

Master thesis

Computing the minimal rebinding effect for nonreversible processes

**Zur Erlangung des akademischen Grades
Master of Science**

Susanne Röhl

Berlin, den 9. Januar 2017

Contents

1	Introduction	3
1.1	Drug Design	3
1.2	Rebinding Effect in Drug Design	3
1.3	Minimizing the Rebinding Effect	3
1.4	Markov State Models	3
1.5	Literature	4
1.6	Structure	4
2	Markov State Models	5
2.1	Markov Process	5
2.2	Transfer Operator	7
2.3	Metastability / Dominant Spectrum	8
2.4	Markov State Model	9
2.4.1	Galerkin Projection	9
2.4.2	Recrossing Effect	10
3	Dominant Structures	11
3.1	Dominant metastable sets	11
3.2	Dominant cycles	11
3.2.1	Schur Decomposition	11
4	Rebinding Effect in a Given Kinetics	12
4.1	Receptor-Ligand System	12
4.2	Rebinding Effect	12
4.3	Molecular Kinetics as a Projection	13
4.4	Minimizing the Rebinding Effect	14
5	Illustrative Examples	15
5.1	Artificial Process	15
	Bibliography	i

1 Introduction

1.1 Drug Design

A drug becomes good if it has a high binding affinity to the virus/.. . The binding affinity of the drug can be affected by different aspects. Like - shape of the drug molecule (favorable spatial preorganisation of the ligands leads to an entropy loss at ligand binding) - thermodynamical reasons - rebinding effect

Eine Randnotiz

1.2 Rebinding Effect in Drug Design

drug design: minimize rebinding effect to make drugs better

1.3 Minimizing the Rebinding Effect

In [10] Weber solved the problem of minimizing the rebinding effect for reversible stochastic processes. We extend his approach in generalizing it to the case of nonreversible processes.

In that context the main difference between reversible and nonreversible processes is their spectrum. While reversible processes possess only real eigenvalues, the eigenvalues of nonreversible processes can be complex. So a spectral decomposition is not immediately possible for nonreversible processes. Instead we use the Schur decomposition as it has been shown in [1] how to find dominant structures of nonreversible processes.

1.4 Markov State Models

Thermodynamical Ensemble/ Boltzmann Distribution, (Potential) Energy Landscape: High Potential Energy = transition region, Low Potential Energy = (stable/ metastable/ invariant) region

Clustering in metastable conformations - Mixture Model Clustering - Hidden Markov Model - PCCA+ = Clustering Algorithm using Membership Functions

Deuffhard/ Schuette: set-based clustering approach Weber (PCCA+): function-based clustering (meshless function), (continuous and overlapping)membership functions, partial density dunction/ total density function, complexity reduction

1.5 Literature

Basis theory of Metastability and Markov State Models is from Schuette (MSM book), including notation, definitions, transfer operator, dominant spectrum, Galerkin projection...

1.6 Structure

First: What is a Markov Process, what is a MSM (Markov State Model), Why is the (dominant) spectrum of importance, What is Metastability, including important definitions (transfer operator, Galerkin Projection,...)

Second: Metastability, the importance of the dominant spectrum, the difference of reversible and irreversible MC (dominant metastable sets vs dominant cycles), Schur Decomposition, Schwerpunkt on irreversible processes, Schwierigkeiten von irreversible processes

Third: Application of the previous chapters to MD (Molecular Dynamics): Re-binding Effect, What is the Rebinding Effect, Molecular Kinetics as a Projection, Minimize the Rebinding Effect

Fourth: An Artificial Example: Computing the Minimal Rebinding Effect, Some Pictures!

Eine Randno-
tiz

2 Markov State Models

In this chapter, we introduce the mathematical fundamentals that we need to describe the behaviour of molecular systems.

We suppose that the basics of probability theory (see ...) are known. The theory of (discrete) Markov chains is not necessarily a prerequisite to understand the definitions in this chapter, but still useful to know (see..). Here we will present a generalized (continuous) concept of this topic, so from time to time we will refer to the special case of a Markov chain to exemplify/illustrate our results/notions/definitions.

