Konrad-Zuse-Zentrum
für Informationstechnik Berlin

Takustraße 7
D-14195 Berlin-Dahlem
Germany

PETER DEUFLHARD WILHELM HUISINGA
ALEXANDER FISCHER CHRISTOF SCHÜTTE

Identification of Almost Invariant Aggregates in Reversible Nearly Uncoupled Markov Chains

Preprint SC 98-03 (February 1998)

Identification of Almost Invariant Aggregates in Reversible Nearly Uncoupled Markov Chains

P. Deuflhard W. Huisinga A. Fischer Ch. Schütte

Abstract

The topic of the present paper has been motivated by a recent computational approach to identify chemical conformations and conformational changes within molecular systems. After proper discretization, the conformations show up as almost invariant aggregates in reversible nearly uncoupled Markov chains. Most of the former work on this subject treated the *direct* problem: given the aggregates, analyze the loose coupling in connection with the computation of the stationary distribution (aggregation/disaggregation techniques). In contrast to that the present paper focuses on the *inverse* problem: given the system as a whole, identify the almost invariant aggregates together with the associated transition probabilities. A rather simple and robust algorithm is suggested and illustrated by its application to the n-pentane molecule.

Key words. essential molecular dynamics, nearly reducible, nearly completely decomposable, nearly uncoupled Markov chain, almost invariant aggregates, transition probability.

Mathematics subject classification. 15A18, 15A51, 60J10, 60J20, 65U05.

1 Introduction

The present investigation has been motivated by a novel approach to compute the *essential* features of molecular dynamical systems, as suggested first in [3] and, in a much improved form, more recently in [12]. In these approaches *chemical conformations* are interpreted as almost invariant sets either in the phase space [3] or in the position space [12] of the associated Hamiltonian dynamical system. Both conformations and rates of conformational changes need to be identified to obtain information of chemical interest. Any discretization gives rise to finite dimensional *Markov chains* and to almost invariant aggregates as discrete analogs of the above mentioned almost invariant sets. Given such a Markov chain (usually in terms of the corresponding transition matrix), the task is to identify an *unknown* number k of almost invariant aggregates in nearly uncoupled Markov chains.

In the setting of [12], the obtained Markov chain is *reversible* and *regular*, or, equivalently, the corresponding transition matrix is *reversible* and *primitive*—an assumption, which will be crucial for the construction of the algorithm to be presented below. The reversibility implies that all eigenvalues λ are real and $|\lambda| \leq 1$. In the *irreducible* case, Perron–Frobenius theory states that $\lambda = 1$ is simple and a corresponding left eigenvector can be chosen to have only *positive* components. From [3] we conjecture that eigenvectors corresponding to an eigenvalue cluster close to $\lambda = 1$ contain the relevant information of how to decompose the discrete state space into reasonable almost invariant aggregates: there, the case $k = 2$ has already been worked out in some "Gedankenexperiment" based on *left* eigenvectors corresponding to two eigenvalues $\lambda_1 = 1$ and $\lambda_2 \approx 1$. A possible treatment of the general case $k \geq 2$ has been demonstrated in [2] for the different scenario of *almost cyclic* aggregates, whereas the case of almost invariant aggregates is only indicated. In contrast to that, the present paper tries to generalize the concept of [3] to $k > 2$ aiming at a comparable simplicity of both theory and algorithm.

In Section 2, we first treat the *reducible* case of an *uncoupled* Markov chain in terms of the block structure of the corresponding transition matrix. In our application context, such a structure is only present in a perturbed form and, additionally, hidden due to permutations. That is why, in Section 3, we study the case of *nearly uncoupled* Markov chains in terms of a linear *perturbation analysis* for the transition matrix—following closely former work of STEWART [15]. On this basis, we are able to extend the above mentioned "Gedankenexperiment" [3] based on *right* eigenvectors corresponding to k eigenvalues close to 1. From this, we derive a rather simple and robust algorithm in Section 4. Its application to the n-pentane molecule is finally illustrated in Section 5.

2 Uncoupled Markov Chains

Throughout the paper, the term Markov chain will be used to denote a finite homogeneous Markov chain defined by a finite set of states $\{s_1, \ldots, s_n\}$ and a (row) stochastic matrix P, the transition matrix. Sets of states are called *aggregates*.

As stated above, we are concerned with reversible regular Markov chains, which bear some hidden structure of almost invariant aggregates. In order to classify the term "almost invariant", we first treat the situation of *invariant* aggregates and *uncoupled* Markov chains here. For this purpose, we need to collect some important definitions together with spectral properties of the transition matrix.

A Markov chain is called *regular*, if the corresponding transition matrix P is *primitive* (irreducible and aperiodic), i.e., if there exists a natural number $m > 0$ such that P^m is componentwise positive [1, 13]. For primitive stochastic matrices the Perron–Frobenius theorem gives the following insight about the eigenvalue of largest modulus:

Theorem 2.1 [1, 13] *Let P be a primitive stochastic matrix. Then*

1. *the so–called Perron root $\lambda = 1$ is simple and dominant, i.e., $|\lambda| < 1$ for any other eigenvalue $\lambda \neq 1$.*

2. *there are positive left and right eigenvectors corresponding to $\lambda = 1$, which are unique up to constant multiples.*

The *left* eigenvector corresponding to $\lambda = 1$ represents the *stationary distribution* $\pi = (\pi_1, \ldots, \pi_n)^T$, which in our application context [12] is known, while the corresponding *right* eigenvector is $e = (1, \ldots, 1)^T$.

