Probabilistic Approximations of Signaling Pathway Dynamics: Supplementary Materials

Bing Liu[1], P.S. Thiagarajan[1,2] and David Hsu[1,2]

[1] NUS Graduate School for Integrative Sciences and Engineering, National University of Singapore
[2] Department of Computer Science, National University of Singapore
{liubing, dyhsu, thiagu}@comp.nus.edu.sg

1 Appendix: Technical Background and Details

Most of the material reviewed here is standard and has been assembled from [1–3].

1.1 Continuity

Assume that X and Y are metric spaces. A function $f : X \to Y$ is said to be of class C^k, where $k \in \mathbb{N}$, if the derivatives f', f'', ..., $f^{(k)}$ exist and are continuous. Thus, the class C^0 consists of all continuous functions and the class C^1 consists of all continuously differentiable functions. A function $f : X \to Y$ is called **Lipschitz continuous**, if there exists a constant $\lambda \in \mathbb{R}_+$ such that

$$\forall x, y \in X, d_Y(f(x), f(y)) \leq \lambda d_X(x, y),$$

where d_X and d_Y denote the respective metrics in X and Y. The class of Lipschitz continuous functions is denoted as C^{1-} and $C^1 \subseteq C^{1-} \subseteq C^0$ holds. [1].

1.2 Probability and Measure Theory

A **σ-algebra** over a set X is a nonempty collection of subsets of X that is closed under complementation and countable unions. The **Borel σ-algebra** on a topological space X, denoted as $\mathcal{B}_X$, is the minimal σ-algebra containing all the open sets of X.

A **probability space** is a triple $(\Omega, \mathcal{F}, P)$ consisting of a set Ω, a σ-algebra $\mathcal{F}$ over Ω, and a function $P : \mathcal{F} \to [0, 1]$ such that: (i) $P(\Omega) = 1$; (ii) if $\{A_w\}_{w \in W}$ is a countable family of pairwise disjoint sets in $\mathcal{F}$, then $P(\cup_w A_w) = \sum_w P(A_w)$.

Let X and Y be nonempty sets and $\mathcal{M}$ and $\mathcal{N}$ be σ-algebras of subsets of X and Y respectively. A function $f : X \to Y$ is said to be **$(\mathcal{M}, \mathcal{N})$-measurable** if

$$E \in \mathcal{N} \Rightarrow f^{-1}(E) \equiv \{x \in X | f(x) \in E\} \in \mathcal{M}.$$

1.3 Flows and Probability Distributions

We assume a set of ODEs $\dot{x_i}(t) = f_i(\mathbf{x}(t), \mathbf{p})$ involving the continuous real-valued variables $\{x_1, x_2, \ldots, x_n\}$ and real-valued parameters $\{p_1, p_2, \ldots, p_m\}$. We treat parameters as time-invariant variables and implicitly assume the given system of ODEs to be augmented with m additional differential equations of the form $\dot{p_j}(t) = 0$ with j ranging over $\{1, 2, \ldots, m\}$. In what follows, we will let $\mathbf{x}$ range over $\mathbb{R}^n$, $\mathbf{k}$ range over $\mathbb{R}^m$ and $\mathbf{z}$ range over $\mathbb{R}^{n+m}$.

In vector form, our system of ODEs may be represented as $\mathbf{Z}' = F(\mathbf{Z})$. We assume $F : \mathcal{D} \to \mathcal{D}$ is a C^1 function, where the domain $\mathcal{D}$ is a bounded region of $\mathbb{R}^{n+m}$. Given $\mathbf{z}_0 = (\mathbf{v}_0, \mathbf{k})$ where $\mathbf{v}_0$ specifies the initial values of the variables and $\mathbf{k}$ specifies the parameters values, the system of ODEs will have a unique solution if $F \in C^{1-}$ which will be the case since $F \in C^1$ by assumption and $C^1 \subseteq C^{-1}$ [1]. We shall denote this solution by $\mathbf{Z}(t)$ with $\mathbf{Z}(0) = \mathbf{z}_0$ and $\mathbf{Z}'(t) = F(\mathbf{Z}(t))$.

The flow $\Phi : \mathbb{R} \times \mathcal{D} \to \mathcal{D}$ of $\mathbf{Z}' = F(\mathbf{Z})$ will be a C^0 function and it can be defined via: $\Phi(t, \mathbf{z}) = \mathbf{Z}(t)$ with $\Phi(0, \mathbf{z}) = \mathbf{Z}(0) = \mathbf{z}$. The following fact is crucial for our purposes.

