



Aalto University

The degradation reaction $\mathcal{A} \rightarrow \emptyset$

Stochastic analysis and simulation of reactive and diffusive systems

Francesco Corona ($\sim\!\!\sim$) and Augusto Magalhães

Chemical and Metallurgical Engineering
School of Chemical Engineering

$$\mathcal{A} \rightarrow \emptyset$$

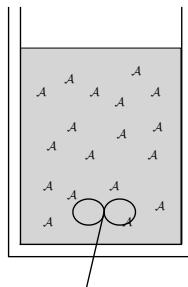
Probabilistic formulation

$$\mathcal{A} \rightarrow \emptyset$$

We start with the degradation of some chemical species \mathcal{A} to an uninteresting form \emptyset

$$\mathcal{A} \xrightarrow{\kappa} \emptyset$$

The equation does not state that \mathcal{A} is degraded into nothing, but rather that it is degraded into uninteresting species (or, that there is an outflow of \mathcal{A})



These reactions are intended to capture the change of a molecule of species \mathcal{A} into one or more other molecules, such as radioactive decay or dissociation of complex molecules

\rightsquigarrow They are not intended to model conversions in the presence of a catalyst

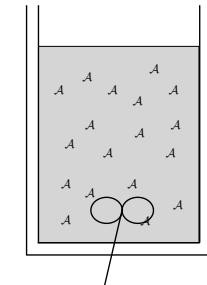
In systems biology, they are used as simplified models for RNA and protein degradation

$$\mathcal{A} \rightarrow \emptyset \text{ (cont.)}$$

$$\mathcal{A} \xrightarrow{\kappa} \emptyset$$

Reactions occurs in a some vessel of fixed volume V

- The content of the container is well mixed
- The system is in thermal equilibrium



The rate constant κ is defined in such a way that (κdt) corresponds to the probability that a randomly chosen molecule of \mathcal{A} is degraded in the infinitesimally small interval

$$[t, t + dt)$$

$\mathcal{A} \rightarrow \emptyset$ (cont.)

Let $N_{\mathcal{A}}(t)$ be the number of molecules of species \mathcal{A} present in the system at time t

Between time t and time $t + dt$, we assume that one of three things can occur

~~ No reaction occurs (no molecules of \mathcal{A} degrade)

$$\rightsquigarrow N_{\mathcal{A}}(t + dt) = N_{\mathcal{A}}(t) - 0$$

~~ One reaction occurs (only one molecule of \mathcal{A} degrades)

$$\rightsquigarrow N_{\mathcal{A}}(t + dt) = N_{\mathcal{A}}(t) - 1$$

~~ Two or more reactions occur (two or more molecules of \mathcal{A} degrade)

$$\rightsquigarrow N_{\mathcal{A}}(t + dt) = N_{\mathcal{A}}(t) - 2 \text{ (or more)}$$

We are interested in the probabilities that these events occur in the interval $[t, t + dt]$

 $\mathcal{A} \rightarrow \emptyset$ (cont.)

We can assume that the molecules of \mathcal{A} are indistinguishable and that act independently

~~ The probability that one reaction occurs in dt is the same for each of them

$$\mathbb{P}(\text{one reaction occurs}) = \mathbb{P}(\text{EITHER molecule } \mathcal{A}_1 \text{ OR } \dots \text{ OR } \mathcal{A}_{N_{\mathcal{A}}(t)} \text{ reacts}) \quad (1a)$$

$$= \mathbb{P}(\mathcal{A}_1 \text{ reacts} \cup \dots \cup \mathcal{A}_{N_{\mathcal{A}}(t)} \text{ reacts}) \quad (1b)$$

$$= \mathbb{P}(\mathcal{A}_1 \text{ reacts}) + \dots + \mathbb{P}(\mathcal{A}_{N_{\mathcal{A}}(t)} \text{ reacts}) \quad (1c)$$

$$= \mathbb{P}(\text{a molecule of } \mathcal{A} \text{ reacts}) \times N_{\mathcal{A}}(t) \quad (1d)$$

$$= kdt \times N_{\mathcal{A}}(t) \quad (1e)$$

(We implicitly assumed that reactions of individual molecules are mutually exclusive)

If the time interval dt is small enough, then the probability that two or more reactions occur is very small (much smaller than that of the other events) and it can be neglected

$$\mathbb{P}(\text{two or more reactions occur}) = 0 \quad (2)$$

The probability that no reactions occur in $[t, t + dt]$ equals the remaining probability

$$\mathbb{P}(\text{no reactions occur}) = 1 - [N_{\mathcal{A}}(t)\kappa dt + 0] \quad (3)$$

 $\mathcal{A} \rightarrow \emptyset$ (cont.)

Suppose that at initial time $t = 0$ there are exactly $n_{\mathcal{A}}(0)$ molecules of \mathcal{A} in the system

$$N_{\mathcal{A}}(t = 0) = n_{\mathcal{A}}(0)$$

We are interested in the number of molecules in the system, as time $t > 0$ progresses

Because reaction events occur randomly with certain probabilities, now we do not know when they will occur and we cannot determine precisely how the system will evolve

Let us denote with $(N_{\mathcal{A}}(t))_{t \geq 0}$ the evolution in time of the copy-numbers of species \mathcal{A}

$$(N_{\mathcal{A}}(t))_{t \geq 0}$$

Each $N_{\mathcal{A}}(t)$ is a **random variable**, thus we do not know its value nor when it will change

- We know that they do not vary continuously with time
- (Copy numbers can only take on integer values)

The collection $(N_{\mathcal{A}}(t))_{t \geq 0}$ of random variables $N_{\mathcal{A}}(t)$ indexed by t is a **random process**

 $\mathcal{A} \rightarrow \emptyset$ (cont.)

We can use the probabilistic definition of reaction events to design a procedure that simulates a **realisation** $n_{\mathcal{A}}(t)$ of the number $N_{\mathcal{A}}(t)$ of molecules of \mathcal{A} , over times $t > 0$

Given a large number of realisations of the system's evolution, we do statistics with it

That is,

- ~~ Firstly, simulate the reaction system to replicate the random events over time
- ~~ Secondly, calculate the average behaviour and variability among replicates

The simulations of the system are probabilistic, thus no two simulations will be identical

$\mathcal{A} \rightarrow \emptyset$ (cont.)

We choose a small interval Δt and evaluate the number of molecules $N_{\mathcal{A}}(t)$, recursively

More explicitly, we define a partitioning of the time axis

$$t \in \{t_k = k\Delta t\}_{k=1,2,\dots,K}$$

That is, graphically

$$0 \cdots t_1 \cdots \underbrace{t_{k-1} \cdots t_k}_{\Delta t} \underbrace{t_k \cdots t_{k+1}}_{\Delta t} \cdots t_{K-1} \cdots t_K$$

We want to get a realisation of the process $(N_{\mathcal{A}}(t))_{t \geq 0}$, at those discrete points in time

$$(N_{\mathcal{A}}(t))_{t \geq 0} \rightsquigarrow (N_{\mathcal{A}}(t_k))_{k=1}^K$$

That is,

$$N_{\mathcal{A}}(0) \cdots N_{\mathcal{A}}(t_1) \cdots \cdots N_{\mathcal{A}}(t_{k-1}) \cdots N_{\mathcal{A}}(t_k) \cdots N_{\mathcal{A}}(t_{k+1}) \cdots \cdots N_{\mathcal{A}}(t_{K-1}) \cdots N_{\mathcal{A}}(t_K)$$

 $\mathcal{A} \rightarrow \emptyset$ (cont.)

Suppose that we know the precise number of molecules of \mathcal{A} at the initial time $t = 0$

We can start a realisation of the process from $n_{\mathcal{A}}(0)$ molecules of \mathcal{A} at time $t = 0\Delta t$

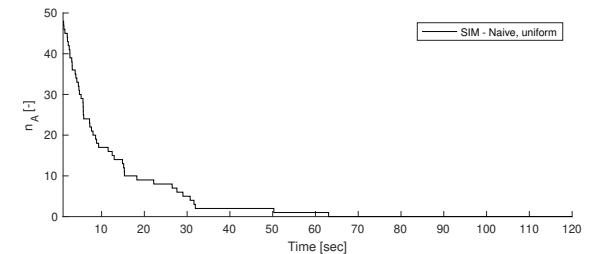
$$n_{\mathcal{A}}(0) \cdots N_{\mathcal{A}}(t_1) \cdots \cdots N_{\mathcal{A}}(t_k) \cdots \cdots N_{\mathcal{A}}(t_{K-1}) \cdots N_{\mathcal{A}}(t_K)$$

To move forward to $t = 1\Delta t$, we would need to evaluate (fix) the copy-number $N_{\mathcal{A}}(t_1)$

$$n_{\mathcal{A}}(0) \cdots n_{\mathcal{A}}(t_1) \cdots \cdots N_{\mathcal{A}}(t_k) \cdots \cdots N_{\mathcal{A}}(t_{K-1}) \cdots N_{\mathcal{A}}(t_K)$$

We firstly need to determine whether a reaction event occurred between time 0 and t_1

