Supplementary Material for "Simulation and inference algorithms for stochastic biochemical reaction networks: from basic concepts to state-of-the-art"

David J. Warne[1], Ruth E. Baker[2], and Matthew J. Simpson[1]

[1]School of Mathematical Sciences, Queensland University of Technology, Brisbane, Queensland 4001, Australia
[2]Mathematical Institute, University of Oxford, Oxford, OX2 6GG, United Kingdom

January 29, 2019

Contents

Appendix A Derivation of mono-molecular chain mean and variances using the chemical master equation

In this section, we provide an example on how to derive moments of the chemical master equation (CME) solution without explicit CME evaluation. The presented analysis is specific to the two species mono-molecular chain model as presented in the main text. Our approach is based on the examples from Erban et al. [2], however, the result is more complex since we deal with a two chemical species, A and B.

For convenience, we restate the model. Here we consider a two species mono-molecular chain,

$$\emptyset \xrightarrow{k_1} A \xrightarrow{k_2} B \xrightarrow{k_3} \emptyset, \tag{A.1}$$

with known kinetic rate parameters k_1, k_2 and k_3. Given the state vector, $\mathbf{X}(t) = [A(t), B(t)]^T$, the respective propensity functions are

$$a_1(\mathbf{X}(t)) = k_1, \quad a_2(\mathbf{X}(t)) = k_2 A(t), \quad a_3(\mathbf{X}(t)) = k_3 B(t). \tag{A.2}$$

The stoichiometric vectors are

$$\nu_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \nu_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \nu_3 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}. \tag{A.3}$$

For $P(\mathbf{x}, t \mid \mathbf{x}_0) = \mathbb{P}(\mathbf{X}(t) = \mathbf{x} \mid \mathbf{X}(0) = \mathbf{x}_0)$, the general form of the CME is

$$\frac{\mathrm{d}P(\mathbf{x}, t \mid \mathbf{x}_0)}{\mathrm{d}t} = \sum_{j=1}^{M} a_j(\mathbf{x} - \nu_j) P(\mathbf{x} - \nu_j, t \mid \mathbf{x}_0) - P(\mathbf{x}, t \mid \mathbf{x}_0) \sum_{j=1}^{M} a_j(\mathbf{x}). \tag{A.4}$$

After substituting the propensity functions (Equation (A.2)) and stoichiometric vectors (Equation (A.3)) into Equation (A.4), we obtain the CME specific to the mono-molecular chain model (Equation (A.1))

$$\begin{aligned} \frac{\mathrm{d}P(a, b, t \mid a_0, b_0)}{\mathrm{d}t} &= k_1 P(a-1, b, t \mid a_0, b_0) + k_2(a+1) P(a+1, b-1, t \mid a_0, b_0) \\ &\quad + k_3(b+1) P(a, b+1, t \mid a_0, b_0) - (k_1 + k_2 a + k_3 b) P(a, b, t \mid a_0, b_0). \end{aligned} \tag{A.5}$$

Henceforth, we will denote $p_{a,b}(t)$ as the solution to the mono-molecular CME (Equation (A.5)).

Rather than solve the full CME, we seek a solution to the mean copy number of A at time t,

$$M_a(t) = \sum_{a=0}^{\infty} \sum_{b=0}^{\infty} a p_{a,b}(t), \tag{A.6}$$

the mean copy number of B at time t,

$$M_b(t) = \sum_{a=0}^{\infty} \sum_{b=0}^{\infty} b p_{a,b}(t), \tag{A.7}$$

the variance of A at time t,

$$V_a(t) = \sum_{a=0}^{\infty} \sum_{b=0}^{\infty} \left(a - M_a(t)\right)^2 p_{a,b}(t), \tag{A.8}$$

the variance of B at time t,

$$V_b(t) = \sum_{a=0}^{\infty} \sum_{b=0}^{\infty} \left(b - M_b(t)\right)^2 p_{a,b}(t), \tag{A.9}$$

and the covariance of A and B at time t,

$$C_{a,b}(t) = \sum_{a=0}^{\infty} \sum_{b=0}^{\infty} \left(a - M_a(t)\right)\left(b - M_b(t)\right) p_{a,b}(t). \tag{A.10}$$

We will derive a system of ODEs that describe the evolution of $M_a(t)$,$M_b(t)$, $V_a(t)$, $V_b(t)$, and $C_{a,b}(t)$ without explicitly solving the CME in Equation (A.5). Instead we exploit the linearity of the derivative along with the property,

$$\sum_{a=0}^{\infty}\sum_{b=0}^{\infty} p_{a,b}(t) = 1, \tag{A.11}$$

for all t.

To derive an ODE for $M_a(t)$, we multiply Equation (A.5) by a and sum over all a and b.

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left[\sum_{a=0}^{\infty}\sum_{b=0}^{\infty} a p_{a,b}(t)\right] &= \sum_{a=1}^{\infty}\sum_{b=0}^{\infty} k_1 a p_{a-1,b}(t) + \sum_{a=0}^{\infty}\sum_{b=1}^{\infty} k_2 a(a+1) p_{a+1,b-1}(t) \\ &\quad + \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} k_3 a(b+1) p_{a,b+1}(t) - \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} a(k_1 + k_2 a + k_3 b) p_{a,b}(t).\end{aligned}$$

After changing indices ($a-1 \to a$ in the first term, $a+1 \to a$ and $b-1 \to b$ in the second term, and $b+1 \to b$ in the third term), we obtain

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left[\sum_{a=0}^{\infty}\sum_{b=0}^{\infty} a p_{a,b}(t)\right] &= \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} k_1(a+1) p_{a,b}(t) + \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} k_2(a-1) a p_{a,b}(t) \\ &\quad + \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} k_3 ab p_{a,b}(t) - \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} a(k_1 + k_2 a + k_3 b) p_{a,b}(t).\end{aligned}$$

We simplify the right hand side,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\sum_{a=0}^{\infty}\sum_{b=0}^{\infty} a p_{a,b}(t)\right] = k_1 \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} p_{a,b}(t) - k_2 \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} a p_{a,b}(t),$$

and apply property (A.11) to give

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\sum_{a=0}^{\infty}\sum_{b=0}^{\infty} a p_{a,b}(t)\right] = k_1 - k_2 \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} a p_{a,b}(t).$$

Using the definition of $M_a(t)$ (Equation (A.6)), we obtain the ODE for the mean of A,

$$\frac{\mathrm{d}M_a(t)}{\mathrm{d}t} = k_1 - k_2 M_a(t). \tag{A.12}$$

Similarly, we derive an ODE for $M_b(t)$ by muliplying Equation (A.5) by b and proceed

in the same manner as we did for $M_a(t)$

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\left[\sum_{a=0}^{\infty}\sum_{b=0}^{\infty} bp_{a,b}(t)\right] &= \sum_{a=1}^{\infty}\sum_{b=0}^{\infty} k_1 bp_{a-1,b}(t) + \sum_{a=0}^{\infty}\sum_{b=1}^{\infty} k_2 b(a+1)p_{a+1,b-1}(t) \\
&\quad + \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} k_3 b(b+1)p_{a,b+1}(t) - \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} b(k_1+k_2a+k_3b)p_{a,b}(t) \\
&= \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} k_1 bp_{a,b}(t) + \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} k_2(b+1)ap_{a,b}(t) \\
&\quad + \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} k_3(b-1)bp_{a,b}(t) - \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} b(k_1+k_2a+k_3b)p_{a,b}(t) \\
&= k_2\sum_{a=0}^{\infty}\sum_{b=0}^{\infty} ap_{a,b}(t) - k_3\sum_{a=0}^{\infty}\sum_{b=0}^{\infty} bp_{a,b}(t).
\end{aligned}$$

