

# **Simultaneous Stochastic Simulation of Multiple Perturbations in Biological Network Models**

Werner Sandmann

University of Bamberg

Department of Information Systems & Applied Computer Science

CMSB 2007, Edinburgh, Scotland, 20 September 2007

# Outline

- Stochastic Chemical Kinetics
  - ☞ Chemical Master Equation
  - ☞ Direct Stochastic Simulation Algorithm
  
- Foundations for Simultaneous Stochastic Simulation
  - ☞ Importance Sampling and Likelihood Ratios
  - ☞ Formulation for Biological Networks
  - ☞ Appropriate Interpretation
  
- Simultaneous Stochastic Simulation Algorithm
  
- Numerical Results and Runtime Comparison
  
- Conclusion and Further Research

# Coupled Chemical Reactions

- ➔ Molecular Species  $S_1, \dots, S_d$
- ➔ Reaction channels  $R_1, \dots, R_M$  mit  $R_m, m \in \{1, \dots, M\}$
- ➔ For each reaction channel  $R_m$ 
  - ☞ reaction rate constant  $c_m \rightsquigarrow$  reaction rate
  - ☞ state change vector  $v_m = (v_{m_1}, \dots, v_{m_d})$  where
    - $v_{m_k}$  denotes change of molecules of species  $S_k$  due to reaction  $R_m$
    - determined by stoichiometric coefficients

## Stoichiometry of chemical reaction $R_m$



- ➔ Involved species  $S_{m_1}, \dots, S_{m_\ell}, \ell \in \mathbb{N}, m_1, \dots, m_\ell \in \{1, \dots, d\}$
- ➔ Stoichiometric coefficients  $s_{m_1}, \dots, s_{m_\ell} \in \mathbb{N}$

# Stochastic Model

- ➔  $X_k(t)$ : number of molecules of species  $S_k$  at time  $t \geq 0$
- ➔  $X(t) = (X_1(t), \dots, X_d(t))$ : system state at time  $t \geq 0$
- ➔ State transitions occur due to chemical reaction  $R_m$ ,  $m \in \{1, \dots, M\}$
- ➔ Reaction rate of each  $R_m$  given by **propensity function**  $\alpha_m$
- ➔  $\alpha_m(x)h$  is the conditional probability that  $R_m$  reaction occurs in infinitesimal time interval  $[t, t + h)$ , given state  $x$  at time  $t$ :

$$\alpha_m(x)h = P(\text{Reaction of type } R_m \text{ in } [t, t + h) \mid X(t) = x)$$

- ➔  $\alpha_m$  given by  $c_m$  times the number of possible combinations of the required reactants

$$\alpha_m(x) = c_m \cdot \prod_{i=1}^{m_r} \binom{x_{m_j}}{s_{m_j}}$$

where  $x_{m_j}$  denotes the number of molecules of species  $S_{m_j}$  present in state  $x$

Temporal evolution described by stochastic process  $(X(t), t \geq 0)$

# Chemical Master Equation

Transient state probabilities (initial state  $x_0$  at time  $t_0$ )

$$p^{(t)}(x) := p^{(t)}(x|x_0, t_0) = P(X(t) = x \mid X(t_0) = x_0)$$

Temporal evolution, complete description of  $(X(t), t \geq 0)$

$$p^{(t+h)}(x) = \underbrace{p^{(t)}(x) \cdot \left(1 - \sum_{m=1}^M \alpha_m(x)h\right)}_{\text{At time } t \text{ in } x \text{ and no reaction}} + \underbrace{\sum_{m=1}^M p^{(t)}(x - v_m|x_0, t_0)\alpha_m(x - v_m)h}_{\text{At time } t \text{ in } x - v_m \text{ and reaction of type } R_m}$$

**Chemical Master Equation (CME)**

$$\frac{\partial p^{(t)}(x)}{\partial t} = \sum_{m=1}^M \left( \alpha_m(x - v_m)p^{(t)}(x - v_m) - \alpha_m(x)p^{(t)}(x) \right)$$

# Kolmogorov Differential Equations

CTMC with transition probability matrix  $\mathbf{P}(h) = (p_{ij}(h))_{i,j \in \mathbb{N}}$  and generator  $\mathbf{Q} = (q_{ij})_{i,j \in \mathbb{N}}$ .

