A Hybrid Monte Carlo Method for Essential Molecular Dynamics

Ch. Schütte[1,2], A. Fischer[1], W. Huisinga[1], and P. Deuflhard[1,2]

[1] Konrad Zuse Zentrum Berlin, Takustr. 7, 14195 Berlin, Germany
[2] Freie Universität Berlin, Fachbereich Mathematik, Arnimallee 2–6, 14195 Berlin, Germany

March 2, 1998

Abstract. Recently, a novel concept for the computation of *essential* features of Hamiltonian systems (such as those arising in molecular dynamics) has been proposed. The realization of that concept was based on subdivision techniques applied to the Frobenius–Perron operator for the dynamical system. The present paper suggests an alternative but related concept based on statistical mechanics, which allows to attack *realistic* molecular systems. In a first step, the frequency of conformational changes is characterized in statistical terms leading to the definition of some Markov operator T that describes the corresponding transition probabilities within the canonical ensemble. In a second step, a discretization of T via hybrid Monte Carlo techniques (based on *short* term subtrajectories only) is shown to lead to a *stochastic* matrix P. With these theoretical preparations, an identification algorithm for conformations is applicable (to be presented elsewhere). Numerical results for the n-pentane molecule are given and interpreted.

1 Introduction

The classical microscopic description of molecular processes leads to a mathematical model in terms of Hamiltonian differential equations. In principle, the discretization of such systems permits a simulation of the dynamics. However, direct simulation is even today restricted to relatively short time spans and to comparatively small discretization steps. Fortunately, most questions of chemical relevance just require the computation of *averages* of physical observables, of *stable conformations*, or of *conformational changes*. In a conformation, the large scale geometric structure of the molecule is understood to be conserved, whereas on smaller scales the system may well rotate, oscillate or fluctuate. The computational characterization of a conformation via direct simulation thus often requires inaccessibly long time spans. Therefore, unlike some former approaches (e.g. [2,16]), we herein advocate a different line of method: we suggest to *directly* attack the determination of conformations and their stability time spans, which means some *global* approach clearly differing from any kind of statistical analysis based on long term trajectories. In a recent article [9], such a global approach based just on *short* *term simulations* has first been presented. The idea is to directly solve a discretized eigenvalue problem for the Frobenius–Perron operator, an operator which describes the propagation of probability within the system. However, it has turned out that, even though the numerical results are intriguing, this approach it suffers both from a (yet) unclear theoretical justification and from the "curse of dimension" of the suggested subdivision algorithm.

Herein, we will proceed to an alternative but related concept based on a deeper understanding of the physical interpretation of the Frobenius–Perron operator in the context of statistical mechanics (Sec. 2). On this basis, a *spatial transition operator* will be defined, which describes the probability of fluctuations within the canonical ensemble (Sec. 3.1). This operator replaces the Frobenius–Perron operator in computing the conformations, again via the eigenmodes of the operator. The corresponding eigenvalue problem is naturally discretized by means of a Galerkin procedure (Sec. 3.2), which results in a reversible stochastic transition matrix. The entries of this matrix describe transition probabilities between certain subsets of phase space induced by the spatial discretization of the system and can be evaluated by means of hybrid Monte Carlo methods (Sec. 3.3). In our new approach, chemical "conformations" now arise as subsets in *position space* rather than in phase space. Once the transition matrix is assembled, the conformational subsets can directly be determined by our recently developed identification algorithm [11]. The performance of the resulting algorithm is illustrated by numerical results for the n-pentane molecule (Sec. 4).

With this novel concept, both of the above named difficulties of the former approach [9] can be overcome. In particular, a well-suited theoretical basis can be given and realistic molecular systems come into reach.

2 Conformations as Almost Invariant Sets

In classical MD (cf. textbook [1]) a molecule is modeled by a Hamiltonian function

$$H(q,p) = \tfrac{1}{2}\, p^T M^{-1} p + V(q), \tag{1}$$

where q and p are the corresponding positions and momenta of the atoms, M the diagonal mass matrix, and V a differentiable potential. The Hamiltonian H is defined on the phase space $\Gamma \subset \mathbb{R}^{6N}$. The corresponding canonical equations of motion

$$\dot{q} = M^{-1} p, \qquad \dot{p} = -\operatorname{grad} V \tag{2}$$

describe the dynamics of the molecule. The formal solution of (2) with initial state $x_0 = (q(0), p(0))$ is given by $x(t) = (q(t), p(t)) = \Phi^t x_0$, where Φ^t denotes the flow.

On the smallest time scales (around 1 femtosecond) the dynamics described by the flow Φ^t consists of fast oscillations around equilibrium positions (bond length or bond angle vibrations). In contrast to these fast fluctuations the phrase "conformations" describes quasi-stable global configurations of the molecule. *Conformational changes* are therefore rare events, which will show up only in long term simulations of the dynamics (e.g. on a nano- or millisecond time scale). From a mathematical point of view, conformations are special "almost invariant" subsets of phase space: *Invariant sets* correspond to infinite durations of stay (or relaxation times) and contain all subsets associated with different conformations. *Almost invariant sets* correspond to finite relaxation times and consist of conformational subsets. In order to characterize the conformational dynamics of the molecular system, these subsets are the objects of interest.

More precisely, a subset $B \subset \Gamma$ is called *invariant* under the flow Φ^t iff, for all $t > 0$,

$$\Phi^t(B) = B \quad \text{and, thus,} \quad \Phi^{-t}(B) = B.$$

We now aim at a precise mathematical understanding of "almost invariance" of a subset $B \subset \Gamma$ in a certain time span τ. Therefore, we have to introduce a measure for quantizing the fraction $B \cap \Phi^\tau(B)$ that remains in B under the action of the flow during this time span. Let μ be any suitable probability measure on Γ. The degree of invariance of B with respect to μ is given by the corresponding conditional probability

$$\delta(B,\tau) = \frac{\mu\left(B \cap \Phi^\tau(B)\right)}{\mu(B)} \leq 1, \qquad B \ \mu\text{-measurable.} \tag{3}$$

In particular, if B is invariant, then $\delta(B,\tau) = 1$ independent of the choice of μ. We are interested in subsets B with $\delta(B,\tau)$ sufficiently close to $\delta = 1$, to be denoted as *almost invariant* subsets.

From now on, let us fix a suitable time span τ. Thus, we have reduced the continuous dynamical system (2) to a discrete dynamical system

$$x_{k+1} = g(x_k), \quad k = 0,1,2,\ldots \quad \text{with} \quad g = \Phi^\tau. \tag{4}$$

The long term behavior of the system (4) is described by so-called *invariant measures*: a probability measure μ is invariant, iff $\mu(\Phi^\tau(B)) = \mu(B)$ for all measurable subsets $B \subset \Gamma$. Thus, $\mu(B)$ may be interpreted as the probability of finding the molecular system in B at an arbitrary instant $t = k\tau$, $k \in \mathbb{Z}$.

