A Structured Approach…
Part IV

Model Checking in Systems and Synthetic Biology

Robin Donaldson

Bioinformatics Research Centre University of Glasgow, Glasgow, UK

Outline

- Introduction to model checking
- Our approach
  - Simulation-based Model Checking
  - PLTLc
  - MC2(PLTc) model checker

- Results
  - Stochastic analysis of MAPK pathway
  - Continuous parameter scan analysis

Model Checking

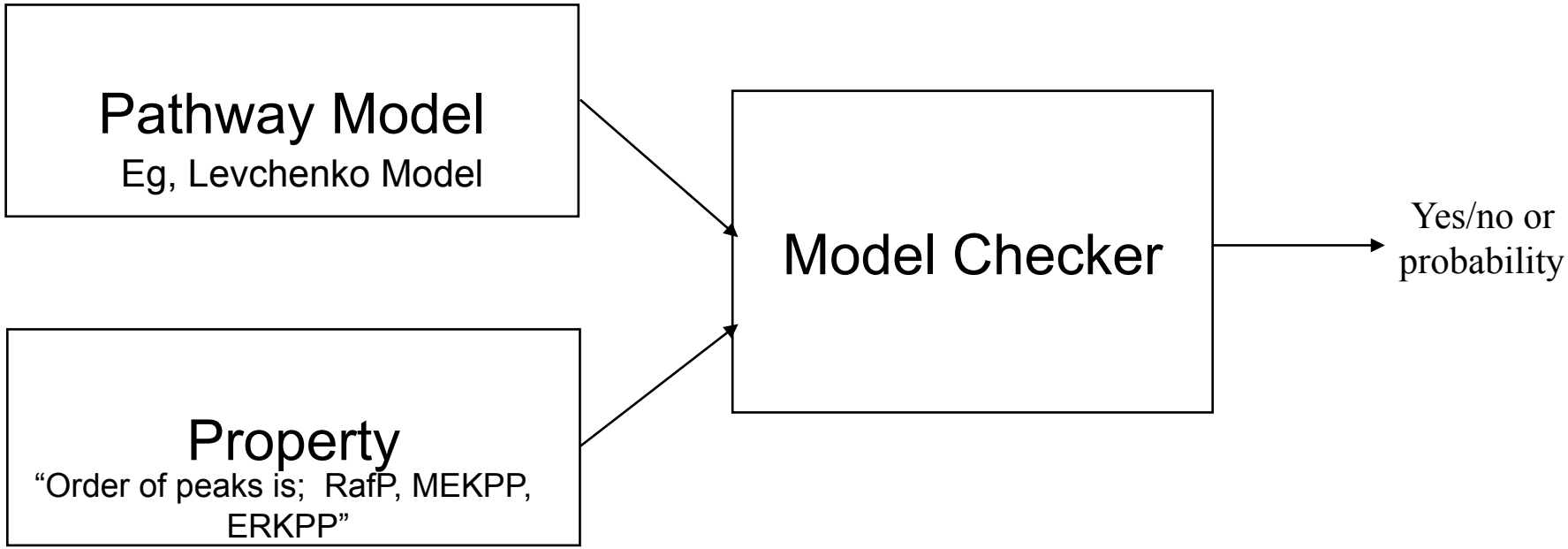
In a sentence:

- "Formally check whether a model of a biochemical system does what we want"

Components:

- A model
  - the current description of a biochemical system of interest
- A property
  - a property which we think the system should have
- A model checker
  - a program to test whether the model has the property

Model Checking Biochemical Pathways



Why model check in Systems Biology

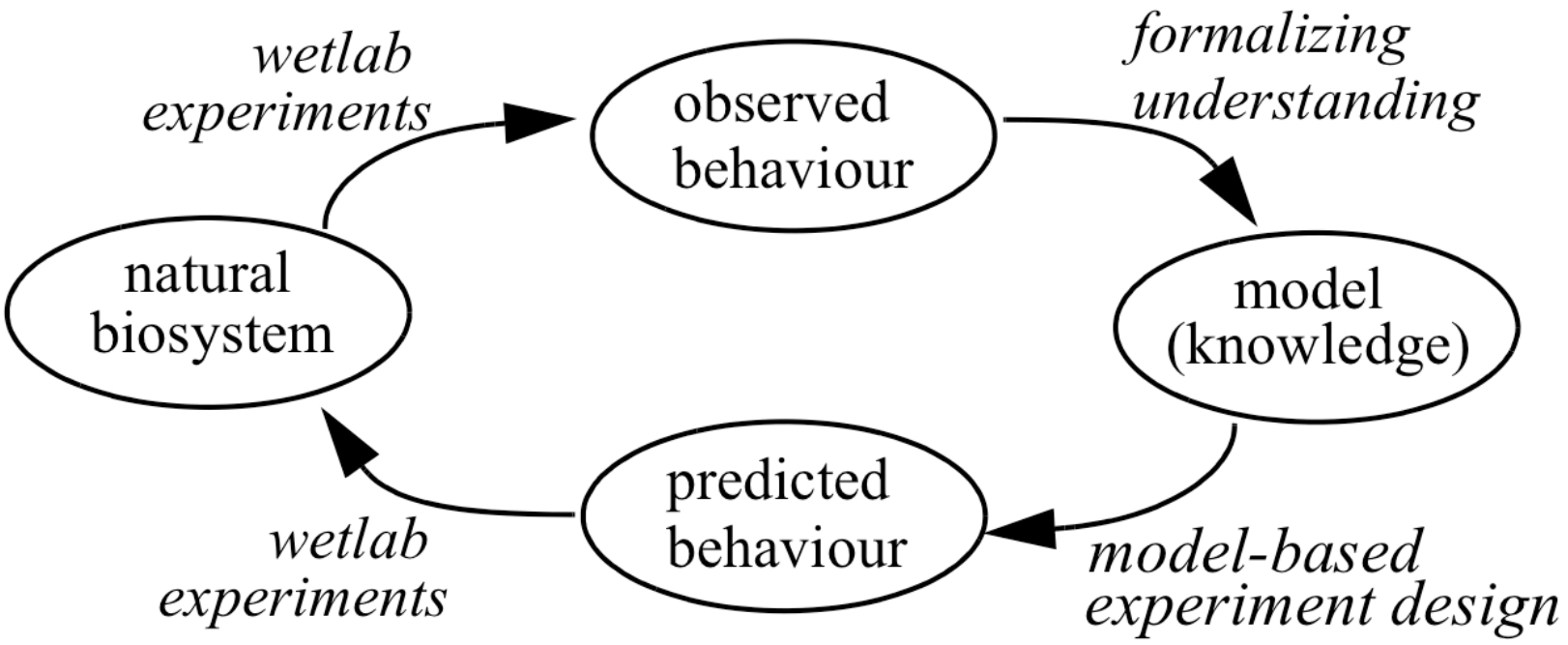
- Biologists will often talk in qualitative or semi-quantitative language (trends).
  - "this protein peaks after 5 seconds, then falls to half concentration"
  - Often quite certain about time.

- Systems biology; Part of model design process, validate the model conforms to the **observed** data.

- Synthetic biology; Make sure the model and constructed bio system conform to the **desired** behaviour.