Using the definitions of $M_a(t)$ (Equation (A.6)) and $M_b(t)$ (Equation (A.13)) we obtain the ODE for the mean of B,

$$\frac{\mathrm{d}M_b(t)}{\mathrm{d}t} = k_2M_a(t) - k_3M_b(t). \tag{A.13}$$

To derive the ODE for $V_a(t)$, first note that through expanding Equation (A.8) it can be shown that

$$V_a(t) + M_a(t)^2 = \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} a^2p_{a,b}(t). \tag{A.14}$$

Thus, we multiply Equation (A.5) by a^2, sum over all a and b, change indices, and simplify as follows,

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\left[\sum_{a=0}^{\infty}\sum_{b=0}^{\infty} a^2p_{a,b}(t)\right] &= \sum_{a=1}^{\infty}\sum_{b=0}^{\infty} k_1 a^2p_{a-1,b}(t) + \sum_{a=0}^{\infty}\sum_{b=1}^{\infty} k_2 a^2(a+1)p_{a+1,b-1}(t) \\
&\quad + \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} k_3 a^2(b+1)p_{a,b+1}(t) - \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} a^2(k_1+k_2a+k_3b)p_{a,b}(t) \\
&= \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} k_1(a+1)^2p_{a,b}(t) + \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} k_2(a-1)^2ap_{a,b}(t) \\
&\quad + \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} k_3 a^2bp_{a,b}(t) - \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} a^2(k_1+k_2a+k_3b)p_{a,b}(t) \\
&= k_1\sum_{a=0}^{\infty}\sum_{b=0}^{\infty} p_{a,b}(t) + (2k_1+k_2)\sum_{a=0}^{\infty}\sum_{b=0}^{\infty} ap_{a,b}(t) - 2k_2\sum_{a=0}^{\infty}\sum_{b=0}^{\infty} a^2p_{a,b}(t) \\
&= k_1 + (2k_1+k_2)\sum_{a=0}^{\infty}\sum_{b=0}^{\infty} ap_{a,b}(t) - 2k_2\sum_{a=0}^{\infty}\sum_{b=0}^{\infty} a^2p_{a,b}(t).
\end{aligned}$$

Using the definition of $M_a(t)$ (Equation (A.12)) and property (A.14), we have the ODE,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[V_a(t) + M_a(t)^2\right] = k_1 + (2k_1+k_2)M_a(t) - 2k_2\left(V_a(t) + M_a(t)^2\right).$$

Apply the chain rule to obtain,

$$\frac{\mathrm{d}V_a(t)}{\mathrm{d}t} = -2M_a(t)\frac{\mathrm{d}M_a(t)}{\mathrm{d}t} + k_1 + (2k_1 + k_2)M_a(t) - 2k_2\left(V_a(t) + M_a(t)^2\right), \tag{A.15}$$

then substitute Equation (A.12) into Equation (A.15) and simplify to arrive at the ODE for the variance of A,

$$\frac{\mathrm{d}V_a(t)}{\mathrm{d}t} = k_1 + k_2M_a(t) - 2k_2V_a(t). \tag{A.16}$$

Similarly, to derive the ODE for $V_b(t)$ we note that through expanding Equation (A.9) and Equation (A.10) it can be shown that

$$V_b(t) + M_b(t)^2 = \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} b^2 p_{a,b}(t), \tag{A.17}$$

and

$$C_{a,b}(t) + M_a(t)M_b(t) = \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} abp_{a,b}(t). \tag{A.18}$$

Thus, we multiply Equation (A.5) by b^2, sum over all a and b, change indices, and simplify as follows,

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\left[\sum_{a=0}^{\infty}\sum_{b=0}^{\infty} b^2 p_{a,b}(t)\right] &= \sum_{a=1}^{\infty}\sum_{b=0}^{\infty} k_1 b^2 p_{a-1,b}(t) + \sum_{a=0}^{\infty}\sum_{b=1}^{\infty} k_2 b^2 (a+1) p_{a+1,b-1}(t) \\
&\quad + \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} k_3 b^2 (b+1) p_{a,b+1}(t) - \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} b^2 (k_1 + k_2 a + k_3 b) p_{a,b}(t) \\
&= \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} k_1 b^2 p_{a,b}(t) + \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} k_2 (b+1)^2 a p_{a,b}(t) \\
&\quad + \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} k_3 (b-1)^2 b p_{a,b}(t) - \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} b^2 (k_1 + k_2 a + k_3 b) p_{a,b}(t) \\
&= k_2 \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} a p_{a,b}(t) + k_3 \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} b p_{a,b}(t) \\
&\quad + 2k_2 \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} ab p_{a,b}(t) - 2k_3 \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} b^2 p_{a,b}(t).
\end{aligned}$$

Using the definitions of $M_a(t)$ (Equation (A.6)) and $M_b(t)$ (Equation (A.7)), and properties (A.17) and (A.18) we have the ODE

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[V_b(t) + M_b(t)^2\right] = k_2M_a(t) + k_3M_b(t) + 2k_2\left(C_{a,b}(t) + M_a(t)M_b(t)\right) - 2k_3\left(V_b(t) + M_b(t)^2\right).$$

We apply the chain rule

$$\begin{aligned}
\frac{\mathrm{d}V_b(t)}{\mathrm{d}t} &= -2M_b(t)\frac{\mathrm{d}M_b(t)}{\mathrm{d}t} + k_2M_a(t) + k_3M_b(t) \\
&\quad + 2k_2\left(C_{a,b}(t) + M_a(t)M_b(t)\right) - 2k_3\left(V_b(t) + M_b(t)^2\right),
\end{aligned} \tag{A.19}$$

then substitute Equation (A.13) into Equation (A.19) and simplify to obtain the ODE for the variance of B

$$\frac{\mathrm{d}V_b(t)}{\mathrm{d}t} = k_2 M_a(t) + k_3 M_b(t) + 2k_2 C_{a,b} - 2k_3 V_b(t). \tag{A.20}$$

