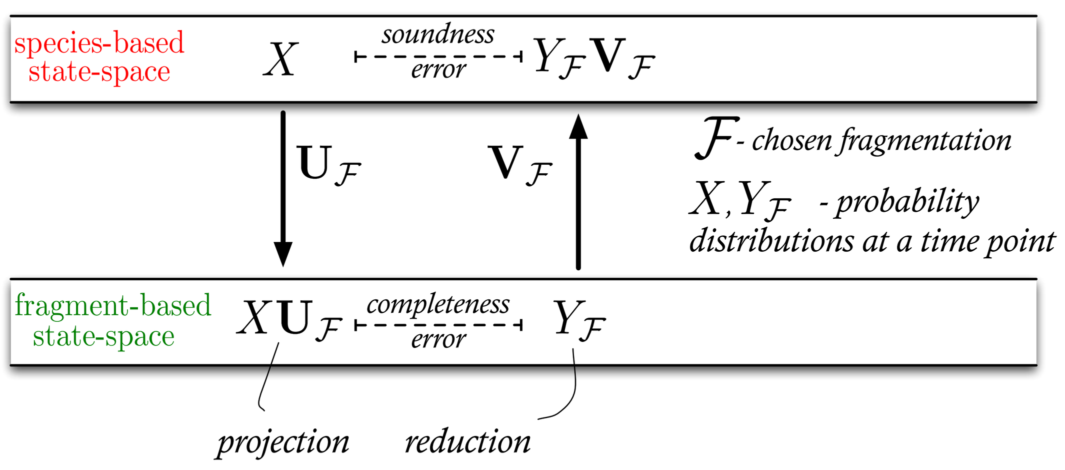
Approximate reductions of rule-based models

Tatjana Petrov, Heinz Koeppl, Jérôme Feret

Problem

In recent works, general algorithms for the exact reductions of rule-based models were established. However, especially in the stochastic setting, the reduced state space often remains combinatorially large.
Can we quantitatively study the effect of approximate reductions of rule-based models?



Contributions

1. A procedure for running approximate reductions over a formalism of site-graph-rewrite rules;
2. Error measure based on KL divergence
   - ✓ point-wise and trace distributions for discrete time (can be extended to continuous-time)
3. Case studies:
   - ✓ A sound and complete reduction with exponentially smaller state-space,
   - ✓ Convergence of completeness error for random initial distributions.

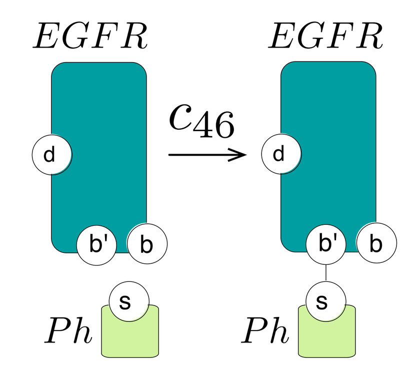
Method

Site-graph-rewrite rules

The cellular signalling program is executed by applying the rules according to Gillespie's algorithm.

