Reaction Rate theory SOSE24: ISOKANN

Luca Donati

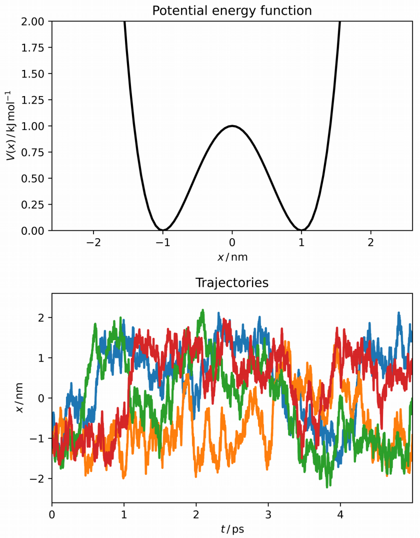
Zuse Institute Berlin
Modeling and Simulation of Complex Processes
Computational Molecular Design

30 May 2024

We consider a dynamical system driven by the SDE

$$dx_t = -\nabla V(x_t)dt + \sqrt{2D}\,dW_t\,, \qquad (1)$$

with $V(x) : \Gamma \to \mathbb{R}$ potential energy function, D diffusion constant and W_t a standard Wiener process.



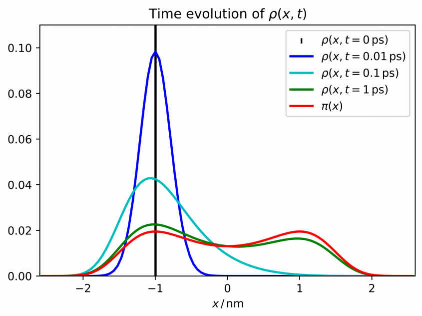
Associated to eq. 1, there exists the PDE

$$\frac{\partial \rho}{\partial t} = \mathcal{Q}^* \rho_t(x) \,, \tag{2}$$

where $\mathcal{Q}^*$ is the *Kolmogorov forward operator*.
The formal solution is

$$\begin{aligned} \rho_{t+\tau}(x) &= e^{\tau \mathcal{Q}^*} \rho_t(x) \\ &= \mathcal{P}_\tau \rho_t(x) \,, \end{aligned} \tag{3}$$

where $\mathcal{P}_\tau$ is the *propagator*.



Stationary distribution:

$$\lim_{t \to +\infty} \rho_t(x) = \pi(x) \,.$$

Similarly, we introduce the adjoint equation for arbitrary observable functions $f(x)$

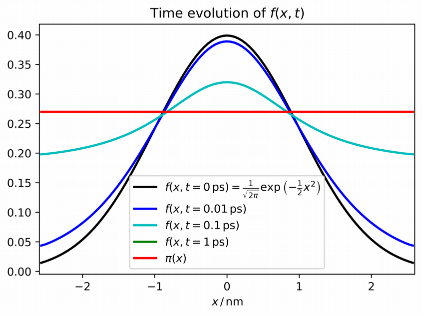
$$\frac{\partial f}{\partial t} = \mathcal{Q} f_t(x) \,, \tag{4}$$

where $\mathcal{Q}$ is the *Kolmogorov backward operator*. The formal solution is

$$\begin{aligned} f_{t+\tau}(x) &= e^{\tau \mathcal{Q}} f_t(x) \\ &= \mathcal{K}(\tau) f_t(x) \,, \end{aligned} \tag{5}$$

where $\mathcal{K}(\tau)$ is the *Koopman operator*:

$$\mathcal{K}(\tau) = (\mathcal{P}_\tau)^*_\pi \,. \tag{6}$$



$$\lim_{t \to +\infty} f_t(x) = \text{const.}$$

- *Membership functions*

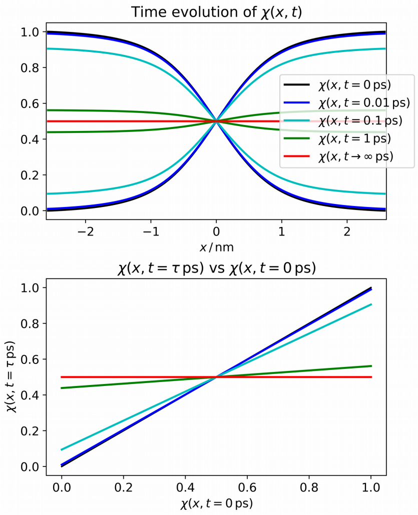
$$\chi(x) : \Gamma \to [0,1].$$

- The Koopman operator acts on $\chi(x)$ as

$$\begin{aligned}\chi_{t+\tau}(x) &= \mathcal{K}(\tau)\chi_t(x) \\ &= a_1\chi_t(x) + a_2\,.\end{aligned}$$

