BioModel Engineering for Systems and Synthetic Biology

David Gilbert

School of Information Systems, Computing & Mathematics
Centre for Systems & Synthetic Biology
Brunel University, London UK

Brunel UNIVERSITY LONDON

Workshop motivation

(Lectures at BioModLat "**Modeling in systems biology and synthetic biology" – Erasmus intensive programme, June 2013)**

Constructing, analysing and applying models of biochemical networks for

- prediction (systems biology) or
- design (synthetic biology),

is a major challenge that can benefit from the application of methods originating in computer science and software engineering.

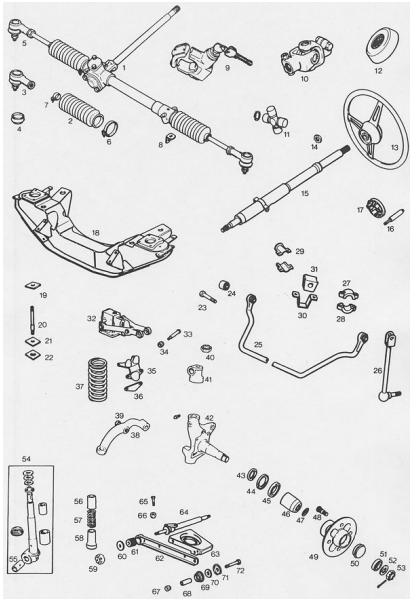
Workshop Synopsis

- Modelling framework
  - Qualitative: Qualitative Petri nets
  - Quantitative: Continuous Petri nets and Stochastic Petri nets.
  - Colour
- BioModel engineering
- Modular modelling
- Model checking
- Synthetic biology
- Case studies
- Computational tools

Materials

- people.brunel.ac.uk/~csstdrg/workshops/biomod2013/
- people.brunel.ac.uk/~csstdrg/workshops/biomod2013/papers/readinglist.html

Bioinformatics, Systems Biology, Synthetic Biology

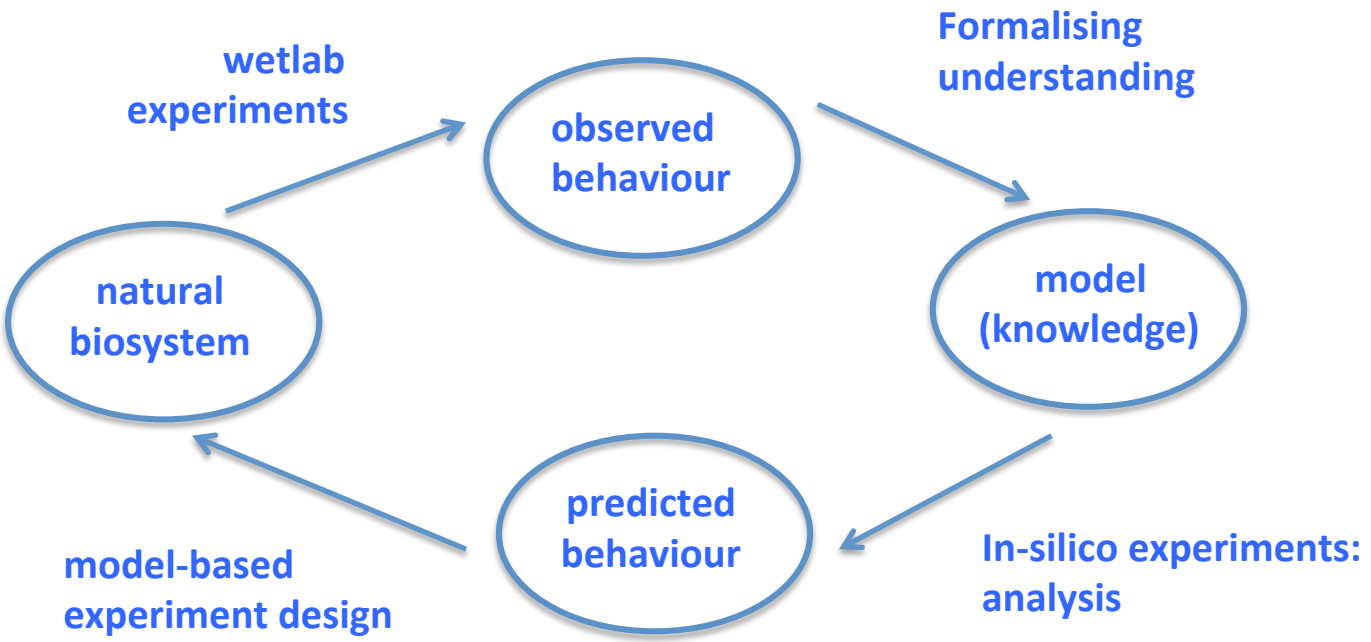


But how do these work together?

Build me a better one!

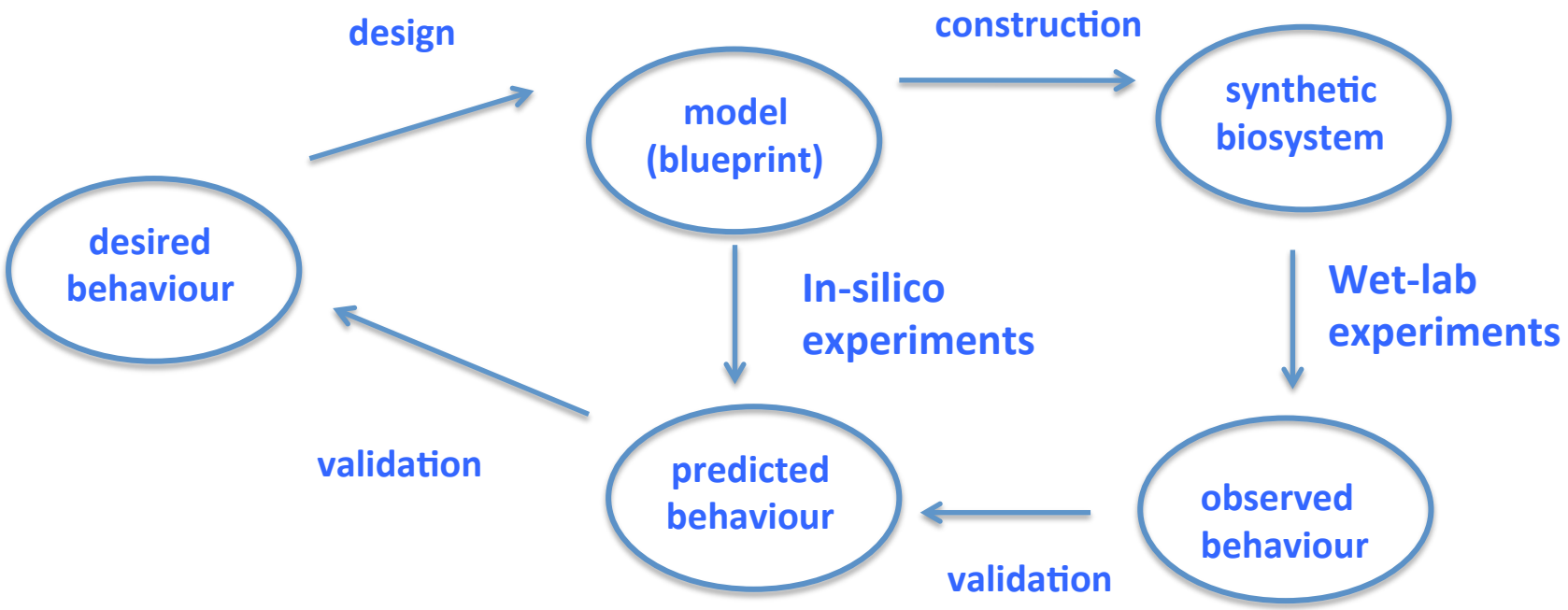
Brunel UNIVERSITY LONDON

Systems biology



Brunel UNIVERSITY LONDON

Synthetic Biology



Brunel UNIVERSITY LONDON

FM for molecular biology

- ordinary differential equation models
- process calculi
- state machines
- process algebras
- logics
- constraints-based modeling
- P-system
- …directly executable specifications

Engineering

The discipline, art and profession of

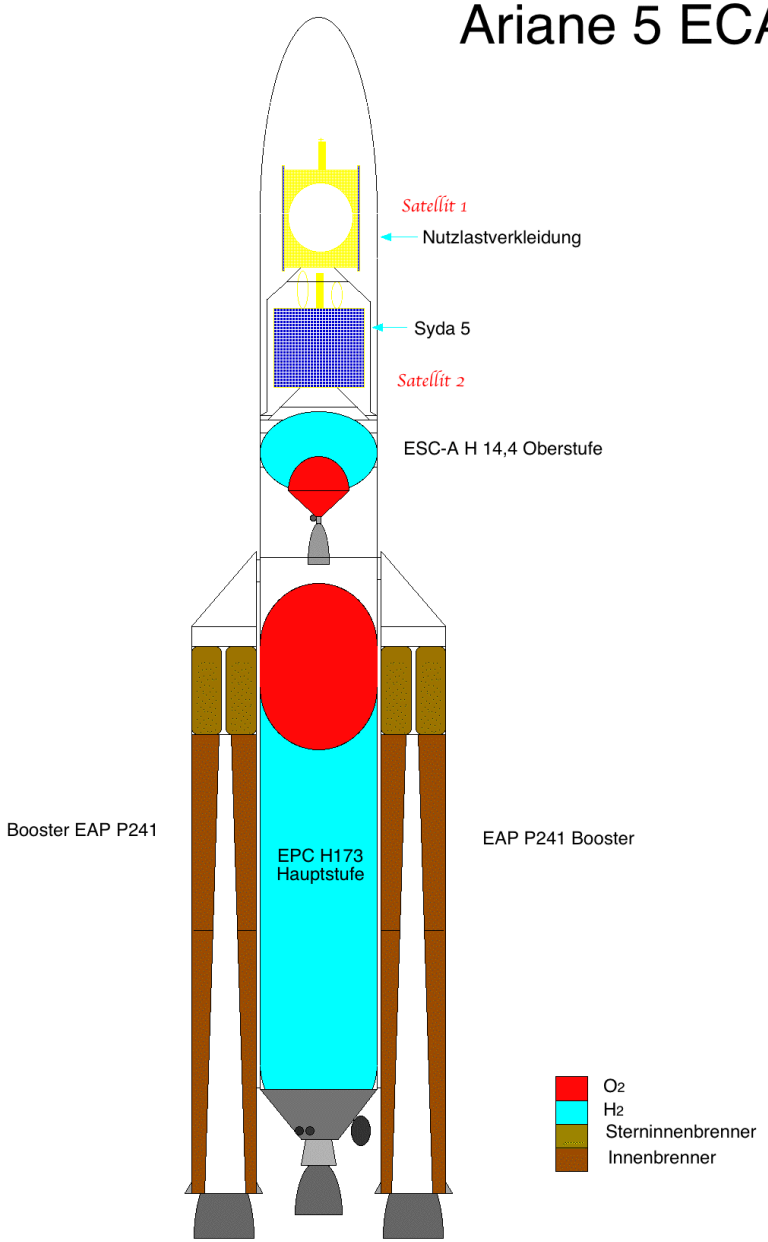
acquiring and applying technical, scientific, and mathematical knowledge to

- design, and
- implement

materials, structures, machines, devices, systems, and processes

that safely realize a desired objective or invention.

Ariane 5 ECA



Brunel UNIVERSITY LONDON

Brunel UNIVERSITY LONDON

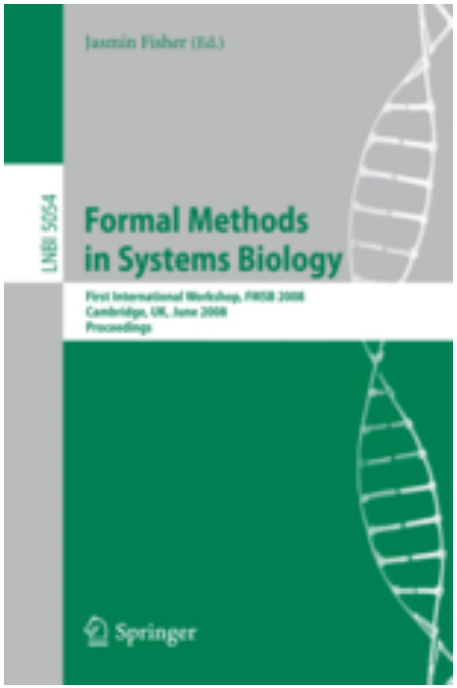
- 1996: the first launch of the Ariane 5 booster ended with a spectacular crash off the coast of French Guiana. The cause was traced to a variable overflow that affected software running in both channels of its dual redundant inertial reference system.

- 2005: European Space Agency's Huygens probe successfully beamed back only half of its image data. The other half was lost because of a single missing line of code.

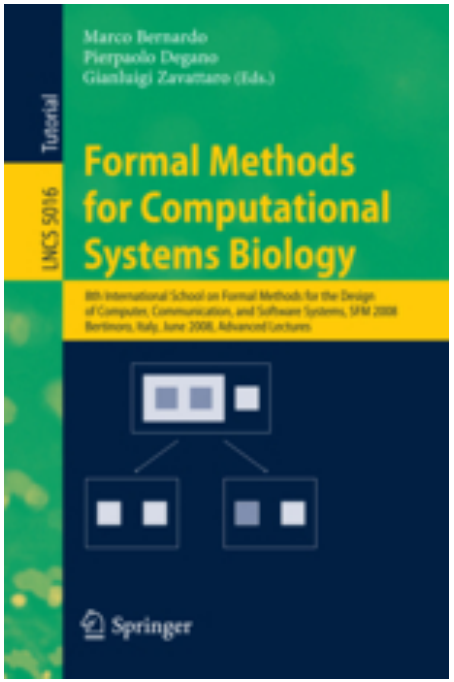
```
...
declare
  vertical_veloc_sensor: float;
  horizontal_veloc_sensor: float;
  vertical_veloc_bias: integer;
  horizontal_veloc_bias: integer;
  ...
begin
  declare
    pragma suppress(numeric_error, horizontal_veloc_bias);
  begin
    sensor_get(vertical_veloc_sensor);
    sensor_get(horizontal_veloc_sensor);
    vertical_veloc_bias := integer(vertical_veloc_sensor);
    horizontal_veloc_bias := integer(horizontal_veloc_sensor);
    ...
  exception
    when numeric_error => calculate_vertical_veloc();
    when others => use_irs1();
  end;
end irs2;
```

Formal Methods for molecular biology

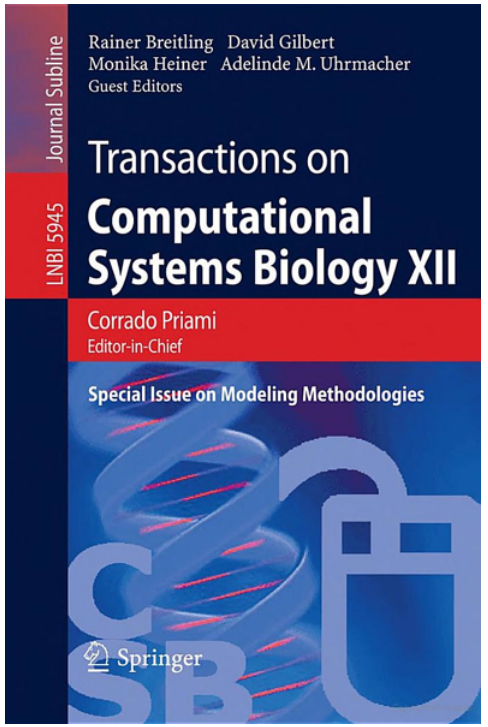
- Classical, well-established approach of modeling biological processes using continuous and stochastic differential equations,
- Formal logical models advantages:
  - **Easy compositionality**, which allows the generation and management of large cellular models from a number of pre-defined and reliably manipulated building blocks;
  - **model checking** for the rigorous exploration of model consistency, including the comprehensive exploration of state-space and the identification of necessary additions to an existing system description;
  - **unambiguous visualization** based on the strictly enforced syntax of the modeling language.
- Combination of formal logical models with continuous and stochastic differential equation models, showing important relationships between the two approaches and further expanding the expressivity of the resulting models.



Formal Methods in Systems Biology Cambridge, UK, 2008,
Lecture Notes in Computer Science / Lecture Notes in Bioinformatics, Vol. 5054



Formal Methods for Computational Systems Biology
8th International School on Formal Methods for the Design of Computer, Communication, and Software Systems, Bertinoro, Italy, 2008
Lecture Notes in Computer Science /
Programming and Software Engineering, Vol. 5016



Formal Methods in Molecular Biology
Dagstuhl Seminar 09091
February 2009
Transactions on Computational Systems Biology XII:
Special Issue on Modeling Methodologies. Springer LNBI 5945 2010

Formal Methods s for synthetic biology?

- Formal description techniques
- Formal languages
- Model construction
- Model checking
- System design
- System verification & validation

Validation & verification

- **Validation** – '***You built the right product?***'. (Quality control)
  - Product / system accomplishes its intended requirements.
  - But you didn't tell me you wanted a **red** bus!

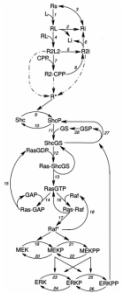
- **Verification** - '***You built the product right?***'. (Quality assurance)
  - System complies with its specification
- Possible for the product to produce the required outcome, but not in accord with its specification. (?)

Brunel UNIVERSITY LONDON

Current Synthetic Biology development cycle

Model construction

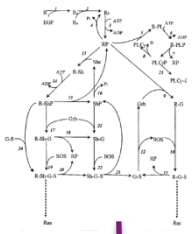
Biosystem



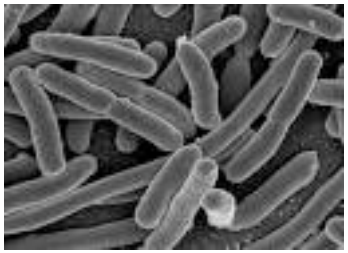
observations

validate

modify



construct



verify

modify

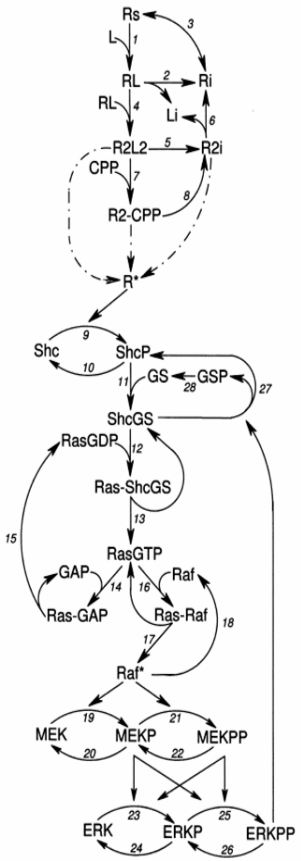
construct

verify

Brunel UNIVERSITY LONDON

Future vision

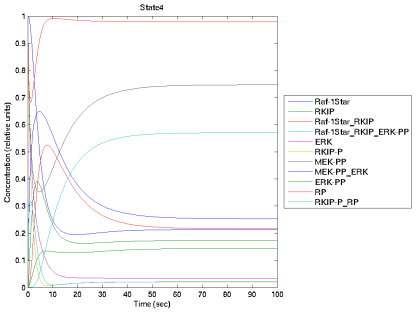
Model construction



observations

validate

Desired behaviour



construct

verify

Biological system

Component library

Standard biological parts

BioModel Engineering

- Takes place at the interface of computing science, mathematics, engineering & biology.
- A systematic approach for **designing**, **constructing** and **analyzing** computational models of biological systems.
- Inspiration from efficient software engineering strategies.
- Not engineering biological systems *per se,* but
  - describes their structure and behaviour,
  - in particular at the level of intracellular molecular processes,
  - using computational tools and techniques in a principled way.

Rainer Breitling, David Gilbert, Monika Heiner, Richard Orton (2008). A structured approach for the engineering of biochemical network models, illustrated for signalling pathways. Briefings in Bioinformatics

David Gilbert, Rainer Breitling, Monika Heiner, and Robin Donaldson (2009). An introduction to BioModel Engineering, illustrated for signal transduction pathways, 9th International Workshop, WMC 2008, Edinburgh, UK LNCS Volume 539, pp13-28

Rainer Breitling, Robin Donaldson, David Gilbert, Monika Heiner (2010): Biomodel Engineering - From Structure to Behavior; : Trans. Comp Systems Biology XII, Springer LNBI 5945, pp. 1-12

Monika Heiner and David Gilbert (2012): BioModel Engineering for Multiscale Systems Biology; Progress in Biophysics and Molecular Biology.

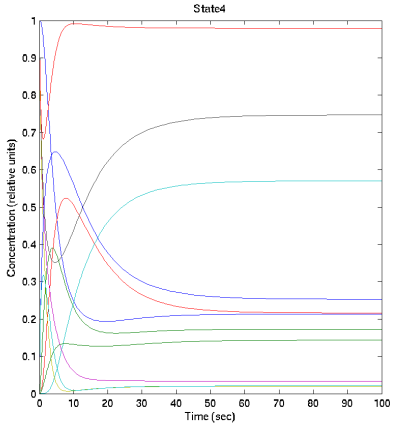
Biomodel engineering

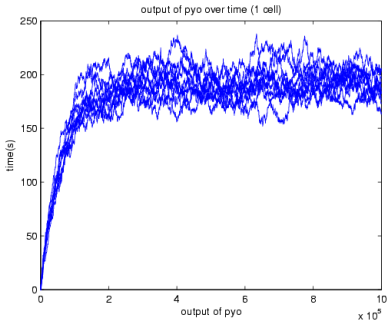
1. Problem identification
2. Construction
3. Simulation
4. Analysis & interpretation
5. Management & development

Biomodel engineering

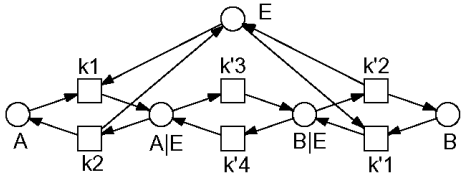
1. Problem identification
2. **Construction**
3. **Simulation**
4. **Analysis & interpretation**
5. Management & development

Where to start?





$$\frac{d[A]}{dt} = -k_1 \times [A]$$

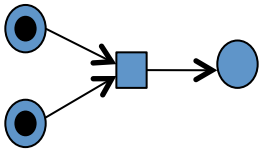


Brunel UNIVERSITY LONDON

We need a Framework!