Proposition 1. [2] *If X and Y are metric spaces and $f : X \to Y$ is continuous, then f is $(\mathcal{B}_X, \mathcal{B}_Y)$-measurable.*

As observed earlier, the flow Φ is continuous. In addition, we have $\mathcal{D} \subseteq \mathbb{R}^{n+m}$ is a metric space. Thus, $\Phi(t, \cdot)$ is $(\mathcal{B}_\mathcal{D}, \mathcal{B}_\mathcal{D})$-measurable by Proposition 1. In what follows, we use Φ_t to denote $\Phi(t, \cdot)$. Then for all $t \in \mathbb{R}$, we have

$$B \in \mathcal{B}_\mathcal{D} \Rightarrow \Phi_t^{-1}(B) = \{\mathbf{z} \in \mathcal{D} \mid \Phi(t, \mathbf{z}) \in B\} \in \mathcal{B}_\mathcal{D}.$$

We assume we are given a prior distribution in the form of a probability density function Υ^0 capturing the distribution of initial values. This will induce the probability space $(\mathcal{D}, \mathcal{B}_\mathcal{D}, P^0)$, where P^0 is given by:

$$P^0(B) = \int_B \Upsilon^0(\mathbf{z}) d\mathbf{z}, \text{ for every } B \in \mathcal{B}_\mathcal{D}.$$

We can now define a probability distribution P^t over $\mathcal{B}_\mathcal{D}$ for every t as:

$$P^t(B) = P^0(\Phi_t^{-1}(B)), \text{ for every } B \in \mathcal{B}_\mathcal{D}.$$

1.4 $\mathcal{MC}_{ideal}$ and $\mathcal{MC}_{approx}$

Suppose $\mathcal{MC}_{ideal} = (\mathcal{M}, \{p_{ij}\})$ is the Markov chain we wish to approximate. We then compute a Markov chain $\mathcal{MC}_{approx} = (\mathcal{M}, \{\hat{p}_{ij}\})$ from N trajectories.

Proposition 2. *Each transition probability $\hat{p}_{ij}$ in $\mathcal{MC}_{approx}$ differs from the corresponding p_{ij} in $\mathcal{MC}_{ideal}$ by an error less than or equal to ϵ.*

$$\epsilon = \phi^{-1}(\frac{c+1}{2})\sqrt{\frac{p_{ij}(1-p_{ij})}{N}} \text{ with probability } c,$$

where $\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-x}^{x} e^{-y^2/2} dy$.

Proof. Let X be a random variable such that $X = 1$ ($X = 0$ resp.) denote the event that a sample trajectory passes (not passes resp.) a discrete state $\mathbf{s}$ at time $d \cdot \Delta t$. Suppose X has a *Bernoulli distribution* with parameter p_{ij}, then we have $\mu = p_{ij}$ and $\sigma^2 = p_{ij}(1 - p_{ij})$. If $X_1, X_2, \ldots, X_N$ are the N measurements, by *Central Limit Theorem*, we have:

$$P\{-\epsilon \leq \frac{\sum_{i=1}^{N} X_i}{N} - \mu \leq \epsilon\} \approx 2\phi(\epsilon\frac{\sqrt{N}}{\sigma}) - 1$$

where ϵ is the error and $\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-x}^{x} e^{-y^2/2} dy$. Thus,

$$\epsilon = \phi^{-1}(\frac{c+1}{2})\sqrt{\frac{p_{ij}(1-p_{ij})}{N}} \text{ with probability } c.$$

□

Let $M = (\mathbf{s} = (I_1, I_2, \ldots, I_{n+m}), d) \in \mathcal{M}$, where $I_i \in \mathcal{I}$. We define $Pr(M) = P^{d \cdot \Delta t}(\{\mathbf{z} \mid \mathbf{z} \in I_1 \times I_2 \times \ldots \times I_{n+m}\})$. Similar to proposition 2, we have:

Proposition 3. *For each $M = (\boldsymbol{s}, d) \in \mathcal{M}$, its probability $\hat{Pr}(M)$ computed from $\mathcal{MC}_{approx}$ differs from the one $Pr(M)$ defined in $\mathcal{MC}_{ideal}$ by an error less than or equal to ϵ:*

$$\epsilon = \phi^{-1}(\frac{c+1}{2})\sqrt{\frac{Pr(M)(1-Pr(M))}{N}} \textit{ with probability } c,$$

where $\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-x}^{x} e^{-y^2/2} dy$.