2.1 Dynamical Systems Approach and the Frobenius–Perron Operator

The numerical computation of invariant measures is equivalent to the solution of an *eigenvalue problem* for the so-called *Frobenius–Perron operator* $U : \mathcal{M} \to \mathcal{M}$ defined on the set $\mathcal{M}$ of probability measures on Γ by virtue of

$$(U\mu)(B) = \mu(g^{-1}(B)) \quad \text{for measurable } B \subset \Gamma \text{ and } \mu \in \mathcal{M}. \tag{5}$$

Invariant measures correspond to fixed points of U which means that $U\mu = \mu$ iff $\mu \in \mathcal{M}$ is invariant.

It has been discovered [8], that for many discrete dynamical systems the

almost invariant sets are related to eigenmodes of the Frobenius–Perron operator for eigenvalues $\lambda \approx 1$ *inside* the unit circle ($|\lambda| < 1$). (6)

One strategy for identification of almost invariant sets is to discretize the Frobenius–Perron operator in order to approximate these eigenvalues $\lambda \approx 1$. In a sequence of articles (cf. [7,8]), M. Dellnitz and coworkers established numerical techniques realizing this strategy for different non-Hamiltonian systems. The Frobenius–Perron operator is discretized via a multi-level subdivision process, which generates a box covering of the system's relative global attractor. Recently, this approach has been extended to Hamiltonian systems [9]. In this case, the collection of discretization boxes covers the special *energy surface* under consideration. This numerical approach seems to reproduce precisely the correct results from the dynamical point of view. In fact, one obtains the expected almost invariant sets together with reasonable corresponding eigenvalues.

This "dynamical systems approach", however, has two crucial difficulties. First, this approach turns out to be useful only for small molecular systems, since it suffers from *combinatorial explosion* of the necessary number of discretization boxes already for moderate size molecules. Second, the theoretical justification of the approach is unclear: The crucial properties of the Frobenius–Perron operator U depend on the spaces under consideration.

- If U is acting on the entire phase space Γ, then there exist *infinitely many* invariant measures and the Frobenius–Perron operator possibly has *no eigenvalues inside the unit circle*. Thus, there may be no eigenvalues $\lambda < 1$ allowing for the identification of almost invariant sets. As we will see below, this results from the fact that $g = \Phi^\tau$ preserves energy and volume and can already be understood by reformulating U as an operator on *phase space densities* instead of measures.
- The situation changes if we consider U as an operator acting on measures on a certain *energy surface* (as intended in [9]) and not on the entire phase space. In this case, there might be no invariant measure with an density (see [9], Sec. 3.1), i.e., the invariant measure may indeed be unique. But even then, the existence of eigenvalues *inside* the unit circle is still questionable.

We herein want to take into account the typical fluctuation of total energy in molecular systems, i.e., we anyway have to work on the entire phase space. To simplify, we now reformulate the Frobenius–Perron operator U as an operator on *phase space densities*. We consider the function spaces

$$L^p(\Gamma) \;=\; \{f : \Gamma \to \mathbb{C}, \quad \int_\Gamma |f(x)|^p \, dx \;<\; \infty\}, \quad \text{with } p = 1, 2,$$

and denote the corresponding function norms by $\|\cdot\|_p$. For defining U on these L^p spaces, let us assume that the measure μ in the definition (5) has a density $f \in L^p(\Gamma)$, so that $\mu(B) = \int_B f(x)dx$. Moreover, let $U\mu$ also have a density denoted by Uf. Then, the above defined (volume preserving) transformation $y = g(x)$ leads us to

$$\int\limits_B Uf(x)\,dx \;=\; \int\limits_{g^{-1}(B)} f(x)\,dx \;=\; \int\limits_B f(g^{-1}(y))\,dy, \quad \text{for measurable } B \subset \Gamma.$$

Thus, the definition of the Frobenius–Perron operator on $L^p(\Gamma)$ is

$$Uf \;=\; f \circ g^{-1} \;=\; f \circ \Phi^{-\tau}, \tag{7}$$

for more details see [20]. For these operators the literature on ergodic theory contains the following result (cf. [28], Thm. 1.25; [20], Prop. 3.1.2; or [25]):

$$\|Uf\|_p \;=\; \|f\|_p, \qquad \forall f \in L^p(\Gamma), \quad p = 1, 2. \tag{8}$$

As a consequence, the spectrum of our operators lies *on* the unit circle.

2.2 Reformulation in Terms of Statistical Mechanics

To better understand the physical meaning of the Frobenius–Perron operator we have to consider the role of phase space densities in the statistical formulation of classical mechanics. In statistical mechanics the evolution of a *statistical ensemble* of identically prepared systems is described by a time dependent probability density $f = f(x,t)$ in phase space. The propagation of the probability density is described by the Liouville equation for the Hamiltonian H:

$$\partial_t f \;=\; i\mathcal{L}f \;=\; \{H, f\}, \qquad f(t=0) \;=\; f_0, \tag{9}$$

where $\{\cdot,\cdot\}$ denotes the well–known Poisson bracket and $\mathcal{L} = -i\{H,\cdot\}$ the associated Liouville operator (cf. [19]). The density f_0 describes the initial probability distribution in the statistical ensemble, i.e., $f_0(x)$ is interpreted as the relative frequency in the ensemble of systems in state x at time $t = 0$. Therefore, the density f_0 is to be *defined* in accordance with the *initial experimental preparation* of the ensemble.

On the one hand, the solution of (9) is given by the flow corresponding to H,

$$f(x,t) \;=\; f_0(\Phi^{-t}x),$$

on the other hand, it can be denoted using the semi-group generated by $\mathcal{L}$ on the Hilbert space $L^2(\Gamma)$:

$$f(\cdot,t) \;=\; \exp(it\mathcal{L})\, f_0.$$

From this equality and a comparison with the definition (7), we may write our Frobenius–Perron operator U as

$$U \;=\; \exp(i\tau\mathcal{L}) \qquad \text{in } L^2(\Gamma).$$

This result has some intriguing consequences: In $L^2(\Gamma)$, U is unitary, since $\mathcal{L}$ is self-adjoint [18]. This explains equation (8) for $p = 2$. All stationary solutions of the Liouville equations are invariant densities of U, i.e., eigenvectors for the eigenvalue $\lambda = 1$. In particular, for *arbitrary* smooth functions $F : \mathbb{R} \to [0,1]$, the associated densities $f(x) = F(H(x))$ are stationary solutions of the Liouville equation. Consequently, there are infinitely many invariant densities (and associated invariant measures) for U — as long as we are working on the entire phase space Γ.

