Discrete-time leap method for stochastic simulation

Christian Rohr

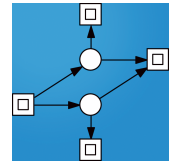
Department of Computer Science
Brandenburg University of Technology Cottbus-Senftenberg

Workshop on Biological Processes & Petri Nets 2016

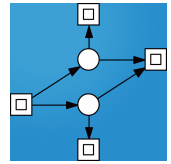
Torun

June 20, 2016

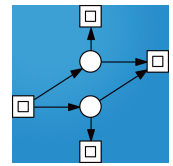
Outline



Outline



Introduction



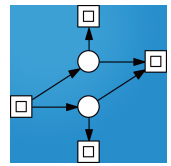
- Biological models are steadily increasing in their size and complexity
  - advances in systems and molecular biology, in particular through the high-throughput omic technologies
  - arise of multi-scale modelling[1]
- Smallbone and Mendes[2] raised issues of large-scale metabolic models:

  "... the inherent stiffness of genome-scale models."

$\Rightarrow$ demands for faster and more efficient simulation algorithms

[1] Heiner et al. 2013.

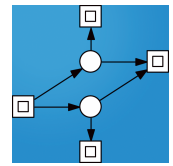
[2] Smallbone et al. 2013.



Definition

A stochastic Petri net $\mathcal{SPN} = (P, T, f, v, m_0)$ is defined as followed

- P a finite, non empty set of places ◯
- T a finite, non empty set of transitions □
- f: $((P \times T) \cup (T \times P)) \to \mathbb{N}_0$ (weighted directed arcs)
- m_0: initial state
- v: $T \to H$ (stochastic firing rate functions) with
  $H := \bigcup_{t \in T} \left\{ h_t | h_t : \mathbb{N}_0^{|\bullet t|} \to \mathbb{R}^+ \right\}$
  $v(t) = h_t$ for all transitions $t \in T$
- Semantic: Continuous Time Markov Chain (CTMC)



Definition

CTMC is a 3-tuple $(R, \mathbf{Q}, m_0)$ with R denoting the state space of the underlying net and m_0 the initial state.

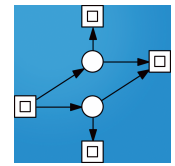
$$\mathbf{Q}: R \times R \longrightarrow \mathbb{R}_{\geq 0}$$

$$\mathbf{Q}(m, m') = \begin{cases} h_t(m) & \exists\, t \in T : m \xrightarrow{t} m' \\ 0 & \text{otherwise .} \end{cases}$$

$$E(m) = \sum_{m' \in R} \mathbf{Q}(m, m'), \text{ exit rate of state s}$$

The probability of a transition t enabled in marking m to fire (which results in marking m') within n time units is

$$1 - e^{-\mathbf{Q}(m,m') \cdot n} .$$

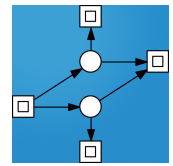


Definition

The change in the marking induced by firing transition t is denoted by

$$\Delta t = \{p \in {}^\bullet t \cup t^\bullet : \Delta t(p) = f(t,p) - f(p,t)\} \ .$$

When t in m fires, a new marking $m' = m + \Delta t$ is reached. This is denoted by $m \xrightarrow{t} m'$. The firing itself does not consume any time and takes place atomically.



A transition t may still be enabled after firing once, i.e. it may fire in a sequence $m \xrightarrow{t} m' \xrightarrow{t} m'' \xrightarrow{t} \ldots \xrightarrow{t} m^n$ until it is not enabled any more.

Definition

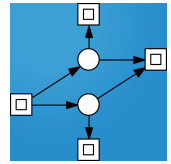
The number of occurrences of transition t in such a sequence starting in marking m is named *enabledness (concurrency) degree*

$$\mathrm{ed}_t = min_{p \in {}^\bullet t} \left(\left\lfloor \frac{m(p)}{f(p,t)} \right\rfloor \right).$$

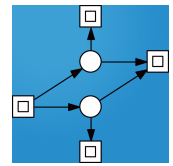
The standard firing rule can be extended by the enabledness degree

$$m' = m + \Delta t \cdot \mathrm{ed}_t(m).$$

Outline



Discrete-time Leap Method



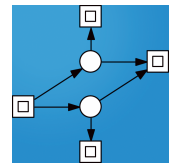
Basic idea

- converting the CTMC into a DTMC by uniformization[3]
  $\hookrightarrow$ CTMC $\equiv$ DTMC with implicit Poisson process[4]
- generating paths through the DTMC is as expensive as for the CTMC
  $\hookrightarrow$ leaping over several states
- all enabled and not mutually exclusive transitions are forced to fire in a leap
- let each transition fire concurrently to itself

[3] Jensen 1953; Stewart 1994.
[4] Sandmann 2008.

Discrete-time Leap Method

Transition firing



$$\text{firing rate} \approxeq random[0, \text{enablness degree}]$$

- Uniformized DTMC: the number of transitions in any time interval δ has a Poisson distribution with rate λ
- Poisson limit theorem

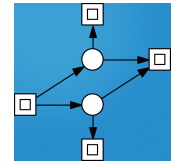
$$\text{Pois}(\lambda) \Longrightarrow \lambda = k \cdot pr \Longrightarrow \text{B}(k, pr)$$

- Binomial distribution

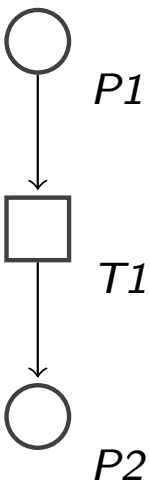
$$k = \text{ed}_t\,,\ pr = \begin{cases} 1 - e^{-\frac{h_t(m)}{\text{ed}_t(m)} \cdot \delta} & \text{ed}_t(m) > 0 \\ 0 & otherwise. \end{cases}$$

