State and Parameter Estimation from Partial State Observations in Stochastic Reaction Networks

Muruhan Rathinam and *Mingkai Yu*

Department of Mathematics and Statistics
University of Maryland, Baltimore County

Sayas Numerics Day 2022, Sept. 17, 2022

Outline

- Stochastic reaction networks
- Filtering problem: Continuous in time observations.
  - Evolution equations for conditional probabilities.
  - Unnormalized equations and stochastic interpretation.
  - Monte Carlo algorithm: Particle Filter.
  - Numerical examples.
- Filtering problem: Discrete in time observations.

Stochastic reaction networks

n molecular **species** and m **reaction channels**. $Z(t) \in \mathbb{Z}_+^n$: the copy number vector of the species at time t is a CTMC.

- **Stoichiometric vector**: $\nu_j \in \mathbb{Z}^n$, $j = 1, ...m$.
- **Propensity functions** $a_j(z)$:

$$\begin{aligned} a_j(z)h + o(h) = &\text{Probability that reaction } j \text{ will occur in infinitesimal} \\ &\text{time interval } (t, t+h], \text{given} Z(t) = z. \end{aligned}$$

Example: $A + B \longrightarrow C$, $C \longrightarrow A + B$, $\emptyset \longrightarrow C$.

$$\nu_1 = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}, \ \nu_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \ \ \nu_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

Evolution of the p.m.f.

The time evolution of the probability mass function $p(t,z) = P\{Z(t) = z\}$ is given by the *Kolmogorov's forward equations* also known as the *chemical master equations* (CME):

$$p'(t,z) = \sum_{j=1}^{m} a_j(z-\nu_j)\, p(t, z-\nu_j) - \sum_{j=1}^{m} a_j(z)\, p(t,z), \quad \forall t > 0, z \in \mathbb{Z}_+^n$$

It is hard to compute when state space is large or infinite. Necessitates Monte Carlo methods.

Stochastic Simulation Algorithm [Gillespie,1977]

Input: Initial state z_0, terminal time T
Initialize $z = z_0$, $t = 0$
while $t < T$ **do**
    Set $\lambda_i = a_j(z)$ for $j = 1, ..., m$, and $\lambda_0 = \sum_{j=1}^{m} \lambda_j$
    Generate $\tau \sim$ Exponential(λ_0)
    Generate $u \sim$ Uniform$[0, 1]$
    if $t + \tau < T$ **then**
        Choose j that jth reaction occurs with probability $\frac{\lambda_j}{\lambda_0}$
        Set $t = t + \tau$, $z = z + \nu_j$
    else
        Set $t = T$
    end if
end while

Continuous observations: Problem statement

- **Assumption:** Exact (noiseless) observations of the copy number of the last n_2 species in continuous time.
- $Z(t) = (X(t), Y(t))$, where $X(t) \in \mathbb{Z}_+^{n_1}$ is the **unobserved state** and $Y(t) \in \mathbb{Z}_+^{n_2}$ is the **observed state**.
- **Goal:** Estimate the conditional probability distribution

$$\boxed{\pi(t,x) = P\{X(t) = x \,|\, Y(s) = y(s),\, 0 \le s \le t\} \quad \forall x \in \mathbb{Z}_+^{n_1}}.$$

Observable and unobservable reactions

- $\nu_j = (\nu'_j, \nu''_j)$, where $\nu'_j \in \mathbb{Z}^{n_1}$, $\nu''_j \in \mathbb{Z}^{n_2}, j = 1, ..., m$.
- **Observable reactions**: $\mathcal{O} = \{j : \nu''_j \neq 0\}$.
  Unobservable reactions: $\mathcal{U} = \{j : \nu''_j = 0\}$.
- $$a^{\mathcal{O}}(x,y) = \sum_{j \in \mathcal{O}} a_j(x,y), \quad a^{\mathcal{U}}(x,y) = \sum_{j \in \mathcal{U}} a_j(x,y).$$