Clearly, we need a *unique* invariant density in order to induce the measure μ used to quantify the "almost invariance" of sets according to (3). Thus, we have to *choose* one of the possible invariant densities of U. As already stated above, in terms of the statistical formulation this choice corresponds to the *initial preparation* of a stationary ensemble, i.e., the initial conditions of the experiment determine the invariant density f_0. Hence, due to (3), the degree of invariance of a subset $B \subset \Gamma$ is given by:

$$\delta(B,\tau) \;=\; \frac{1}{\int_B f_0(x)\,dx} \int\limits_B \chi_B(\Phi^\tau x)\, f_0(x)\, dx,$$

with χ_B denoting the characteristic function of B, i.e., $\chi_B(x) = 1$ iff $x \in B$ and $\chi_B(x) = 0$ otherwise. Hence, $\delta(B,\tau)$ is the conditional probability in the ensemble that systems being at time t in the phase space region B are found there again at time $t+\tau$. The following generalization of this definition is obvious: For two subsets $B, C \subset \Gamma$ let

$$w(B,C,\tau) \;=\; \frac{1}{\int_B f_0(x)\,dx} \int\limits_B \chi_C(\Phi^\tau x)\, f_0(x)\, dx \tag{10}$$

denote the *transition probability* from B to C. Of course, $\delta(B,\tau) = w(B,B,\tau)$, which is called the probability to stay within B.

Most experiments on molecular systems are performed under the conditions of constant temperature and volume. The corresponding stationary density is the *canonical density* associated with the Hamiltonian H

$$f_0(x) \;=\; \frac{1}{Z}\exp\left(-\beta\, H(x)\right), \quad \text{with} \quad Z = \int_\Gamma \exp\left(-\beta\, H(x)\right) dx,$$

where $\beta = 1/k_B\mathcal{T}$, with $\mathcal{T}$ being the system's temperature $\mathcal{T}$ and k_B Boltzmann's constant. Since H was assumed to be separable, f_0 is a product

$$f_0(x) \;=\; \underbrace{\frac{1}{Z_p}\exp\left(-\frac{\beta}{2}\,p^T M^{-1} p\right)}_{=\mathcal{P}(p)}\; \underbrace{\frac{1}{Z_q}\exp\left(-\beta\, V(q)\right)}_{=\mathcal{Q}(q)}, \tag{11}$$

where we normalize $\mathcal{P}$ and $\mathcal{Q}$ such that

$$\int \mathcal{P}(p)dp = \int \mathcal{Q}(q)dq = 1.$$

In the following we always consider this *canonical ensemble*, i.e., f_0 will always be given by (11).

We have defined conformations as special subsets of phase space. However, the chemical phrase "conformation" does not refer to any momentum information. In other words, the information of chemical interest is *spatial*, i.e., given by the geometry of positions. In order to consider this aspect in our context, assume that the phase space is given by $\Gamma = \Omega \times \mathbb{R}^{3N}$ with $\Omega \subset \mathbb{R}^{3N}$ being the position space. For a spatial subset $B \subset \Omega$ we then define the corresponding phase space cylinder with base B as

$$\Gamma(B) \;=\; \{(q,p) \in \Gamma, \quad q \in B\}.$$

From now on, we are interested only in subsets of this form. For $B, C \subset \Omega$ we use the simplified notation $w(B,C,\tau) \;=\; w(\Gamma(B),\Gamma(C),\tau)$ for the corresponding transition probabilities and denote the probability to be within $B \subset \Omega$ by

$$\pi(B) \;=\; \int\limits_B \mathcal{Q}(q)\,dq \;=\; \int\limits_{\Gamma(B)} f_0(x)\,dx. \tag{12}$$

Summarizing, we have introduced three important modelling steps: determination of the invariant density by preparation of a stationary ensemble, choice of the canonical ensemble representing the typical experimental preparation, and restriction to spatial conformations.

3 Transition Probabilities and Associated Markov Chains

As a consequence of the considerations of the previous section, we now have to replace the Frobenius–Perron operator acting on Γ by another stochastic operator that represents the restriction to purely *spatial* structures within the canonical ensemble. After that, we proceed to the identification of almost invariant sets via the eigenvalue problem of this operator.

3.1 Definition of a Spatial Transition Operator

As will turn out subsequently, an appropriate choice for a stochastic operator is the *spatial transition operator* T defined via momentum weighting due to

$$Tu(q) \;=\; \int u(\Phi^{-\tau}(q,p))\,\mathcal{P}(p)\,dp, \tag{13}$$

where $u = u(q)$ is a function $u : \Omega \to \mathbb{C}$ and $u(\Phi^{-\tau}(q,p))$ means $u(q_1)$ if $(q_1, p_1) = \Phi^{-\tau}(q,p)$. We consider T as an operator on the weighted spaces

$$L^p_{\mathcal{Q}}(\Omega) \;=\; \{u : \Omega \to \mathbb{C}, \quad \int_\Omega |u(q)|^p \, \mathcal{Q}(q)\, dq < \infty\}, \qquad p = 1, 2.$$

Obviously, $L^2_{\mathcal{Q}}(\Omega)$ is a Hilbert space with scalar product

$$\langle u, v \rangle_{\mathcal{Q}} \;=\; \int_\Omega u^*(q)\, v(q)\, \mathcal{Q}(q)\, dq$$

and induced norm $\|u\|^2_{\mathcal{Q}} = \langle u, u \rangle_{\mathcal{Q}}$. With respect to these spaces, the important properties of T are the following (cf. [25]):

1. T is a Markov operator on $L^1_{\mathcal{Q}}(\Omega)$.
2. T is bounded: $\|Tu\|_{\mathcal{Q}} \le \|u\|_{\mathcal{Q}}$.
3. In $L^2_{\mathcal{Q}}(\Omega)$, T is *selfadjoint*, since Φ^τ is *reversible*. Hence, the spectrum $\sigma(T)$ of T is real-valued and bounded: $\sigma(T) \subset [-1, 1]$.
4. For subsets $B, C \subset \Omega$ we find:

$$\langle T\chi_B, \chi_C \rangle \;=\; \int_{\Gamma(B)} \chi_{\Gamma(C)}(\Phi^\tau x)\, f_0(x)\, dx, \tag{14}$$

   showing that T represents the transition probabilities of our interest.

On the basis of Section 2.2, the Frobenius–Perron operator must be replaced by T. In analogy to (6), we once again may hope that the eigenmodes of T for eigenvalues near $\lambda = 1$ allow for an identification of almost invariant sets.