$\mathtt{EGFR(b,b',d),Ph(s)} \rightarrow \mathtt{EGFR(b,b'!1,d),Ph(s!1)}@c_{46}$

rule condition — rule update — rule rate



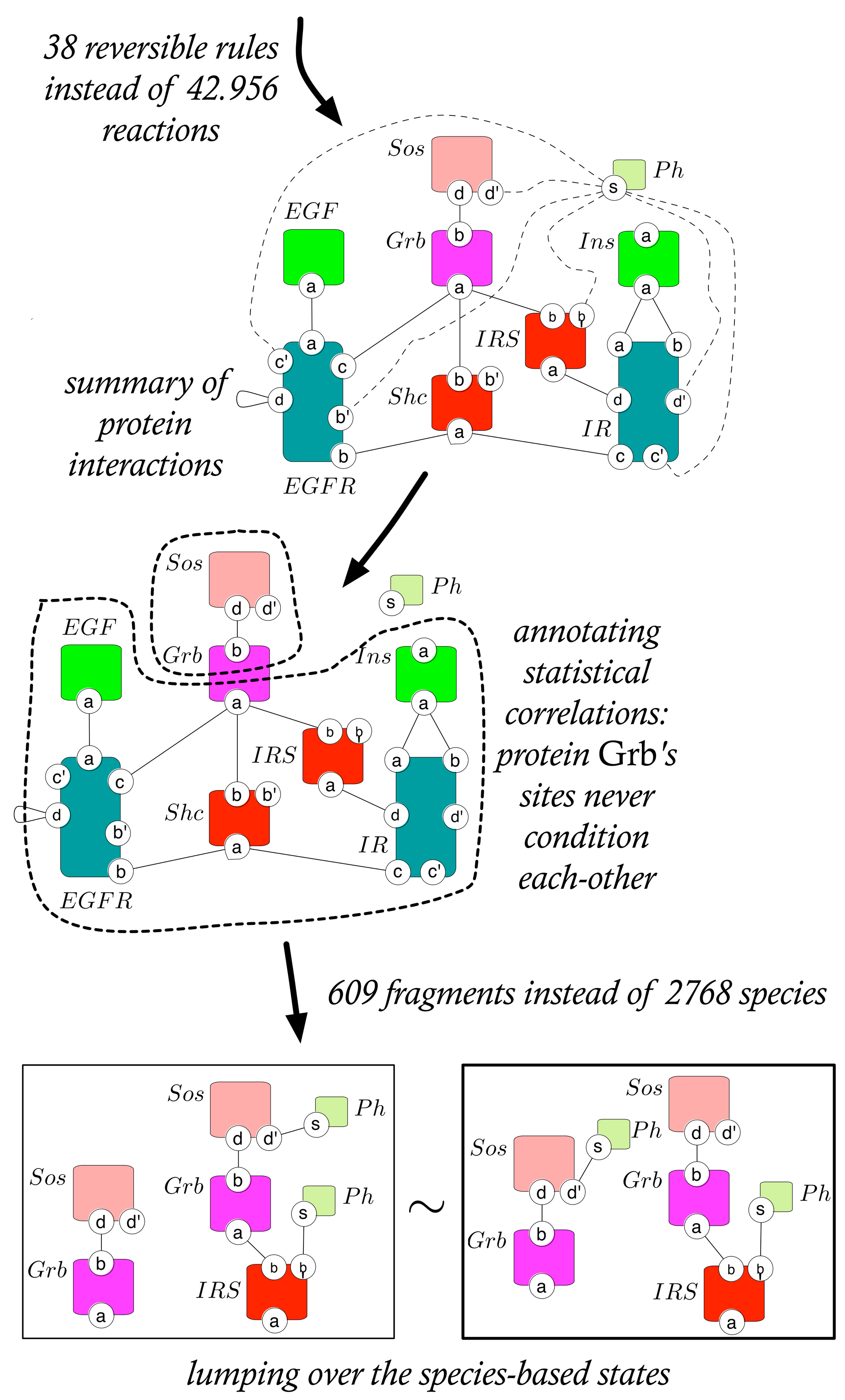
nodes: EGFR, Ph;
sites: s,d,b,b';
edges: ((Ph,s),EGFR,b')

The internal sites are encoded as binding events. For example, phosphorylation is incorporated in the model by binding a phosphate group Ph, which is assumed to be highly abundant.