~ Then, we change the copy-number from $n_{\mathcal{A}}(0)$ to $n_{\mathcal{A}}(t_1)$, accordingly

 $\mathcal{A} \rightarrow \emptyset$ (cont.)

$$0 \cdots t_1 \cdots \underbrace{t_{k-1} \cdots t_k}_{\Delta t} \underbrace{t_k \cdots t_{k+1}}_{\Delta t} \cdots t_{K-1} \cdots t_K$$

At times t_k with $k = 1, 2, \dots, K$, we firstly evaluate whether a reaction event occurs

- ~ For degradation systems, the probability that a reaction occurs is $n_{\mathcal{A}}(t_k)\kappa\Delta t$
- (This probability depends on the abundance of \mathcal{A} and on the rate constant)
- (This probability is approximated well, so long as Δt is small enough)

Once we established whether a reaction has occurred, we update the molecule count

$$n_{\mathcal{A}}(t_k) \rightsquigarrow n_{\mathcal{A}}(t_{k+1})$$

To move forward, repeat the procedure from $n_{\mathcal{A}}(t_{k+1})$ and probability $n_{\mathcal{A}}(t_{k+1})\kappa\Delta t$

$$\underbrace{n_{\mathcal{A}}(0) \cdots n_{\mathcal{A}}(t_1)}_{\Delta n_{\mathcal{A}}(0)} \cdots \underbrace{n_{\mathcal{A}}(t_{k-1}) \cdots n_{\mathcal{A}}(t_k)}_{\Delta n_{\mathcal{A}}(t_{k-1})} \underbrace{\cdots N_{\mathcal{A}}(t_{k+1}) \cdots N_{\mathcal{A}}(t_{K-1}) \cdots N_{\mathcal{A}}(t_K)}_{\Delta N_{\mathcal{A}}(t_{K-1})}$$

 $\mathcal{A} \rightarrow \emptyset$ (cont.)

Each simulation of the system generates a **sample path**, a **realisation** of the process

- The procedure can be repeated multiple (many) times
- No two realisations of the process will be equal

The objective is to generate a large number R of sample paths $\{(n_{\mathcal{A}}^{(r)}(t_k))_{t_k=1}^K\}_{r=1}^R$

~ We use them to compute statistics

~ To get insight about the process

$\mathcal{A} \rightarrow \emptyset$
 Simulations
 Master equation
 Statistics

$$\mathcal{A} \rightarrow \emptyset$$

Stochastic simulation

$\mathcal{A} \rightarrow \emptyset$
 Simulations
 Master equation
 Statistics

$\mathcal{A} \rightarrow \emptyset | \text{Simulation}$

$$\underbrace{n_{\mathcal{A}}(0) \dots n_{\mathcal{A}}(t_1) \dots \dots}_{\Delta n_{\mathcal{A}}(0)} \underbrace{\dots n_{\mathcal{A}}(t_{k-1}) \dots n_{\mathcal{A}}(t_k) \dots}_{\Delta n_{\mathcal{A}}(t_{k-1})} \underbrace{N_{\mathcal{A}}(t_{k+1}) \dots N_{\mathcal{A}}(t_{K-1}) \dots N_{\mathcal{A}}(t_K)}_{\Delta N_{\mathcal{A}}(t_k)} \underbrace{\dots}_{\Delta N_{\mathcal{A}}(t_{K-1})}$$

Realisations are obtained using computer routines able to generate random numbers

- This is needed to decide whether a reaction occurs or not, but how?

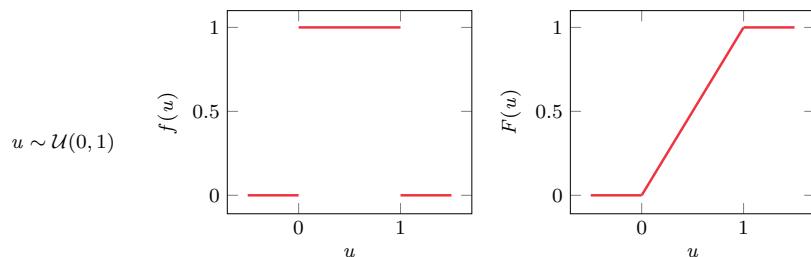
Easy to generate a random number u uniformly distributed in the unit interval $[0, 1]$

~ Number u can then be used to determine whether a reaction occurs or not

$\mathcal{A} \rightarrow \emptyset | \text{Simulation (cont.)}$

Definition (Uniform distribution)

$\mathcal{A} \rightarrow \emptyset$
 Simulations
 Master equation
 Statistics



The probability density function

$$f(u) = \begin{cases} 1, & u \in [0, 1] \\ 0, & \text{elsewhere} \end{cases}$$

The cumulative distribution function

$$F(u) = \begin{cases} 0, & u < 0 \\ u, & u \in [0, 1] \\ 1, & u > 1 \end{cases}$$

$\mathcal{A} \rightarrow \emptyset | \text{Simulation (cont.)}$

$\mathcal{A} \rightarrow \emptyset$
 Simulations
 Master equation
 Statistics

For any $a, b \in [0, 1]$ such that $b > a$, we can compute the probability that $u \in (a, b)$

$$\begin{aligned} \mathbb{P}(a \leq U \leq b) &= \mathbb{P}(U \leq b) - \mathbb{P}(U \leq a) \\ &= \int_{u=0}^b f(u)du - \int_{u=0}^a f(u)du \\ &= b - a \end{aligned}$$

We can use this result to determine whether a reaction event occurs or not

$\mathcal{A} \rightarrow \emptyset$ | Simulation (cont.)

We have defined the probability that a reaction occurs in the interval $[t_k, t_k + \Delta t)$

$$n_{\mathcal{A}}(t_k) \kappa \Delta t$$

Let $a = 0$ and $b = n_{\mathcal{A}}(t_k) \kappa \Delta t$, we have

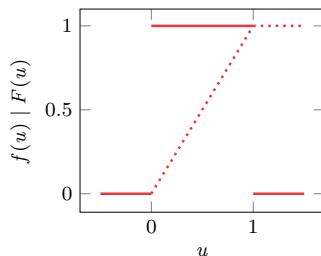
$$\mathbb{P}(0 \leq U \leq n_{\mathcal{A}}(t_k) \kappa \Delta t) = n_{\mathcal{A}}(t_k) \kappa \Delta t - 0$$

- If we draw a $u \in [0, n_{\mathcal{A}}(t_k) \kappa \Delta t]$, then we say that the reaction does occur

$$n_{\mathcal{A}}(t_{k+1} = t_k + \Delta t) = n_{\mathcal{A}}(t_k) - 1$$

- If we draw a $u \in [n_{\mathcal{A}}(t_k) \kappa \Delta t, 1]$, then we say that the reaction does not occur

$$n_{\mathcal{A}}(t_{k+1} = t_k + \Delta t) = n_{\mathcal{A}}(t_k)$$

 $\mathcal{A} \rightarrow \emptyset$ | Simulation (cont.)

Drawing a number $u \sim \mathcal{U}(0, 1)$ and checking whether it is smaller than $n_{\mathcal{A}}(t_k) \kappa \Delta t$ correctly implements the definition of reaction event, provided that Δt is very small

- To obtain valid results, we must ensure that Δt is small
- More precisely, it is required that $n_{\mathcal{A}}(t_k) \kappa \Delta t \ll 1$

 $A \rightarrow \emptyset$ | Simulation (cont.)

Starting at time $t = 0$ with $N_A(0) = n_A(0)$ molecules of A , we can simulate the system

Algorithm 1 Degradation | Stochastic simulation algorithm, naive (uniform)

```

1: procedure DEGRADATION | SSA (NAIVE FORMULATION, UNIFORM)
  Input:  $N_A(t = 0) = n_A(0)$ ,  $\kappa$ , and  $\Delta t$ 
  Output:  $(n_A(t_k))_{k=0,1,\dots}$ 
2:   Set  $k = 0$ 
3:   Set  $t_k = k \Delta t$ 
4:   Generate a random number  $u \sim \mathcal{U}(0, 1)$ 
5:   if  $u < n_A(t_k) \kappa \Delta t$  then
6:      $n_A(t_{k+1}) = n_A(t_k) - 1$ 
7:   else
8:      $n_A(t_{k+1}) = n_A(t_k)$ 
9:   end if
10:  Set  $k \rightsquigarrow k + 1$  and repeat
11: end procedure

```

 $A \rightarrow \emptyset$ | Simulation (cont.)

