Received: 17 June 2021 | Accepted: 20 September 2021

DOI: 10.1002/pamm.202100102

MSM building and projection for the analysis of time-resolved spectra

Renata Sechi[1,*], **Marcus Weber**[1,**], and **Karsten Heyne**[2,***]

[1] Zuse Institute Berlin, Takustraße 7, 14195 Berlin
[2] Free University Berlin, Arnimallee 14, 14195 Berlin

Understanding the kinetics between the components of time-resolved spectra is a crucial step in the study of photo-activated processes. However, modeling the kinetics requires usually some a priori knowledge about the system. In our approach, we build a Markov State Model (MSM) from the spectral data, and obtain a Koopman transition matrix $K(\tau)$. With genPCCA, an invariant subspace projection, we project the process into its metastable components. The result of the application of genPCCA is a transition matrix $K^c(\tau)$, from which we can read the transition probability between the metastable components of the reaction. We discuss the application of this analysis method to the transient absorption spectrum of brominated Al-corrole.

1 Markov State Modeling (MSM) of time-resolved spectral data

This work applies the Markov State Modeling (MSM) to the analysis of experimental time-resolved spectral data. The construction of a MSM enables the computation of a transition matrix $K^c(\tau)$, which describes the transition probability between the metastable components of the spectrum and therefore the probability that a dynamical pathway might occur.
As a difference to the broadly applied analysis procedures [2], which involve a model assumption for the kinetic of the metastable components, the construction of a MSM and its genPCCA projection yield to know (a) the probability of multiple kinetic pathways for the reaction, (b) to which extent a metastable conformation χ_i characterizes the system at time t.
The spectra are considered as a N-dimensional, autonomous Markovian trajectory M, $M \in \mathbf{R}^{N\times T}$, with N being the number of measured wavelengths, and T being the number of delay times at which the spectrum has been detected. Considering the spectra as trajectories means that they are functions of the state space Ω.
The evolution of functions f (observables) in a dynamical system is governed by the Koopman operator $\mathcal{K}(\tau)$. This operator is defined as the expectation value of the evolution of a system's observable f for a given, fixed lag time τ. In order words, $\mathcal{K}(\tau) : L^\infty(\Omega) \to L^\infty(\Omega)$,

$$f_{t+\tau}(a) = [\mathcal{K}(\tau) f_t](a) = \mathbb{E}[f_t(\tilde{X}_{t+\tau}) \mid \tilde{X}_t = a] = \int_\Omega p_\tau(b|a) f_t(b) d\mu(b).$$

To treat real data, the Koopman operator $\mathcal{K}(\tau)$ is approximated via Galerkin projection into the transition matrix $K(\tau)$. The Galerkin projection is computed with a Voronoi tessellation of the state space, by picking some spectra at a given delay time as centers of the Voronoi cells.
In order to interpret the dataset, we aim to identify the metastable components of the spectra and their kinetics transition pathways. The Koopman matrix $K(\tau)$ does not describe the transition probability between the metastable components; rather, it describes the transition probabilities between fast micro components, which may not have alone a physical or chemical meaning for the reaction.
GenPCCA [4] is a fuzzy spectral clustering algorithm that relates micro-components to metastable ones. genPCCA is based on invariant subspaces, given by the Schur decomposition of $K(\tau)$. The Schur decomposition allows to treat non-reversible processes ($K(\tau)$ has complex eigenvalues) and numerically non-feasible eigendecompositions (singular matrix).
The genPCCA projection is based on the selection of the leading Schur-vectors X of the Koopman matrix $K(\tau)$, which are invariant subspaces of $K(\tau)$. Each Schur-vector in X represents a metastable component. GenPCCA uses X to compute by a linear transformation the membership values χ of the system to the metastable components. Finding the membership values χ means that the linear transformation is a matrix $\mathcal{A} \in \mathbb{R}^{r\times r}$ such that [3,4]

$$\chi = X\mathcal{A} \quad \chi_j(i) \in [0,1], i \in 1...n; \quad \sum_{j=1}^{r} \chi_j(i) = 1.$$

With the membership functions in χ the transition matrix of the metastable reaction components $K^c(\tau)$ is obtained by Galerkin projection on basis χ :

$$K^c(\tau) = \langle \chi, \chi \rangle_\pi^{-1} \langle \chi, K(\tau)\chi \rangle_\pi. \tag{1}$$

* Corresponding author: email sechi@zib.de, phone +49 30 841 85 534
** email weber@zib.de, phone +49 30 8 41 85 189
*** email karsten.heyne@physik.fu-berlin.de, phone +49 30 838 56107

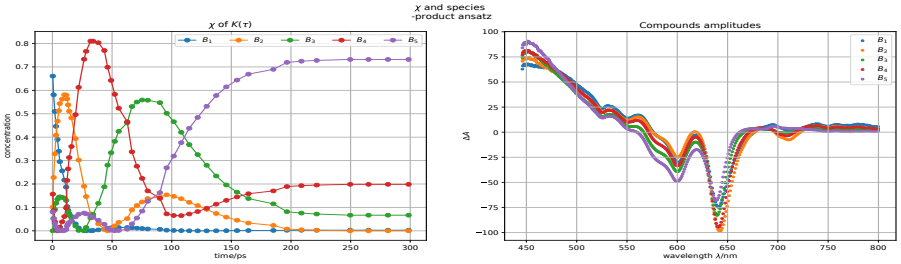

Fig. 1: The spectral amplitudes W_i on the right are obtained with a product ansatz $M = \chi W$.