Due to the discretization [12] we may assume that the the Markov chain is *reversible*, i.e., the corresponding transition matrix $P = (p_{ij})$ satisfies[1]

$$\pi_i \, p_{ij} \quad = \quad \pi_j \, p_{ji} \qquad \text{for all } i, j. \tag{1}$$

Then, P is then called *reversible* [8]. The spectral structure of reversible stochastic matrices is most evident with respect to the inner product $\langle \cdot, \cdot \rangle_\pi$ induced by the stationary distribution:

$$\langle x, y \rangle_\pi \quad = \quad x^T \mathrm{diag}(\pi_i) \, y,$$

which corresponds to the finite dimensional *weighted* Euclidean space $l_\pi^2(n)$. Two vectors x, y satisfying $\langle x, y \rangle_\pi = 0$ will be called *π–orthogonal*.

Lemma 2.2 *Let P be a reversible primitive stochastic matrix. Then P is symmetric with respect to the inner product $\langle \cdot, \cdot \rangle_\pi$.*

Proof. It is shown in [6] that $P = (p_{ij})$ is equivalent to the symmetric matrix

$$P_{\mathrm{sym}} \quad = \quad DPD^{-1}, \tag{2}$$

where the transformation is given by $D = \mathrm{diag}(\sqrt{\pi_i})$. In terms of D, equation (1) can be written as $D^2 P = P^T D^2$, thus

$$\langle x, Py \rangle_\pi \; = \; x^T \mathrm{diag}(\pi_i) \, Py \; = \; x^T D^2 P y \; = \; x^T P^T D^2 y \; = \; \langle Px, y \rangle_\pi,$$

which concludes the assertion. □

As a consequence of Lemma 2.2 the stochastic matrix P possesses the following properties:

1. There exists a basis of π–orthogonal right eigenvectors, which diagonalizes P.

2. All eigenvalues of P are real and contained in the interval $[-1, +1]$.

3. For every *right* eigenvector x there is an associated *left* eigenvector $y = \mathrm{diag}(\pi_i) \, x$, which corresponds to the same eigenvalue.

[1] Equation (1) is also called the *detailed balance* condition.

To keep the notion as simple as possible, from now on eigenvectors are understood to be right eigenvectors, unless stated otherwise.

We are now ready to introduce the term *transition probabilities* between aggregates for an arbitrary transition matrix $P = (p_{ij})$. Interpreting the entries p_{ij} as conditional probabilities to be in state s_i and to change to state s_j, the natural generalization is the following

Definition 2.3 *Given a Markov chain by its transition matrix P (not necessary primitive) and a stationary distribution $\pi > 0$. Let A and B be two arbitrary aggregates. Then the* transition probability *between A and B with respect to π, i.e., the probability that the system will move (in a single step) from A to B, is given by*

$$w(A, B) \quad = \quad \frac{\sum\limits_{a \in I_A, b \in I_B} \pi_a \, p_{ab}}{\sum\limits_{a \in I_A} \pi_a},$$

where I_A and I_B denote the index sets corresponding to A and B, respectively. For the special case $A = B$, we call $w(A, A)$ the probability to stay within A *(in a single step).*

A statistical characterization of uncoupled Markov chains (UMC) will be based on transition probabilities between aggregates. An aggregate A is said to be *invariant*, if $w(A, A) = 1$[2]. A Markov chain is then called *uncoupled*, if it allows to decompose the state space into disjoint *invariant* aggregates $A_1, \ldots, A_k$ such that

$$w(A_i, A_j) \quad = \quad \delta_{ij}.$$

As a consequence the states of an UMC with k aggregates *can be ordered* such that the transition matrix P is of *block–diagonal* form

$$P \; = \; D \; = \begin{pmatrix} D_{11} & 0 & \cdots & 0 \\ 0 & D_{22} & \cdots & 0 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 0 & 0 & \cdots & D_{kk} \end{pmatrix},$$

where each block D_{ii} is a square stochastic matrix.

Assume that each of these matrices D_{ii} is primitive. Then, due to the Perron–Frobenius theorem, each block possesses a unique eigenvector $e_i = (1, \ldots, 1)^T$ of length $\dim(D_{ii})$ corresponding to its Perron root $\lambda_i = 1$. Therefore in terms of the

[2]Note that in the case of an UMC, the stationary distribution is not unique, because the corresponding transition matrix is not irreducible. However, in this special case the probabilities are independent of the chosen stationary distribution. If the Markov chain is not reversible, it is necessary to require $w(A^c, A) = 0$, too.