2 Supplementary Figures and Tables

Table 1. DBN Structure of EGF-NGF signaling pathway

Name	Variable	Parents
EGF	x_1	k_1, x_1, x_3, k_2, x_4
NGF	x_2	k_3, x_2, x_5, k_4, x_6
free EGF Recepter	x_3	k_1, x_1, x_3, k_2, x_4
bound EGF Recepter	x_4	k_1, x_1, x_3, k_2, x_4
free NGF Recepter	x_5	k_3, x_2, x_5, k_4, x_6
bound NGF Recepter	x_6	k_3, x_2, x_5, k_4, x_6
inactive Sos	x_7	$k_9, x_{10}, x_8, x_8, k_{10}, k_5, x_4, x_7, x_7, k_6, k_7, x_6, x_7, x_7, k_8$
active Sos	x_8	$k_9, x_{10}, x_8, x_8, k_{10}, k_5, x_4, x_7, x_7, k_6, k_7, x_6, x_7, x_7, k_8$
inactive P90Rsk	x_9	$k_{27}, x_{21}, x_9, x_9, k_{28}$
active P90Rsk	x_{10}	$k_{27}, x_{21}, x_9, x_9, k_{28}$
inactive Ras	x_{11}	$k_{11}, x_{11}, x_{11}, k_{12}, k_{13}, x_{13}, x_{12}, x_{12}, k_{14}$
active Ras	x_{12}	$k_{11}, x_{11}, x_{11}, k_{12}, k_{13}, x_{13}, x_{12}, x_{12}, k_{14}$
active RasGap	x_{13}	x_{13}
inactive Raf	x_{14}	$k_{15}, x_{12}, x_{14}, x_{14}, k_{16}, k_{45}, x_{32}, x_{15}, x_{15}, k_{46}, k_{35}, x_{25}, x_{15}, x_{15}, k_{36}$
active Raf	x_{15}	$k_{15}, x_{12}, x_{14}, x_{14}, k_{16}, k_{45}, x_{32}, x_{15}, x_{15}, k_{46}, k_{35}, x_{25}, x_{15}, x_{15}, k_{36}$
inactive B-Raf	x_{16}	$k_{43}, x_{29}, x_{16}, x_{16}, k_{44}, k_{47}, x_{32}, x_{17}, x_{17}, k_{20}$
active B-Raf	x_{17}	$k_{43}, x_{29}, x_{16}, x_{16}, k_{44}, k_{47}, x_{32}, x_{17}, x_{17}, k_{20}$
inactive Mek	x_{18}	$k_{17}, x_{15}, x_{18}, x_{18}, k_{18}, k_{19}, x_{17}, x_{18}, x_{18}, k_{48}, k_{21}, x_{31}, x_{19}, x_{19}, k_{22}$
active Mek	x_{19}	$k_{17}, x_{15}, x_{18}, x_{18}, k_{18}, k_{19}, x_{17}, x_{18}, x_{18}, k_{48}, k_{21}, x_{31}, x_{19}, x_{19}, k_{22}$
inactive Erk	x_{20}	$k_{23}, x_{19}, x_{20}, x_{20}, k_{24}, k_{25}, x_{31}, x_{21}, x_{21}, k_{26}$
active Erk	x_{21}	$k_{23}, x_{19}, x_{20}, x_{20}, k_{24}, k_{25}, x_{31}, x_{21}, x_{21}, k_{26}$
inactive PI3K	x_{22}	$k_{29}, x_4, x_{22}, x_{22}, k_{30}, k_{31}, x_{12}, x_{22}, x_{22}, k_{32}$
active PI3K	x_{23}	$k_{29}, x_4, x_{22}, x_{22}, k_{30}, k_{31}, x_{12}, x_{22}, x_{22}, k_{32}$
inactive Akt	x_{24}	$k_{33}, x_{23}, x_{24}, x_{24}, k_{34}$
active Akt	x_{25}	$k_{33}, x_{23}, x_{24}, x_{24}, k_{34}$
inactive C3G	x_{26}	$k_{37}, x_6, x_{26}, x_{26}, k_{38}$
active C3G	x_{27}	$k_{37}, x_6, x_{26}, x_{26}, k_{38}$
inactive Rap1	x_{28}	$k_{39}, x_{27}, x_{28}, x_{28}, k_{40}, k_{41}, x_{30}, x_{29}, x_{29}, k_{42}$
active Rap1	x_{29}	$k_{39}, x_{27}, x_{28}, x_{28}, k_{40}, k_{41}, x_{30}, x_{29}, x_{29}, k_{42}$
active RapGap	x_{30}	x_{30}
active PP2A	x_{31}	x_{31}
active RafPP	x_{32}	x_{32}