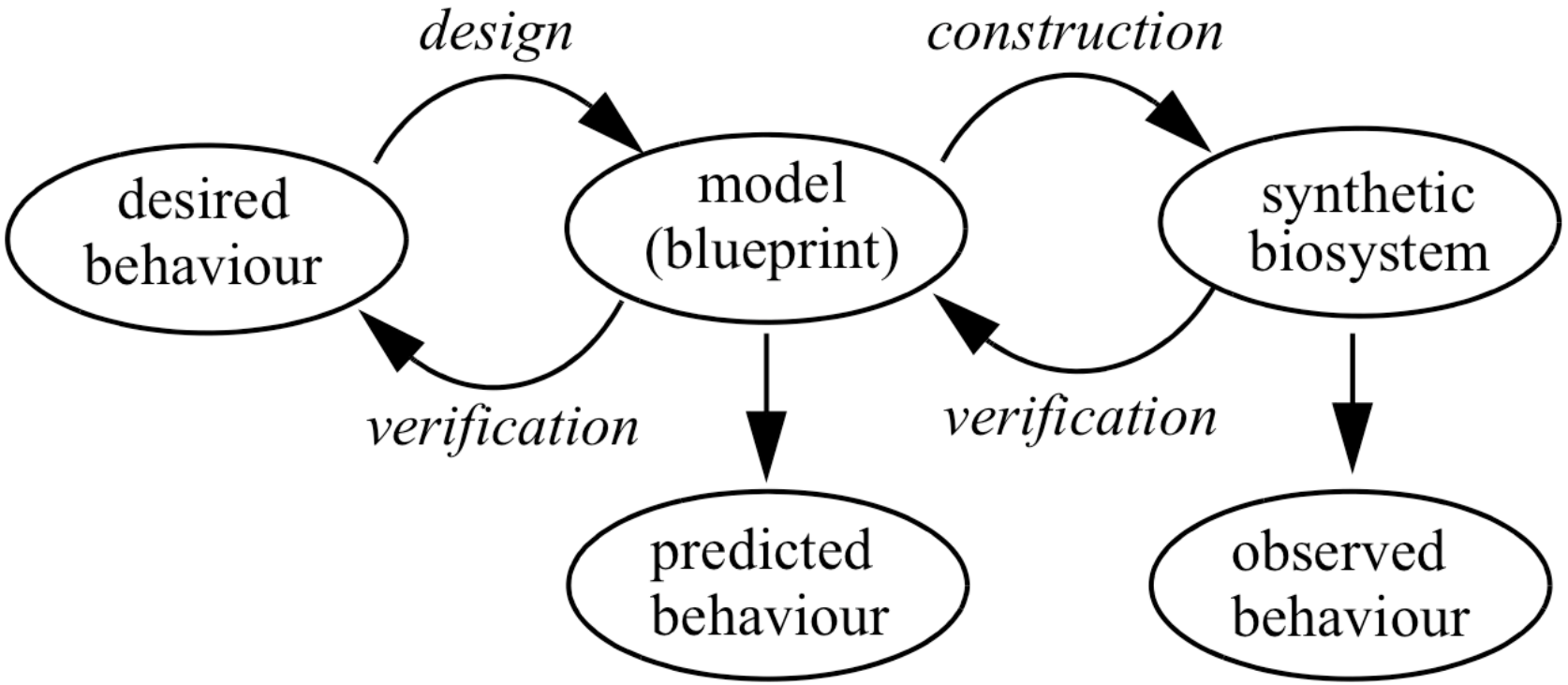
Systems Biology

systems biology: modelling as formal knowledge representation



Synthetic Biology

synthetic biology: modelling for system construction



Analytical vs Simulative Model Checking

- Analytical:
  - Exact probabilities
  - Prove properties
  - A model which is large or has many molecules has a large state space
  - State space = possible model states = every combination of # molecules to species.
  - State space grows exponentially
  - Model checking with even as little as 12 molecules can be impossible with today's technology

- Solve this problem using…
- Simulative:
  - Analyse a subset of the state space
  - Approximate probabilities
  - Can't prove

Our approach

- Language
  Uses a Probabilistic language called Probabilistic Linear-time Temporal Logic (PLTL)

  Extend this language with numerical constraints (essentially variables in a property). New language PLTL with Numerical Constraints (PLTLc)

- Model Checker:
  Developed an ***offline*** Monte Carlo Model Checker for PLTLc properties, MC2(PLTLc) for short.

Our approach

- Operates on a finite set of simulations – simulative approach
- Typically, many stochastic simulations (1,000 in this presentation)

  Probability = fraction of simulations which satisfy the property over the #simulations
- Monte Carlo approximation – 2 approximations made:
  - finite paths
  - finite set of paths

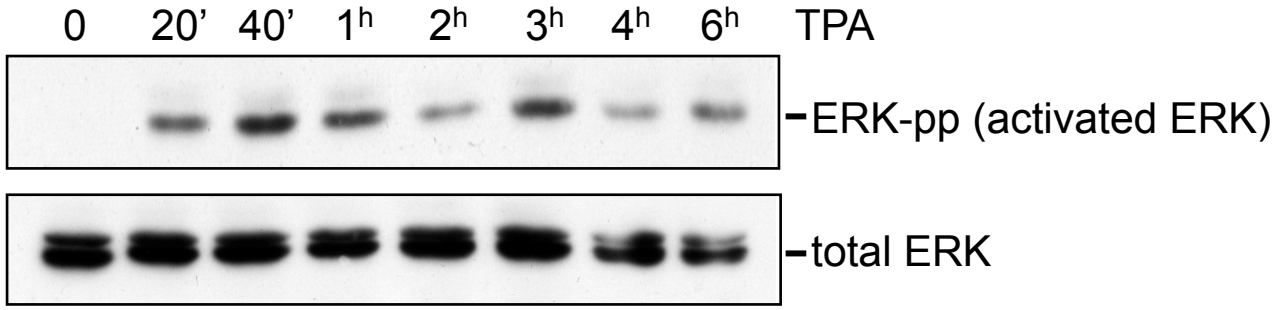
Our approach

- The set of traces can be:
  - Set of stochastic runs
  - A single deterministic run
  - A parameter scan
  - Lab data!