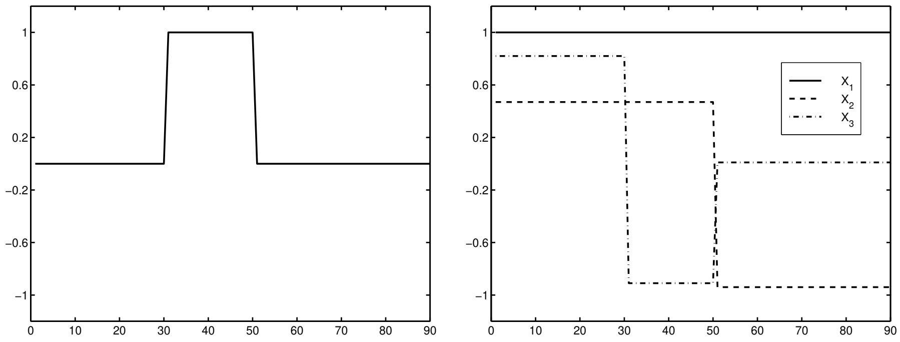

Figure 1: Uncoupled Markov chain with $k = 3$ aggregates. The state space $\{s_1, \ldots, s_{90}\}$ divides into the aggregates $A_1 = \{s_1, \ldots, s_{29}\}$, $A_2 = \{s_{30}, \ldots, s_{49}\}$ and $A_3 = \{s_{50}, \ldots, s_{90}\}$. Left: Characteristic function χ_{A_2}. Right: a possible basis of the eigenspace corresponding to $\lambda = 1$. Observe that each eigenvector is constant on each aggregate. The sign structure associated, e.g., with the state s_{69} is $(+, -, 0)$.

transition matrix P the eigenvalue $\lambda = 1$ is k–fold and the corresponding eigenspace is spanned by the vectors

$$\chi_{A_i} \quad = \quad (0, \ldots, 0, e_i^T, 0, \ldots, 0)^T, \qquad i = 1, \ldots, k.$$

Here our notation deliberately emphasizes that the eigenvectors can be interpreted as *characteristic functions* of the uncoupled aggregates (see Fig. 1, left).

In general, any basis $\{X_i\}_{i=1,\ldots,k}$ of the eigenspace corresponding to $\lambda = 1$ can be written as linear combinations of the characteristic functions χ_{A_i}, i.e., there are coefficients $\alpha_{ij} \in \mathbf{R}$ such that

$$X_i \quad = \quad \sum_{j=1}^{k} \alpha_{ij}\, \chi_{A_j}, \qquad i = 1, \ldots k.$$

As a consequence, eigenvectors corresponding to $\lambda = 1$ are *constant on each aggregate* (see Fig. 1, right). With these preparations we are now ready to derive the key tool for our algorithm to be presented in Section 4.

Lemma 2.4 *Given a block–diagonal transition matrix P consisting of primitive blocks and a π–orthogonal basis $\{X_i\}_{i=1,\ldots,k}$ of its eigenspace corresponding to $\lambda = 1$. Associate with every state s_i its sign structure*

$$s_i \quad \longmapsto \quad (\mathrm{sign}((X_1)_i), \ldots, \mathrm{sign}((X_k)_i)). \tag{3}$$

Then

1. *invariant aggregates are collections of states with common sign structure,*
2. *different aggregates exhibit different sign structures.*

Proof. In order to prove statement 1, recall that each eigenvector corresponding to $\lambda = 1$ is constant on each of the aggregates, which implies that states belonging to the same aggregate must share the same sign structure.

As for statement 2, let, without loss of generality, every aggregate consist of only one state. In a first step, we demonstrate the assertion for an orthogonal eigenvector basis $\{Q_i\}_{i=1,\dots,k}$ of the symmetric matrix $P_{\text{sym}} = DPD^{-1}$ (see (2)). In a second step, we then generalize it to the assertion stated in the proposition.

Define the $k \times k$ matrix $Q = [Q_1 \cdots Q_k]$. Since Q is orthogonal, i.e., $Q^T = Q^{-1}$, the transpose Q^T is an orthogonal matrix, too. Thus, the rows of Q are orthogonal, a fact that we will exploit in the following.

Now consider a π–orthogonal eigenvector basis $\{X_i\}_{i=1,\dots,k}$ of P as stated in the proposition. In view of (2), we get the following relation between the eigenvector bases: $X_i = D^{-1}Q_i$ for $i = 1, \dots, k$. Since the transformation matrix D^{-1} has positive diagonal entries, the sign structures of X_i and Q_i, $i = 1, \dots, k$, are the same.

In view of (3), the sign structure of the mth aggregate is equal to the sign structure of the mth row of $X = [X_1 \cdots X_k]$. Now suppose there exist two aggregates A_i and A_j with the same sign structure. Then the ith and jth row of X, and thus of Q, are equal in sign, which is in contradiction to the orthogonality of Q. □

Summarizing, Lemma 2.4 states that the set of all k eigenvectors can be used to *identify* all aggregates via sign structures. Note that this can also be done by using *left* eigenvectors instead of *right* eigenvectors, since their sign structures are the same: For every left eigenvector $y = (y_i)$ there exits an associated right eigenvector $x = (x_i)$ with $y_i = \pi_i x_i$, hence $\text{sign}(y_i)=\text{sign}(x_i)$.

3 Perturbation Analysis

In the context of our molecular dynamics applications, nearly uncoupled Markov chains[3] (NUMC) will arise such that the corresponding transition matrix P bears some hidden structure of an (unknown) number k of almost invariant aggregates. Despite the unknown permutations and perturbations, we will show in this section that eigenvectors of P can nevertheless be used to identify such aggregates. As for the associated perturbation analysis, we closely follow the framework of Stewart [15].

In this section, we assume the Markov chain to be reversible and primitive. Therefore, in particular, its stationary distribution π is unique (see the note following Theorem 2.1), and all transition probabilities are well–defined with respect to this π.

As in the uncoupled case, a statistical characterization of NUMCs will be based on *transition probabilities* between aggregates. An aggregate A is said to be *almost invariant* , if the probability to stay in A under the condition of being in A is close to 1, i.e., $w(A, A) \approx 1$[4]. A Markov chain is then called *nearly uncoupled*, if it allows

[3] Also known as nearly completely decomposable or nearly reducible Markov chains.