Table 2. Prior (initial) probability distribution of variables

Probability distribution
$x_1 \sim U(8801760.0, 1.10022 \times 10^7)$
$x_2 \sim U(401280.0, 501600.0)$
$x_3 \sim U(70400.0, 88000.0)$
$x_4 \sim U(0.0, 17600.0)$
$x_5 \sim U(8800.0, 11000.0)$
$x_6 \sim U(0.0, 2200.0)$
$x_7 \sim U(105600.0, 132000.0)$
$x_8 \sim U(0.0, 26400.0)$
$x_9 \sim U(105600.0, 132000.0)$
$x_{10} \sim U(0.0, 26400.0)$
$x_{11} \sim U(105600.0, 132000.0)$
$x_{12} \sim U(0.0, 26400.0)$
$x_{13} \sim U(105600.0, 132000.0)$
$x_{14} \sim U(105600.0, 132000.0)$
$x_{15} \sim U(0.0, 26400.0)$
$x_{16} \sim U(105600.0, 132000.0)$
$x_{17} \sim U(0.0, 26400.0)$
$x_{18} \sim U(528000.0, 660000.0)$
$x_{19} \sim U(0.0, 132000.0)$
$x_{20} \sim U(528000.0, 660000.0)$
$x_{21} \sim U(0.0, 132000.0)$
$x_{22} \sim U(105600.0, 132000.0)$
$x_{23} \sim U(0.0, 26400.0)$
$x_{24} \sim U(105600.0, 132000.0)$
$x_{25} \sim U(0.0, 26400.0)$
$x_{26} \sim U(105600.0, 132000.0)$
$x_{27} \sim U(0.0, 26400.0)$
$x_{28} \sim U(105600.0, 132000.0)$
$x_{29} \sim U(0.0, 26400.0)$
$x_{30} \sim U(105600.0, 132000.0)$
$x_{31} \sim U(105600.0, 132000.0)$
$x_{32} \sim U(105600.0, 132000.0)$

Table 3. The range and nominal probability distributions of parameters. For unknown parameters (marked with *), we assume the their prior are uniform distributions over their ranges.