3.2 Spatial Discretization

If we restrict our attention to the weighted Hilbert space $L^2_{\mathcal{Q}}(\Omega)$, we can (as in [8,9]) naturally derive a special Galerkin procedure to discretize the eigenvalue problem $Tu = \lambda u$. Let $B_1, \ldots, B_n \subset \Omega$ be a covering of Ω so that $B_k \cap B_l$ is of (Lebesgue) measure zero for $k \neq l$ and $\cup_{k=1}^n B_k = \Omega$. Then, the sets $\Gamma(B_k)$, $k = 1, \ldots, n$, are a covering of Γ. Our finite dimensional ansatz space $\mathcal{V}_n = \operatorname{span}\{\chi_1, \ldots, \chi_n\}$ is spanned by the associated characteristic functions $\chi_k = \chi_{B_k}$. The Galerkin projection $\Pi_n : L^2_{\mathcal{Q}}(\Omega) \to \mathcal{V}_n$ of $u \in L^2_{\mathcal{Q}}(\Omega)$ is defined by

$$\Pi_n u = \sum_{k=1}^n \frac{1}{\pi(B_k)} \langle \chi_k, u \rangle_{\mathcal{Q}}\, \chi_k.$$

The resulting discretized transition operator $\Pi_n T \Pi_n$ induces the approximate eigenvalue problem $\Pi_n T \Pi_n u = \lambda u$ in $\mathcal{V}_n$. Let λ be one of the corresponding eigenvalues and let the related eigenvector be $u = \sum_{k=1}^n \alpha_k \chi_k$. Then, the discretized eigenvalue problem has the form

$$\sum_{l=1}^n \langle T\chi_k, \chi_l \rangle_{\mathcal{Q}}\; \alpha_l \;=\; \lambda\, \pi(B_k)\, \alpha_k, \qquad \forall k = 1, \ldots, n.$$

After division by $\pi(B_k)$ (known to be positive), we end up with the convenient form

$$P\alpha = \lambda\alpha \quad \text{with} \quad \alpha = (\alpha_1, \ldots, \alpha_n),$$

where in fact the entries of the $n \times n$ matrix P are given by the spatial transition probabilities from B_k to B_l:

$$P_{kl} = \frac{\langle T\chi_k, \chi_l\rangle_Q}{\pi(B_k)} = w(B_k, B_l, \tau). \tag{15}$$

This result finally confirms that (13) was the correct choice of a transition operator in the statistical context.

Since T is a Markov operator, its Galerkin discretization P is a (row) stochastic matrix, i.e., $P_{kl} \geq 0$ and $\sum_{l=1}^{n} P_{kl} = 1$ for all $k = 1, \ldots, n$ (for details about stochastic matrices see [3]). Hence, all its eigenvalues λ satisfy $|\lambda| \leq 1$. Moreover, we have the following three important properties:

1. The row vector $\pi = (\pi_1, \ldots, \pi_n)$, $\pi_k = \pi(B_k)$ denotes the discretized invariant density. Simple calculus reveals that π is a left eigenvector to the eigenvalue $\lambda = 1$, i.e., that $\pi P = \pi$.
2. As shown in [25], P is *irreducible and aperiodic*, which implies, that the eigenvalue $\lambda = 1$ is *simple*. Hence, the discretized invariant density π is the *unique* stationary distribution of P.
3. P is *reversible*, since T is self-adjoint. In other words, P fulfills the condition of *detailed balance*:

$$\pi_k\, P_{kl} = \pi_l\, P_{lk}, \qquad \forall k, l \in \{1, \ldots, n\}.$$

   Therefore, all eigenvalues of P are real-valued: $\sigma(P) \subset [-1, 1]$.

This means that, for arbitrary coverings $B_1, \ldots, B_n \subset \Omega$, the discretization matrices P are inheriting the most important properties of the operator T.

As any stochastic matrix, our discretization matrix P also defines a *discrete Markov chain*, i.e., the stochastic (random) walk of a single system through phase space. The associated statistical interpretation is as follows: If at instance $j \in \mathbb{N}$ the system is in B_k, the probability of finding the system in B_l at instance $j + 1$ is $P_{kl} = w(B_k, B_l, \tau)$. With $j \to \infty$ the system visits all subset B_k with the probability π_k, the value given by the stationary distribution of P.

According to our definition of "almost invariance", we are interested in such unions $B = \cup_{k \in I} B_k$ of our "discretization boxes" B_k, for which the probability $w(B, B, \tau)$ to stay within is sufficiently close to $\delta = 1$. In other words, we are looking for a nontrivial index set $I \subset \{1, \ldots, n\}$ so that the discrete system almost certainly stays within $B = \cup_{k \in I} B_k$ within one single step $j \to j + 1$. As derived in [11], such index sets ("almost invariant aggregates") can be identified via the right eigenvectors of P for eigenvalues close to $\lambda = 1$. After a conformational subset B is identified, the probability

$\delta(B,\tau) = w(B,B,\tau)$ to *stay within* B can easily be computed by virtue of the relation:

$$\delta(B,\tau) = \frac{1}{\sum\limits_{k\in I} \pi_k} \sum_{k,l\in I} \pi_k \, P_{kl}. \tag{16}$$

3.3 Realization via Hybrid Monte Carlo

Up to now, the remaining question is how to compute the matrix P for given boxes B_k. According to (15) we have to determine the transition probabilities between the B_k. This task includes two subproblems:

1. "Sampling of the canonical density": That is, we have to generate a sequence of states $S = \{x_k, \quad k = 1,\ldots,M\} \subset \Gamma$ that is approximately distributed according to f_0.
2. Approximation of the transition probabilities: We will see below that this reduces to counting all such $x_j \in S$ for which $x_j \in B_k$ and $\Phi^\tau x_j \in B_l$. For checking the last condition, sufficient approximations $\tilde{x}_j \approx \Phi^\tau x_j$ of all M subtrajectories starting from S are needed.

The typical approach to sampling the canonical density is via Monte Carlo (MC) techniques. The literature on this topic is extremely rich and varied [5] and we surely will not give particular merits to any special MC variant. We merely suggest to apply a hybrid Monte Carlo (HMC) technique because it seems to be particularly appropriate for linking the above mentioned subproblems 1 and 2. In order to explain this advantage and the basic idea of HMC we have to shortly recall the basic steps of a Metropolis MC approach to sampling f_0 (for additional details see, e.g., [26]).

Each update step $x_j \to x_{j+1}$ of the Metropolis construction of a sampling set S consists of two parts:

1. The *proposal step* $x_j \to \tilde{x}_j$: The numerical realization of the proposal step should exclude any evaluation of f_0 and must yield a final update step which satisfies the detailed balance condition.
2. The *acceptance step*: evaluate $\Delta E = H(\tilde{x}_j) - H(x_j)$ and compute r randomly equidistributed from $[0,1]$. The state $\tilde{x}_j$ is accepted as x_{j+1} if $r \leq \min\{1, \exp(-\beta \Delta E)\}$, otherwise we set $x_{j+1} = x_j$.

