Chasing Errors by Diagnostics

Peter Schuster

Institut für Theoretische Chemie, Universität Wien, Austria
and
The Santa Fe Institute, Santa Fe, New Mexico, USA

TBI Seminar

Wien, 25.11.2015

Web-Page for further information:

http://www.tbi.univie.ac.at/~pks

$$P_n(t) \;=\; \mathrm{Prob}\big(\mathcal{X}_{\mathsf{A}}(t) \;=\; n\big)$$

$$\frac{\mathrm{d}P_n(t)}{\mathrm{dt}} \;=\; \sum_{m=0}^{\infty}\Big(W(n\,|\,m,t)\,P_m(t) - W(m\,|\,n,t)\,P_n(t)\Big)\,;\;\; n,m \in \mathbb{N}$$

$$\frac{\mathrm{d}P_n(t)}{\mathrm{dt}} \;=\; w_{n-1}^{+}\,P_{n-1}(t) \;+\; w_{n+1}^{-}\,P_{n+1}(t) \;-\; (w_n^{+} + w_n^{-})\,P_n(t)$$

$$\begin{aligned} w_n^{+} &\equiv \lambda_n \quad \text{for} \quad n \;\to\; n+1 \quad \text{and} \\ w_n^{-} &\equiv \mu_n \quad \text{for} \quad n \;\to\; n-1 \end{aligned}$$

Master equation in chemical reaction kinetics

$$P(\tau, \boldsymbol{\mu}) \;=\; P_1(\tau) \cdot P_2(\boldsymbol{\mu}|\tau)$$

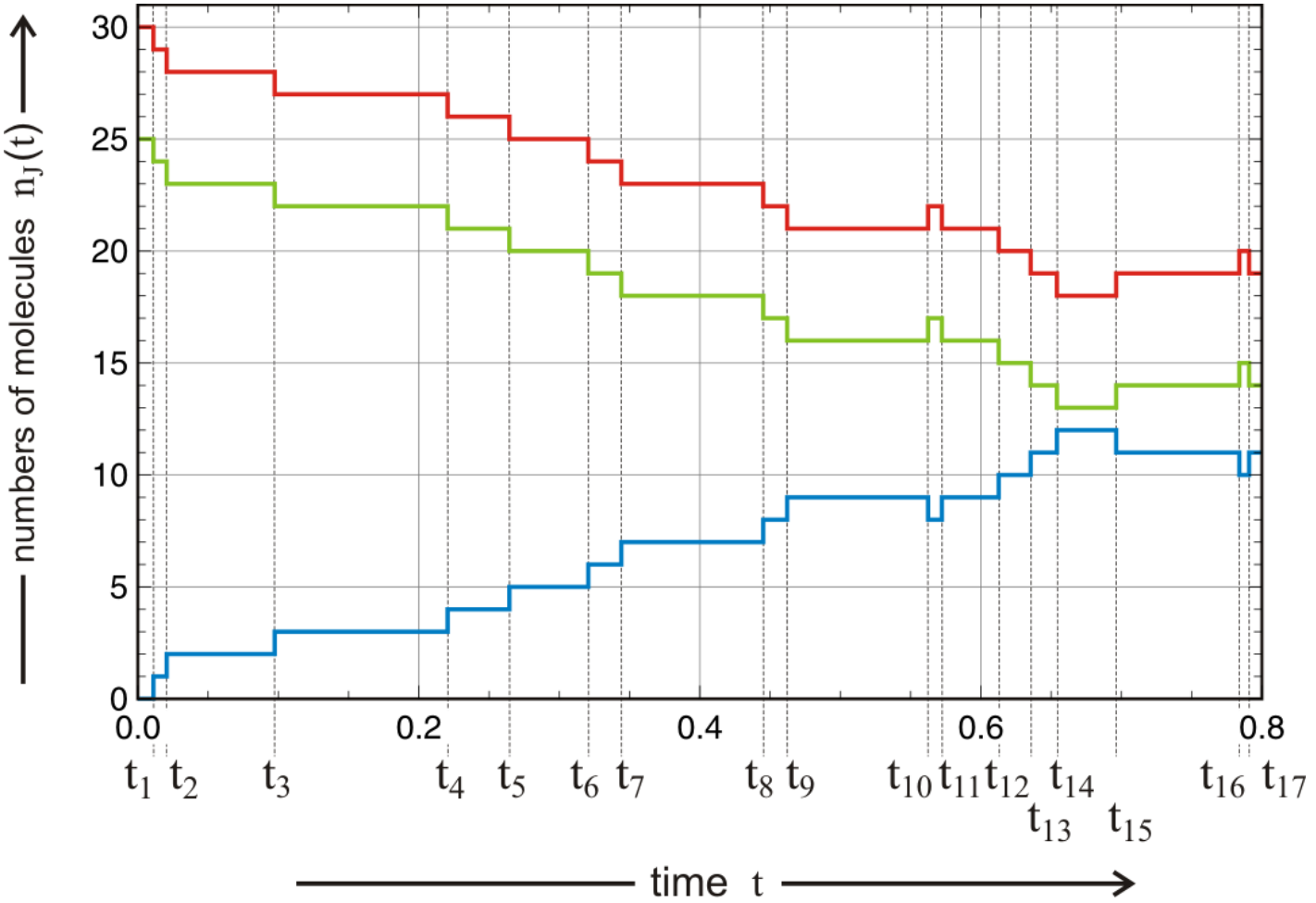
$$P_1(\tau) \;=\; \sum_{\boldsymbol{\mu}=1}^{K} P(\tau, \boldsymbol{\mu}) \qquad P_2(\boldsymbol{\mu}|\tau) \;=\; P(\tau, \boldsymbol{\mu}) \;/\; \sum_{\boldsymbol{\nu}}^{K} P(\tau, \boldsymbol{\nu})$$

$$P_1(\tau) \;=\; \alpha \, \exp\left(-\alpha \, \tau\right), \quad 0 \leq \tau < \infty \quad \text{and}$$

$$P_2(\boldsymbol{\mu}|\tau) \;=\; P_2(\boldsymbol{\mu}) \;=\; \alpha_{\boldsymbol{\mu}} \;/\; \alpha \,, \quad \boldsymbol{\mu} = 1, \ldots, K$$

$$\alpha_{\boldsymbol{\mu}} \;\equiv\; \gamma_{\boldsymbol{\mu}} h_{\boldsymbol{\mu}}(\mathbf{n}) \quad \text{and} \quad \alpha \;=\; \sum_{\boldsymbol{\mu}=1}^{K} \alpha_{\boldsymbol{\mu}} \;\equiv\; \sum_{\boldsymbol{\mu}=1}^{K} \gamma_{\boldsymbol{\mu}} h_{\boldsymbol{\mu}}(\mathbf{n})$$

The stochastic simulation algorithm (SSA, Gillespie et al.)