Molecules/Levels
CTL, LTL

Qualitative



Time-free

Timed, Quantitative

Abstraction

Abstraction

Monika Heiner

Markov chain
Molecules/Levels
Stochastic rates
CSL

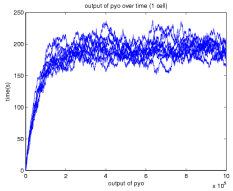
Stochastic

Approximation

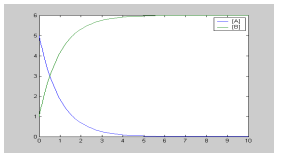
Approximation

Continuous

ODES
Concentrations
Deterministic rates
LTLc



$$\frac{d[A]}{dt} = -k_1 \times [A]$$



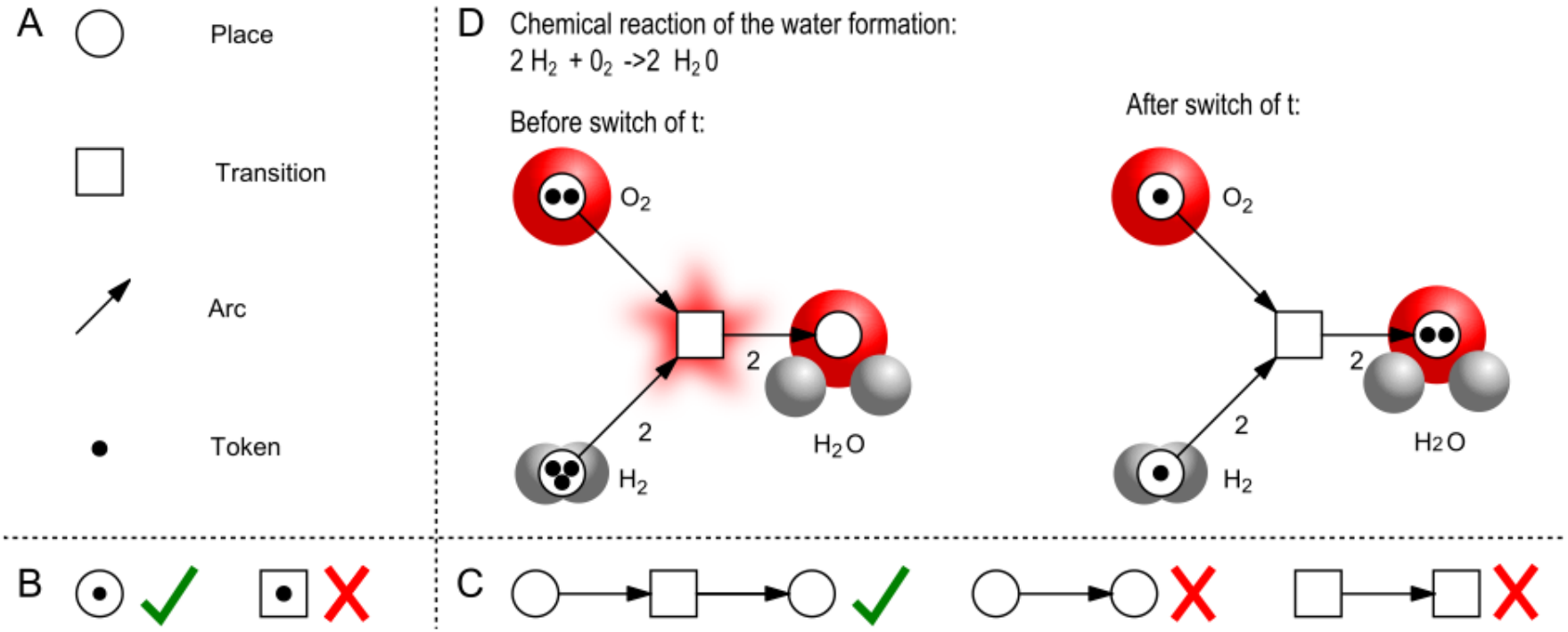
DiscreteState Space

Continuous State Space

Gilbert, Heiner and Lehrack. ``A Unifying Framework for Modelling and Analysing Biochemical Pathways Using Petri Nets.'' Proc CMSB 2007

Brunel UNIVERSITY LONDON

Petri nets

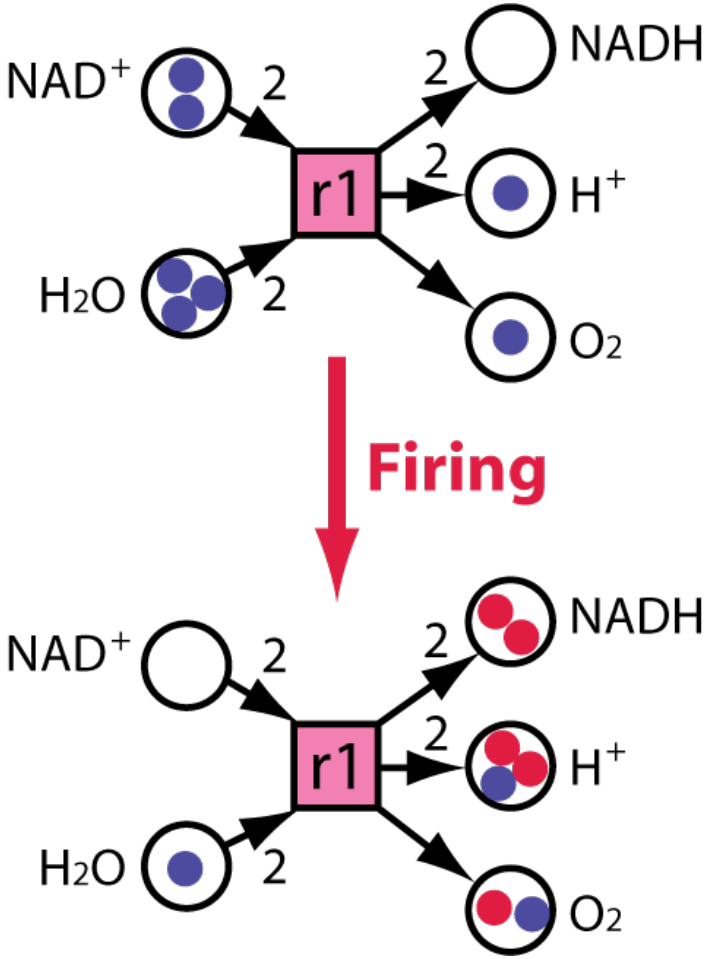


Marking: numbers of tokens in places

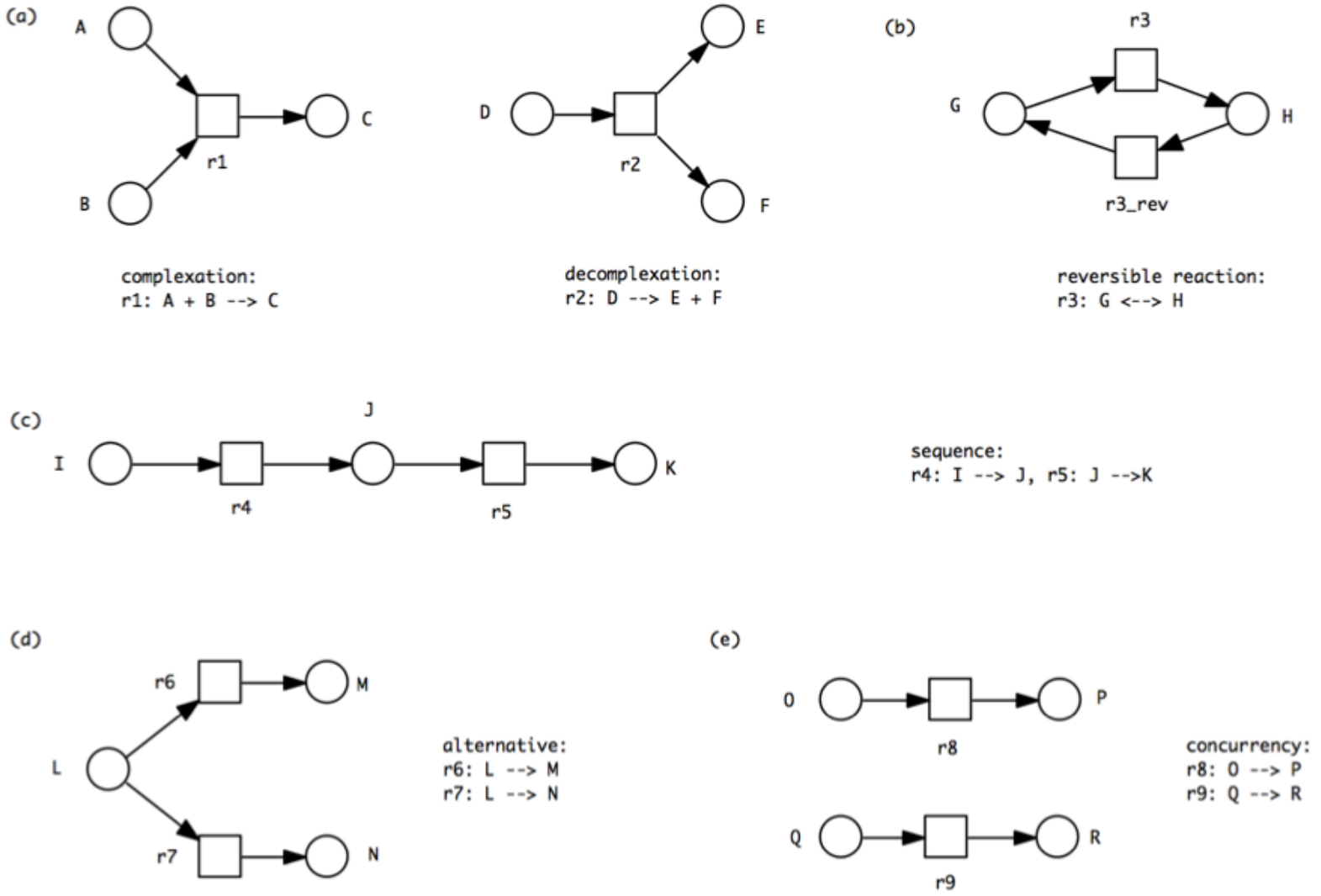
MA Blätke, M Heiner, and W Marwan (2011), **Tutorial - Petri Nets in Systems Biology. Otto von Guericke University Magdeburg, Magdeburg Centre for Systems Biology**

Brunel UNIVERSITY LONDON

Petri nets



Some biochemical reactions



Brunel UNIVERSITY LONDON

Petri net demo!

simple1
simple2
simple3

H2O

NADH

web animation

Petri net

A standard Petri net is a quadruple
N = (P, T, f, m0), where:

P, T are finite, non-empty, disjoint sets.
P is the set of places. T is the set of transitions.

f: $((P \times T) \cup (T \times P)) \rightarrow N0$ defines the set of directed arcs, weighted by non-negative integer values.

m0: $P \rightarrow N0$ gives the initial marking.

Marking

m(p) - the number of tokens on place p in the marking m.

- Place p is clean (empty, unmarked) in m if m(p) = 0, otherwise place p is marked in m.
- A set of places is called clean if all places are clean, otherwise the set is marked.
- The postset and preset of a node $x \in P \cup T$, is defined as:
- Preset: $\bullet x := \{y \in P \cup T \mid f(y,x) \neq 0\}$
- Postset: $x \bullet := \{y \in P \cup T \mid f(x,y) \neq 0\}$
- For places and transitions, we get four types of sets:
- $\bullet t$ - preplaces of transition t (reaction's precursor)
  $t \bullet$ - postplaces of transition t (reaction's products)
  $\bullet p$ - pretransitions of place p (all producing reactions of a component)
- $p \bullet$ - posttransitions of place p (all consuming reactions of a component)

Firing rule

Let N = (P, T, f, m0) be a Petri net:

- A transition is enabled in marking m, written as m [t⟩, if $\forall p \in {\bullet}t : m(p) \geq f(p, t)$, else disabled.
- A transition t, which is enabled in m, may fire. When t in m fires, a new marking m′ is reached, written as m[t⟩m′, with $\forall p \in P : m'(p) =$
- $m(p) - f(p,t) + f(t,p)$.
  The firing happens atomically and does not consume any time.

MA1: mass-action enzymatic reaction

a

$$A \xrightarrow{E} B$$

b

$$E + A \underset{k_2}{\overset{k_1}{\rightleftarrows}} E \mid A \xrightarrow{k_3} E + B$$

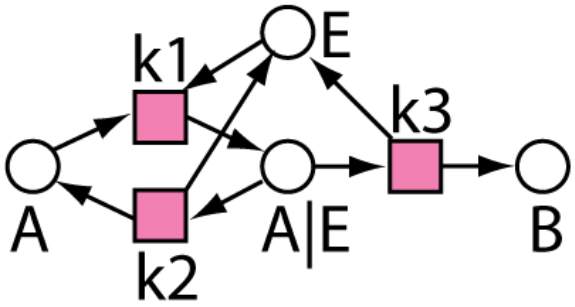
A: substrate

B: product

E: enzyme

E|A complex

c



d

$$\frac{d[A]}{dt} = -k_1[A][E] + k_2[A \mid E]$$

$$\frac{d[E]}{dt} = -k_1[A][E] + k_2[A \mid E] + k_3[A \mid E]$$

$$\frac{d[A \mid E]}{dt} = k_1[A][E] - k_2[A \mid E] - k_3[A \mid E]$$

$$\frac{d[B]}{dt} = k_3[A \mid E]$$

MA1: mass-action enzymatic reaction

A: substrate

B: product

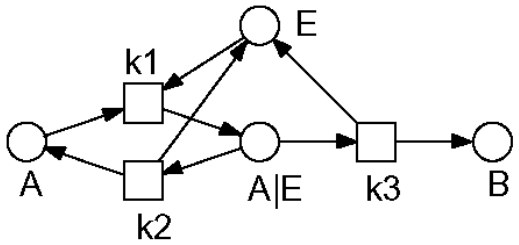
E: enzyme

E|A complex

$$A \xrightarrow{E} B$$

$$A + E \underset{k_2}{\overset{k_1}{\rightleftharpoons}} A \mid E \xrightarrow{k_3} B + E$$

Biochemistry



Petri net

$$\frac{d[A]}{dt} = -k_1 \times [A] \times [E] + k_2 \times [A \mid E]$$

$$\frac{d[A \mid E]}{dt} = +k_1 \times [A] \times [E] - k_2 \times [A \mid E] - k_3 \times [A \mid E]$$

$$\frac{d[B]}{dt} = + k_3 \times [A \mid E]$$

$$\frac{d[E]}{dt} = -k_1 \times [A] \times [E] + k_2 \times [A \mid E] + k_3 \times [A \mid E]$$

Ordinary Differential Equations

MA1 Petri net

N = (P,T,f,m0)

Set of places P = {Enzyme, Substrate, EnzymeSubstrateComplex, Product}

Set of transitions T = {Association, Dissociation, Synthesis}

Set of directed arcs: $((P \times T) \cup (T \times P))$ is the combination of the following subsets

Places connected with (→) transitions:

$(P \times T)$ = {Substrate × Association, Enzyme × Association, EnzymeSubstrateComplex × Dissociation, EnzymeSubstrateComplex × Synthesis}

Transitions connected with (→) places:

$(T \times P)$ = {Assoication × EnymeSubstrateComplex, Dissociation × Substrate, Dissociation × Product, Synthesis × Product, Synthesis × Enzyme}

Initial Marking m0 = {Enzyme = 1, Substrate = 1, EnymeSubstrateComplex = 0, Product = 0};

The amount of tokens must be expressed as an integer variable.

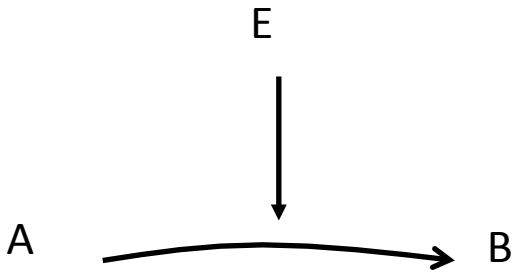
Petri net demo!

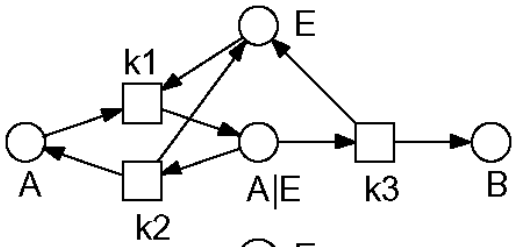
enzymatic reaction

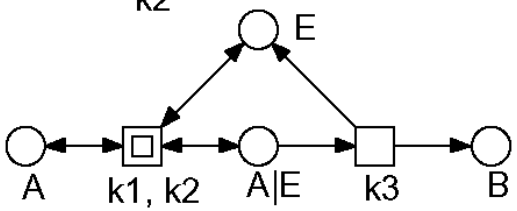
MA1: Mass action for enzymatic reaction

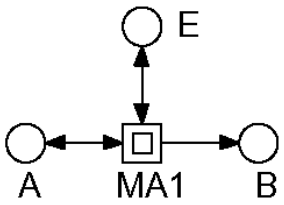
$$E + A \underset{k_2}{\overset{k_1}{\rightleftharpoons}} E \mid A \xrightarrow{k_3} E + B$$

- A: substrate
- B: product
- E: enzyme
- E|A substrate-enzyme complex





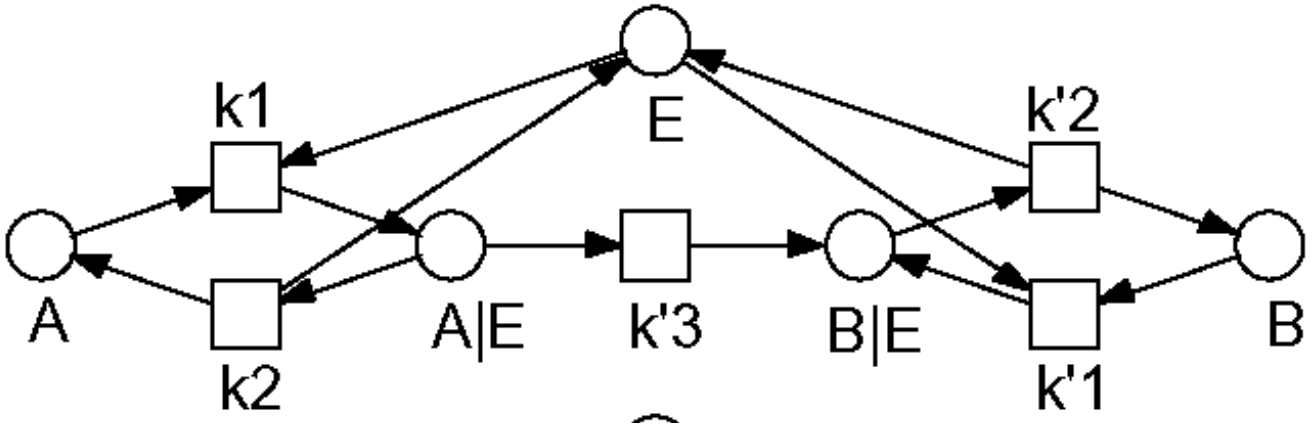




Brunel UNIVERSITY LONDON

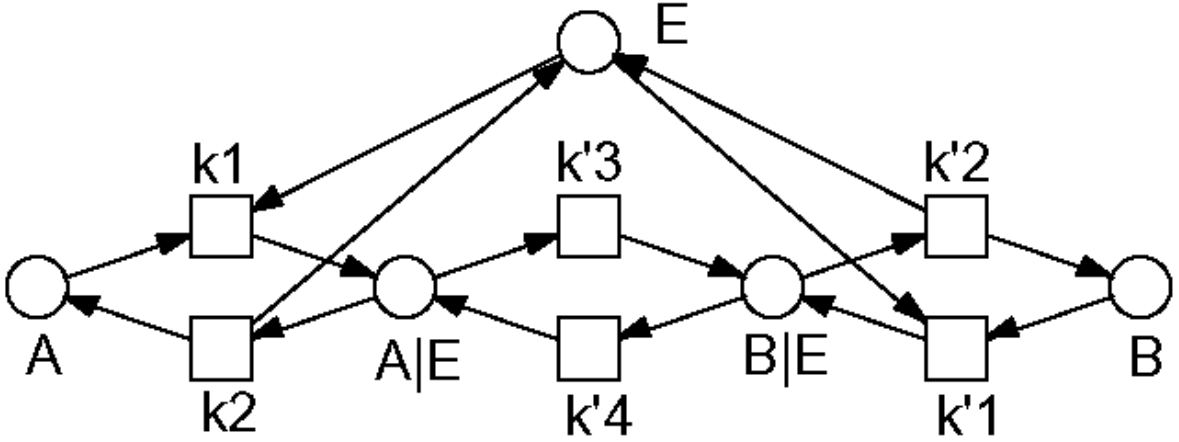
MA2 model

$$A + E \underset{k_2}{\overset{k_1}{\rightleftarrows}} A \mid E \xrightarrow{k_3} B \mid E \underset{k5}{\overset{k_4}{\rightleftarrows}} B + E$$



MA3 model

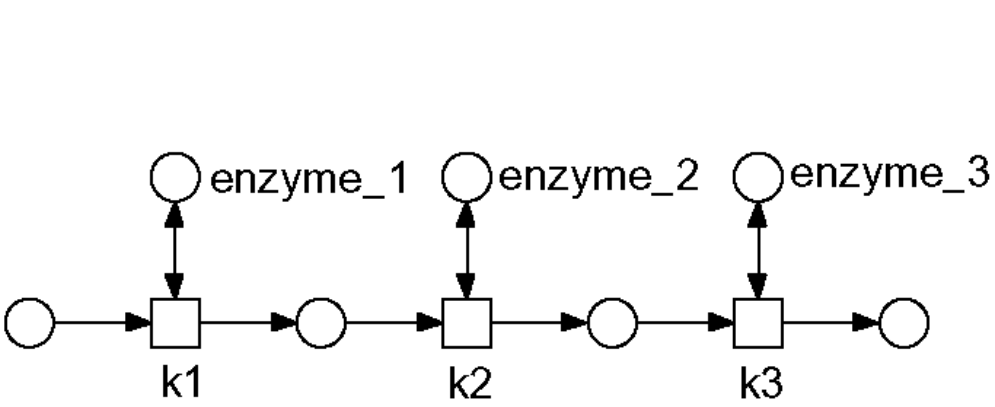
$$A + E \underset{k_2}{\overset{k_1}{\rightleftharpoons}} A \mid E \underset{k_6}{\overset{k_3}{\rightleftharpoons}} B \mid E \underset{k_5}{\overset{k_4}{\rightleftharpoons}} B + E$$

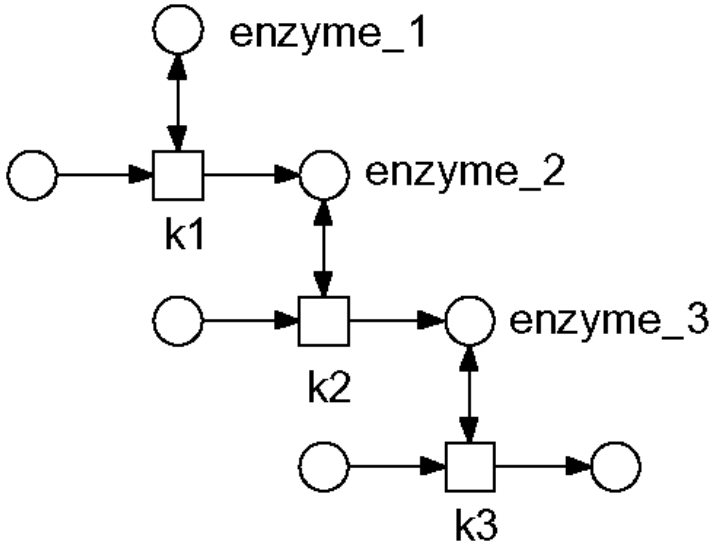


Multiple substrates

$$A_1 + A_2 + E \underset{k_{2_{A_1}}}{\overset{k_{1_{A_1}}}{\rightleftharpoons}} A_1|E + A_2 \underset{k_{2_{A_2}}}{\overset{k_{1_{A_2}}}{\rightleftharpoons}} A_1|A_2|E \xrightarrow{k_3} B_1 + B_2 + E$$

Metabolic pathways vs Signalling Pathways

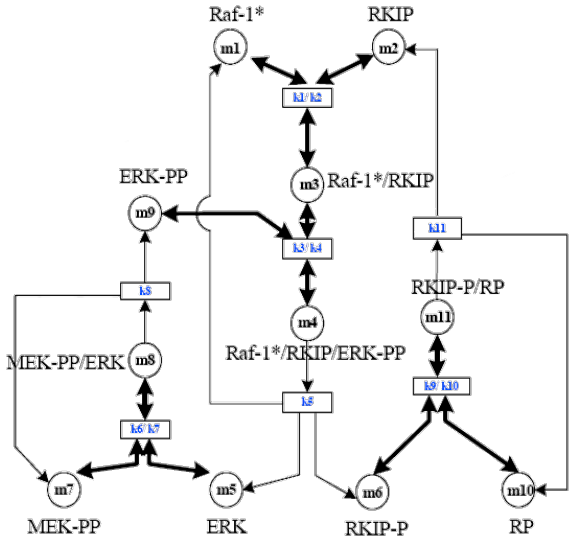




demo

What is a biochemical network model?