Example (Degradation - Stochastic simulation, naive)

An evolution of $(N_A(t))$ is obtained by stochastic simulation for given parameter values

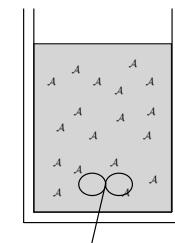


Kinetic parameters and initial condition

$$\begin{aligned} \kappa &= 0.1 \text{ sec}^{-1} \\ N_A(0) &= n_A(0) \\ &= 20 \text{ molecules} \end{aligned}$$

Simulation parameters

$$\Delta t = 0.001 \text{ sec}$$

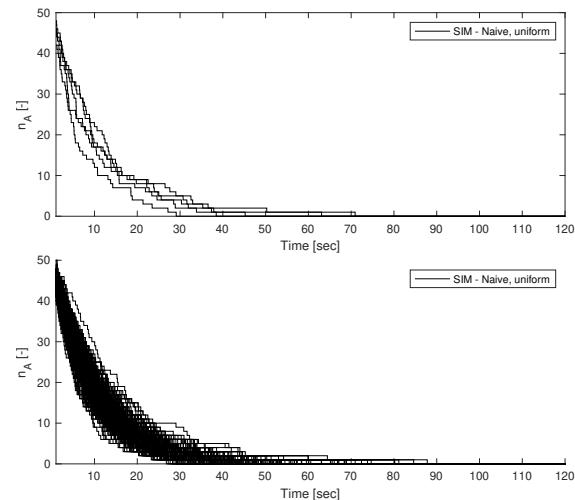


Each time the simulation is repeated, a different realisation of $(N_A(t))_{t>0}$ is obtained

$$\{n_A^{(r)}(t_k)\}_{k=1}^K \quad (r = 1, 2, \dots, R)$$

$\mathcal{A} \rightarrow \emptyset$ | Simulation (cont.)

$\mathcal{A} \rightarrow \emptyset$
Simulations
Master equation
Statistics

 $\mathcal{A} \rightarrow \emptyset$ | Simulation (cont.)

$\mathcal{A} \rightarrow \emptyset$
Simulations
Master equation
Statistics

Because $N_{\mathcal{A}}(0) = 20$, $\kappa = 0.1\text{sec}^{-1}$, and $\Delta t = 0.001\text{sec}$, we always have $n_{\mathcal{A}}(t)\kappa\Delta t \ll 1$

The worst case scenario occurs at the initial time

$$\underbrace{n_{\mathcal{A}}(t)}_{20} \underbrace{\kappa}_{0.1} \underbrace{\Delta t}_{0.001} = 0.002 \ll 0$$

The largest probability that a single reaction event occurs is at $t = 0$, when $n_{\mathcal{A}} = 20$

■

 $\mathcal{A} \rightarrow \emptyset$ | Simulation (cont.)

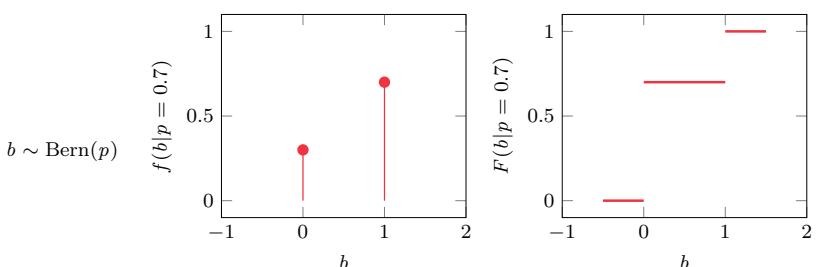
$\mathcal{A} \rightarrow \emptyset$
Simulations
Master equation
Statistics

Instead of uniform numbers u , we can generate a number b with a Bernoulli distribution
The probability of success is $p \in (0, 1)$ is equal to the probability that a reaction occurs

$$p = n_{\mathcal{A}}(t_k)\kappa\Delta t$$

 $\mathcal{A} \rightarrow \emptyset$ | Simulation (cont.)

Definition (Bernoulli distribution)



The probability mass function

$$f(b | p) = \begin{cases} p, & b = 1 \\ q = 1 - p, & b = 0 \end{cases}$$

The cumulative distribution function

$$F(b | p) = \begin{cases} 0, & b < 0 \\ 1 - p & b = 1 \\ 1, & b > 0 \end{cases}$$

■

$\mathcal{A} \rightarrow \emptyset$ | Simulation (cont.)

Starting at time $t = 0$ with $N_{\mathcal{A}}(0) = n_{\mathcal{A}}(0)$ molecules of \mathcal{A} , we can simulate the system

Algorithm 2 Degradation | Stochastic simulation algorithm, naive (Bernoulli)

```

1: procedure DEGRADATION | SSA (NAIVE FORMULATION, BERNOUILLI)
   Input:  $N_{\mathcal{A}}(t = 0) = n_{\mathcal{A}}(0)$ ,  $\kappa$ , and  $\Delta t$ 
   Output:  $\{n_{\mathcal{A}}(t_k)\}_{k=0,1,\dots}$ 
2:   Set  $k = 0$ 
3:   Set  $t_k = k\Delta t$ 
4:   Generate a random number  $b \sim \text{Bern}(n_{\mathcal{A}}(t_k)\kappa\Delta t)$ 
5:   if  $b = 1$  then
6:      $n_{\mathcal{A}}(t_{k+1}) = n_{\mathcal{A}}(t_k) - 1$ 
7:   else
8:      $n_{\mathcal{A}}(t_{k+1}) = n_{\mathcal{A}}(t_k)$ 
9:   end if
10:  Set  $k \rightsquigarrow k + 1$  and repeat
11: end procedure
```

 $\mathcal{A} \rightarrow \emptyset$ | Simulation (cont.)**Example (Degradation - Stochastic simulation, naive)**

An evolution of $(N_{\mathcal{A}}(t))$ is obtained by stochastic simulation for given parameter values



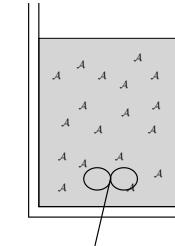
Kinetic parameters and initial condition

$$\kappa = 0.1 \text{ sec}^{-1}$$

$$\begin{aligned} N_{\mathcal{A}}(0) &= n_{\mathcal{A}}(0) \\ &= 20 \text{ molecules} \end{aligned}$$

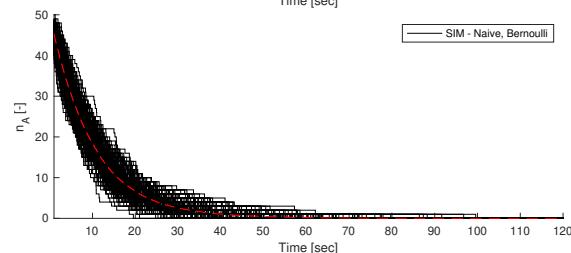
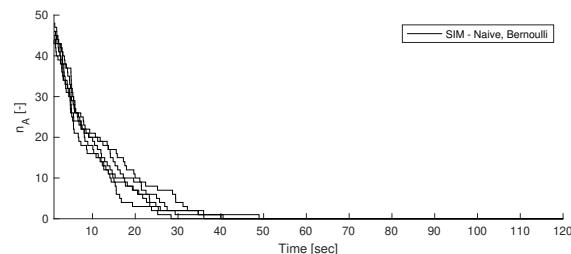
Simulation parameters

$$\Delta t = 0.001 \text{ sec}$$



Each time the simulation is repeated, a different realisation of $(N_{\mathcal{A}}(t))_{t>0}$ is obtained

$$\{n_{\mathcal{A}}^{(r)}(t_k)\}_{k=1}^K \quad (r = 1, 2, \dots, R)$$

 $\mathcal{A} \rightarrow \emptyset$ | Simulation (cont.) $\mathcal{A} \rightarrow \emptyset$ | Simulation, by jumps

To avoid risk of multiple events, the simulation accuracy is increased by decreasing Δt

$$0 \cdots t_1 \cdots \underbrace{t_{k-1} \cdots t_k}_{\Delta t} \cdots \underbrace{t_{k+1} \cdots t_{K-1}}_{\Delta t} \cdots t_K$$

~ Unfortunately, this increases the computational requirements of the simulation

- At each time node in $\{k\Delta t\}_{k=1,2,\dots}$, we draw a random number and check

However, at most time nodes in $\{k\Delta t\}_{k=1,2,\dots}$ no reaction event is ever found to occur

$\mathcal{A} \rightarrow \emptyset$ | Simulation, by jumps

$\mathcal{A} \rightarrow \emptyset$
Simulations
Master equation
Statistics

To improve efficiency, we can approach the simulation task from a different viewpoint
We could proceed differently and ask at what time the next reaction event will occur

- If the time now is t , we are interested in the time $t + s$ of the next reaction

We do not know how to determine $t + s$ precisely, we can model it as a random variable

- That is, we would need to determine its distribution

 $\mathcal{A} \rightarrow \emptyset$ | Simulation, by jumps (cont.)

Let $f(s | N_{\mathcal{A}}(t)) ds$ be the probability that the next reaction occurs in $[t, t + s + ds)$

- Given that we are at time t and that there are $N_{\mathcal{A}}(t)$ copies of \mathcal{A}

$$\cdots t \underbrace{\cdots t + s}_{f(s | N_{\mathcal{A}}(t)) ds} \cdots t + s + ds \cdots$$

For a reaction to occur in $[t + s, t + s + ds)$, no reaction must have occurred in $[t, t + s)$

We let $G(s | N_{\mathcal{A}}(t))$ be the probability that no reaction occurred in $[t, t + s)$

- Again, given that we are at time t and there are $N_{\mathcal{A}}(t)$ copies of \mathcal{A}

$$\cdots t \underbrace{\cdots t + s}_{G(s | N_{\mathcal{A}}(t))} \cdots t + s + ds \cdots$$

We know the probability that one reaction occurs in $[t + s, t + s + ds)$, given $N_{\mathcal{A}}(t + s)$

$$\cdots t \cdots t + s \underbrace{\cdots t + s + ds}_{N_{\mathcal{A}}(t+s)\kappa ds} \cdots$$

 $\mathcal{A} \rightarrow \emptyset$ | Simulation, by jumps (cont.)