Discrete-time Leap Method

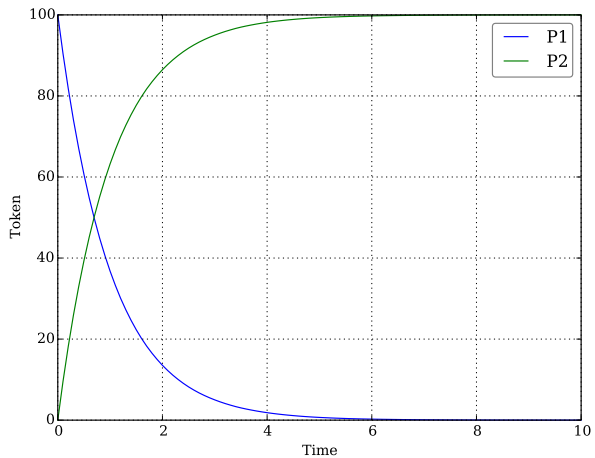
Transition firing



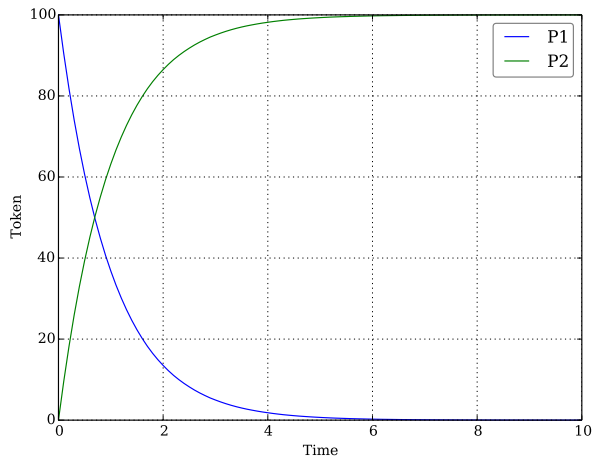
First order reaction: $P1 \rightarrow P2$



(a) $\mathcal{PN}$



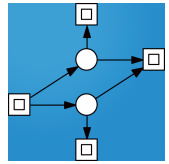
(b) δ-leaping



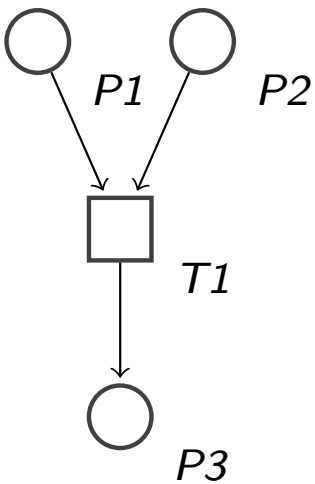
(c) direct method

Discrete-time Leap Method

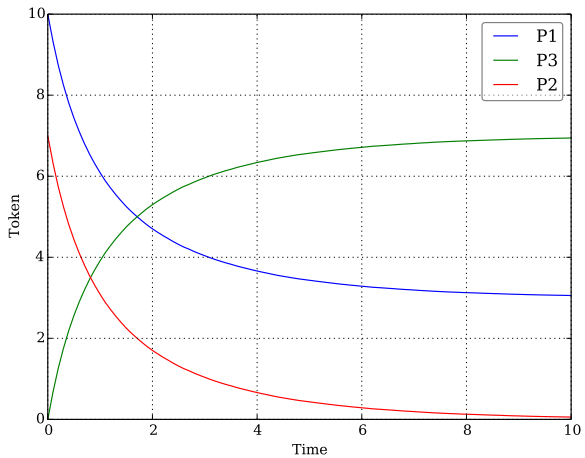
Transition firing



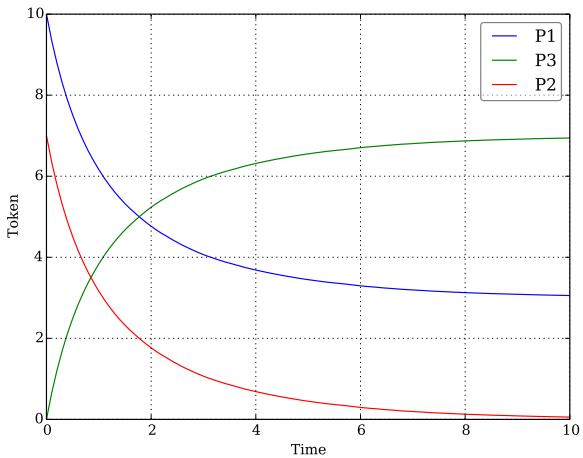
Second order reaction: $P1 + P2 \rightarrow P3$



(a) $\mathcal{PN}$



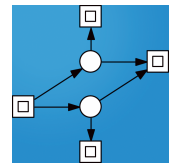
(b) δ-leaping



(c) direct method

Discrete-time Leap Method

Dependent Subnets



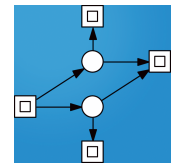
- conflicts and sequences
  $\hookrightarrow$ order of transition firings is important
- generate a weighted random sequence of all transitions in each step and let them fire (if enabled) sequentially
- modern version of the Fisher–Yates shuffle[5]
- Bernoulli sampling to realize a shuffling in accordance to transition weights
- transition weight approximates the expected firing rate, if the transition would be enabled

$$w_t = \sum_{p \in {}^\bullet t} f(p,t)^{f(p,t)}$$

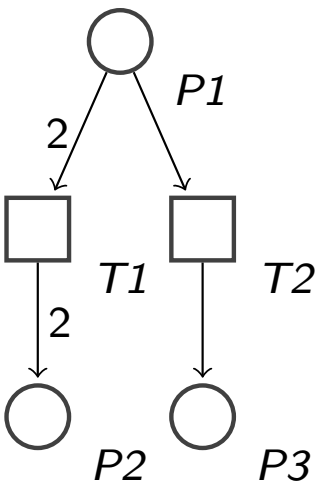
[5] Durstenfeld 1964; Fisher et al. 1963.

