

FIETE HAACK

## **Properties of Reduced Reversible Markov Chains**



# Properties of Reduced Reversible Markov Chains

Fiete Haack

November 13, 2008

## Abstract

The enormous time lag between fast atomic motion and complex protein folding events makes it almost impossible to compute molecular dynamics on a high resolution. A common way to tackle this problem is to model the system dynamics as a Markov process. Yet for large molecular systems the resulting Markov chains can hardly be handled due to the curse of dimensionality. Coarse graining methods can be used to reduce the dimension of a Markov chain, but it is still unclear how far the coarse grained Markov chain resembles the original system. In order to answer this question, two different coarse-graining methods were analysed and compared: a classical set-based reduction method and an alternative subspace-based approach, which is based on membership vectors instead of sets. On the basis of a small toy system, it could be shown, that in contrast to the subset-based approach, the subspace-based reduction method preserves the Markov property as well as the essential dynamics of the original system.

**AMS MSC 2000:** 62-07, 60J22, 65C40

**Keywords:** Markov chain, PCCA+, Conformation dynamics

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	Molecular dynamics as stochastic process . . . . .	3
1.2	Coarse graining . . . . .	3
<b>2</b>	<b>Reduction of the state space</b>	<b>5</b>
2.1	Set-based reduction . . . . .	5
2.1.1	Identification of metastable sets . . . . .	5
2.1.2	Calculation of transition frequencies . . . . .	5
2.2	Subspace-based reduction . . . . .	6
2.2.1	Identification of metastable sets . . . . .	6
2.2.2	Galerkin discretization . . . . .	7
<b>3</b>	<b>Prove of the Markov property</b>	<b>8</b>
3.1	Propagation of densities . . . . .	8
3.2	Chapman-Kolmogorov equation . . . . .	9
3.3	Exponential decay of eigenvalues . . . . .	9
3.4	Time scales of relaxation processes . . . . .	10
<b>4</b>	<b>Results and Discussion</b>	<b>11</b>
4.1	Coarse grained transition matrices . . . . .	11
4.2	Propagation of densities . . . . .	11
4.3	Chapman-Kolmogorov equation . . . . .	12
4.4	Eigenvalue analysis . . . . .	13
<b>5</b>	<b>Conclusion and Summary</b>	<b>15</b>
<b>6</b>	<b>Acknowledgements</b>	<b>15</b>
	<b>Literature</b>	<b>15</b>

# 1 Introduction

Atomic motion is extremely fast. Thermal vibrations like rotation, oscillation or fluctuation have a timescale of only a few femtoseconds. On the contrary, most essential dynamics, like complex protein-folding or protein-ligand binding processes, have a timescale of several seconds. The resulting time lag makes it almost impossible to simulate molecular motion with high resolution. Therefore a model is needed, which allows the simulation of molecular motion on a coarse level without losing the essential dynamics.

## 1.1 Molecular dynamics as stochastic process

Most molecules have metastable conformations, i.e. on large scales the molecule has the same geometric structure, whereas on small scales the system may rotate, oscillate or fluctuate. The essential dynamic of most molecules can therefore be described as a jump process with the molecule staying in one conformation for long periods of time and rare switches between these conformations. A physical explanation for these conformations can be given by the free energy landscape, which consists of deep wells, representing local minima. In general, each well is surrounded by large barriers, which separate each well from another. Due to these high energy barriers, transitions between different local minima are rare events [2].

Consider now the dynamic of a molecular system in equilibrium, with the conformational space decomposed into a set of  $N$  disjoint but contiguous states. By observing a trajectory of this system at discrete time steps  $t = 0, 1, \dots, i$  the trajectory can be seen as discrete stochastic process with the system having a specific state  $q^{(0)}, q^{(1)}, \dots, q^{(i)} \in \Omega$  at each of these discrete timesteps [1; 6]. This stochastic process defines a Markov chain, if the probability of the current state  $X_i = q^{(i)}$  only depends on its previous state  $X_{i-1} = q^{(i-1)}$ :

$$P(X_i = q^{(i)} | X_{i-1} = q^{(i-1)}, \dots, X_0 = q^{(0)}) = P(X_i = q^{(i)} | X_{i-1} = q^{(i-1)}),$$

i.e. it fulfils the Markov property. For a Markov chain with finite state space  $|\Omega| = N$ , the transition matrix  $P$  contains the conditional probabilities  $P(a, b)$  for each pair of states  $q^{(i-1)} = a$  and  $q^{(i)} = b$ .  $P$  is a stochastic matrix with non-negative elements and row sum 1.