Finally, we derive the ODE for $C_{a,b}(t)$ by multiplying Equation (A.5) by ab, summing over all a and b, changing indices, and simplifying as follows:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\left[\sum_{a=0}^{\infty}\sum_{b=0}^{\infty} abp_{a,b}(t)\right] &= \sum_{a=1}^{\infty}\sum_{b=0}^{\infty} k_1 abp_{a-1,b}(t) + \sum_{a=0}^{\infty}\sum_{b=1}^{\infty} k_2 ab(a+1)p_{a+1,b-1}(t) \\
&\quad + \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} k_3 ab(b+1)p_{a,b+1}(t) - \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} ab(k_1 + k_2 a + k_3 b)p_{a,b}(t) \\
&= \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} k_1(a+1)bp_{a,b}(t) + \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} k_2(a-1)(b+1)ap_{a,b}(t) \\
&\quad + \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} k_3 a(b-1)bp_{a,b}(t) - \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} ab(k_1 + k_2 a + k_3 b)p_{a,b}(t) \\
&= k_1 \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} bp_{a,b}(t) - k_2 \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} bp_{a,b}(t) \\
&\quad - (k_2 + k_3)\sum_{a=0}^{\infty}\sum_{b=0}^{\infty} abp_{a,b}(t) + k_2 \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} a^2 p_{a,b}(t).
\end{aligned}$$

Using the definition of $M_a(t)$ (Equation (A.6)) and $M_b(t)$ (Equation (A.7)), and properties (A.14) and (A.18) we obtain the ODE

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\left[C_{a,b}(t) + M_a(t)M_b(t)\right] &= k_1 M_a(t) - k_2 M_b(t) - (k_2 + k_3)\left(C_{a,b}(t) + M_a(t)M_b(t)\right) \\
&\quad + k_2\left(V_a(t) + M_a(t)^2\right).
\end{aligned}$$

Apply the chain rule and product rule

$$\begin{aligned}
\frac{\mathrm{d}C_{a,b}(t)}{\mathrm{d}t} &= -M_a(t)\frac{\mathrm{d}M_b(t)}{\mathrm{d}t} - M_b(t)\frac{\mathrm{d}M_a(t)}{\mathrm{d}t} + k_1 M_a(t) - k_2 M_b(t) \\
&\quad - (k_2 + k_3)\left(C_{a,b}(t) + M_a(t)M_b(t)\right) + k_2\left(V_a(t) + M_a(t)^2\right),
\end{aligned} \tag{A.21}$$

then substitute Equation (A.12) and Equation (A.13) into Equation (A.21) and simplify to obtain the ODE for the covariance of A and B

$$\frac{\mathrm{d}C_{a,b}(t)}{\mathrm{d}t} = k_2 V_a(t) - k_2 M_a(t) - (k_2 + k_3)C_{a,b}(t). \tag{A.22}$$

Therefore, Equations (A.12),(A.13),(A.16),(A.20), and (A.22) form a non-homogeneous

linear system of ODEs,

$$\begin{aligned}
\frac{\mathrm{d}M_a(t)}{\mathrm{d}t} &= k_1 - k_2 M_a(t), \\
\frac{\mathrm{d}M_b(t)}{\mathrm{d}t} &= k_2 M_a(t) - k_3 M_b(t), \\
\frac{\mathrm{d}V_a(t)}{\mathrm{d}t} &= k_1 + k_2 M_a(t) - 2k_2 V_a(t), \\
\frac{\mathrm{d}V_b(t)}{\mathrm{d}t} &= k_2 M_a(t) + k_3 M_b(t) + 2k_2 C_{a,b}(t) - k_3 V_b(t), \\
\frac{\mathrm{d}C_{a,b}(t)}{\mathrm{d}t} &= k_2 V_a(t) - k_2 M_a(t) - (k_2 + k_3) C_{a,b}(t).
\end{aligned}$$

After solving for the homogeneous solution, a particular solution may be obtained through using the method of undetermined coefficients. Given the initial conditions $A(0) = a_0$ and $B(0) = b_0$ with probability one, the solution, in the case when $k_2 \neq k_3$, is

$$M_a(t) = \frac{k_1}{k_2} + \left(a_0 - \frac{k_1}{k_2}\right) e^{-k_2 t}, \tag{A.23}$$

$$M_b(t) = \frac{k_1}{k_3} + \frac{k_2 a_0 - k_1}{k_3 - k_2} e^{-k_2 t} + \left(b_0 - \frac{k_2 a_0 - k_1}{k_3 - k_2} - \frac{k_1}{k_3}\right) e^{-k_3 t}, \tag{A.24}$$

$$V_a(t) = \frac{k_1}{k_2} + \left(a_0 - \frac{k_1}{k_2}\right) e^{-k_2 t} - a_0 e^{-2k_2 t}, \tag{A.25}$$

$$\begin{aligned}
V_b(t) = {} & \frac{k_1}{k_3} + \frac{k_2 a_0 - k_1}{k_3 - k_2} e^{-k_2 t} + \left(b_0 - \frac{k_2 a_0 - k_1}{k_3 - k_2} - \frac{k_1}{k_3}\right) e^{-k_3 t} + \frac{2a_0 k_2^2}{k_3^2 - k_2^2} e^{-(k_3 + k_2)t} \\
& - \frac{a_0 k_2}{k_3 - k_2} e^{-2k_2 t} + \left[\frac{a_0 k_2}{k_3 - k_2}\left(1 - \frac{2k_2}{k_3 + k_2}\right) - b_0\right] e^{-2k_3 t},
\end{aligned} \tag{A.26}$$

$$C_{a,b}(t) = -\frac{a_0 k_2}{k_3 - k_2} e^{-(k_3 + k_2)t} + \frac{a_0 k_2}{k_3 - k_2} e^{-2k_2 t}. \tag{A.27}$$

Equations (A.23)–(A.27) are the time dependent solutions for the moments and from these solutions we can evaluate the long time limit, $t \to \infty$, to give simple expression for the associated stationary solutions,

$$\begin{aligned}
\lim_{t\to\infty} M_a(t) &= \frac{k_1}{k_2}, \\
\lim_{t\to\infty} V_a(t) &= \frac{k_1}{k_2}, \\
\lim_{t\to\infty} M_b(t) &= \frac{k_1}{k_3}, \\
\lim_{t\to\infty} V_b(t) &= \frac{k_1}{k_3}, \\
\lim_{t\to\infty} C_{a,b}(t) &= 0.
\end{aligned}$$

See the example codes `DemoCMEMeanVar.m` and `DemoStationaryDist.m` for the evaluation of this solution.