Example of observable and unobservable reactions

Example:

$$A + B \longrightarrow C, \quad C \longrightarrow A + B,$$
$$\emptyset \longrightarrow C, \ \ \emptyset \longrightarrow A, \ \ C \longrightarrow \emptyset.$$

- Suppose we observe species A. Then observable reactions $\mathcal{O} = \{1, 2, 4\}$. Unobservable reactions $\mathcal{U} = \{3, 5\}$.
- Suppose we observe species B. Then observable reactions $\mathcal{O} = \{1, 2\}$. Unobservable reactions $\mathcal{U} = \{3, 4, 5\}$.

Evolution of $\pi(t, x)$

It can be shown that $\pi(t, x)$ satisfies nonlinear system of ODEs with jumps:

- Let $t_k, k = 1, 2, \ldots$ be the successive jump times of $y(t)$.
- On interval $[t_k, t_{k+1})$

$$\pi'(t,x) = \sum_{j \in \mathcal{U}} \pi(t, x - \nu'_j)\, a_j(x - \nu'_j, y(t_k)) - \sum_{j \in \mathcal{U}} \pi(t,x)\, a_j(x, y(t_k))$$
$$- \pi(t,x) \left(a^{\mathcal{O}}(x, y(t_k)) - \sum_{\tilde{x}} a^{\mathcal{O}}(\tilde{x}, y(t_k)) \pi(t, \tilde{x}) \right) \quad \forall x \in \mathbb{Z}_+^{n_1}. \tag{1}$$

- At jump times t_k

$$\pi(t_k, x) = \frac{\sum_{l \in \mathcal{O}_k} a_l(x - \nu'_l, y(t_{k-1}))\, \pi(t_k-, x - \nu'_l)}{\sum_{\tilde{x}} \sum_{l \in \mathcal{O}_k} a_l(\tilde{x}, y(t_{k-1}))\, \pi(t_k-, \tilde{x})} \quad \forall x \in \mathbb{Z}_+^{n_1}, \tag{2}$$

where $\mathcal{O}_k = \{ j \in \mathcal{O} \mid \nu''_j = y(t_k) - y(t_k-) \}$.

Issues:

- Hard to compute when state space of X is large or infinite.
- Nonlinear evolution does not correspond to a Markov process.

Unnormalized conditional distribution $\rho(t,x)$

Define $\rho(t,x)$ which satisfies linear ODEs with jumps:

- On interval $[t_k, t_{k+1})$,

$$\begin{aligned}\rho'(t,x) = \sum_{j\in\mathcal{U}} \rho(t, x-\nu_j')\, a_j(x-\nu_j', y(t_k)) - \sum_{j\in\mathcal{U}} \rho(t,x)\, a_j(x, y(t_k)) \\ - \rho(t,x)\, a^{\mathcal{O}}(x, y(t_k)) \quad \forall x \in \mathbb{Z}_+^{n_1}.\end{aligned} \quad (3)$$

- At jump times t_k,

$$\rho(t_k, x) = \frac{1}{|\mathcal{O}_k|} \sum_{j\in\mathcal{O}_k} a_j(x-\nu_j', y(t_{k-1}))\, \rho(t_k-, x-\nu_j'), \quad x \in \mathbb{Z}_+^{n_1}, \quad (4)$$

where $\mathcal{O}_k = \{j \in \mathcal{O} \,|\, \nu_j'' = y(t_k) - y(t_k-)\}$.

Advantages of $\rho(t,x)$

- $\pi(t,x) = \rho(t,x)/\sum_{\tilde{x}} \rho(t,\tilde{x})$.
  $\rho(t,x)$ is called **unnormalized conditional distribution**.
- We can relate $\rho(t,x)$ to a Markov process.

