

Applied Numerics SS'17  
Lecture 4

Max von Kleist, Vikram Sunkara, Vitaly Belik

May 17, 2017



# Chapter 1

## Lecture 4: Stochastic simulation

### 1.1 Recap

#### 1.1.1 Definitions

A simple definition of a stochastic process is that it is a random variable distributed according to a probability distribution which is evolving through time.

**Definition 1.** Let  $\Omega$  be the state space. Let  $T$  be a finite subset of  $[0, \infty)$ . If we have a family of probability distributions indexed by time  $\{p_t : t \in T\}$ . Then the collection of all random variables  $X_t \sim p_t$  is called a stochastic process.

**Definition 2.** We define  $\mathcal{P}(\lambda t)$  to be the Poisson Process at time  $t > 0$  with rate  $\lambda > 0$ . That is, for  $x \in \Omega$ ,

$$p(\mathcal{P}(\lambda t) = x; t) := \frac{(\lambda t)^x e^{-\lambda t}}{x!}.$$

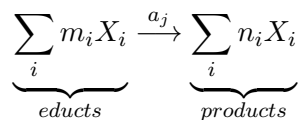
**Theorem 1.** The sum of two independent Poisson processes is a Poisson process,

$$\mathcal{P}(\lambda_1 t) + \mathcal{P}(\lambda_2 t) = \mathcal{P}((\lambda_1 + \lambda_2)t),$$

with  $\lambda_1, \lambda_2 > 0$ .

#### 1.1.2 Reactions

Let  $\Omega = \mathbb{N}_0$ . We are given a set of reactions  $j$  with



has  $m_1 \dots m_q$  educt molecules and produces  $n_1 \dots n_q$  product molecules of species  $1 \dots q$ .

- Its *stoichiometric change vector* is  $s_{\cdot,j} = n_{\cdot,j} - m_{\cdot,j}$ .
- Each reaction has reaction rate  $a_j(X_t) = k_j \cdot \prod_i X_{i,t}^{m_{i,j}} \dots$  (elementary/simple reactions).

Our Poisson process is thus defined by:

$$X_t = X_0 + \sum_j s_{\cdot,j} \mathcal{P} \left( \int_0^t a_j(X_s) ds \right). \quad (1.1)$$

**Remark 1.** We denote by

$$N_{t,j} := \mathcal{P} \left( \int_0^t a_j(X_s) ds \right)$$

a process which counts how many reaction events of type  $j$  occurred by time  $t$ .

## 1.2 Gillespie

Now let

$$N_t := \sum_j N_{t,j} = \sum_j \mathcal{P} \left( \int_0^t a_j(X_s) ds \right)$$

be the process which counts how many events occurred in *total*. Using Theorem 1, we can deduce that,

$$N_t = \mathcal{P} \left( \sum_j \int_0^t a_j(X_s) ds \right).$$

Now we wish to sample for a time at which **one** reaction occurs. That is  $t^{(1)}$  such that  $N_{t^{(1)}} = 1$ . Since  $N_t$  is a Poisson Process, the first firing time is exponentially distributed. In our case,

$$a_0 t = \sum_j \int_0^t a_j(X_s) ds$$

where  $\lambda$  is the rate for the process  $N_t$  fires. Since we are trying to compute the first time at which an event occurs, the process  $X_\bullet$  has not changed, it is still  $X_0$ . Hence, substituting this into the integral gives us

$$\lambda t^{(1)} = a_0 t^{(1)} = \sum_j a_j(X_{t_0}) t^{(1)}$$

Then, a sample for the time of a reaction firing is

$$t^{(1)} = \frac{1}{a_0} \log \left( \frac{1}{\tau} \right),$$

where  $\tau \sim U[0, 1)$ .

### 1.2.1 Efficacy

**Note:** We have to consider the time until the first reaction fires, which can become inefficient, as

$$\lim_{a_0 \rightarrow \infty} t^{(1)} \rightarrow 0.$$

**Note:** The above condition is true for a system with first- and second order reactions, in which the number of reactants/educts of at least one reaction tend to become very large  $X_i \rightarrow \infty$ .