Appendix B Evaluation of the mono-molecular chain chemical master equation solution

Jahnke and Huisinga [5] derive an analytic solution to the CME for a general mono-molecular BCRN. Applying their general solution to the species mono-molecular chain (Equation (A.1)) results in the following general solution to the CME (Equation (A.5)),

$$P(a,b,t \mid a_0,b_0) = \mathcal{P}\left(a,b,\lambda_a(t),\lambda_b(t)\right) * \mathcal{M}\left(a,b,a_0,\alpha_a(t),\alpha_b(t)\right) * \mathcal{M}\left(a,b,b_0,\beta_a(t),\beta_b(t)\right), \tag{B.1}$$

where $*$ is the discrete convolution operation [5]. $\mathcal{P}\left(a,b,\lambda_a(t),\lambda_b(t)\right)$ is a product Poisson distribution, given by

$$\mathcal{P}\left(a,b,\lambda_a(t),\lambda_b(t)\right) = \begin{cases} \dfrac{\lambda_a(t)^a}{a!}\dfrac{\lambda_b(t)^b}{b!}e^{-(|\lambda_a(t)|+|\lambda_b(t)|)}, & \text{if } a \geq 0, b \geq 0, \\ 0 & \text{otherwise,} \end{cases} \tag{B.2}$$

where the functions $\lambda_a(t)$ and $\lambda_b(t)$ are obtained through the initial value problem (IVP)

$$\frac{\mathrm{d}\lambda_a(t)}{\mathrm{d}t} = k_1 - k_2\lambda_a(t), \quad \frac{\mathrm{d}\lambda_b(t)}{\mathrm{d}t} = k_2\lambda_a(t) - k_3\lambda_b(t), \quad t > 0, \tag{B.3}$$

with initial conditions $\lambda_a(0) = \lambda_b(0) = 0$.$\mathcal{M}\left(a,b,a_0,\alpha_a(t),\alpha_b(t)\right)$ and $\mathcal{M}\left(a,b,b_0,\beta_a(t),\beta_b(t)\right)$ are multinomial distributions, given by

$$\mathcal{M}\left(a,b,a_0,\alpha_a(t),\alpha_b(t)\right) = \begin{cases} a_0!\dfrac{\left(1-|\alpha_a(t)|-|\alpha_b(t)|\right)^{a_0-|a|-|b|}}{(a_0-|a|-|b|)!}\dfrac{\alpha_a(t)^a}{a!}\dfrac{\alpha_b(t)^b}{b!} & \text{if } |a|+|b| \leq a_0, \\ 0 & \text{otherwise,} \end{cases} \tag{B.4}$$

and

$$\mathcal{M}\left(a,b,b_0,\beta_a(t),\beta_b(t)\right) = \begin{cases} b_0!\dfrac{\left(1-|\beta_a(t)|-|\beta_b(t)|\right)^{b_0-|a|-|b|}}{(b_0-|a|-|b|)!}\dfrac{\beta_a(t)^a}{a!}\dfrac{\beta_b(t)^b}{b!} & \text{if } |a|+|b| \leq b_0, \\ 0 & \text{otherwise.} \end{cases} \tag{B.5}$$

The functions $\alpha_a(t)$, $\alpha_b(t)$, $\beta_a(t)$ and $\beta_b(t)$ are obtained through the IVPs

$$\frac{\mathrm{d}\alpha_a(t)}{\mathrm{d}t} = -k_2\alpha_a(t), \quad \frac{\mathrm{d}\alpha_b(t)}{\mathrm{d}t} = k_2\alpha_a(t) - k_3\alpha_b(t), \quad t > 0, \tag{B.6}$$

and

$$\frac{\mathrm{d}\beta_a(t)}{\mathrm{d}t} = -k_2\beta_a(t), \quad \frac{\mathrm{d}\beta_b(t)}{\mathrm{d}t} = k_2\beta_a(t) - k_3\beta_b(t), \quad t > 0, \tag{B.7}$$

with initial conditions $\alpha_a(0) = 1$, $\alpha_b(0) = 0$, $\beta_a(0) = 0$, and $\beta_b(0) = 1$.

Equation (B.1) represents a direct substitution of the two species mono-molecular chain into the general solution by Jahnke and Huisinga [5]. However, direct point-wise evaluation of this solution is not feasible. Specifically, there are two challenges: (i) the two convolutions are taken over an infinite two-dimensional integer lattice; and (ii) the non-zero probabilities in the product Poisson and Multinomial distribution can be so small that numerical underflow/overflow is almost certain. The first issue can be solved

by determining the finite set of lattice sites that do not contribute to the convolutions, this can be achieved by invoking specific features of Equation (A.1). The second issue requires that we perform calculations using logarithms of probabilities rather than the true probabilities. Extra care must be taken in the convolution summations.

We first simplify the convolution operations to ensure finite computations. Solving the IVPs B.3,B.6, and B.7 yields, for $k_2 \neq k_3$,

$$\lambda_a(t) = \frac{k_1}{k_2}\left(1 - e^{-k_2 t}\right), \tag{B.8}$$

$$\lambda_b(t) = \frac{k_1}{k_3} + \frac{k_1}{k_3 - k_2}\left[e^{-k_2 t} + (k_2 - 2k_3)e^{-k_3 t}\right], \tag{B.9}$$

$$\alpha_a(t) = e^{-k_2 t}, \tag{B.10}$$

$$\alpha_b(t) = \frac{k_2}{k_3 - k_2}\left(e^{-k_2 t} - e^{-k_3 t}\right), \tag{B.11}$$

$$\beta_a(t) = 0, \tag{B.12}$$

$$\beta_b(t) = e^{-k_3 t}. \tag{B.13}$$

A key result is that $\beta_a(t)$ is zero for all time (Equation (B.12)). Through substitution of Equation (B.12) into Equation (B.5), we have

$$\mathcal{M}\left(a, b, b_0, 0, \beta_b(t)\right) = \begin{cases} b_0! \dfrac{\left(1 - |\beta_b(t)|\right)^{b_0 - |a| - |b|}}{(b_0 - |a| - |b|)!} \dfrac{0^a}{a!} \dfrac{\beta_b(t)^b}{b!} & \text{if } |a| + |b| \leq b_0, \\ 0 & \text{otherwise.} \end{cases} \tag{B.14}$$

This implies that $\mathcal{M}\left(a, b, b_0, \beta_a(t), \beta_b(t)\right) = 0$ if $a \neq 0$. That is,

$$\mathcal{M}\left(a, b, b_0, 0, \beta_b(t)\right) = \begin{cases} b_0! \dfrac{\left(1 - |\beta_b(t)|\right)^{b_0 - |b|}}{(b_0 - |b|)!} \dfrac{\beta_b(t)^b}{b!} & \text{if } a = 0, \text{ and } |b| \leq b_0, \\ 0 & \text{otherwise.} \end{cases} \tag{B.15}$$

We can now make a significant simplification of the second convolution in Equation (B.1). Let $\mathcal{M}_a(a, b, t) = \mathcal{M}\left(a, b, a_0, \alpha_a(t), \alpha_b(t)\right)$ and $\mathcal{M}_b(a, b, t) = \mathcal{M}\left(a, b, b_0, 0, \beta_b(t)\right)$, we have

$$\begin{aligned} \mathcal{M}_a(a, b, t) * \mathcal{M}_b(a, b, t) &= \sum_{a_w \in \mathbb{N}} \sum_{b_w \in \mathbb{N}} \mathcal{M}_a(a_w, b_w, t) \mathcal{M}_b(a - a_w, b - b_w, t) \\ &= \sum_{b_w \in \mathbb{N}} \mathcal{M}_a(a, b_w, t) \mathcal{M}_b(0, b - b_w, t), \end{aligned}$$

where $\mathbb{N} = \mathbb{Z}^+ \cup \{0\}$. By Equation (B.15), $\mathcal{M}_b(0, b, t) = 0$ if $|b| > b_0$. Furthermore, we have $b \geq 0$ from the definition of the BCRN (Equation (A.1)). It follows that only terms with $b \geq b_w \geq \max(0, b - b_0)$ can contribute to the convolution, that is,

$$\mathcal{M}_a(a, b, t) * \mathcal{M}_b(a, b, t) = \sum_{b_w = \max(0, b - b_0)}^{b} \mathcal{M}_a(a, b_w, t) \mathcal{M}_b(0, b - b_w, t).$$

