Master thesis

Computing the minimal rebinding effect for nonreversible processes

Zur Erlangung des akademischen Grades Master of Science

Susanne Röhl

Berlin, den 8. Januar 2017

Contents

1 Introduction **5**

1.1 Drug Design . . . 5
1.2 Rebinding Effect in Drug Design . . . 5
1.3 Minimizing the Rebinding Effect . . . 5
1.4 Markov State Models . . . 5
1.5 Literature . . . 6
1.6 Structure . . . 6

2 Markov State Models **7**

2.1 Markov Process . . . 7
2.2 Transfer Operator . . . 7
2.3 Metastability / Dominant Spectrum . . . 8
2.4 Markov State Model . . . 8
2.4.1 Galerkin Projection . . . 8
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 . . . 12

4 Rebinding Effect in a Given Kinetics **13**

4.1 Receptor-Ligand System . . . 13
4.2 Rebinding Effect . . . 13
4.3 Molecular Kinetics as a Projection . . . 14
4.3.1 Galerkin Projection . . . 14
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 Potentiel Energy = (stable/ metastable/ invariant) region

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

Deuflhard/ 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): Rebinding 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 Randnotiz

2 Markov State Models

In this chapter, we will introduce the mathematical fundamentals which are necessary to describe the behaviour of molecular systems.

The notation of this chapter is 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

Definition 2.1 (Markov Process)**.** *A stochastic process X_t is a Markov process if... state space $\mathbb{X}$*

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...

Proposition 2.2. *bla*

Lemma 2.3. *bla*

Satz 2.4. *bla*

Reversibility

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

2.2 Transfer Operator

transition function

To describe the evolution/ propagation of a Markov process, we introduce the transfer operator.

skillnad transfer operator generator

Definition 2.5 (Transfer Operator)**.** *The semigroup of propagators or forward transfer operators $P^t : L^r(\mu) \to L^r(\mu)$ with $t \in T$ and $1 \leqslant r \leqslant \infty$ is defined as follows:*

$$\int_A P^t v(y)\mu(dy) = \int_{\mathbb{X}} v(x)p(t,x,A)\mu(dx) \tag{2.1}$$

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

Definition 2.6 (Infinitesimal Generator)**.** *The semigroup of propagators or forward transfer operators $P^t : L^r(\mu) \to L^r(\mu)$ with $t \in T$ and $1 \leqslant r \leqslant \infty$ is defined as follows:*

$$L\nu = \lim_{t \to \infty} \frac{P^t \nu - \nu}{t}. \tag{2.2}$$

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

2.3 Metastability / Dominant Spectrum

Der Spektralansatz ist also hervorragend geeignet, um Metastabilität bei Markovprozes- sen zu charakterisieren, bringt aber zwei Nachteile mit sich. Zum einen ist das Resultat nur auf reversible Prozesse angewendbar, 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. Da- her 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 □

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 chose an appropriate/ desired/ favoured ansatz space $\chi = \{\chi_1, \ldots, \chi_n\} \subset L^2(\mu)$.

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 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, al- lerdings 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 ge- eigneten Basis des Unterraums ein. Sowohl für die Funktionen einer Teilung der Eins als auch für die charakteristischen Funktionen wird im Folgenden der Buchstabe verwen- det, was jeweils gemeint ist, sollte allerdings immer aus dem Kontext klar sein.

Im Falle einer vollständigen Partition lässt sich also über den projizierten Transfer- operator eine Markovkette definieren, welche sich auf den Mengen der Partition bewegt. Diese Markovkette behält einige wichtige Eigenschaften des Prozesses auf X,unter an- derem 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 Pro- jektion 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 Uberschreitung der Barriere bedeutet, wird auch als Recrossing Problem bezeichnet und ist in...

-¿ 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?

$$\begin{tikzcd}
P(\tau) \arrow[r, "\tau \to \tau k"] \arrow[d, "\text{projection}"] & (P(\tau))^k \arrow[d, "\text{projection}"] \\
P_C(\tau) \arrow[r, "\tau \to \tau k"] & (P_C(\tau))^k
\end{tikzcd}$$

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

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 .

what is a (metastable) cycle?

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

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

4.2 Rebinding Effect

In fact, a stochastic process describing a receptor-ligand molecular system IS NOT necesarrily 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 rebing again thain to stay dissolved.

Zusammenhang zum Absatz davor?

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 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 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.

4.3 Molecular Kinetics as a Projection

Mathematical description of molecular kinetics

Boltzmann distribution

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

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 $V(q)$.

membership functions, overlapping partial densities

Transfer Operator

Galerkin Projection

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

But this is spoiling the Markov Property of the process.

4.4 Minimizing the Rebinding Effect

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

With the tools from chapter 2 (Schur Decomposition) we will solve this problem here for nonreversible processes (NESS processes).

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.