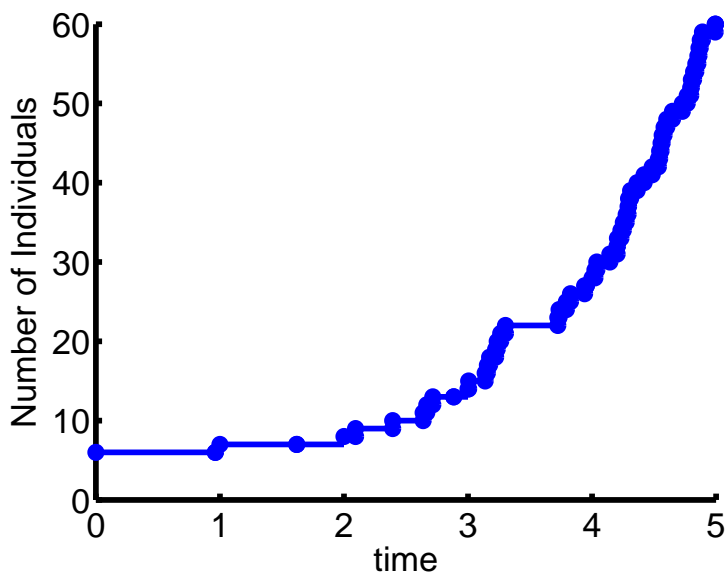


Figure 1.1: Pure birth process. The time steps tend to zero as the process evolves.

## 1.3 Tau-leaping

### 1.3.1 Idea

**Idea:** The idea was proposed by Dan Gillespie (2001), J. Chem Phys. 115:1716. Like in an *explicit Euler method*, we fix a time-step  $\Delta t$  (called *tau-leap*) and make the following assumption

$$X_{t+\Delta t} \approx X_t + \sum_j s_{\cdot,j} \mathcal{P}(a_j(X_t)\Delta t). \quad (1.2)$$

This computational procedure is known as the *tau-leaping* approximation. We denote with

$$\hat{N}_{\Delta t, j} := \mathcal{P}(a_j(X_t)\Delta t)$$

the number of firings of reaction channel  $j$  when we applied the *tau-leaping* approximation with time step  $\Delta t$ .

### 1.3.2 When the leap condition is good

**If each reaction fires at most once and reaction channels  $j$  are independent** (in case of multiple reaction channels). The condition holds that  $a_j(X_t)\Delta t = \int_0^t a_j(s) ds$  and thus

$$N_{t, j} = \hat{N}_{\Delta t, j}$$

**Note:** If reaction rates do not depend on  $X$  (zero order reaction), leap conditions are exact.

**Note:** If the reactant molecule populations are very large, it will take a very large number of reaction events to change the propensity functions “noticeably”.

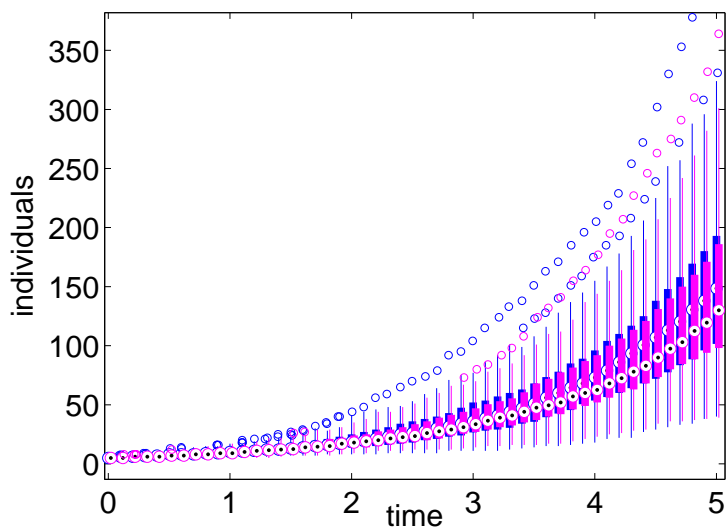


Figure 1.2: Statistics of the SSA (blue) vs. the Poisson simulations (magenta) with fixed  $\Delta t = 0.1$  for the pure birth process.