The notation of this chapter is mainly based on the book *Metastability and Markov State Models in Molecular Dynamics* from Schütte and Sarich[5] which gives a good overview over the basic concepts of Markov State Models.

2.1 Markov Process

A Markov process is a certain/particular stochastic process with some nice properties which makes it easy to compute with. They are defined on a continuous state space and continuous time, in opposite to a Markov chain where both these properties are discrete (finite state space?, transition matrix = stochastic matrix).

transition function, transition kernel, transition probability

From now on we will denote by (E, Σ) a measurable space and with $(\Omega, \mathcal{A}, \mathbb{P})$ a measure space.

A random variable is a function from a measure space into a measurable space s.t. preimages of measurable sets are measurable. That definition induced a measure on the measurable space.

Definition 2.1 (Stochastic Process). *A stochastic process on an index set I and state space E is a family of random variables $(X_i)_{i \in I}$, $X_i : \Omega \rightarrow E$.*

Definition 2.2 (Transition function). *A function $p : \mathbb{T} \times \mathbb{X} \times \mathcal{B}(\mathbb{X}) \rightarrow [0, 1]$ is a transition function if it fullfills the following properties:*

- i) $x \mapsto p(t, x, A)$ is measurable for all $t \in \mathbb{T}$ and $A \in \mathcal{B}(\mathbb{X})$
- ii) $A \mapsto p(t, x, A)$ is a probability measure for all $t \in \mathbb{T}$ and $x \in \mathbb{X}$

transfer operator: propagation of distributions?

iii) $p(0, x, \mathbb{X} \setminus x) = 0$ for all $x \in \mathbb{X}$

iv) p fulfills the Chapman-Kolmogorov equation, namely for all $t, s \in \mathbb{T}, x \in \mathbb{X}$ and $A \subset \mathbb{X}$:

$$p(t + s, x, A) = \int_{\mathbb{X}} p(t, x, z) p(s, z, A) dz. \quad (2.1)$$

So the transition function gives us the probability to get into a certain subset A in a time interval t starting from a point x . That means that we can describe the time evolution of a stochastic process by a transition function. In particular, the transition matrix of a Markov chain (time discrete, state space finite) is a special case of the transition function since it fulfills the above properties.

?

why is that immediate? Chapman?

Now we can define Markov processes, a certain stochastic process.

Definition 2.3 (Markov Process). *A stochastic process $(X_t)_{t \in \mathbb{T}}$ on a state space \mathbb{X} is a Markov process if...*

$$p(t, x, A) = \mathbb{P}(X_{t+s} \in A \mid X_s = x). \quad (2.2)$$

for all $s, t \in T, x \in \mathbb{X}$ and $A \subset \mathbb{X}$.

So the time evolution of a Markov process is described by its transition function. From this definition, it is clear that such a process possesses no "memory". The change of the process after an arbitrary time step t depends only on the last known state of the process at time s and not of any previous states, as we can see on the right side of equation

So a Markov process is a process without memory; its evolution depends only on the current space and is independent of the history. We are only interested in (time-)homogeneous Markov processes, i.e. processes which...

So the transition function for a Markov process plays the same role as the transition matrix for a Markov chain; propagating distributions?

Definition 2.4 (Invariant measure). *Let $(X_t)_{t \in \mathbb{T}}$ be a Markov process. The probability measure μ is invariant w.r.t. $(X_t)_{t \in \mathbb{T}}$ if for all $A \subset \mathbb{X}$ we have that*

$$\int_{\mathbb{X}} p(t, x, A) \mu(dx) = \mu(A). \quad (2.3)$$

Ergodicity

Definition 2.5 (ergodic process). *Let $(X_t)_t$ be a Markov process with invariant probability measure μ . Then $(X_t)_t$ is ergodic w.r.t. μ if for all functions $u : \mathbb{X} \rightarrow \mathbb{R}$ with $\mathbb{E}[|u|] := \int_{\mathbb{X}} |u| \mu(dx) < \infty$*

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T u(X_t) dt = \mathbb{E}_{\mu}[u]. \quad (2.4)$$

Reversibility

A process is reversible if it fulfills the detailed balance equation...; it means that they keep the same probability law even if their movement is considered backwards in time.