Discrete-time Leap Method

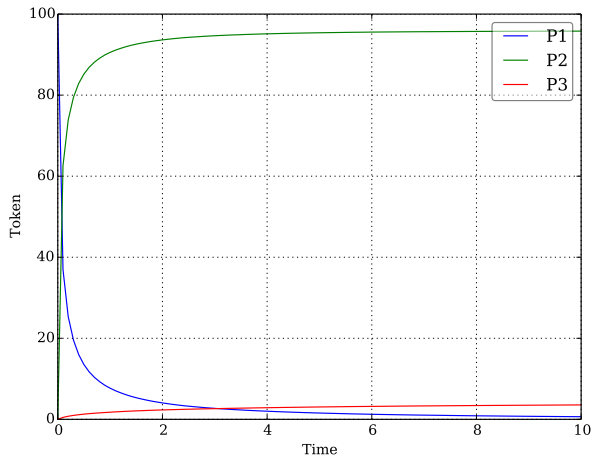
Dependent Subnets



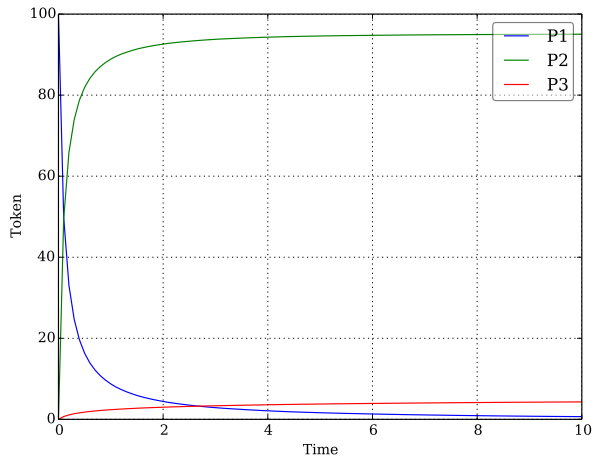
Conflict: $P1 \rightarrow P2$ and $P1 \rightarrow P3$



(a) $\mathcal{PN}$



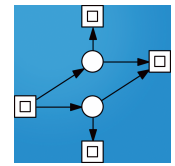
(b) δ-leaping



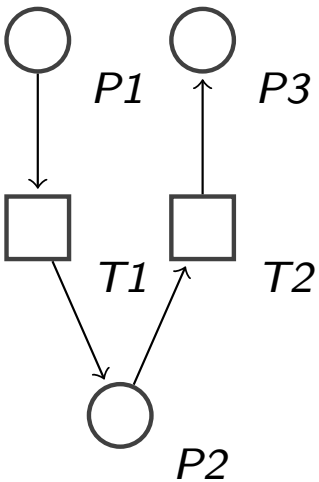
(c) direct method

Discrete-time Leap Method

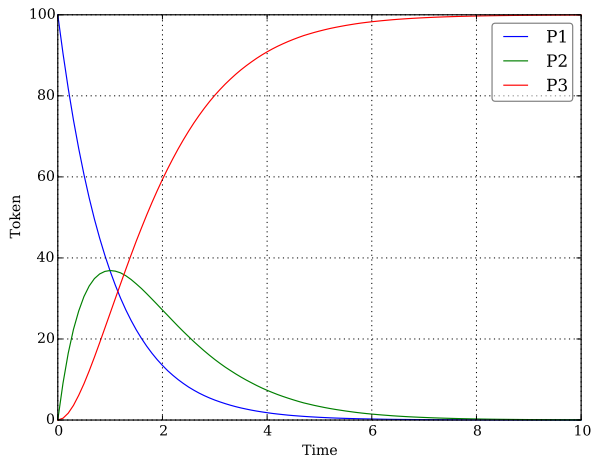
Dependent Subnets



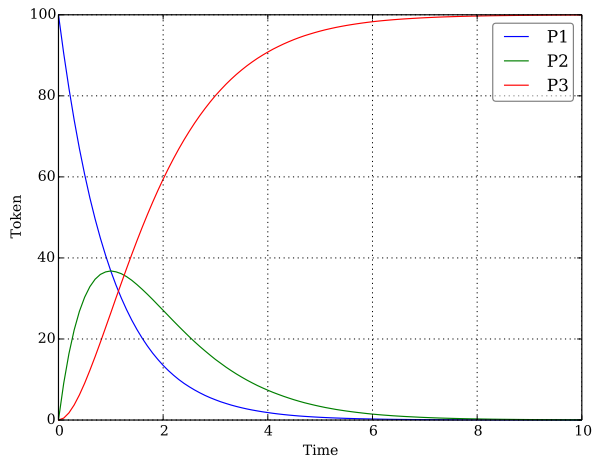
Sequence: $P1 \rightarrow P2 \rightarrow P3$



(a) $\mathcal{PN}$



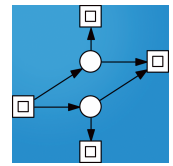
(b) δ-leaping



(c) direct method

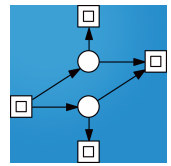
Discrete-time Leap Method

Algorithm

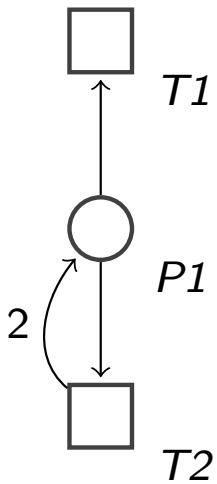


Require: $\mathcal{SPN}$ with initial marking m_0, time interval $[\tau_0, \tau_{max}]$, time step δ
Ensure: marking m at time point τ_{max}