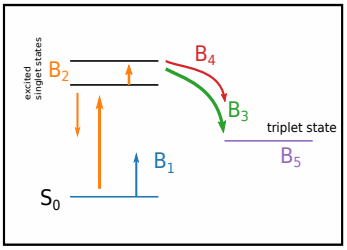

Fig. 2: Most probable kinetic scheme, color coding as in fig. 1

2 Results

In this application example, the VIS pump–supercontinuum probe spectrum of a Al(tpfc-Br8)(py)2 (brominated Al-corrole) is modelled with MSM and PCCA+ projection. This experimental dataset has been analyzed and interpreted in [1]; the system in ground state is excited to higher singlet electronic-states, then the excitation energy relaxes to the first singlet excited states (first 20 ps); the system has a spin transition to the triplet state (intersystem crossing) with a time constant of 95 ps.
The MSM Koopman matrix $K(\tau)$ is based on a discretization of the state space Ω into 50 Voronoi cells and the model is projected via genPCCA into 5 metastable conformations selected by the leading Schur values of $K(\tau)$. The projected transition matrix for this process is computed by equation (1). The resulting membership vectors are displayed in Fig. 1, on the left side. By assuming a product ansatz $M = \chi W$, it is possible to estimate the the amplitudes W_i of the metastable components (Fig 1, right).
The obtained coarse-grained transition matrix $K^c(\tau)$ helps in understanding the relation between the metastable components. The combined interpretation of the χ, the amplitudes, and the most probable transition pathway is schematized in figure 2. From figure 1 (left), the metastable components alternates in sequential way $B_1 \to B_2 \to B_4 \to B_3 \to B_5$. B_1, B_2, B_4 arise and full decay in the first 70 ps. The first state, B_1 is the start conformation and represents the excitation of the system, since it arises only at the beginning and decays completely after 25ps. B_2 shows again an increase of the negative signal at 650 nm and a negative signal around 710 nm. Its amplitude curve is also slightly red-shifted w.r.t. the other metastable components at 650 nm and 710 nm, perhaps because of a broader negative signal given by stimulated emission and ground state bleaching together. The metastable conformation B_4 shows decreased negative signal at 710 nm and increased positive signal (ESA) in the $>$ 710 nm region. B_4 is a transition state between some of the molecules in the sample moving to the T_1 state and others still in the singlet-system. B_3 has a similar shape as B_4, but shows a loss in the ESA-signal in the $>$ 750 nm region. Finally, B_5 shows a positive signal in the 700-710 nm region, a decreased negative bleaching signal and an increased bleaching signal in the 570 nm region, due to the loss of ESA. This suggests that the population in B_5 is mainly in triplet state and it is starting to transfer back to the S_0, since the bleaching is starting to recover. However, what one observe is not the recovery of the bleaching signal, but the recovery of the SE signal. This piece of information can be gained by observing the transition matrix. B_5 is a sink state for the dynamics and the transition probability to B_1 is zero. The fact that B_3 and B_5 have an absorbing-state character suggests that they represent to different degrees the population in triplet.

3 Conclusion

The construction of a MSM can be used to unravel the kinetics of time-resolved spectra. A MSM can be built by considering each measured spectrum at a given time t as a point of multidimensional trajectory. After discretization, the process is projected onto its metastable components by genPCCA. The metastable components represent meaningful processes of the reaction in the spectra. Our method is feasible for treating real data, because it does not require an a priori kinetic model assumption: The relation between the components is given by the projected transition matrix $K^c(\tau)$. Nevertheless, the model and the Schur-decomposition can be influenced by the noise in the data and some components, identified as discerned ones, can actually represent the same process. However, the combined analysis of the membership vectors χ, their amplitudes W, and the transition matrix allows for the determination of the physical meaning of the metastable components.

Acknowledgements Open access funding enabled and organized by Projekt DEAL.

References

[1] T., Stensitzki, Y., Yang, A., Berg, A. Mahammed, Z., Gross, and K., Heyne ,Structural Dynamics **3**, 043210 (2016).
[2] I. H.M., van Stokkum, D. S., Larsen, R., van Grondelle, BBA - Bioenergetics, **1657**,82-104 (2004).
[3] S., Röblitz, M., Weber, Adv Data Anal Classif **7**, 147–179 (2013).
[4] K., Fackeldey, and M., Weber. WIAS Rep **29**, 70-80 (2017).