However, for large molecular systems this simplified model still suffers from the curse of dimensionality, due to the enormous amount of local minima located on the free energy landscape. Therefore one has to find a way to further reduce the dimension of the Markov chain. This can be done by coarse graining.

## 1.2 Coarse graining

In order to reduce a given Markov chain, the easiest and most intuitive way would be a set-based approach. The standard practice contains the following steps:

- identify all metastable sets
- assign all states uniquely to one set
- calculate the transition frequencies between all pairs of different sets

- finally rescale the reduced transition matrix  $P_c$  to row sum 1.

Unfortunately, the reduced transition matrix  $P_c$  is not a Markov chain, since the Markov property does not hold anymore. This can be demonstrated with the following example:

A system with state space  $\Omega$  containing  $N = 36$  states is given. Assume the system can be decomposed into 3 metastable regions A, B and C with large transition probabilities between states of the same metastable set and low transition probabilities between states of different metastable sets. The system is shown in figure 1 with lines indicating transition probabilities larger than zero. Furthermore consider some states in set C with only low transition probabilities from their neighbors to these states, indicated by white circles. These states form a barrier inside

Legen Sie bitte den fertiggestellten Report ab in ein temporäres Verzeichnis /zibtmp/bibliothek mit dem Namen des Reports. In dieses Verzeichnis kopieren Sie bitte unbedingt das pdf-file - falls vorhanden möglichst auch das ps-file und das TeX-file - sowie die dazugehörigen Abbildungen. Achtung: Der Ordner und die files müssen für die Welt lesbar sein ( $\geq 755$ ). Danach benachrichtigen Sie bitte die Bibliothek, die alles weitere übernimmt.

set C, such that the probability of the transitions  $C \rightarrow A$  and  $C \rightarrow B$  depends on the previous step: if the previous transition was  $A \rightarrow C$  then the probability of  $C \rightarrow A$  is much higher than it is for  $C \rightarrow B$ . Therefore the set-based reduction of the Markov chain does not preserve the Markov property.

In contrast to this method, a subspace-based reduction preserves the Markov

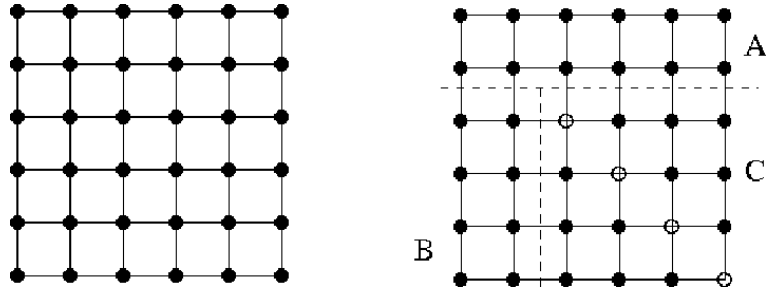


Figure 1: System of  $N=36$  states and three metastable sets A, B and C. The Markov property holds for single states (left) but can not be transferred to subsets of states (right). The probability  $C \rightarrow A$  depends on whether the system has been in B or in A before it enters C [8].

property [8]. The subspace-based reduction method presented in the cited and the present report is based on a Galerkin Discretization of the Markov chain represented by  $P$ . For the Galerkin Discretization, a basis  $\chi$  of an invariant subspace of  $P$  has to be determined, which represents the decomposition of the state space  $\Omega$  into metastable regions. As the number of metastable sets is equal to the number of dominant eigenvalues  $\theta$  of  $P$  near the perron root  $\theta = 1$ , the corresponding eigenvectors  $X$  span exactly the invariant subspace of  $P$  that we were looking for. So we have to find a transformation matrix  $A$  that transforms the eigenvector  $X$  to non-negative membership functions  $\chi$ :  $\chi = XA$ , with a regular matrix  $A \in \mathbb{R}^{n \times n}$ . The transformation matrix  $A$  can be generated via Robust Perron Cluster Analysis (PCCA+) [3; 7].

## 2 Reduction of the state space

In order to capture the main dynamics of a molecule in feasible time, one has to reduce the complexity of the system or, in terms of Markov chains, reduce the dimension of the transition probability matrix. A good way is to find all metastable sets and measure the transition frequencies between these sets. As a result, the number of states is reduced from  $N$  states to  $n$  metastable states, where  $n \ll N$ , and thus the full Markov chain with its transition matrix  $P \in \mathbb{R}^{N \times N}$  is reduced to  $P_c \in \mathbb{R}^{n \times n}$ .