Fragmentation: overview

A mechanistic model of the network of EGF and insulin receptor crosstalk is taken from [2].

$$
\begin{array}{ll}
01: & \mathtt{IR(a^1,b),Ins(a^1)} \rightarrow \mathtt{IR(a,b),Ins(a)}\\
02: & \mathtt{IR(a,b),Ins(a)} \rightarrow \mathtt{IR(a^1,b),Ins(a^1)}\\
03: & \mathtt{IR(a,b^1),Ins(a^1)} \rightarrow \mathtt{IR(a,b),Ins(a)}\\
 & \ldots\\
75: & \mathtt{EGFR(c_p^1,d),Grb(a^1)} \rightarrow \mathtt{EGFR(c_p,d),Grb(a)}\\
76: & \mathtt{EGFR(c_p,d),Grb(a)} \rightarrow \mathtt{EGFR(c_p^1,d),Grb(a^1)}
\end{array}
$$



Error measure

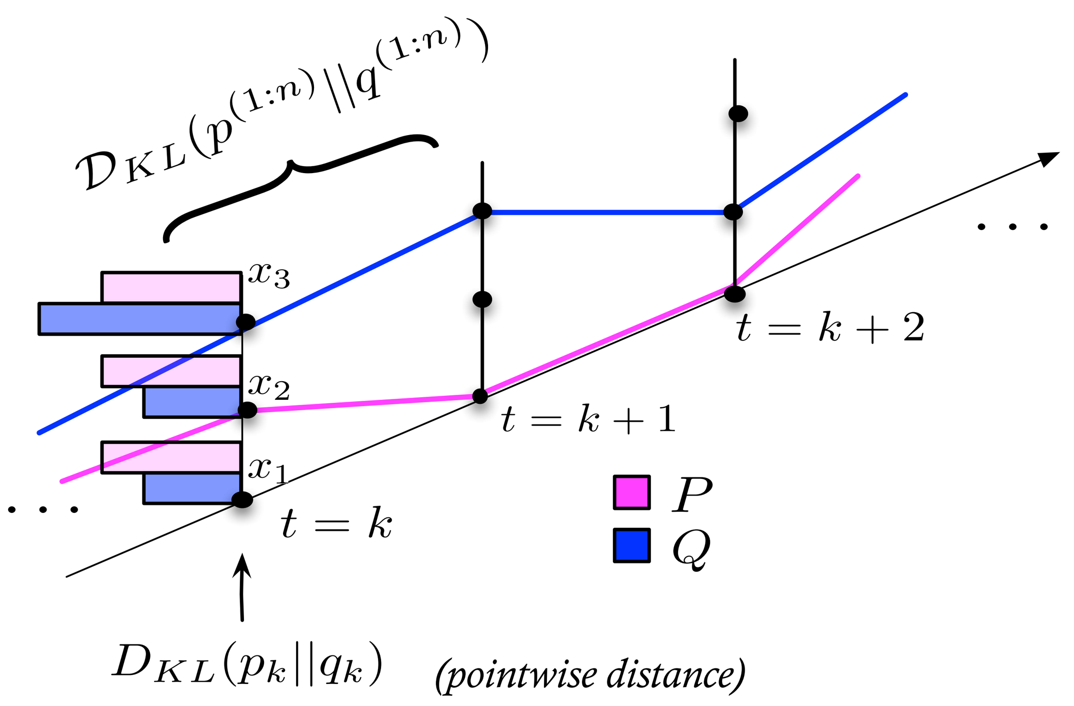
Given a Markov chain and a lumping relation on its state space, **lumpability** [1] guarantees that the projected process remains Markovian, and that the sound (and complete) reduction is possible.
KL-divergence (relative entropy) between probability distributions p and q, over the state space $\mathcal{S} = \{x_1, \ldots, x_M\}$:

$$D_{KL}(p||q) = \sum_{x\in\mathcal{S}} p(x)\log\frac{p(x)}{q(x)}.$$

KL-divergence between distributions over traces of length n, generated by two discrete-time Markov chains (DTMC's) with transition matrices P and Q, such that P is absolutely continuous with respect to Q:

$$\mathcal{D}_{KL}(p^{(1:n)}||q^{(1:n)}) = \sum_{t=1}^{n}\sum_{x\in\mathcal{S}} p_t(x)f(x)$$