- From linear regression we can estimate a_1 and a_2, and exit rates [1]:

$$\kappa = -\frac{1}{\tau}\log(a_1)\left(1 + \frac{a_2}{a_1 - 1}\right)$$



[1] Marcus Weber, and Natalia Ernst. *arXiv:1708.00679*.

Given an initial function $f_0(x)$:

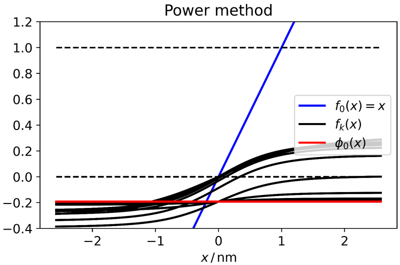
$$\begin{aligned} f_{k+1} &= \frac{\mathcal{K}_\tau f_k}{\|\mathcal{K}_\tau f_k\|} \\ \lim_{k\to\infty} f_{k+1} &= \phi_0(x) \end{aligned}$$

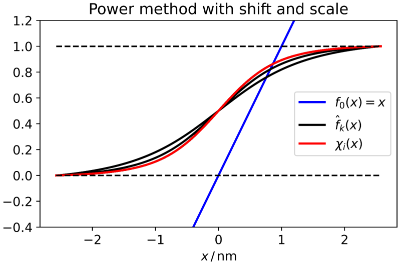
If we apply an appropriate linear transformation $\mathcal{S}$:

$$\begin{aligned} f_{k+1} &= \mathcal{S}\mathcal{K}_\tau f_k \\ \lim_{k\to\infty} f_{k+1} &= \chi_i(x) \quad i = 1, 2 \end{aligned}$$

For a two-state system, $\mathcal{S}$ is given by

$$\mathcal{S}\mathcal{K}_\tau f_k = \frac{\mathcal{K}_\tau f_k - \min\left(\mathcal{K}_\tau f_k\right)}{\max\left(\mathcal{K}_\tau f_k\right) - \min\left(\mathcal{K}_\tau f_k\right)}$$





- In the previous examples, $\mathcal{K}_\tau$ was estimated by SqRA. Alternatively, it can be estimated as

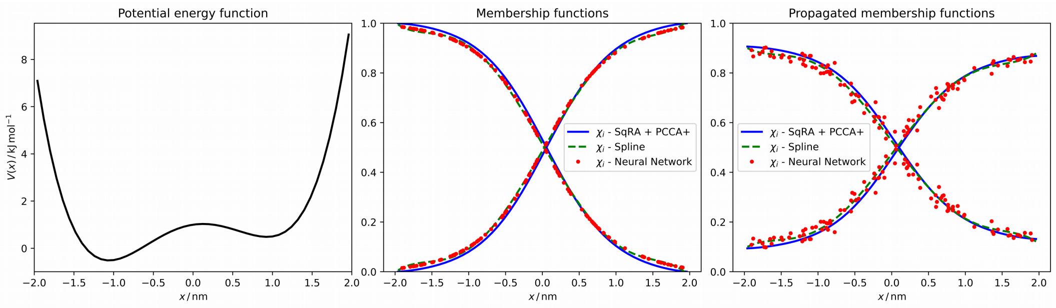
$$
\begin{aligned}
f_{t+\tau}(x) &= \mathbb{E}\left[f_t(x_{t+\tau})|x_t = x\right] \\
&\approx \frac{1}{N}\sum_{n=1}^{N} f_t(x_{t+\tau,n}|x_t = x) \\
&= \bar{f}_t(x_{t+\tau}|x_t = x)\,,
\end{aligned}
$$