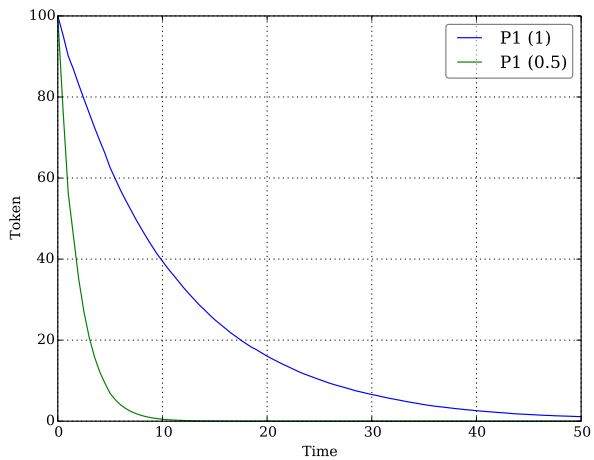
1: $time\ \tau \leftarrow \tau_0$
2: $marking\ m \leftarrow m_0$
3: $T_r \leftarrow T$
4: **while** $\tau \leq \tau_{max}$ **do**
5: $\quad T_r \leftarrow \textsc{weightedRandomShuffle}(T_r)$
6: $\quad$ **for all** transitions $t_j \in T_r$ **do**
7: $\quad\quad k \leftarrow \textsc{enablednessDegree}(t_j, m)$
8: $\quad\quad h \leftarrow \textsc{transitionRate}(t_j, m)$
9: $\quad\quad$ **if** $k > 0$ **then**
10: $\quad\quad\quad f \leftarrow \textsc{binomialSampling}(h, (1 - e^{-\frac{h}{k} \cdot \delta}))$
11: $\quad\quad\quad m \leftarrow m + f \cdot \Delta t_j$
12: $\quad\quad$ **end if**
13: $\quad$ **end for**
14: $\quad \tau \leftarrow \tau + \delta$
15: **end while**



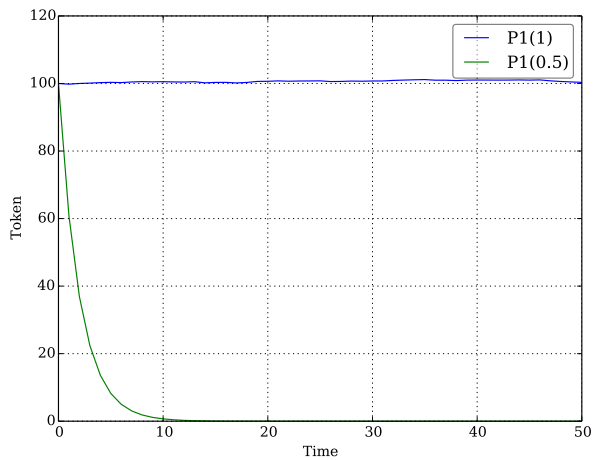
Simplified birth-death process



(a) $\mathcal{PN}$



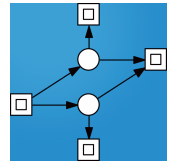
(b) δ-leaping



(c) direct method

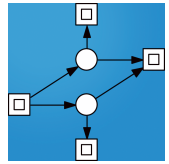
The results for different rate constants of *T2*, i.e. $c_{T2} = 1$ (blue) and $c_{T2} = 0.5$ (green).

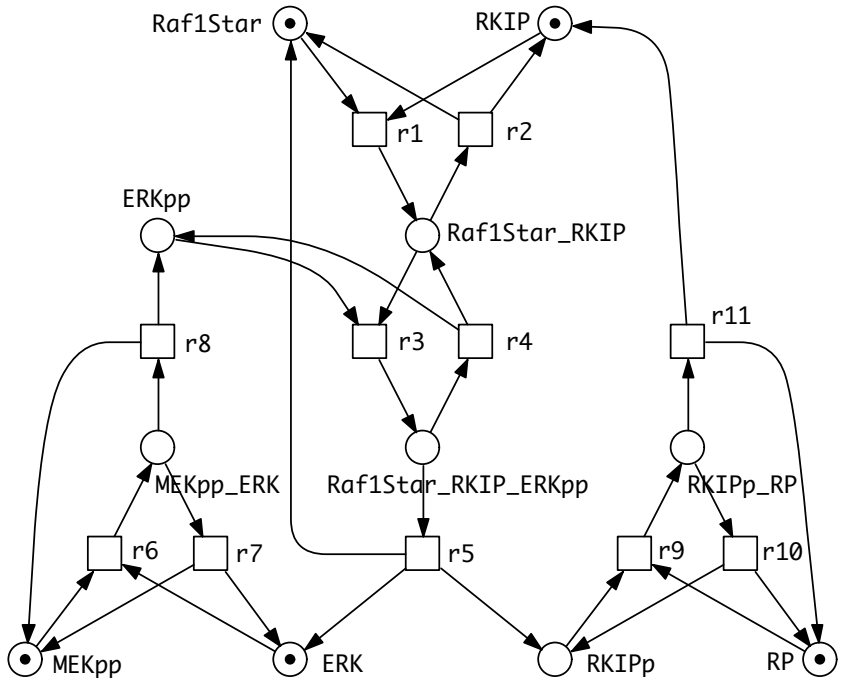
Outline



Case Studies

RKIP inhibited ERK pathway





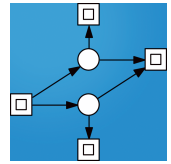
- non-linear ODE model[6], Petri net model[7]
- stochastic Petri net with 11 places and 11 transitions connected by 34 arcs
- mass action kinetics with original parameter values