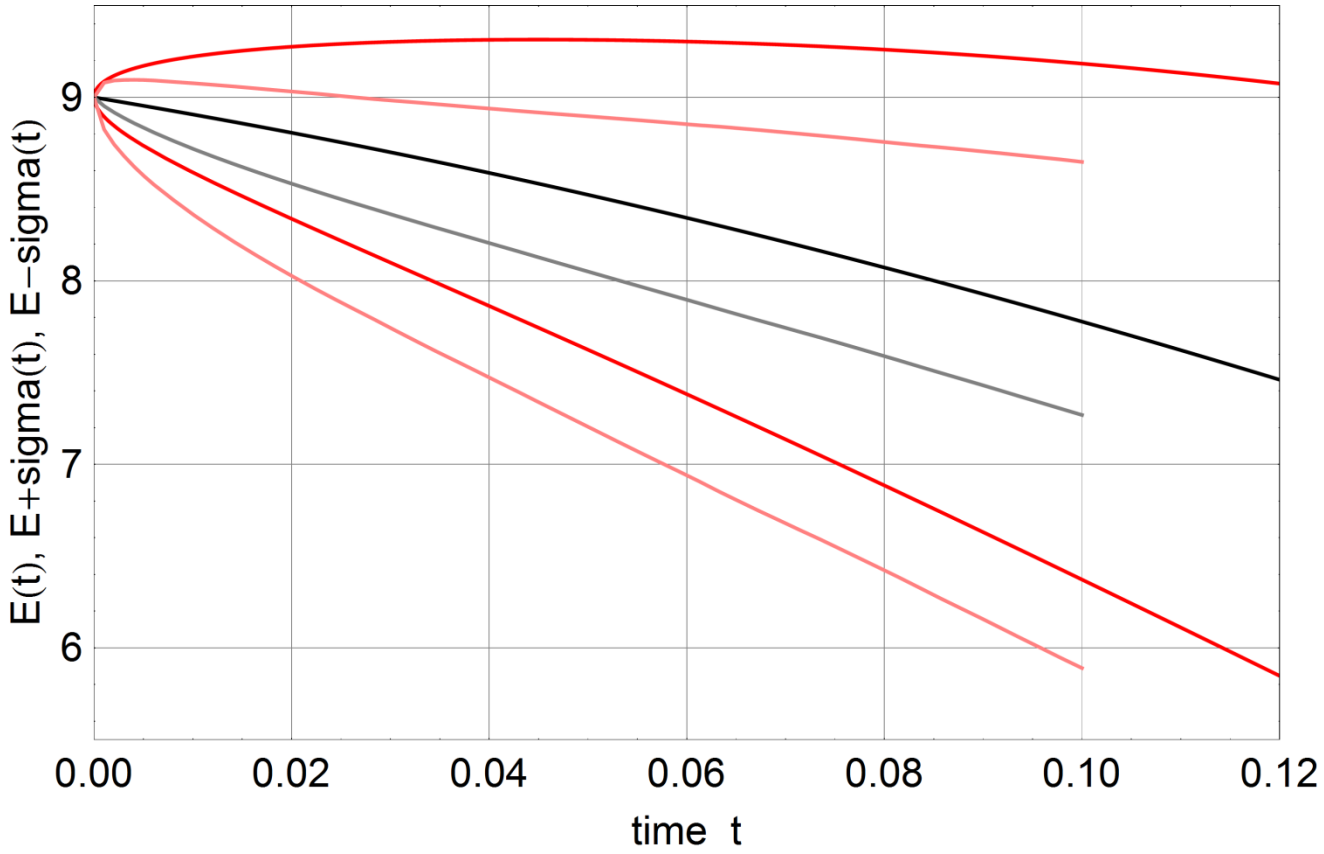


A + B $\leftrightarrow$ C + D: $k = 0.06$, $l = 0.01$, $N = 55$

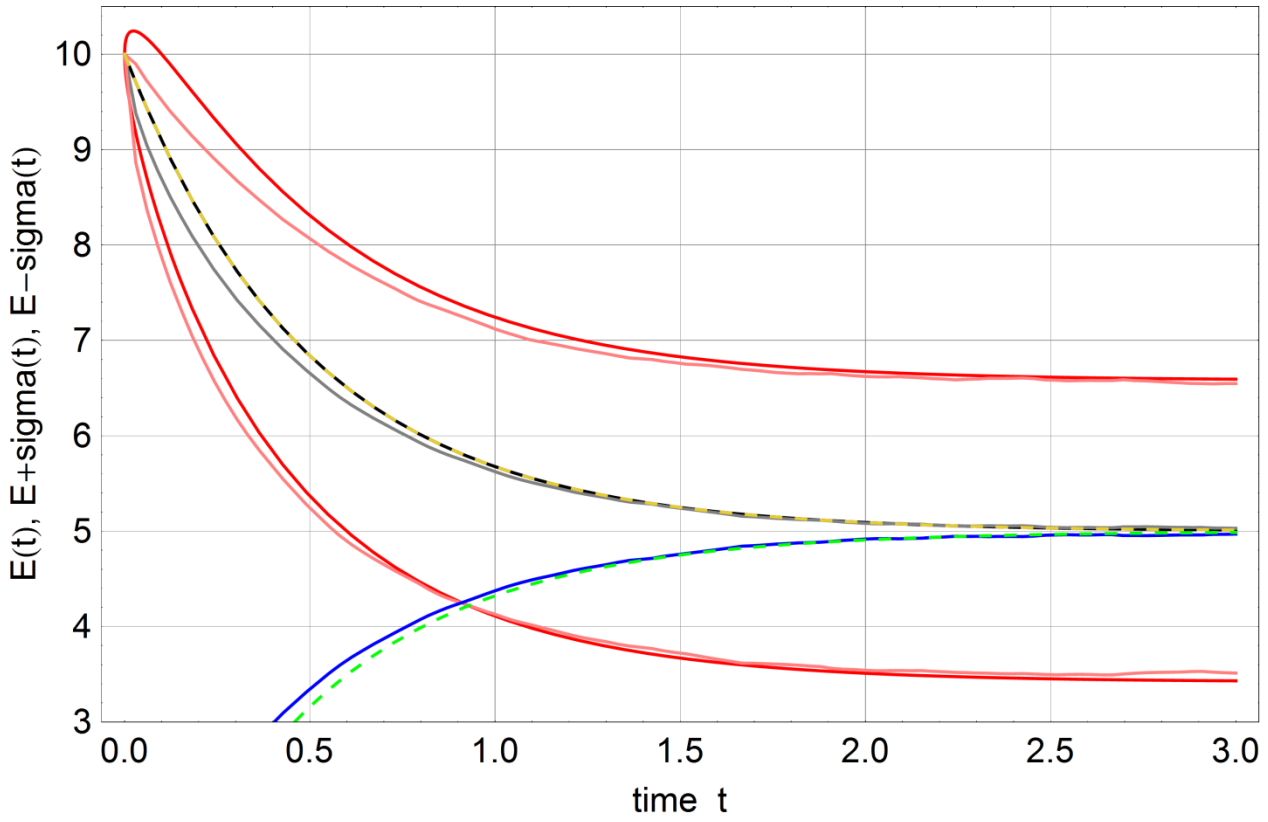
Simulation tool: xSSA-lite

xSSA is a module of the xCellerator simulation package for systems biology of whole cells.

Shapiro, B.E., Levchenko, A., World, E.M.M.B.J., Mjolsness, E.D.: Cellerator: Extending a computer algebra system to include biochemical arrows for signal transduction simulations. Bioinformatics **19**, 677–678 (2003)



A + X $\rightarrow$ 2 X: $k = 1.0, N = 10$



A $\leftrightarrow$ B: $k = 5$, $l = 5$, $N = 10\,000$

Table 1: **Simulation errors of xSSA-lite in the xCellerator package for the reactions** $\mathsf{A} \rightleftharpoons \mathsf{B}$ **.** Shown are the equilibrium expectation values of the random variable $\bar{E}\big(\mathcal{X}_{\mathsf{A}}(t)\big)$ and the width of the one-σ confidence interval $2\sigma\big(\mathcal{X}_{\mathsf{A}}(t)\big)$ obtained by simulation and calculation. Sample size: $N = 1000$ trajectories.

parameters			simulation			calculation		
n_0	k	l	$\bar{\mathrm{E}}(\mathcal{X}_{\mathsf{A}})$	$\bar{\mathrm{E}}(\mathcal{X}_{\mathsf{B}})$	$2\,\bar{\sigma}(\mathcal{X})$	$\bar{\mathrm{E}}(\mathcal{X}_{\mathsf{A}})$	$\bar{\mathrm{E}}(\mathcal{X}_{\mathsf{B}})$	$2\,\bar{\sigma}(\mathcal{X})$
10	5	5	5.005	4.995	3.068	5	5	3.162
10	4	6	5.995	4.005	3.011	6	4	3.098
10	3	7	6.977	3.023	2.778	7	3	2.898
10	2	8	7.949	2.051	2.382	8	2	2.529
10	1	9	8.883	1.117	1.676	9	1	1.897