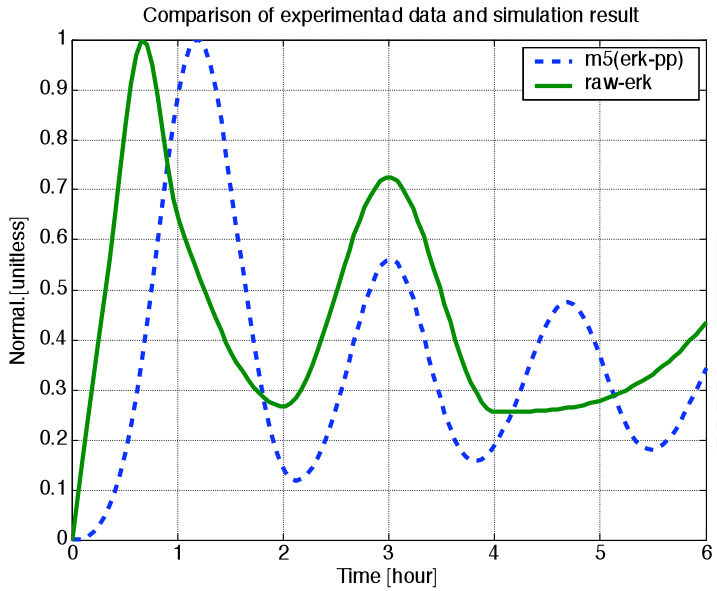
- We could use simulation output from;
  - ODE, SDE, CTMC, Gillespie, hybrid approaches