### 2.1 Set-based reduction

The most intuitive and easiest way to reduce the dimension of  $P$  is a set-based approach. By identifying the metastable sets, all states  $q$  of the state space  $\Omega$  will be uniquely assigned to one metastable set, i.e. one conformation. The transition frequencies between two conformations can now be calculated by the weighted sum of all transition frequencies between the states belonging to these two conformations.

#### 2.1.1 Identification of metastable sets

For a dynamic process with metastable sets, the transition matrix  $P$  can be arranged to a nearly block-diagonal structure. Hence, the problem of identifying metastable sets in Markov chains is similar to the problem of locating the hidden block-diagonal structure in the transition matrix. Algorithms that reorder the rows and columns of the transition matrix such that states belonging to the same conformation appear in consecutive order already exist [3; 7]. Since each block represents a metastable set, the block-diagonal structure of the transition matrix can be used to define a membership matrix  $\chi \in \{0, 1\}^{N \times n}$  where  $\chi(i, j) = 1$  if and only if state  $i$  belongs to subset  $j$ .

#### 2.1.2 Calculation of transition frequencies

In order to construct the reduced transition matrix  $P_c \in \mathbb{R}^{n \times n}$  based on the given metastable sets, we first need to transform  $P$  into a symmetric matrix  $\bar{P}$  containing the relative transition frequencies,

$$\bar{P} = DP,$$

where  $D$  is the diagonal matrix  $D = \text{diag}(w)$  with the stationary density  $w = (w_1, w_2, \dots, w_n)$ . For a given set of  $n$  clusters, i.e. metastable sets, one can now identify the frequencies between each pair of clusters  $(k, l)$  with  $k, l = 1 \dots n$  by simply summing up those rows and columns of the transition frequency matrix  $\bar{P}$  that belong to the same cluster:

$$\bar{P}_c(k, l) = \sum_{i \in C_k, j \in C_l} \bar{P}(i, j).$$

Having characterized all metastable sets by membership functions  $\chi$ , one can also write:

$$\overline{P}_c = \chi^T \overline{P} \chi = \chi^T D P \chi.$$

Finally we have to rescale  $\overline{P}_c$  to row sum 1:

$$P_c = \tilde{D}^{-1} \chi^T D P \chi \quad (1)$$

with the diagonal matrix  $\tilde{D} = \text{diag}(w_c)$ , where  $w_c = \chi^T w$  is the reduced stationary density.

$P_c$  is a set-based transition matrix, with the conditional probability for the transition between each pair of subsets of the reduced system. However,  $P_c$  does not represent a Markov chain, since in a time series realization the Markov property is not given anymore.

## 2.2 Subspace-based reduction

In contrast to the subset-based reduction, as explained above, the representation of metastable conformations in a subspace-based reduction is based on soft membership functions. For each metastable set a weight is assigned to each state, describing how much the particular state belongs to a certain conformation.

### 2.2.1 Identification of metastable sets

It has been shown that the number of metastable sets  $n$  corresponds to the number of dominant eigenvalues of the matrix  $P$  near the Perron root  $\lambda = 1$  [2; 5]:

$$P X = X \Lambda, \quad \Lambda = \text{diag}(\lambda_i), \quad \lambda_i \approx 1, \quad X^T D X = Id.$$

Each dominant eigenvector therefore represents a metastable set. If the dominant eigenvectors  $X$  could now be transformed to membership functions  $\chi$ , we had a fuzzy decomposition of the state space  $\Omega$  that corresponds to the metastable regions of the system. The transformation of the eigenvectors  $X$  can be done by the cluster algorithm PCCA+, which has been developed by Marcus Weber and Peter Deuffhard [3; 7]. By applying PCCA+, a nonsingular matrix  $A$  will be constructed,

$$\chi = X A, \quad A \text{ nonsingular}$$

such that the transformed vectors  $\chi = [\chi_1, \chi_2, \dots, \chi_n]$  are almost characteristic functions, i.e.

$$\chi(i, j) \in [0, 1].$$

and sum to 1:

$$\sum_{j=1}^n \chi(i, j) = 1, \quad \forall i = 1, \dots, N$$

The construction of  $A$  is a constrained optimization problem. For further information see [3].