We will only consider ergodic Markov processes, i.e. processes that are irreducible and aperiodic (no periodic behaviour). A very nice property of Markov processes is reversibility.

only one
eigenvalue 1

Definition 2.6 (reversible process). *Let $(X_t)_t$ be a Markov process with invariant probability measure μ . Then $(X_t)_t$ is reversible w.r.t. μ if*

$$\int_A p(t, x, B) \mu(dx) = \int_B p(t, x, A) \mu(dx) \quad (2.5)$$

for all $t \in \mathbb{R}$ and $A, B \subset \mathbb{X}$.

2.2 Transfer Operator

With the previously defines transition function, we have a tool to describe to propagation of distributions(?) of stochastic processes. Now we are going to introduce an operator that propagates probability densities of Markov processes.

transition
function
skillnad
transfer
operator
generator

Definition 2.7 (L^r -Spaces).

$$L^r(\mu) = \{u : \mathbb{X} \rightarrow \mathbb{R} \mid \int_{\mathbb{X}} |u(x)|^r \mu(dx) < \infty\} \quad (2.6)$$

for $1 \leq r < \infty$ and

$$L^\infty(\mu) = \{u : \mathbb{X} \rightarrow \mathbb{R} \mid \text{ess sup}_{x \in \mathbb{X}} |u(x)|^r \mu(dx) < \infty\}. \quad (2.7)$$

The space L^2 is the only of the L^r -spaces which can be equipped with a canonical scalar product and thus becomes a Hilbert space (see..). For $u, v \in L^2$ we define

$$\langle u, v \rangle_\mu := \int_{\mathbb{X}} u(x)v(x)\mu(dx). \quad (2.8)$$

Now let ν_0 be the density function corresponding to the given start distribution (function). Then the density function of a subset $A \subset \mathbb{X}$ at time t is given by

$$\nu_t(A) = \int_{\mathbb{X}} \nu_0 p(t, x, A) \mu(dx). \quad (2.9)$$

On the other hand, the density ν_t is given by

$$\nu_t(A) = \int_A \nu_t(x) \mu(dx). \quad (2.10)$$

Definition 2.8 (Transfer Operator). *Let $p : \mathbb{T} \times \mathbb{X} \times \mathcal{B}(\mathbb{X}) \rightarrow [0, 1]$ be the transition function of a Markov Process $(X_t)_{t \in \mathbb{T}}$ and μ be an invariant measure of $(X_t)_{t \in \mathbb{T}}$. The semigroup of propagators or (forward) transfer operators $P^t : L^r(\mu) \rightarrow L^r(\mu)$ with $t \in T$ and $1 \leq r \leq \infty$ is defined via*

transfer operator: propagation of densities

$$\int_A P^t v(y) \mu(dy) = \int_{\mathbb{X}} v(x) p(t, x, A) \mu(dx) \quad (2.11)$$

for measurable $A \subset \mathbb{X}$.

Another thing to describe the process is the infinitesimal generator.

backwards operator important?

Definition 2.9 (Infinitesimal Generator). *The semigroup of propagators or forward transfer operators $P^t : L^r(\mu) \rightarrow L^r(\mu)$ with $t \in T$ and $1 \leq r \leq \infty$ is defined as follows:*

$$L\nu = \lim_{t \rightarrow \infty} \frac{P^t \nu - \nu}{t}. \quad (2.12)$$

Then the operator $L : \dots$ is called the infinitesimal generator corresponding to the semigroup P^t .

We have $P^t = \exp(tL)$ in $L^2(\mu)$, so L "generates" the semigroup of transfer operators.

So the eigenvalues $\lambda_1, \dots, \lambda_m$ of the propagator P^t are related to the eigenvalues $\Lambda_1, \dots, \Lambda_m$ of the propagator L in the following way:

$$\lambda_k = \exp(t\Lambda_k) \quad (2.13)$$

for all $1 \leq k \leq m$. Their corresponding (associated) eigenfunctions are identical.