While this convolution never involves more that b terms, we can apply a further constraint on the upper bound of the index. By Equation (B.4) we have $\mathcal{M}_a(a, b, t) = 0$ if $|a| + |b| >$

a_0. Since $a \geq 0$ and $b \geq 0$ from the definition of the BCRN (Equation (A.1). That is, terms with $b_w \geq a_0 - a \geq 0$ will not contribute to the convolution. Therefore, the multinomial convolution term in Equation (B.1) is

$$\mathcal{M}_a(a,b,t) * \mathcal{M}_b(a,b,t) = \sum_{b_w=\max(0,b-b_0)}^{\min(b,\max(0,a_0-a))} \mathcal{M}_a(a,b_w,t)\mathcal{M}_b(0,b-b_w,t). \tag{B.16}$$

Let $\mathcal{P}(a,b,t) = \mathcal{P}\left(a,b,\lambda_a(t),\lambda_b(t)\right)$ and substitute Equation (B.16) into Equation (B.1) to yield

$$\begin{aligned} P(a,b,t \mid a_0,b_0) &= \mathcal{P}\left(a,b,t\right) * \left[\sum_{b_w=\max(0,b-b_0)}^{\min(b,\max(0,a_0-a))} \mathcal{M}_a(a,b_w,t)\mathcal{M}_b(0,b-b_w,t)\right] \\ &= \sum_{a_z\in\mathbb{N}}\sum_{b_z\in\mathbb{N}} \mathcal{P}\left(a-a_z,b-b_z,t\right)\left[\sum_{b_w=\max(0,b_z-b_0)}^{\min(b_z,\max(0,a_0-a_z))} \mathcal{M}_a(a_z,b_w,t)\mathcal{M}_b(0,b_z-b_w,t)\right]. \end{aligned}$$

By definition of the product Poisson distribution (Equation (B.2)), $\mathcal{P}\left(a,b,t\right) = 0$ for $a < 0$ or $b < 0$. Hence, only terms with $a \geq a_z$ and $b \geq b_z$ contribute to the convolution. Therefore, we obtain the following expression for Equation (B.1)

$$\begin{aligned} P(a,b,t \mid a_0,b_0) = \sum_{a_z=0}^{a}\sum_{b_z=0}^{b} &\mathcal{P}\left(a-a_z,b-b_z,t\right) \\ &\times \left[\sum_{b_w=\max(0,b_z-b_0)}^{\min(b_z,\max(0,a_0-a_z))} \mathcal{M}_a(a_z,b_w,t)\mathcal{M}_b(0,b_z-b_w,t)\right], \end{aligned} \tag{B.17}$$

which requires $\mathcal{O}(ab^2)$ evaluations of either Equation (B.2), Equation (B.4) or Equation (B.15).

Now that we have bounded the number of operations required to evaluate the solution of the CME, we now address the problem of numerical overflow/underflow. There are two possible sources for this type of numerical error. Firstly, the factorials and products of powers involved in the evaluation of Equation (B.2), Equation (B.4), and Equation (B.15) can be very large, causing overflow. Secondly, the probabilities in the convolution terms can be very small, causing underflow.

To avoid these issues we work with logarithms of probabilities. For the non-zero cases of Equations (B.2), (B.4) and (B.15), we have

$$\ln \mathcal{P}\left(a,b,t\right) = a\ln\lambda_a(t) + b\ln\lambda_b(t) - \left(|\lambda_a(t)| + |\lambda_b(t)|\right) - \sum_{a_i=1}^{a}\ln a_i - \sum_{b_i=1}^{b}\ln b_i, \tag{B.18}$$

$$\begin{aligned} \ln \mathcal{M}_a\left(a,b,t\right) = {} & a\ln\alpha_a(t) + b\ln\alpha_b(t) + (a_0-a-b)\ln\left(1-\alpha_a(t)-\alpha_b(t)\right) \\ & + \sum_{a_i=a_0-a-b}^{a_0}\ln a_i - \sum_{a_i=1}^{a}\ln a_i - \sum_{b_i=1}^{b}\ln b_i, \end{aligned} \tag{B.19}$$

$$\ln \mathcal{M}_b\left(a,b,t\right) = b\ln\beta_b(t) + (b_0-b)\ln\left(1-\beta_b(t)\right) + \sum_{b_i=b_0-b}^{b_0}\ln b_i - \sum_{b_i=1}^{b}\ln b_i. \tag{B.20}$$

Equations (B.18)–(B.20) enable the computation to proceed with overflow or underflow being significantly less likely. Therefore, we take the logarithm of Equation (B.17) to obtain

$$\ln P(a,b,t \mid a_0,b_0) = \ln\left[\sum_{a_z=0}^{a}\sum_{b_z=0}^{b} e^{\ln \mathcal{P}(a-a_z,b-b_z,t)+\ln \mathcal{F}(a_z,b_z,t)}\right], \tag{B.21}$$

where

$$\ln \mathcal{F}(a_z,b_z,t) = \ln\left[\sum_{b_w=\max(0,b_z-b_0)}^{\min(b_z,\max(0,a_0-a_z))} e^{\ln \mathcal{M}_a(a_z,b_w,t)+\ln \mathcal{M}_b(0,b_z-b_w,t)}\right]. \tag{B.22}$$

Computing the logarithms of summations of exponential functions in Equation (B.21) and Equation (B.22) is still prone to overflow and underflow since the probabilities will be very small in practice. A common solution to numerically stable logarithm of summations of exponential functions is known as the "log-sum-exp trick". This works by noting, for any $x, y \in \mathbb{R}$, that

$$\begin{aligned}\ln\left[e^x + e^y\right] &= \ln\left[\left(e^{x-\max(x,y)} + e^{y-\max(x,y)}\right) e^{\max(x,y)}\right] \\ &= \ln\left[e^{x-\max(x,y)} + e^{y-\max(x,y)}\right] + \max(x,y).\end{aligned}$$

Thus, computations are re-scaled to the natural scale of the terms in the summation, thus terms that do underflow would not have affected the result significantly. Now, let

$$R(a_z,b_z) = \max_{b_w\in[\max(0,b_z-b_0),\min(b_z,\max(0,a_0-a_z))]} \left\{\ln \mathcal{M}_a(a_z,b_w,t) + \ln \mathcal{M}_b(0,b_z-b_w,t)\right\},$$

and

$$S(a,b) = \max_{[a_z,b_z]\in[0,a]\times[0,b]} \left\{\ln \mathcal{P}\left(a-a_z,b-b_z,t\right) + \ln \mathcal{F}\left(a_z,b_z,t\right)\right\}.$$