### 2.2.2 Galerkin discretization

$\chi$  represents a basis of the same invariant subspace of  $P$  as it is spanned by  $X$ . Thus, the reduced propagator  $P_c$  can now be obtained correctly by adapting Eq. (1):

$$P_c = \tilde{D}^{-1} \chi^T D P \chi \quad (2)$$

with  $\tilde{D} = \chi^T D \chi$ . Eq. (2) is the Galerkin discretization to the set of basis vectors  $\chi$  with restriction operator  $R$ ,

$$R := \chi^T$$

and interpolation operator  $I$ ,

$$I := D \chi (\chi^T D \chi)^{-1}.$$

With the Galerkin discretization of  $P$  in Eq. (2) we have mapped the dynamic of the system into an invariant subspace of  $\Omega$  which is spanned by  $\chi$ . Consequently the coarse system has the same properties as the original one [4]. Further the reduced transition matrix  $P_c^T$  can be interpreted as a Markov chain and can be used for time series realizations with a correctly reduced stationary density vector  $v_c^{(0)} = D \chi a$  with arbitrary vector  $a \in \mathbb{R}^n$  [8].

### 3 Prove of the Markov property

In order to evaluate the consequences of the different reduction methods for the simulation, a transition matrix of the system shown in Figure 1 was constructed to perform different tests on the Markov chain. The Markov chain was constructed such that the subsets A, B and C are metastable and set C contains a small barrier.

#### 3.1 Propagation of densities

Markov chains cannot only be modeled as a time series of states, but also as a density propagation. In contrast to the state-based time series, with the system having a defined state  $q^{(i)} \in \Omega$  at each timestep  $i$ , the density propagation approach is based on an ensemble of states. Given an initial distribution of states

$$v^{(0)} \in \mathbb{R}_+^N, ||v^{(0)}|| = 1,$$

the system evolves in time by

$$v^{(i)} = P^T v^{(i-1)}.$$

Here  $v^{(i)}$  contains the probabilities to reach any of the  $N$  states at time-step  $i$  and only depends on the previous time step  $(i - 1)$ .

If one applies coarse graining on a given Markov chain, the reduced propagator  $P_c^T$  can be used for a time series realization

$$v_c^{(0)} \xrightarrow{P_c^T} v_c^{(1)} \xrightarrow{P_c^T} \dots \xrightarrow{P_c^T} v_c^{(i)},$$

where  $v_c^{(i)}$  are the reduced density vectors.

Instead of propagating an initial density vector  $v^{(0)}$  in the high-dimensional space and subsequently restricting the result  $v^{(i)}$  to a reduced density vector  $v_c^{(i)}$ , the reduced propagator  $P_c^T$  can be used to directly propagate a reduced initial density  $v_c^{(0)}$ , i.e. create a distribution time series in the low-dimensional space.

However, for a correctly reduced propagator  $P_c$  and a valid restriction operator  $R = \chi^T$ , the two different approaches:

- propagation of  $v$  with  $P^T$  first, then reduction to  $v_c$
- reduction of  $v$  to  $v_c$  first, then propagation of  $v_c$  with  $P_c^T$

commute, i.e. lead to the same result [8; 4]. We can therefore use the two approaches to evaluate the quality of the particular reduction methods presented in Chapter 2 and check whether the Markov property is also given in the reduced Markov chain  $P_c$ .

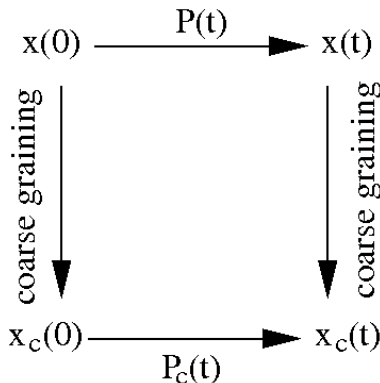


Figure 2: For a correctly reduced propagator  $P_c^T$  the propagation of the densities via  $P^T$  reduced to the subsets defined by  $\chi$  should be equal to the propagation of the reduced densities via  $P_c^T$ .

### 3.2 Chapman-Kolmogorov equation

For time homogeneous Markov chains the Chapman-Kolmogorov equation

$$P(t + s) = P(t)P(s)$$

holds. The equation can be extended to

$$P(kt) = P(t)^k.$$

Based on the idea of lag time sampling (see [1] and [6]) it is possible to prove the validity of the Chapman-Kolmogorov equation for a given transition matrix  $P(t)$ .

A number of samplings with different lag time  $k$ <sup>1</sup> will be performed. Based on the resulting trajectories, the transition matrix  $P(kt)$  will be reconstructed.