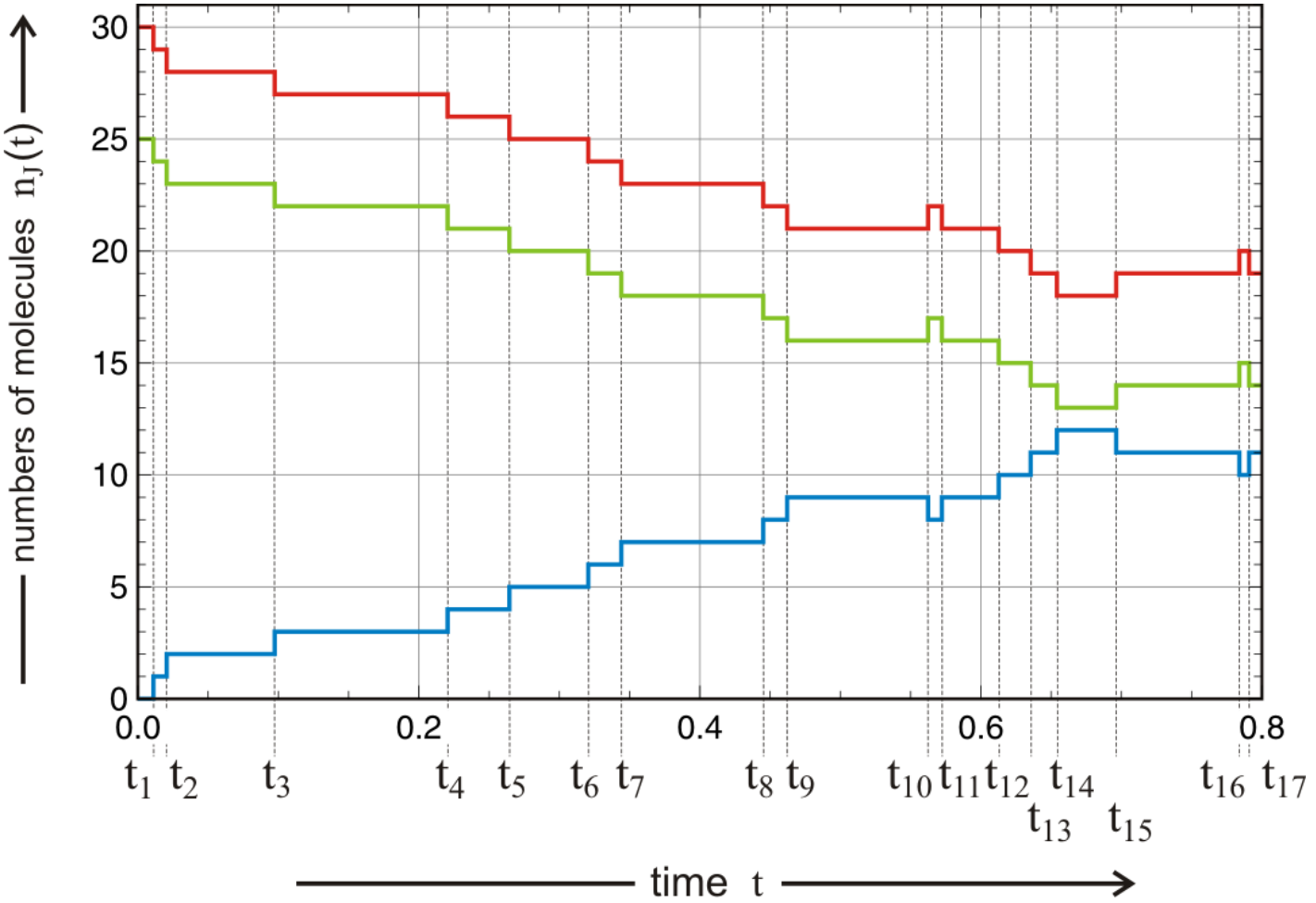
Binomial distribution: $\bar{\mathrm{E}}\big(\mathcal{X}_{\mathsf{A}}\big) = n_0 \dfrac{l}{l+k}$ and $\overline{\mathrm{var}}\big(\mathcal{X}_{\mathsf{A}}\big) = n_0 \dfrac{l \cdot k}{(l+k)^2}$

Expectation values are wrong for asymmetric reaction rates

equilibrium constants $K \neq 1$.

Variances or standard deviations are calculated too small.