- Or experimental data from the wet lab

Our approach



Western blots COS1 cell lysates



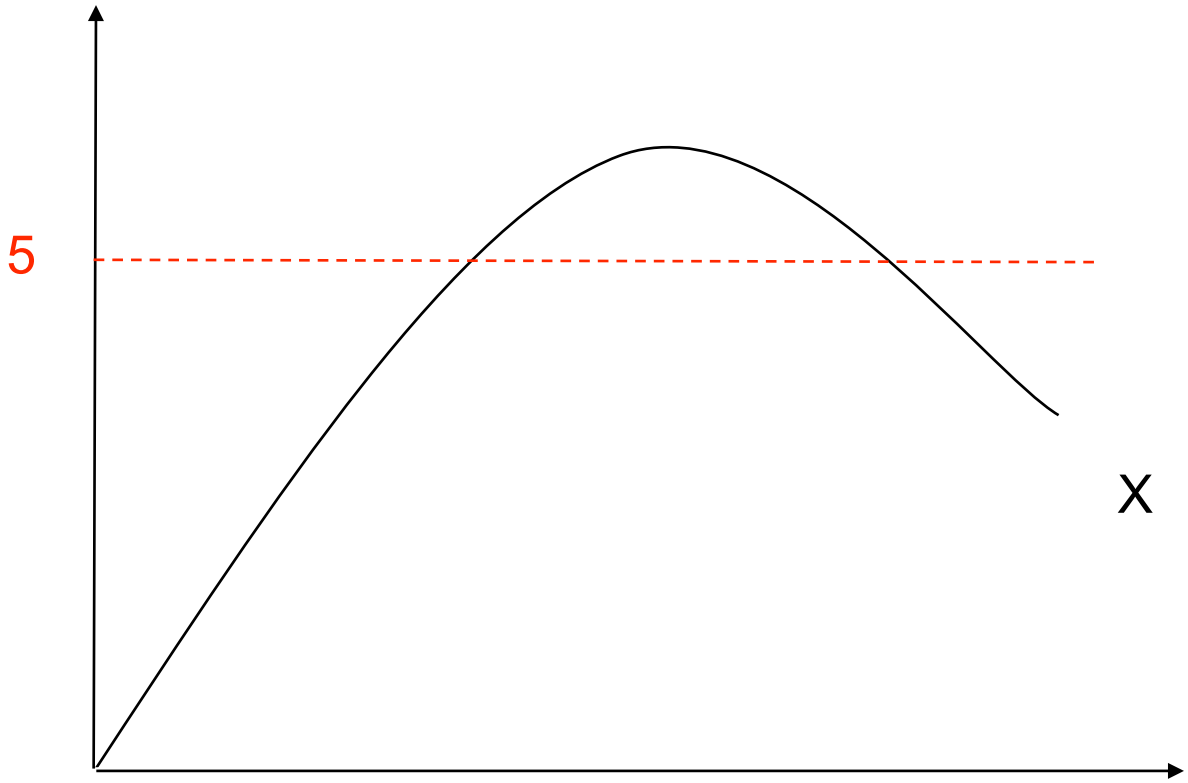
Experiment

Simulation

MC2 with ODE Output

$P_{=?}[\ F(\ X > 5\ )\ ]$

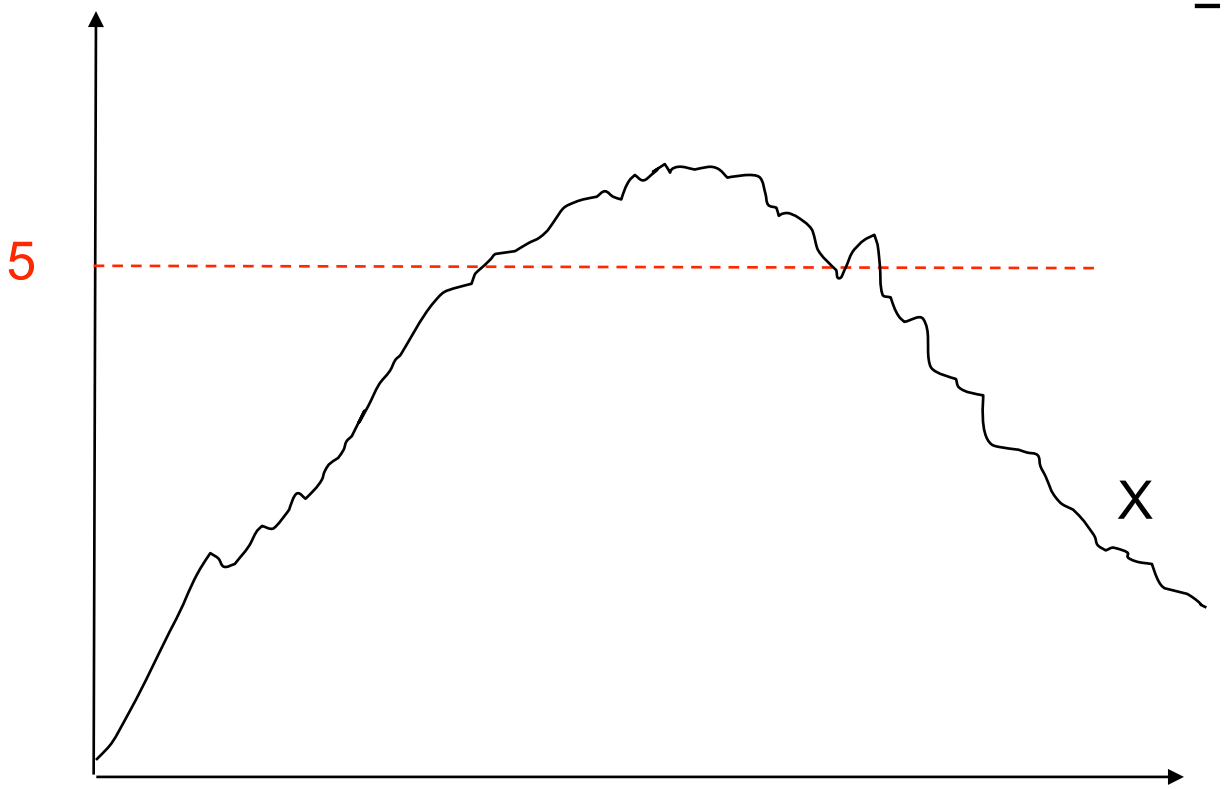
=> P = 1



MC2 with Gillespie Output

$P_{=?}[\ F(\ X > 5\ )\ ]$

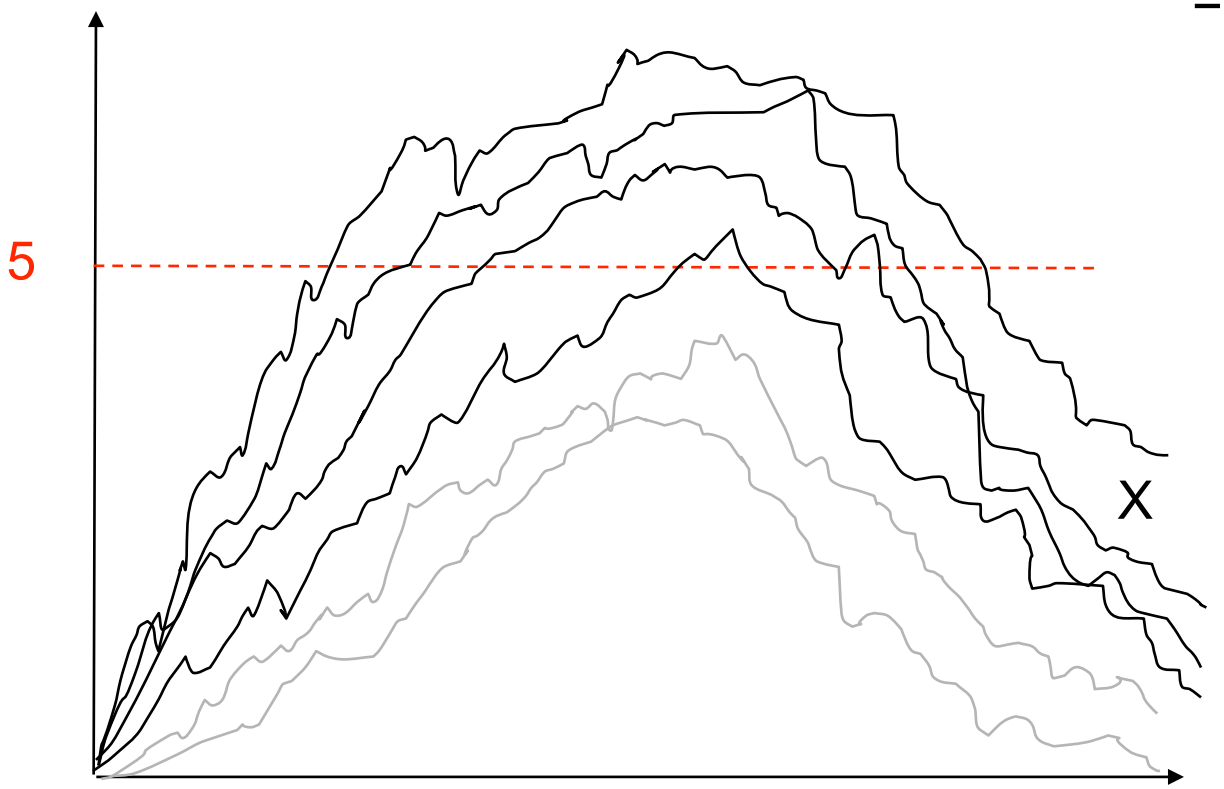
=> P = 1



MC2 with Gillespie Output

$P_{=?}[\ F(\ X > 5\ )\ ]$

=> P = 4/6



Implementation

- MC2 implemented in Java, available for download now.
- Model checks a set of simulation traces – parallelisable!
- Operate over the BRC cluster comprising 45 Sun X2200 servers (180 cores)
- 180 simulation traces checked in the same time as 1

RESULTS

Two Reaction Model

- First a simple model of two reactions:

$$A \xrightarrow{0.01} B$$

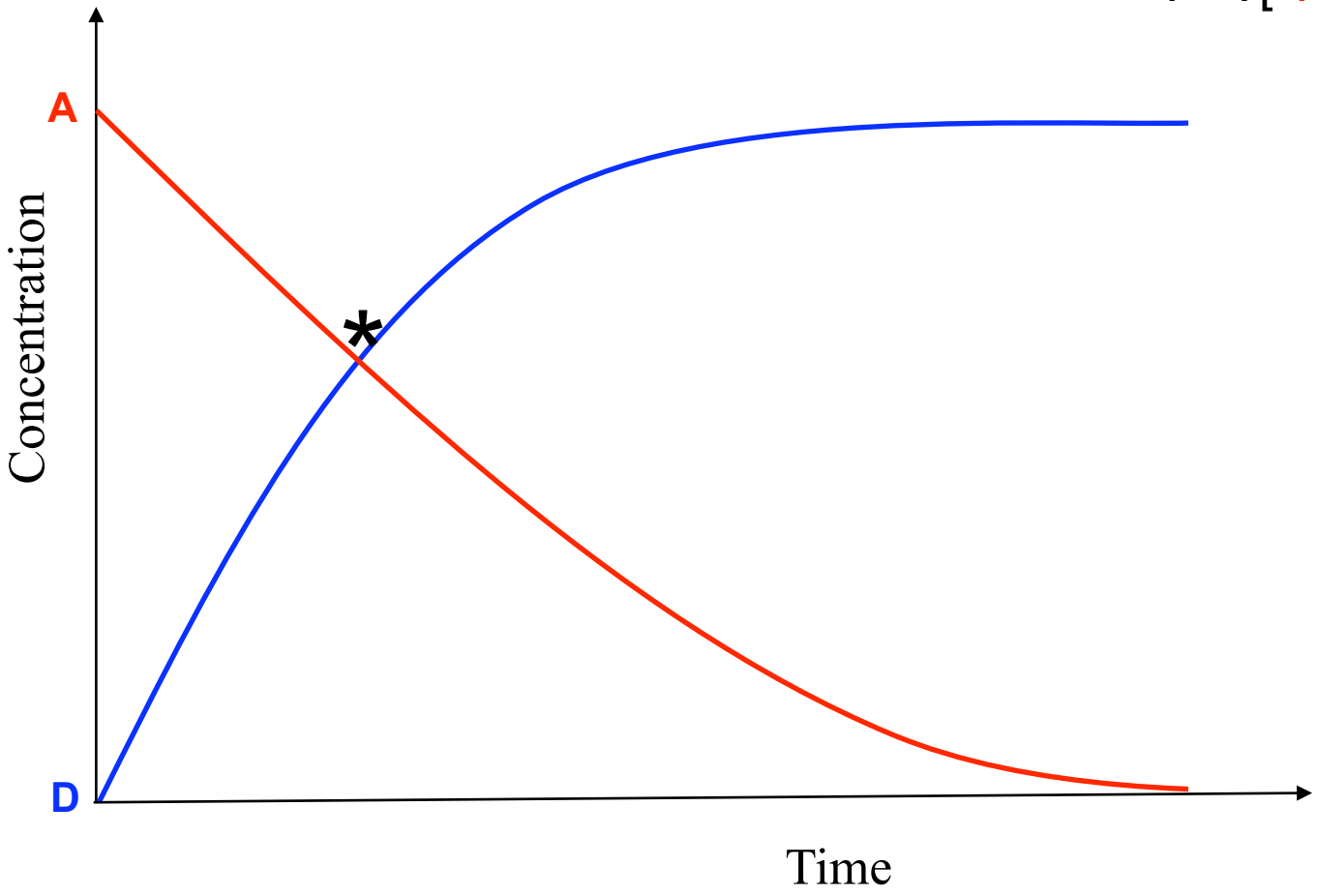
$$C \xrightarrow{0.1} D$$

- Assess property:
  P=?[A = $X { A = D }]

- “What is the probability that when A and D first equal each other, they both have X number of molecules?”