Then use $S(a,b)$ and $R(a_z,b_z)$ with the "log-sum-exp trick" to yield a numerically robust form of Equation (B.21). That is,

$$\ln P(a,b,t \mid a_0,b_0) = \ln\left[\sum_{a_z=0}^{a}\sum_{b_z=0}^{b} e^{\ln \mathcal{P}(a-a_z,b-b_z,t)+\ln \mathcal{F}(a_z,b_z,t)-S(a,b)}\right] + S(a,b), \tag{B.23}$$

where

$$\ln \mathcal{F}(a_z,b_z,t) = \ln\left[\sum_{b_w=\max(0,b_z-b_0)}^{\min(b_z,\max(0,a_0-a_z))} e^{\ln \mathcal{M}_a(a_z,b_w,t)+\ln \mathcal{M}_b(0,b_z-b_w,t)-R(a_z,b_z)}\right] + R(a_z,b_z). \tag{B.24}$$

The example code, `CMEsolMonoMol.m` provides a numerical implementation of the CME solution (Equation (B.1)) using Equation (B.23) and Equation (B.24).

Appendix C Synthetic data

The synthetic data used in the main manuscript and example code is provide in Table C.1 for the mono-molecular chain model and in Table C.2 for the enzyme kinetics model.

Table C.1: Data, $\mathbf{Y}_{\text{obs}}$, used for inference on the mono-molecular chain. Generated using true parameter values $k_1 = 1.0$, $k_2 = 0.1$, and $k_3 = 0.05$ and initial conditions, $A(0) = 100$ and $B(0) = 0$.

	$\mathbf{Y}(t_1)$	$\mathbf{Y}(t_2)$	$\mathbf{Y}(t_3)$	$\mathbf{Y}(t_4)$
t	25	50	75	100
$A(t)$	14	12	17	15
$B(t)$	68	34	14	14

Table C.2: Data, $\mathbf{Y}_{\text{obs}}$, used for inference on the enzyme kinetic model. Generated using true parameter values $k_1 = 0.001$, $k_2 = 0.005$, and $k_3 = 0.01$ and initial conditions $S(0) = E(0) = 100$ and $C(0) = P(0) = 0$.

	$\mathbf{Y}(t_1)$	$\mathbf{Y}(t_2)$	$\mathbf{Y}(t_3)$	$\mathbf{Y}(t_4)$	$\mathbf{Y}(t_5)$
t	0	20	40	60	80
$P(t)$	0	5	16	28	39
$P(t) + \xi$	2.04	6.99	14.30	28.71	38.14

Appendix D Additional ABC results

In the main manuscript only marginal probability densities are used to demonstrate ABC convergence. Here we present plot matrices with the bivariate marginals also.

Since the mono-molecular chain model has three rate parameters, we have three univariate marginals posteriors and three bivariate marginal posteriors. Through application of the ABC with acceptance threshold, ϵ, the equivalent marginals are

$$p(k_1 \mid \rho(\mathbf{Y}_{\text{obs}}, \mathbf{S}_{\text{obs}}) \leq \epsilon) = \iint_{\mathbb{R}^2} p(k_1, k_2, k_3 \mid \rho(\mathbf{Y}_{\text{obs}}, \mathbf{S}_{\text{obs}}) \leq \epsilon)\, \mathrm{d}k_2\, \mathrm{d}k_3, \tag{D.1}$$

$$p(k_2 \mid \rho(\mathbf{Y}_{\text{obs}}, \mathbf{S}_{\text{obs}}) \leq \epsilon) = \iint_{\mathbb{R}^2} p(k_1, k_2, k_3 \mid \rho(\mathbf{Y}_{\text{obs}}, \mathbf{S}_{\text{obs}}) \leq \epsilon)\, \mathrm{d}k_1\, \mathrm{d}k_3, \tag{D.2}$$

$$p(k_3 \mid \rho(\mathbf{Y}_{\text{obs}}, \mathbf{S}_{\text{obs}}) \leq \epsilon) = \iint_{\mathbb{R}^2} p(k_1, k_2, k_3 \mid \rho(\mathbf{Y}_{\text{obs}}, \mathbf{S}_{\text{obs}}) \leq \epsilon)\, \mathrm{d}k_1\, \mathrm{d}k_2, \tag{D.3}$$

$$p(k_1, k_2 \mid \rho(\mathbf{Y}_{\text{obs}}, \mathbf{S}_{\text{obs}}) \leq \epsilon) = \int_{\mathbb{R}} p(k_1, k_2, k_3 \mid \rho(\mathbf{Y}_{\text{obs}}, \mathbf{S}_{\text{obs}}) \leq \epsilon)\, \mathrm{d}k_3, \tag{D.4}$$

$$p(k_1, k_3 \mid \rho(\mathbf{Y}_{\text{obs}}, \mathbf{S}_{\text{obs}}) \leq \epsilon) = \int_{\mathbb{R}} p(k_1, k_2, k_3 \mid \rho(\mathbf{Y}_{\text{obs}}, \mathbf{S}_{\text{obs}}) \leq \epsilon)\, \mathrm{d}k_2, \tag{D.5}$$

$$p(k_2, k_3 \mid \rho(\mathbf{Y}_{\text{obs}}, \mathbf{S}_{\text{obs}}) \leq \epsilon) = \int_{\mathbb{R}} p(k_1, k_2, k_3 \mid \rho(\mathbf{Y}_{\text{obs}}, \mathbf{S}_{\text{obs}}) \leq \epsilon)\, \mathrm{d}k_1. \tag{D.6}$$

The exact univariate and bivariate marginal posteriors are plotted against the ABC posterior for $\epsilon = [50, 25, 12.5, 0]$ (with $\epsilon = 0$ meaning the exact posterior is sampled using the CME-based likelihood). Equations (D.1)–(D.6) are plotted in Figure D.1.

Reducing ϵ further than 12.5 is prohibitive, even for the mono-molecular chain model. Both Barber et al. [1] and Fearnhead and Prangle [3] provide an asymptotic result for the computation time, $\mathcal{C}$, as a function of ϵ, that is, $\mathcal{C} = \mathcal{O}(\epsilon^{-d})$, where d is the dimensionality of the data used in the ABC inference. For the synthetic data we have from Table C.1, we have $d = n_t N$. Figure D.2 demonstrates that the computation times we obtain are consistent with this.