$\mathcal{A} \rightarrow \emptyset$
Simulations
Master equation
Statistics

$$\cdots t \underbrace{\cdots t + s}_{G(s | N_{\mathcal{A}}(t))} \underbrace{\cdots t + s + ds}_{N_{\mathcal{A}}(t+s)\kappa ds} \cdots$$

$$f(s | N_{\mathcal{A}}(t)) ds$$

Probability $f(s | N_{\mathcal{A}}(t)) ds$ is computed as product of $G(s | N_{\mathcal{A}}(t))$ and $N_{\mathcal{A}}(t + s)\kappa ds$

- No reactions in $[t, t + s)$ AND one reaction in $[t + s, t + s + ds)$
- The events must jointly occur and not affect each other

That is,

$$f(s | N_{\mathcal{A}}(t)) ds = G(s | N_{\mathcal{A}}(t)) \times N_{\mathcal{A}}(t + s)\kappa ds$$

Because no reactions have occurred in $[t, t + s)$, we have that $N_{\mathcal{A}}(t + s) = N_{\mathcal{A}}(t)$

$$f(s | N_{\mathcal{A}}(t)) ds = G(s | N_{\mathcal{A}}(t)) \times N_{\mathcal{A}}(t)\kappa ds$$

We need to determine the probability $G(s | N_{\mathcal{A}}(t))$ of no reactions in $[t, t + s)$

 $\mathcal{A} \rightarrow \emptyset$ | Simulation, by jumps (cont.)

$\mathcal{A} \rightarrow \emptyset$
Simulations
Master equation
Statistics

$$f(s | N_{\mathcal{A}}(t)) ds = \underbrace{G(s | N_{\mathcal{A}}(t))}_{\text{To be determined}} \underbrace{N_{\mathcal{A}}(t)\kappa ds}_{\text{known}}$$

We first consider the probability that no reaction occurs in some interval $[t, t + \sigma + \Delta\sigma)$

$$\cdots t \cdots t + \sigma \underbrace{\cdots t + \sigma + \Delta\sigma}_{G(\sigma + \Delta\sigma | N_{\mathcal{A}}(t))} \cdots$$

It is given by the product of the probabilities of two events that must jointly occur

- The probability that no reaction occurs in $[t, t + \sigma)$
- The probability that no reaction occurs in $[t + \sigma, t + \sigma + \Delta\sigma)$

As this is true for any arbitrary time $\sigma > 0$ (say, $\sigma = s$), for some small $\Delta\sigma$ we have

$$\cdots t \underbrace{\cdots t + \sigma}_{G(\sigma | N_{\mathcal{A}}(t))} \underbrace{\cdots t + \sigma + \Delta\sigma}_{1 - N_{\mathcal{A}}(\sigma + \Delta\sigma)\kappa \Delta\sigma} \cdots$$

$$G(\sigma + \Delta\sigma | N_{\mathcal{A}}(t))$$

For any value of $N_{\mathcal{A}}(t)$, we can write

$$G(\sigma + \Delta\sigma | N_{\mathcal{A}}(t)) = G(\sigma | N_{\mathcal{A}}(t)) \times [1 - N_{\mathcal{A}}(\sigma + \Delta\sigma)\kappa \Delta\sigma]$$

$\mathcal{A} \rightarrow \emptyset$ | Simulation, by jumps (cont.)

$$G(\sigma + \Delta\sigma \mid N_{\mathcal{A}}(t)) = G(\sigma \mid N_{\mathcal{A}}(t)) [1 - N_{\mathcal{A}}(t + \sigma) \kappa \Delta\sigma]$$

As no reaction occurs during the interval $t, t + \sigma$, we have that $N_{\mathcal{A}}(t + \sigma) = N_{\mathcal{A}}(t)$

After substituting and rearranging, we can thus write

$$G(\sigma + \Delta\sigma \mid N_{\mathcal{A}}(t)) = G(\sigma \mid N_{\mathcal{A}}(t)) - G(\sigma \mid N_{\mathcal{A}}(t)) N_{\mathcal{A}}(t) \kappa \Delta\sigma$$

Similarly, we can write the change in probability

$$G(\sigma + \Delta\sigma \mid N_{\mathcal{A}}(t)) - G(\sigma \mid N_{\mathcal{A}}(t)) = -G(\sigma \mid N_{\mathcal{A}}(t)) N_{\mathcal{A}}(t) \kappa \Delta\sigma$$

Rearranging again, we get

$$\frac{\underbrace{G(\sigma + \Delta\sigma \mid N_{\mathcal{A}}(t)) - G(\sigma \mid N_{\mathcal{A}}(t))}_{\Delta G(\sigma \mid N_{\mathcal{A}}(t))}}{\Delta\sigma} = -N_{\mathcal{A}}(t) G(\sigma \mid N_{\mathcal{A}}(t)) \kappa$$

In the limit for $\Delta\sigma \rightarrow 0$,

$$\frac{dG(\sigma \mid N_{\mathcal{A}}(t))}{d\sigma} = -N_{\mathcal{A}}(t) G(\sigma \mid N_{\mathcal{A}}(t)) \kappa$$

 $\mathcal{A} \rightarrow \emptyset$ | Simulation (cont.)

$$\frac{dG(\sigma \mid N_{\mathcal{A}}(t))}{d\sigma} = -N_{\mathcal{A}}(t) G(\sigma \mid N_{\mathcal{A}}(t)) \kappa$$

The probability of no reaction in $[t, t + s]$, given $N_{\mathcal{A}}(t)$ molecules at t , changes in time

Because $N_{\mathcal{A}}(t)$ is fixed, we can re-write the differential relation more clearly

$$\frac{dG(\sigma \mid N_{\mathcal{A}}(t))}{d\sigma} = -N_{\mathcal{A}}(t) \kappa G(\sigma \mid N_{\mathcal{A}}(t))$$

This is a common ordinary differential equation of the general form

$$\frac{dy(t)}{dx} = \lambda y(t) \quad (y(x=0) = y_0)$$

The solution is the exponential decay or growth, depending on λ

$$y(t) = y_0 \exp(\lambda x)$$

 $\mathcal{A} \rightarrow \emptyset$ | Simulation, by jumps (cont.)

$$\frac{dG(\sigma \mid N_{\mathcal{A}}(t))}{d\sigma} = -N_{\mathcal{A}}(t) \kappa G(\sigma \mid N_{\mathcal{A}}(t))$$

We solve the ordinary differential equation from initial condition $G(\sigma = 0 \mid N_{\mathcal{A}}(t)) = 1$

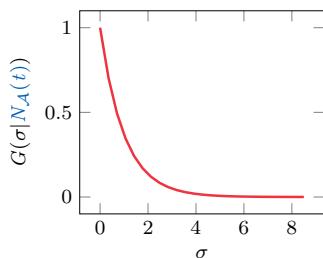
- That is, the probability that no reaction occurs in zero time is one
- This is true whatever the copy-number $N_{\mathcal{A}}(t)$ and time t

$$\rightsquigarrow G(\sigma \mid N_{\mathcal{A}}(t)) = \exp(-N_{\mathcal{A}}(t) \kappa \sigma)$$

The solution is the probability that no reaction occurs in the interval $[t, t + \sigma]$

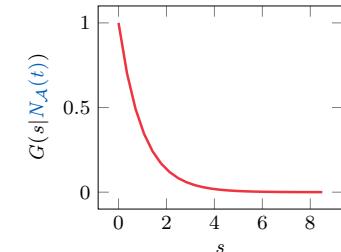
\rightsquigarrow Given $N_{\mathcal{A}}(t)$ molecules of \mathcal{A} at t

It is a function of the random time σ

 $\mathcal{A} \rightarrow \emptyset$ | Simulation, by jumps (cont.)

$$\rightsquigarrow G(s \mid N_{\mathcal{A}}(t)) = \exp(-N_{\mathcal{A}}(t) \kappa s)$$

$$\begin{aligned} f(s \mid N_{\mathcal{A}}(t)) ds &= G(s \mid N_{\mathcal{A}}(t)) N_{\mathcal{A}}(t) \kappa \\ &= \exp(-N_{\mathcal{A}}(t) \kappa s) N_{\mathcal{A}}(t) \kappa ds \end{aligned}$$



After substituting, we get the probability (density) of the time s to the next reaction

$$f(s \mid N_{\mathcal{A}}(t)) = N_{\mathcal{A}}(t) \kappa \exp(-N_{\mathcal{A}}(t) \kappa s)$$

Again, given that there are $N_{\mathcal{A}}(t)$ molecules of \mathcal{A} at time t

$\mathcal{A} \rightarrow \emptyset$ | Simulation, by jumps (cont.)