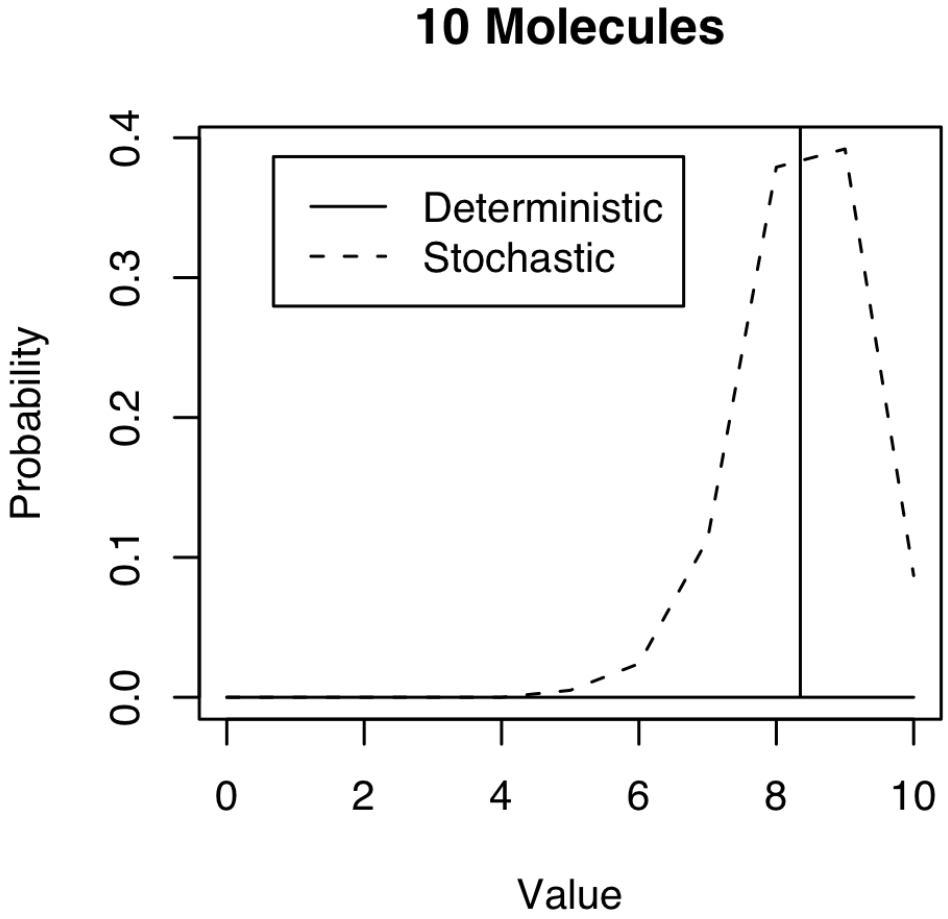
Two Reaction Model

Property:
P=?[A = X { A = D }]

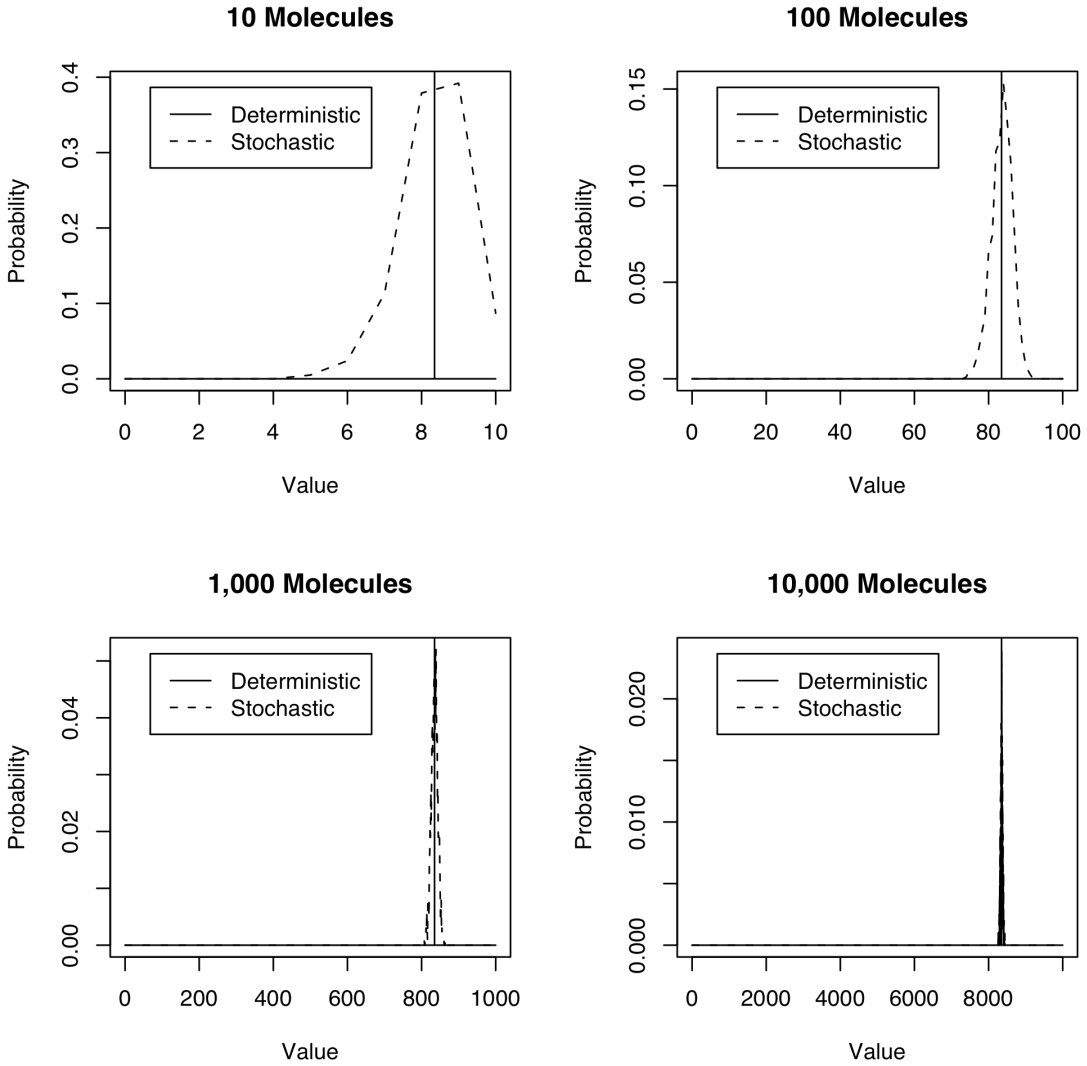


Two Reaction Model

- Set reactants to 10 molecules (model bound to 10 molecules)
- Simulate with Gillespie 1,000 times and model check each output
- Number of simulations which are true over total number of simulations is the probability.