More on $\pi(t, x)$ and $\rho(t, x)$

- Filtering formula for discrete time Markov chains [*Fristedt, Jain and Krylov, 2007*].
- Rigorous derivation of evolution equation for $\pi(t, x)$ in the finite state CTMC case is in [*Confortola and Fuhrman, 2013*].
- Intuitive derivation for reaction networks based on the discrete time formula was derived in [*Rathinam and Yu, 2021*].
- The unnormalized equations were introduced in [*Rathinam and Yu, 2021*]. Can be derived by setting $\rho(t, x) = g(t)\, \pi(t, x)$ and and looking for a linear evolution equation for ρ.

Monte Carlo Algorithms

We can relate $\rho(t,x)$ to a Markov process $(V(t), w(t))$:

- $V(t)$ is a proxy for $X(t)$,
- $w(t)$ is a weight process,
- $\rho(t,x) = \mathbb{E}[w(t)\, 1_{\{x\}}(V(t))]$.

In practice, we use an estimator

$$\hat{\rho}(t,x) = \frac{1}{N_s} \sum_{i=1}^{N_s} 1_{\{x\}}(V^{(i)}(t))\, w^{(i)}(t), \tag{5}$$

with i.i.d. samples $V^{(i)}(t), w^{(i)}(t), i = 1, ..., N_s$.

Evolution of $(V(t), w(t))$

In between jump times, $t_k \leq t < t_{k+1}$.

- The process V is evolved according to the underlying reaction network with all observable reactions removed (Gillespie Algorithm).
- The weight process $w(t)$ is evolved according to

$$w(t) = w(t_k)\, \exp\left\{ -\int_{t_k}^{t} a^{\mathcal{O}}(V(s), y(t_k))\, ds \right\}. \qquad (6)$$

Evolution at jump times t_k

The process $(V(t), w(t))$ jumps as follows. Randomly pick one $j \in \mathcal{O}_k$ with uniform probability $1/|\mathcal{O}_k|$, and set

$$V(t_k) = V(t_k-) + \nu'_j, \tag{7}$$

$$w(t_k) = w(t_k-)\, a_j(V(t_k-), y(t_{k-1})), \tag{8}$$

where $\mathcal{O}_k = \{j \in \mathcal{O} \,|\, \nu''_j = y(t_k) - y(t_k-)\}$.

Main Result [Rathinam and Yu, 2021]

The process $(V(t), w(t))$ described by the above procedure satisfies

$$\mathbb{E}[1_{\{x\}}(V(t))\, w(t)] = \rho(t, x),$$

for all $t \geq 0$ and x, where $\rho(t, x)$ is the solution to the unnormalized filtering equations (3) and (4).

An analytically tractable example

We consider the simple reaction network

$$
\begin{aligned}
S &\xrightarrow{c_1} S + A, \\
\varnothing &\xrightarrow{c_2} S, \\
S &\xrightarrow{c_3} \varnothing,
\end{aligned}
\tag{9}
$$

with mass action form of propensities $a_1(z) = c_1 z_2$, $a_2(z) = c_2$, $a_3(z) = c_3 z_2$. Species S is observed exactly while the species A is unobserved. Thus, $Z(t) = (X(t), Y(t)) = (\#A(t), \#S(t))$.

$$
\pi(t, x) = \frac{\lambda^x e^{-\lambda}}{x!}
$$

where $\lambda = \int_0^t c_1 y(t) dt$.