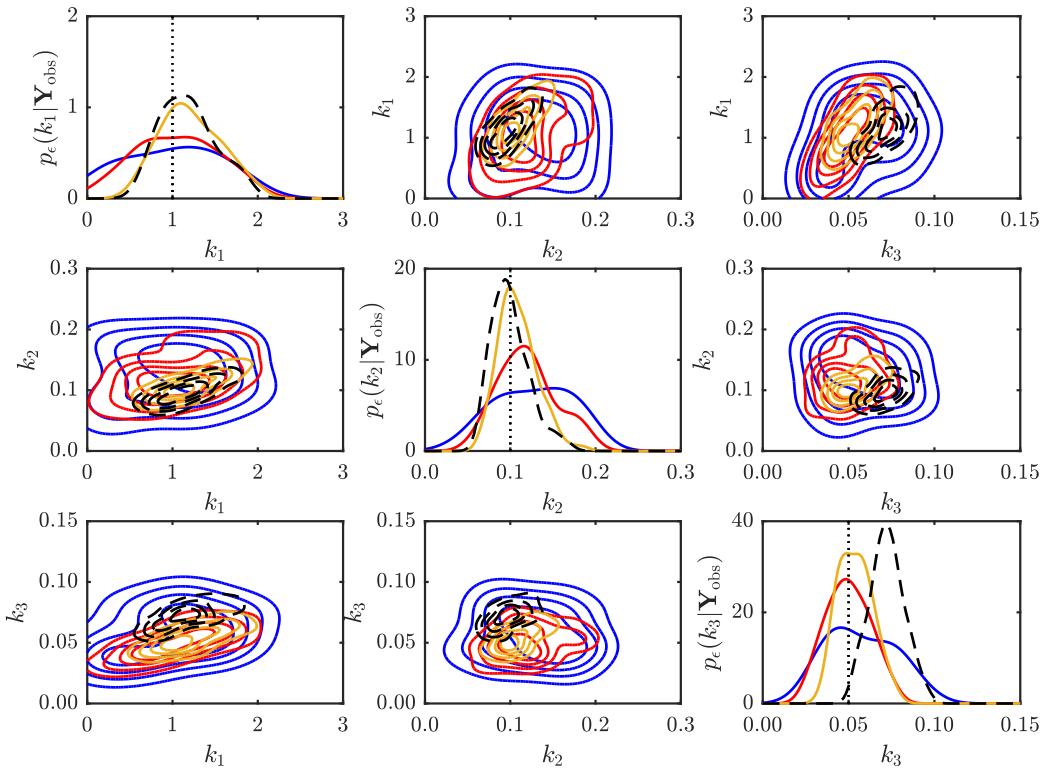

Figure D.1: Convergence of ABC posterior to the true posterior as $\epsilon \to 0$ for the monomolecular chain inference problem. Marginal posteriors are plotted for $\epsilon = 50$ (blue solid), $\epsilon = 25$ (red solid), $\epsilon = 12.5$ (yellow solid), and $\epsilon = 0$ (black dashed). Here, the $\epsilon = 0$ case corresponds to the exact likelihoods using the CME solution. Univariate marginals are plotted on the diagonals and bivariate marginals on off diagonal elements. Contour lines in bivariate marginal plots are selected such that six equal probability density intervals are shown. The true parameter values (black dotted) are $k_1 = 1.0$, $k_2 = 0.1$ and $k_3 = 0.05$. Note that the exact Bayesian posterior does not recover the true parameter for k_3.

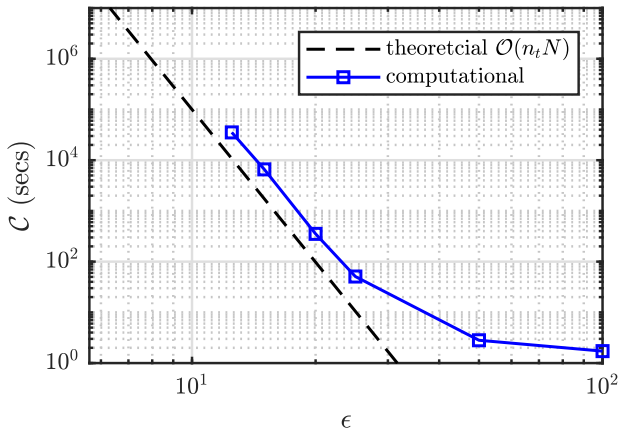

Figure D.2: Computation time growth of ABC rejection sampling against theoretical result. Computations are performed using an Intel® Core™ i7-5600U CPU (2.6 GHz).

Appendix E ABC Multilevel Monte Carlo

Here we provide the ABC Multilevel Monte Carlo scheme (ABCMLMC) (see Warne et al. [6] for more details and the derivation). This particular implementation computes an estimate of the posterior mean. Given a sequence of acceptance thresholds, $\epsilon_0 > \epsilon_1 > \cdots > \epsilon_L = \epsilon$, and a sequence of sample numbers $m_0 > m_1 > \cdots > m_L$ (see Giles [4] and Warne et al. [6] for details on optimally computing the sample numbers), ABCMLMC proceeds as follows:

1. initialise $\ell = 0$;
2. set $i = 1$;
3. generate a prior sample $\boldsymbol{\theta}^* \sim p(\boldsymbol{\theta})$;
4. generate simulated data, $\mathbf{S}^*_{\text{obs}} \sim s(\mathbf{S}_{\text{obs}}; \boldsymbol{\theta}^*)$;
5. if $\rho(\mathbf{Y}_{\text{obs}}, \mathbf{S}^*_{\text{obs}}) \leq \epsilon_\ell$, accept $\boldsymbol{\theta}^{(i)}_{\epsilon_\ell} = \boldsymbol{\theta}^*$ and set $i = i + 1$, otherwise continue;
6. if $i \leq m_\ell$, go to step 3, otherwise continue;
7. set $\hat{F}_{\ell,j}(z) = \sum_{i=1}^{m_\ell} \mathbb{1}_z \left(k^{(i)}_{\epsilon_\ell,j}\right)/m_\ell$ for $j = 1, 2, \ldots, M$;
8. if $\ell = 0$, then set $\hat{\boldsymbol{\theta}}_\epsilon = \sum_{i=1}^{m_\ell} \boldsymbol{\theta}^{(i)}_{\epsilon_\ell}/m_\ell$, set $\ell = \ell + 1$, and go to step 2, otherwise continue;
9. set $i = 1$;
10. set $k^{(i)}_{\epsilon_{\ell-1},j} = \hat{F}^{-1}_{\ell-1,j}(\hat{F}_{\ell,j}(k^{(i)}_{\epsilon_\ell,j}))$ for $j = 1, 2, \ldots, M$;
11. set $\boldsymbol{\theta}^{(i)}_{\epsilon_{\ell-1}} = [k^{(i)}_{\epsilon_{\ell-1},1}, k^{(i)}_{\epsilon_{\ell-1},2}, \ldots, k^{(i)}_{\epsilon_{\ell-1},M}]$ and set $i = i + 1$;
12. if $i \leq m_\ell$, then go to step 9, otherwise continue;
13. set $\hat{F}_{\ell,j}(z) = \hat{F}_{\ell-1,j}(z) + \sum_{i=1}^{m_\ell} \left(\mathbb{1}_z \left(k^{(i)}_{\epsilon_\ell,j}\right) - \mathbb{1}_z \left(k^{(i)}_{\epsilon_{\ell-1},j}\right)\right)/m_\ell$;
14. set $\hat{\boldsymbol{\theta}}_\epsilon = \hat{\boldsymbol{\theta}}_\epsilon + \sum_{i=1}^{m_\ell} \left(\boldsymbol{\theta}^{(i)}_{\epsilon_\ell} - \boldsymbol{\theta}^{(i)}_{\epsilon_{\ell-j}}\right)/m_\ell$;
15. if $\ell = L$, then terminate, otherwise set $\ell = \ell + 1$ and go to step 2;

Appendix F ABC algorithm configurations and additional results

The following algorithm configurations for the ABC rejection sampler, ABCMCMC, ABCSMC and ABCMLMC are used to generate the results in Tables 3 and 4 of the main manuscript. The parameters are also contained in the code examples, `DemoABCMethodsMonoMol.m` and `DemoABCMethodsMichMent.m` .