Kolmogorov Forward

$$\frac{\partial}{\partial t} \mathbf{P}(t) = \mathbf{P}(t) \mathbf{Q}$$

Kolmogorov Backward

$$\frac{\partial}{\partial t} \mathbf{P}(t) = \mathbf{Q} \mathbf{P}(t)$$

Kolmogorov Global

$$\frac{\partial}{\partial t} p^{(t)} = p^{(t)}$$

**Kolmogorov Global Differential Equation** explicit form

$$\frac{\partial p_i^{(t)}}{\partial t} = \sum_{j:j \neq i} p_j^{(t)} q_{ji} - \sum_{j:j \neq i} p_i^{(t)} q_{ij} = \sum_{j:j \neq i} \left( p_j^{(t)} q_{ji} - p_i^{(t)} q_{ij} \right)$$

Equivalence of CME and Kolmogorov differential equations easily seen by interpreting  $i \in \mathbb{N}$  as the number assigned to state  $x \in \mathcal{S}$ ,  $p_i^{(t)} = p^{(t)}(x)$ :

➔  $q_{ij} = \alpha_m(x)$  if  $j$  is the number assigned to state  $x + v_m$

➔  $q_{ji} = \alpha_m(x - v_m)$  if  $j$  is the number assigned to state  $x - v_m$

# Stochastic Simulation

Temporal evolution of number of molecules is CTMC

Hence, simulation of coupled molecular reactions is simulation of a CTMC.

→ Simple direct generation of trajectories.

➔ Formulation in terms of the CME due to Gillespie (1976,1977):

Init  $t := t_0$  and  $x := x_0$

**repeat**

1. Compute all  $\alpha_m(x)$  and  $\alpha_0(x) := \alpha_1(x) + \dots + \alpha_M(x)$
2. Generate two random numbers  $u_1, u_2$ , uniformly distributed on  $(0, 1)$
3. Generate time  $\tau$  to next reaction:  $\tau = -\ln(u_1)/\alpha_0(x)$
4. Determine reaction type:  $m = \min\{k : \alpha_1(x) + \dots + \alpha_k(x) > u_2\alpha_0(x)\}$
5. Update  $t := t + \tau$ ;  $x := x + v_m$
6. Store/Collect/Handle Data

**until** terminating condition

# Importance Sampling Foundations

- ➔ Two probability measures  $P$  and  $P^*$  on a measurable space  $(\Omega, \mathcal{A})$
- ➔  $P$  absolutely continuous with respect to  $P^*$

$$\forall A \in \mathcal{A} : P^*(A) = 0 \Rightarrow P(A) = 0$$

- ➔ Radon-Nikodym theorem guarantees that the **Radon-Nikodym derivative** aka **Likelihood Ratio**  $L = dP/dP^*$  exists, and that

$$\forall A \in \mathcal{A} : P(A) = \int_A L(\omega) dP^*$$

- ➔ For random variable  $Y$  on  $(\Omega, \mathcal{A})$  :

$$E_P[Y] = \int Y(\omega) dP = \int Y(\omega) L(\omega) dP^* = E_{P^*}[Y L]$$

This is exploited by Importance Sampling.

# Importance Sampling Simulation

General setting of Importance Sampling Simulation for estimating  $E_P[Y]$

- ➔ **Change of Measure** Choose measure  $P^*$ , only requirement: absolute continuity
- ➔ Sample  $Y_1, \dots, Y_N$ , iid according to  $P^*$

## Importance Sampling Estimator

$$\hat{\gamma}_{IS} := \frac{1}{N} \sum_{i=1}^N Y_i L(Y_i)$$

unbiased estimator for  $E_P[Y]$

Specific applications require according distribution/density

→ of trajectories in case of Markov chains

→ of “reaction paths” (trajectories) in case of coupled molecular reactions

↔ likelihood ratios

# Molecular Reaction Path Density

Time evolution completely described by sequence of states at reaction times and times between reactions

→  $(x(t_0), \tau_0), (x(t_1), \tau_1), (x(t_2), \tau_2), \dots$  describes a trajectory.