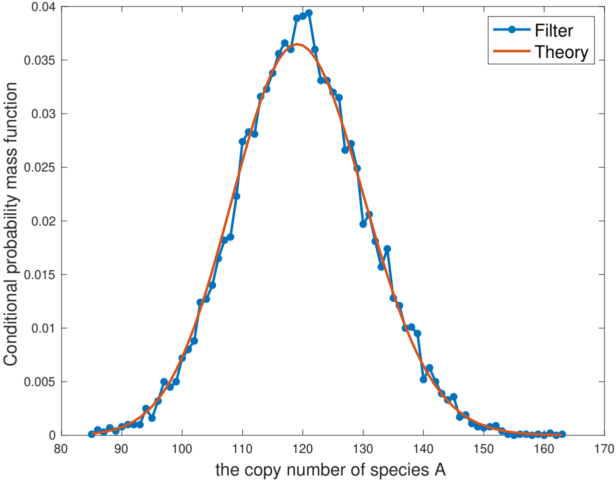

Figure: Estimated conditional probability mass function of the copy number of species A compared with the exact theoretical function at time $T = 20$. $z_0 = (0, 5)$, parameter vector $c = (1, 5, 1)$, filter sample size $N_s = 10,000$.

A point estimate of conditional expectation $\mathbb{E}[X(t) \mid \mathcal{Y}_t]$

$$\mathcal{E}(t) = \frac{\sum_{i=1}^{N_s} V^{(i)}(t)\, w^{(i)}(t)}{\sum_{i=1}^{N_s} w^{(i)}(t)}.$$

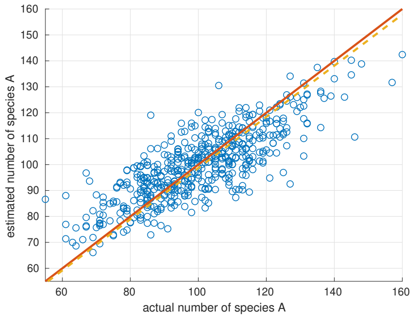

Figure: Scatter plot of the point estimates $\mathcal{E}(T)$ of the copy number of species A (based on the observation of species S) against the actual copy numbers $X(T)$ with $T = 20$. Trial sample size $N_r = 500$, filter particle number $N_s = 1000$.

Genetic toggle switch example

$$\varnothing \xrightarrow{a_1} S_1,\ S_1 \xrightarrow{a_2} \varnothing,\ \varnothing \xrightarrow{a_3} S_2,\ S_2 \xrightarrow{a_4} \varnothing. \tag{10}$$

Let $Z(t) = (X(t), Y(t)) = (\#S_1(t), \#S_2(t))$,
with $a_1(z) = \frac{\alpha_1}{1+y^\beta}$, $a_2(z) = x$, $a_3(z) = \frac{\alpha_2}{1+x^\gamma}$, $a_4(z) = y$.

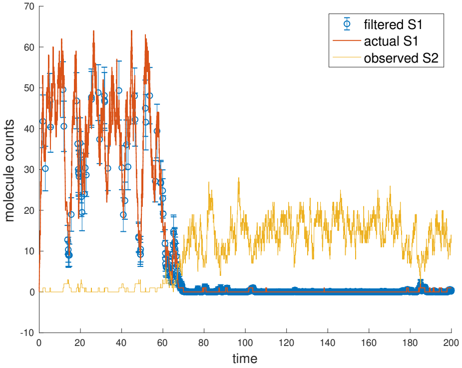

Figure: Estimation of copy number of S_1 at time t based on one observed trajectory of S_2 over $[0, t]$. Estimates are shown at jump times of the observed species S_2 along with their 68% confidence intervals. Filter sample size $N_s = 10,000$. $Z(0) = (0, 0)$, $\alpha_1 = 50$, $\alpha_2 = 16$, $\beta = 2.5$, $\gamma = 1$.

Parameter estimation

Propensity functions $a_j = a_j(x, y, c)$ where $c \in \mathbb{R}^p$.