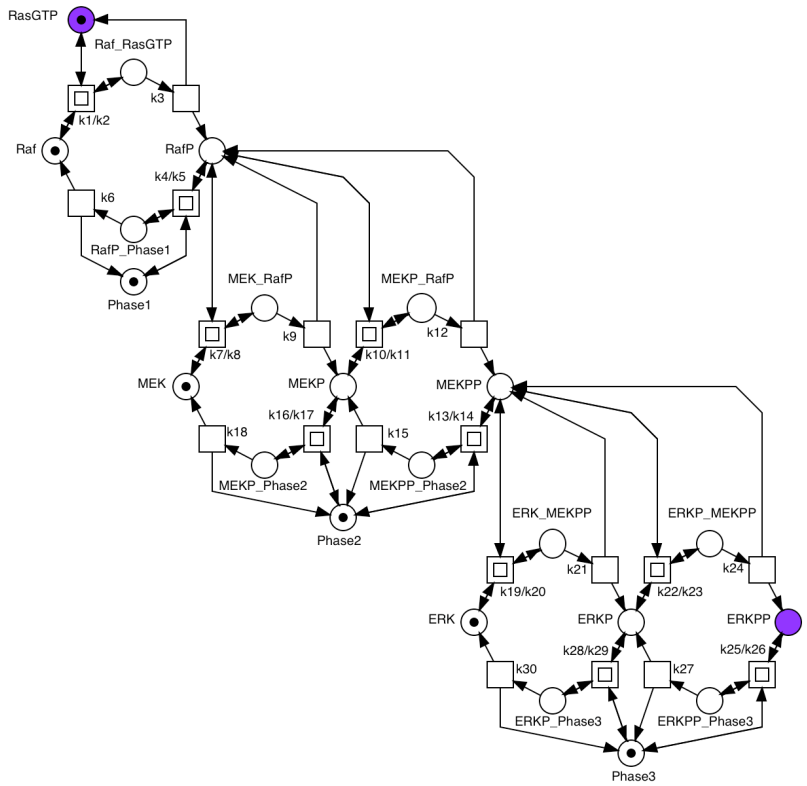


Two Reaction Model



MAPK Pathway - Levchenko

- Mitogen Activated Protein Kinase (MAPK) Pathway
- Obligatory pathway for model checking
- Use the Levchenko *et al.* model of the pathway and check two properties, S1 and S2.
- Have used MC2(PLTLc) for continuous (ODE) analysis too.

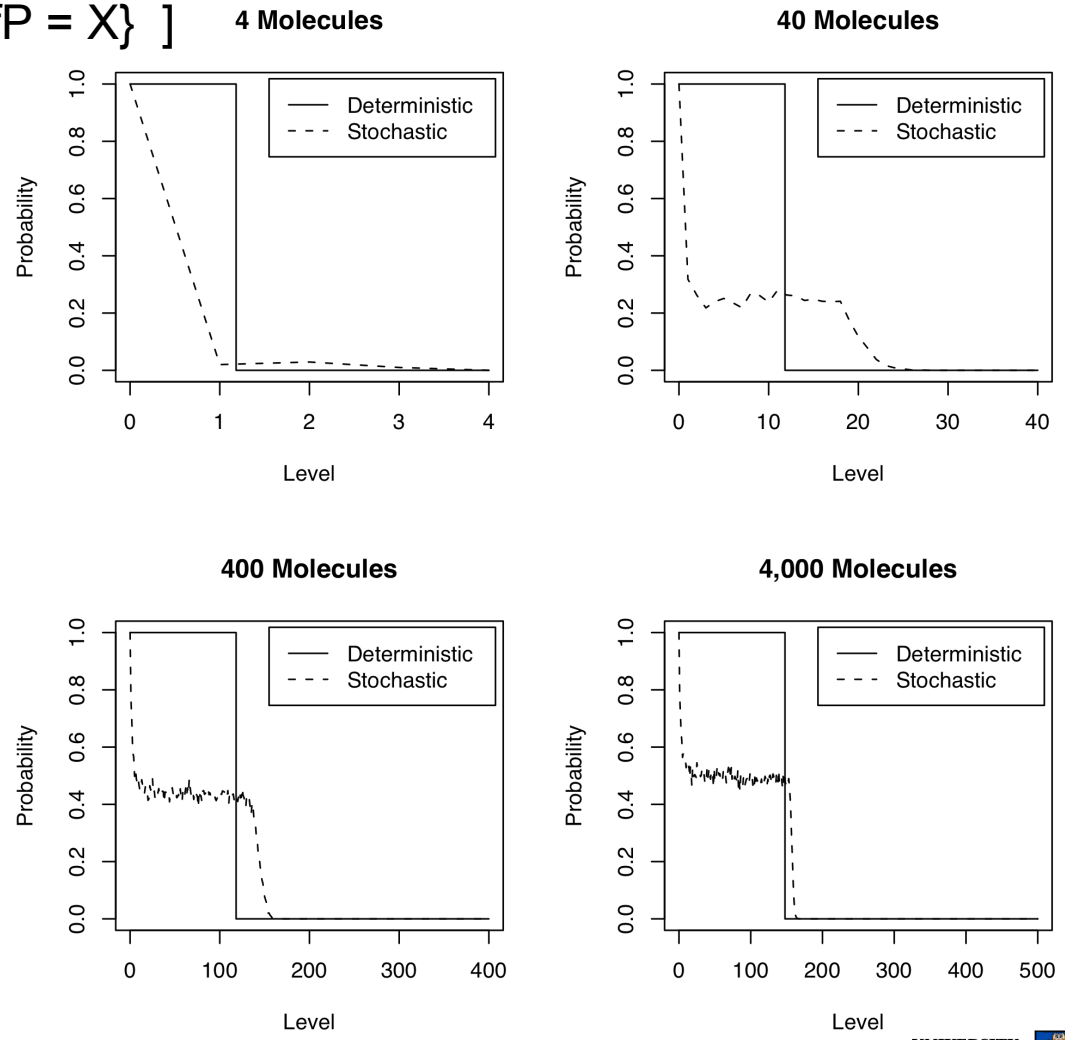


Levchenko *et al*. "Scaffold proteins may biphasically affect the levels of mitogen-activated protein kinase signaling and reduce its threshold properties". Proc Natl Acad Sci USA, 97(11):5818–5823, 2000.