If the underlying process is an homogeneous Markov process, the transition matrix  $P(kt)$  reconstructed out of a trajectory with lag time  $k$  should be identical to the original transition matrix raised to the power of  $k$   $P(t)^k$ .

Concerning the sampling method, it is important to note that we need a special density-based sampling method for a subspace-based reduced transition matrix, as a standard state-based sampling approach would generate wrong results.

### 3.3 Exponential decay of eigenvalues

Another way to prove, if the Chapman-Kolmogorov property holds for the particular Markov chain, is to analyse the second largest eigenvalue of the reconstructed matrices. If the previously tested equation  $P(kt) = P(t)^k$  holds, then the eigenvalues should be similarly related, i.e. if  $\lambda$  is an eigenvalue of  $P(t)$ , then  $P(kt)$  is expected to have an eigenvalue  $\theta = \lambda^k$ . If a process is Markov, the eigenvalues of the reconstructed transition matrices thus have an exponential decay.

<sup>1</sup>Lag time is the number of time steps that are left out until the state is accepted.

However, the aim of coarse-graining is not only to preserve the Markov property, but also to maintain the essential dynamics of the original system. In order to evaluate this, we can also use the eigenvalues: While the largest eigenvalue of the transition matrix  $P(t)$  is always one, the second largest eigenvalue determines the maximum timescale of the Markov process, i.e. the rate of convergence. Thus for a good reduction method, the differences of the dominant eigenvalues of the reduced and the original system should be as small as possible.

### 3.4 Time scales of relaxation processes

The eigenvalues can also be seen as time rates for relaxation processes of the system. Therefore, the expected exponential decay of the eigenvalues can also be illustrated as a timescale of the corresponding relaxation process. In the case of a Markov chain, the time scales of the relaxation processes are in direct proportion to the lag times. The relative decay time rate  $\tau(k) = -k/\ln(\theta)$  as a function of lag time  $n$  is thus expected to be constant for all eigenvalues of  $P(kn)$ ,  $\forall k = 1, \dots, n$ .

## 4 Results and Discussion

### 4.1 Coarse grained transition matrices

Based on the methods described in Chapter 2, coarse graining was applied to the transition matrix  $P$  of the constructed Markov chain (see Figure 1). For the set-based reduction according to Eq. (1) we get:

$$P_{c1} = \begin{pmatrix} 0.9917 & 0.0025 & 0.0058 \\ 0.0016 & 0.9945 & 0.0039 \\ 0.0025 & 0.0025 & 0.9950 \end{pmatrix}$$

In contrast to that, the subspace-based reduction according to Eq. (2) leads to a slightly different reduced transition matrix:

$$P_{c2} = \begin{pmatrix} 0.9924 & 0.0023 & 0.0053 \\ 0.0017 & 0.9947 & 0.0036 \\ 0.0025 & 0.0023 & 0.9952 \end{pmatrix}$$

In the following, we will use  $P_{c1}$  for the set-based reduced matrix and  $P_{c2}$  for the subspace-based reduced matrix.

### 4.2 Propagation of densities

As already stated in 3.1, with a correct reduced propagator there should be no difference if the density is first propagated and then reduced or if a reduced density is propagated by a reduced propagator:

$$RP^T v = P_c R v$$

To test the preservation of the Markov property, the difference between the two diversely propagated densities at the same time step was measured.

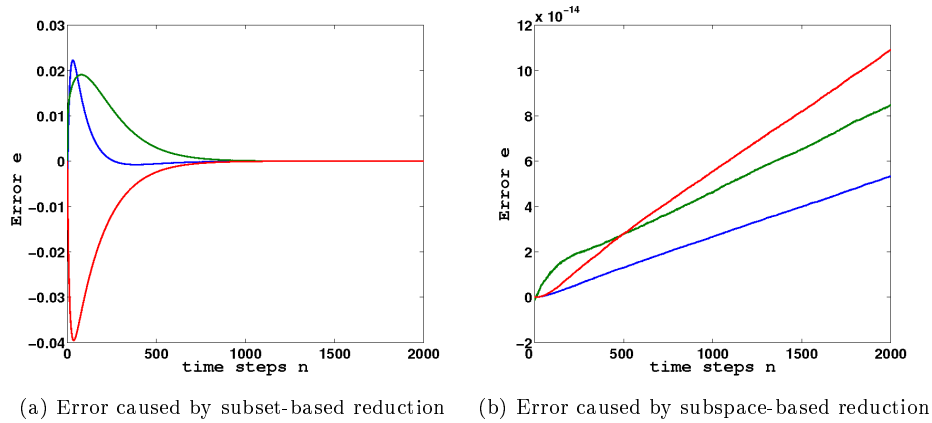


Figure 3: Propagation error caused by different reduction approaches. For each set of the system the difference  $e = \|RP^T v - P_c^T R v\|_2$  is shown.

