Collective variables in complex systems: from molecular dynamics to agent-based models and fluid dynamics

Péter Koltai

Department of Mathematics, University of Bayreuth

Brin MRC workshop on *Rare Events: Analysis, Numerics, and Applications*
Feb 28, 2023

Variously joint work with
R. Banisch, A. Bittracher, M. Dellnitz, B. Hamzi, J. Heitzig, S. Klus, M. Lücke, P. Maity, N. Molkenthin, J. Schumacher, Ch. Schütte, S. Weiss, S. Winkelmann

Dimension reduction

General idea:
search for "embedding" $\Phi : \mathbb{D} \subset \mathbb{X} \to \mathbb{Y}$, where $\dim \mathbb{Y} \ll \dim \mathbb{X}$, and Φ discards irrelevant information.

$\dim \approx 10.000$

Image: [Lafon 2004]

$\longrightarrow$

$\dim = 1$

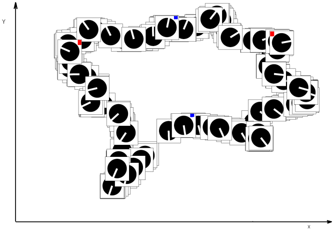
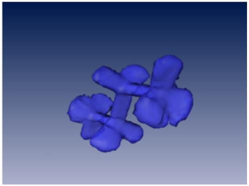
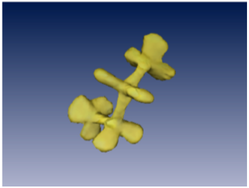


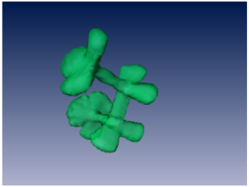
Image: [dela Porte et al 2008]

Manifold learning (E.g. ISOMAP, Laplacian eigenmaps, diffusion maps [Coifman, Lafon 2006], and many others)

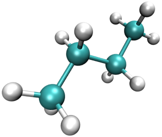
Conformation dynamics

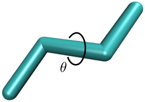


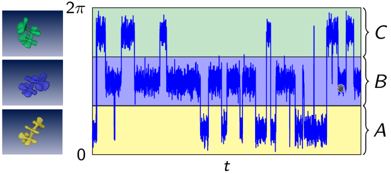




- Metastable conformations? (high-dim.!)
- Transitions between them (timescales)?

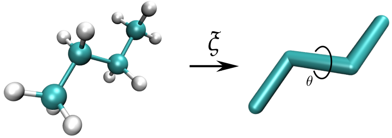






Collective variable (CV)

Transition manifold / Effective dynamics



Collective variables (CV) describe progress between metastable sets
Desire:

Slow timescales of $\xi(X_t) \approx$ Slow timescales of X_t

[Bittracher, K., Klus, Banisch, Dellnitz, Schütte 2018]

$$dX_t = -\nabla V(X_t)\, dt + \sigma\, dW_t \qquad (*)$$

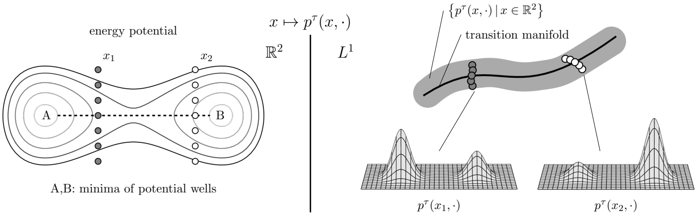
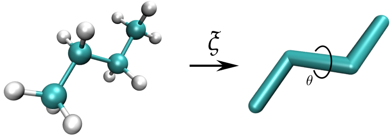


Image: [Bittracher, Schütte 2020]

Transition manifold / Effective dynamics



Collective variables (CV) describe progress between metastable sets
Desire:

Slow timescales of $\xi(X_t) \approx$ Slow timescales of X_t

[Bittracher, K., Klus, Banisch, Dellnitz, Schütte 2018]

$$dX_t = -\nabla V(X_t)\,dt + \sigma\,dW_t \qquad (*)$$

Theorem: For $(*)$, if $\{p^\tau(x,\cdot) \,|\, x \in \mathbb{X}\} \subset L^1$ is ε-close to a r-dim. manifold, then there is a r-dim. CV reproducing slow timescales up to $\mathcal{O}(\varepsilon)$.