  where $x_{t+\tau,n}$ is the end-point of the nth simulation of length τ starting in $x_t = x$.
- The shift-and-scale power method algorithm becomes

$$
f_{k+1}(x_0) = \mathcal{S}\bar{f}_k(x_\tau|x_0 = x)\,.
$$

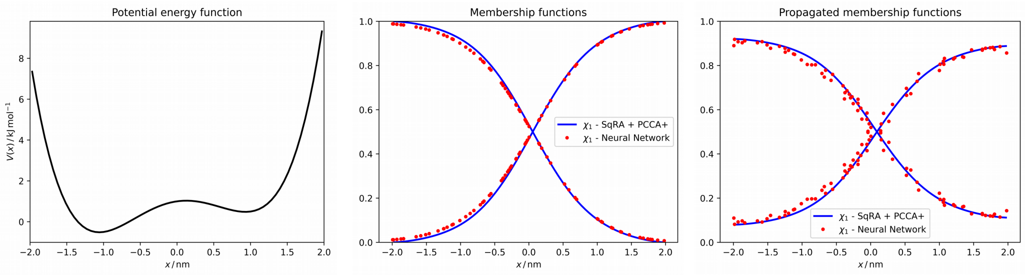
- The initial function is arbitrary, it is useful to consider an interpolating function such as a *spline* or *neural network*.

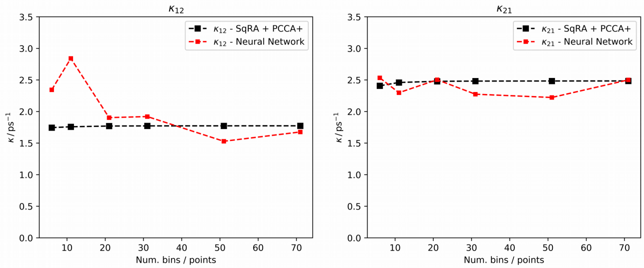
Illustrative example 1D

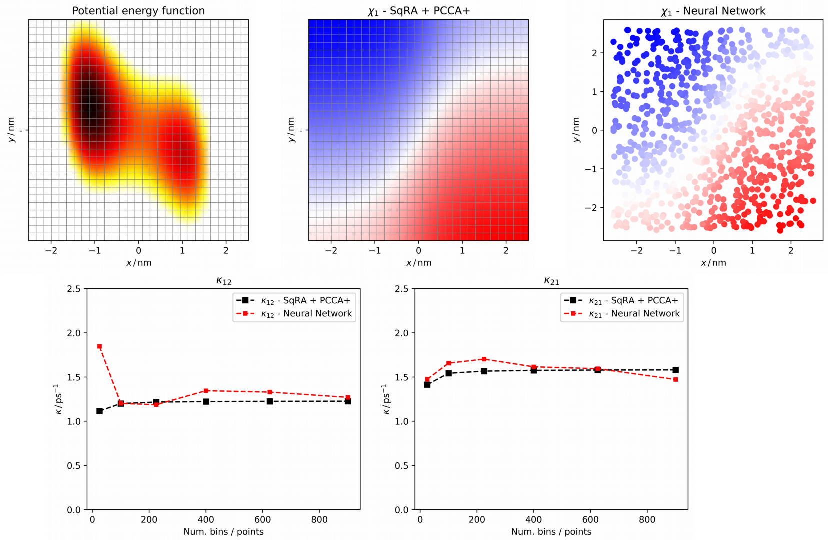


-	SqRA + PCCA+	Spline	Rel. Err.	Neural Network	Rel. Err.
κ_{12} ps^{-1}	1.773	1.811	0.021	1.837	0.035
κ_{21} ps^{-1}	2.483	2.391	0.037	2.414	0.027

ISOKANN vs SqRA 1D






Potential energy function
χ_1 - SqRA + PCCA+
χ_1 - Neural Network
y / nm
x / nm
−2
−1
0
1
2
κ_{12}
κ_{12} - SqRA + PCCA+
κ_{12} - Neural Network
κ_{21}
κ_{21} - SqRA + PCCA+
κ_{21} - Neural Network
κ / ps^{-1}
0.0
0.5
1.0
1.5
2.0
2.5
0
200
400
600
800
Num. bins / points

ISOKANN vs SqRA 3D