Parameter	Range	Nominal probability distribution
k_1*	$[0, 4.37006 \times 10^{-5}]$	$k_1 \sim U(1.748024 \times 10^{-5}, 2.622036 \times 10^{-5})$
k_2*	$[0, 0.0242016]$	$k_2 \sim U(0.00968064, 0.01452096)$
k_3*	$[0, 2.76418 \times 10^{-7}]$	$k_3 \sim U(1.105672 \times 10^{-7}, 1.658508 \times 10^{-7})$
k_4*	$[0, 0.01447622]$	$k_4 \sim U(0.005790488, 0.008685732)$
k_5	$[0, 1389.462]$	$k_5 \sim U(555.7848, 833.6772)$
k_6	$[0, 1.217214 \times 10^7]$	$k_6 \sim U(4868856.0, 7303284.0)$
k_7	$[0, 778.856]$	$k_7 \sim U(311.5424, 467.3136)$
k_8	$[0, 4225.32]$	$k_8 \sim U(1690.128, 2535.192)$
k_9	$[0, 3223.94]$	$k_9 \sim U(1289.576, 1934.364)$
k_{10}	$[0, 1793792.0]$	$k_{10} \sim U(717516.8, 1076275.2)$
k_{11}*	$[0, 64.688]$	$k_{11} \sim U(25.8752, 38.8128)$
k_{12}*	$[0, 71908.6]$	$k_{12} \sim U(28763.44, 43145.16)$
k_{13}	$[0, 3018.72]$	$k_{13} \sim U(1207.488, 1811.232)$
k_{14}	$[0, 2864820.0]$	$k_{14} \sim U(1145928.0, 1718892.0)$
k_{15}*	$[0, 1.768192]$	$k_{15} \sim U(0.7072768, 1.0609152)$
k_{16}	$[0, 124929.2]$	$k_{16} \sim U(49971.68, 74957.52)$
k_{17}*	$[0, 371.518]$	$k_{17} \sim U(148.6072, 222.9108)$
k_{18}	$[0, 9536700.0]$	$k_{18} \sim U(3814680.0, 5722020.0)$
k_{19}	$[0, 250.178]$	$k_{19} \sim U(100.0712, 150.1068)$
k_{20}	$[0, 315896.0]$	$k_{20} \sim U(126358.4, 189537.6)$
k_{21}	$[0, 5.66486]$	$k_{21} \sim U(2.265944, 3.398916)$
k_{22}	$[0, 1037506.0]$	$k_{22} \sim U(415002.4, 622503.6)$
k_{23}*	$[0, 19.70734]$	$k_{23} \sim U(7.882936, 11.824404)$
k_{24}	$[0, 2014680.0]$	$k_{24} \sim U(805872.0, 1208808.0)$
k_{25}	$[0, 17.7824]$	$k_{25} \sim U(7.11296, 10.66944)$
k_{26}	$[0, 6992980.0]$	$k_{26} \sim U(2797192.0, 4195788.0)$
k_{27}*	$[0, 0.0427394]$	$k_{27} \sim U(0.01709576, 0.02564364)$
k_{28}*	$[0, 1527046.0]$	$k_{28} \sim U(610818.4, 916227.6)$
k_{29}*	$[0, 21.3474]$	$k_{29} \sim U(8.53896, 12.80844)$
k_{30}	$[0, 369824.0]$	$k_{30} \sim U(147929.6, 221894.4)$
k_{31}	$[0, 0.1542134]$	$k_{31} \sim U(0.06168536, 0.09252804)$
k_{32}	$[0, 544112.0]$	$k_{32} \sim U(217644.8, 326467.2)$
k_{33}*	$[0, 0.1132558]$	$k_{33} \sim U(0.04530232, 0.06795348)$
k_{34}*	$[0, 1307902.0]$	$k_{34} \sim U(523160.8, 784741.2)$
k_{35}	$[0, 30.2424]$	$k_{35} \sim U(12.09696, 18.14544)$
k_{36}	$[0, 238710.0]$	$k_{36} \sim U(95484.0, 143226.0)$
k_{37}*	$[0, 293.824]$	$k_{37} \sim U(117.5296, 176.2944)$
k_{38}*	$[0, 25752.4]$	$k_{38} \sim U(10300.96, 15451.44)$
k_{39}*	$[0, 2.8029]$	$k_{39} \sim U(1.12116, 1.68174)$
k_{40}	$[0, 21931.2]$	$k_{40} \sim U(8772.48, 13158.72)$
k_{41}*	$[0, 54.53]$	$k_{41} \sim U(21.812, 32.718)$
k_{42}	$[0, 591980.0]$	$k_{42} \sim U(236792.0, 355188.0)$
k_{43}*	$[0, 4.4199]$	$k_{43} \sim U(1.76796, 2.65194)$
k_{44}*	$[0, 2050920.0]$	$k_{44} \sim U(820368.0, 1230552.0)$
k_{45}	$[0, 0.252658]$	$k_{45} \sim U(0.1010632, 0.1515948)$
k_{46}	$[0, 2123.42]$	$k_{46} \sim U(849.368, 1274.052)$
k_{47}	$[0, 882.574]$	$k_{47} \sim U(353.0296, 529.5444)$
k_{48}	$[0, 2.1759 \times 10^7]$	$k_{48} \sim U(8703600.0, 1.30554 \times 10^7)$