- Constructive† approach to learn CVs
- Quantitative goodness measure

---

† Cf. also [Mardt, Pasquali, Wu, Noé 2018], [Lusch, Kutz, Brunton 2018], [Chen, Tan, Ferguson 2018], ...

Existence of CVs

- Systems with multiple time scales

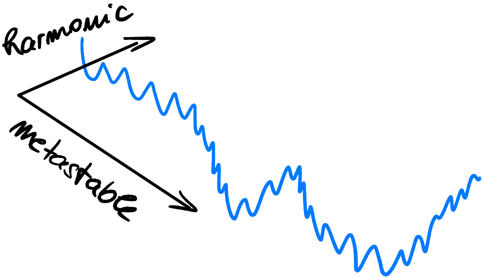
$$
\begin{aligned}
dX_t &= f(X_t, Y_t)dt + \sigma\, dW_t \\
\varepsilon dY_t &= g(X_t, Y_t)dt + \sqrt{\varepsilon}\sigma\, dB_t
\end{aligned}
$$

- Metastable systems

$$
\sigma(\mathcal{L}): \quad \lambda_0 \geqslant \ldots \geqslant \lambda_K \gg \lambda_{K+1} \geqslant \ldots
$$

[Bittracher, Mollenhauer, K., Schütte (to appear)]

- In between...



...many metastabilities replaced by slow effective diffusion

Computing collective variables

1) For given anchor points x^i, compute realizations of $X^i_\tau \sim p^\tau(x^i,\cdot)$

2a) Embed $p^\tau(x^i,\cdot)$ in finite dimensions [Hunt, Kaloshin 1999]

2b) Kernel embedding of $p^\tau(x,\cdot)$'s: [Bittracher, Klus, Hamzi, K., Schütte 2021]

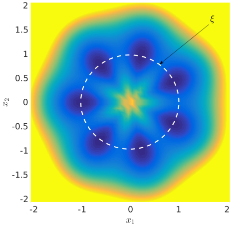
Low-dimensional variation $\rightsquigarrow$ less samples

[Bittracher, Mollenhauer, K., Schütte (to appear)]

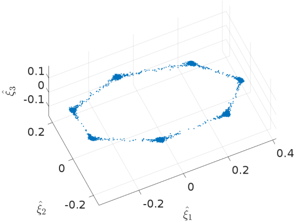
3) Manifold learning (diffusion maps) to find geometry [Coifman, Lafon 2006]

Example in $\mathbb{R}^{10}$:

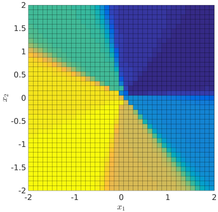
- Multi-well potential in x_1, x_2, quadratic potential in $x_3, \ldots, x_{10}$



potential



embedded TM

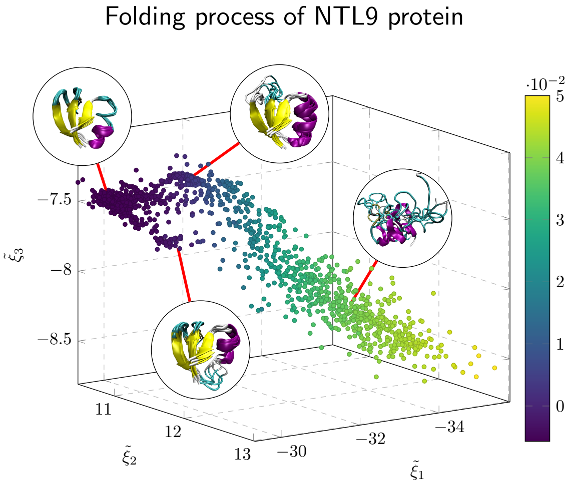


CV

Properties

- Short simulations on intermediate time scale τ
- Anchor points in transition region $\rightsquigarrow$ rare events
- Evaluation at new anchor points w/o recomputing the CV
  (out-of-sample extension / Nystrom method)
- Local method: charts the "visible" part of state space
- Exploration / Towards rare event detection: continue from boundary