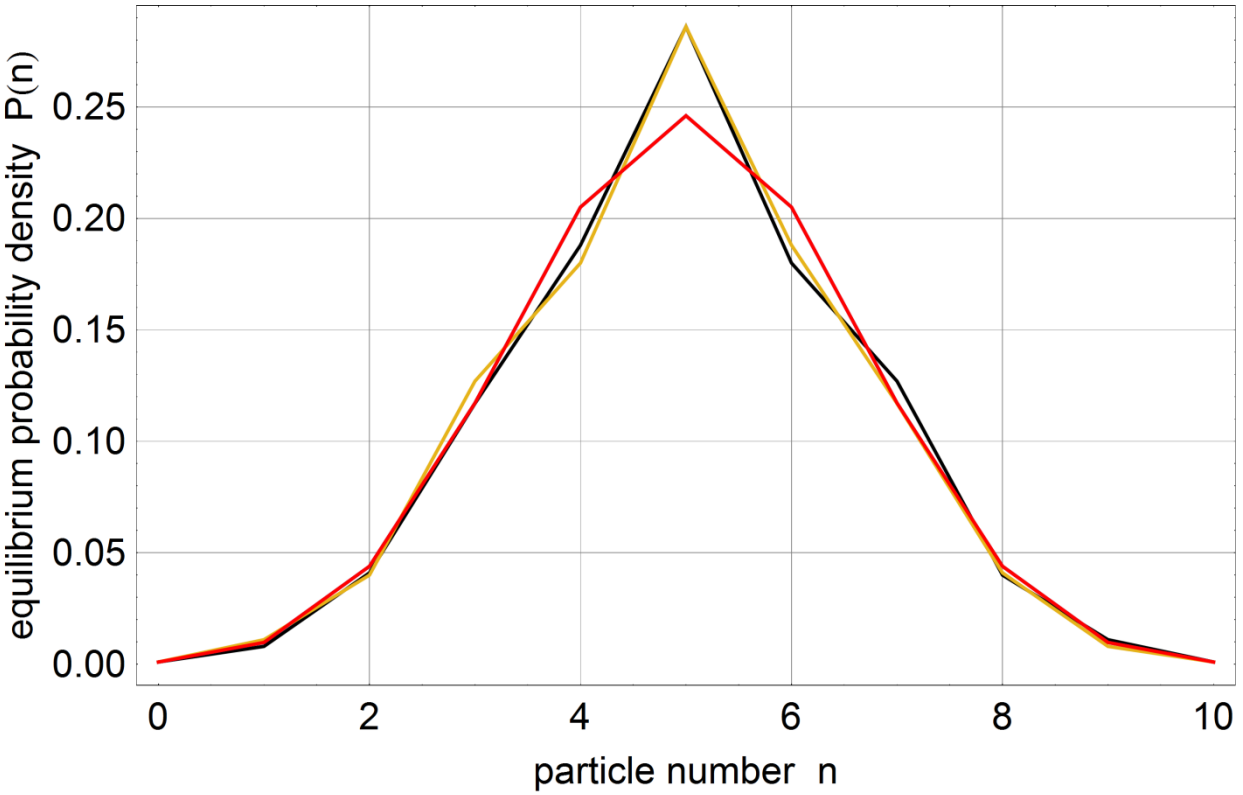
Diagnostics of simulation errors



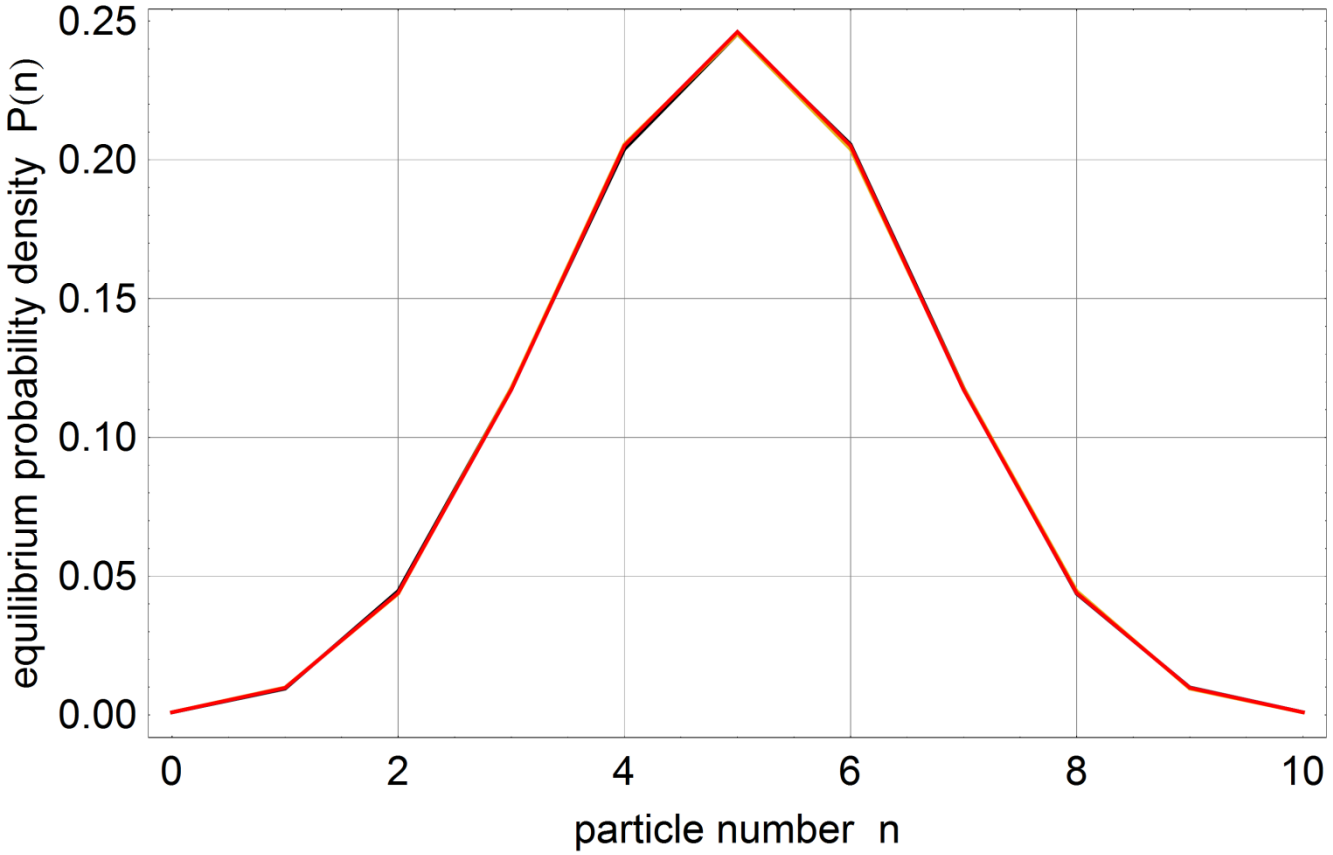
Endpoint statistics

Table 2: **Endpoint statistics of xSSA-lite in the xCellerator package for the reactions** $\mathsf{A} \rightleftharpoons \mathsf{B}$ **.** Shown are the equilibrium expectation values of the random variable $\bar{E}\big(\mathcal{X}_{\mathsf{A}}(t)\big)$ and the width of the one-σ confidence interval $2\sigma\big(\mathcal{X}_{\mathsf{A}}(t)\big)$ obtained by simulation and calculation. Sample size: $N = 1000$ trajectories.

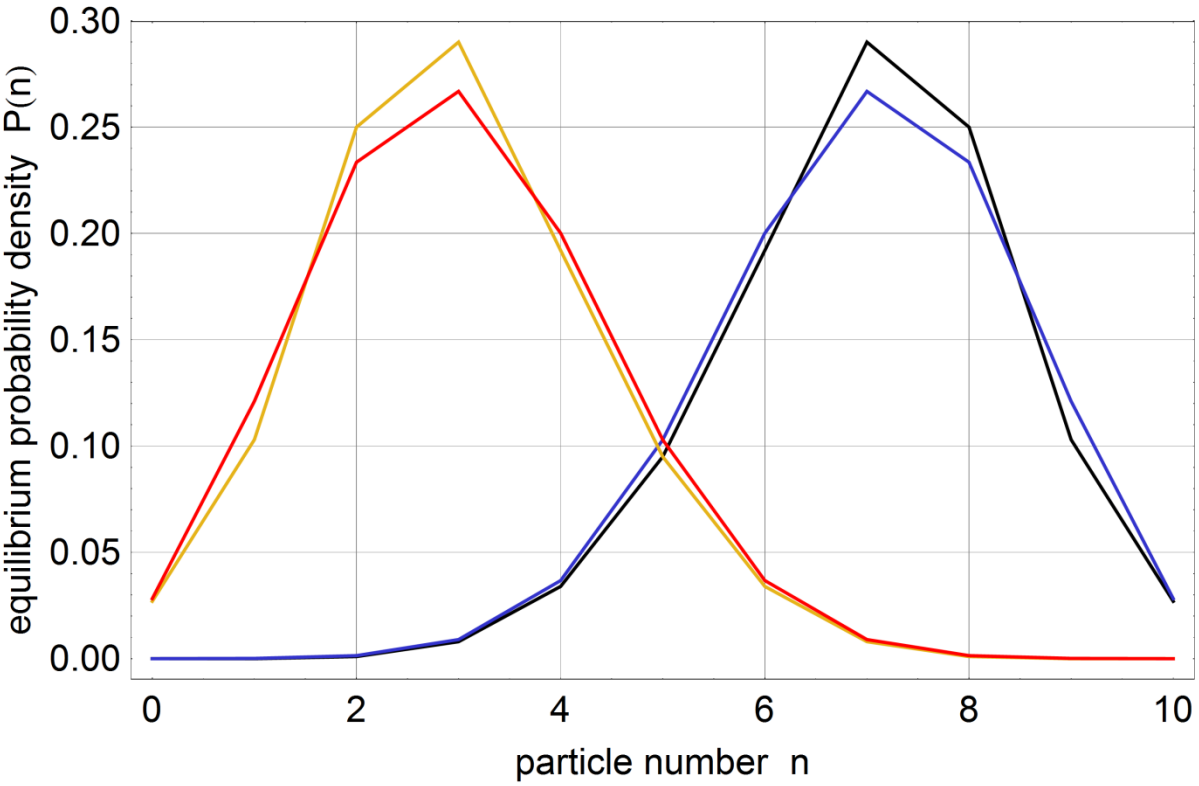
parameters			simulation			calculation		
n_0	k	l	$\bar{\mathrm{E}}(\mathcal{X}_{\mathsf{A}})$	$\bar{\mathrm{E}}(\mathcal{X}_{\mathsf{B}})$	$2\,\bar{\sigma}(\mathcal{X})$	$\bar{\mathrm{E}}(\mathcal{X}_{\mathsf{A}})$	$\bar{\mathrm{E}}(\mathcal{X}_{\mathsf{B}})$	$2\,\bar{\sigma}(\mathcal{X})$
10	5	5	5.001	4.999	3.166	5	5	3.162
10	3	7	6.998	3.002	2.899	7	3	2.898
10	1	9	8.996	1.004	1.903	9	1	1.897



