits the systems' symmetry, since also the kinetic energy is invariant under parity $(\phi, \psi) \mapsto -(\phi, \psi)$. Intriguingly we find that, other than the free energy, the kinetic energy carries a higher (approximate) symmetry that stems from the (almost) invariance with respect to the transformations $\phi \mapsto -\phi$ or $\psi \mapsto -\psi$. The slight perturbation of the symmetry stems from the non-uniform mass distribution along the peptide's backbone.

The kinetic energy plots reveal an interesting feature of the system: Let us assume that no extra potential were present in the optimal prediction Hamiltonian, i.e., $G = 0$, while keeping the shape of the kinetic energy as is (which is not so far off as the

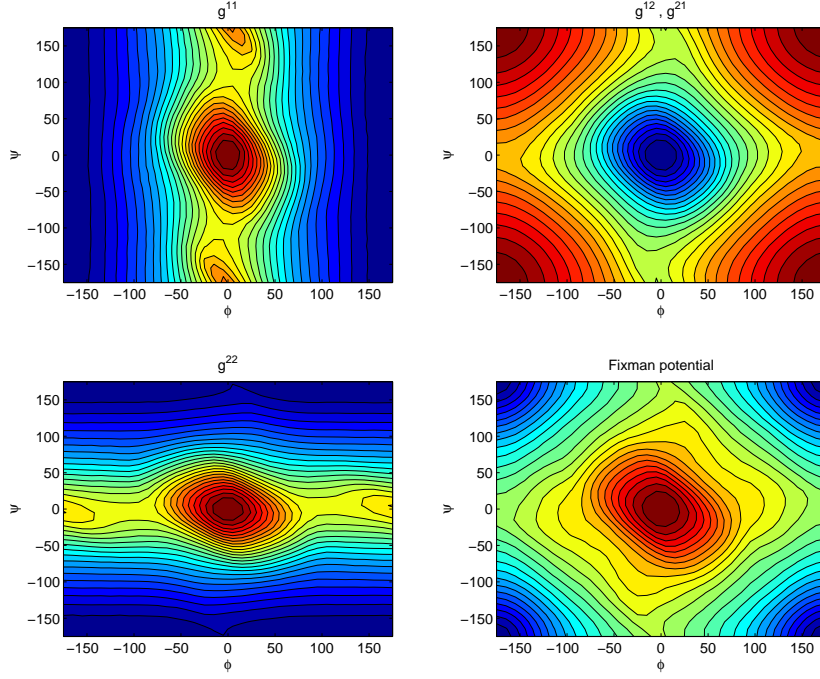


Figure 38. Effective inverse mass $g^{-1} = \mathbf{E}_{\Sigma}(J_{\Phi}^T M^{-1} J_{\Phi})$ of the two central backbone angles at $T = 300\text{K}$. The Fixman potential $A = -\beta^{-1} \ln \sqrt{\det g}$ is shown in the lower right corner (cf. also Figure 36).

comparison with the Ryckaert-Bellemans example shows). Then the system is (in a loose sense) geodesic, but is still likely to remain in the C5 conformation whenever the momentum is nonzero; only for $\eta = 0$ (inertia-free motion) the total energy is identically zero in the whole Ramachandran plane. Note that this comes up as a property of the effective Riemannian metric in the expression of the kinetic energy which therefore acts as a dynamical force that is mainly induced by the structure of the backbone (cf. the discussion in Example 3.20). In this sense the extrinsic geometry of the reaction coordinate renders conformation *dynamics* of a molecule.