1. Structure



graph
QUALITATIVE

2. Kinetics (if you can)

reaction rates

$$d[Raf1^*]/dt = k1*m1*m2 + k2*m3 + k5*m4$$

$$k1 = 0.53;\ k2 = 0.0072;\ k5 = 0.0315$$

QUANTITATIVE

3. Initial conditions

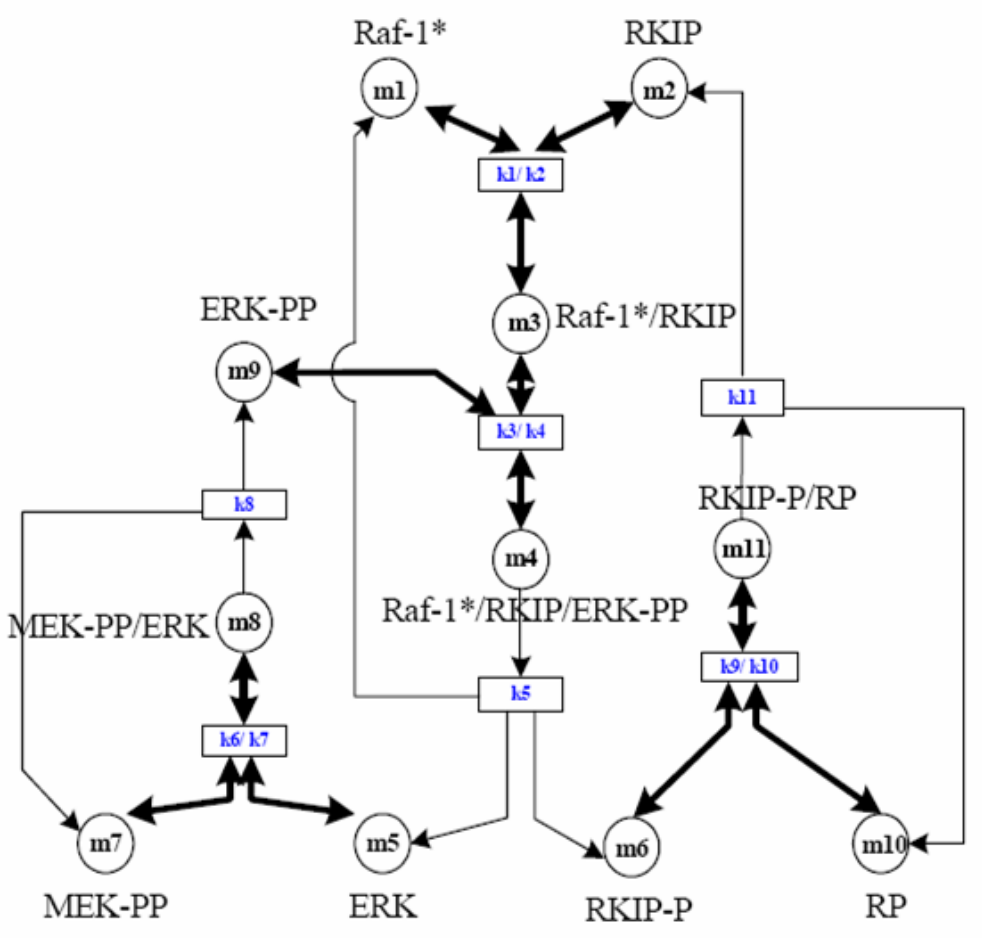
marking , concentrations

$[Raf1^*]_{t=0}$= 2 μMolar

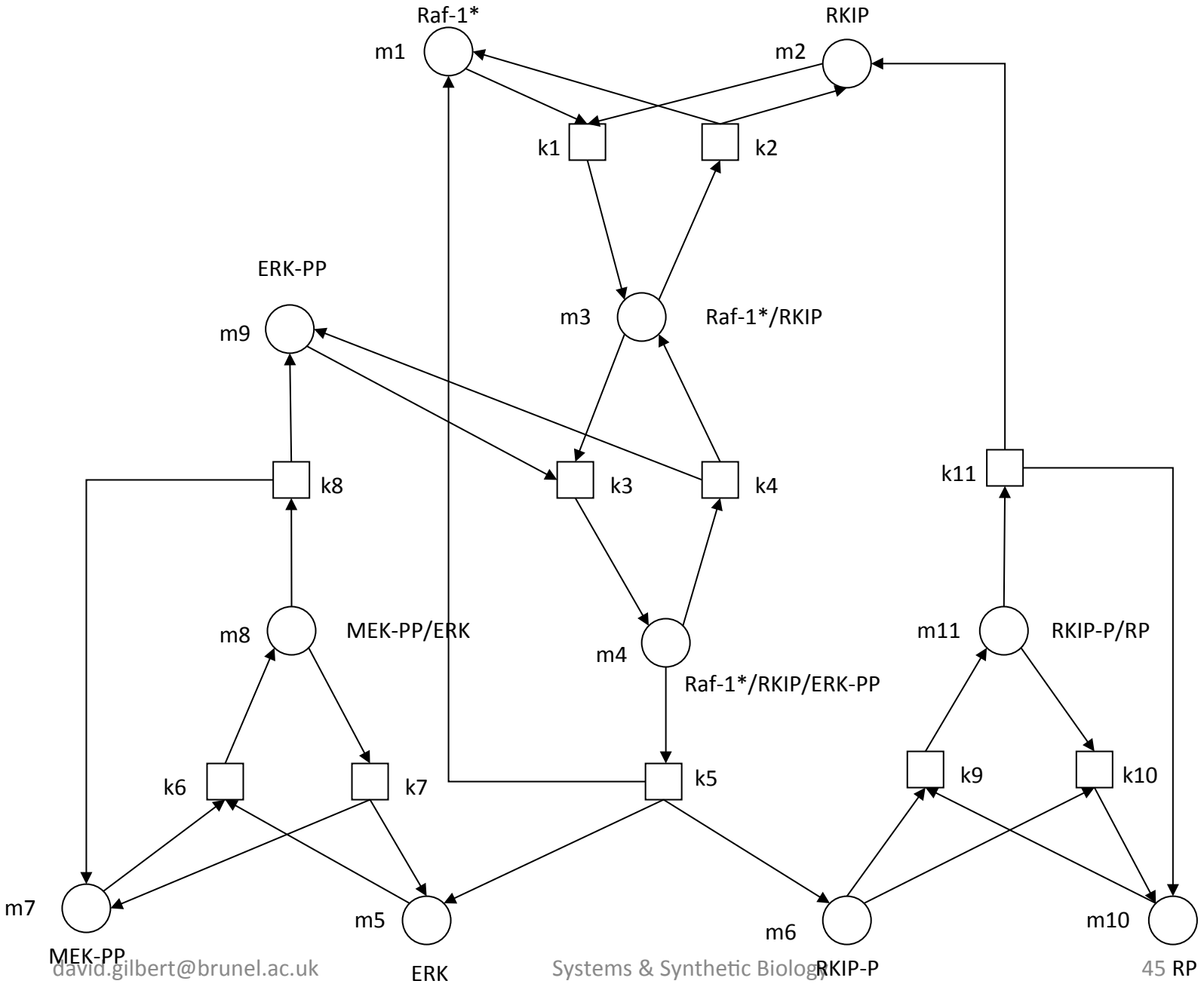
QUANTITATIVE

Brunel UNIVERSITY LONDON

The Raf-1/RKIP/ERK pathway

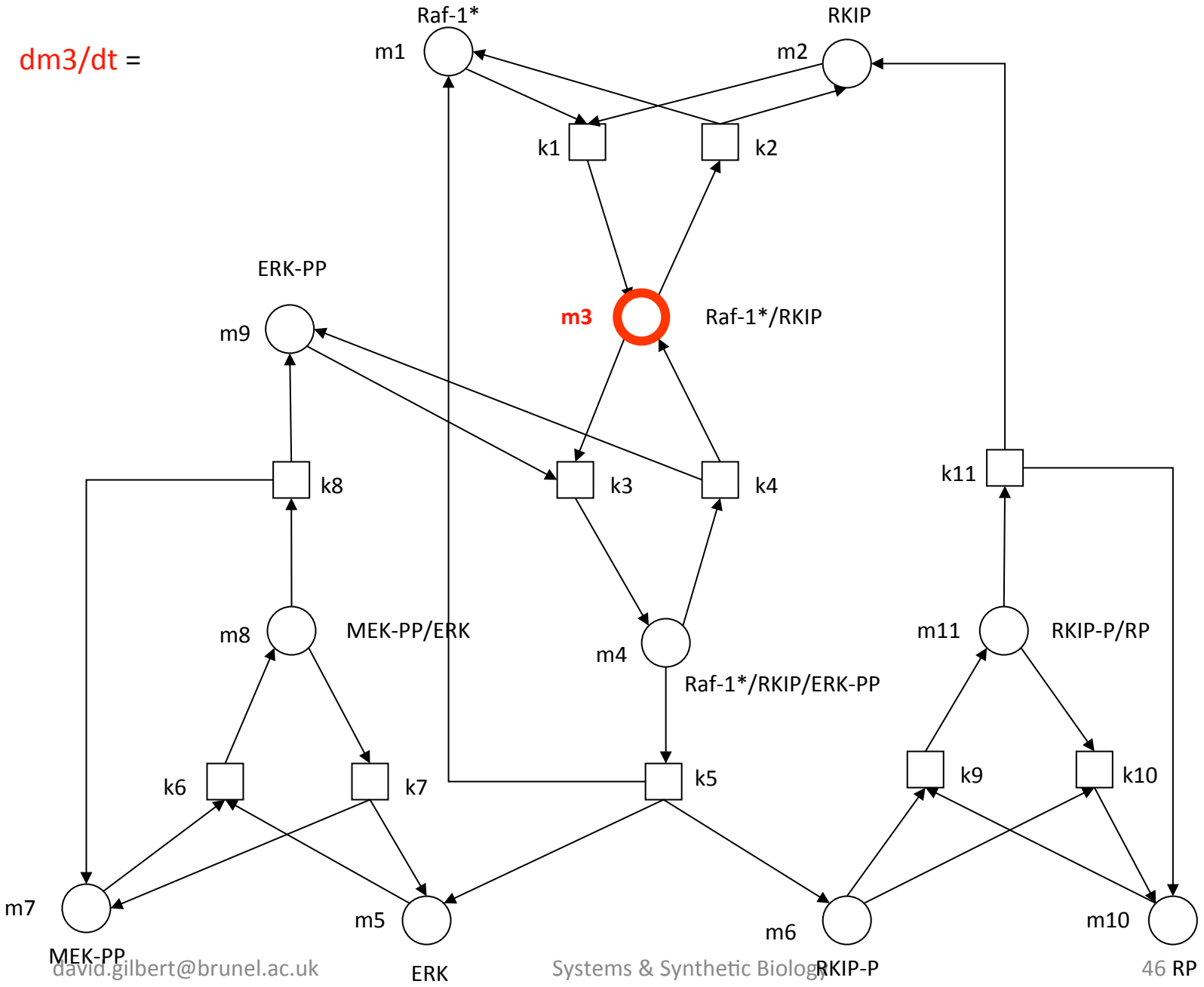


Brunel UNIVERSITY LONDON



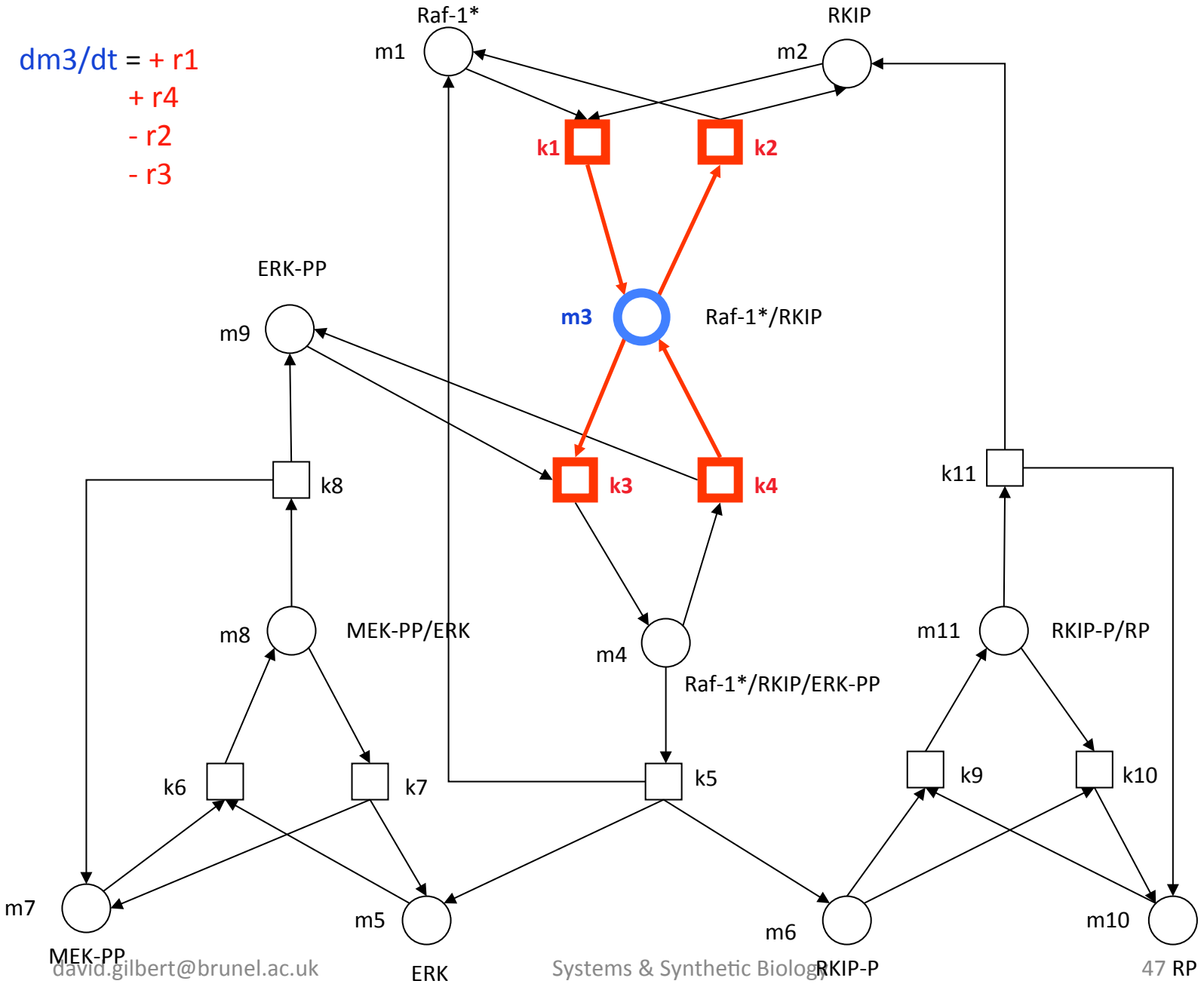
Brunel UNIVERSITY LONDON

dm3/dt =



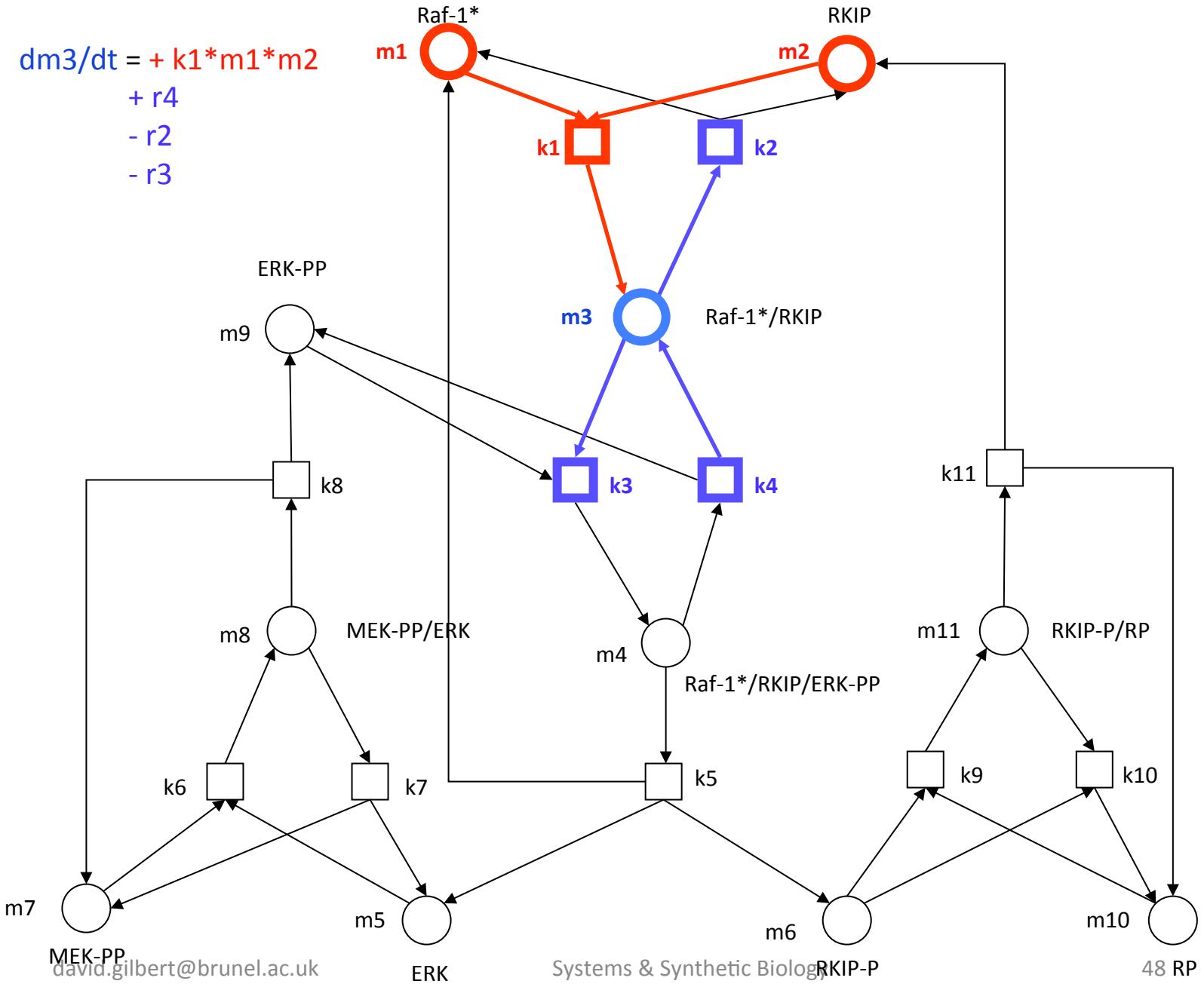
Brunel UNIVERSITY LONDON

dm3/dt = + r1
+ r4
- r2
- r3



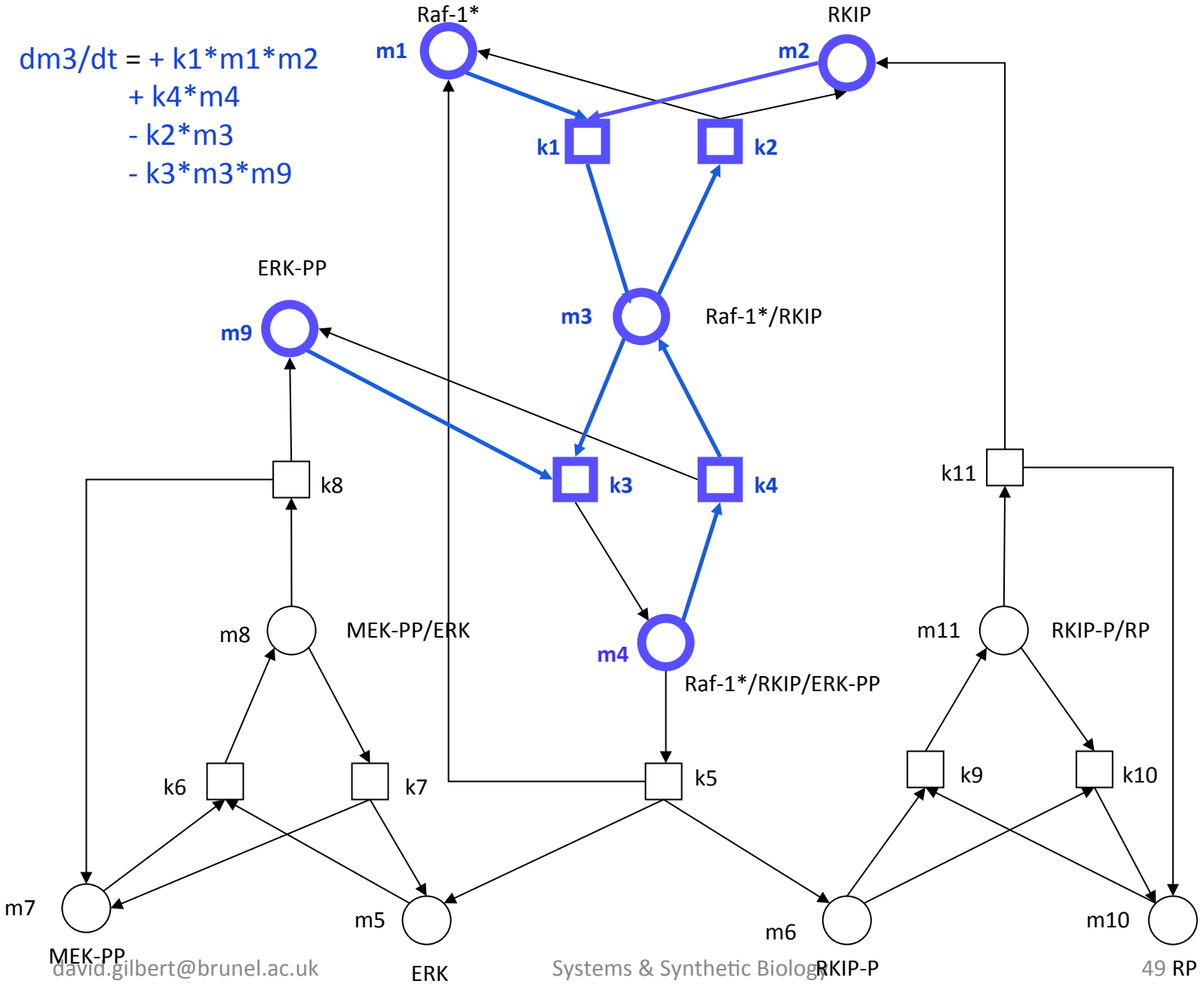
Brunel UNIVERSITY LONDON

dm3/dt = + k1*m1*m2
+ r4
- r2
- r3



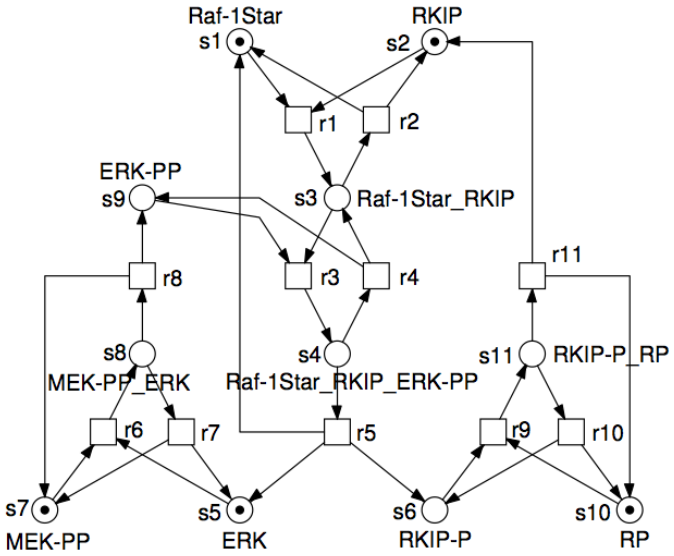
Brunel UNIVERSITY LONDON

dm3/dt = + k1*m1*m2
+ k4*m4
- k2*m3
- k3*m3*m9



Brunel UNIVERSITY LONDON

Continuous to stochastic



First order reactions: The stochastic and deterministic rate constants are equivalent

Second order reactions: *fs* is a scaling factor to map the given *mass* in the continuous concentration onto a finite number of levels (i.e tokens), with N being the highest level number, i.e. *fs* = *mass/N*.

#	reaction equation	rate function v_i	rate constant [a] stochastic	deterministic
r1	s1 + s2 → s3	$c_1 \cdot s1 \cdot s2$	$c_1' = c_1 \cdot f_s$	$c_1 = 0.53$
r2	s3 → s1 + s2	$c_2 \cdot s3$	$c_2' = c_2$	$c_2 = 0.0072$
r3	s3 + s9 → s4	$c_3 \cdot s3 \cdot s9$	$c_3' = c_3 \cdot f_s$	$c_3 = 0.625$
r4	s4 → s3 + s9	$c_4 \cdot s4$	$c_4' = c_4$	$c_4 = 0.00245$
r5	s4 → s1 + s5 + s6	$c_5 \cdot s4$	$c_5' = c_5$	$c_5 = 0.0315$
r6	s5 + s7 → s8	$c_6 \cdot s5 \cdot s7$	$c_6' = c_6 \cdot f_s$	$c_6 = 0.8$
r7	s8 → s5 + s7	$c_7 \cdot s8$	$c_7' = c_7$	$c_7 = 0.0075$
r8	s8 → s7 + s9	$c_8 \cdot s8$	$c_8' = c_8$	$c_8 = 0.071$
r9	s6 + s10 → s11	$c_9 \cdot s6 \cdot s10$	$c_9' = c_9 \cdot f_s$	$c_9 = 0.92$
r10	s11 → s6 + s10	$c_{10} \cdot s11$	$c_{10}' = c_{10}$	$c_{10} = 0.00122$
r11	s11 → s2 + s10	$c_{11} \cdot s11$	$c_{11}' = c_{11}$	$c_{11} = 0.87$

Brunel UNIVERSITY LONDON

Petri net demo!