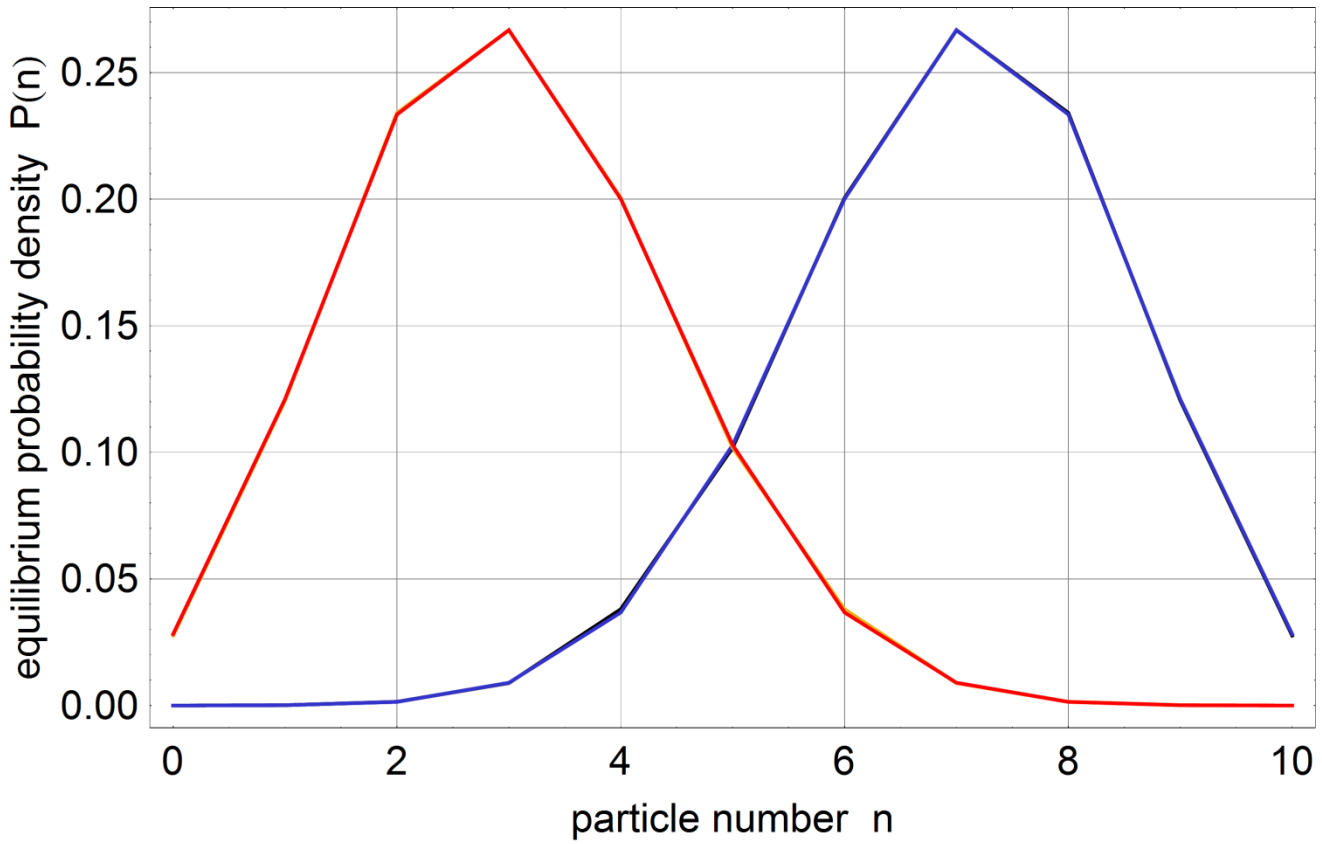
A $\leftrightarrow$ B: $k = 5$, $l = 5$, $N = 1000$



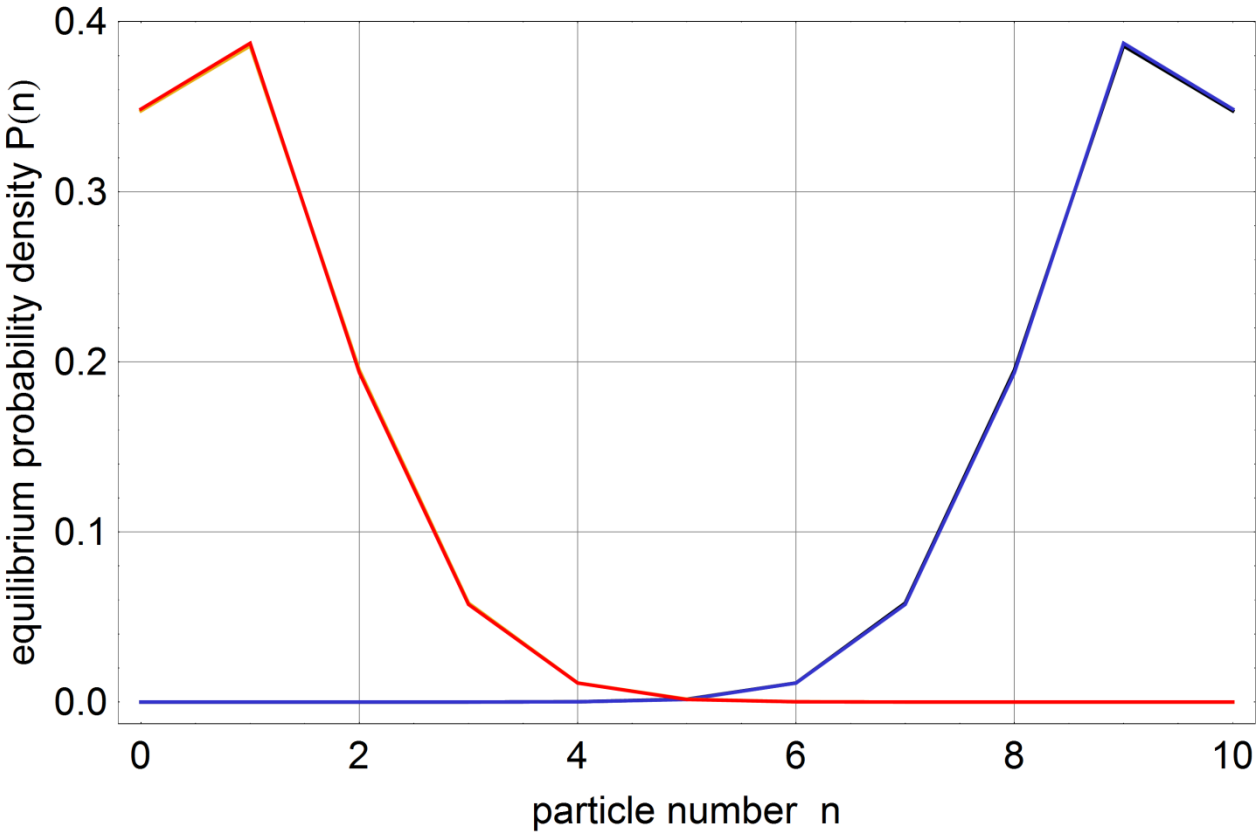
A $\leftrightarrow$ B: $k = 5$, $l = 5$, $N = 100\,000$