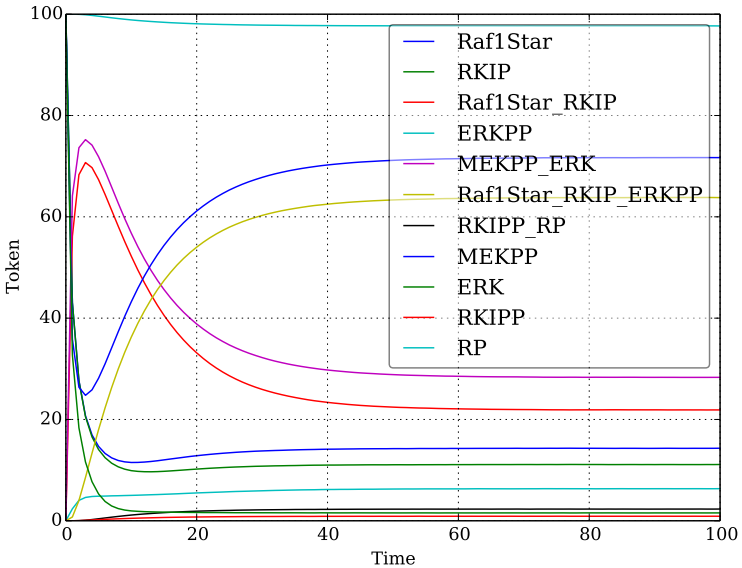
[6] Cho et al. 2003.

[7] Gilbert et al. 2006; Heiner et al. 2010.