Enzymatic reactions

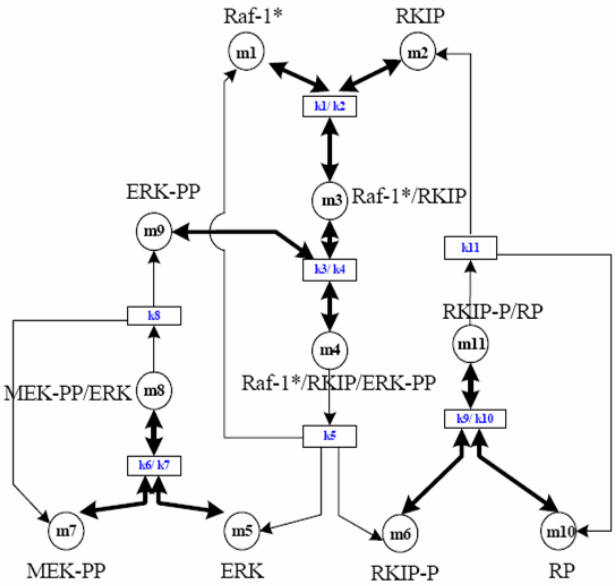
- MA1 (qualitative)
- MA1 (continuous)
- MA1 (stochastic)

- RKIP (continuous)
- RKIP (stochastic)

Modules modelling:

Cell signalling

Construction



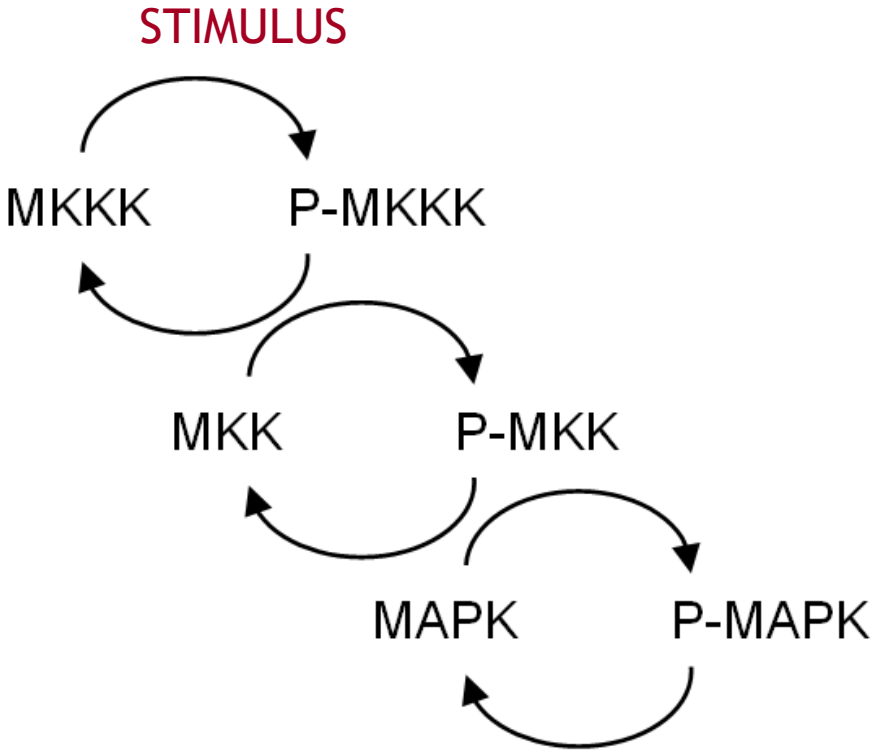
- **Construct topology**
  - Define all the proteins/molecules involved
  - Define the reactions they are involved in
  - Where do you draw the model boundary line?

- **Check literature**
  - What is known about the pathway and proteins?
  - What evidence is there that protein A binds directly to protein B?
  - Protein C also binds directly to protein B: does it compete with protein A or do they bind to protein B at different sites?
  - Trust & Conflicts: it is important to recognize which evidence to trust and which to discard (talk to the people in the wet lab)

- **Simplifying assumptions**
  - Many biological processes are very complex and not fully understood
  - Therefore, developing a model often involves making simplifying assumptions
  - For example, the activation of Raf by Ras is very complicated and not fully understood but it is often modelled as:
    - Raf + Ras-GTP = Raf/Ras-GTP -> Raf-x + Ras-GTP
  - Although this is a simplification, it is able to explain the observed data

MAPK Pathway

- Responds to wide range of stimuli: cytokines, growth factors, neurotransmitters, cellular stress and cell adherence,…
- Pivotal role in many key cellular processes:
  - growth control in all its variations,
  - cell differentiation and survival
  - cellular adaptation to chemical and physical stress.
- Deregulated in various diseases: cance immunological, inflammatory and degenerative syndromes,
- Represents an important drug target.

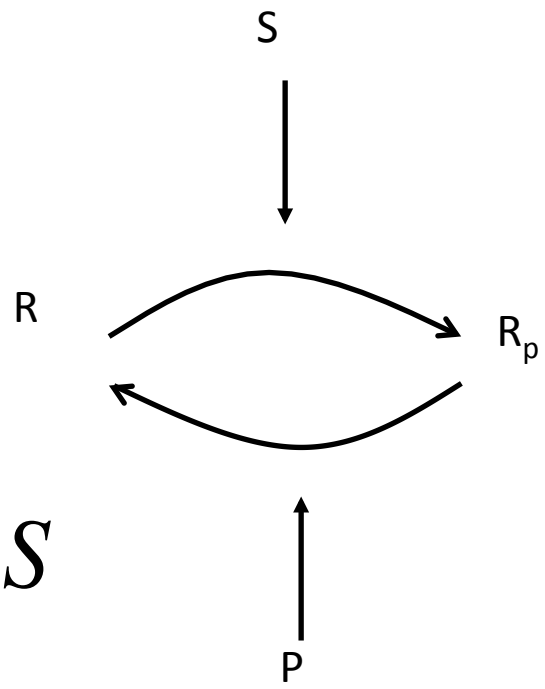


Gilbert et al (2006) Computational methodologies for modelling, analysis and simulation of signalling networks, Briefings in Bioinformatics 2006

david.gilbert@brunel.ac.uk

Phosphorylation - dephosphorylation step
Mass action

- R: unphosphorylated form
- R_p: phosphorylated form
- S: kinase
- P: phosphotase
- R|S unphosphorylated+kinase complex
- R|P unphosphorylated+phosphotase complex

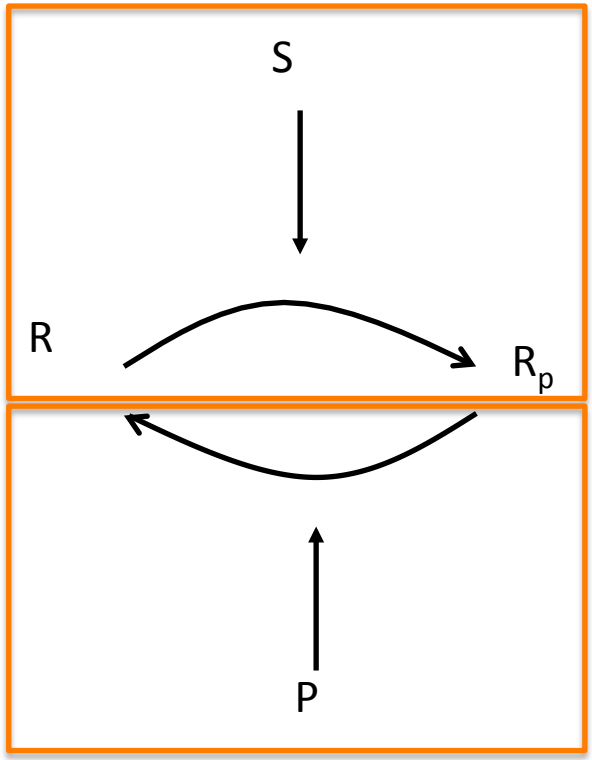


$$R+S \underset{k_2}{\overset{k_1}{\rightleftarrows}} R|S \xrightarrow{k_3} R_p + S$$

$$R+P \xleftarrow{kr_3} R_p|P \underset{kr_2}{\overset{kr_1}{\rightleftarrows}} R_p + P$$

Phosphorylation - dephosphorylation step
Mass action model 1

- R: unphosphorylated form
- R_p: phosphorylated form
- S: kinase
- P: phosphotase
- R|S unphosphorylated+kinase complex
- R|P unphosphorylated+phosphotase complex



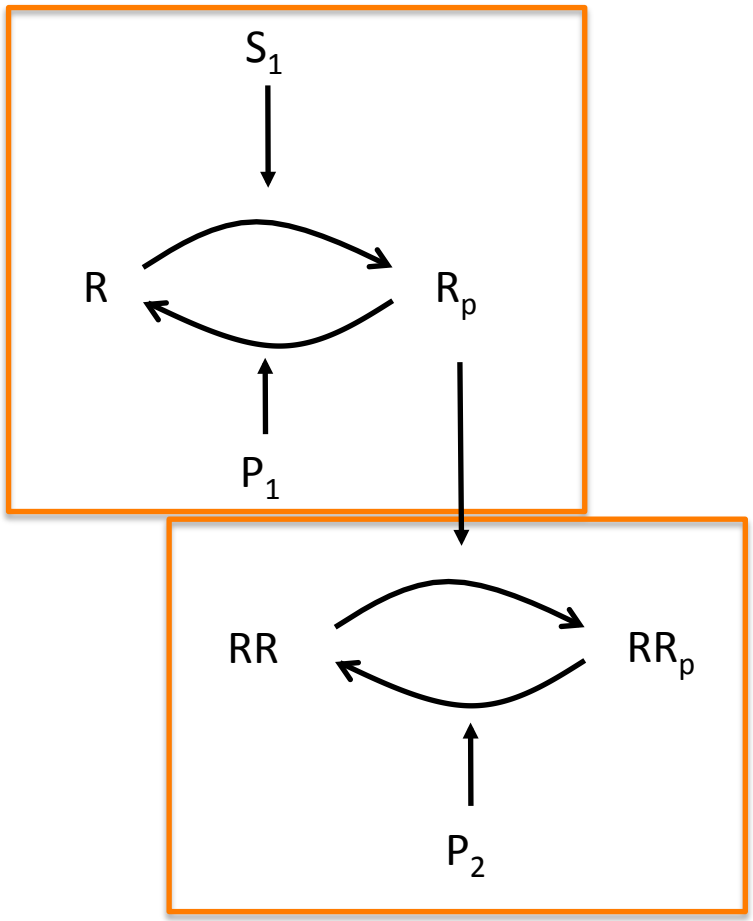
$$R+S \underset{k_2}{\overset{k_1}{\rightleftarrows}} R|S \xrightarrow{k_3} R_p + S$$

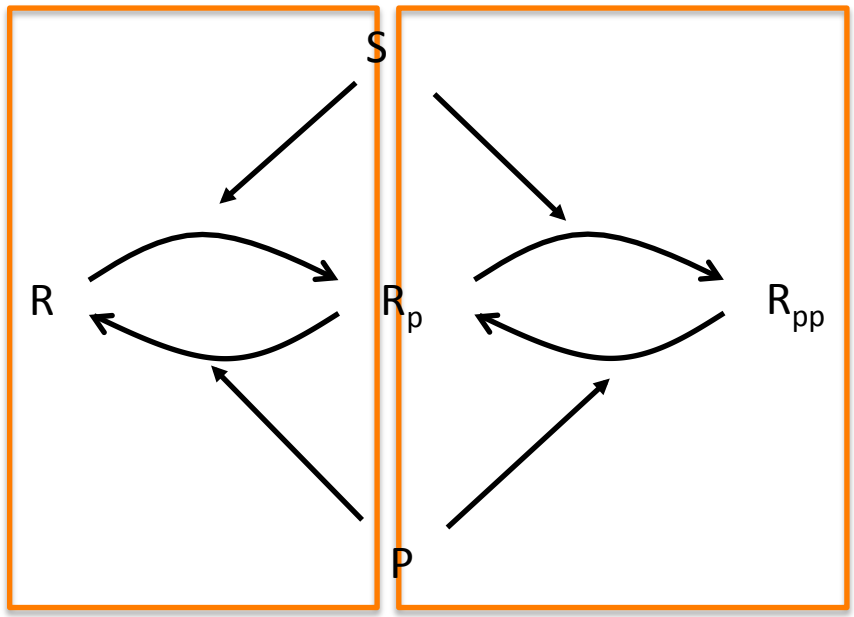
$$R+P \xleftarrow{kr_3} R_p|P \underset{kr_2}{\overset{kr_1}{\rightleftarrows}} R_p + P$$

Breitling, Gilbert, Heiner, Orton. A structured approach for the engineering of biochemical network models, illustrated for signalling pathways. Briefings in Bioinformatics, 2008

Composition
Vertical & horizontal

2-stage cascade

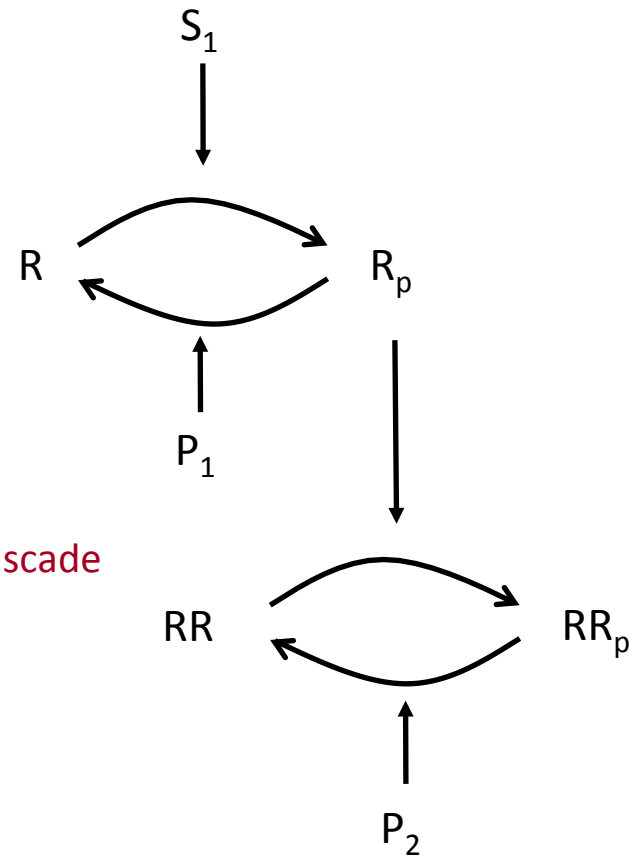




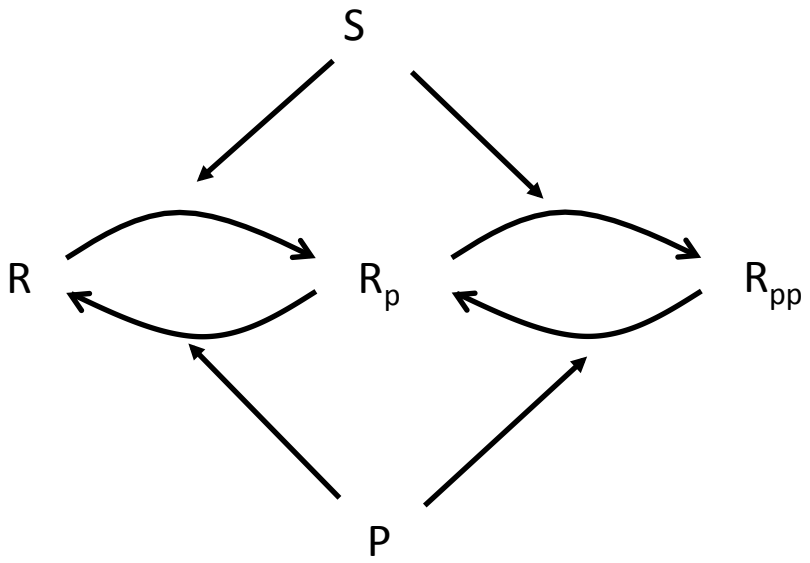
1-stage cascade
double phosphorylation step

Brunel UNIVERSITY LONDON

Composition
Vertical & horizontal



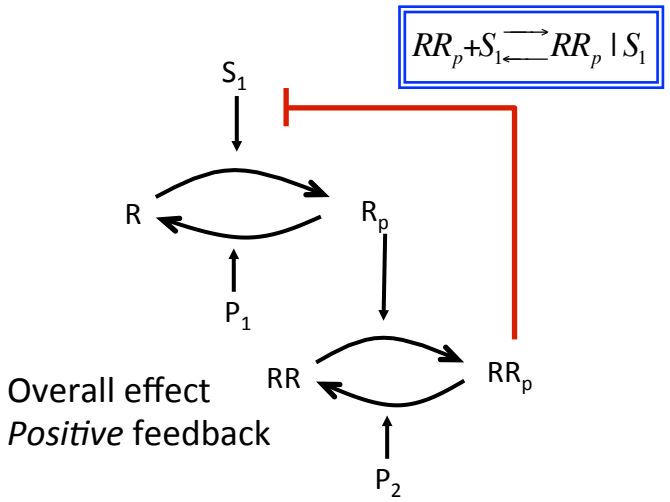
2-stage cascade

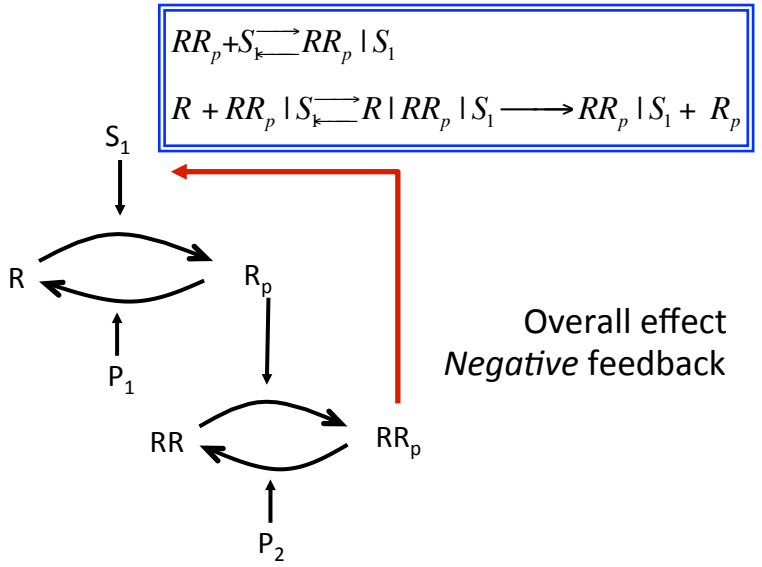


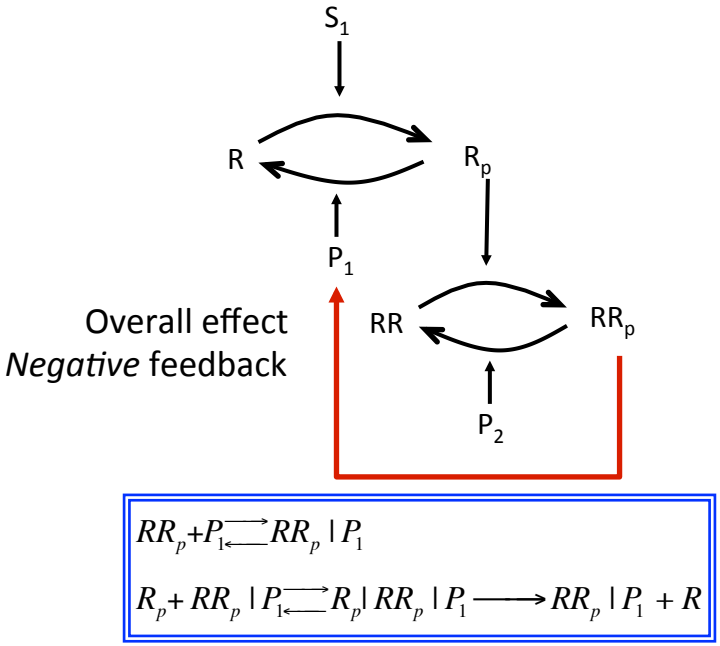
1-stage cascade
double phosphorylation step

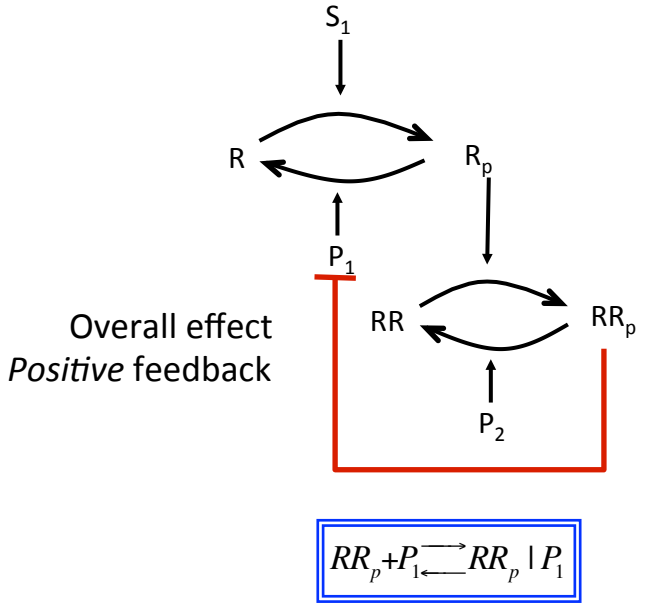
Brunel UNIVERSITY LONDON

Phosphorylation cascade + feedback









Brunel UNIVERSITY LONDON

Phosphorylation cascade, negative feedback. Inhibitor on 2nd stage

$$RRR_p + S_1 \underset{ki'}{\overset{ki}{\rightleftarrows}} RRR_p \mid S_1$$

$$R + S_1 \underset{k_2}{\overset{k_1}{\rightleftarrows}} R \mid S_1 \xrightarrow{k_3} R_p + S_1$$

$$R + P_1 \xleftarrow{k_3'} R_p \mid P_1 \underset{k_2'}{\overset{k_1'}{\rightleftarrows}} R_p + P_1$$

$$RR + R_p \underset{kk_2}{\overset{kk_1}{\rightleftarrows}} RR \mid R_p \xrightarrow{kk_3} RR_p + R_p$$

$$RR + P_2 \xleftarrow{kk_3'} RR_p \mid P_2 \underset{kk_2'}{\overset{kk_1'}{\rightleftarrows}} RR_p + P_2$$

$$RRR + RR_p \underset{kkk_2}{\overset{kkk_1}{\rightleftarrows}} RRR \mid RR_p \xrightarrow{kkk_3} RRR_p + RR_p$$

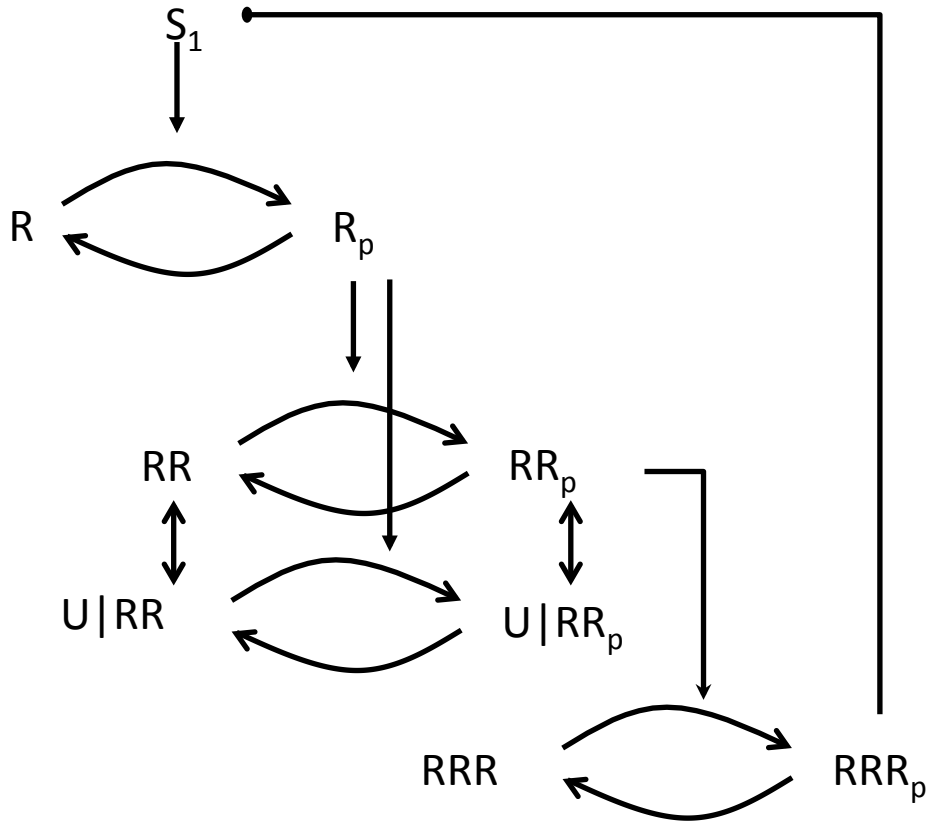
$$RRR + P_3 \xleftarrow{kkk_3'} RRR_p \mid P_3 \underset{kkk_2'}{\overset{kkk_1'}{\rightleftarrows}} RRR_p + P_3$$

$$U + RR \underset{ku_2}{\overset{ku_1}{\rightleftarrows}} U \mid RR$$

$$U + RR_p \underset{ku_2}{\overset{ku_1}{\rightleftarrows}} U \mid RR_p$$

$$U \mid RR + R_p \underset{kk_2}{\overset{kk_1}{\rightleftarrows}} U \mid RR \mid R_p \xrightarrow{kk_3} U \mid RR_p + R_p$$

$$U \mid RR + P_2 \xleftarrow{kk_3'} U \mid RR_p \mid P_2 \underset{kk_2'}{\overset{kk_1'}{\rightleftarrows}} U \mid RR_p + P_2$$



Brunel UNIVERSITY LONDON

Petri net demo!