The result for a trajectory with 2000 states is shown in figure 3 for the set-based and the subspace-based reduction approach, respectively. The difference  $e = \|RP^T v - P_c^T Rv\|_2$  is shown for each set, i.e. each conformation of the system.

Although it has been theoretically proven that  $P_{c2}^T$  is a correct propagator of the coarse grained system, there is still a difference of  $10^{-14}$  between the two diversely calculated reduced density  $v_c$ . This difference is caused by an amplification of the initial error, which has the order of machine precision  $10^{16}$ . For the set-based reduction the error of each state has a size of  $10^{-2}$  and reaches its maximum between 0 and 200 time steps. After that, the error for each state decreases until it converges to zero at time step 1000.

This is not surprising, as the stationary density of the reduced transition matrix  $P_{c1}$  corresponds to the reduced stationary density of the original Markov chain  $P$ . Hence, the density vectors approach the stationary density with increasing time steps.

### 4.3 Chapman-Kolmogorov equation

The following tests focus on the dynamics of the reduced Markov chains. First, the Chapman-Kolmogorov property was directly tested. For this purpose, the transition matrices  $P(kt)$ , reconstructed from time series with lagtime  $k$ , were compared to the original transition matrix  $P$  raised to the power of  $k$ .

The error  $e = \|P(kt) - P(t)^k\|_2$  as a function of lag time is shown in Figure 4 for the original Markov chain  $P$  and the two differently reduced Markov chains  $P_{c1}$  and  $P_{c2}$ . All three error curves are, except for small fluctuations, constant after some initial time. The error for the original Markov chain has a value of 0.12 and for the reduced systems it has a value of 0.02.

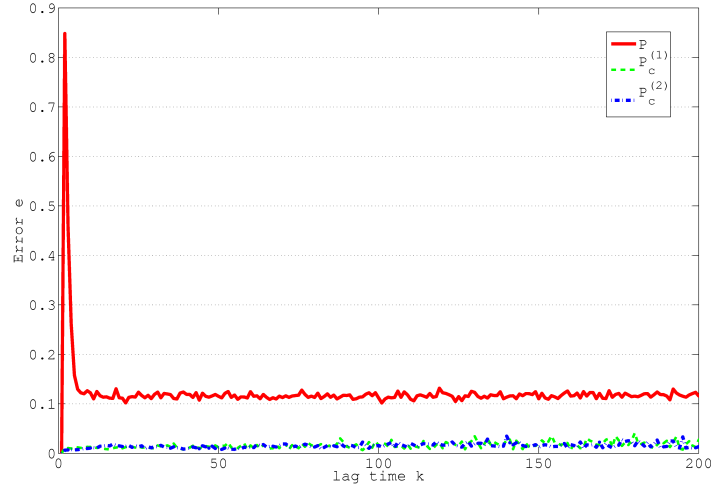


Figure 4: Error  $e = \|P(kt) - P(t)^k\|_2$  as a function of lag time.  $P(kt)$  is the transition matrix, that is reconstructed out of the trajectory, that is based on the respective propagator  $P^T$  (red),  $P_{c1}^T$  (green) and  $P_{c2}^T$  (blue).

The result seems surprising. As for a true homogeneous Markov chain the Chapman Kolmogorov equation holds, the error  $e = ||P(kt) - P(t)^k||_2$  is expected to be zero. Yet the results show an error of about 0.12 for the original Markov chain  $P(t)$ .

The reason for this difference is most likely the statistical error that is generated by the sampling method. A trajectory with 10.000 states does not sufficiently cover the whole state space  $\Omega$ , and hence the reconstructed transition matrix  $P(kt)$  contains minor variations, i.e.  $P(kt)$  does not completely represent the dynamic. If it was possible to sample infinitely long, one could avoid this statistical error. Consequently, the statistical error increases with the size of the system, which is the reason why the original Markov chain produces a higher error, than the two reduced Markov chains. Unfortunately, the statistical error is also most likely the reason why there's not a significant difference between the remaining two coarse transition matrices. With finer methods applied, we can make the difference visible, as demonstrated in the following section.