$\mathcal{A} \rightarrow \emptyset$
Simulations
Master equation
Statistics

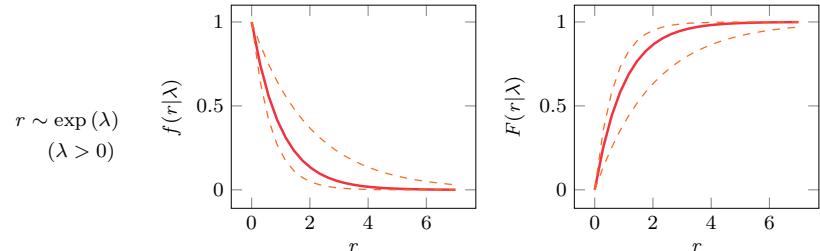
$$f(s | N_{\mathcal{A}}(t)) = \underbrace{\frac{N_{\mathcal{A}}(t) \kappa}{\lambda}}_{\lambda} \exp(-\underbrace{\frac{N_{\mathcal{A}}(t) \kappa}{\lambda} s}_{\lambda \exp(-\lambda s)})$$

This probability density function tells us that the time s until the next reaction is a random number s with an exponential distribution whose parameter is $(N_{\mathcal{A}}(t)\kappa) > 0$

- The larger the parameter (the abundance of \mathcal{A} , the faster the decay)
- Shorter times s are given more probability, thus are more likely

 $\mathcal{A} \rightarrow \emptyset$ | Simulation, by jumps (cont.)

Definition (Exponential distribution)



The probability density function

$$f(r | \lambda) = \begin{cases} \lambda \exp(-\lambda r) & r \geq 0 \\ 0 & r < 0 \end{cases}$$

The cumulative distribution function

$$F(r | \lambda) = \begin{cases} 1 - \exp(-\lambda r) & r \geq 0 \\ 0 & r < 0 \end{cases}$$

■

 $\mathcal{A} \rightarrow \emptyset$ | Simulation, by jumps (cont.)

$\mathcal{A} \rightarrow \emptyset$
Simulations
Master equation
Statistics

The exponential distribution describes the time between events in a ‘Poisson process’

- (A Poisson process models a sequence of events occurring at random times)
- (Poisson processes are continuous-time discrete-state processes)

The most important distribution in the theory of discrete-event stochastic simulation

The expected value of R (and its standard deviation),

$$\begin{aligned} E[R] &= \int_0^\infty r f(r | \lambda) dr \\ &= \int_0^\infty r \lambda \exp(-\lambda r) dr \\ &= \lambda^{-1} \end{aligned}$$

The memoryless property,

$$\mathbb{P}(R > (t + s) | R > t) = \mathbb{P}(X > s) \quad \text{for all } t, s > 0$$

 $\mathcal{A} \rightarrow \emptyset$ | Simulation, by jumps (cont.)

Starting at time $t = 0$ with $N_{\mathcal{A}}(0) = n_{\mathcal{A}}(0)$ molecules of \mathcal{A} , we can simulate the system

Algorithm 3 Degradation | Stochastic simulation algorithm

```

1: procedure DEGRADATION | SSA, EXPONENTIAL
  Input:  $N_{\mathcal{A}}(t = 0) = n_{\mathcal{A}}(0)$  and  $\kappa$ 
  Output:  $(n_{\mathcal{A}}(t + s_z))_{z=1,2,\dots}$ 
2:   Set  $t = 0$ ,  $z = 0$  and  $s_z = 0$ 
3:   Generate a random number for the time  $s_{z+1}$  until next reaction
    $s_{z+1} \sim \text{Exp}(n_{\mathcal{A}}(t + s_z)\kappa)$ 
4:   Set  $z \rightsquigarrow z + 1$ 
5:   Set  $t \rightsquigarrow t + s_{z+1}$ 
6:   Set  $n_{\mathcal{A}}(t + s_{z+1}) = n_{\mathcal{A}}(t + s_z) + 1$ 
7:   Repeat
8: end procedure

```

$\mathcal{A} \rightarrow \emptyset$ | Simulation, by jumps (cont.)

$\mathcal{A} \rightarrow \emptyset$
Simulations
Master equation
Statistics

Example (Degradation - SSA, exponential)

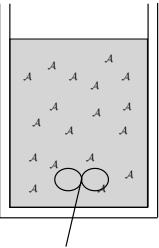
An evolution of ($N_{\mathcal{A}}(t)$) is obtained by stochastic simulation for given parameter values



Kinetic parameters and initial conditions

$$\kappa = 0.1 \text{ sec}^{-1}$$

$$\begin{aligned} N_{\mathcal{A}}(0) &= n_{\mathcal{A}}(0) \\ &= 20 \end{aligned}$$

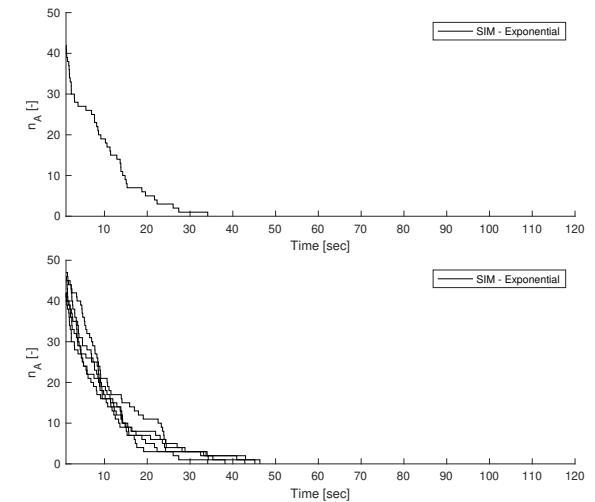


Each time the simulation is repeated, a different realisation of ($N_{\mathcal{A}}(t)$) $_{t>0}$ is obtained

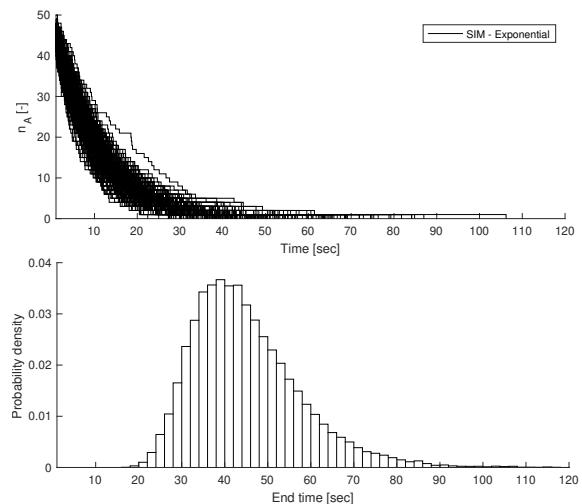
$$\{n_{\mathcal{A}}^{(r)}(t_k)\}_{k=1}^K \quad (r = 1, 2, \dots, R)$$

 $\mathcal{A} \rightarrow \emptyset$ | Simulation, by jumps (cont.)

$\mathcal{A} \rightarrow \emptyset$
Simulations
Master equation
Statistics

 $\mathcal{A} \rightarrow \emptyset$ | Simulation, by jumps (cont.)

$\mathcal{A} \rightarrow \emptyset$
Simulations
Master equation
Statistics

 $\mathcal{A} \rightarrow \emptyset$ | Simulation, by jumps (cont.)

Starting at time $t = 0$ with $N_{\mathcal{A}}(0) = n_{\mathcal{A}}(0)$ molecules of \mathcal{A} , we can simulate the system

Algorithm 4 Degradation | Stochastic simulation algorithm

1: **procedure** DEGRADATION | SSA, EXPONENTIAL FROM UNIFORM
Input: $N_{\mathcal{A}}(t = 0) = n_{\mathcal{A}}(0)$ and κ
Output: $(n_{\mathcal{A}}(t + s_z))_{z=1,2,\dots}$

- 2: Set $t = 0$, $z = 0$ and $s_z = 0$
- 3: Generate a random number $u \sim \mathcal{U}(0, 1)$
- 4: Compute the time s_{z+1} until next reaction

$$s_{z+1} = \frac{1}{n_{\mathcal{A}}(t + s_z)\kappa} \ln \left(\frac{1}{u} \right)$$

- 5: Set $z \rightsquigarrow z + 1$
- 6: Set $t \rightsquigarrow t + s_{z+1}$
- 7: Set $n_{\mathcal{A}}(t + s_{z+1}) = n_{\mathcal{A}}(t + s_z) - 1$
- 8: Repeat
- 9: **end procedure**

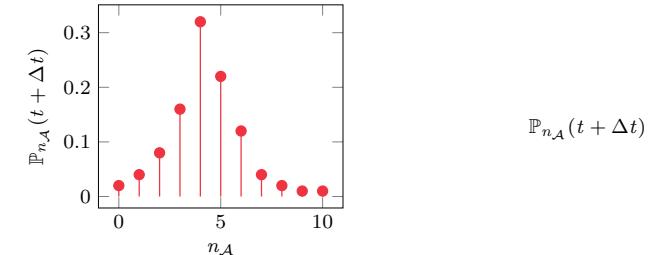
$\mathcal{A} \rightarrow \emptyset$
Simulations
Master equation
Statistics

$\mathcal{A} \rightarrow \emptyset$ Master equation

$\mathcal{A} \rightarrow \emptyset$
Simulations
Master equation
Statistics

$\mathcal{A} \rightarrow \emptyset$ | Master equation

Let $\mathbb{P}_{n_{\mathcal{A}}}(t + \Delta t)$ be the probability that there are $n_{\mathcal{A}}$ molecules of \mathcal{A} at time $(t + \Delta t)$



$\mathcal{A} \rightarrow \emptyset$
Simulations
Master equation
Statistics

$\mathcal{A} \rightarrow \emptyset$ | Master equation (cont.)