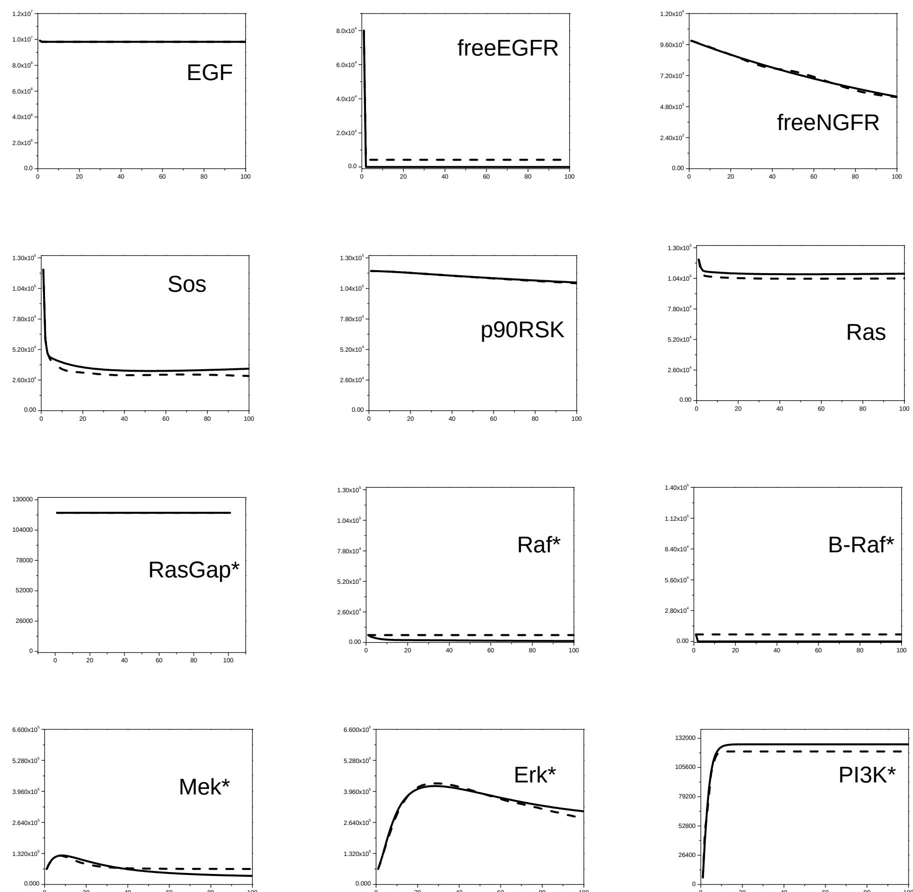

Fig. 1. Simulation results of NGF-EGF signaling pathway **Part I**. Solid lines represent nominal profiles and dashed lines represent BN simulation profiles.

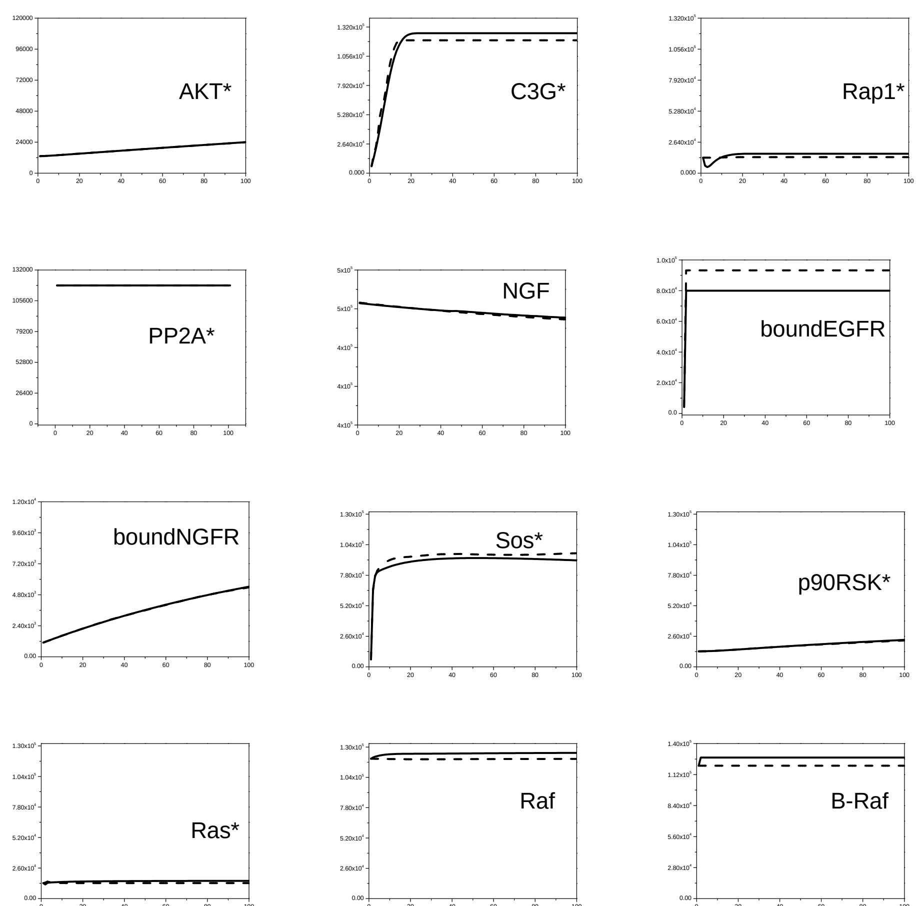

Fig. 2. Simulation results of NGF-EGF signaling pathway **Part II**. Solid lines represent nominal profiles and dashed lines represent BN simulation profiles.