radonald@brc.dcs.gla.ac.uk

MAPK Pathway - Levchenko

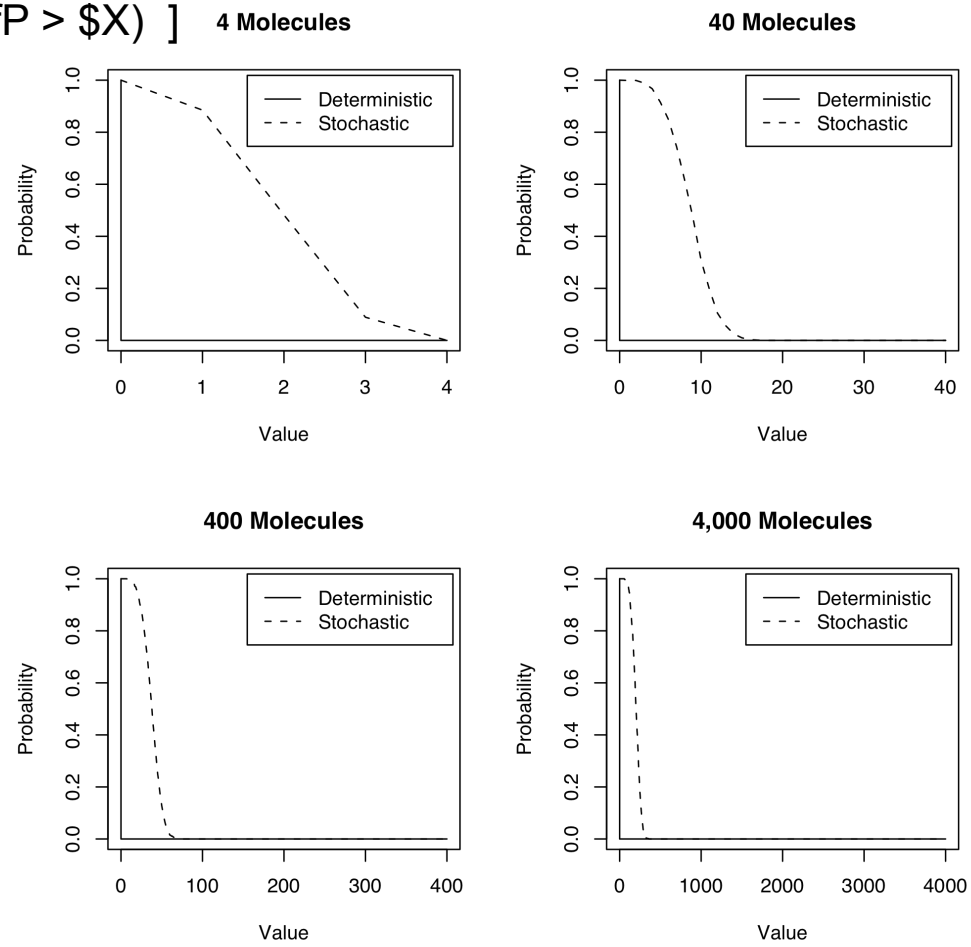
- S1: "What is the probability that RafP will increase when at a concentration X?"
  $P_{=?}[\ (RafP = X)\ U\ (RafP > X)\ \{RafP = X\}\ ]$



radonald@brc.dcs.gla.ac.uk

MAPK Pathway - Levchenko

- S2: “What is the probability that RafP will reach concentration X while MEKPP and ERKPP remain at 0?”
  $P_{=?}[\ (MEKPP = 0\ \wedge\ ERKPP = 0)\ U\ (RafP > \$X)\ ]$

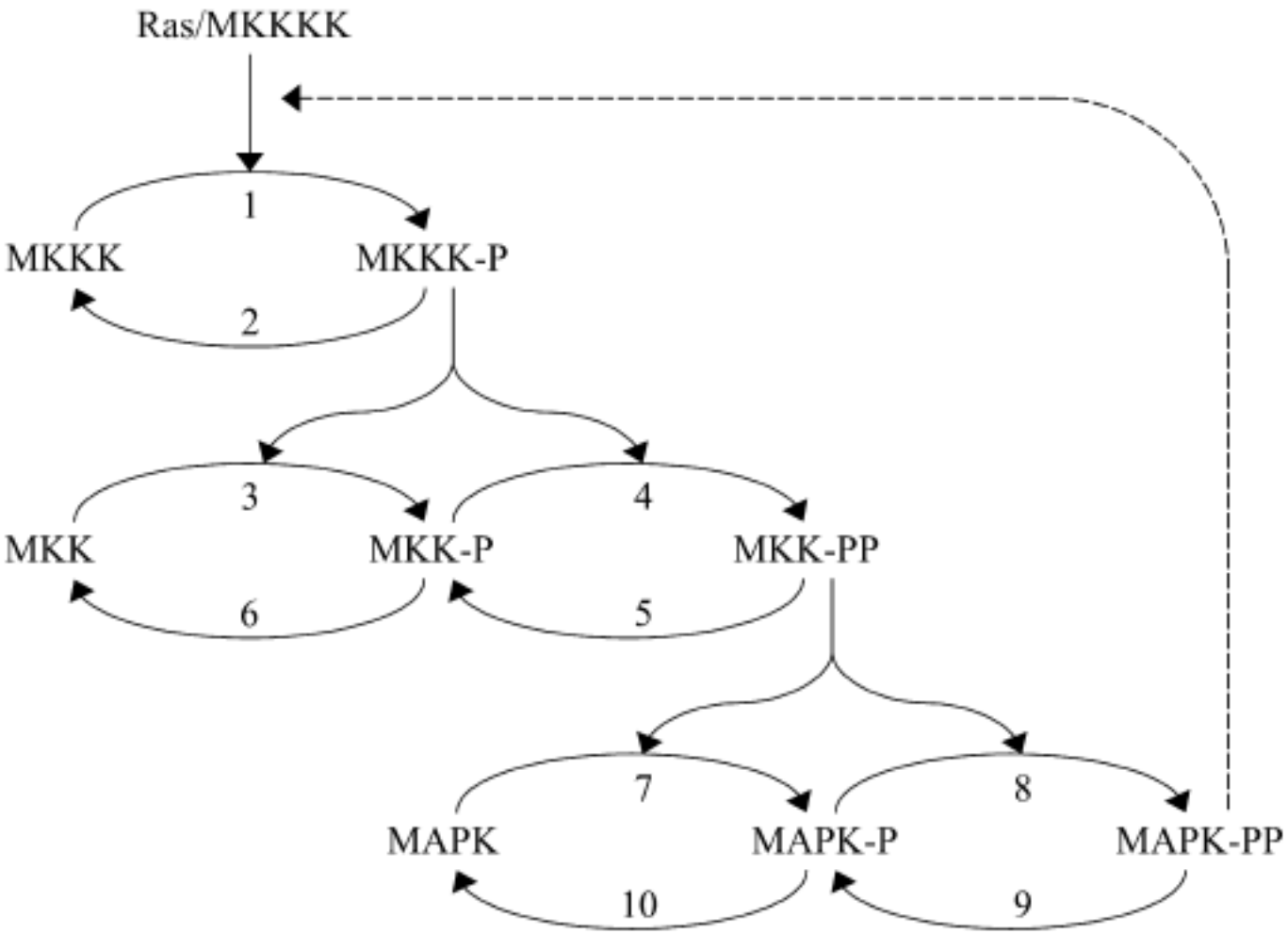


MAPK Pathway - Levchenko

Molecules	Property S1	Property S2
4	6s	10s
40	70s	78s
400	6 mins*	3 mins*
4,000	8 hrs*	20 mins*

- * ran over the cluster
- Property S2 is much faster than S1 because S2 exploits the use of variables.