So-called "hybrid" MC variants have to our knowledge first been introduced in the late 80's (cf. [12]) and have in MD mostly been used for condensed matter and polymer-like systems (cf. [22,17,15]). The technique fits into the general Metropolis MC framework as a specification of the proposal step for separable Hamiltonians of form (1) (cf. [21]). For explaining this, let $\Psi^{\Delta t}$ denote a reversible and volume-preserving one-step discretization of the flow Φ^t, i.e., of the Hamiltonian equations (2). The reader unfamiliar with this notation may think of $\Psi^{\Delta t}$ as denoting, e.g., the well-known Verlet discretization [27,1] with stepsize Δt. Starting with $x_j = (q_j, p_j)$ the HMC proposal step consists of two substeps: Firstly, choose new initial momenta p_j^0 from the normal distribution $\mathcal{P} = \mathcal{P}(p)$, i.e., from the canonical momentum distribution. Secondly, compute the proposal $\tilde{x}_j$ as

$$\tilde{x}_j \;=\; \left(\Psi^{-\Delta t}\right)^m (q_j, p_j^0), \tag{17}$$

i.e., via $m \in \mathbb{N}$ steps of the discretization $\Psi^{\Delta t}$. Obviously, the momenta are determined randomly before each step. In general, they are *not* distributed according to $\mathcal{P}$ afterwards, so that HMC is to be understood as a *pure position sampling* of the spatial canonical distribution $\mathcal{Q} = \mathcal{Q}(q)$.

The main result of the theoretical justification of *any* Metropolis algorithm (cf. [6]) is the convergence of mean values of the sequences $S = \{x_1, \ldots, x_M\}$ to the corresponding expectation values for $M \to \infty$: For any observable $\mathcal{A} : \Gamma \to \mathbb{R}$ we have asymptotically

$$\left| \frac{1}{M} \sum_{j=1}^{M} \mathcal{A}(x_j) \;-\; \underbrace{\int_{\Gamma} \mathcal{A}(x) f_0(x)\, dx}_{=\langle \mathcal{A} \rangle} \right| \;\leq\; C\, M^{-1/2}, \tag{18}$$

with a constant C not explicitly depending on $\dim(\Gamma) = 6N$. We consider the special observables $\mathcal{A}_k = \chi_{\Gamma(B_k)}$ and $\mathcal{A}_{kl}$ with

$$\mathcal{A}_{kl}(x) \;=\; \chi_{\Gamma(B_k)}(x)\; \chi_{\Gamma(B_l)}(\Phi^\tau x)$$

with the corresponding relations

$$\begin{aligned}
\langle \mathcal{A}_k \rangle &\approx \tfrac{1}{M} \sum_{j=1}^{M} \mathcal{A}_k(x_j) \;=\; \tfrac{1}{M} \left(\text{number of } x_j \in \Gamma(B_k)\right), \\
\langle \mathcal{A}_{kl} \rangle &\approx \tfrac{1}{M} \sum_{j=1}^{M} \mathcal{A}_{kl}(x_j) \;=\; \tfrac{1}{M} \left(\text{number of } x_j \in \Gamma(B_k) \text{ with } \Phi^\tau x_j \in \Gamma(B_l)\right).
\end{aligned}$$

This shows, that we are able to approximate the desired transition probabilities $w(B_k, B_l, \tau) = \langle \mathcal{A}_{kl} \rangle / \langle \mathcal{A}_k \rangle$ "simply" by counting. Since $x_j = (q_j, p_j) \in \Gamma(B_k)$ is equivalent to $q_j \in B_k$, no problems will arise when using HMC for a position sampling only. Moreover, the main advantage of HMC in this context is obvious: we need approximations of $\Phi^\tau x_j$ and can have them "for free" if we use $m\Delta t = \tau$ with sufficiently small Δt in the HMC proposal step (17).

Theoretically, the transition matrix P is reversible. In order to reproduce this property for its approximation, we may simply count each transition from B_k to B_l as a transition $B_l \to B_k$, too (thus exploiting the reversibility of the discretization $\Psi^{\Delta t}$).

As is widely known, MC simulations for ensemble averages may suffer from possible "critical slowing down" [21]. This phenomenon occurs when the iteration $x_k \to x_{k+1}$ gets trapped near a local potential minimum due to high energy barriers so that a proper sampling of the phase space within reasonable computing times is prevented. Typically, this also happens to HMC applications. Therefore, a novel approach combining HMC with the reweighting technique [13,4] has been developed [14]. This HMC variant generates the distribution of a mixed-canonical ensemble composed of two canonical ensembles at low and high temperature. Its analysis shows an efficient sampling of the canonical distribution at the low temperature, whereas the high temperature component facilitates crossing of the crucial energy barriers.

3.4 Essential Degrees of Freedom

Up to now, we have presented the toolbox for the identification of conformations in terms of given spatial discretization boxes $B_k \subset \Omega$. It is a main advantage of the algorithmic concept that the HMC sampling techniques do *not* scale exponentially with the size of the molecular system under consideration. Therefore, the crucial part of the algorithm as it stands now is the solution of the eigenvalue problem for the transition matrix P. Fortunately, we only need the cluster of the largest eigenvalues of P near $\lambda = 1$, which permits us to apply subspace oriented iterative techniques (see, e.g., [24] or [10], Sec. 4.1). But even these more sophisticated approaches scale exponentially with the molecular size if the number of discretization boxes explodes. Hence, the question remains of how to choose the boxes without risking a combinatorial explosion of their number.

Our present strategy is based on chemical insight: In the chemical literature the conformation of a molecule is mostly described in terms of some *essential degrees of freedom*, e.g., certain torsion angles which are of crucial importance for the geometry of what is called the conformation. This means, that most of the spatial degrees of freedom are of minor or no importance for the "characterization" of a conformation (even though they may be of importance for its identification). Consequently, we are allowed to choose our discretization boxes with respect to these essential degrees of freedom only, while our HMC process samples the entire (position) phase space. The example studied in Sec. 4 may serve as an illustration of this procedure.

4 Numerical Experiments

In this section, the performance of the above derived algorithm in application to the n-pentane molecule $CH_3(CH_2)_3CH_3$ is presented. It is illustrated that the algorithmically identified almost invariant sets are in perfect agreement with the chemically observed conformations (cf. Fig. 1).