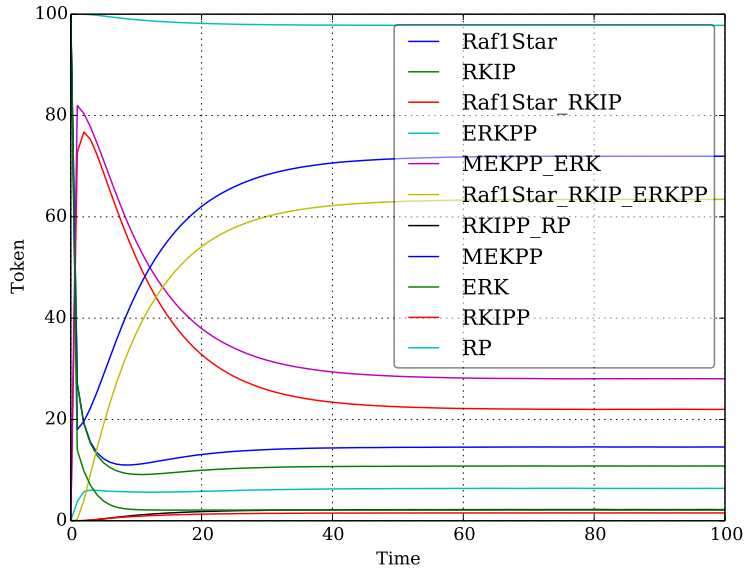
Case Studies

RKIP inhibited ERK pathway





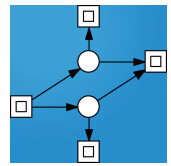
(a) direct method, $N = 100$



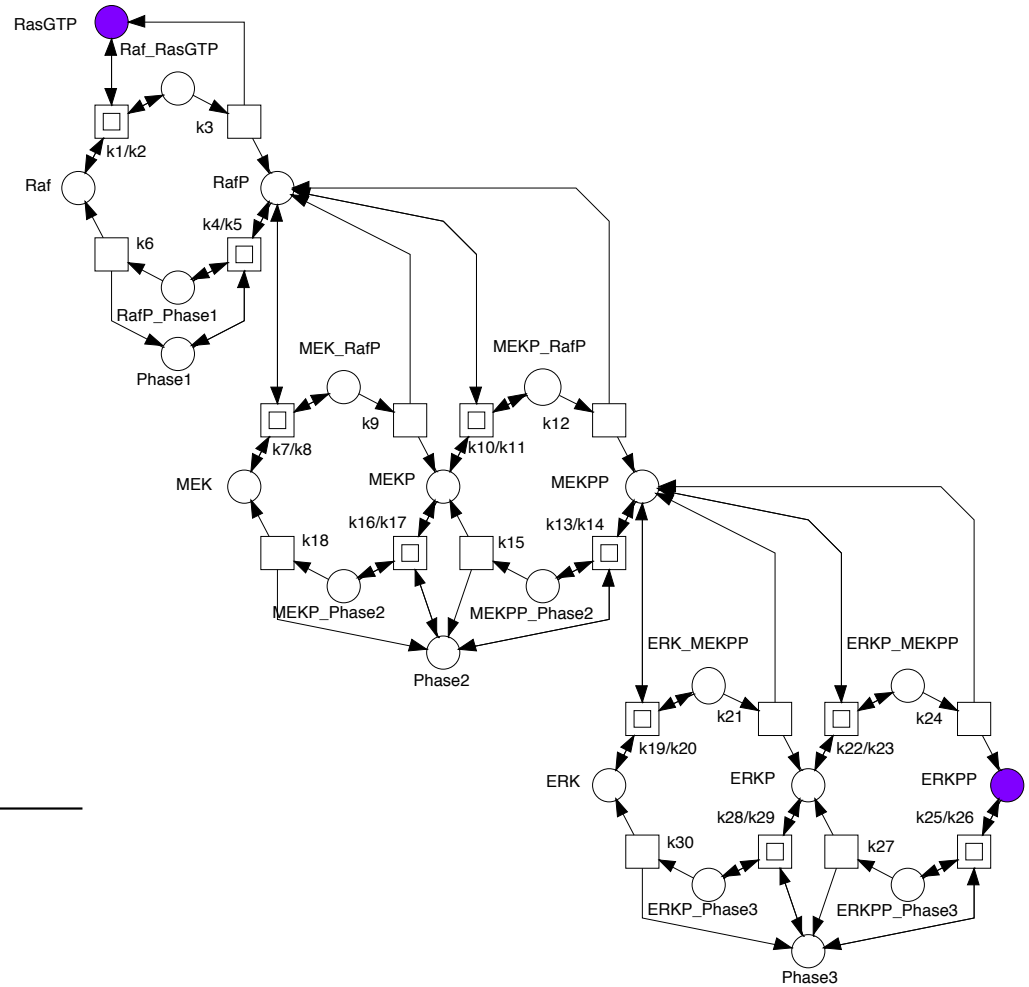
(b) δ-leaping, $N = 100$

Case Studies

Mitogen-activated Protein Kinase



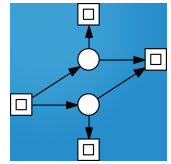
- ODE model[a], Petri net model[b]
- stochastic Petri net with 22 places and 30 transitions connected by 90 arcs
- mass action kinetics with original parameter values

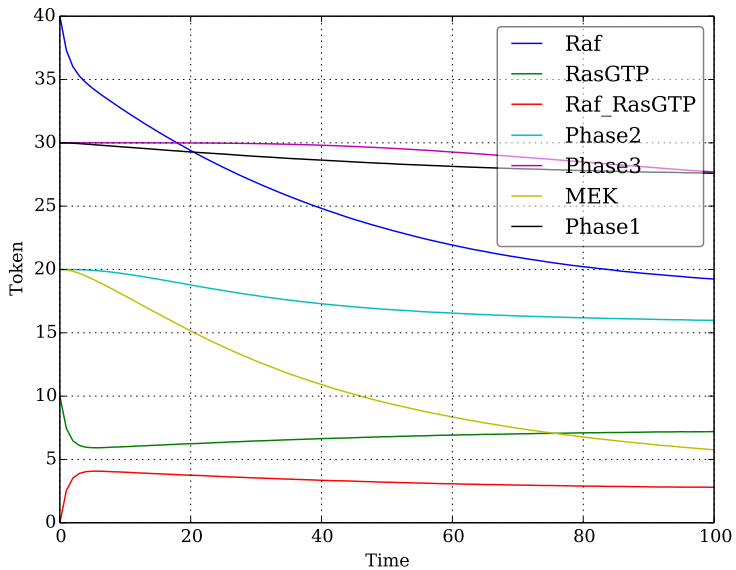


[a] Levchenko et al. 2000.
[b] Heiner et al. 2008.

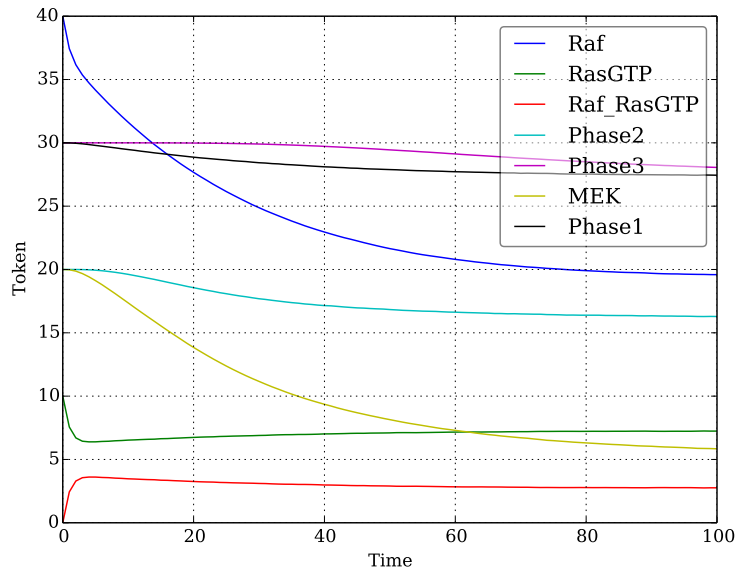
Case Studies

Mitogen-activated Protein Kinase





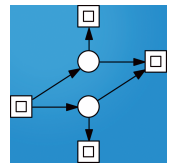
(a) direct method, $N = 10$



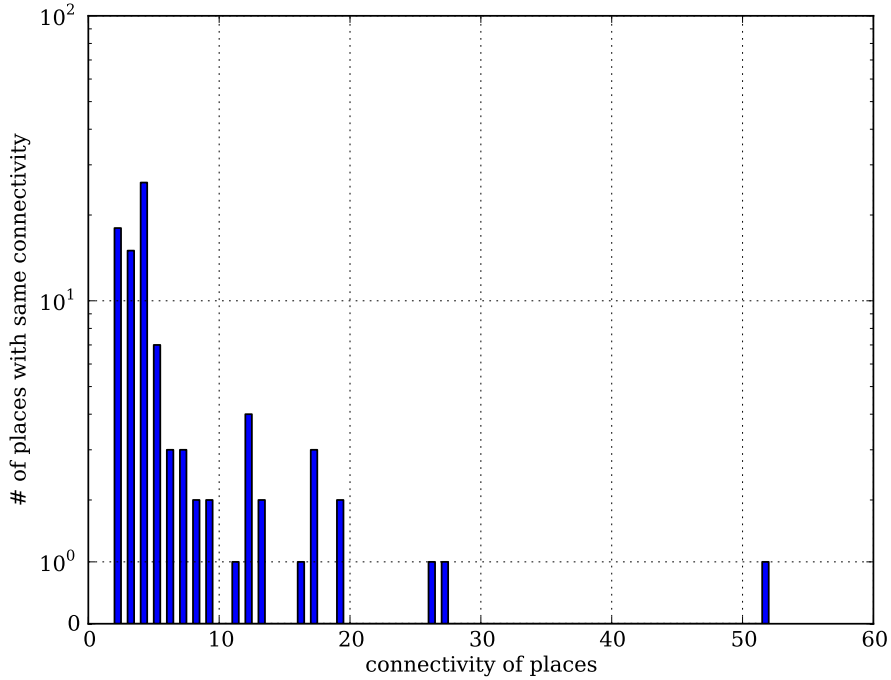
(b) δ-leaping, $N = 10$

Case Studies

E.coli K-12 Core Metabolic model



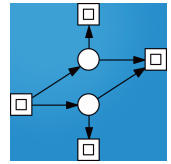
- FBA model[a], Petri net model[b]
- stochastic Petri net with 93 places and 172 transitions connected by 589 arcs
- mass action kinetics with unknown parameter values