### 1.3.3 How to choose the time-step/leap size

**Large leaps:** In case of large leaps the above condition is for most models violated. I.e.

$$N_{t,j} \neq \hat{N}_{\Delta t,j}$$

**Small leaps:** If we were to take  $t$  to be the relatively small value, i.e.  $\Delta t = 1/a_0$ , the tau-leaping is likely more accurate, there is of little benefit as the computational burden is unchanged.

*Example:* For  $\Delta t = 1/a_0 \Rightarrow \mathbb{E}(N_{\Delta t}) = 1$  (on average only one reaction occurs in total).

**Goal:** Select the largest  $\Delta t$  that is compatible with the leap conditions.

**Post-leap check:** Gillespie proposed the following scheme to assess the accuracy of the tau-leaping method:

$$|a_j(X_t + \sum_j \hat{N}_{\Delta t,j} \cdot s_{.,j}) - a_j(X_t)| \rightarrow \text{small}$$

In a post-leap scheme this would result in a computationally inefficient rejection scheme (leap, check leap, accept/reject step, possibly decrease leap size). Also, large fluctuations may be legitimate in some cases (FN rejection).

**Pre-leap check:** We know that

$$\mathbb{E}(\hat{N}_{\Delta t,j}) = a_j(X_t)\Delta t \quad (\text{Poisson distr. random number})$$

Therefore, the *expected net change* in  $[t, t + \Delta t)$  will be

$$\nu = \sum_j s_{.,j} a_j(X_t) \Delta t = \Delta t \xi(X_t)$$

where  $\xi(X_t)$  is the *expected net change* in a *unit* of time. As proposed by Gillespie, we may derive a scheme in which the leap condition may be valid, i.e.

$$|a_j(X_t + \nu) - a_j(X_t)| \leq \epsilon a_0(X_t) \quad (\text{for } j \in [1 \dots J])$$

and with  $0 < \epsilon < 1$ .

We do a first-order Taylor approximation to the left side, i.e.

$$a_j(X_t + \nu) - a_j(X_t) \approx \nu \nabla a_j(X_t) = \sum_i \Delta t \xi_i(X_t) \frac{\partial a_j(X_t)}{\partial X_i} = \sum_i \Delta t \xi_i(X_t) b_{j,i}(X_t)$$

Thus we get

$$\Delta t \left| \sum_i \xi_i(X_t) b_{j,i}(X_t) \right| \leq \epsilon a_0(X_t).$$

Substituting everything together gives

$$\Delta t = \min_{j \in [1 \dots J]} \left( \frac{\epsilon a_0(X_t)}{\left| \sum_i \xi_i(X_t) b_{j,i}(X_t) \right|} \right).$$

If  $\Delta t \leq \frac{c}{a_0(X_t)}$  use the SSA instead, where  $c$  is a small integer value.

### 1.3.4 Negative numbers

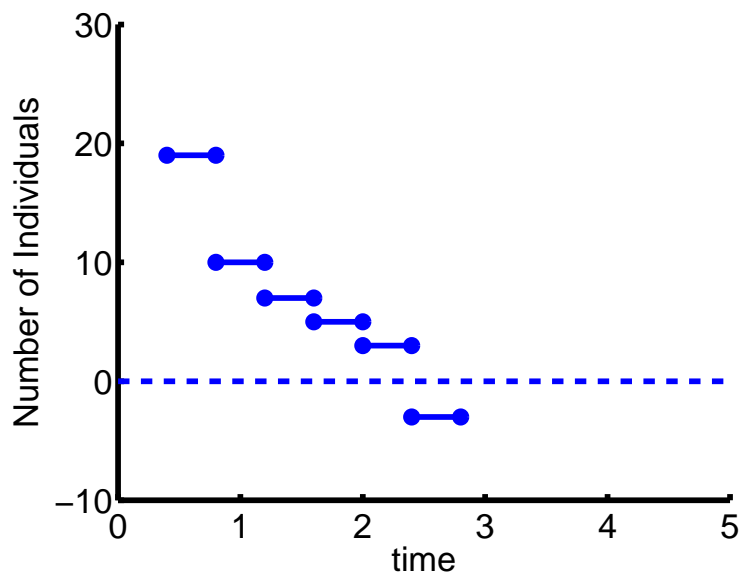


Figure 1.3: Example of a Poisson simulations of a birth-death process with  $\delta > \lambda$  with fixed  $\Delta t = 0.4$ .