When the time interval $[t, t + \Delta t]$ is small, at most one molecule had been degraded
Only one of two events can lead to have $n_{\mathcal{A}}(t + \Delta t)$ molecules of \mathcal{A} at time $t + \Delta t$

~ One reaction occurred in $[t, t + \Delta t]$

$$n_{\mathcal{A}}(t + \Delta t) \text{ was } n_{\mathcal{A}}(t + \Delta t) + 1$$

This occurs with probability $\kappa(n_{\mathcal{A}}(t) + 1)\Delta t$, for each of the $n_{\mathcal{A}}(t) + 1$ copies

~ No reactions occurred in $[t, t + \Delta t]$

$$n_{\mathcal{A}}(t + \Delta t) \text{ was } n_{\mathcal{A}}(t + \Delta t)$$

This occurs with probability $1 - \kappa n_{\mathcal{A}}(t)\Delta t$, for each of the $n_{\mathcal{A}}(t)$ molecules

Reaction events are mutually exclusive (OR) and valid for each molecule of \mathcal{A} (AND)

$\mathcal{A} \rightarrow \emptyset$
Simulations
Master equation
Statistics

$\mathcal{A} \rightarrow \emptyset$ | Master equation (cont.)

The probability of having $n_{\mathcal{A}}$ molecules at time $t + \Delta t$ must encode this information

$$\mathbb{P}_{n_{\mathcal{A}}}(t + \Delta t) = \mathbb{P}_{n_{\mathcal{A}}+1}(t) \underbrace{\times}_{\text{AND}} \underbrace{\kappa(n_{\mathcal{A}}(t) + 1)\Delta t}_{\text{One reaction occurs}}$$

+

$$\underbrace{\mathbb{P}_{n_{\mathcal{A}}}(t) \times [1 - \kappa n_{\mathcal{A}}(t)\Delta t]}_{\text{No reactions occur}}$$

After expanding terms, we obtain

$$\mathbb{P}_{n_{\mathcal{A}}}(t + \Delta t) = \mathbb{P}_{n_{\mathcal{A}}+1}(t) \times \kappa(n_{\mathcal{A}}(t) + 1)\Delta t + \mathbb{P}_{n_{\mathcal{A}}}(t) - \mathbb{P}_{n_{\mathcal{A}}}(t) \times \kappa n_{\mathcal{A}}(t)\Delta t$$

Rearranging and dividing by Δt ,

$$\begin{aligned} \frac{\mathbb{P}_{n_{\mathcal{A}}}(t + \Delta t) - \mathbb{P}_{n_{\mathcal{A}}}(t)}{\Delta t} &= \mathbb{P}_{n_{\mathcal{A}}+1}(t) \times \kappa(n_{\mathcal{A}}(t) + 1) - \mathbb{P}_{n_{\mathcal{A}}}(t) \times (\kappa n_{\mathcal{A}}(t)) \\ &= \kappa(n_{\mathcal{A}}(t) + 1)\mathbb{P}_{n_{\mathcal{A}}+1}(t) - \kappa n_{\mathcal{A}}(t)\mathbb{P}_{n_{\mathcal{A}}}(t) \end{aligned}$$

$\mathcal{A} \rightarrow \emptyset$ | Master equation (cont.)

$\mathcal{A} \rightarrow \emptyset$
Simulations
Master equation
Statistics

$$\frac{\underbrace{\mathbb{P}_{n_A}(t + \Delta t) - \mathbb{P}_{n_A}(t)}_{\Delta \mathbb{P}_{n_A}(t)}}{\Delta t} = \kappa(n_A(t) + 1)\mathbb{P}_{n_A+1}(t) - \kappa n_A(t)\mathbb{P}_{n_A}(t)$$

For each copy-number n_A and at each time t , the rate at which the probability of that number changes depends on an outgoing flow and on an incoming flow of probability

- ~~ The probability of a smaller copy number decreases the rate of change
- ~~ The probability of a larger copy number increases the rate of change

Whenever the two flows are equal, the rate of change in probability is equal to zero

- The system, from the viewpoint of $\mathbb{P}_{n_A}(t)$, is said to be at steady-state

 $\mathcal{A} \rightarrow \emptyset$ | Master equation (cont.)

$\mathcal{A} \rightarrow \emptyset$
Simulations
Master equation
Statistics

$$\frac{\Delta \mathbb{P}_{n_A}(t)}{\Delta t} = \kappa(n_A + 1)\mathbb{P}_{n_A+1}(t) - \kappa n_A(t)\mathbb{P}_{n_A}(t)$$

For $\Delta t \rightarrow 0$, we get the ordinary differential difference equation for copy-number n_A

$$\frac{d\mathbb{P}_{n_A}(t)}{dt} = \kappa(n_A + 1)\mathbb{P}_{n_A+1}(t) - \kappa n_A(t)\mathbb{P}_{n_A}(t)$$

These ordinary differential difference equations are **chemical master equations** (CMEs)

- ~~ In general, the CME is defined for all values $n_A \in \{0, 1, \dots, \infty\}$

For the degradation reaction system, there can be at most $n_A(t=0) = \bar{n}_A$ molecules

- ~~ The corresponding CME is defined for $n_A \in \{0, 1, \dots, \bar{n}_A\}$

 $\mathcal{A} \rightarrow \emptyset$ | Master equation (cont.)

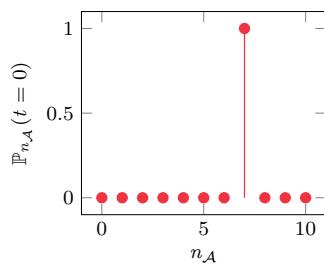
$\mathcal{A} \rightarrow \emptyset$
Simulations
Master equation
Statistics

$$\frac{d\mathbb{P}_{n_A}(t)}{dt} = \kappa(n_A + 1)\mathbb{P}_{n_A+1}(t) - \kappa n_A(t)\mathbb{P}_{n_A}(t) \quad (\text{for } n_A = 0, 1, \dots, \bar{n}_A)$$

Thus, this chemical master equation consists of $\bar{n}_A + 1$ ordinary differential equations

- The initial condition is the set $\{\mathbb{P}_{n_A}(t=0)\}_{n_A=0}^{\bar{n}_A}$ of initial probabilities

For the degradation system, the initial condition can be assumed to be concentrated



$$\mathbb{P}_{n_A}(t=0) = \begin{cases} 1, & \text{for } n_A = \bar{n}_A \\ 0, & \text{elsewhere} \end{cases}$$

(Although, it does not have to be so)

 $\mathcal{A} \rightarrow \emptyset$ | Master equation (cont.)

$\mathcal{A} \rightarrow \emptyset$
Simulations
Master equation
Statistics

$$\frac{d\mathbb{P}_{n_A}(t)}{dt} = \kappa(n_A + 1)\mathbb{P}_{n_A+1}(t) - \kappa n_A(t)\mathbb{P}_{n_A}(t) \quad (\text{with } n_A = 0, 1, \dots, \bar{n}_A)$$

We can inspect the component equations of the CME for the few first values of n_A

For $n_A = 0$, we have

$$\begin{aligned} \frac{d\mathbb{P}_0(t)}{dt} &= \kappa(0 + 1)\mathbb{P}_{0+1}(t) - \kappa(0)\mathbb{P}_0(t) \\ &= \kappa\mathbb{P}_1(t) \end{aligned}$$

For $n_A = 1$, we have

$$\begin{aligned} \frac{d\mathbb{P}_1(t)}{dt} &= \kappa(1 + 1)\mathbb{P}_{1+1}(t) - \kappa(1)\mathbb{P}_1(t) \\ &= 2\kappa\mathbb{P}_2(t) - \kappa\mathbb{P}_1(t) \end{aligned}$$

$\mathcal{A} \rightarrow \emptyset$ | Master equation (cont.)