Kholodenko Model



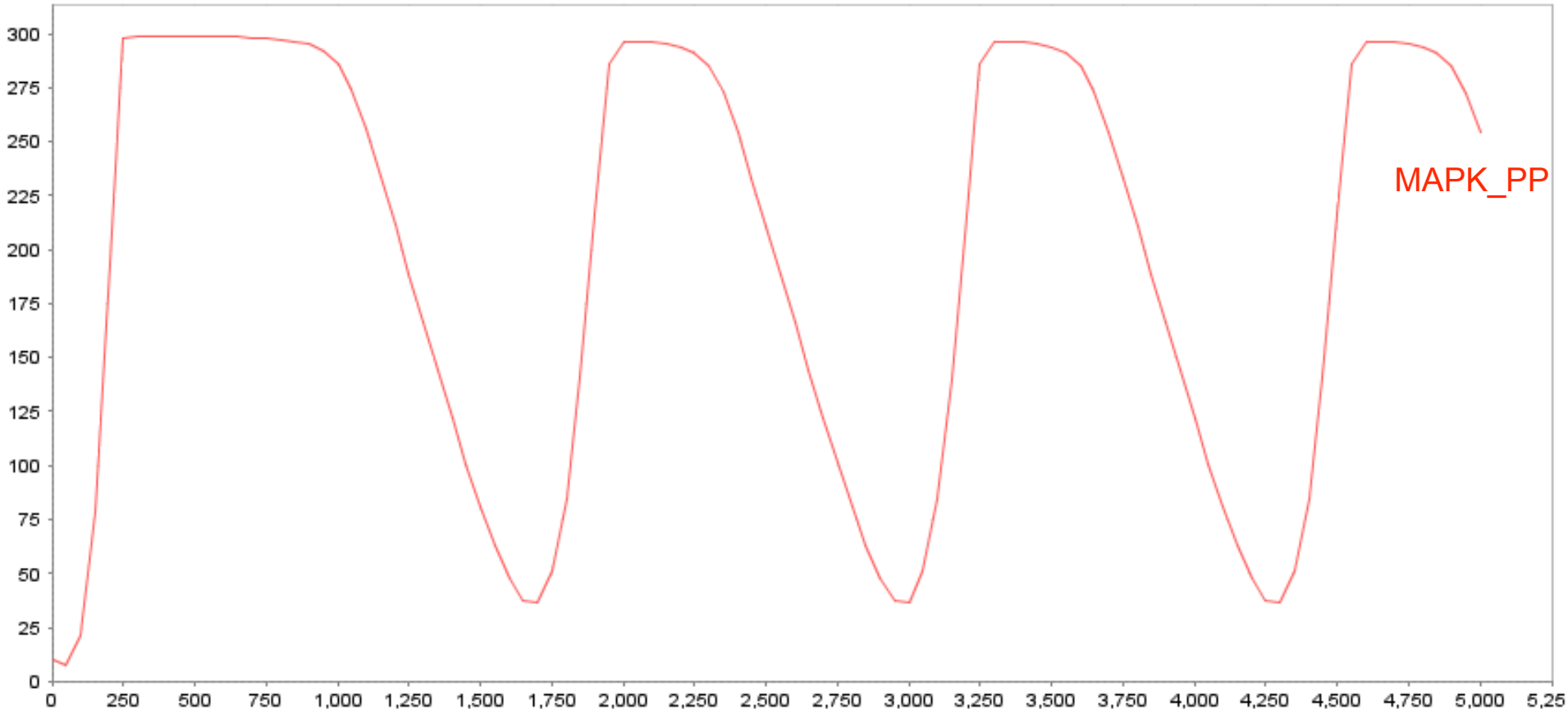
Kholodenko Model

- Kholodenko (2000): The oscillatory behaviour is because of the negative feedback loop.
- We vary the negative feedback effect (parameter Ki) and assess the effect on #oscillations detected:
  - $P_{=?}[\ F(d(MAPK\_PP) > 0\ \hat{}\ F(d(MAPK\_PP) < 0\ \hat{}\ \ldots\ )\ )\ ]$
- Using BioNessie parameter scan
- Can also distribute on the cluster

Kholodenko (2000), "Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades". Eur. J. Biochem. 2000, 267(6): 1583–1588.

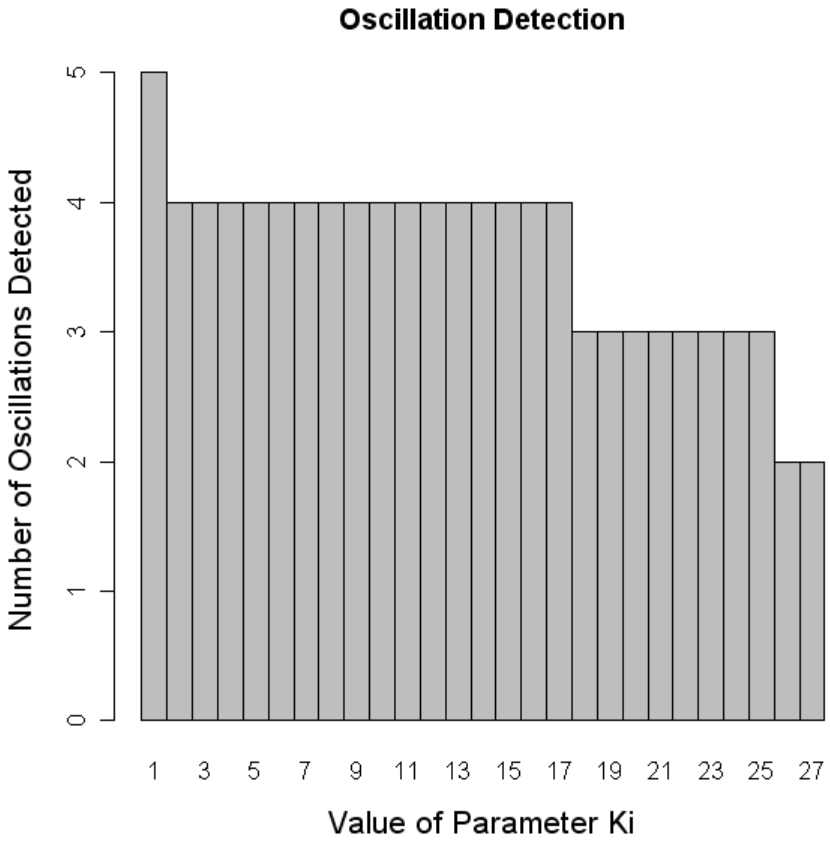
Kholodenko Model

- Oscillatory output:



Running MC2

- Output:
  5 oscillations: Satisfying Parameters: 1
  4 oscillations: Satisfying Parameters: 1 – 17
  3 oscillations: Satisfying Parameters: 1 - 25
  2 oscillations: Satisfying Parameters: 1 - 27



Conclusion

- Simulative approach to model checking
- Advantages of offline approach
- Added constraints to PLTL

- MC2(PLTLc) website (software + examples)
  http://www.brc.dcs.gla.ac.uk/software/mc2/

- MC2(PLTLc) in BioNessie
  Incorporated into the BioNessie simulator

- Funded by the Simulation Modelling of the Map Kinase Pathway (SIMAP) project.

Future Work

- So far, check a single model for behaviours.

  We could select models which have a desired behaviour from a set of models:

  “Give me every model which looks like this”

- We have a language to describe behaviours

  We could use this language to automatically build a model which has our behaviour.