## 1.4 Tau-Leaping Project

- How to do choose the leap size cleverly and avoid negative numbers, i.e. Cao et al. (2006) *The Journal of Chemical Physics* 124 (044109): 1–11

## 1.5 Extrande

**Observation:** The coefficient of variation becomes small as  $a_0 \rightarrow \infty$ .

Since we have

$$\mathbb{E}[t^{(1)}] = \frac{1}{a_0}$$

reactions occur in infinitesimally small time units for  $a_0 \rightarrow \infty$ . For reactions of **first-** or **higher order**, this is also the case when  $X \rightarrow \infty$ .

The effective result from this is that stochastic fluctuations become small when  $t^{(1)} \rightarrow 0$ .



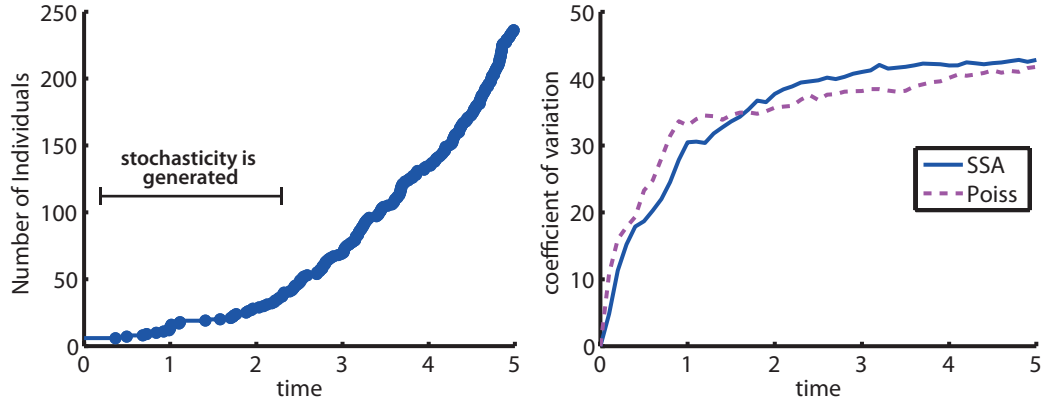


Figure 1.4: Left: Stochastic trajectory for the pure birth process. Right: Coefficient of Variation  $\sigma/\mu$  over 100 simulations. The plot shows that stochasticity is generated in the low copy number regimen and at some point only propagated (it does not increase further).

Thus, if we have a subsystem for which for all time points  $X_i \gg 1$  such that  $a_{0,\mathcal{J}} \gg 1$

$$\lim_{X_i \rightarrow \infty} \frac{X_t}{V} = \mathbb{E}[X_t], \quad (1.3)$$

where  $V$  in the equation above is a scaling parameter (reaction volume usually). The equation above motivates the use of an ODE framework.

### 1.5.1 Idea

We decompose our system into

- 'continuous' species  $Y$  for which the approximation (1.3) is reasonable.
- 'discrete' species  $Z$  whose changes in number require a discrete stochastic description.

The partitioning is disjoint and we have  $X = (Y, Z)^T$ . Within this course we will only focus, unlike in the original article, on subsystems  $Y$  that can be described by ODEs (unlike SDEs as in the original article.)

For  $Y$  we therefore choose

$$\frac{d}{dt} Y_t = S_r \cdot r(Y_t, Z_t)$$

where  $r(Y_t, Z_t)$  denotes the vector of reaction rates that change  $Y$ , but not  $Z$ .  $S_r$  is the corresponding stoichiometric matrix. As is evident from the equation above,  $Y$  changes continuously in time while  $Z$  will remain constant until the next reaction fires that changes

it (next section).