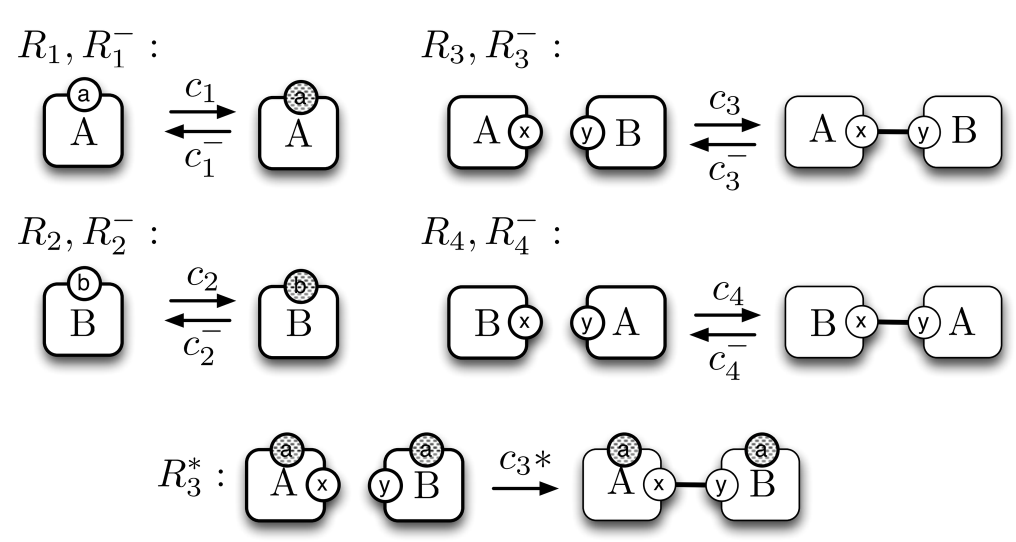
where $f(x) = \sum_{x_j\in\mathcal{S}} P(x,x_j)\log\frac{P(x,x_j)}{Q(x,x_j)}$.



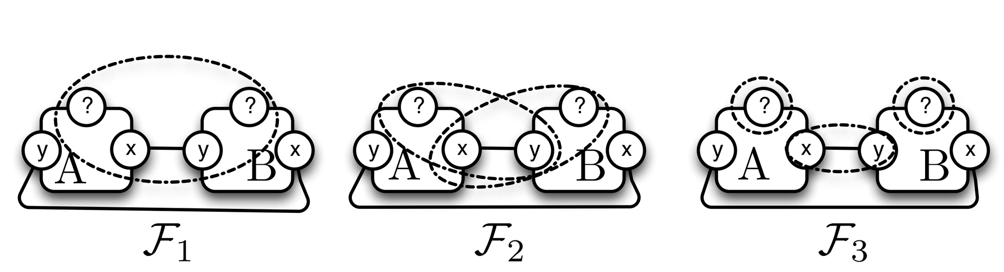
KL-divergance rate [3] between two DTMC's over the state space $\mathcal{S}$, given by transition matrices P and Q, with a unique stationary distribution π, and such that P is absolutely continuous with respect to Q:

$$\mathbf{D}_{KL}(P||Q) = \lim_{n\to\infty}\frac{1}{n}\mathcal{D}_{KL}(p^{(1:n)}||q^{(1:n)}) = \sum_{x\in\mathcal{S}}\pi(x)f(x).$$