[4] If the Markov chain is not reversible, it is necessary to require $w(A^c, A) \approx 0$, too.

to decompose the state space into disjoint *almost invariant* aggregates $A_1, \dots, A_k$ such that

$$w(A_i, A_j) \approx \delta_{ij}.$$

As a consequence, the states of a NUMC with k aggregates *can be ordered* such that the transition matrix P is of *block–diagonally dominant* form

$$P = D + E = \begin{pmatrix} D_{11} & E_{12} & \cdots & E_{1k} \\ E_{21} & D_{22} & \cdots & E_{2k} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ E_{k1} & E_{k2} & \cdots & D_{kk} \end{pmatrix}, \tag{4}$$

where the off–diagonal blocks E_{ij} are small compared with the diagonal blocks D_{ii}. For later reference, we set $\|E\| = \epsilon$ in terms of the spectral norm.

For the transition matrix (4), we assume according to STEWART [15, Condition 1 and 2] the following

Regularity condition. In the limit $\epsilon \to 0$ the number k of blocks of the transition matrix P remains constant.

This condition implies that for $\epsilon \to 0$ the spectrum of P can be divided into three parts [10, 15]:

1. the Perron root $\lambda = 1$,
2. a cluster of $k-1$ eigenvalues approaching 1 in the limit case $\epsilon \to 0$ and
3. the remaining part of the spectrum, which is bounded away from 1.

In particular, for small ϵ there is a spectral gap between the eigenvalue cluster and the remaining part of the spectrum.

Remark. The above characterization of nearly uncoupled Markov chains via transition probabilities between aggregates is compatible both with the *coupling matrix* defined in [10, 15] as well as with the concept of the *conductance* of a Markov chain [14]. It is, however, quite different from the approach of HARTFIEL AND MEYER [7]: their *uncoupling measure* does not admit a statistical interpretation and depends heavily on the dimension of the transition matrix.

The following perturbation theorem is a specification of a theorem due to STEWART [15, Theorem 4.1] reformulated for our present context.

Theorem 3.1 *Let P be a reversible primitive stochastic matrix satisfying the above regularity condition. Then there exists a π–orthogonal basis $\{\tilde{X}_i\}_{i=1,\dots,n}$ of eigenvectors, which can be divided into three parts:*

1. an eigenvector corresponding to the Perron root $\lambda = 1$ given by

$$\tilde{X}_1 \quad = \quad e = (1, \dots, 1)^T,$$

2. a set of $k-1$ eigenvectors corresponding to the eigenvalue cluster near $\lambda = 1$ of the form

$$\tilde{X}_i \quad = \quad \sum_{j=1}^{k} \alpha_{ij} \, \chi_{A_j} + \mathcal{O}(\epsilon) \qquad i = 2, \dots k$$

for appropriate coefficients $\alpha_{ij} \in \mathbf{R}$ and aggregates $A_1, \dots, A_k$ corresponding to the block–diagonally dominant form (4) of P, and

3. the remaining $n-k$ eigenvectors corresponding to the spectrum bounded away from 1, which cannot be interpreted as perturbations of vectors that are constant on aggregates.

Remark. Exploiting the fact that the transition matrix P admits a complete basis of eigenvectors (see Lemma 2.2) the above theorem is just a specification of the more general Theorem 4.1 of [15]. The reader might notice that, because of Lemma 2.2, the present statement may also be proved by exploiting the well–developed spectral theory of symmetric matrices. In particular, this would allow to gain additional information about the $\mathcal{O}(\epsilon)$-error term.

By Theorem 3.1, eigenvectors corresponding to the eigenvalue cluster near $\lambda = 1$ essentially preserve the structure of the unperturbed case. Therefore we may as well use the sign structures to identify almost invariant aggregates as presented in the previous section.

MATLAB–Example. To illustrate Theorem 3.1 we define the following reversible primitive stochastic matrix P with $k = 3$ blocks; the notation corresponds to (4).

Generate a *symmetric* block diagonal matrix D and a *symmetric* perturbation matrix E, both with equidistributed random entries. Now define for $\delta > 0$ the symmetric matrix

$$P_{\text{sym}} \quad = \quad (1-\delta)D + \delta E.$$

A short calculation shows that normalizing the rows of P_{sym} results in a *reversible* stochastic matrix $P = (p_{ij})$; its stationary distribution π is given by the normalized sum of the ith row of $P_{\text{sym}} = ((p_{\text{sym}})_{ij})$:

$$\pi_i \quad = \quad \sum_{j=1}^{n} (p_{\text{sym}})_{ij} / \|P_{\text{sym}}\|_1.$$

If at least one of the diagonal entries p_{ii} is different from zero (which is easy to check) than the matrix is *primitive*, too [1]. Figure 2 shows an associated eigenvector basis $\{X_1, X_2, X_3\}$ of such a model matrix.