For  $Z$  we choose a stochastic description

$$Z_t = Z_0 + \sum_{\varsigma} s_{,\varsigma} \mathcal{P} \left( \int_0^t a_{\varsigma}(Y_s, Z_s) ds \right),$$

where  $a_{\varsigma}(Y_s, Z_s)$  denotes the propensity functions that affect  $Z$  (and which may also affect  $Y$ ) with stoichiometric change vector  $s_{,\varsigma}$ .

Like in the Gillespie algorithm, we want to sample the first time  $t^{(1)}$  at which a reaction  $\varsigma$  occurs that alters  $Z$ . We know that  $Z$  is unchanged until then, but that  $Y$  has changed. We therefore get *time-dependent propensity functions*  $a_{0,t} = \sum_{\varsigma} a_{\varsigma}(t)$ , where the time-dependent component is exclusively determined by  $Y$ . The dependence on time means we cannot sample the process easily.

### 1.5.2 Algorithm

The idea of the extrande algorithm is simple:

1. We determine an upper bound for  $a_{0,t}$ , which we call  $B$  (a save choice is  $a_{0,t} \leq B \forall t$ ).
2. We determine

$$t^{(1)} = \frac{1}{B} \log \left( \frac{1}{\tau} \right),$$

and integrate  $Y$  from  $t^{(0)}$  to  $t^{(1)}$ . This allows us to compute  $a_{0,t}$ .

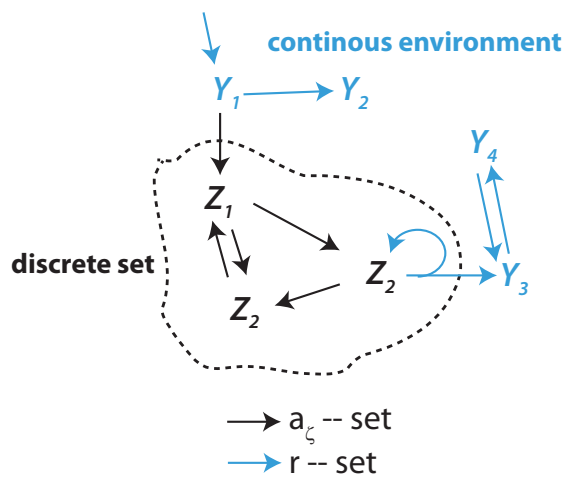
3. With probability  $B - a_{0,t}$  we do not execute any reaction  $\varsigma$  (rejection), else we execute a stochastic reaction  $\varsigma$  that alters  $Z$  (and possibly  $Y$ ).
4. We repeat the procedure until the final time is reached

Obviously, the efficacy of the Extrande algorithm is highly dependent on the right choice of  $B$ . The difficulty is to find  $B$  such that  $a_{0,t} \leq B$ , at the same time if  $B \gg a_{0,t}$  many rejections will occur making the algorithm inefficient.

## 1.6 Extrande project

- Task: How to do choose  $B$  appropriately
- Resource: Voliotis et al. (2016) *Plos Computational Biology* 12(6): e1004923

## A. Hybrid approach



## B. Rejection sampling

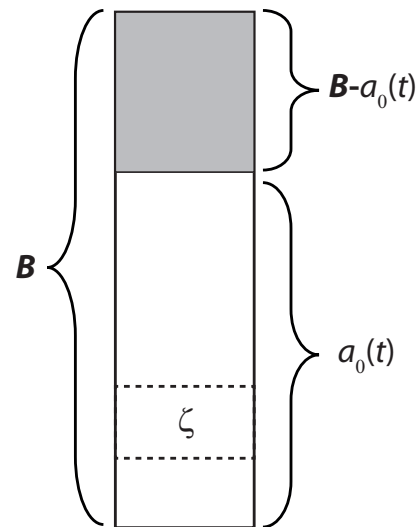


Figure 1.5: Principle of the ExtranDE algorithm. Species are partitioned into continuous and discrete ( $Y, Z$ ), and reactions are partitioned on the basis of whether they solely affect the continuous species (reactions  $r$ ) or whether they affect the stochastic species (propensity functions  $a_\zeta$ ). B: Reactions are chosen according to the Gillespie scheme. With probability  $B - a_{0,t}$  an empty reaction (rejection) is sampled (grey area).