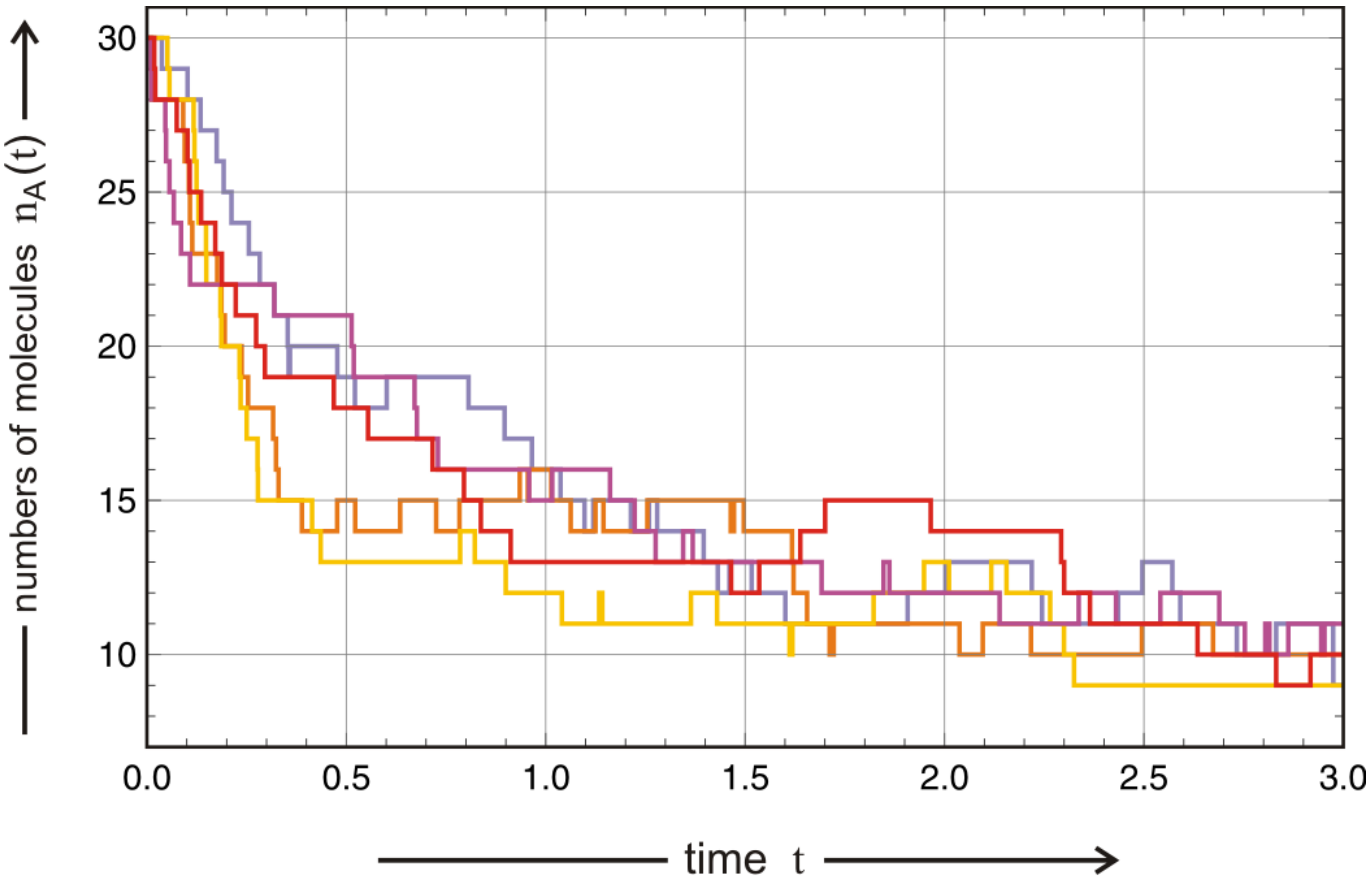
A $\leftrightarrow$ B: $k = 3$, $l = 7$, $N = 1000$



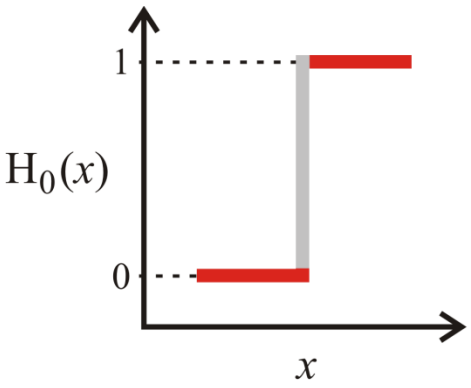
A $\leftrightarrow$ B: $k = 3$, $l = 7$, $N = 100\,000$



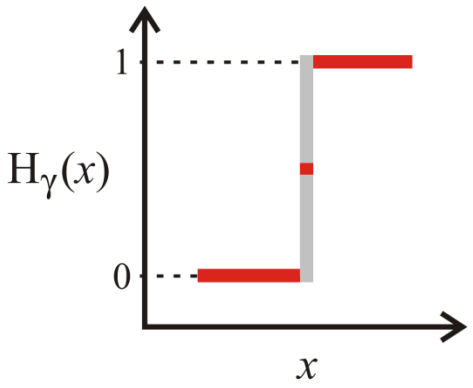
A $\leftrightarrow$ B: $k = 1$, $l = 9$, $N = 100\,000$