For modelling the n-pentane molecule, we use the united atom model (cf. Fig. 2) with the typical bound length and bond angle potentials, and a Lennard-Jones potential modelling the interaction between the first and the

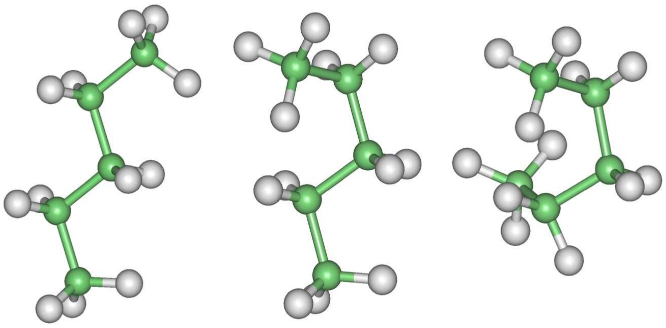

Fig. 1. Different conformations of n-pentane: From the left to the right: trans-trans, trans-gauche, gauche-gauche orientations.

last of the united "atoms". The dihedral angle potentials are chosen according to [23], cf. Fig. 2. The form of the dihedral angle potential shows three different minima corresponding to the trans and gauche orientations of the angles. The vibrational frequencies induced by these potentials are considerably smaller than those induced by the bond interactions. Consequently, the conformations of the n-pentane molecule are described in terms of these dihedral angles, i.e., they are the above mentioned essential degrees of freedom of n-pentane.

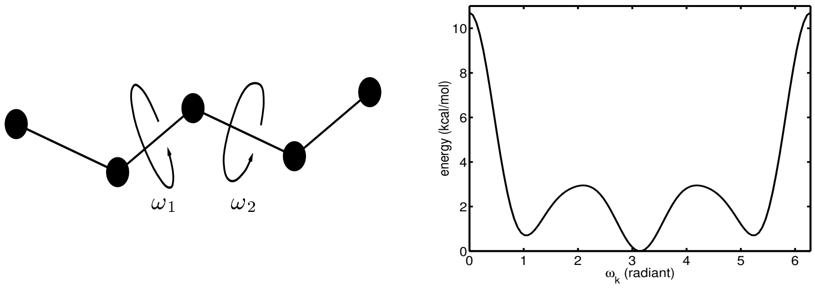

Fig. 2. United atom model of n-pentane with the two dihedral angles ω_1 and ω_2. On the left: Dihedral angle potential due to [23]. The main minimum corresponds to the trans orientation of the angle, the two side minima to the $\pm$gauche orientations.

Figures 3 to 6 below illustrate the execution of the algorithm for the temperature $\mathcal{T} = 300$K. The discretization boxes are constructed via uniform decomposition of the possible values $[0, 2\pi] \times [0, 2\pi]$ of the two dihedral angles ω_1 and ω_2 in $n = 20 \times 20 = 400$ boxes. The HMC sampling has been realized using the Verlet time discretization with a subtrajectory length of $\tau = 160$fs. Fig. 3 shows the resulting sequences of HMC steps in terms of the dihedral angles.

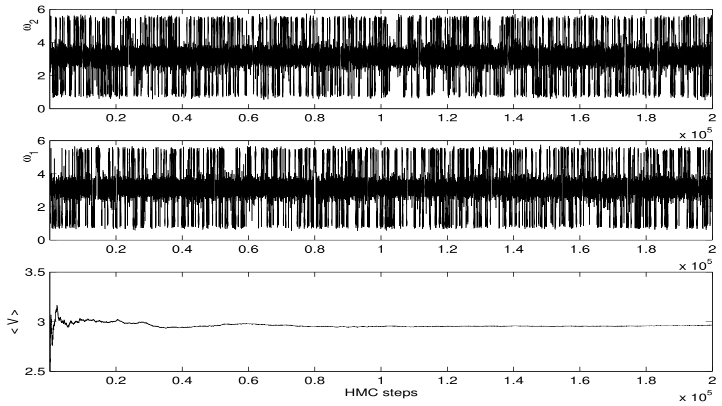

Fig. 3. HMC simulation of n-pentane for $\mathcal{T} = 300$K. From top to bottom: The two dihedral angles (in radiant) versus the step number and the convergence of the potential energy expectation $\langle V \rangle$.

We observe frequent transitions between the different "trans" and "gauche" orientations of both angles. This observation nicely illustrates that it is not sufficient to know the probability to *be within* a particular orientation of the angles but that the essential dynamical information is given by the probability to *stay within* it until a transition into another orientation occurs.

From such an HMC sampling with $M = 200.000$ steps we computed the *transition matrix* P by the procedure explained in Sec. 3.3. Within this sampling length, the HMC method produced a more than sufficient sampling of the canonical density (see the equilibration diagram on bottom of Fig. 3) — the question of whether M could be smaller will be discussed below.

From Sec. 3.2 we know that the discrete invariant density $(\pi(B_k))_{k=1,\dots,n}$ is given by the left eigenvector of P for the largest eigenvalue $\lambda_1 = 1$. The result produced herein is given in Fig. 4. As expected, the invariant density shows distinct local maxima at the minima of the dihedral angle potentials.

Conformations. Following [11], the chemical conformations are analyzed via the right eigenvectors corresponding to an eigenvalue cluster near $\lambda = 1$. A presentation of the derivation of the algorithmic procedure would be beyond the scope of the present paper. We herein can only give a sketch of the construction principle: In a first step, determine the eigenvalue cluster near $\lambda = 1$, which is separated from the remaining part of the spectrum by a significant spectral gap – in our case, these are the seven largest eigenvalues. Fig. 5 shows a schematic plot of the corresponding right eigenvectors. We observe that we may decompose the discretization domain into disjoint regions by distinguishing between different positive, negative, and almost zero values of these eigenvectors. The details of the algorithmic realization are nontrivial,

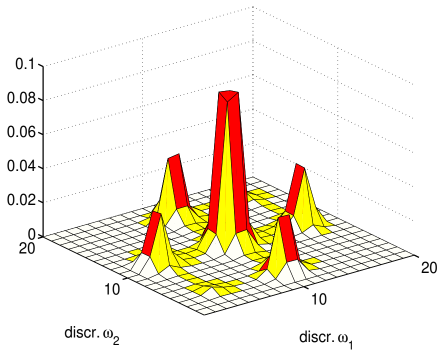

Fig. 4. Discrete canonical distribution for n-pentane versus the indices of the discretization boxes of the two dihedral angles ω_1 and ω_2. $\mathcal{T} = 300$K.

because it has to include an iterative procedure to decide what is "almost zero".