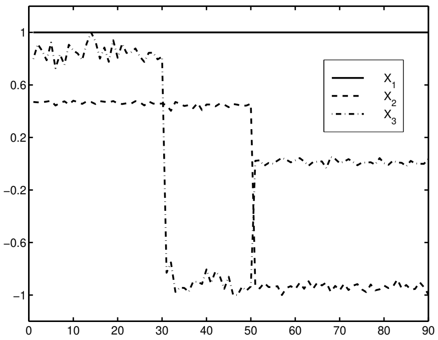

Figure 2: Illustration of Theorem 3.1 (see the example). As a result of the perturbation process the eigenvectors X_1, X_2 and X_3 corresponding to $\lambda = 1, 0.75$ and 0.52, respectively, are almost constant on the aggregates A_1, A_2 and A_3 (cf. Fig. 1). The sign structure of the eigenvectors is the same as in Fig. 1 except for X_3 on the aggregate A_3, where perturbations introduce "erratic" sign structures.

4 Algorithmic Realization

In this section we present the basic concept of an algorithm for the identification of almost invariant aggregates. As derived above, this algorithm explores the special structure of eigenvectors corresponding to an eigenvalue cluster near $\lambda = 1$.

In a first step we have to determine the number k of almost invariant aggregates. This is done by computing the cluster of eigenvalues near $\lambda = 1$ which is well-separated from the remaining part of the spectrum by a gap (Theorem 3.1). Iterative eigenvalue solvers with simultaneous subspace iteration (see e.g. [9],[4, Section 4]) would be a natural way to perform this task. Frankly speaking, however, in our present version of the algorithm we have simply applied MATLAB to calculate all eigenvalues and split off the cluster near 1 by examination.

Having computed the $k - 1$ right eigenvectors (apart from the already known eigenvector e), each of which corresponds to an eigenvalue of the cluster, we are then interested in a decomposition of the state space into k almost invariant aggregates. For an *uncoupled* Markov chain this could be merely done by aggregating states according to the piecewise constant levels in the eigenvectors or due to their sign structure (see Section 2). Unfortunately, for *nearly uncoupled Markov chains* perturbations of the eigenvectors disturb their piecewise constant "level structure". Moreover, we a priori do not know the specific permutation for bringing the transition matrix in a block-diagonally dominant form (4). Both together, unknown permutations and unsuitable numbering of the states, prevent us from exploiting the otherwise intriguing level structure.

The *sign structure*, however, is also perturbed, but only by the "almost zero" entries of some eigenvectors (see Fig. 2). Fortunately, the algorithm to be presented below allows to recover a unique decomposition of the state space by first treating all almost zero entries as optional positive or negative signs (resulting in $k+l$ provisional aggregates) and in a second step by an iterative condensation to k aggregates.

For this purpose, we introduce an ε-threshold and disregard any signs for all values less than ε (in modulus). Therefore, all states which have equal sign structure in all components greater than ε, can be assigned to the same aggregate. This assignment need not be unique, but by repeatedly increasing ε "artificial" aggregates will be removed. This iterative process terminates as soon as the state space is decomposed into exactly k aggregates. Afterwards we search for unique assignments of the ambiguous states by decreasing ε again.

Note that any *scaling* procedure for the eigenvectors will greatly affect the performance of the algorithm. In our numerical experiments it turned out that the assignment process to the final almost invariant aggregates is especially difficult for entries of the eigenvectors near a jump from one almost constant level to another. In order to make our algorithm less sensitive to such assignment problems, we decided to scale the right eigenvectors $\{X_i\}_{i=1,\ldots,k}$ componentwise using a fractional power r with $0 \leq r \leq 1$ ($r = 0.1$ throughout Section 5) of the stationary distribution such that

$$X_i^{\text{scal}} = \text{diag}(\pi_i^r)\, X_i \quad \text{for} \quad i = 1, \ldots, k.$$

For $r = 0$ we then obtain the *right* eigenvectors, whereas for $r = 1$ we get the *left* eigenvectors (see the Remark following Lemma 2.2). The effect of this scaling is that the above mentioned problematic perturbations will get smeared out. Moreover, to be less dependent on aggregate sizes and absolute values of eigenvectors, we partitioned each scaled eigenvector into its positive and negative part and normalized these parts by their maximum norm.

Summarizing, our identification algorithm consists of the following steps:

1. *Compute eigenvectors corresponding to the cluster of eigenvalues near* 1*. This includes the determination of k.*
2. *Scale these eigenvectors.*
3. *Partition the state space into aggregates according to all occurring sign structures, thus generating $k+l$ aggregates.*
4. *If $l > 0$ gradually remove all "artificial" aggregates.*
5. *Compute the probabilities to stay within each of the remaining k aggregates .*

5 Numerical Example

In order to test our algorithm, we applied it to a problem from molecular dynamics. In the so-called united atom representation [11], the n-pentane molecule

$CH_3 - (CH_2)_3 - CH_3$ (see Fig. 5) consists of 5 mass points, which interact with each other due to bond length, bond angle, dihedral angle and Lennard-Jones potentials. The most flexible part is the dihedral angle potential (Fig. 5) with its three minima.

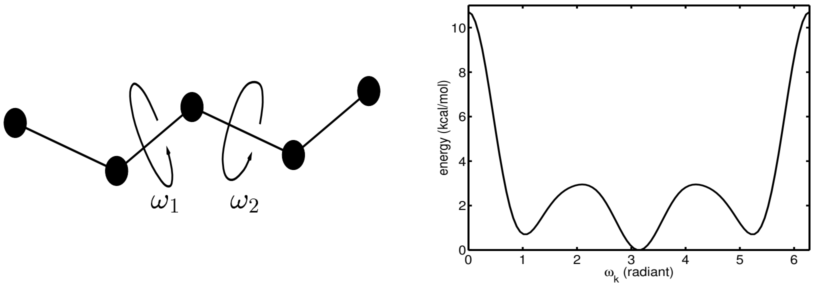

Figure 3: United atom model of n-pentane with the two dihedral angles ω_1 and ω_2. On the right: Dihedral angle potential due to [11]. The main minimum corresponds to the *trans* orientation of the angle, the two side minima to the $\pm$*gauche* orientations.