2.3 Metastability / Dominant Spectrum

Der Spektralansatz ist also hervorragend geeignet, um Metastabilität bei Markovprozessen zu charakterisieren, bringt aber zwei Nachteile mit sich. Zum einen ist das Resultat nur auf reversible Prozesse anwendbar, da sonst der Transferoperator nicht selbst-adjungiert ist und es nicht gewährleistet ist, dass alle Eigenwerte reell sind, und zum anderen lässt sich das Eigenvektorproblem des Transferoperators nur global lösen. Daher beschäftigen wir uns nun mit einem alternativen Ansatz, der diese beiden Probleme nicht hat. Dieser Ansatz wurde in [5] bereits erfolgreich auf Markov-Ketten angewandt, und wir wollen nun analoge Resultate für den kontinuierlichen Fall herleiten.

Proof. Here is my proof

□

A transfer operator has spectral radius 1, which means that all eigenvalues have $\lambda_i \leq 1$. We will consider only ergodic process which means that the eigenvalue 1 is unique. If the transfer operator is self-adjoint, then all eigenvalues are real; otherwise they lie in the complex plane and are possibly nonreal.

What is the dominant structure of a nonreversible process? cycle?

2.4 Markov State Model

p.70

interpret P_Q as a projected transfer operator

2.4.1 Galerkin Projection

Until now we have a process on a continuous state space. Since we are interested in computations with our process, we are now creating a smaller (namely finite) state space which shall inherit (important) properties of our original process. This can be done by a Galerkin projection/ discretization.

need(?) to get a discrete/finite state space

At first, we need to choose an appropriate/ desired/ favoured ansatz space $\chi = \{\chi_1, \dots, \chi_n\} \subset L^2(\mu)$.

Partition of Unity

MA Jo p. 15 - 16

Im folgenden beschäftigen wir uns mit der Situation, dass wir vom Zustandsraum X eines Prozesses übergehen zu einem endlichen Unterraum $D \subset X$. Durch die Projektion des Transferoperators auf den Unterraum lässt sich ein stochastischer Prozess auf dem Unterraum beschreiben, der viele Eigenschaften des ursprünglichen Prozesses behält, allerdings in der Regel nicht mehr die Markov-Eigenschaft besitzt. Dennoch ist der auf einen geeigneten Unterraum projizierte Prozess eine gute Approximation des Markov- Prozesses auf X . Zunächst führen wir die Galerkin-Projektion auf einen Unterraum mit Hilfe einer geeigneten Basis des Unterraums ein. Sowohl für die Funktionen einer Teilung der Eins als auch für die charakteristischen Funktionen wird im Folgenden der Buchstabe verwendet, was jeweils gemeint ist, sollte allerdings immer aus dem Kontext klar sein.

Im Falle einer vollständigen Partition lässt sich also über den projizierten Transferoperator eine Markovkette definieren, welche sich auf den Mengen der Partition bewegt. Diese Markovkette behält einige wichtige Eigenschaften des Prozesses auf X , unter anderem ist die stationäre Verteilung der über P_Q definierten Markovkette gerade die Projektion des invarianten Maßes auf den Unterraum D .

Eine wichtige Eigenschaft des Transferoperators und des Generators geht durch die Projektion auf den Unterraum jedoch verloren, und zwar die Markoveigenschaft

Dieses Beispiel, in welchem das überschreiten der Barriere für den nächsten Zeitraum eine höhere Wahrscheinlichkeit einer erneuten Überschreitung der Barriere bedeutet, wird auch als Recrossing Problem bezeichnet und ist in...

-i. spoling Markov property???

An important question when it comes to projections of stochastic processes onto lower-dimensional state spaces is shown in the following diagram. Does it make a difference if we first project the process and then propagate it and vice versa?

$$\begin{array}{ccc} P(\tau) & \xrightarrow{\tau \rightarrow \tau^k} & (P(\tau))^k \\ \downarrow \text{projection} & & \downarrow \text{projection} \\ P_C(\tau) & \xrightarrow{\tau \rightarrow \tau^k} & (P_C(\tau))^k \end{array}$$

Weber shows in habilitation that under a certain Galerkin Projection using membership functions (not set-based family) leads to a commuting diagram of projection and propagation. + Markov Property is preserved?

non-reversible?

2.4.2 Recrossing Effect

If we consider a (continuous) Markov process (X_t) and project it onto a finite-dimensional state space via a Galerkin projection, then the process can lose his Markov property; that phenomena is called the recrossing effect.