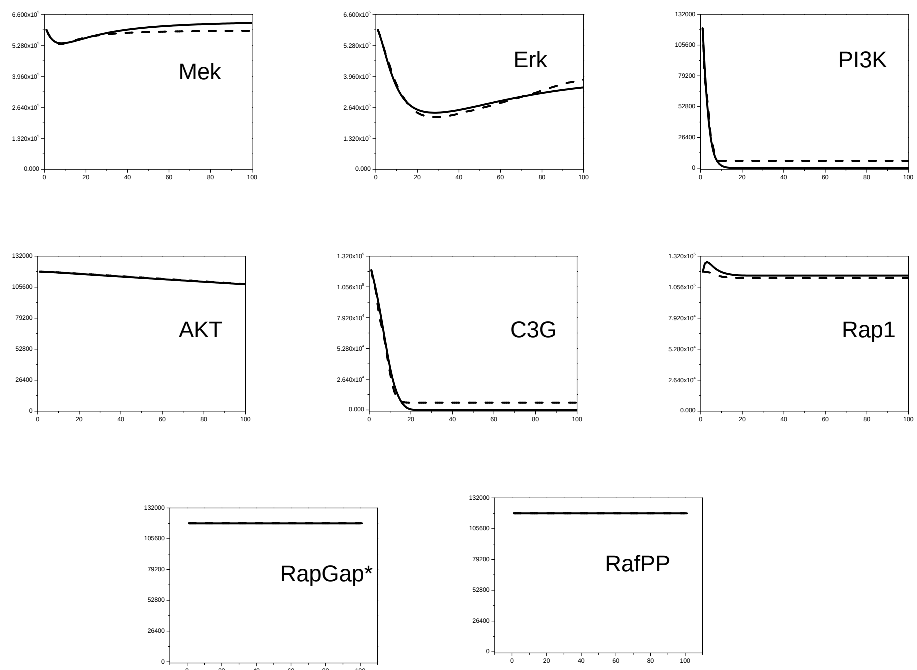

Fig. 3. Simulation results of NGF-EGF signaling pathway **Part III**. Solid lines represent nominal profiles and dashed lines represent BN simulation profiles.

Table 4. Parameter estimation results. The posterior distributions of unknown parameters inferred by our method.

Parameter	Range	Posterior probability distribution
k_1*	$[0, 4.37006 \times 10^{-5}]$	$k_1 \sim U(2.62204E \times 10^{-5}, 3.49605 \times 10^{-5})$
k_2*	$[0, 0.0242016]$	$k_2 \sim U(0.01452096, 0.01936128)$
k_3*	$[0, 2.76418 \times 10^{-7}]$	$k_3 \sim U(1.65851 \times 10^{-7}, 2.21134 \times 10^{-7})$
k_4*	$[0, 0.01447622]$	$k_4 \sim U(0.011580976, 0.01447622)$
k_{11}*	$[0, 64.688]$	$k_{11} \sim U(38.8128, 51.7504)$
k_{12}*	$[0, 71908.6]$	$k_{12} \sim U(28763.44, 43145.16)$
k_{15}*	$[0, 1.768192]$	$k_{15} \sim U(1.4145536, 1.7681922)$
k_{17}*	$[0, 371.518]$	$k_{17} \sim U(74.3036, 148.6072)$
k_{23}*	$[0, 19.70734]$	$k_{23} \sim U(7.882936, 11.824404)$
k_{27}*	$[0, 0.0427394]$	$k_{27} \sim U(0, 0.00854788)$
k_{28}*	$[0, 1527046]$	$k_{28} \sim U(0, 305409.2)$
k_{29}*	$[0, 21.3474]$	$k_{29} \sim U(0, 4.26948)$
k_{33}*	$[0, 0.1132558]$	$k_{33} \sim U(0.06795348, 0.09060464)$
k_{34}*	$[0, 1307902]$	$k_{34} \sim U(784741.2, 1046321.6)$
k_{37}*	$[0, 293.824]$	$k_{37} \sim U(117.5296, 176.2944)$
k_{38}*	$[0, 25752.4]$	$k_{38} \sim U(20601.92, 25752.4)$
k_{39}*	$[0, 2.8029]$	$k_{39} \sim U(2.24232, 2.8029)$
k_{41}*	$[0, 54.53]$	$k_{41} \sim U(43.624, 54.53)$
k_{43}*	$[0, 4.4199]$	$k_{43} \sim U(3.53592, 4.4199)$
k_{44}*	$[0, 2050920]$	$k_{44} \sim U(1230552, 1640736)$