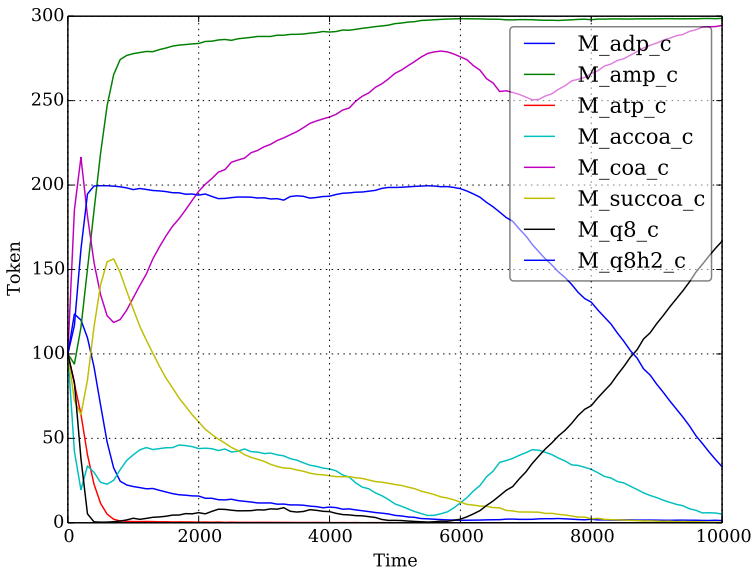


[a] Orth et al. 2010.
[b] Gilbert et al. 2016.