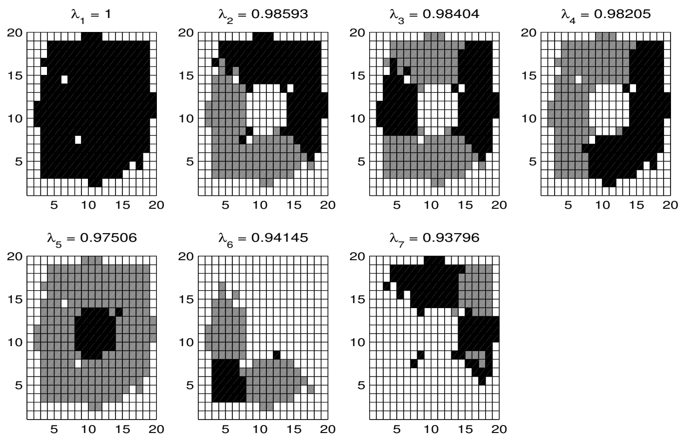

Fig. 5. Schematic plot of the right eigenvectors corresponding to the seven largest eigenvalues $\lambda_1, \ldots, \lambda_7$ of P versus the indices $(1, \ldots, 20) \times (1, \ldots, 20)$ of the discretization boxes of the two dihedral angles ω_1 and ω_2. Positive entries of the eigenvectors are indicated by black boxes, negative entries by gray boxes and white boxes indicate almost zero entries. $\mathcal{T} = 300$K.

By analyzing the eigenvectors as illustrated, the algorithm from [11] identifies the conformational subsets shown in Fig. 6. As can be seen the automatic procedure in fact supplies the chemically expected information. After identifying the conformations, the corresponding probabilities to stay within each conformational subset can be computed due to equation (16). The resulting values p are also given in Fig. 6. We observe that the trans/trans conformation is slightly more stable than the different trans/gauche and gauche/trans conformations. As expected, the two gauche/gauche conformations are clearly less stable.

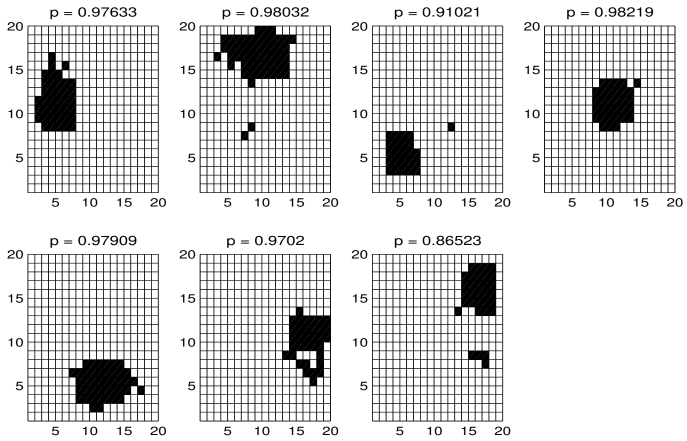

Fig. 6. Almost invariant sets for $\mathcal{T} = 300$K. The numbers p on top of each figure are the probabilities to stay within the corresponding subsets during the time span τ. From the left hand side on top to the right hand side below we see the -gauche/trans, trans/+gauche, -gauche/-gauche, trans/trans, trans/-gauche, +gauche/trans, and +gauche/+gauche conformations (cf. Fig. 1).

As already emphasized above, the probabilities to *stay within* should *not* be confused with the probability to *be within* a conformation, which is already given by the invariant density (cf. Fig. 4). In the table below, these two different probabilities are enlisted for each of the conformational subsets shown in Fig. 6 (±g and t denote the ±gauche and trans orientations):

conformation	-g/t	t/+g	-g/-g	t/t	t/-g	+g/t	+g/+g
prob. to be within	0.120	0.132	0.012	0.473	0.117	0.132	0.013
prob. to stay within	0.976	0.980	0.910	0.982	0.979	0.970	0.865

The slight differences between the probabilities to be within the ±g/t and t/±g orientations may be used as an error indicator for the sampling. The probability to be within the +gauche/-gauche or -gauche/+gauche orientations is less than 0.0005, showing that they are irrelevant in this context.

Parameter Sensitivity. The results presented herein surely depend on a number of crucial parameters, some of them being of physical nature (e.g., the temperature $\mathcal{T}$), others being introduced by the algorithm (e.g., the number n of discretization boxes or the length M of the HMC sampling). We want to emphasize that the algorithm as it stands now is far from being perfectly tuned. We thus can only present some experiences from numerical experiments for the n-pentane molecule and some other comparably small systems.

At first, let us consider the dependence of the conformations on the temperature $\mathcal{T}$. Varying the temperature between $\mathcal{T} = 200$K and $\mathcal{T} = 600$K we do not observe an influence on the identified conformations. But, as to be expected, the probabilities to stay within these conformations are decreasing with increasing $\mathcal{T}$: Fig. 7 shows the corresponding decrease of the nine largest eigenvalues of the transition matrices $P = P(\mathcal{T})$. It also illustrates thatin all cases tested so far there exists a distinct *spectral gap* between the seven largest eigenvalues used to identify the conformational subsets, and the remaining part of the spectrum.

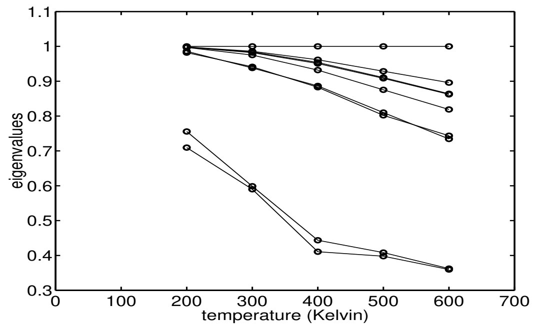

Fig. 7. Temperature dependence of the nine largest eigenvalues of the transition matrix P.

The present version of the algorithmic realization does not include any automatic mechanism for controlling the length M of the HMC sampling. If, for fixed temperature and spatial discretization, the number of steps is decreased from $M = 200.000$ down to $M = 50.000$, we observe that the approximation quality of the invariant density slowly deteriorates. This corresponds to a slowly increasing distortion of the approximate "conformational" subsets.