The dynamics of the molecule is described by the Hamiltonian equations of motion corresponding to these potentials. The typical dynamical behavior is characterized by extremely fast bond length and bond angle vibrations, which are nonlinearly coupled to significantly slower changes in the orientation of the dihedral angles. Therefore the orientations of the dihedral angles describe the "conformations" of the molecule, i.e., those subsets of the position space that are *almost invariant* under the flow of the Hamiltonian system.

As worked out in detail in [12], these almost invariant subsets are implicitly defined via some *spatial Markov operator* T. The invariant density of T is the well–known Boltzmann distribution. A *Galerkin-type discretization* of this operator (by means of a hybrid Monte Carlo method with step size $\tau = 160$fs) generates a stochastic matrix P, which is primitive and reversible, inheriting the special properties of the operator. Hence, after discretization, the almost invariant sets associated with the operator should show up as almost invariant aggregates of the matrix P.

A uniform discretization of each of the dihedral angles into 20 parts leads us to a number of $n = 20 * 20 = 400$ states and a 400×400 stochastic matrix P, the only input required for our herein suggested algorithm (see Section 4).

In our particular case the *stationary distribution* π of P is a priori known *explicitly* (being the spatial discretization of the Boltzmann distribution). It is represented in Fig. 4. This figure also shows the connection between the dihedral angle potential and regions with high probability for the system to be within. For example, the main maximum of the distribution corresponds to a n-pentane structure, where both dihedral angles are in the trans-orientation. The other maxima can be interpreted in the

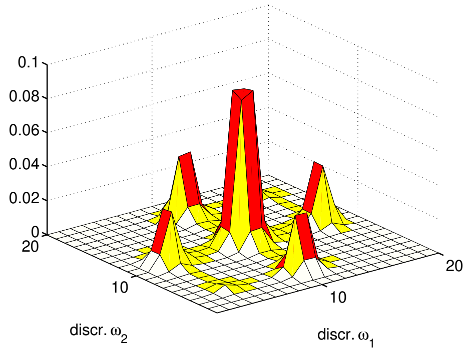

Figure 4: Stationary distribution π versus dihedral angles ω_1 and ω_2.

same way as combinations of +gauche-, $-$gauche- and trans-orientations. However, the stationary distribution does not contain any *dynamical* information. Rather, for the identification of almost invariant aggregates we have to investigate the spectral structure of P. For this purpose, the 10 eigenvalues of P with largest absolute value are arranged:

k	1	2	3	4	5	6	7	8	9	10
λ_k	1	0.986	0.984	0.982	0.975	0.941	0.938	0.599	0.590	-0.562

The first nine ones are positive. From the 10th one on negative eigenvalues appear frequently. As can be seen, a first spectral gap arises between λ_5 and λ_6, and an even more significant one between λ_7 and λ_8. We therefore tested our algorithm both for $k = 5$ and for $k = 7$.

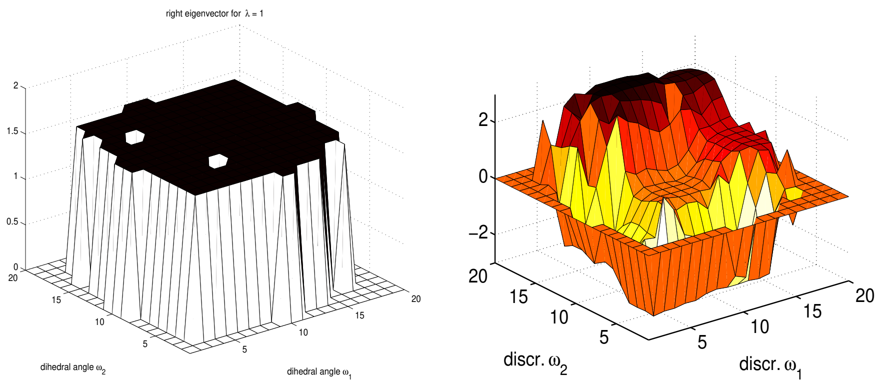

Figure 5: Right eigenvectors for largest eigenvalues $\lambda_1 = 1$ (left) and second largest eigenvalue $\lambda_2 = 0.9859$ (right) versus ω_1 and ω_2.

In Fig. 5 the right eigenvectors corresponding to $\lambda_1 = 1$ and $\lambda_2 = 0.986$ are illustrated. Of course, for $\lambda_1 = 1$, we obtain e, which in grid representation is just a flat plateau (ignoring zeroes for cut-off states). For λ_2, the right eigenvector contains more information. Just as in our model example (see Fig. 2), we can distinguish between different plateau levels, which seem to indicate different almost invariant aggregates.