$\mathcal{A} \rightarrow \emptyset$
Simulations
Master equation
Statistics

$$\frac{d\mathbb{P}_{n_{\mathcal{A}}}(t)}{dt} = \kappa(n_{\mathcal{A}} + 1)\mathbb{P}_{n_{\mathcal{A}}+1}(t) - \kappa n_{\mathcal{A}}\mathbb{P}_{n_{\mathcal{A}}}(t) \quad (\text{with } n_{\mathcal{A}} = 0, 1, \dots, \overline{n_{\mathcal{A}}})$$

For a few last values of $n_{\mathcal{A}}$, we have the following component equations

For $n_{\mathcal{A}} = \overline{n_{\mathcal{A}}}$, we have

$$\begin{aligned} \frac{d\mathbb{P}_{\overline{n_{\mathcal{A}}}}(t)}{dt} &= \kappa(\overline{n_{\mathcal{A}}} + 1)\mathbb{P}_{\overline{n_{\mathcal{A}}}+1}(t) - \kappa(\overline{n_{\mathcal{A}}})\mathbb{P}_{\overline{n_{\mathcal{A}}}}(t) \\ &= -\kappa\overline{n_{\mathcal{A}}}\mathbb{P}_{\overline{n_{\mathcal{A}}}}(t) \end{aligned}$$

For $n_{\mathcal{A}} = \overline{n_{\mathcal{A}}} - 1$, we have

$$\begin{aligned} \frac{d\mathbb{P}_{\overline{n_{\mathcal{A}}}-1}(t)}{dt} &= \kappa(\overline{n_{\mathcal{A}}} - 1 + 1)\mathbb{P}_{\overline{n_{\mathcal{A}}}-1+1}(t) - \kappa(\overline{n_{\mathcal{A}}} - 1)\mathbb{P}_{\overline{n_{\mathcal{A}}}-1}(t) \\ &= \kappa\overline{n_{\mathcal{A}}}\mathbb{P}_{\overline{n_{\mathcal{A}}}}(t) - \kappa(\overline{n_{\mathcal{A}}} - 1)\mathbb{P}_{\overline{n_{\mathcal{A}}}-1}(t) \end{aligned}$$

 $\mathcal{A} \rightarrow \emptyset$ | Master equation (cont.)

$\mathcal{A} \rightarrow \emptyset$
Simulations
Master equation
Statistics

$$\frac{d\mathbb{P}_{n_{\mathcal{A}}}(t)}{dt} = \kappa(n_{\mathcal{A}} + 1)\mathbb{P}_{n_{\mathcal{A}}+1}(t) - \kappa n_{\mathcal{A}}\mathbb{P}_{n_{\mathcal{A}}}(t) \quad (\text{for } n_{\mathcal{A}} = 0, 1, \dots, \overline{n_{\mathcal{A}}})$$

Consider the component equation of the CME corresponding to $n_{\mathcal{A}} = \overline{n_{\mathcal{A}}}$, we have

$$\frac{d\mathbb{P}_{\overline{n_{\mathcal{A}}}}(t)}{dt} = -\kappa\overline{n_{\mathcal{A}}}\mathbb{P}_{\overline{n_{\mathcal{A}}}}(t)$$

We can solve this ordinary differential equation, we have an initial condition

As there are $\overline{n_{\mathcal{A}}}$ molecules of \mathcal{A} at $t = 0$, the associated initial condition

$$\mathbb{P}_{\overline{n_{\mathcal{A}}}}(t = 0) = 1$$

 $\mathcal{A} \rightarrow \emptyset$ | Master equation (cont.)

$\mathcal{A} \rightarrow \emptyset$
Simulations
Master equation
Statistics

$$\frac{d\mathbb{P}_{\overline{n_{\mathcal{A}}}}(t)}{dt} = -\kappa\overline{n_{\mathcal{A}}}\mathbb{P}_{\overline{n_{\mathcal{A}}}}(t)$$

By integrating this component equation over the interval $[0, t]$ from $\mathbb{P}_{\overline{n_{\mathcal{A}}}}(t = 0) = 1$,

$$\begin{aligned} \mathbb{P}_{\overline{n_{\mathcal{A}}}}(t) &= \mathbb{P}_{\overline{n_{\mathcal{A}}}}(t = 0) \exp(-\kappa\overline{n_{\mathcal{A}}}t) \\ &= \exp(-\kappa\overline{n_{\mathcal{A}}}t) \end{aligned}$$

The probability of having $\overline{n_{\mathcal{A}}}$ molecules of \mathcal{A} at time t decays exponentially with t

- The decay rate depends on the kinetics κ and the initial copy-number

 $\mathcal{A} \rightarrow \emptyset$ | Master equation (cont.)

$\mathcal{A} \rightarrow \emptyset$
Simulations
Master equation
Statistics

$$\frac{d\mathbb{P}_{n_{\mathcal{A}}}(t)}{dt} = \kappa(n_{\mathcal{A}} + 1)\mathbb{P}_{n_{\mathcal{A}}+1}(t) - \kappa n_{\mathcal{A}}\mathbb{P}_{n_{\mathcal{A}}}(t) \quad (\text{with } n_{\mathcal{A}} = 0, 1, \dots, \overline{n_{\mathcal{A}}})$$

Now consider the component equation of the CME corresponding to $n_{\mathcal{A}}(t) = \overline{n_{\mathcal{A}}} - 1$,

$$\begin{aligned} \frac{d\mathbb{P}_{\overline{n_{\mathcal{A}}}-1}(t)}{dt} &= \underbrace{\kappa(\overline{n_{\mathcal{A}}} - 1 + 1)}_{\overline{n_{\mathcal{A}}}} \underbrace{\mathbb{P}_{\overline{n_{\mathcal{A}}}-1+1}(t)}_{\mathbb{P}_{\overline{n_{\mathcal{A}}}}(t)} - \kappa(\overline{n_{\mathcal{A}}} - 1)\mathbb{P}_{\overline{n_{\mathcal{A}}}-1}(t) \\ &= \kappa\overline{n_{\mathcal{A}}}\exp(-\kappa\overline{n_{\mathcal{A}}}t) - \kappa(\overline{n_{\mathcal{A}}} - 1)\mathbb{P}_{\overline{n_{\mathcal{A}}}-1}(t) \end{aligned}$$

As there are exactly $\overline{n_{\mathcal{A}}}$ molecules of \mathcal{A} at $t = 0$, we have the associated initial condition

$$\mathbb{P}_{\overline{n_{\mathcal{A}}}-1}(t = 0) = 0$$

Again, we can use this initial condition to solve this individual component equation

$\mathcal{A} \rightarrow \emptyset$ | Master equation (cont.)

$\mathcal{A} \rightarrow \emptyset$
Simulations
Master equation
Statistics

$$\frac{d\mathbb{P}_{\overline{n_A}-1}(t)}{dt} = \underbrace{\kappa \overline{n_A} \exp(-\kappa \overline{n_A} t)}_{\text{constant, at } t} - \kappa(\overline{n_A} - 1)\mathbb{P}_{\overline{n_A}-1}(t)$$

Integrating over the interval $[0, t]$ from the initial condition $\mathbb{P}_{\overline{n_A}-1}(t=0) = 0$, we get

$$\mathbb{P}_{\overline{n_A}-1}(t) = \exp[-\kappa(\overline{n_A} - 1)t] \times \overline{n_A} [1 - \exp(-\kappa t)]$$

 $\mathcal{A} \rightarrow \emptyset$ | Master equation (cont.)

It can be shown by induction that for copy-number n_A , we have the general solution

$$\mathbb{P}_{n_A}(t) = \exp(-\kappa n_A t) \binom{\overline{n_A}}{n_A} [1 - \exp(-\kappa t)]^{\overline{n_A} - n_A}$$

We used the binomial coefficient,

$$\binom{\overline{n_A}}{n_A} = \frac{\overline{n_A}!}{n_A!(\overline{n_A} - n_A)}$$

The probability that there are n_A molecules of \mathcal{A} at t is given by a binomial distribution

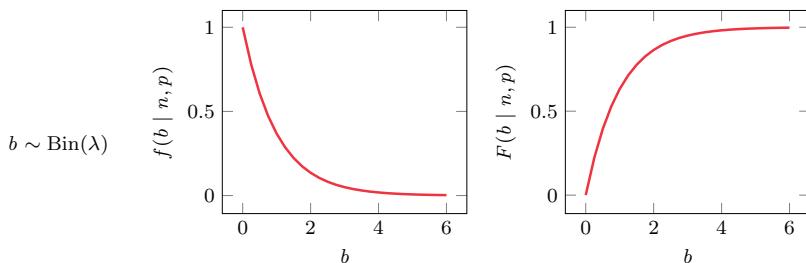
$$\begin{aligned} \mathbb{P}_{n_A}(t) &= \exp(-\kappa n_A t) \binom{\overline{n_A}}{n_A} [1 - \exp(-\kappa t)]^{\overline{n_A} - n_A} \\ &= \underbrace{\exp(-\kappa t)^n}_{p} \underbrace{\binom{\overline{n_A}}{n_A}}_{\binom{n}{k}} \underbrace{\left[1 - \exp(-\kappa t)\right]}_{1-p}^{\overline{n_A} - n_A} \\ &\quad \underbrace{(1-p)^{n-k}}_{(1-p)^{\overline{n_A} - n_A}} \end{aligned}$$

The parameters of the distribution are $p = \exp(-\kappa t)$ and $\overline{n_A}$

 $\mathcal{A} \rightarrow \emptyset$ | Master equation (cont.)