#### 4.4 Eigenvalue analysis

Instead of comparing the matrices directly, the second eigenvalues of the reconstructed matrices were analysed. As already stated, the eigenvalues should be similarly related, i.e. if  $\lambda$  is an eigenvalue of the transition Matrix  $P(t)$ , the eigenvalue for the transition matrix  $P(kt)$  should be equal to the value of  $\theta = \lambda^n$ .

In figure 5 the eigenvalues of the two differently reduced transition matrices  $P_{c1}$  and  $P_{c2}$  are plotted as functions of lag time.

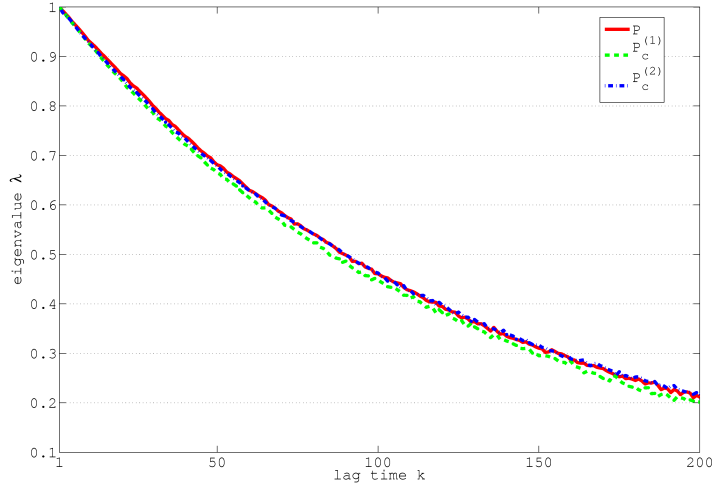


Figure 5: Second largest Eigenvalues  $\theta$  of the reconstructed transition matrices ( $P(kt)$  (red),  $P_{c1}(kt)$ (green) and  $P_{c2}(kt)$ (blue)) as function of lag times.

The eigenvalue curves are not identical. The eigenvalues of  $P_{c2}(kt)$  have higher values and are in general more similar to the second dominant eigenvalue of the original transition matrix  $P(kt)$  than the eigenvalues of  $P_{c1}(kt)$ . As

the second largest eigenvalue shows the maximum relaxation rate of a Markov process, we can state, that  $P_{c2}$  has almost the same essential dynamics as the original process. To analyse the exponential decay, the decay time  $\tau$  is calculated by

$$\tau(k) = -k/\ln(\theta).$$

Figure 6 shows the time scales  $\tau$  for each eigenvalue of the two reduced transition matrices  $P_{c1}$  and  $P_{c2}$ . If the process is Markov, the decay time is supposed to be a constant function [6]. For the set-based reduced transition matrix  $P_{c1}$ , the time scale of the second eigenvalue is not constant and shows a decreasing relaxation time with increasing lag-time  $k$ . Although the difference is rather small, this indicates that the set-based reduced transition matrix  $P_{c1}$  does not have a constant exponential decay and therefore violates the Chapman-Kolmogorov equation. In contrast to that, the time scale for the second eigenvalues of the

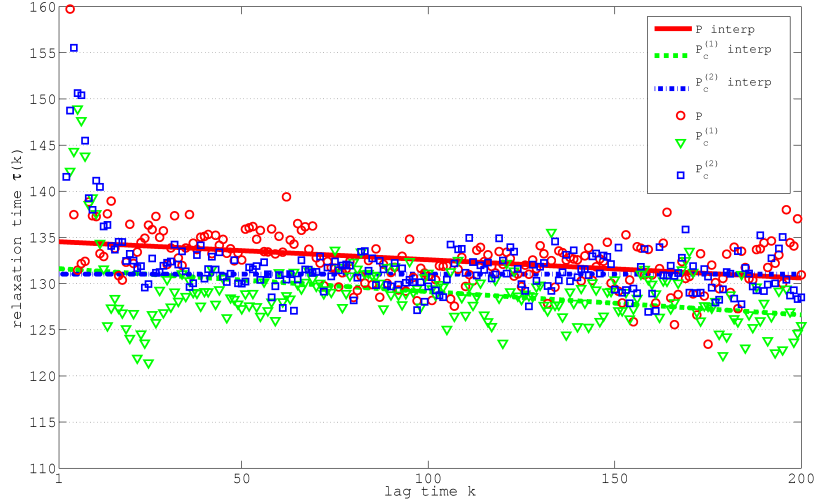


Figure 6: Time scales for relaxation processes implied by the second largest eigenvalues of the reconstructed transition matrices ( $P(kt)$  (red),  $P_{c1}(kt)$ (green) and  $P_{c2}(kt)$ (blue)). The exponential decay constants for the processes are related to the eigenvalues by  $\tau(k) = -k/\ln(\theta)$

subspace-based reduced transition matrix  $P_{c2}$  is almost constant, indicating that the (second) eigenvalue of the transition matrix follows the expected exponential decay.