CVs for MD



[Bittracher, Banisch, Schütte 2018]

Coordinates of transport and mixing

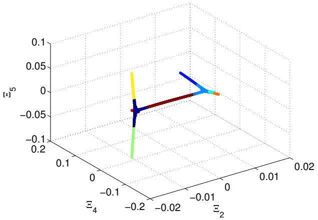
How does [Froyland 2015], [Karrasch, Keller 2020]

$$dX_t^\varepsilon = v(t, X_t^\varepsilon)\, dt + \varepsilon\, dW_t$$

spread across trajectories of

$$\dot{x}_t = v(t, x_t)?$$

CVs

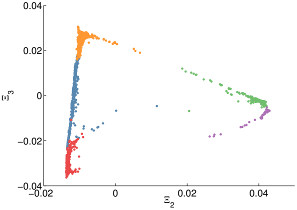


Clusters / coherent sets

Theorem: $\lim_{\#\text{data}\to\infty}$ (trajectories + diffusion maps) = "ε-generator" of $\phi_{-t}X_t^\varepsilon$

[Banisch, K. 2017], [K., Renger 2018]

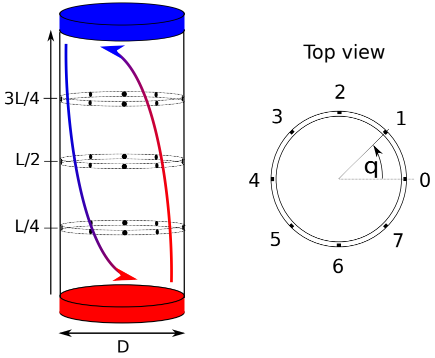
CVs in the ocean

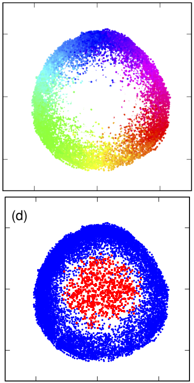


[Banisch, K. 2017]

CVs in Rayleigh–Bénard convection: experiment

- RB convection in cylinder $\rightsquigarrow$ disc
- Learn structure and dynamics (cf. also [Berry, Giannakis, Harlim 2015])





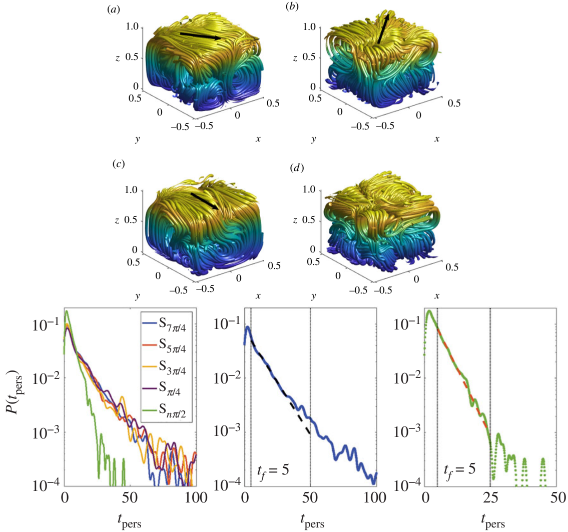
[Weiss, Ahlers 2011]

[K., Weiss 2020]

- Effective state space & dynamics without prior physical knowledge

CVs in Rayleigh–Bénard convection: simulation

RB convection in cube $\rightsquigarrow$ discrete LSCs



[Maity, K., Schumacher 2022]

Noisy voter model

- Why? Macroscopic observable (polls) $\rightsquigarrow$ macroscopic dynamics
- Existence of CVs: stochastic analysis + random graphs (ongoing)