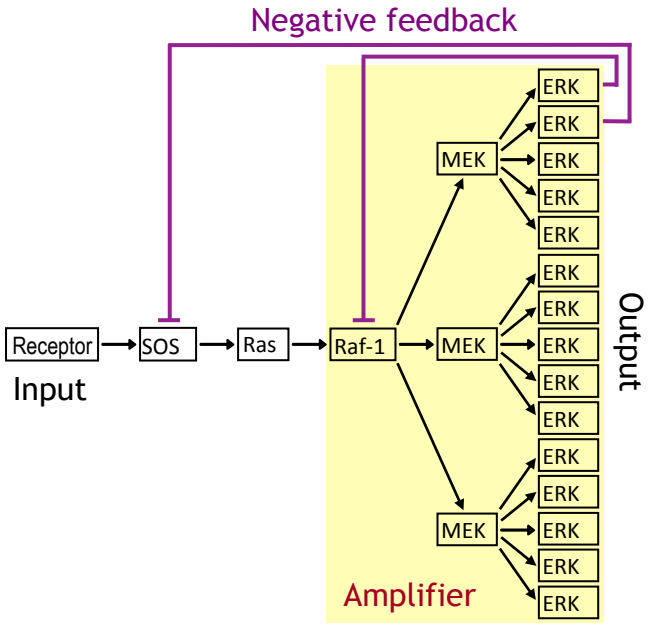
- 3 stage cascade (no dephosphorylation)
- 3 stage cascade
- 3 stage cascade + negative feedback
- Kholodenko model (continuous)
- Kholodenko model (stochastic)
- 3 stage cascade + negative feedback + inhibitor

Negative Feedback Amplifier

- in practice

ERK cascade well known biological amplifier

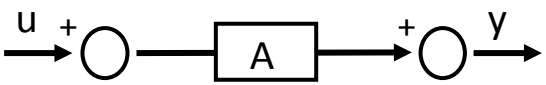
- Amplifies the original signal to create effective cellular responses.
- 1:3:5 are the approximate ratios of Raf-1, MEK and ERK in fibroblasts.
- Well known negative feedback loop: phosphorylation of SOS by ERK-PP (via MAPKAP1) resulting in the dissociation of the Grb2/SOS complex.
- New negative feedback loop: ERK-PP phosphorylates Raf-1 resulting in a hyper-phosphorylated inactive form of Raf (Dougherty *et al*. 2005)



Dougherty *et al*. (2005), Regulation of Raf-1 by Direct Feedback Phosphorylation, Molecular Cell 17 215-224

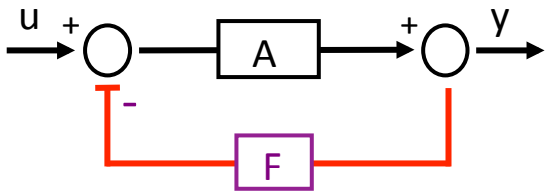
Negative Feedback Amplifier

Standard Amplifier



$y = u*A$

Negative Feedback Amplifier



$$y = u \frac{A}{1+A*F}$$

- Negative feedback amplifier from electronics
- Amplifier with a negative feedback loop from the output of the amplifier to its input.

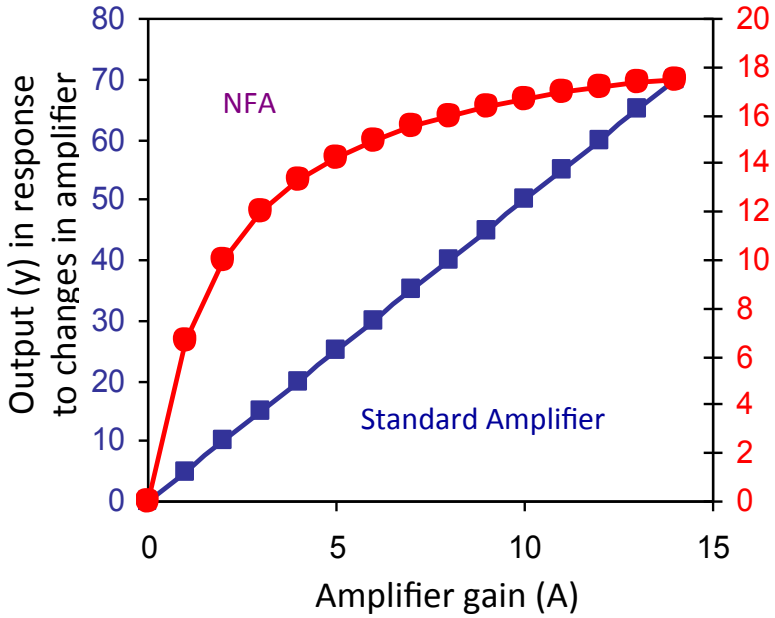
- NF loop → a system much more robust to disturbances in the amplifier.

- NFA was invented in 1927 by Harold Black of Western Electric.
- Originally used for reducing distortion in long distance telephone lines.

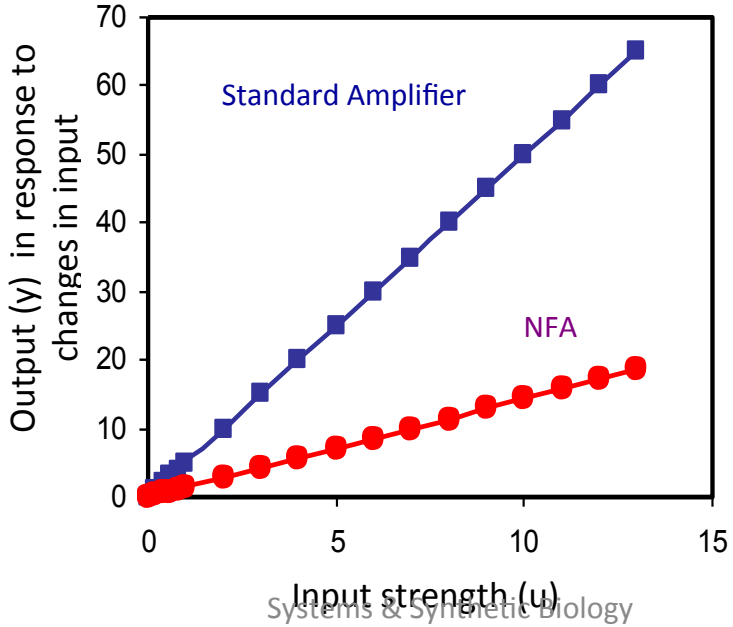
- NFA a key electrical component used in a wide variety of applications

Figure 1

C)

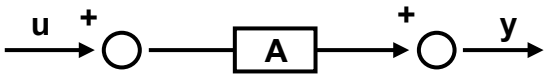

Output (y) in response to changes in amplifier
Amplifier gain (A)
NFA
Standard Amplifier
0
10
20
30
40
50
60
70
80
0
2
4
6
8
10
12
14
16
18
20
0
5
10
15

D)


Output (y) in response to changes in input
Input strength (u)
Standard Amplifier
NFA
0
10
20
30
40
50
60
70
0
5
10
15

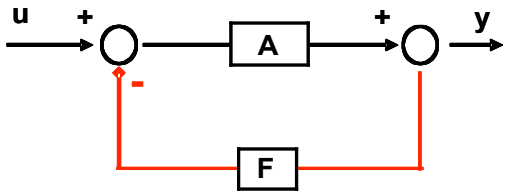
The negative feedback imparts signalling robustness

Standard Amplifier



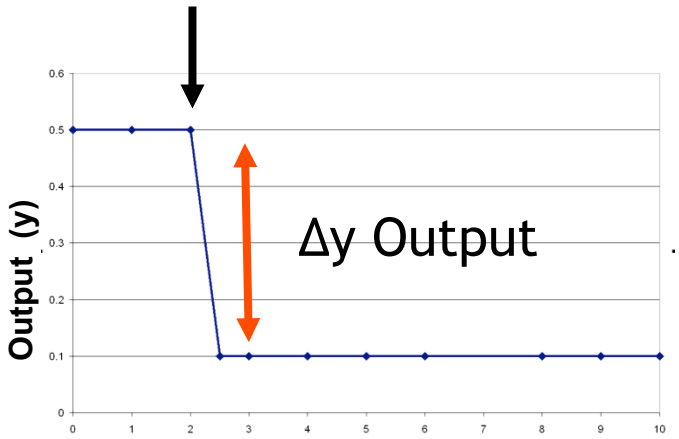
y=A*u

Negative Feedback Amplifier

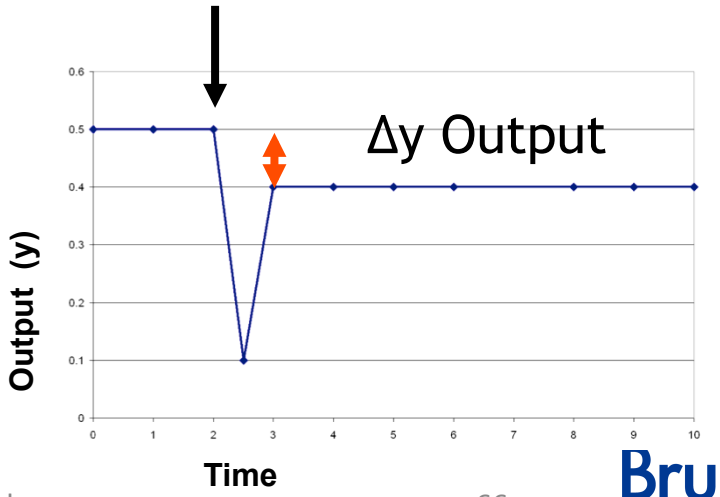


y=A*u/(1+A*F)

Sudden drop in Amplifier (A) gain

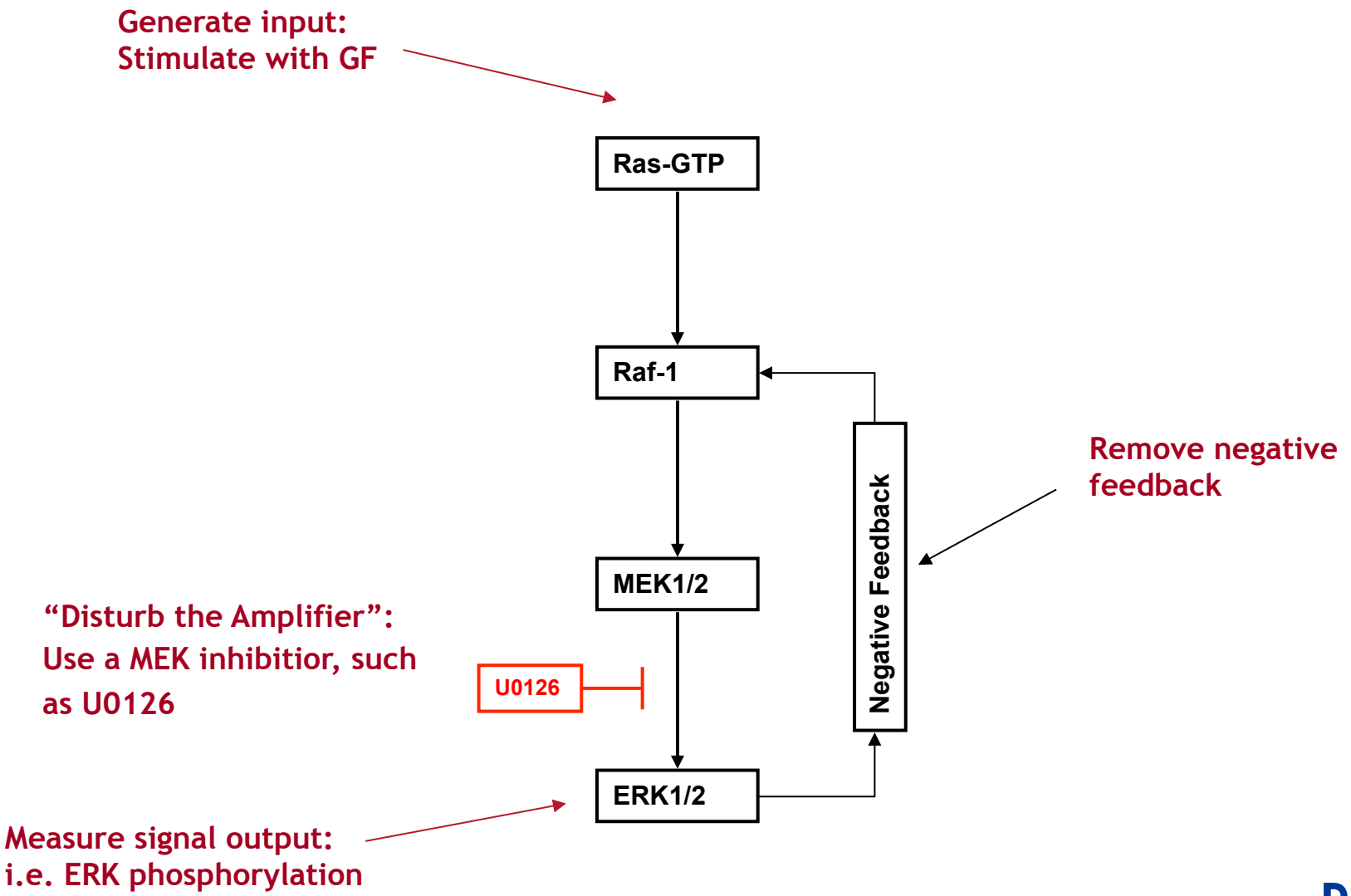


Sudden drop in Amplifier (A) gain



Brunel UNIVERSITY LONDON

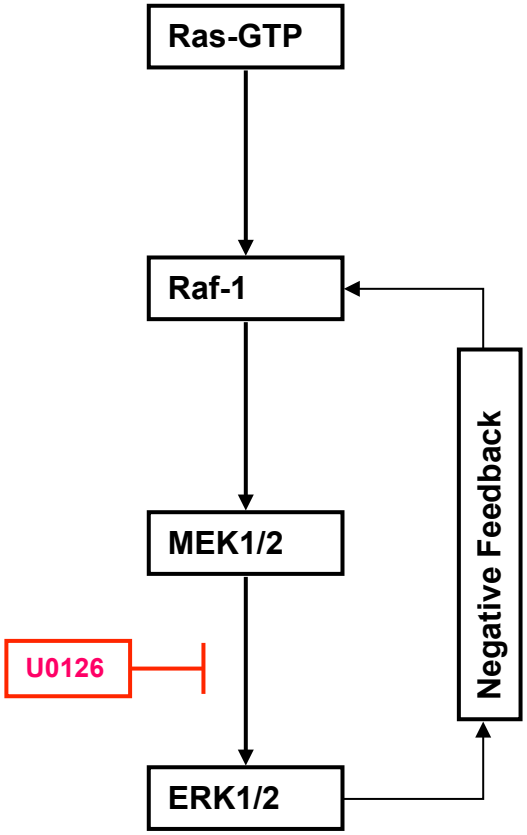
How to test if the ERK pathway is a NFA?

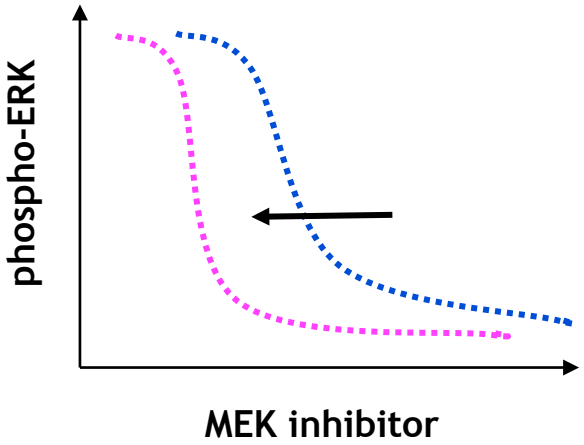


Brunel UNIVERSITY LONDON

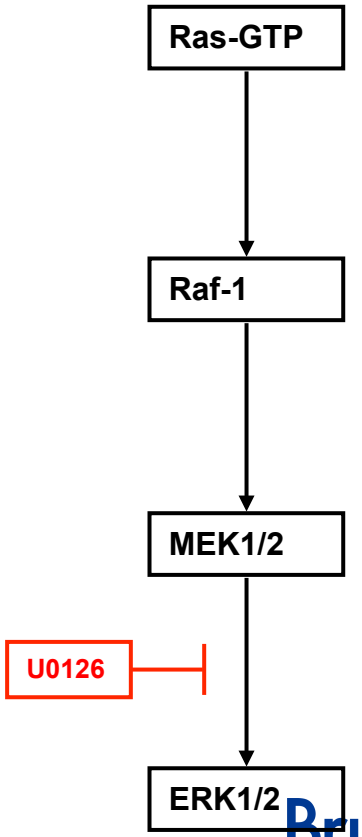
Hypothesis: Breaking the feedback should sensitise the ERK pathway to MEK-inhibitor

Feedback intact





Feedback removed



How to test if the ERK pathway is a NFA?

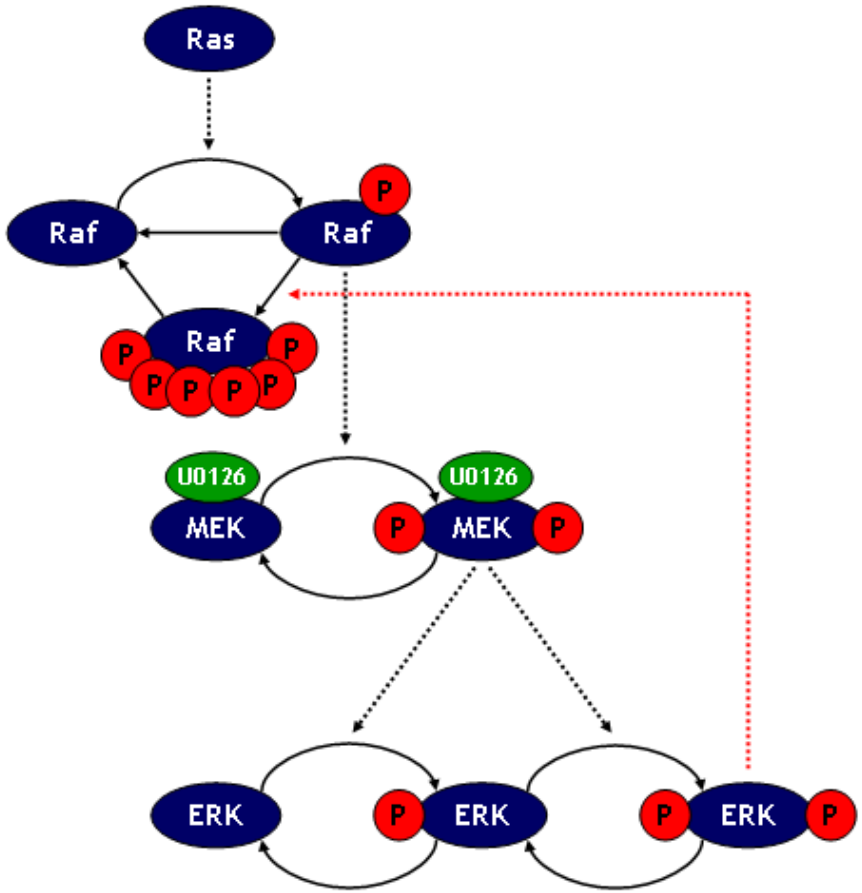
Strategy

In vivo system that allows us to compare feedback broken to feedback intact model.

Computational Model of ERK pathway with/without feedback

Computational Modeling 1: Build the model

- Non-linear ordinary differential equations (ODE's).
- ODE's were solved using Math Lab and Gepasi.
- Models are based on the Schoeberl et al. (2002) model
- Mass Action Kinetics instead of Michaelis Menten
- Kinetic parameters are from literature, previous models and "guesstimates"



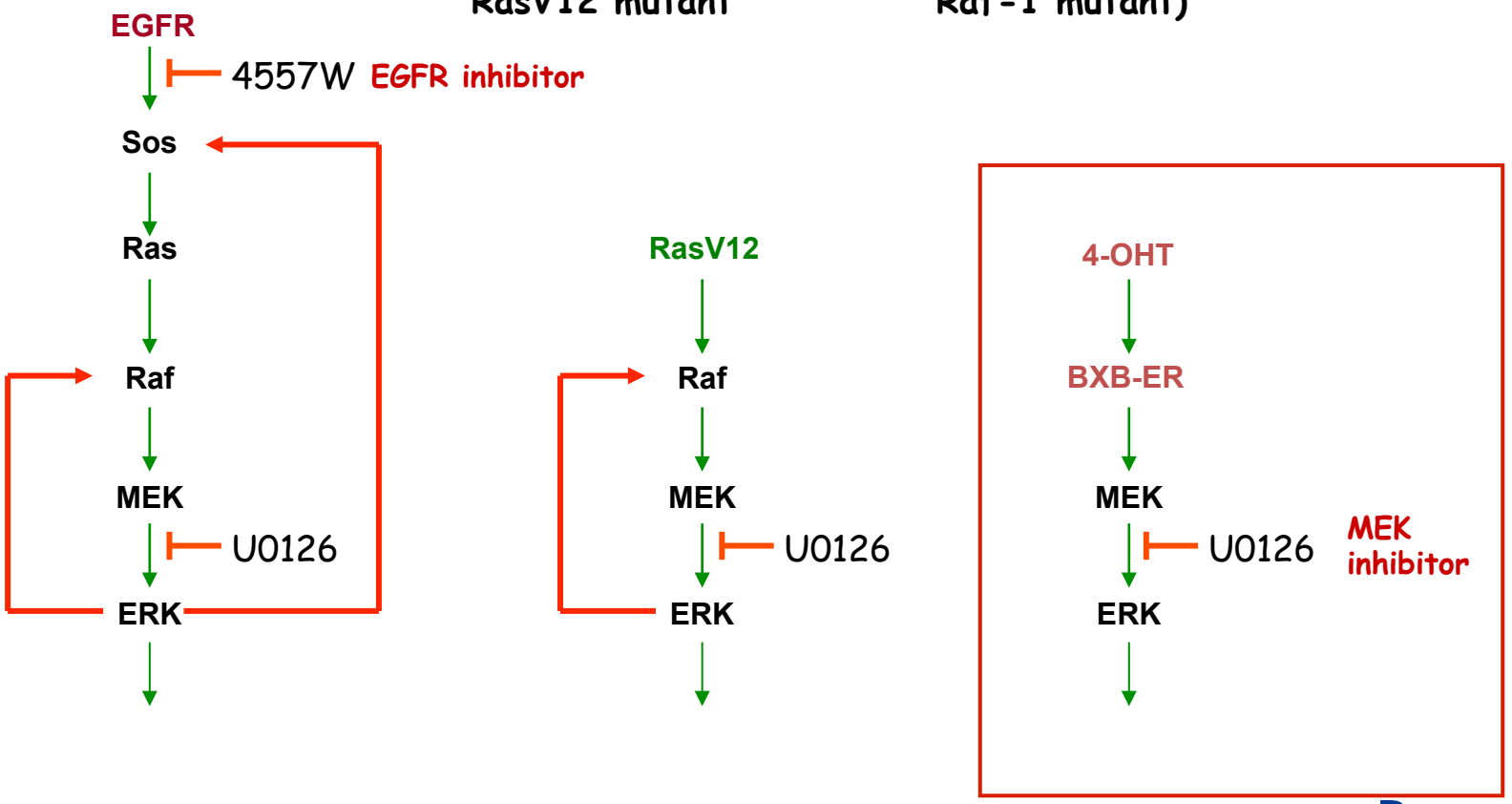
Schoeberl *et al*. (2002), Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors, Nature Biotechnology 20, 370-375

The experimental systems

Negative feedback loops intact

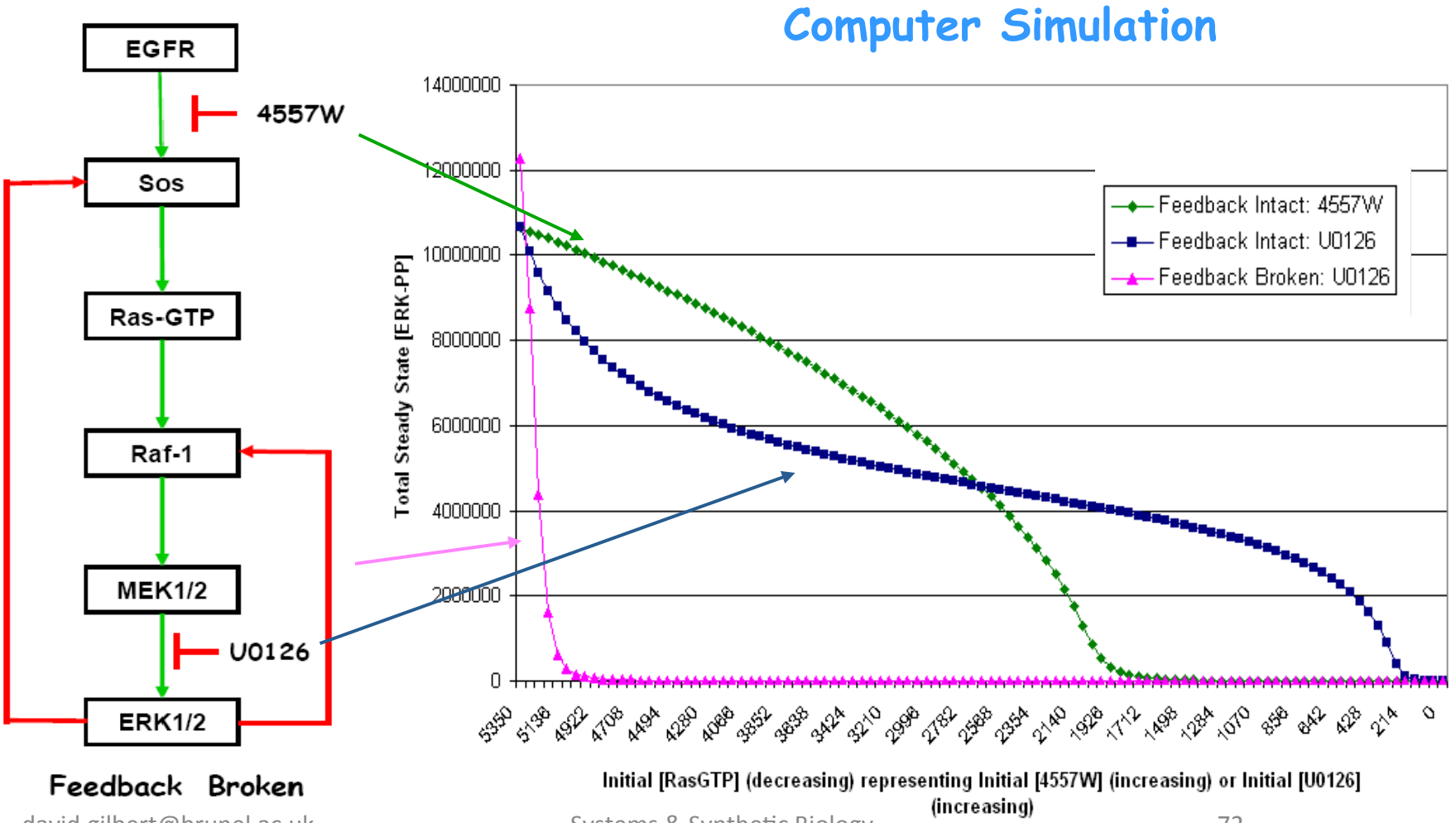
One feedback loop eliminated by constitutively active RasV12 mutant