Dependence on Discretization. Finally, let us illustrate an extremely important property of the presented algorithm, the stability of the results even when significantly coarser discretizations are used. For the n-pentane molecule we indeed can reduce the decomposition of the discretization domain from $n = 20 \times 20$ boxes to $n = 3 \times 3$ boxes but the algorithm still identifies approximately the same conformations and nearly the same probabilities (both to stay and to be within). The reason for this is illustrated in Fig. 8: since the HMC procedure samples the phase space independent of the discretization, the seven largest eigenvalues of the transition matrix P

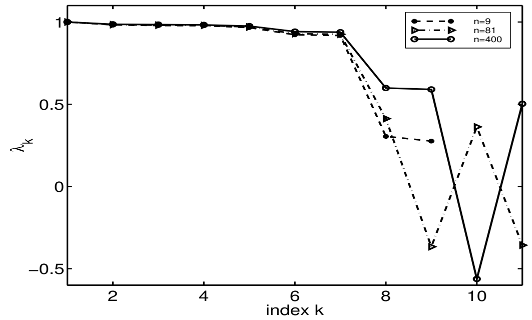

Fig. 8. Sensitivity of the *absolutely* largest eigenvalues of P for different uniform discretizations of $[0, 2\pi]^2$ with $n = 3 \times 3 = 9$ boxes (dashed line), $n = 9 \times 9 = 81$ (dashed-dotted), and $n = 20 \times 20 = 400$ boxes (dense line). Note that the seven largest eigenvalues – only these are used for the identification of the conformations – remain almost unperturbed if the grid gets coarser.

are only insignificantly perturbed when the number of discretization boxes is reduced.

Acknowledgments. It is a pleasure to thank Frank Cordes for many helpful discussions. W.H. was supported within the DFG–Schwerpunkt "Ergodentheorie, Analysis und effiziente Simulation dynamischer Systeme" under Grant De 293/2-1. Moreover, we are grateful to the "éminence grise" Folkmar Bornemann for his suggestion of the more elucidatory representation in Fig. 5.

References

1. M.P. Allen and D.J. Tildesley. *Computer Simulations of Liquids*. Clarendon Press, Oxford, 1990.
2. A. Amadei, A.B.M. Linssen, and H.J.C. Berendsen. Essential dynamics on proteins. *Proteins*, 17, 1993.
3. A. Berman and R. J. Plemmons. *Nonnegative Matrices in the Mathematical Sciences*. Academic Press, New York, 1979. Reprinted by SIAM, Philadelphia, 1994.
4. B.J. Berne and J.E. Straub. Novel methods of sampling phase space in the simulation of biological systems. *Current Opinion in Structural Biology*, 7:181–189, 1997.
5. K. Binder[Hsg.]. *The Monte Carlo method in condensed matter physics*, volume Bd. 71 of *Topics in applied physics*. Springer Verlag, Berlin u.a., 1992.
6. K.L. Chung. *Markov chains with stationary transition probabilities*, volume 104 of *Grundlehren der mathematischen Wissenschaften*. Springer, Berlin, Göttingen, Heidelberg, 1960.

7. M. Dellnitz and A. Hohmann. The computation of unstable manifolds using subdivision and continuation. In H.W. Broer, S.A. van Gils, I. Hoveijn, and F. Takens, editors, *Nonlinear Dynamical Systems and Chaos*, pages 449–459. Birkhäuser, *PNLDE* **19**, 1996.
8. M. Dellnitz and O. Junge. On the approximation of complicated dynamical behavior. *To appear in SIAM J. Num. Anal.*, 1998. Available via http://www.uni-bayreuth.de/departments/math/org/mathe5/publ/papers/.
9. P. Deuflhard, M. Dellnitz, O. Junge, and Ch. Schütte. Computation of essential molecular dynamics by subdivision techniques. To appear in Proceedings of the 2nd International Symposium "Algorithms for Macromolecular Modelling", P. Deuflhard and J. Hermans and B. Leimkuhler and A. Mark and B. Skeel and S. Reich (eds.), Springer-Verlag, 1998. Available via http://www.zib.de/schuette/.
10. P. Deuflhard, T. Friese, and F. Schmidt. A nonlinear multigrid eigenproblem solver for the complex Helmholtz equation. Preprint SC 97-55, Konrad Zuse Zentrum, Berlin, 1997. Available via http://www.zib.de/bib/pub/pw/.
11. P. Deuflhard, W. Huisinga, A. Fischer, and Ch. Schütte. Identification of almost invariant aggregates in reversible nearly uncoupled Markov chains. Preprint SC 98-03, Konrad Zuse Zentrum, Berlin, 1998. Available via http://www.zib.de/bib/pub/pw/.
12. S. Duane, A.D. Kennedy, B.J. Pendleton, and D. Roweth. Hybrid Monte Carlo. *Phys. Letters B*, 195(2):216–222, 1987.
13. A.M. Ferrenberg and R.H. Swendsen. New Monte Carlo technique for studying phase transitions. *Phys. Rev. Letters*, 61(23):2635–2638, 1988.
14. A. Fischer, F. Cordes, and Ch. Schütte. A hybrid Monte Carlo method in a mixed-canonical ensemble for efficient conformational analysis of RNA. Preprint SC 97-67, Konrad Zuse Zentrum, Berlin, 1997. Available via http://www.zib.de/bib/pub/pw/.
15. B.M. Forrest and U.W. Suter. Hybrid Monte Carlo simulations of dense polymer systems. *J. Chem. Phys.*, 101(3):2616–2629, 1994.
16. H. Grubmueller and P. Tavan. Molecular dynamics of conformational substates for a simplified protein model. *J. Chem. Phys.*, 101, 1994.
17. D.W. Heermann and L. Yixue. A global-update simulation method for polymer systems. *Makromol. Chem., Theory Simul.*, 2:299–308, 1993.
18. B.O. Koopman. Hamiltonian dynamics and transformations in Hilbert space. *Proc. Nat. Acad. Sci.*, 17, 1931.
19. R. Kurt. *Axiomatics of Classical Statistical Mechanics.* Pergamon Press, Oxford, New York, 1980.
20. A. Lasota and C. Mackey. *Chaos, Fractals and Noise.* Springer, 1994.
21. E. Leontidis, B.M. Forrest, A.H. Widmann, and U.W. Suter. Monte Carlo algorithms for the atomistic simulation of condensed polymer phases. *J. Chem. Soc. Faraday Trans.*, 91(16):2355–2368, 1995.
22. B. Mehlig, D.W. Heermann, and B.M. Forrest. Hybrid Monte Carlo method for condensed-matter systems. *Phys. Review B*, 45(2):679–685, 1992.
23. J.-P. Ryckaert and A. Bellemans. Molecular dynamics of liquid n-butane near its boiling point. *Chem. Phys. Letters*, 30(1):123–125, 1975.
24. Y. Saad. *Numerical Methods for Large Eigenvalue Problems.* Manchester University Press, Manchester, 1992.
25. Ch. Schütte. A spatial transition operator for essential molecular dynamics. Preprint in preparation, 1998.

26. A.D. Sokal. Monte Carlo methods in statistical mechanics. Lecture note, Department of Physics, New York University, 1989.
27. L. Verlet. Computer experiments on classical fluids. Part I. *Phys. Rev.*, 159:98–103, 1967.
28. P. Walters. *An Introduction to Ergodic Theory.* Springer, 1991.