Case study

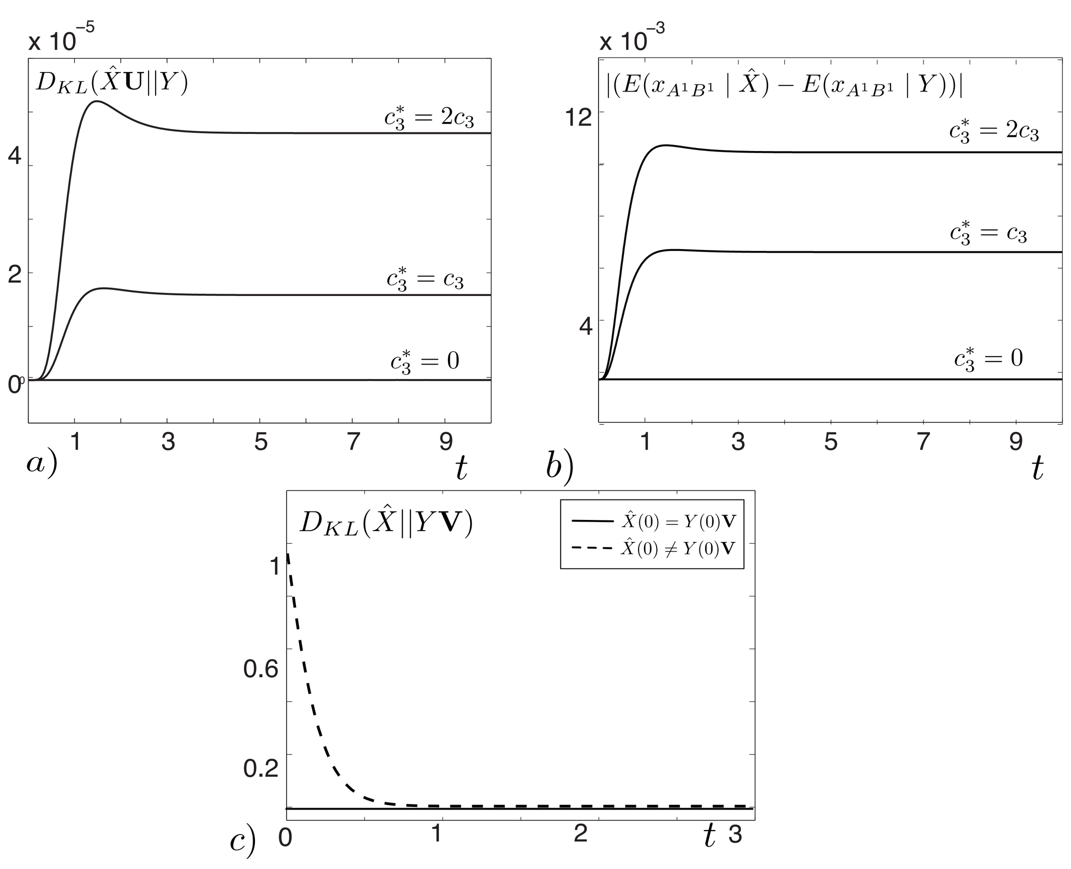


Three candidate fragmentations:



Summary of reduction properties for $c_3^* \in \{0, c_3\}$.

$n_A = n_B = n$	$\mathcal{S}$	$\mathcal{F}_1$	$\mathcal{F}_2$
dim. of st. sp.	$O(4^n)$	7	6
size of st. sp.	$O(4^n)$	$O(n^7)$	$O(n^6)$
soundness error, $c_4 = 0$			
$c_3^* = 0, \Omega = 10$	0	0	0
$c_3^* = c_3, \Omega = 10$	0	0([1])	(a)
$c_3^* = c_3, \Omega \to \infty$	0	0([1])	0([4])



Simulation results: a) the soundness error, b) the absolute error between the correct and approximated rate, c) the decay of completeness error.

References

[1] J. Feret, T. Henzinger, H. Koeppl, T. Petrov: Lumpability Abstractions of Rule-based Systems. In *MeCBIC, 2010*

[2] H. Conzelmann, D. Fey, E. D. Gilles: Exact model reduction of combinatorial reaction networks. In *BMC Systems Biology 2, 2008*

[3] K. Deng, P. G. Mehta, S. P. Meyn: Optimal Kullback-Leibler Aggregation via Spectral Theory of Markov Chains. In *IEEE Transactions on Automatic Control, 2011*

[4] J. Feret. V. Danos, J. Krivine, R. Harmer, W. Fontana: Internal Coarse-graining of Molecular Systems. In *PNAS, 2009*

Future work

Extension of the error measure to the continuous-time case; Developing methods for error estimation without executing the species-based model; Estimating error for ODE fragments.