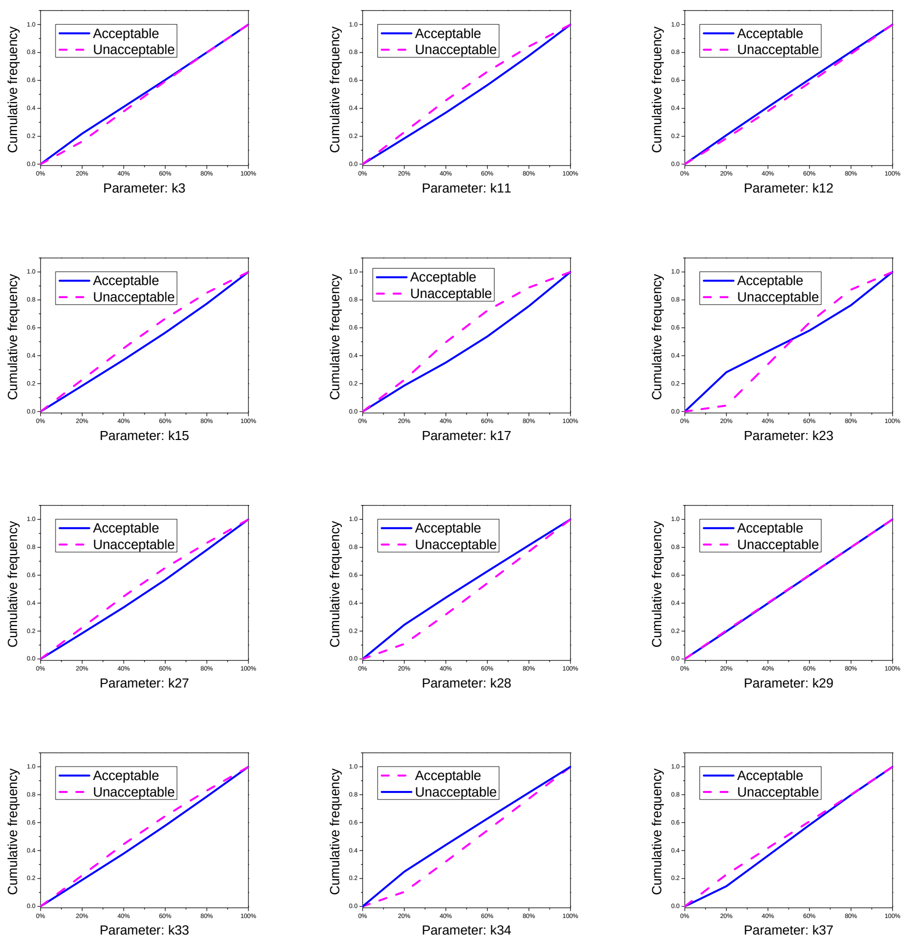

Fig. 4. Sensitivity Analysis Results: Cumulative frequency distributions of the MPSA with respect to the unknown parameters. Solid line denotes the acceptable samples and the dashed line indicates the unacceptable samples. The sensitivity of a parameter is defined as the maximum vertical difference between its two curves (K-S statistic) for the parameter.

3 Space Complexity

The space complexity of our BN model is $O(\hat{d}\hat{n}K^{P+1})$, where $\hat{d}$ is the number of time points, $\hat{n}$ is the number of nodes, K is the maximum discrete value and P is the maximal number of parents a node can have. In particular, the EGF-NGF model requires 147MB space for explicit storage. However, the conditional probabilities tables are sparse. When we accordingly adopt a sparse implementation, it only requires 283KB.

Furthermore, to deal with nodes with a large number of parents P, we can reduce the space required by breaking down 'fat' nodes as shown in Figure 5.

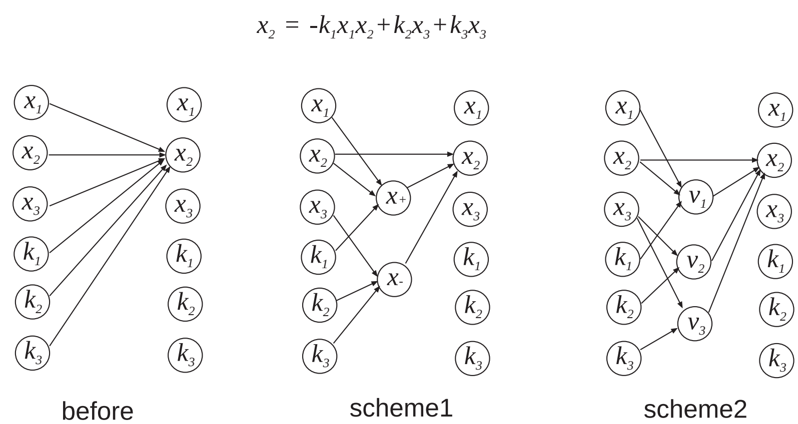

Fig. 5. Reduction examples.

References

1. Ammann, H.: Ordinary Differential Equations: An Introduction to Nonlinear Analysis. Walter de Gruyter (1990)
2. Feldman, J.: Review of measurable functions. University of Brithish Columbia
3. Durrett, R.: Probability: Theory and Examples. Duxbury Press (2004)