Both feedback loops eliminated by BXB-ER (4-OHT regulatable Raf-1 mutant)

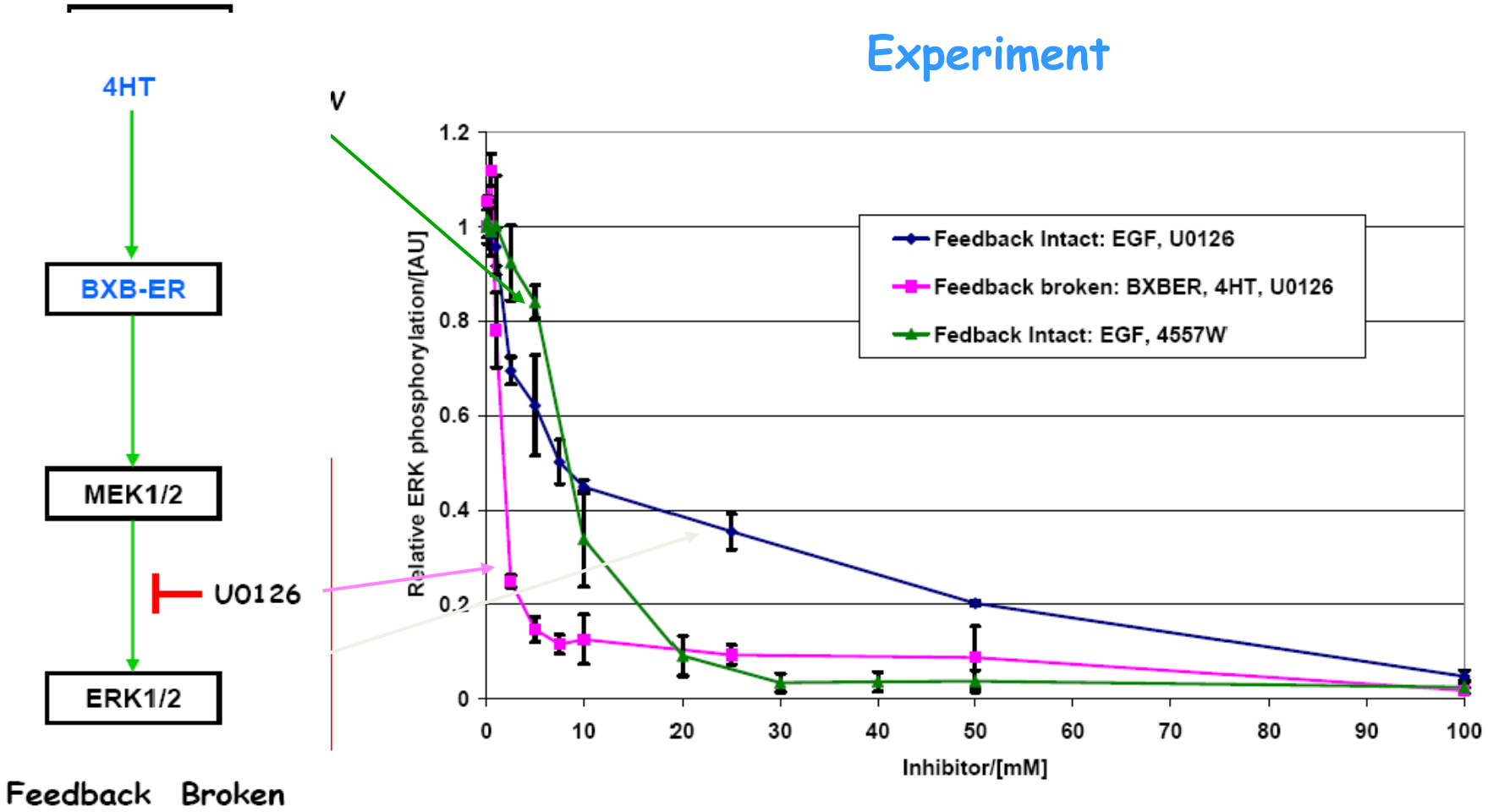


Brunel UNIVERSITY LONDON

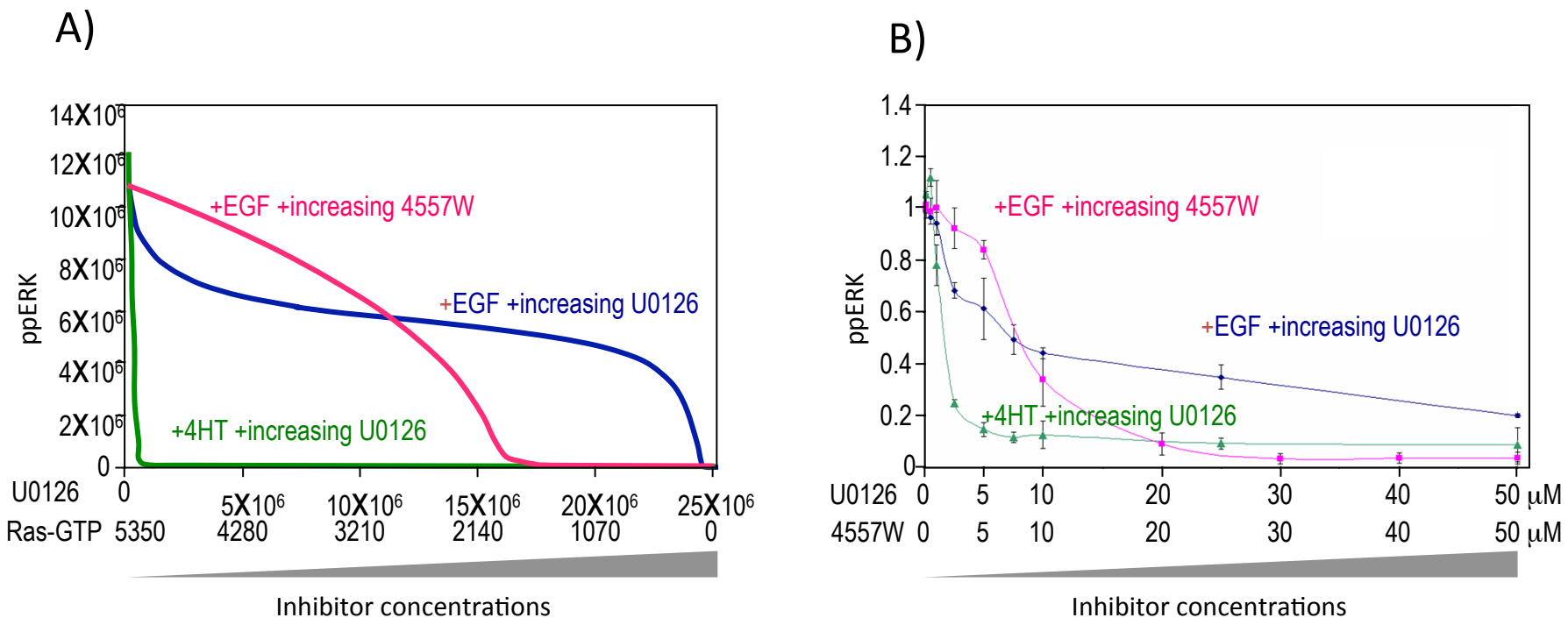
Ablation of feedback by BXBER decreases robustness to MEK-inhibitor U0126



Ablation of feedback by BXBER decreases robustness to MEK-inhibitor U0126



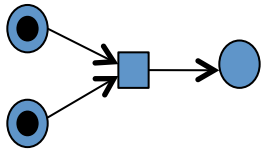
(A) Model prediction (B) Biochemical validation



Model checking

Molecules/Levels
CTL, LTL

Qualitative



Time-free

Timed, Quantitative

Abstraction

Abstr

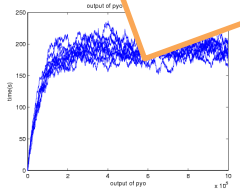
How can we check that models 'do what we want'?

Molecul

Stochas

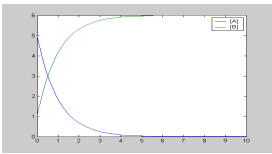
Concentrations
Deterministic rates

Continuous LTLc



Approximation

$$\frac{d[A]}{dt} = -k_1 \times [A]$$



DiscreteState Space

Continuous State Space

Gilbert, Heiner and Lehrack. ``A Unifying Framework for Modelling and Analysing Biochemical Pathways Using Petri Nets.'' Proc CMSB 2007

Brunel UNIVERSITY LONDON

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

To formally express time properties we use a temporal logic

"*I am hungry*."

"*I am always hungry*", "*I will eventually be hungry*",

"I will be hungry until I eat something".

Linear time logics restricted to single time line.

Branching logics can reason about multiple time lines.

"*There is a possibility that I will stay hungry forever*."

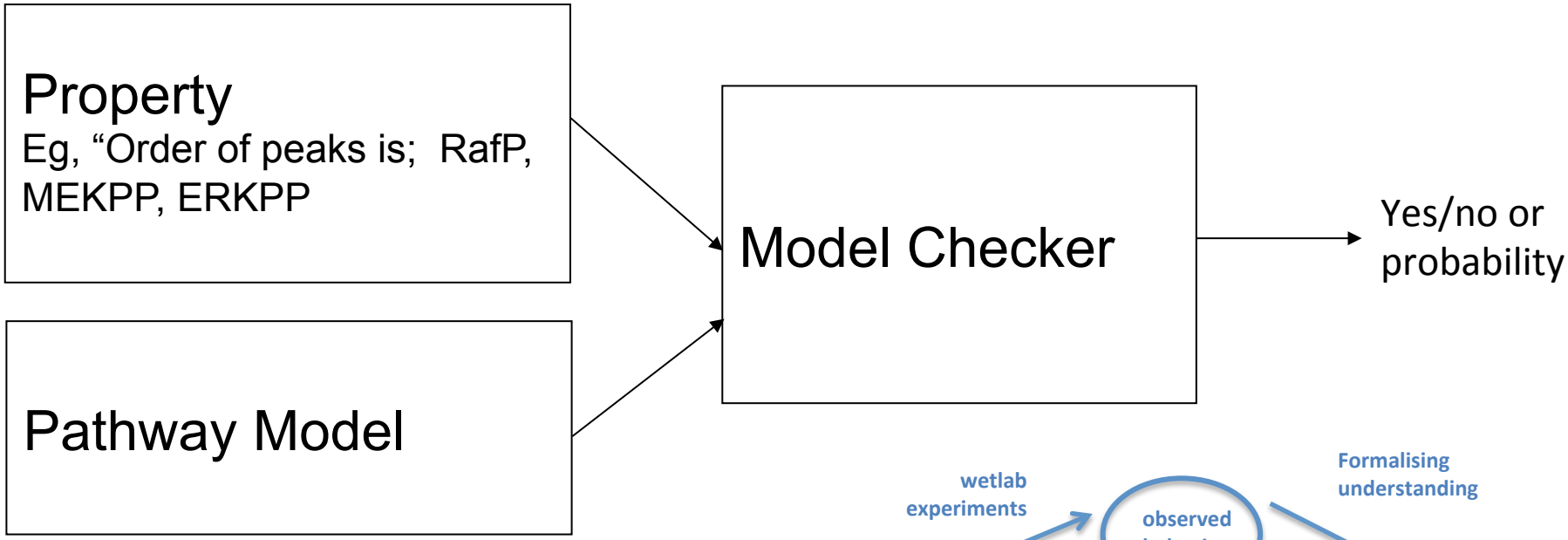
"*There is a possibility that eventually I am no longer hungry*."

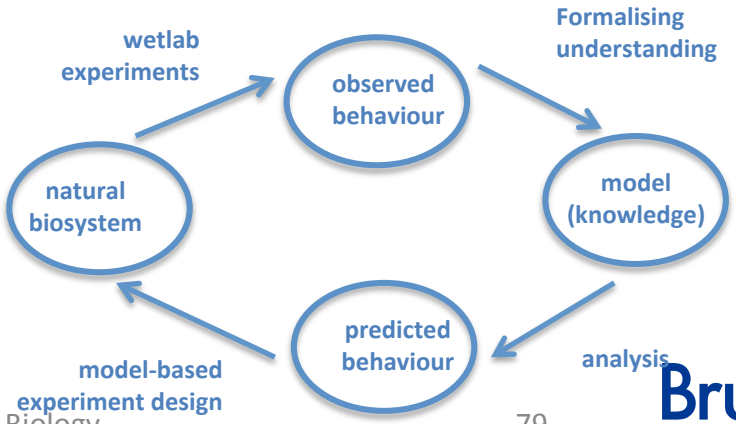
Various logics :

- Computational Tree Logic (CTL)
- Continuous Stochastic Logic (CSL)
- Linear-time Temporal Logic (LTL)

each with different expressivity.

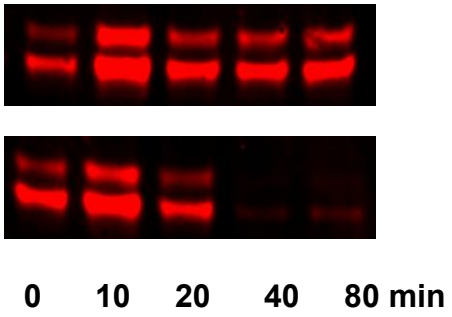
Model Checking Biochemical Pathways





Why model check in Synthetic Biology?

- Biologists will often talk in qualitative or semi-quantitative language (trends).
  - "this protein peaks after 5 minutes, then falls to half concentration"
  - Often quite certain about time,
  - But not about concentrations



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

Properties…

Examples:

- After 100 seconds the concentration of Protein1 is stable
- Protein1 peaks and falls
- Protein1 peaks and stays constant
- Protein1 peaks before Protein2
- Protein1 oscillates 4 times in 5,000 seconds
- Molecules of Protein2 are required for molecules of Protein1 to be created

Analytical vs Simulative Model Checking

- **Analytical**:
  - Exact probabilities & prove properties
  - A model state is an association of #molecules/levels to each of the species
    - **Protein1** has 10 molecules & **Protein2** has 20 molecules
  - Analytical assesses every state that the model can be in (reachable states)
  - State space can grow even worse than exponentially with increasing molecules, or even be infinite!
  - Stochastic model checking with even as little as 12 molecules/levels can be impossible with today's technology

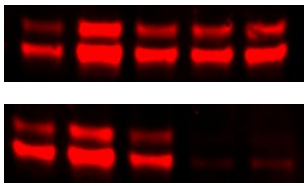
- **Simulative**:
  - Instead of analysing the constructed state space, analyse simulation outputs
  - Simulate the model X times and check these simulations
  - Simulation run = finite path through the state space
  - Can't prove probabilities

Simulative Model Checking

- **In-line**: check the observations as they arrive
  - Requires complex computational machinery: 'combine' simulator & model checker
  - Good for biochemical observations
  - Don't always need to finish the experimental run

- **Off-line**: check the observations after all have been generated
  - Easier to implement computationally (simulate then check)
  - Need to always define when to 'stop' generating observations

Simulation-based Model Checking Biochemical Pathways



Property

Eg, “Order of peaks is RafP, MEKPP, ERKPP”

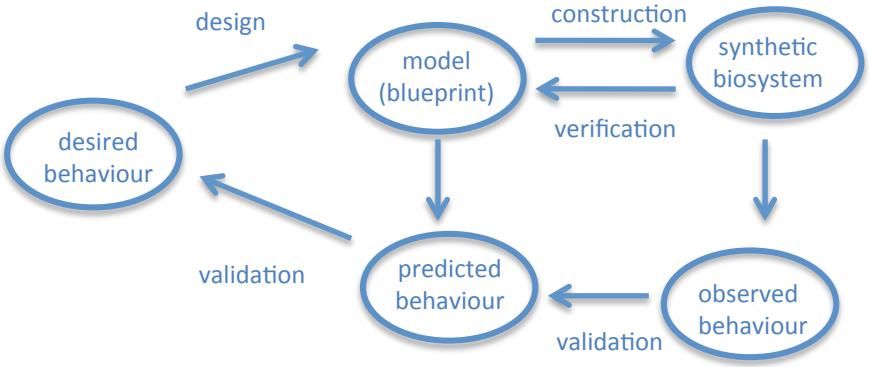
Time series data

Model

Lab

Behaviour Checker

Yes/no or probability



Brunel UNIVERSITY LONDON

MC2: Monte Carlo Model Checker

- ***Offline*** Monte Carlo Model Checker for PLTLc properties.
- Operates on a finite set of simulations – simulative approach
- Typically, many stochastic simulations to approximate probabilities
  - Approximate probability = fraction of simulations which satisfy the property over the #simulations
- Monte Carlo approximation – 2 approximations made:
  - finite number of simulations
  - Simulations of finite length

(P)LTL Linear Temporal Logic

- G (ϕ) : ϕ always happens
- F (ϕ) : ϕ happens at some time
- X (ϕ) : ϕ happens in the next time point
- ϕ_1 U ϕ_2 : ϕ_1 happens until ϕ_2 happens

- Protein stability:

  $P_{=?}$ [time >= 100 → ([Protein] >= 4 ^ [Protein] <= 6)]

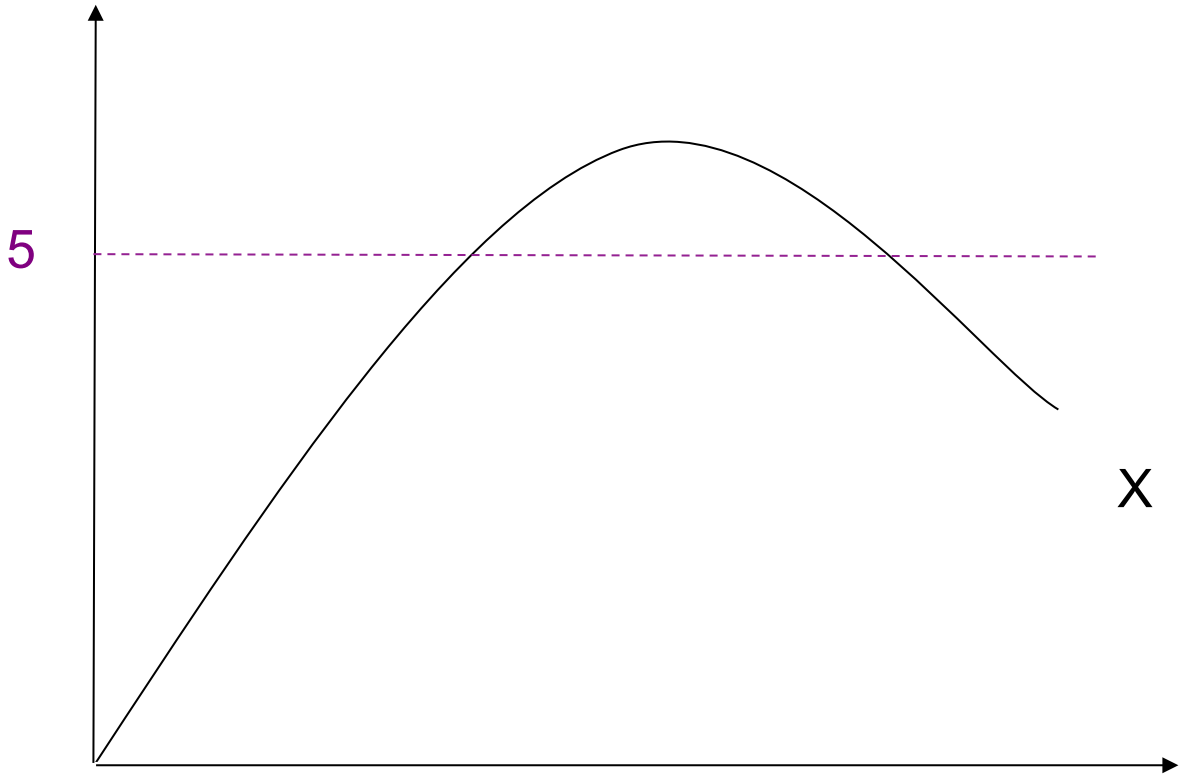
- Protein concentration rises to a maximum value and then remains constant:

  $P_{=?}$ [(d[Protein] > 0) U (G([Protein] >= 0.99*max[Protein]))]

MC2 with ODE Output

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

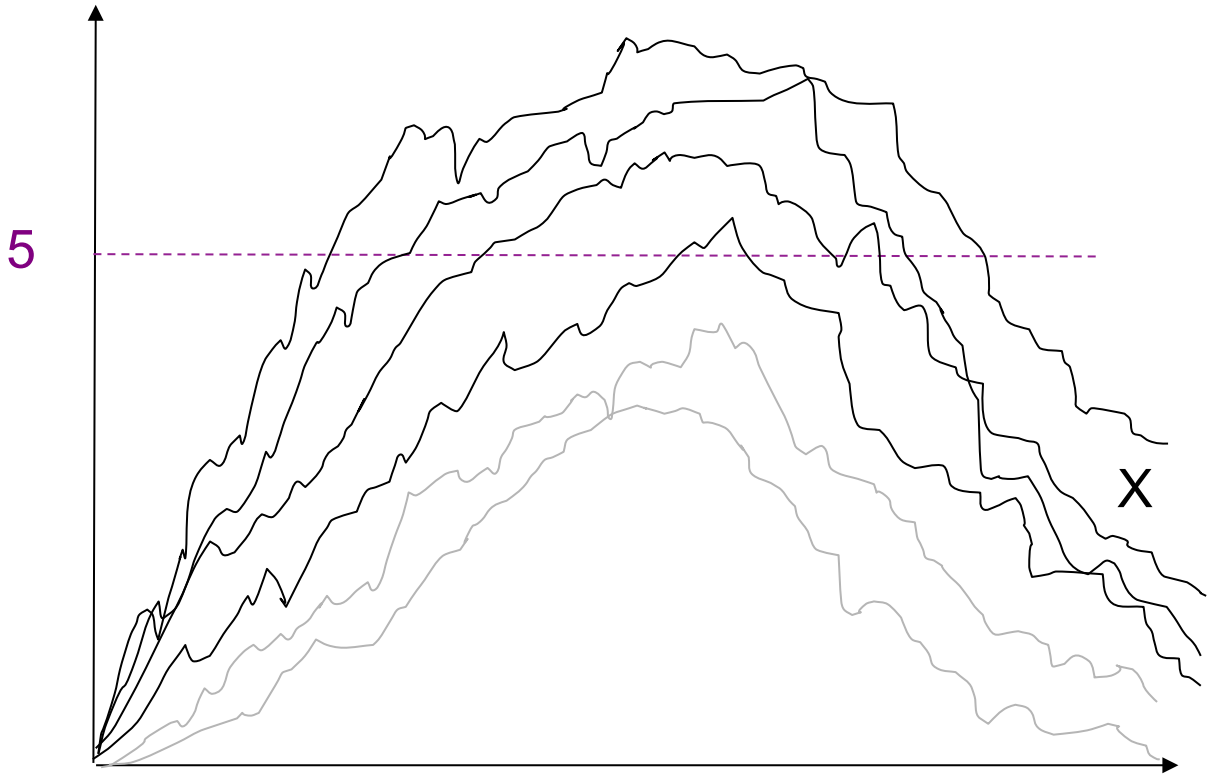
=> P = 1



MC2 with Gillespie Output

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

=> P = 4/6

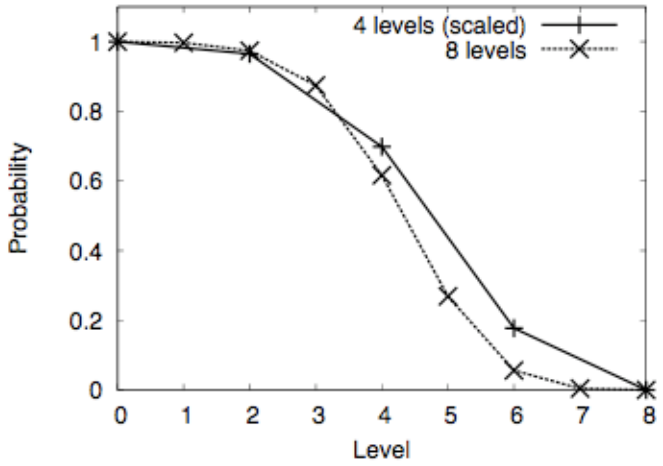


Stochastic Analysis

Check the property S2:

"*What is the probability that RafP will reach concentration X while MEKPP and ERKPP remain at 0?*"

$P_{=?}$ [(MEKPP = 0 ^ ERKPP = 0) U (RafP > X)
{RafP = 0 ^ MEKPP = 0 ^ ERKPP = 0}]

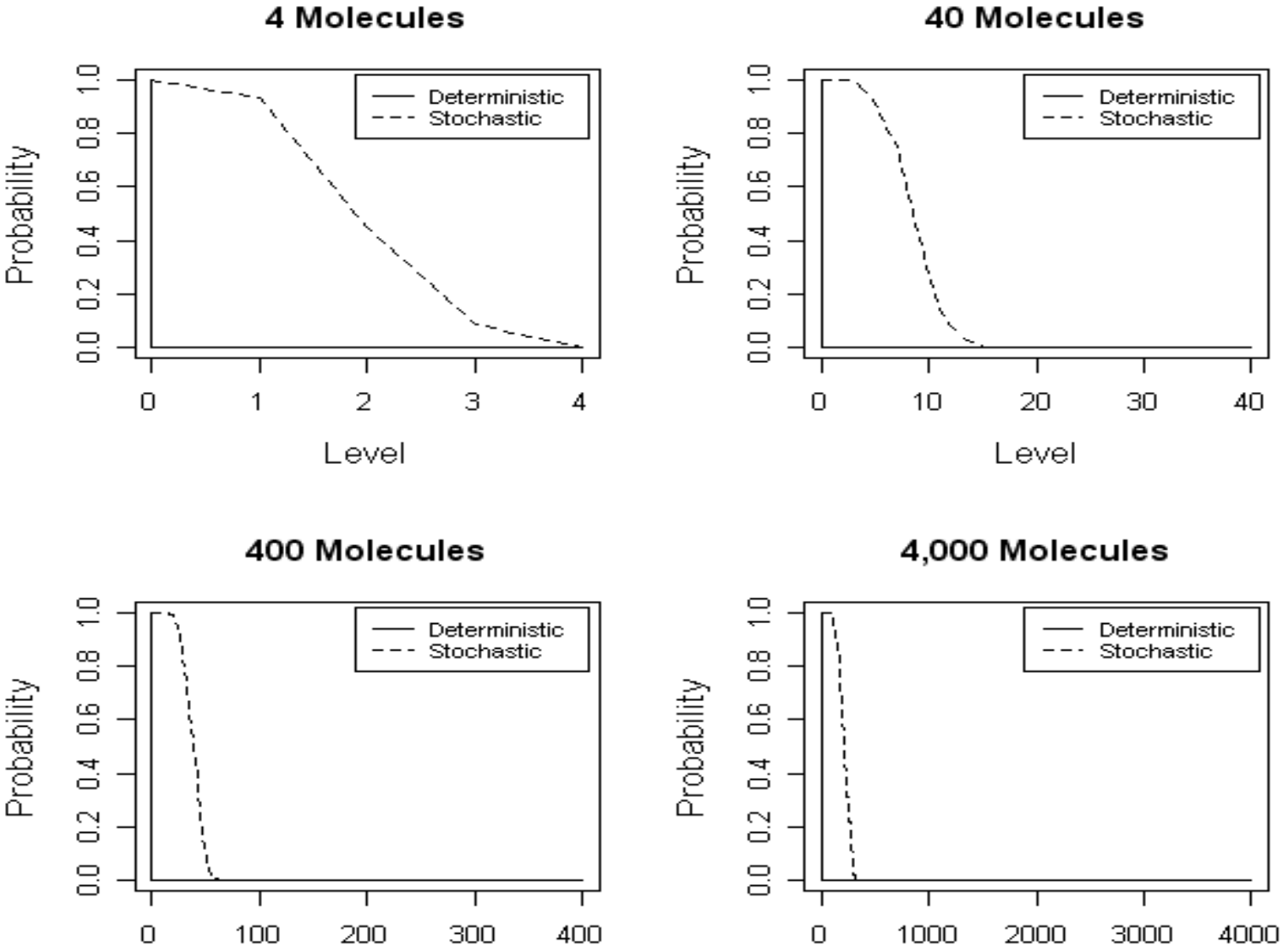


4 LEVELS : 4 HOURS
8 LEVELS : 24 HOURS
16 LEVELS: ??

David Gilbert, Monika Heiner and Sebastian Lehrack (2007). A Unifying Framework for Modelling and Analysing Biochemical Pathways Using Petri Nets. Proc CMSB 2007

Brunel UNIVERSITY LONDON

Checking the property at varying levels



Monika Heiner, David Gilbert, and Robin Donaldson (2008), Petri Nets for Systems and Synthetic Biology. In M Bernardo, P Degano, and G Zavattaro (Eds.): Formal Methods for Systems Biology SFM 2008, Springer LNCS 5016

Model searching

Peaks at least once

(rises then falls below 50% max concentration)

$P_{>=1}[\ ErkPP <= 0.50*max(ErkPP) \wedge d(ErkPP) > 0\ U\ (\ ErkPP = max(ErkPP) \wedge F(\ ErkPP <= 0.50*max(ErkPP)\ )\ )\ ]$

- ***Brown***
- ***Kholodenko***
- ***Schoeberl***

Rises and remains constant

(99% max concentration)

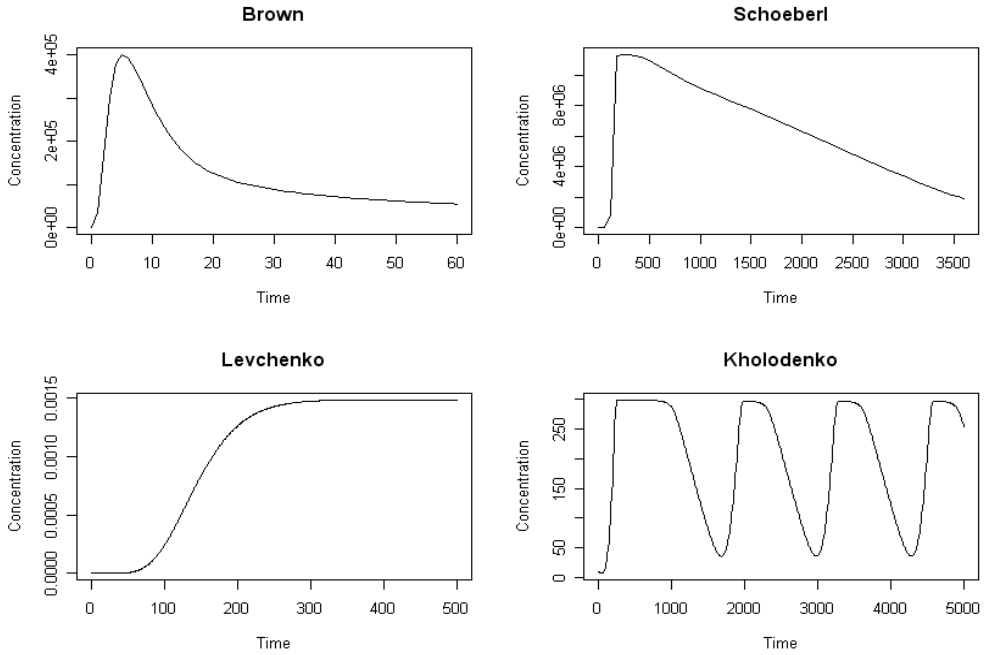
$P_{>=1}[ErkPP <= 0.50*max(ErkPP) \wedge (\ d(ErkPP) > 0\ )\ U\ (\ G(ErkPP >= 0.99*max(ErkPP))\ )\ ]$

- ***Levchenko***

Oscillates at least 4 times

$P_{>=1}[\ F(\ d(ErkPP) > 0 \wedge F(\ d(ErkPP) < 0 \wedge \ldots\ )\ )\ ]$

- **Kholodenko**



Brunel UNIVERSITY LONDON

Qualitative to quantitative descriptions in PLTL

- **Qualitative**:
  Protein rises then falls
  P=? [(d(Protein) > 0) U (G(d(Protein) < 0))]

- **Semi-qualitative**:
  Protein rises then falls to less than 50% of peak concentration
  P=? [(d(Protein) > 0) U (G(d(Protein) < 0) ⋀ F ([Protein] < 0.5 ∗ max[Protein]))]

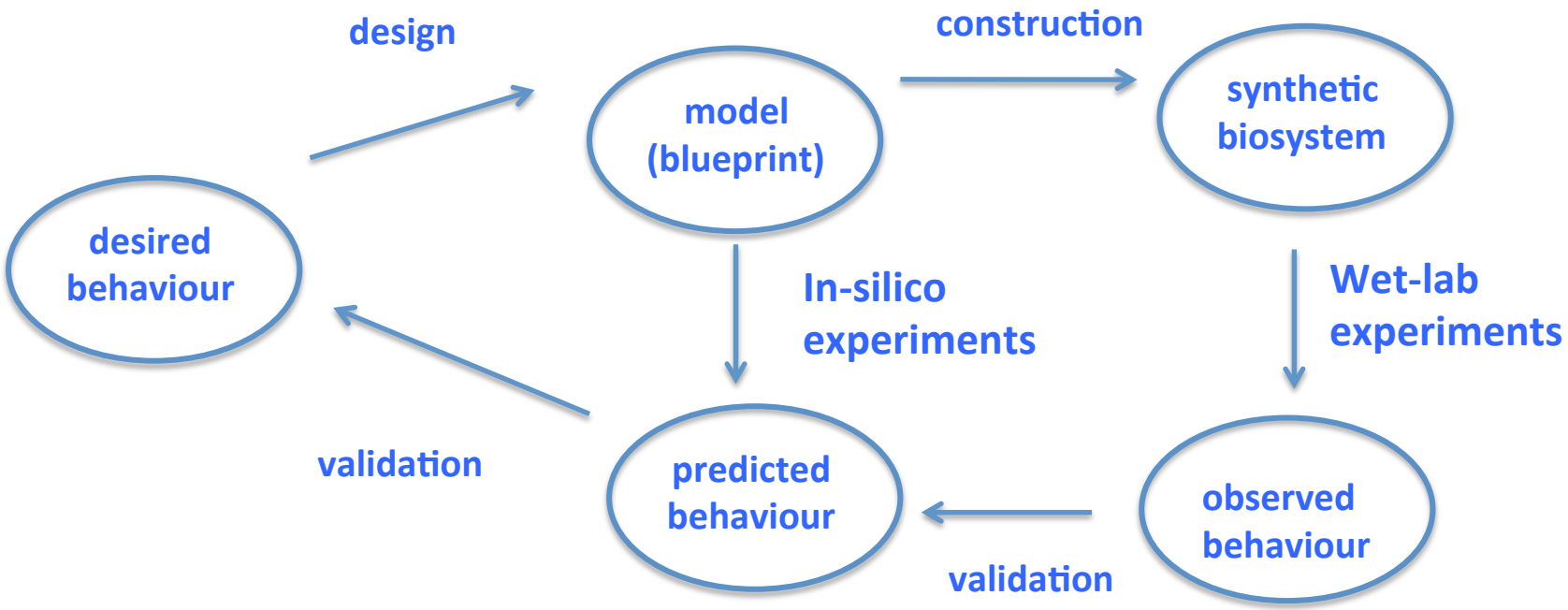
- **Semi-quantitative**:
  Protein rises then falls to less than 50% of peak concentration by 60 minutes
  P=? [(d(Protein) > 0) U (G(d(Protein) < 0) ⋀ F (time = 60 ⋀ Protein < 0.5 ∗ max(Protein)))]

- **Quantitative**:
  Protein rises then falls to less than 100µMol by 60 minutes
  P=? [(d(Protein) > 0) U (G(d(Protein) < 0) ⋀ F (time = 60 ⋀ Protein < 100))]

Model checking demo

- Queries – deterministic
- Queries - stochastic
- Deterministic model checking
- Stochastic model checking:w

So - Synthetic Biology

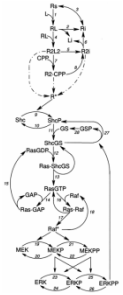


Brunel UNIVERSITY LONDON

Current Synthetic Biology development cycle

Model construction

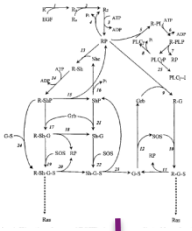
Biosystem



observations

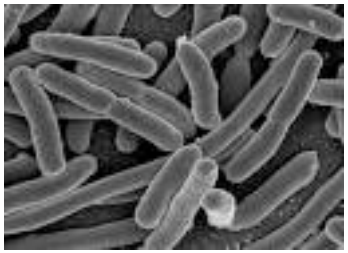
validate

modify



construct

verify



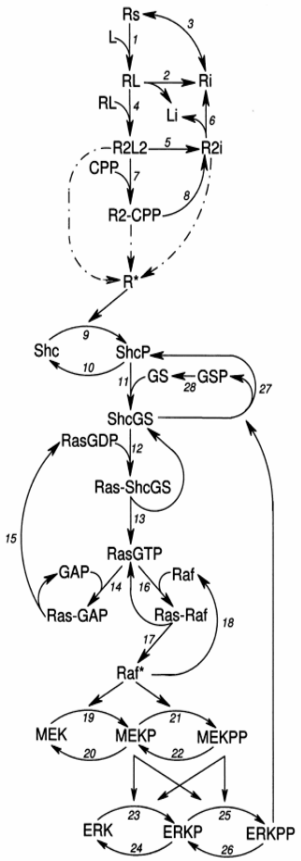
modify

construct

verify

Our future vision

Model construction



Component library

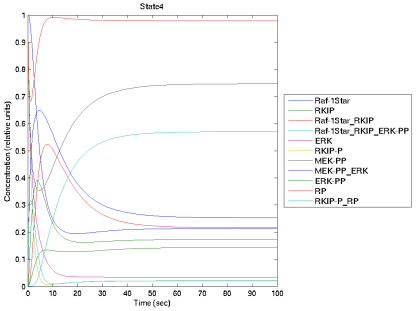
observations

validate

construct

verify

Desired behaviour



Biological system

Standard biological parts

Synthetic biology compilation

- From behaviour to gene sequence…

Automated construction of Models / Designs

Given some target behaviour of desired system

Build model (from scratch or a suitable skeleton):

- Topology (hard!)
- Rates (easier)

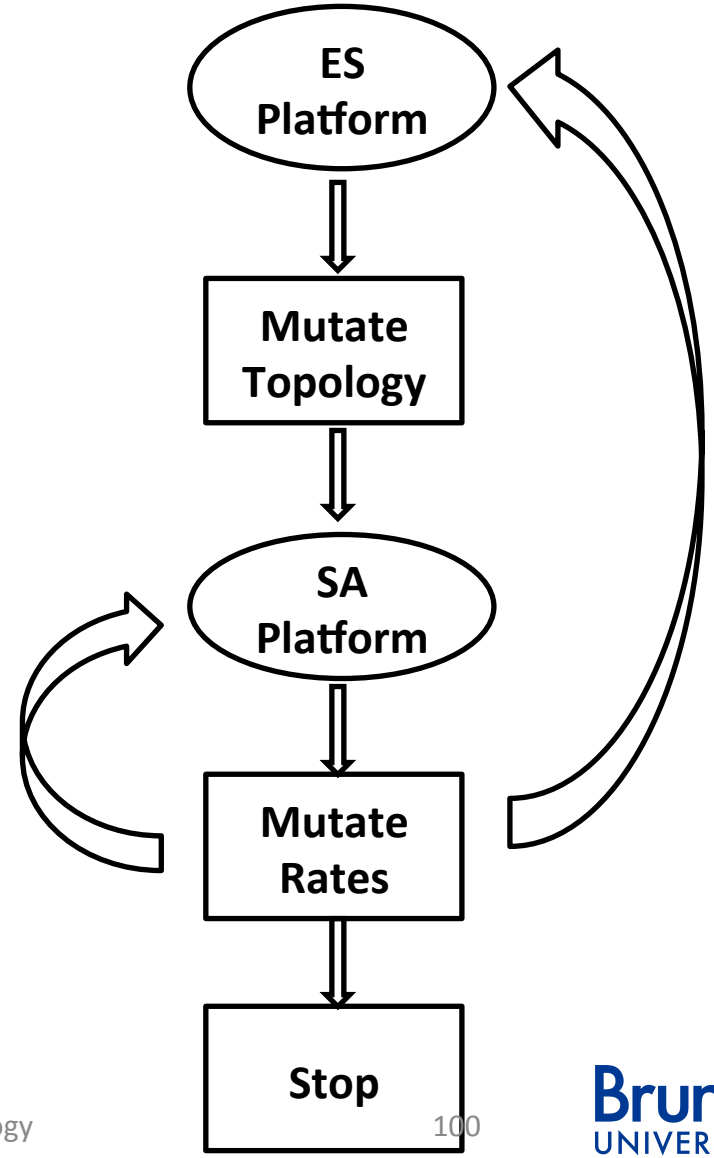
Topologies

Approaches – no knowledge of connectivities or complexes

- **Piece-wise (bottom-up)**
  - Template patterns from library
- **(Large) Network based**
  - Generate random possible networks, then mutate towards target behaviour

Hybrid Optimization of Topology and Kinetic Rates based on Evolution Strategy(ES) and Simulated Annealing(SA)

Algorithm: A hybrid piecewise modelling framework
Require: CompLib, Composition Rules
Ensure: BioNbest
1: Initiate the population;
2: **while** Not reached maximum generation (ES layer) **do**
3: **for** Each individual in the population **do**
4: Mutate the topology of individual by Addition or Subtraction;
5: Check the mutated topology of the individual;
6: Evaluate the mutated individual;
7: **if** The kinetic rates are required to be optimized **then**
8: **while** Not reach minimum temperature (SA layer) **do**
9: Optimize the kinetic rates by Gaussian distribution;
10: Evaluate the mutated kinetic rates;
11: **end while**
12: **end if**
13: **end for**
14: Crossover the individuals;
15: Select offspring for next generation;
16: **end while**
17: Return BioNbest



Biochemical Models
Construction Based on Reuse of Components

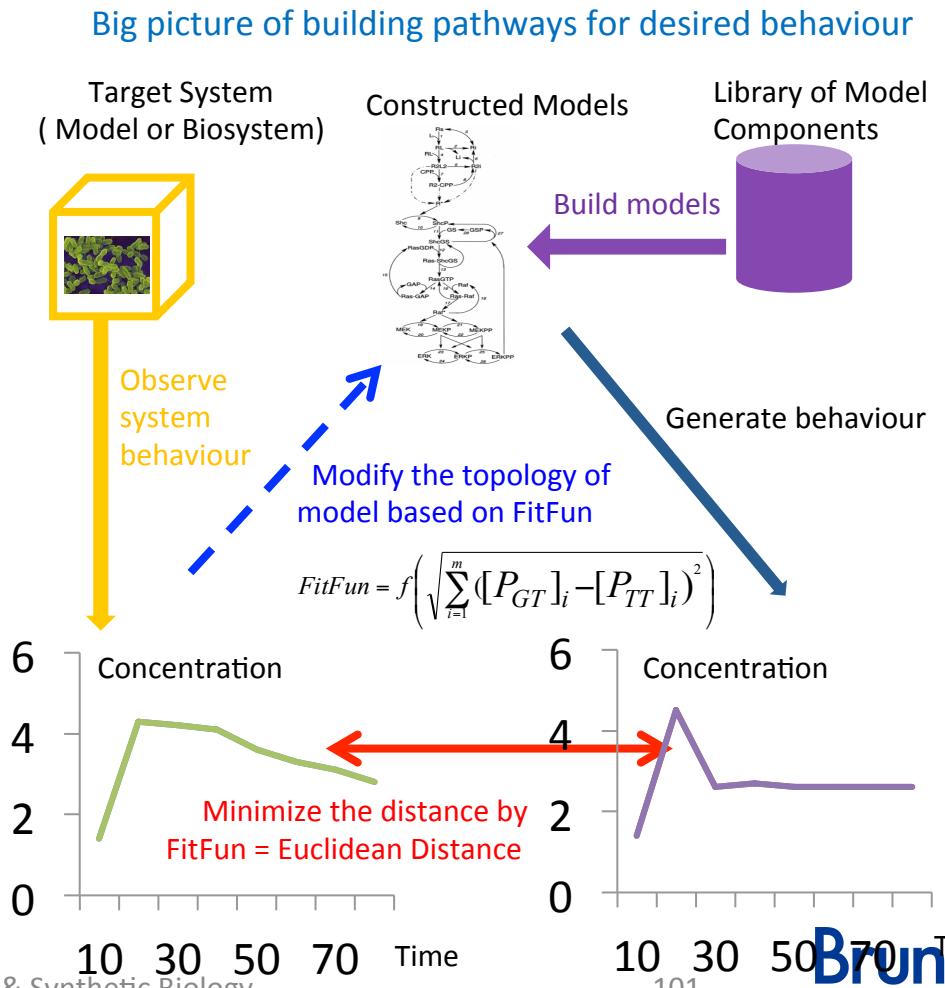
Zujian Wu

Aims:

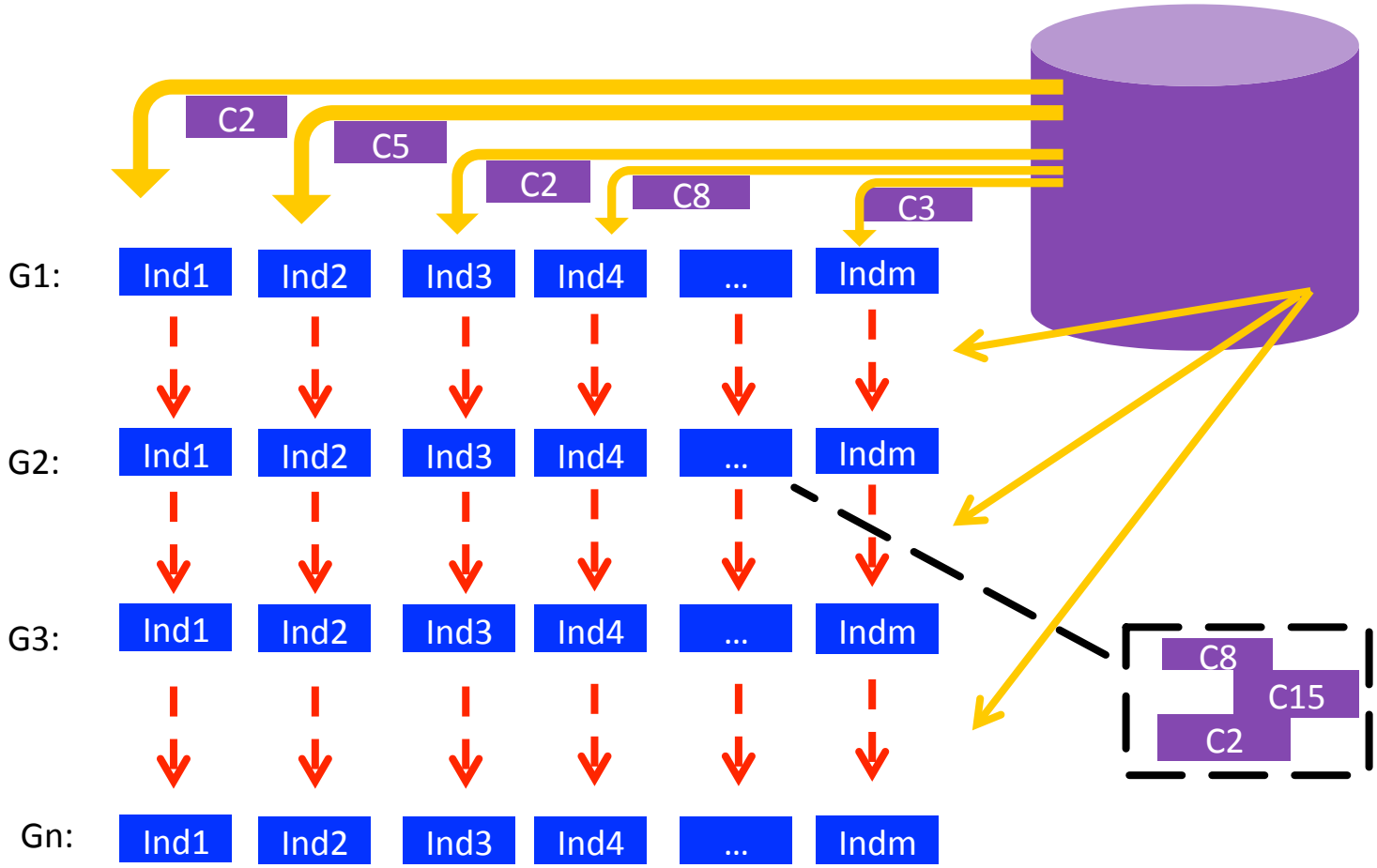
- To achieve the target-driven construction of biochemical models by reference to their desired behaviours
- To address the construction problem by
  - building a library for storing reliable biochemical functional submodels (as components)
  - intelligently selecting, combining and mutating these submodels to generate complex systems