It depends on the chosen partition of unity (the new finite-dim. state space) if the recrossing effect occurs. If we chose a full partition for that ansatz space, then our projected process will maintain its Markovianity.

3 Dominant Structures

As we have seen in chapter 2, knowing the dominant spectrum of the transfer operator is important to detect the (or a) metastable decomposition of the state space of the Markov process.

When we are considering reversible processes, the transfer operator is self-adjoint (?), so it has only real eigenvalues. Then we can just apply PCCA+ to get the best decomposition into metastable sets. (using membership functions bla).

But when it comes to non-reversible processes, the eigenvalues of the transfer operator are not necessarily real; they are lying in the complex plane with spectral radius 1. Since we consider only ergodic processes, the eigenvalue 1 is also here unique. In non-reversible processes, metastability is given by cycles instead of sets .

The basics of cycle representations stem from Kalpazidou[3] and its application to nonreversible Markov processes from Djurdjevac, Weber and Schütte[1].

what is a (metastable) cycle?

3.1 Dominant metastable sets

When a Markov model is reversible then its metastable sets can be found with the spectral approach. If the spectrum of the transition matrix P has m dominant eigenvalues, i.e. m eigenvalues close to 1, then there exists a decomposition of the state space of the process into m metastable sets.

PCCA+

3.2 Dominant cycles

Definition dominant cycle

3.2.1 Schur Decomposition

Dominant structures will be defined utilizing the dominant Schur vectors of the transition matrix instead of its eigenvectors.

A membership matrix can be defined as a linear combination of these Schur vectors.

4 Rebinding Effect in a Given Kinetics

In this chapter we are going to examine a special type of a molecular (kinetics?) system, namely receptor-ligand systems. To describe these systems we can use all the previously defined mathematical objects.

To give a short overview about what is going to happen here. A molecular system can be described via a differential equation. The solution of this differential equation is a (Markov?) process which can be described via a transfer operator (chapter 2). This operator will be projected onto a finite-dimensional state space (Galerkin projection, chapter 2) which (maybe?) spoils the Markov Property of the process.

Here we will try to tackle this subject for nonreversible (NESS) processes also.

This chapter is mathematically based on Weber (Quantifying Rebinding Effect)..

4.1 Receptor-Ligand System

Ligand (L) can bind to a receptor (R) and form a receptor-ligand complex (LR)

together with reb. eff.

What is a molecular system? State space, Phase Space, Ensemble, Configurational Space, Conformation Space, reduced density

4.2 Rebinding Effect

The rebinding effect has been characterized as a memory effect which leads to an additional thermodynamic weight of the bound state.

In fact, a stochastic process describing a receptor-ligand molecular system IS NOT necessarily Markovian. The Markovianity can be spoiled by the Rebinding Effect. If a Receptor-Ligand system dissolves, due to the favorable spatial situation (?) it is more likely to rebind again than to stay dissolved.

There are several papers (...) describing the rebinding effect on a chemical and a mathematical view. In chemistry, there are several reasons/ factors for the rebinding

impact of multivalency on the rebinding effect (Weber, Chem.)
Zusammenhang zum Absatz davor?

effect discussed.

An important application of receptor-ligand processes is drug design. In short: A drug consists of ligands which should bind to the receptors of the virus. If the drug creates many bindings, the virus is "bound" and cannot attack the human (cell?) anymore. So many bindings are a favorable thing. So a high rebinding effect enhances the (overall?) binding affinity of the process/ system which is good for the efficiency of a drug. We want a high rebinding effect. So in this chapter, we examine the minimal rebinding effect for a given Kinetics. This task has been solved by Weber... for reversible processes.

How: bound -> unbound

4.3 Molecular Kinetics as a Projection

In general, to describe molecular systems we can distinguish between two point of views: considering micro states or macro states of the system.

Mathematical
description
of molecular
kinetics

Micro States

A micro state of a molecular system with N atoms can be represented in a $6N$ -dimensional phase space $\Gamma = \Omega \times \mathbb{R}^{3N}$, consisting of the configurational space $\Omega = \mathbb{R}^{3N}$ and the momentum space \mathbb{R}^{3N} .