A + B $\leftrightarrow$ C + D: $k = 0.06$, $l = 0.01$, $N = 55$

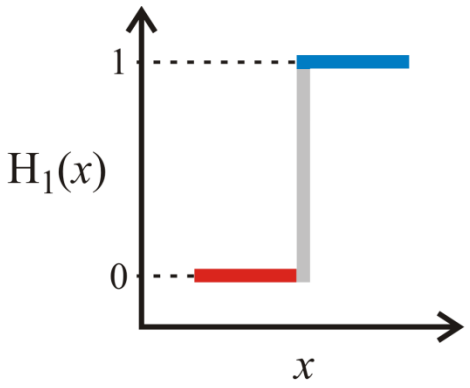


left-hand continuity



no continuity

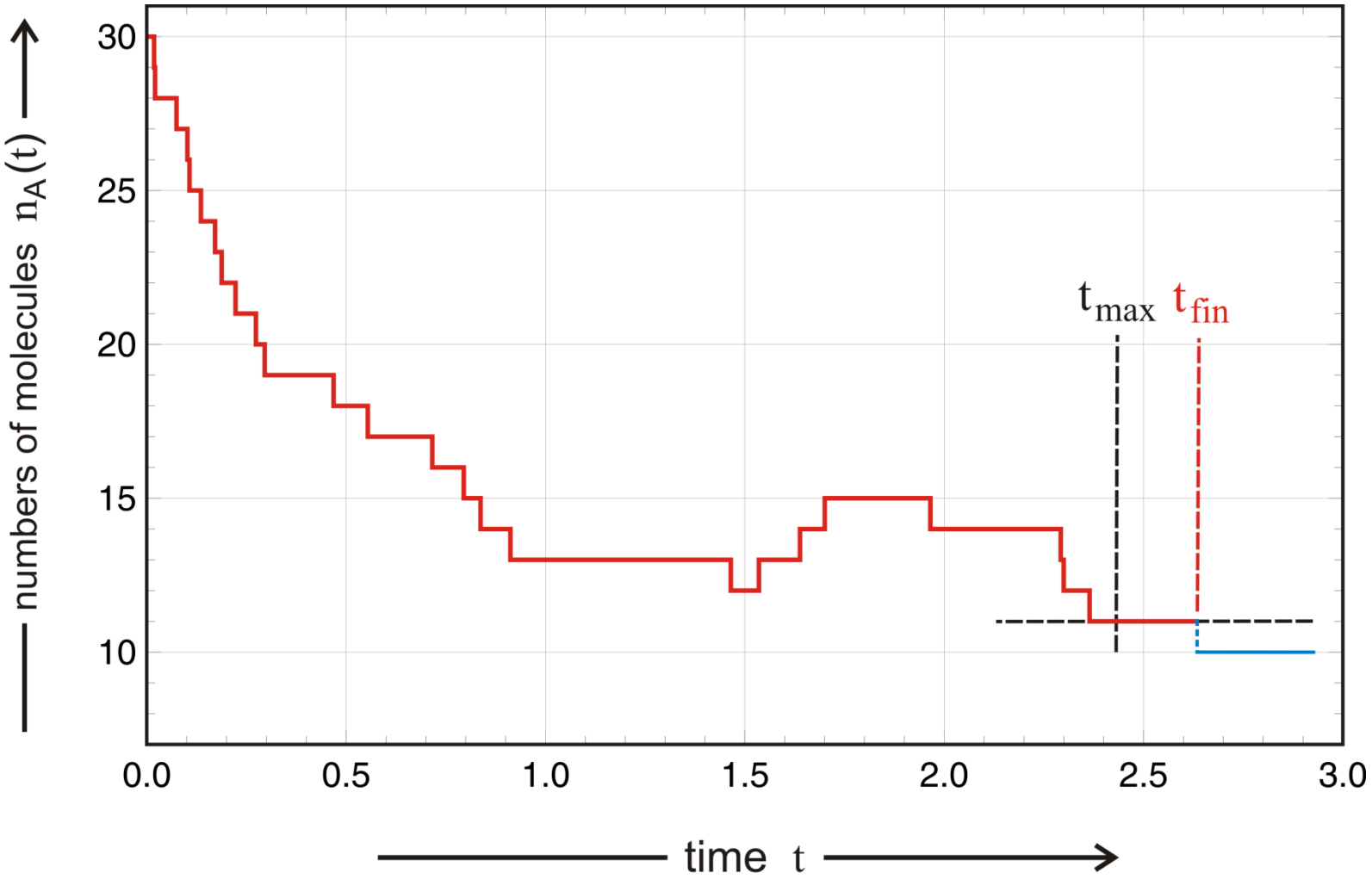
sign function



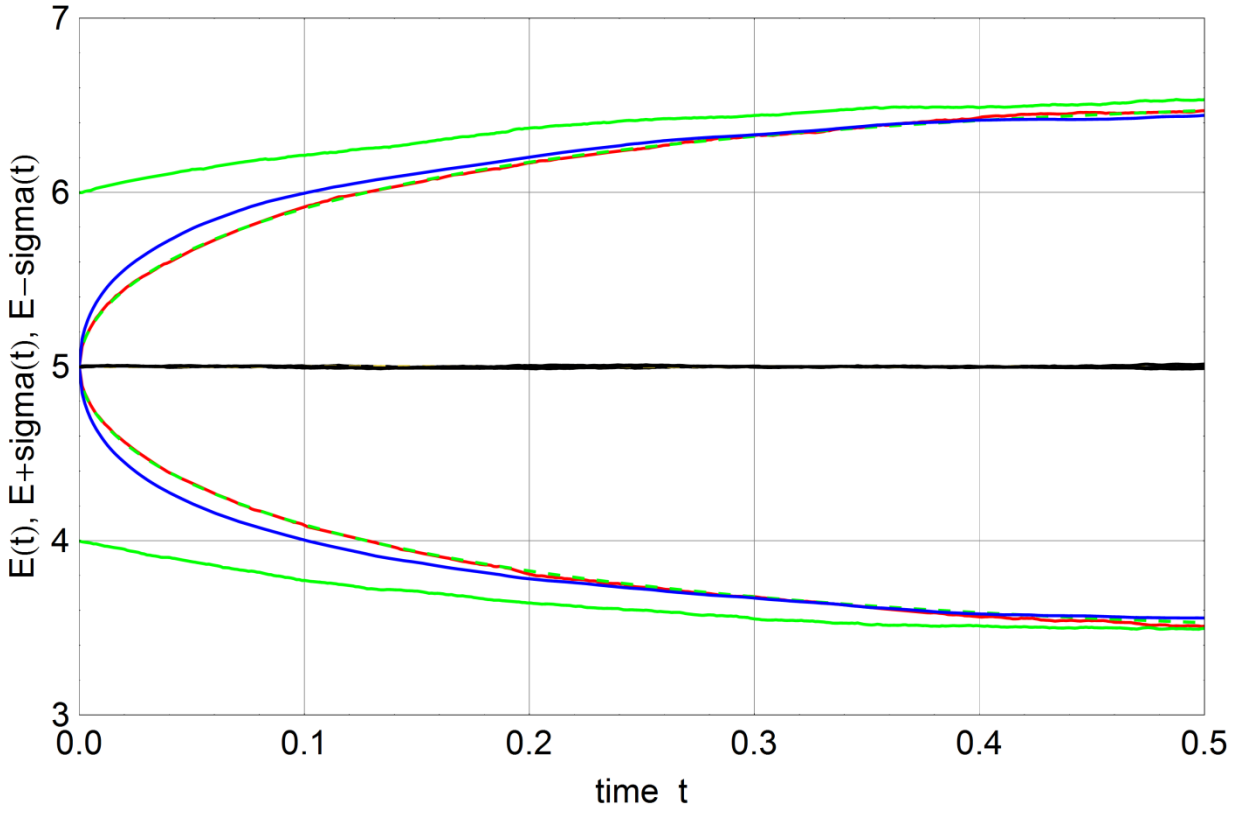
right-hand continuity

càdlàg convention

Continuity in step functions

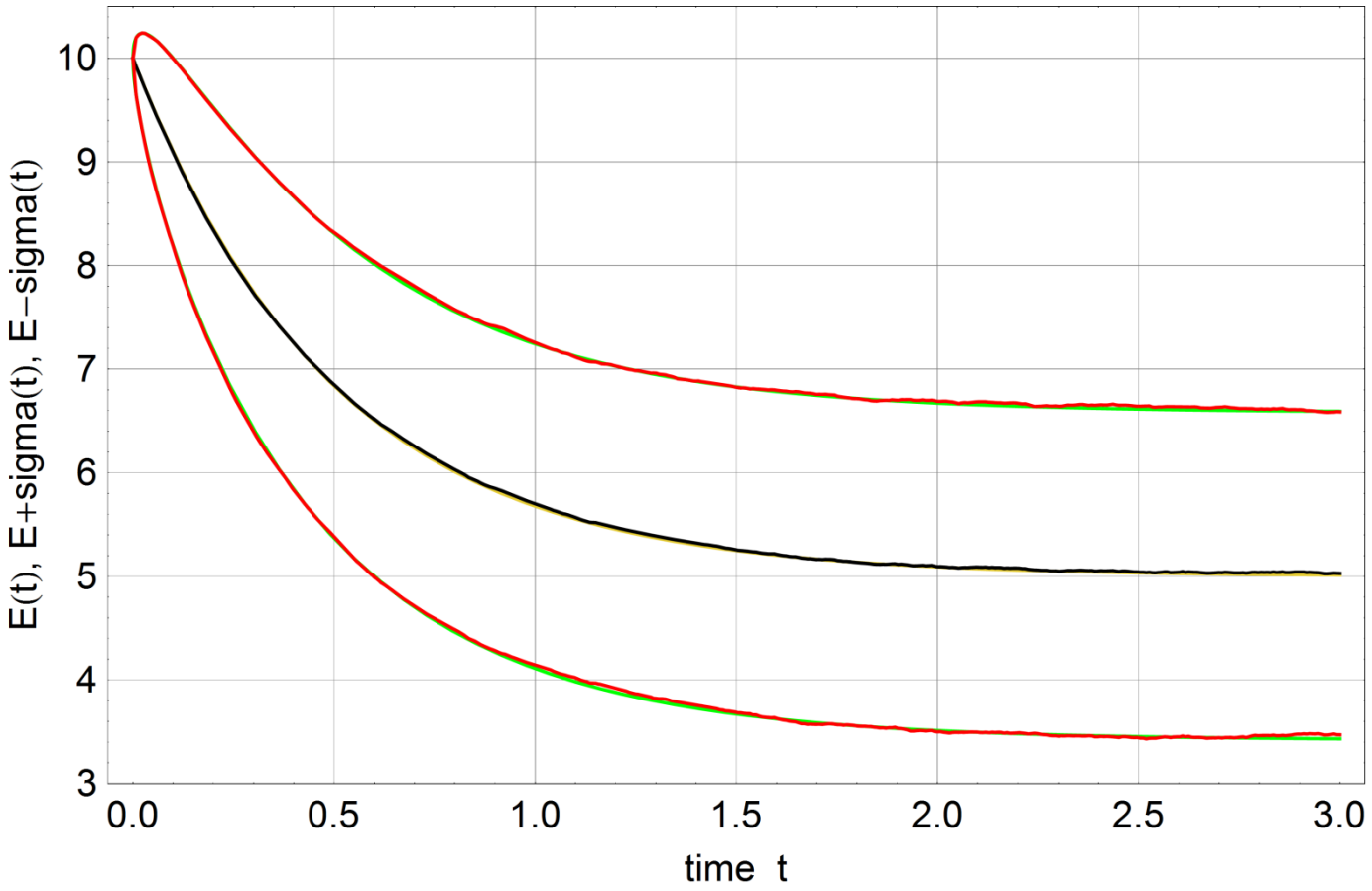


A + B $\leftrightarrow$ C + D: $k = 0.06$, $l = 0.01$, $N = 55$