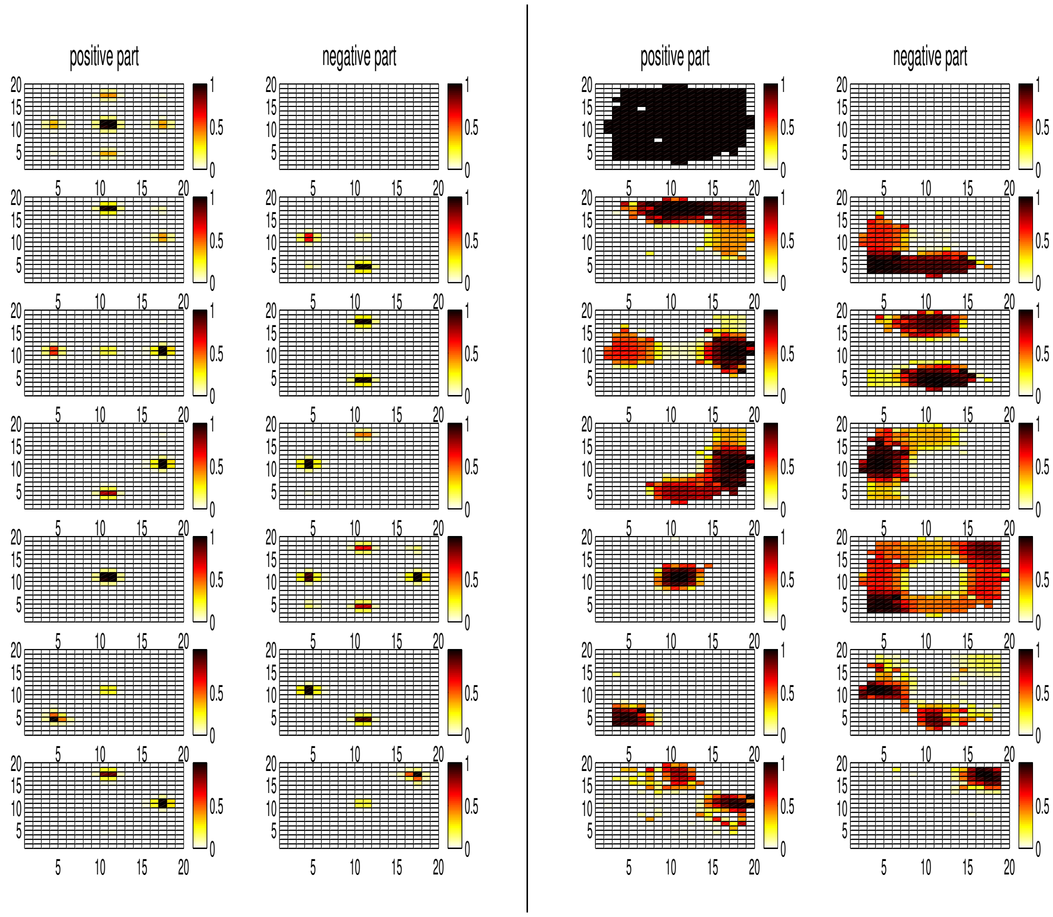

Figure 6: Positive and negative parts of the first seven *left* eigenvectors (left) and *right* eigenvectors (right). Positive and negative parts are scaled with respect to the maximum norm.

Fig. 6 represents the first seven eigenvectors split into their positive and negative parts (as described in Section 4). The *right* eigenvectors show the expected almost constant level structure, which allows to decompose the state space due to the algorithm explained in the previous section. In contrast to this, the *left* eigenvectors have distinct maxima only at the center of each constant level.

For $k = 5$, our identification algorithm started with 13 different sign structures ($l = 8$) and ended up with the five aggregates as illustrated in Fig. 7. All five aggre-

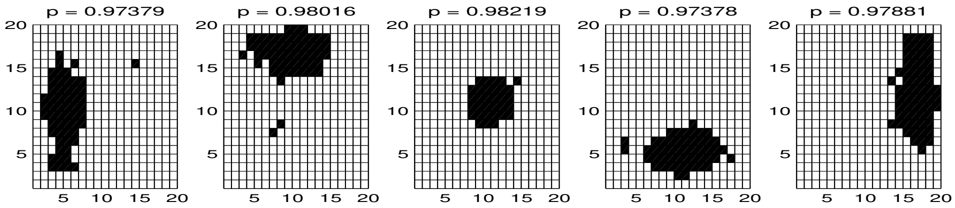

Figure 7: The resulting almost invariant aggregates for $k = 5$. The values for p denote the the probabilities to stay within these aggregates within the underlying discrete time step $\tau = 160$ fs.

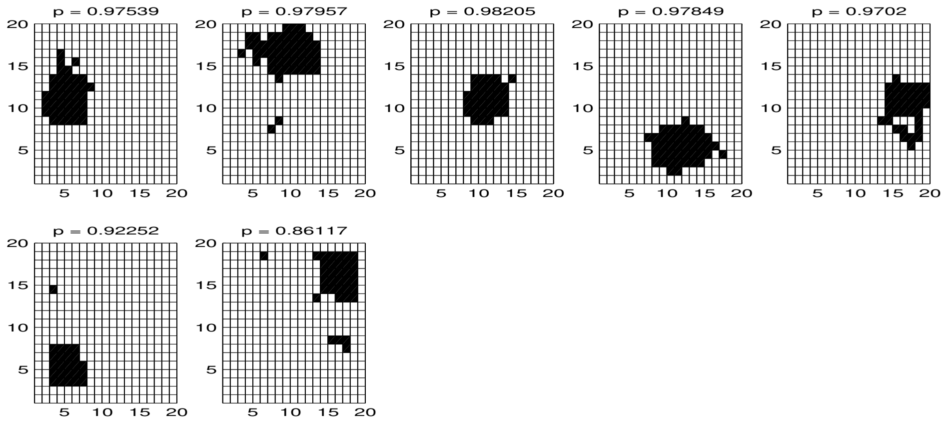

Figure 8: The resulting almost invariant aggregates for $k = 7$.