Bayesian parameter estimation framework:

- Start with a prior distribution $\bar{\mu}$ on the parameter space
- Treat c as a constant-valued stochastic process $C(t)$
- Apply our algorithm to Markov process $Z(t) = (X(t), C(t), Y(t))$

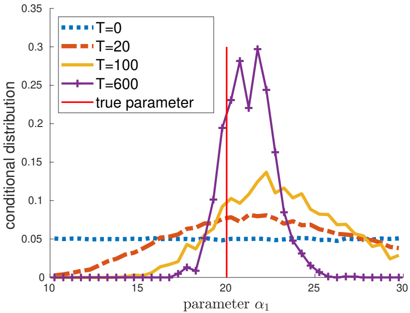

Figure: The posterior conditional probability density function of parameters α_1 and in the genetic toggle example. The nominal parameter values were $\alpha_1 = 20, \alpha_2 = 9, \beta = 2.5, \gamma = 1.$

Discrete time observations: Problem statement

- **Assumption:** Exact (noiseless) observations of the copy number of the last n_2 species at (discrete) snapshots in time.
- $Z(t) = (X(t), Y(t))$, where $X(t) \in \mathbb{Z}_+^{n_1}$ is the **unobserved state** and $Y(t_k) \in \mathbb{Z}_+^{n_2}$ for

$$0 = t_0 < t_1 < t_2 < \dots$$

  are the **observed states**.
- **Focus on two observations** at $t_0 = 0$ and $t_1 = T$.
- **Goal:** Estimate the conditional probability distribution

$$\pi(t, z) = P\{Z(t) = z \,|\, Y(0) = y_0, Y(T) = y\} \quad \forall z \in \mathbb{Z}_+^n.$$

- **Special case:** Chemical bridge process. Given end states $Z(0) = z_0$ and $Z(T) = z_T$, generate sample trajectories according to the conditional distribution.

Approach outline: Focus on reaction count vector $R(T)$

- Observations: $Y_0 = y_0$ and $Y(T) = y$. Assume x_0 is known. $R(t) = (R_1(t), \ldots, R_m(t))$ is the reaction event count vector.

$$X(t) = x_0 + \nu' R(t),$$
$$Y(t) = y_0 + \nu'' R(t).$$

- Change in observed species: $y - y_0 = \nu'' R(T)$.
- Split $R(T) = (R'(T), R''(T))$, where $R'(T)$ is "free" and $R''(T)$ is determined by $R'(T)$.
- Assume that (under simulation probability measure P_0) process R is m-variate Poisson (independent components) with possibly time dependent intensity $\lambda(t, x_0, y_0, y)$.
- Proxy processes: I is a proxy for R, $\tilde{Z}(t) = z_0 + \nu I(t)$. $\tilde{Z} = (U, V)$ is a proxy for $Z = (X, Y)$.

Main Result

Compute the Girsanov change of measure process $\tilde{L}(t)$:

$$\tilde{L}_j(t) = \left(\prod_{i \geq 1} \frac{a_j(U(\tilde{T}_i^j-), V(\tilde{T}_i^j-))}{\lambda_j(T_i^j-, y_0, y)} 1_{\{\tilde{T}_i^j \leq t\}} \right) \times \exp\left(\int_0^t (\lambda_j(s, y_0, y) - a_j(U(s), V(s)))\, ds \right),$$

where $\tilde{T}_i^j$ for $i = 1, 2, \ldots$ are the ith jump time of I_j.

Main Result : Rathinam and Yu

$$P\{Z(t) = z \,|\, O\} = \frac{\mathbb{E}_0[1_{\{z\}}(\tilde{Z}(t))\, \tilde{L}(T)\, W |\, O]}{\mathbb{E}_0[\tilde{L}(T)\, W |\, O]}$$

where $O = \{Y_0 = y_0, Y(T) = y\}$ is the observed event.

References

- F. Confortola and M. Fuhrman, *Filtering of continuous-time Markov chains with noise-free observation and applications*, Stochastics, 85:2, 216-251, 2013.
- B. Fristedt, N. Jain and N. Krylov, *Filtering and prediction: A primer*, American Mathematical Society, 2007.
- M. Rathinam and M. Yu, *State and parameter estimation from exact partial state observation in stochastic reaction networks*, Journal of Chemical Physics, 154, 034103, 2021.