LABORATOIRE
JEAN KUNTZMANN
MATHÉMATIQUES APPLIQUÉES - INFORMATIQUE

Site Campus – Bâtiment IMAG
700 avenue centrale
38400 Saint Martin d'Hères

CS40700 - 38058 Grenoble Cedex 9
France

Sergei Grudinin
CNRS researcher, LJK Laboratory, Grenoble, France

e.mail : Sergei.Grudinin@univ-grenoble-alpes.fr

Grenoble, Sep 29 2024

To the Politechnika Wrocławska University,

The thesis of M. Daniel Wiczew presents a novel methodology and computational technique for the [redacted] with molecular dynamics. The proposed method enhances the sampling by [redacted] This allows us to [redacted] energy landscapes, construct their low-dimensional representations and Markov-state models, and finally compute transition rates and kinetics constants. The thesis demonstrated the applications of the method on test systems and also on an atomistic simulation of a voltage-gated channel Kv 1.2.

The thesis is composed of 12 main chapters. **It begins** with the introduction, where the author reviews the fields of molecular dynamics simulations and artificial intelligence. **Then** the author also presents it in French and **continues** the discussion of the limitations of [redacted] using classical molecular dynamics and a review of [redacted] methods. The author also introduces the main innovation of his work, the [redacted] technique. **In the next chapter,** the author introduces [redacted] and its main architectures, discusses training the models in [redacted] spaces and its limitations, and also its application to molecular dynamics data, more precisely to the discovery of reaction coordinates. **Chapter 5** is devoted to the challenges of studying ion channel kinetics using molecular simulations. The author introduces voltage-gated ion channels, Markov state models, and the main system of study in the thesis, the Kv1.2 potassium voltage-gated ion channel. The author also reviews the main challenges in the kinetics discovery of Kv1.2. **Chapter 6** is devoted to the technical description of the [redacted] algorithm using [redacted] The chapter contains a description of the [redacted] and their experimental validation. It also describes an artificial test system for the validation of the method and molecular systems that have increasing difficulty. **Chapter 7** presents [redacted] applied to the dynamics of Kv1.2. It starts with a discussion of the system features that are chosen. Then, the authors present the molecular dynamics setup and the force field. He also discusses the choice of solvent parameters, the lipid membrane, and general

simulation parameters. **Chapter 8** describes the methodology for the kinetic model of VSD Kv1.2. It presents [redacted] and their training for the discovery of slow coordinates, feature selection in Kv1.2, the loss function, cross-validation analysis, and the selection of the final model. Then, the author presents the Markov state model in the reduced space for Kv1.2 and its analysis and validation. **Chapter 9** presents the results of the [redacted] when applied to the toy 2D potential. It also compares the novel technique with a few state-of-the-art methods and demonstrates its advantages. The author then continues with the results of small molecular systems, [redacted]. Then the author presents the discovered kinetics of Chignolin and its analysis. **Chapter 10** continues with the results of kinetic data of Kv1.2. The author starts with the analysis of Kv1.2 sampling showing the gating charge estimation results and free energy landscape estimation. Then, in **chapter 11**, the author discusses the kinetic model of Kv1.2. He starts with a discussion of the feature's importance and continues with the analysis of the [redacted] Then, the author presents the Markov state model results and its validation and optimization. In the next subchapter, the author delves into the kinetics of Kv1.2 using the Perron Cluster Cluster Analysis and the mean passage time estimation. **Chapter 12** concludes the work by listing the novel developments in the thesis, its main results, and the discovery of the kinetics of Kv1.2.

The thesis is overall very pedagogically written with lots of background information and many high-quality illustrations. I truly enjoyed the organization and the format of the manuscript. I will certainly recommend the thesis, if it becomes open, as background material to Ph.D. students who start their projects in molecular dynamics simulations. I have just a few major questions and also some minor remarks regarding the typos in the manuscript.

The thesis will be defended to obtain a doctorate from the Université de Lorraine and I strongly support the defence.

A few major comments:

– I am still curious how the unbiased transition rates are estimated from the biased sampling simulations. Could you please explain this point better?

– Can you comment on other types of [redacted] methods, for example, those used in robotics applications based on rapidly exploring random trees (RRT)?

– A minor comment in the introduction: Can the Curse of Dimensionality be explained better? What is the origin of the data in Figure 4.1? I suppose the behavior of curves depends very much on the distribution of data... What about the importance of [redacted] which requires a linear amount of [redacted]

Regarding other minor points:

- The current electronic version of the manuscript has misplaced (miscompiled) sidenotes, for example, on page 8"
- In fact, [redacted] in MD is an NP-hard problem., with the size of the protein. This is however not always the case, as there are large proteins with few well-defined conformations." etc
- It will also be helpful to see visual guidance for hyperlinks in the text.
- Reference 51 has an incorrect author list :)
- Section 4 - maybe you can discuss [redacted] and [redacted]?
- Page 31 -- sc can help
- page 34 "one has to construct a [redacted] (three for reversible [redacted] so there are 350002 parameters in fact. While some [redacted] can reduce " - this statement is not strictly correct. One can use incomplete [redacted] factorization, which will be linear in the number of features times the number of components.
- Pages 42,43 Subsection ?? (??)).
- "This allow to escape "-> allows
- Page 50: "If the lag time is too short, the system might not be memoryless, violating the Markovian assumption" - can you please explain more, why it won't be memoryless? Can you still derive the stationary states etc using the eigendecomposition?
- Page 58 "these test may" -> these tests
- page 65 - "Network, $f(S)$ Figure 6.2." parentheses missing?
- "the work of B et al. [37], "- provide the full name?
- Page 85 - several typos in the introduction, e.g. "Where the theory be- hind MSM is described in the in detail. "," objective of section Section 8.1", etc
- "low-dimensional k macrostates or N dimensional vectors" -> of N-dimensional vectors?
- Page 87 - can you please explain once again what is lambda in $g(\lambda) = -\lambda 2$?
- Page 88 - "Nevertheless, data regarding the [redacted] search "[redacted] tuning? Or [redacted] optimization?
- "that they suffers " - suffer
- page 89 - Did you use Batch Normalization? Can you please comment on this?
- Page 91 - [redacted] - please once more specify to which [redacted] Please also comment if the gradient wrt to [redacted]
- Page 99 - What happens if the matrix A is not invertible? Can the matrix be rank-deficient? Is there a theory about this?
- Figure 9.5 caption - please provide a description for each color.
- page 109 [redacted]
- Caption 11.3 "Subsection ??"
- page 120 - a part of the sentence is missing, "A low percentage of change (below 0.1 or 10"

Yours faithfully,

Sergei Grudinin