Boltzmann distribution

$$\bar{\rho}(q, p) \propto \exp(-\beta H(q, p)), \quad (4.1)$$

where $\beta = 1/(k_B T)$ is the inverse of the temperature T multiplied with the Boltzmann constant k_B . The Hamilton function denoted by H is given by $H(q, p) = K(p) + V(q)$, so by the sum of the kinetic energy $K(p)$ and the potential energy $V(q)$.

Macro States

Group/cluster a collection of the micro states having the same or similar values in one observable (e.g. bound, unbound state of receptor-ligand system). By grouping micro states, the (corresponding?) macro states yield **TODO:** statistical weights and entropic information (?). Macro states need not be distinct sets. Approach here: overlapping partial densities, which can be identified as membership functions

χ_1, \dots, χ_n .

partition of unity

membership
functions,
overlap-
ping partial
densities

Transfer Operator

For reversible(!) processes Weber [8] defines a transfer operator which is self-adjoint! due to the detailed balance condition.

$$\mathcal{P}(\tau)f(q) = \int_{\mathbb{R}^{3N}} \left(\int_{\Omega} f(\tilde{q}) \Psi^{-r}(\tilde{q} | (q, p)) d\tilde{q} \right) \pi_p(p) dp. \quad (4.2)$$

So for the nonreversible approach we cannot use this operator? Or: we can use this operator but it is not self-adjoint anymore? -> so complex eigenvalues..

Galerkin Projection

The Galerkin Projection helps us to reduce our continuous stochastic process to a discrete process by a projection.

$$P(\tau) = (\langle \chi, \chi \rangle_{\pi})^{-1} (\langle \chi, \mathcal{P}(\tau) \chi \rangle_{\pi}) \quad (4.3)$$

Normally the Markov Property is spoiled by a projection on a finite-dimensional space, but with this Galerkin Projection Markovianity is preserved....

4.4 Minimizing the Rebinding Effect

For reversible processes, this problem is solved by Weber and Fackeldey [10] with the spectral approach.

With the tools from chapter 2 (Schur Decomposition) we give an approach how this problem can be solved for nonreversible processes (NESS processes) using Schur Decomposition to get rid of the possibly nonreal eigenvalues, see Djurdevac et al [1](2016).

5 Illustrative Examples

5.1 Artificial Process

Bibliography

- [1] N. DJURDJEVAC CONRAD, M. WEBER, AND C. SCHÜTTE, *Finding dominant structures of nonreversible markov processes*, Multiscale Modeling & Simulation, 14 (2016), pp. 1319–1340.
- [2] W. HUISINGA, *Metastability of markovian systems*, Ph.D. thesis, Freie Universität Berlin, (2001).
- [3] S. L. KALPAZIDOU, *Cycle representations of Markov processes*, vol. 28, Springer Science & Business Media, 2007.
- [4] A. NIELSEN, *Computation schemes for transfer operators*, Ph.D. thesis, Freie Universität Berlin, (2015).
- [5] C. SCHÜTTE AND M. SARICH, *Metastability and Markov State Models in Molecular Dynamics: Modeling, Analysis, Algorithmic Approaches*, vol. 24 of Courant Lecture Notes, American Mathematical Soc., 2013.
- [6] G. VAUQUELIN, *Rebinding: or why drugs may act longer in vivo than expected from their in vitro target residence time*, Expert Opinion on Drug Discovery, 5 (2010), pp. 927–941.
- [7] M. WEBER, *Meshless methods in conformation dynamics*, Ph.D. thesis, Freie Universität Berlin, (2006).
- [8] M. WEBER, *A subspace approach to molecular markov state models via a new infinitesimal generator*, Habilitation thesis, Freie Universität Berlin, (2011).
- [9] M. WEBER, A. BUJOTZEK, AND R. HAAG, *Quantifying the rebinding effect in multivalent chemical ligand-receptor systems*, The Journal of Chemical Physics, 137 (2012), 054111.
- [10] M. WEBER AND K. FACKELDEY, *Computing the minimal rebinding effect included in a given kinetics*, Multiscale Modeling & Simulation, 12 (2014), pp. 318–334.