For the mono-molecular chain model inference problem each algorithm is configured as follows: for the ABC rejection sampler we set $m = 100$ and $\epsilon = 15$; for ABCMCMC we set $m_n = 500,000$, $m_b = 100,000$, $m_h = 10,000$, the proposal kernel is a Gaussian random walk with covariance matrix, $\Sigma = \text{diag}(1 \times 10^{-3}, 1 \times 10^{-5}, 2.5 \times 10^{-5})$, and $\epsilon = 15$; for ABCSMC we use $m_p = 100$, the proposal kernel is a Gaussian random walk with covariance matrix, $\Sigma = \text{diag}(1 \times 10^{-3}, 1 \times 10^{-5}, 2.5 \times 10^{-5})$, and the discrepancy threshold sequence is $\epsilon_1 = 100$ with $\epsilon_{r+1} = \epsilon_r/2$ for $r = 2, 3, \ldots, 5$; for ABCMLMC we use the discrepancy threshold sequence $\epsilon_0 = 100$ with $\epsilon_{\ell+1} = \epsilon_\ell/2$ for $\ell = 1, 2, \ldots, 4$, and the sample number sequence, $m_0 = 800$, with $m_{\ell+1} = M_\ell/2$ for $\ell = 1, 2, \ldots, 4$. For the prior, we assume all parameters are independent of each other and uniformly distributed with $k_1 \sim \mathcal{U}(0, 2)$, $k_2 \sim \mathcal{U}(0, 0.2)$, and $k_3 \sim \mathcal{U}(0, 0.1)$.

Similarly for the enzyme kinetics model inference problem each algorithm is configured as follows: for the ABC rejection sampler we set $m = 100$ and $\epsilon = 2.5$; for ABCMCMC we set $m_n = 500,000$, $m_b = 100,000$, $m_h = 10,000$, the proposal kernel is a Gaussian random walk with covariance matrix, $\Sigma = \text{diag}(2.25 \times 10^{-8}, 5.625 \times 10^{-7}, 6.25 \times 10^{-6})$, and $\epsilon = 2.5$; for ABCSMC we use $m_p = 100$, the proposal kernel is a Gaussian random walk with covariance matrix, $\Sigma = \text{diag}(2.25 \times 10^{-8}, 5.625 \times 10^{-7}, 6.25 \times 10^{-6})$, and the discrepancy threshold sequence is $\epsilon_1 = 40$ with $\epsilon_{r+1} = \epsilon_r/2$ for $r = 2, 3, \ldots, 5$; for ABCMLMC we use the discrepancy threshold sequence $\epsilon_0 = 40$ with $\epsilon_{\ell+1} = \epsilon_\ell/2$ for $\ell = 1, 2, \ldots, 4$, and the sample number sequence, $m_0 = 800$, with $m_{\ell+1} = M_\ell/2$ for $\ell = 1, 2, \ldots, 4$. For the prior, we assume all parameters are independent of each other and uniformly distributed with $k_1 \sim \mathcal{U}(0, 0.003)$, $k_2 \sim \mathcal{U}(0, 0.015)$, and $k_3 \sim \mathcal{U}(0, 0.05)$.

The resulting marginal posterior distributions are presented in Figure F.1. ABCSMC and ABCMLMC recover the true parameters effectively and as less computationally intensive. For the enzyme kinetic inference problem, more tuning and samples are required

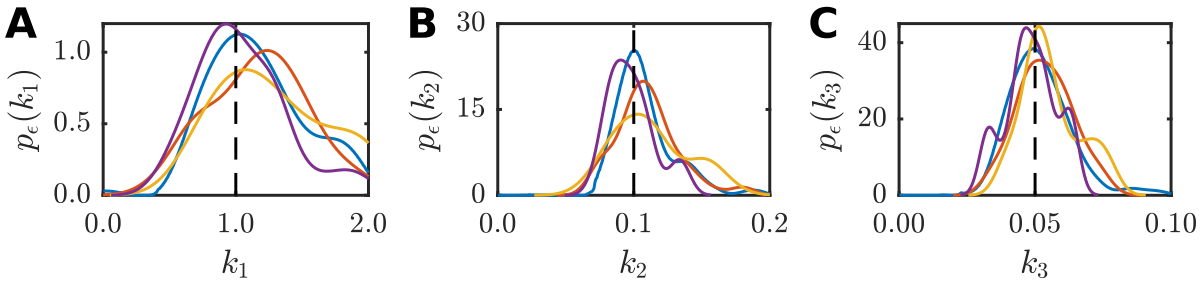

Figure F.1: Comparison of ABC posteriors generated by the ABC rejection sampler (red solid), ABCMCMC (yellow solid), ABCSMC (purple solid) and ABCMLMC (blue solid) for the mono-molecular chain inference problem. The true parameter values (black dashed) are $k_1 = 1.0$, $k_2 = 0.1$ and $k_3 = 0.05$.

to obtain good estimates of the full marginal posterior distributions. Especially, since the ABCMCMC trajectory undergoes a long excursion into the low density tails.

References

[1] S. Barber, J. Voss, and M. Webster. The rate of convergence for approximate Bayesian computation. *Electronic Journal of Statistics*, 9:80–195, 2015.

[2] R. Erban, S. J. Chapman, and P. K. Maini. A practical guide to stochastic simulation of reaction-diffusion processes. *ArXiv e-prints*, 2007. `arXiv:0704.1908 [q-bio.SC]`

[3] P. Fearnhead and D. Prangle. Constructing summary statistics for approximate Bayesian computation: semi-automatic approximate Bayesian computation. *Journal of the Royal Statistical Society Series B (Statistical Methodology)*, 74:419–474, 2012.

[4] M. B. Giles. Multilevel Monte Carlo path simulation. *Operations Research*, 56:607–617, 2008.

[5] T. Jahnke and W. Huisinga. Solving the chemical master equation for monomolecular reaction systems analytically. *Journal of Mathematical Biology*, 54:1–26, 2007.

[6] D. J. Warne, R. E. Baker, and M. J. Simpson. Multilevel rejection sampling for approximate Bayesian computation. *Computational Statistics & Data Analysis*, 124:71–86, 2018.