Rather surprising is the fact that the eigenvalues corresponding to the original transition matrix  $P$  have a decreasing time scale, too. Most likely the statistical error that occurs due to finite sampling is the reason why the time-scale of  $P$  is not perfectly constant and also has a small descent with increasing lag-time  $k$ . Probably the same statistical error enforces the error that occurs in the set-based reduction method.



## 5 Conclusion and Summary

This research project focused on the accuracy of reduction methods for reversible Markov chains. The long-time behaviour of reduced Markov chains was analysed as well as the preservation of the system's dynamic on different time scales.

In a first step it could be shown that the propagation in the coarse space is only possible with a subspace-based reduced propagator (see Sec. 4.2). When reduced density vectors are propagated directly by a set-based reduced propagator we observe an error of size  $10^{-2}$ .

Considering the rather small size of the original system (36 states), the estimated propagation error indicates that the long-time behaviour of the reduced system cannot be transferred to time-independent transition rates. In contrast to that the subspace-based reduced transition matrix has exactly the same long-time behaviour as the original system over the course of time-steps and can thus be generalized to arbitrary times  $t > 0$  [4]. However, the results of the lag time samplings and the following eigenvalue analysis (see Sec. 4) did not show a significant violation of the Chapman-Kolmogorov equation in any of the reduced systems. Obviously, most of the deviation of the exponential decay constants of the eigenvalues in Figure 6 were caused by the finite sampling method. Though, the set-based approach showed particular variations beyond that statistical error, which might indicate a violation of Chapman-Kolmogorov. It would be of further interest to perform additional tests on systems of different size and with varying parameters, e.g. longer sampling periods.

The analysis of the eigenvalues showed another important fact. In contrast to the subset-based approach, the subspace-based reduction preserves the essential dynamics of the original system. This means that the reduced transition matrix contains the correct transition probabilities between the three metastable sets. For the application in conformation dynamics this is a very important fact. With the subspace-based reduction approach, we have a new method, that preserves the Markov property during reduction and maintains the essential dynamics of the original system. In contrast to that we could show exemplary that a subset-based coarse-graining method leads to an invalid reduced transition matrix, that does neither fulfil the Markov property nor represent the essential dynamics of the high-dimensional Markov chain.

## 6 Acknowledgements

I'd like to deeply thank Dr. Marcus Weber for supervising this research project and giving all the helpful comments and suggestions. I'm also very grateful to Susanna Röblitz for help with Matlab, especially for providing an implementation of the density-based sampling method and further for proofreading this report.

## References

- [1] J. D. Chodera, W. C. Swope, J. W. Pitera, and K. A. Dill. Long-time protein folding dynamics from short-time molecular dynamics simulations. *Multiscale Model. Simul.*, 5:1214–1226, 2006.
- [2] P. Deuffhard and C. Schütte. Molecular conformation dynamics and computational drug design. *ZIB-Report*, 03-20, 2003.
- [3] P. Deuffhard and M. Weber. *Robust Perron cluster analysis in conformation dynamics*, volume 398C, pages 161–184. Elsevier, New York, 2005.
- [4] S. Kube and M. Weber. A coarse graining method for the identification of transition rates between molecular conformations. *J. Chem. Phys.*, 126, 2007.
- [5] C. Schütte. *Conformational Dynamics: Modelling, Theory, Algorithm, and Application to Biomolecules*. Habilitation thesis, Freie Universität Berlin, 1999.
- [6] W. C. Swope and J. W. Pitera. Describing protein folding kinetics by molecular dynamics simulation. 1. Theory. *J. Phys. Chem. B*, 108:6571–6581, 2004.
- [7] M. Weber. Improved perron cluster analysis. *ZIB-Report*, 03-04, 2003.
- [8] M. Weber and S. Kube. Preserving the Markov property of reduced reversible Markov chains. In T. E. Simos, G. Psihoyios, and Ch. Tsitouras, editors, *Numerical Analysis and Applied Mathematics*, volume 1048, pages 593–596. Int. Conf. on Num. Analy. and Appl. Math. 2008, American Institute of Physics, September 2008.