Due to the Markov property ( $\Rightarrow$  exponentially distributed times between reactions) the reaction path density up to  $R$ -th reaction is given by

$$p^{(t_0)}(x_0) \cdot \prod_{i=1}^R \alpha_{m_{i-1}}(x(t_{i-1})) \exp(-\alpha_0(x(t_{i-1}))\tau_{i-1})$$

where  $R_{m_i}$  is the type of  $i$ -th reaction with according propensity  $\alpha_{m_{i-1}}$  and

$$\alpha_0(x(t_{i-1})) := \alpha_1(x(t_{i-1})) + \dots + \alpha_M(x(t_{i-1}))$$

# Requirements for Importance Sampling

- Change of measure, only requirement absolute continuity  $\Rightarrow$  great freedom
  - ☞ All reaction paths that are possible under original measure must remain possible
  - ☞ Each measure on the sample path space that meets this is allowed (need not be Markovian)
- In practice: avoid large increase in trajectory generation efforts
  - ☞ Usual and most natural: stay in Markovian world
  - ☞ Change propensity functions  $\alpha_m \rightsquigarrow \alpha_m^*$
  - ☞ Absolute continuity by condition  $\alpha_m^*(x) = 0 \Rightarrow \alpha_m(x) = 0$  for all  $x$   
equivalently  $\alpha_m(x) > 0 \Rightarrow \alpha_m^*(x) > 0$  for all  $x$
- Thus (keeping initial distribution/state unchanged)

$$L(\omega) = \prod_{i=1}^R \frac{\alpha_{m_{i-1}}(x(t_{i-1})) \exp(\alpha_0(x(t_{i-1}))\tau_{i-1})}{\alpha_{m_{i-1}}^*(x(t_{i-1})) \exp(\alpha_0^*(x(t_{i-1}))\tau_{i-1})}$$

- Of course, more general change of measure possible (see paper)

# Reversing Roles in Importance Sampling

- ➔➔ Originally, Importance Sampling was not intended for comparing multiple models but for reducing the variance of the simulation estimator for a single model.
- ➔➔ **Reversing the roles of the original distribution and the IS distribution**
  - ☞ elegant method for obtaining estimates for arbitrarily many models simultaneously from one single simulation experiment (all consisting of independent runs).

## How to achieve this

- ➔➔ Nominal model (unperturbed parameters) as the Importance Sampling model.
  - ☞ Perform simulation with the dynamics of the nominal model.
- ➔➔ Each perturbed model takes the role of the original model in usual IS.
  - ☞ Simulate the nominal model and weight the outcomes by the likelihood ratio to obtain unbiased estimates for the perturbed model.
  - ☞ Keep track of multiple likelihood ratios resulting from the perturbed models and update them all after each simulated reaction for the nominal model.

## Some Notes

➔ **Valid for any perturbed model where the stoichiometry is the same as for the unperturbed model.**

☞ Arbitrary changes/perturbations of the rate constants and thus the propensity functions can be considered.

➔ Algorithm involves trajectory generation of the nominal model

☞ relies on an appropriate algorithm for this purpose,

☞ extends this algorithm by incorporating the likelihood ratio computations.

➔ Any algorithm for trajectory generation can be used and equipped with the feature of Importance Sampling.

**Here, we choose the direct method for trajectory generation.**

➔  $N$  perturbed models should be simulated simultaneously.

➔ Propensity functions of the nominal model are denoted by  $\alpha_m^*$ ,  $m = 1, \dots, M$ .

➔ Propensity functions of the perturbed models denoted by  $\alpha_m^{(1)}, \dots, \alpha_m^{(N)}$ ,  $m = 1, \dots, M$ .

# Simultaneous Stochastic Simulation

Init  $t := t_0$  and  $x := x_0$  and  $L_1 = \dots = L_N = 1$

**repeat**

1. Compute all  $\alpha_m^*(x)$  and  $\alpha_0^*(x) := \alpha_1^*(x) + \dots + \alpha_M^*(x)$

2. **for all**  $i = 1, \dots, N$

    Compute all  $\alpha_m^{(i)}(x)$  and  $\alpha_0^{(i)}(x) := \alpha_1^{(i)}(x) + \dots + \alpha_M^{(i)}(x)$

3. Generate two random numbers  $u_1, u_2$ , uniformly distributed on  $(0, 1)$

4. Generate time  $\tau$  to next reaction:  $\tau = -\ln(u_1)/\alpha_0^*(x)$

5. Determine reaction type:  $m = \min\{k : \alpha_1^*(x) + \dots + \alpha_k^*(x) > u_2\alpha_0^*(x)\}$

6. **for all**  $i = 1, \dots, N$

    Update  $L_i = L_i \cdot \alpha_m^{(i)}(x)/\alpha_m^*(x) \cdot \exp((\alpha_0^*(x) - \alpha_0^{(i)}(x)) \cdot \tau)$