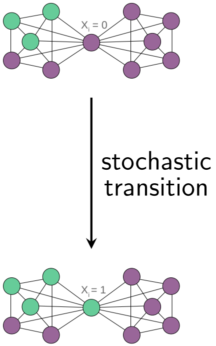
- Random opinion switch

$$\mathbb{P}[X_i(t+dt) = m' \mid X_i(t) = m] =$$
$$f(\text{opinion fraction of agent } i\text{'s neighbors})$$

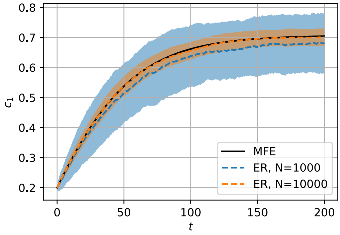
- Macroscopic observable (collective variable)

$$c_m(t) = \frac{1}{N}\,\#\{i \mid X_i(t) = m\}$$

- Continuous-time versions available
- CVs for agents:
  - Opinion-behavior model: [Helfmann, Heitzig, K., Kurths, Schütte 2021]
  - Emergent space in agent-based models: [Kemeth, Bertalan, Thiem, Dietrich, Moon, Laing, Kevrekidis 2022]



Concentration in voter models



- Complete graph: [Kurtz 1978]

$$c_n'(t) = \sum_{m \neq n} c_m(t) \Big(r_{m,n} c_n(t) + \tilde{r}_{m,n} \Big) (e_n - e_m)$$

almost surely as $N \to \infty$

- Random graph: variation decreasing w/ graph size N

Theorem. Convergence to mean field model in probability for Erdős–Rényi[1] random graphs ($N \to \infty$) for

$$\text{edge prob} = \omega\left(\frac{\log N}{N}\right).$$

Theorem. Convergence to mean field model in probability for random regular graphs ($N \to \infty$) for

$$\text{degree} = \omega(1).$$

[Lücke, Heitzig, K., Molkenthin, Winkelmann (preprint)]

[1] Also: stochastic block model and heterogeneous population

Summary & Outlook:

- Framework for collective variables in molecular/fluid/agent systems
- When do CVs exist? Is there a universal framework? (~~reversibility~~?)

Acknowledgments: to my colleagues & funding:

Berlin Mathematics Research Center MATH+

DFG Turbulent Superstructures

Scaling Cascades in Complex Systems

R. Banisch and P. Koltai. *Understanding the geometry of transport: diffusion maps for Lagrangian trajectory data unravel coherent sets*. **Chaos** 27(3): 035804, 2017.

A. Bittracher, P. Koltai, S. Klus, R. Banisch, M. Dellnitz, and Ch. Schütte. *Transition manifolds of complex metastable systems: Theory and data-driven computation of effective dynamics*. **Journal of Nonlinear Science**, 28(2): 471–512, 2018.

P. Koltai and S. Weiss. *Diffusion maps embedding and transition matrix analysis of the large scale flow structure in turbulent Rayleigh–Bénard convection*. **Nonlinearity** 33(4): 1723, 2020.

A. Bittracher, S. Klus, B. Hamzi, P. Koltai, and Ch. Schütte. *Dimensionality Reduction of Complex Metastable Systems via Kernel Embeddings of Transition Manifolds*. **Journal of Nonlinear Science** 31, 3. 2021.

L. Helfmann, J. Heitzig, P. Koltai, J. Kurths, and Ch. Schütte. *Statistical Analysis of Tipping Pathways in Agent-Based Models*. **The European Physical Journal Special Topics** 230, 3249–3271. 2021.

P. Maity, P. Koltai, and J. Schumacher. *Large-scale flow in a cubic Rayleigh–Bénard cell: long-term turbulence statistics and Markovianity of macrostate transitions*. **Philosophical Transactions of the Royal Society A** 380:20210042. 2022.

M. Lücke, J. Heitzig, P. Koltai, N. Molkenthin, S. Winkelmann. *Large population limits of Markov processes on random networks*. **arXiv:2210.02934**, 31 pp. 2022.