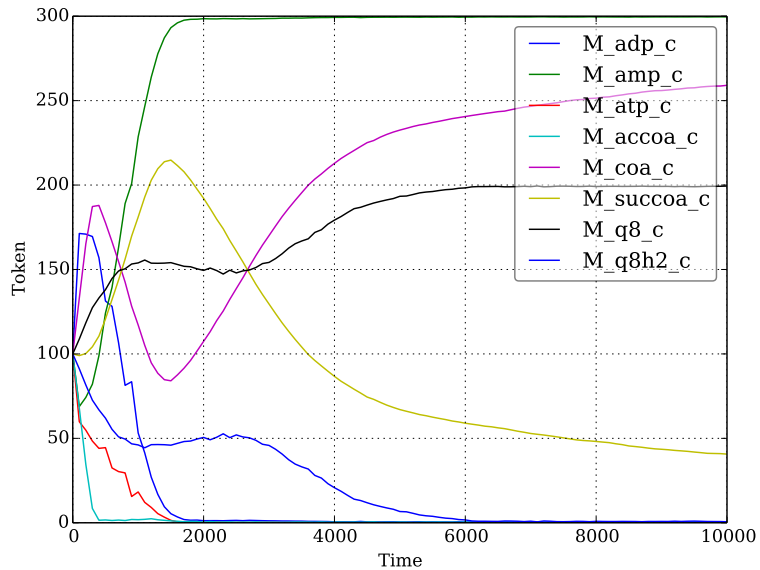
Case Studies

E.coli K-12 Core Metabolic model





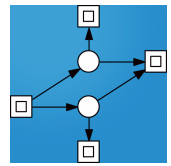
(a) direct method, $N = 100$



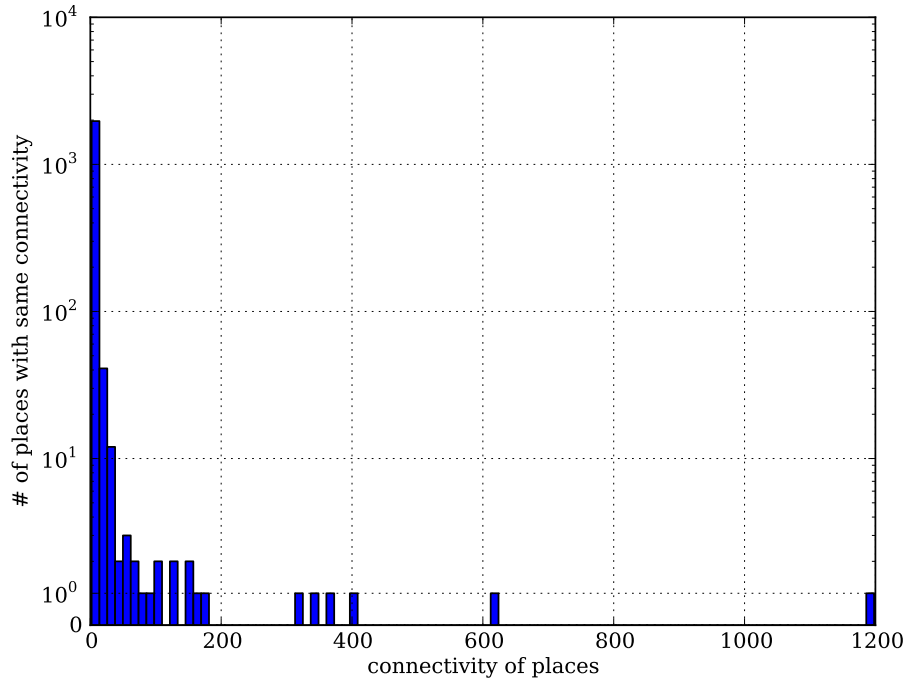
(b) δ-leaping, $N = 100$

Case Studies

E.coli K-12 Genome Scale Metabolic model



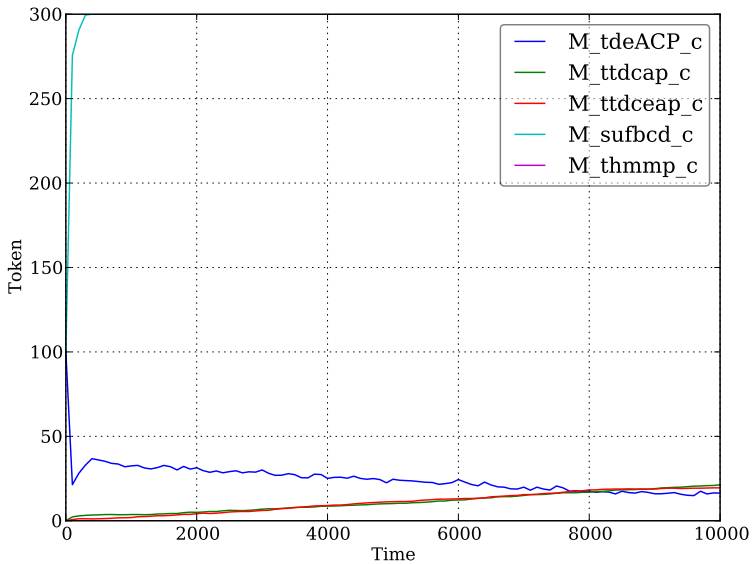
- FBA model[a], Petri net model[b]
- stochastic Petri net with 2046 places and 3703 transitions connected by 13001 arcs
- mass action kinetics with unknown parameter values



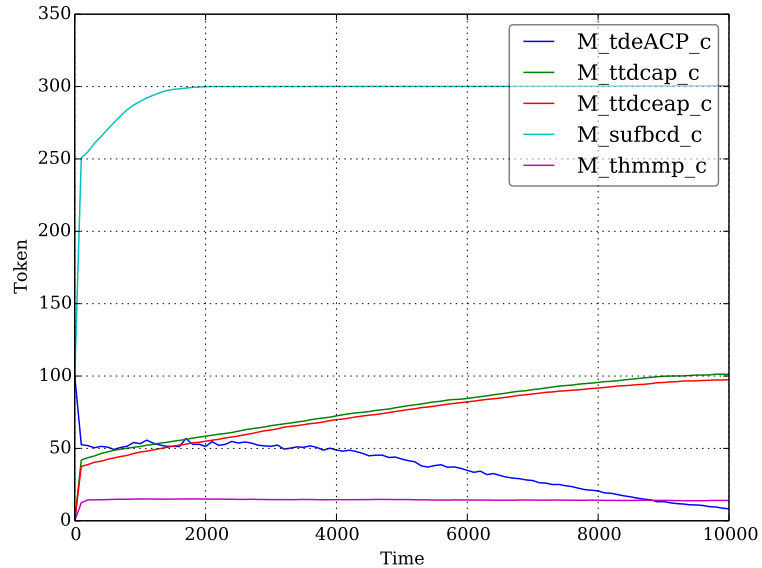
[a] Monk et al. 2013.
[b] Gilbert et al. 2016.

Case Studies

E.coli K-12 Genome Scale Metabolic model



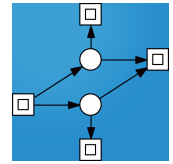
(a) direct method, $N = 100$



(b) δ-leaping, $N = 100$

Case Studies

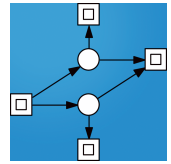
Comparison



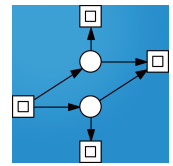
Run-times for the direct method (a) and δ-leaping (b).
† is placed, if the simulation did not finish in reasonable time (>40 days).

model	N	1 run		10 runs		100 runs		100 000 runs	
		a	b	a	b	a	b	a	b
ERK	10	<1s	<1s	<1s	<1s	<1s	<1s	9s	35s
ERK	100	<1s	<1s	<1s	<1s	<1s	<1s	1m21s	46s
ERK	1000	<1s	<1s	<1s	<1s	<1s	<1s	13m27s	1m5s
MAPK	1	<1s	<1s	<1s	<1s	<1s	<1s	12s	1m7s
MAPK	10	<1s	<1s	<1s	<1s	<1s	<1s	1m47s	2m1s
MAPK	100	<1s	<1s	<1s	<1s	<1s	<1s	16m50s	2m38s
E.coli core	50	13s	<1s	2m1s	3s	50m8s	50s	†	11h20m
E.coli core	100	1m47s	<1s	18m4s	3s	6h56m	56s	†	11h50m
E.coli core	500	46m30s	<1s	10h31m	4s	7d10h	1m12s	†	14h10m
E.coli K-12	50	18h33m	9s	10d16h	1m31s	†	33m1s	†	12d13h
E.coli K-12	100	2d17h	9s	40d7h	1m32s	†	33m4s	†	12d14h
E.coli K-12	500	†	9s	†	1m32s	†	33m20s	†	13d12h

Outline

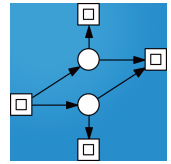


Conclusion



- efficient simulation algorithm for larger and dense networks
- less sensitive to higher number of tokens and thus higher transition rates
- suitable for in silico experiments, i.e. knock-out scenarios of certain species or reactions
- implemented and available in Snoopy and MARCIE

Outlook



- improve transition weight heuristic
- improve accuracy by adaptive δ selection
- apply on further case studies

Thank you for your attention!

http://www-dssz.informatik.tu-cottbus.de