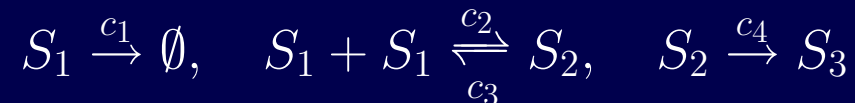
7. Update  $t := t + \tau$ ;  $x := x + v_m$

8. Store/Collect/Handle Data

**until** terminating condition

# Representative Example

## Systematical Studies of Decaying Dimerization



➔➔ Initial numbers of molecules  $X_1(0) = 400$ ,  $X_2(0) = 798$ ,  $X_3(0) = 0$

➔➔ Nominal rate constants  $c_1 = 1$ ,  $c_2 = 10$ ,  $c_3 = 1000$ ,  $c_4 = 0.1$

➔➔ Nominal propensity functions

$$\alpha_1(x) = x_1, \quad \alpha_2(x) = 5x_1(x_1 - 1), \quad \alpha_3(x) = 1000x_2, \quad \alpha_4(x) = 0.1x_2$$

➔➔ Simulations from time  $t_0 = 0$  to time  $t_{end} = 0.2$  and to time  $t_{end} = 0.5$

➔➔ For both time horizons two series of simulations of perturbed models

☞ series with exactly one of the reaction rates perturbed by  $\pm 3\%$   
altogether 8 parameter settings were considered

☞ series with all reaction rates potentially perturbed by  $\pm 3\%$   
including the nominal one, altogether  $3^4 = 81$  parameter settings were considered  
(80 differently perturbed parameter settings)

# Runtimes

**Runtimes of matching simulations** with direct method and simultaneous method.

- ➔ 8 and 80, respectively, separate simulation experiments with direct method, each consisting of  $10^4$  simulation runs up to the given time horizons
- ➔ 8 and 80, resp., simultaneous simulations of all perturbed parameter settings with  $10^4$  simulation runs up to the given time horizons

	$t = 0.2$		$t = 0.5$	
# Perturbed Models	Direct	Simult	Direct	Simult
8	28928	8178	72504	18528
80	290155	51449	725871	99303

- ➔ Not surprisingly, the runtime of the direct method for  $N$  perturbed parameter settings is roughly  $N$  times the runtime for one single setting.
- ➔ With the simultaneous method the runtime grows significantly less with increasing number of parameter settings.

**The extra effort according to the likelihood ratio computation is far less than the effort for a separate simulation.**

## Runtime Analysis in a Nutshell

- ➔ Let  $r$  denote the runtime for one separate simulation experiment (with sufficiently many independent runs) for the nominal parameter settings with the direct method.
  - ☞ For small parameter perturbations roughly the same for all parameter settings.
  - ☞ Runtime for simulating  $N$  parameter settings separately is  $R_{\text{sep}} \approx Nr$ .
- ➔ Let  $v$  denote the overhead (extra effort) for one additional parameter setting.
  - ☞ Runtime for the simultaneous method is  $R_{\text{simult}} = r + Nv$ .

Obviously, when only one parameter setting is considered  $R_{\text{sep}} = r < r + v = R_{\text{simult}}$  which simply means that in this case of course no extra effort should be introduced.

### Condition under which the simultaneous method is faster

$$R_{\text{simult}} < R_{\text{sep}} \Leftrightarrow v < r - \frac{r}{N}$$

is very likely to be met for almost all models, of any size and complexity.

In fact, we studied many more models.

# Conclusion and Further Research

Algorithm for simultaneous simulation of multiple parameter settings within one single simulation experiment

- ➔ inspired by Importance Sampling, appropriately adapted
- ➔ likelihood ratios appropriately weight the results to obtain unbiased estimates
- ➔ particularly useful for comparisons of a large number of parameter settings
- ➔ large amount of runtime gain compared to multiple separate simulations
- ➔ no approximation and thus no loss in statistical validity and accuracy (with respect to underlying algorithm for trajectory generation)

Further Research

- ➔ analytical investigations of the algorithm performance
- ➔ other methods for trajectory generation
  - 👉 approximate methods, e.g. tau-leaping, small/multiscale SSA, . . .
  - 👉 uniformization-based stochastic simulation (Sandmann, FOSBE 2007)