gates possesses a high probability to stay within. In addition, we can also compute all *transition probabilities* $w(A_i, A_j)$ between the identified almost invariant aggregates. The resulting coupling matrix is given below. Its entries may be interpreted as the probability that the system moves from A_i to A_j within the underlying discrete time step $\tau = 160$fs:

$$(w(A_i, A_j))_{i,j=1,\ldots,5} \quad = \quad \begin{pmatrix} 0.9738 & 0.0006 & 0.0163 & 0.0082 & 0.0011 \\ 0.0006 & 0.9802 & 0.0145 & 0.0008 & 0.0040 \\ 0.0044 & 0.0042 & 0.9822 & 0.0043 & 0.0049 \\ 0.0084 & 0.0009 & 0.0165 & 0.9738 & 0.0005 \\ 0.0010 & 0.0038 & 0.0161 & 0.0004 & 0.9788 \end{pmatrix},$$

where the numbering of the aggregates corresponds to Fig. 7.

For $k = 7$, there were 29 sign structures in the beginning ($l = 22$), which were reduced to the seven aggregates in Fig. 8. Observe, that the eigenvectors corresponding to λ_6 and λ_7 (see Fig. 6) contain the additional information about the separation of the +gauche/+gauche- and the −gauche/−gauche-conformation. Therefore, we obtain a more detailed partitioning of the state space, even though the probabilities to stay within the additional conformations are much lower. This example demonstrates that the algorithm can even produce good results, if some of the almost invariant aggregates possess a further substructure, which will be explored by eigenvectors corresponding to smaller eigenvalues. Both results, for $k = 5$ and for $k = 7$, are in good accordance with chemically observed conformations.

Acknowledgments. It is a pleasure to thank E. Behrends (FU Berlin) for helpful discussions and for pointing out reference [14]. One of us (W.H.) was supported within the DFG–Schwerpunkt "Ergodentheorie, Analysis und effiziente Simulation dynamischer Systeme" under Grant De 293/2-1.

References

[1] A. Berman and R. J. Plemmons. *Nonnegative Matrices in the Mathematical Sciences.* Academic Press, New York, 1979. Reprinted by SIAM, Philadelphia, 1994.

[2] M. Dellnitz and O. Junge. On the approximation of complicated dynamical behavior. To appear in SIAM J. Numer. Anal., 1996.

[3] P. Deuflhard, M. Dellnitz, O. Junge, and Ch. Schütte. Computation of essential molecular dynamics by subdivision techniques. To appear in [5], 1998. Available via http://www.zib.de/schuette/.

[4] P. Deuflhard, T. Friese, and F. Schmidt. A nonlinear multigrid eigenproblem solver for the complex Helmholtz equation. Konrad-Zuse-Zentrum Berlin, Preprint SC 97-55, 1997.

[5] P. Deuflhard, J. Hermans, B. Leimkuhler, A. Mark, B. Skeel, and S. Reich, editors. *2nd International Symposium "Algorithms for Macromolecular Modelling"*, Lecture Notes in Computational Science and Engineering. Springer-Verlag, 1998.

[6] George S. Fishman. *Monte Carlo — Concepts, Algorithms, and Applications.* Series in Operations Research. Springer, 1995.

[7] D. J. Hartfiel and C. D. Meyer. On the structure of stochastic matrices with a subdominant eigenvalue near 1. *Linear Algebra Appl.*, 272:193–203, 1998.

[8] A.F. Karr. Markov processes. In D.P. Heyman and M.J. Sobel, editors, *Stochastic Models*, volume 2, pages 95–123. North-Holland, Amsterdam, 1990.

[9] R. B. Lehoucq, D. C. Sorensen, and C.Yang. *ARPACK User's Guide: Solution of Large Eigenvalue Problems by Implicit Restartet Arnoldi Methods.* Rice University Houston, 1998.

[10] C. D. Meyer. Stochastic complementation, uncoupling Markov chains, and the theory of nearly reducible systems. *SIAM Review*, 31:240–272, 1989.

[11] J.-P. Ryckaert and A. Bellemans. Molecular dynamics of liquid n-butane near its boiling point. *Chem. Phys. Letters*, 30(1):123–125, 1975.

[12] Ch. Schütte, A. Fischer, W. Huisinga, and P. Deuflhard. A hybrid Monte Carlo method for essential molecular dynamics. Preprint SC-98-04, Konrad-Zuse-Zentrum, Berlin. Available via http://www.zib.de/bib/pub/pw/, 1998.

[13] E. Seneta. *Non–negative Matrices and Markov Chains.* Springer Series in Statistics. Springer, second edition edition, 1981.

[14] Alistair Sinclair. *Algorithms for Random Generation and Counting – A Markov Chain Approach.* Progress in Theoretical Computer Science. Birkhäuser, 1993.

[15] G. W. Stewart. On the structure of nearly uncoupled Markov chains. In G. Iazeolla, P. J. Courtois, and A. Hordijk, editors, *Mathematical Computer Performance and Reliability*, pages 287–302, New York, 1984. Elsevier.