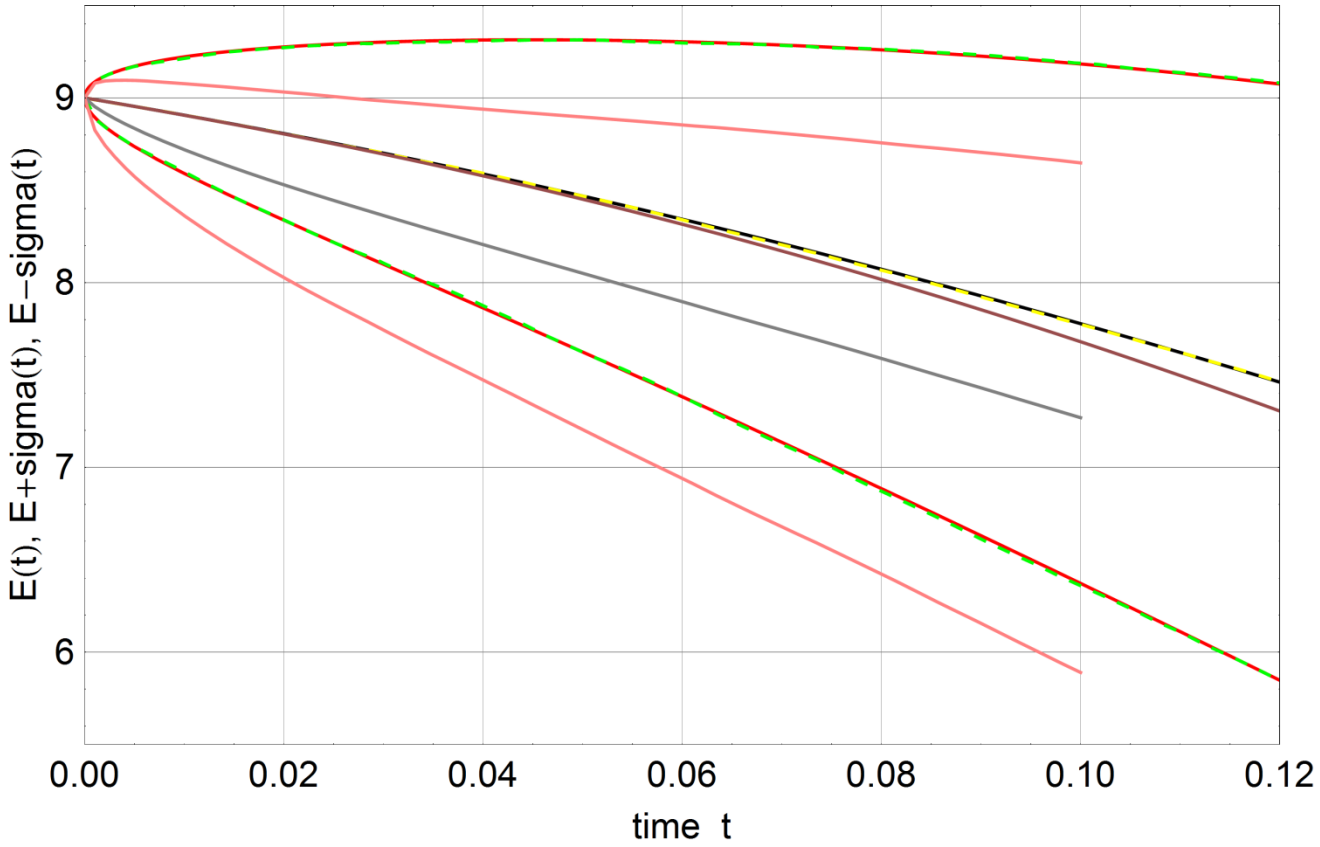


comparison of interpolation order zero and one

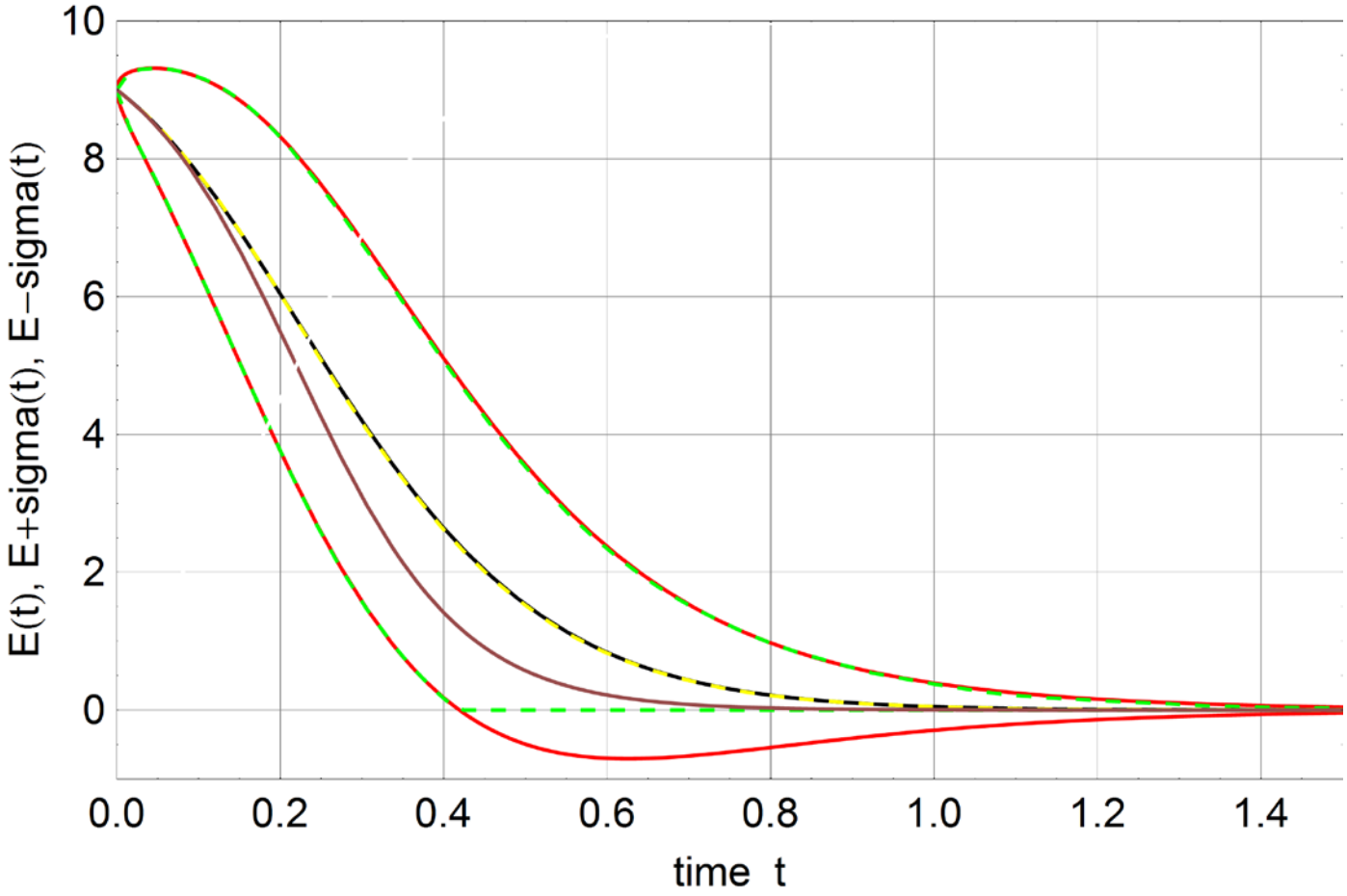
A $\leftrightarrow$ B: $k = 5$, $l = 5$, $N = 10\,000$



A $\leftrightarrow$ B: $k = 5$, $l = 5$, $N = 10\,000$



A + X $\rightarrow$ 2 X: $k = 1.0, N = 10$



A + X $\rightarrow$ 2 X: $k = 1.0, N = 10$

Thank you for your attention!

Web-Page for further information:

http://www.tbi.univie.ac.at/~pks