Definition (Binomial distribution)

$\mathcal{A} \rightarrow \emptyset$
Simulations
Master equation
Statistics



The probability mass function

$$f(b | n, p) = \binom{n}{b} p^b (1-p)^{n-b}$$

The cumulative probability function

$$F(b | n, p) = \sum_{l=0}^{\lfloor b \rfloor} \binom{n}{l} p^l (1-p)^{n-l}$$

 $\mathcal{A} \rightarrow \emptyset$ | Master equation (cont.)

Example (Degradation - Master equation)

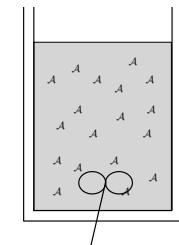
The solution to the chemical master equation



Kinetic parameters and initial condition

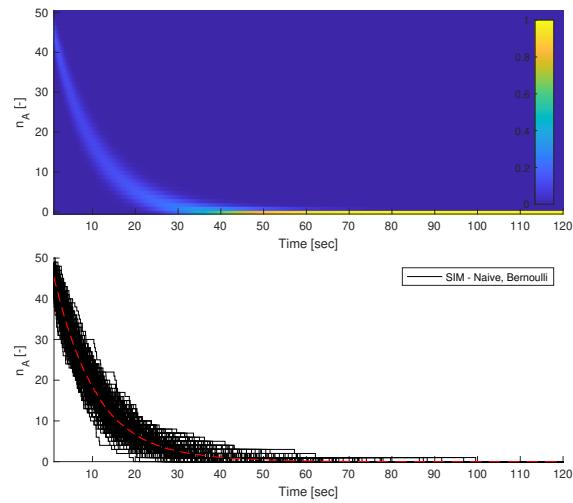
$$\kappa = 0.1 \text{ sec}^{-1}$$

$$\mathbb{P}_{n_A}(0) = \begin{cases} 1, & n_A = 20 \\ 0, & \text{elsewhere} \end{cases}$$



$\mathcal{A} \rightarrow \emptyset$ | Master equation (cont.)

$\mathcal{A} \rightarrow \emptyset$
Simulations
Master equation
Statistics

 $\mathcal{A} \rightarrow \emptyset$ | Master equation (cont.)

$\mathcal{A} \rightarrow \emptyset$
Simulations
Master equation
Statistics

$$\frac{d\mathbb{P}_{n_A}(t)}{dt} = \kappa(n_A + 1)\mathbb{P}_{n_A+1}(t) - \kappa n_A \mathbb{P}_{n_A}(t) \quad (\text{for } n_A = 0, 1, \dots, \overline{n_A})$$

In matrix-vector form,

$$\frac{d}{dt} \begin{bmatrix} \vdots \\ \mathbb{P}_{n_A-1} \\ \mathbb{P}_{n_A} \\ \mathbb{P}_{n_A+1} \\ \vdots \end{bmatrix}_t = \begin{bmatrix} \cdots & 0 & -\kappa(n_A - 1) & \kappa n_A & 0 & \cdots & 0 & \cdots \\ \cdots & \cdots & 0 & -\kappa n_A & \kappa(n_A + 1) & 0 & -\kappa(n_A + 1) & \kappa(n_A + 2) & 0 & \cdots \\ & & & \cdots & 0 & & & & & \cdots \end{bmatrix}_t \begin{bmatrix} \vdots \\ \mathbb{P}_{n_A-1} \\ \mathbb{P}_{n_A} \\ \mathbb{P}_{n_A+1} \\ \vdots \end{bmatrix}_t$$

 $\mathcal{A} \rightarrow \emptyset$ | Master equation (cont.)

$\mathcal{A} \rightarrow \emptyset$
Simulations
Master equation
Statistics

Any attempt to solve the CMEs analytically or even numerically will be unpractical

- CMEs can be solved exactly and explicitly in only a very few cases [1]

As there is a potentially large number of possible states (copy-number of molecules)

For realistic systems, simulation is the approach to gaining insight into the dynamics

Analytic solution for mass-action models with only zeroth and first order reactions [2]

- Analytical solutions are important for testing stochastic simulation algorithms

$\mathcal{A} \rightarrow \emptyset$
Simulations
Master equation
Statistics

$\mathcal{A} \rightarrow \emptyset$
Statistics

¹D.A. McQuarrie (1967). Stochastic approach to chemical kinetics. *Journal of Applied Probability*, 4:413–478.

²T. Jahnke and W. Huisinga (2007). Solving the chemical master equation for monomolecular reactions systems analytically. *Physical Review E*, 80: 066106.

$\mathcal{A} \rightarrow \emptyset$ | Statistics

We use the master equation to determine the expected evolution of process $(N_{\mathcal{A}}(t))_{t \geq 0}$

$$(E[N_{\mathcal{A}}(t)])_{t \geq 0}$$

At a point t in time, we can determine the expected value of the process

- ~ The expected number of molecules of species \mathcal{A} at time t

$$\begin{aligned} E[N_{\mathcal{A}}(t)] &= \sum_{n_{\mathcal{A}}=0}^{\infty} n_{\mathcal{A}} \mathbb{P}_{n_{\mathcal{A}}}(t) \\ &= \sum_{n_{\mathcal{A}}=0}^{\overline{n}_{\mathcal{A}}} n_{\mathcal{A}} \mathbb{P}_{n_{\mathcal{A}}}(t) \\ &= M_{\mathcal{A}}(t) \end{aligned}$$

From the CME, the probability of $n_{\mathcal{A}}$ molecules of \mathcal{A} at t

$$\mathbb{P}_{n_{\mathcal{A}}}(t) = \exp(-\kappa n_{\mathcal{A}} t) \binom{\overline{n}_{\mathcal{A}}}{n_{\mathcal{A}}} (1 - \exp(-\kappa t))^{\overline{n}_{\mathcal{A}} - n_{\mathcal{A}}}$$

 $\mathcal{A} \rightarrow \emptyset$ | Statistics (cont.)

By substituting and rearranging, we get the expected number of molecules at time t

$$\begin{aligned} M_{\mathcal{A}}(t) &= \sum_{n_{\mathcal{A}}=0}^{\overline{n}_{\mathcal{A}}} n_{\mathcal{A}} \underbrace{\exp(-\kappa n_{\mathcal{A}} t) \binom{\overline{n}_{\mathcal{A}}}{n_{\mathcal{A}}} (1 - \exp(-\kappa t))^{\overline{n}_{\mathcal{A}} - n_{\mathcal{A}}}}_{\mathbb{P}_{n_{\mathcal{A}}}(t)} \\ &= \overline{n}_{\mathcal{A}} \exp(-\kappa t) \end{aligned}$$

 $\mathcal{A} \rightarrow \emptyset$ | Statistics (cont.)

Example (Degradation - Statistics)

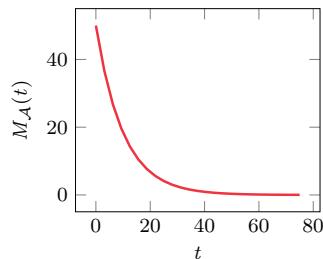
The expected evolution $(E[N_{\mathcal{A}}(t)])_{t \geq 0} = M_{\mathcal{A}}(t)$ of the degradation process $\mathcal{A} \xrightarrow{\kappa} \emptyset$

$$M_{\mathcal{A}}(t) = \overline{n}_{\mathcal{A}} \exp(-\kappa t)$$

Kinetic parameters and initial condition

$$\kappa = 0.1 \text{ sec}^{-1}$$

$$\begin{aligned} N_{\mathcal{A}}(0) &= n_{\mathcal{A}}(0) \\ &= 20 \end{aligned}$$



The approximation of the expected process evolution, based on a number of realisations

- The empirically evaluated spread around the expected evolution

 $\mathcal{A} \rightarrow \emptyset$ | Statistics (cont.)

Consider the deterministic representation of the degradation system in a compartment,

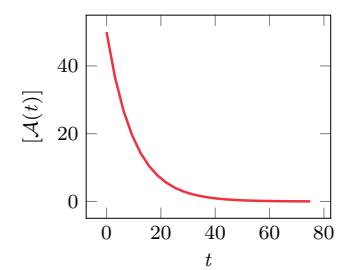
$$\frac{d[\mathcal{A}(t)]}{dt} = -\kappa [\mathcal{A}(t)]$$

The concentration of the chemical species \mathcal{A} in the reactor is $[\mathcal{A}](t) = n_{\mathcal{A}}(t)/V$

After solving this equation with initial condition $[\mathcal{A}(0)] = \overline{n}_{\mathcal{A}}/V$, we get

$$\begin{aligned} [\mathcal{A}(t)] &= [\mathcal{A}(0)] \exp(-\kappa t) \\ &= \overline{n}_{\mathcal{A}}/V \exp(-\kappa t) \\ &= M_{\mathcal{A}}(t)/V \end{aligned}$$

(We used a volume of size one)



For this case, the mean process can be obtained by solving the deterministic model