Methods:

- Petri Nets: Components & Model
- Simulated Annealing: Kinetic rates
- Evolutionary Algorithms: Topology



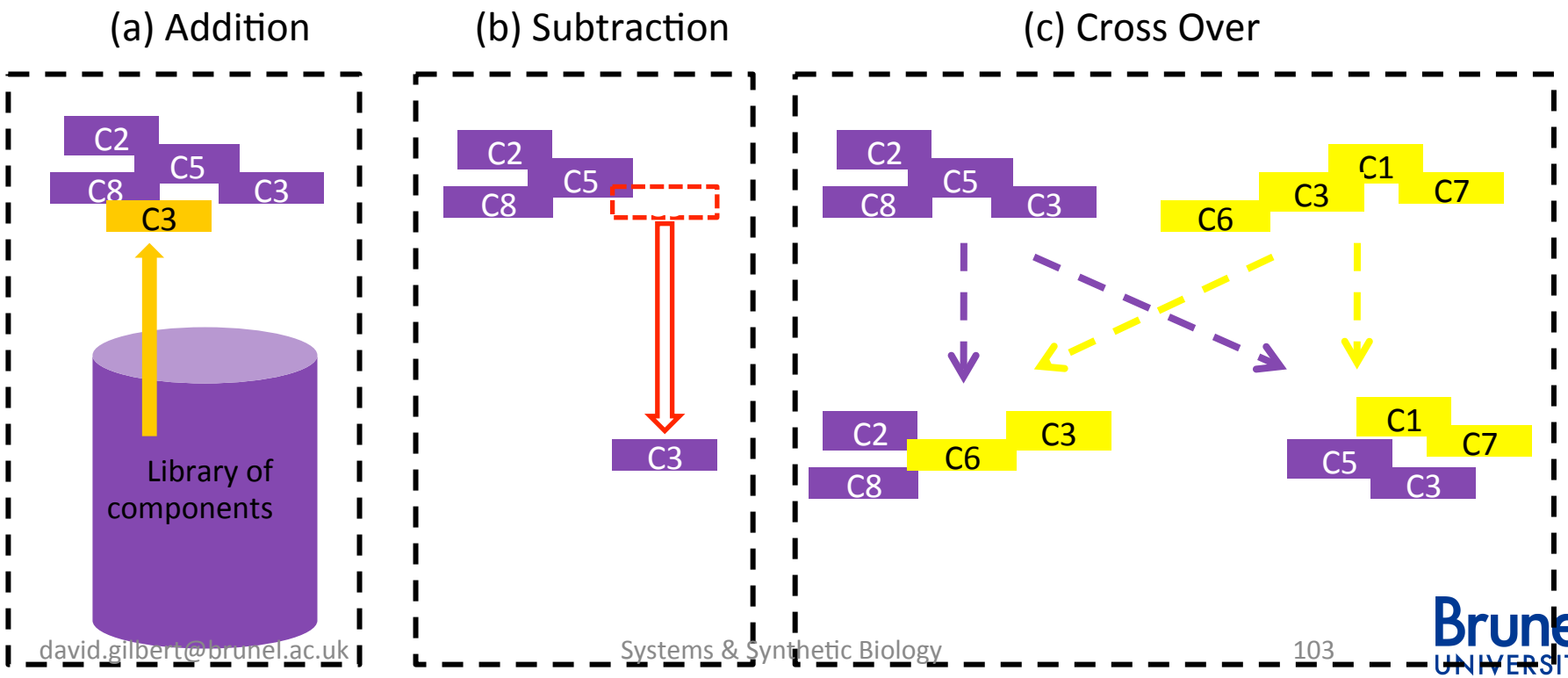
Evolution Strategy – Topology Optimization



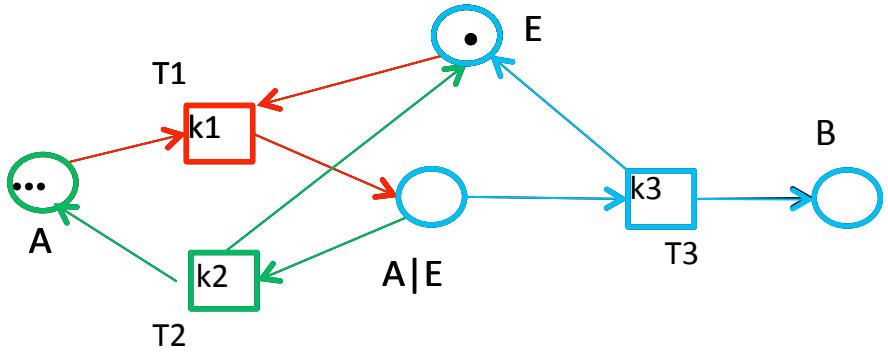
Brunel UNIVERSITY LONDON

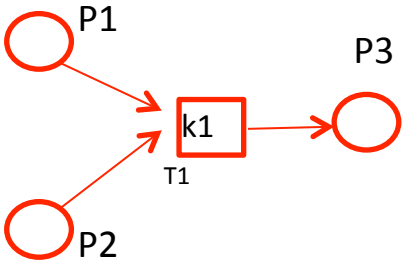
Genetic operators

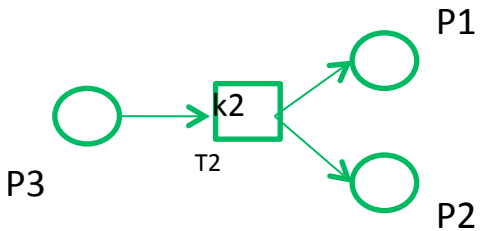
- Addition – composing component to an existing network
- Subtraction – removing component from an existing network
- Cross Over – recombining two networks by 'cut and splice'

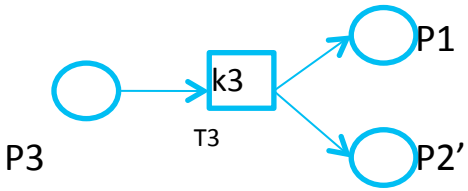


Basic building block patterns





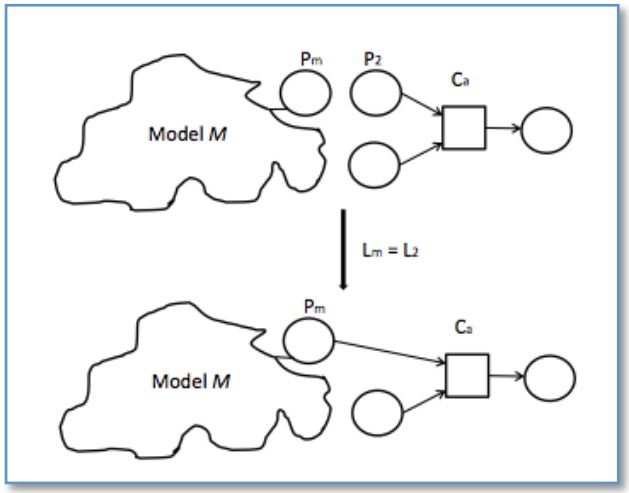




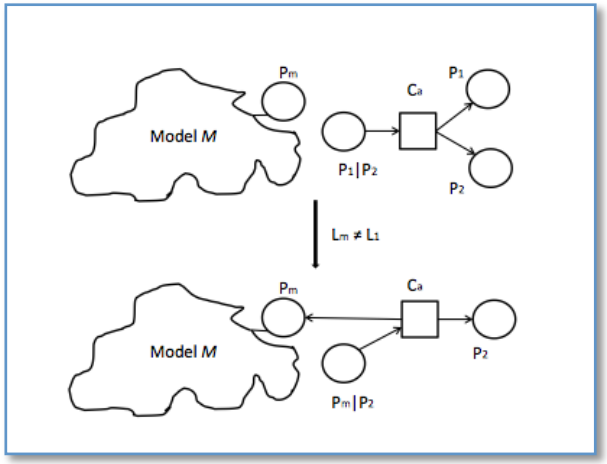
Brunel UNIVERSITY LONDON

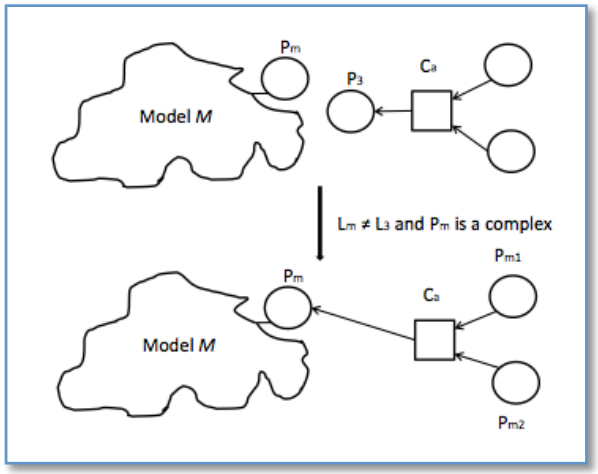
Addition

Component Ca is added to model M without modification

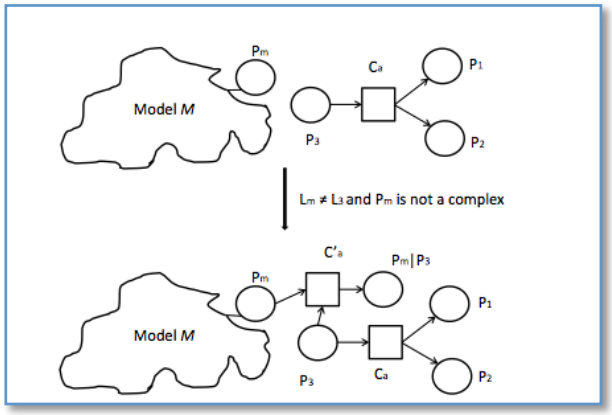


Component Ca is modified by replacing labels and added to model M





Replacement of labels in Ca by the species in M

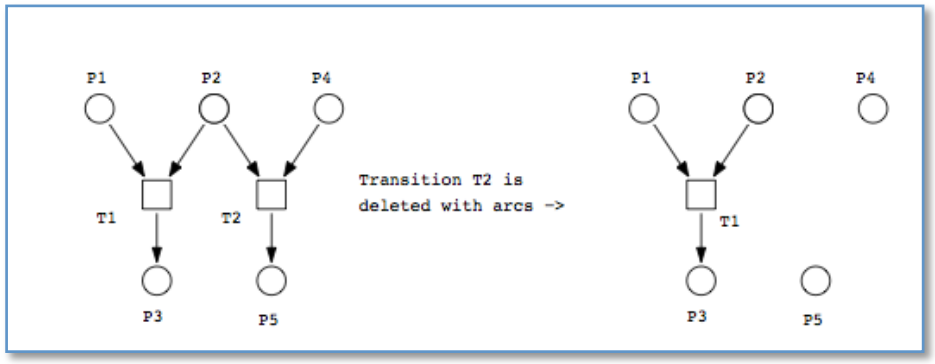


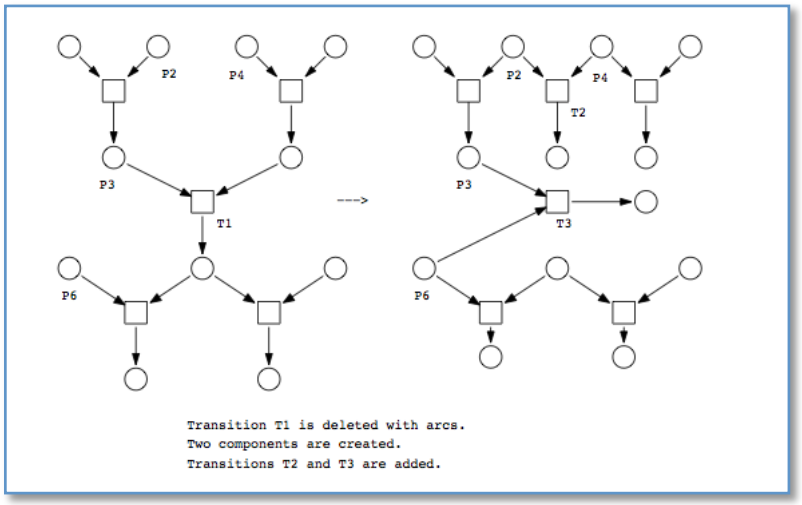
Creation of a new component Ca' by P3 and Pm

Brunel UNIVERSITY LONDON

Subtraction

Component Cm is removed directly from the model M

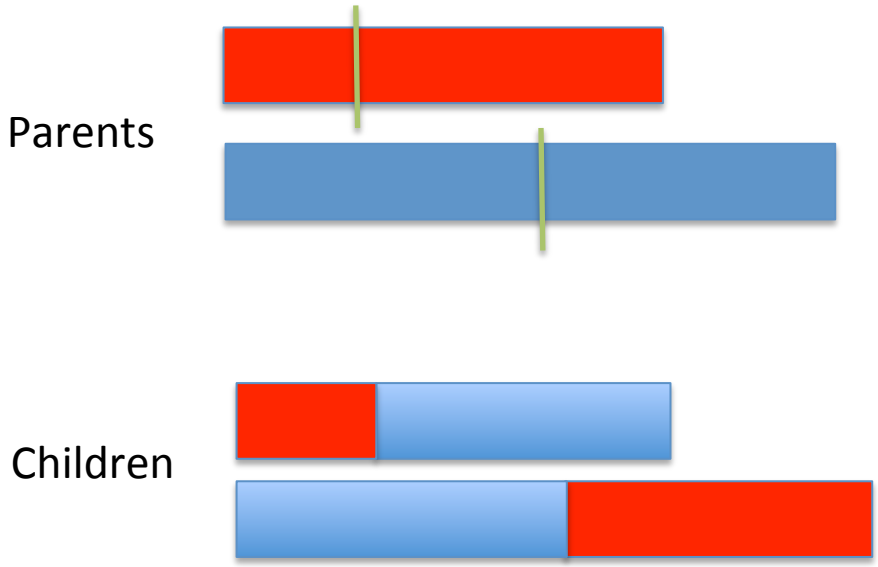




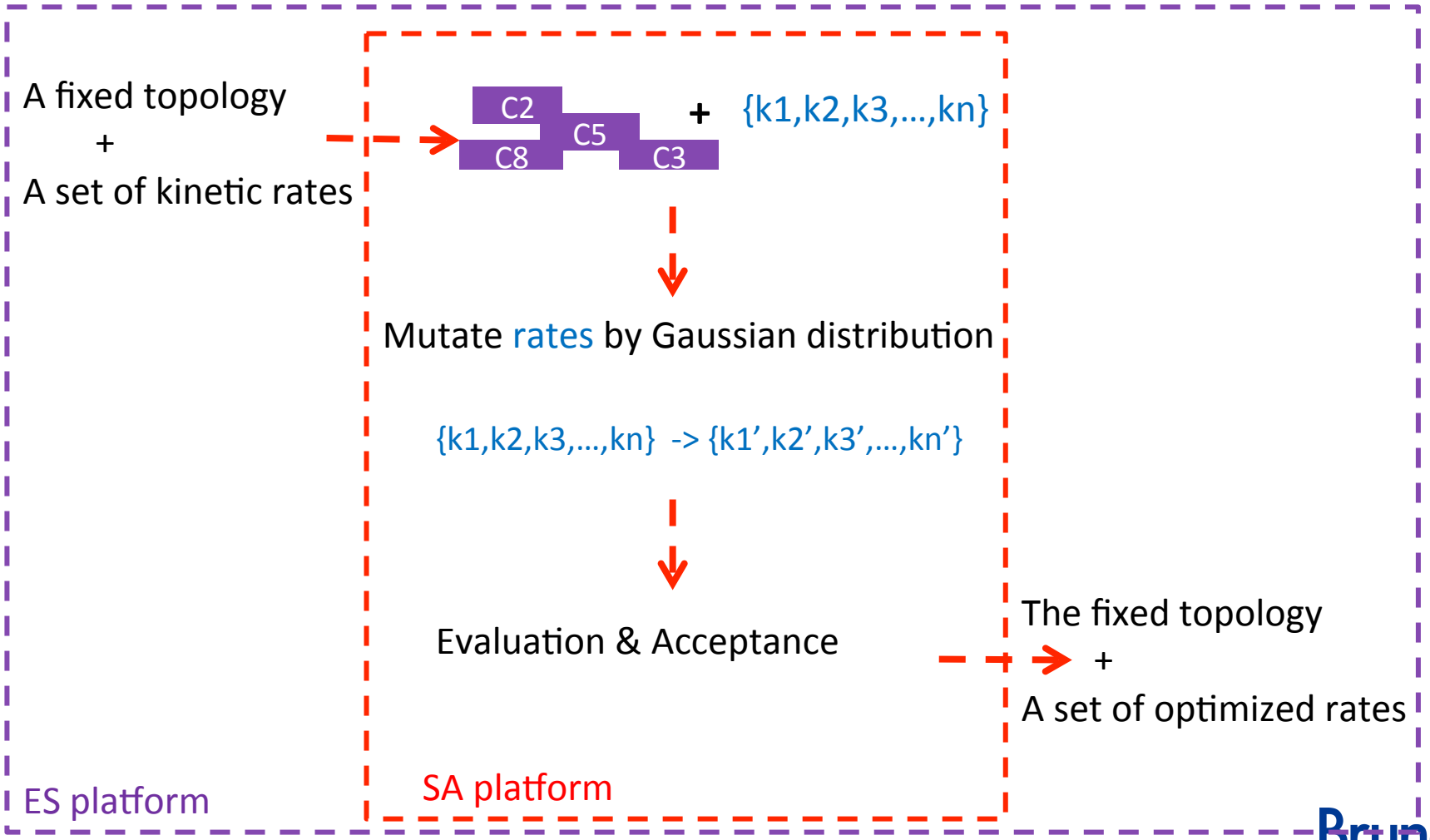
Removal of Cm and linkages of isolated components

Crossover

Cut and splice:



Simulated Annealing – Kinetic Rates Optimization



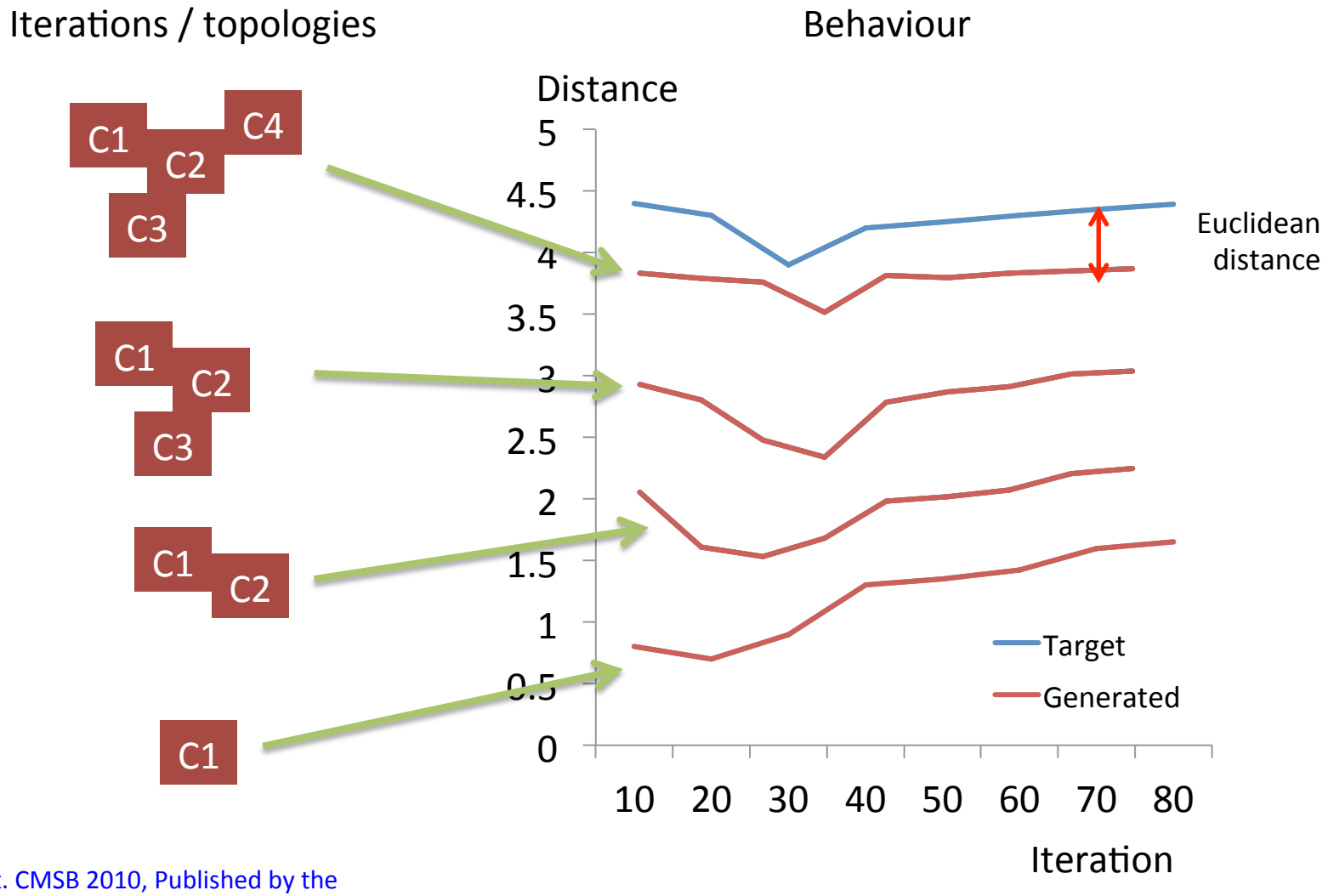
Algorithm *KROSA*

Kinetic rates optimization based on simulated annealing

Require: $T_0, T_{min}, CoolingRate, M_0, N, M$ and $K^N_{t=T_0}(M)$
Ensure: $K(M)$

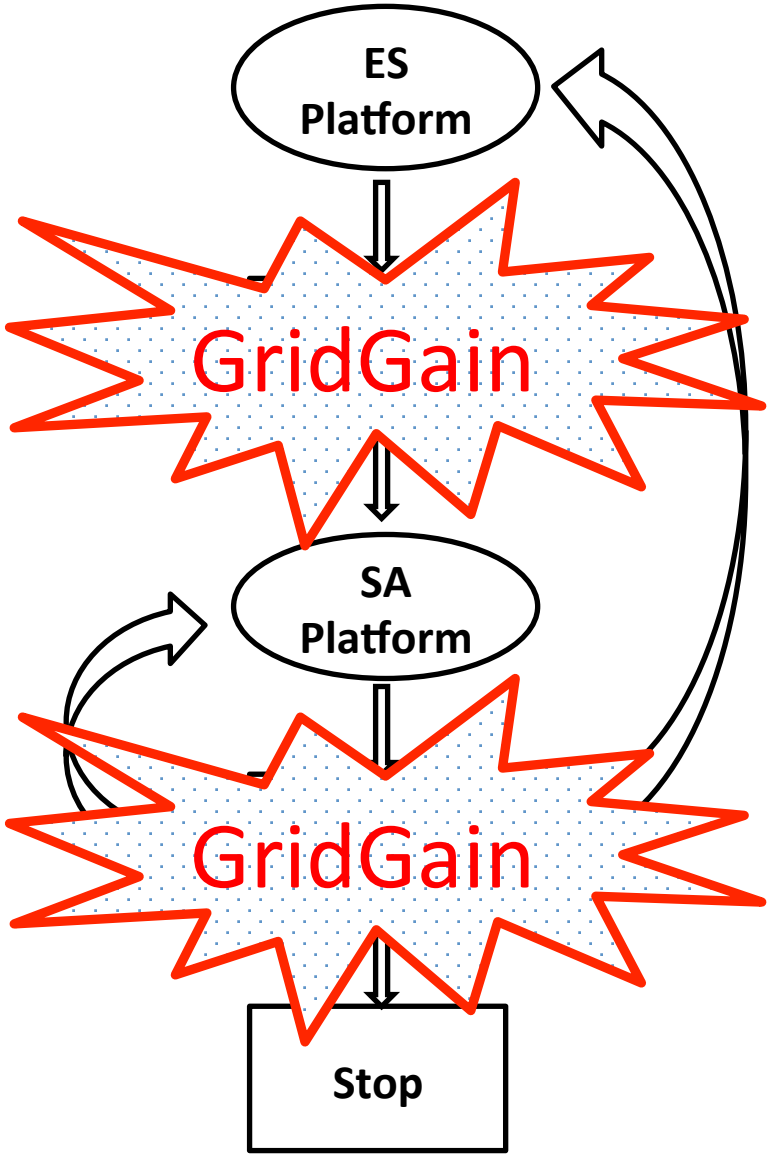
while $T_0 > T_{min}$ **do**
  while $N \neq 0$ **do**
    $K^N_{t=T_0}(M)' \leftarrow$ Modify($K^N_{t=T_0}(M), N(\mu, \sigma)$);
    Δ C = Cost($K^N_{t=T_0}(M)'$)-Cost($K^N_{t=T_0}(M)$);
    if $\Delta C < 0$ **then**
      $K^N_{t=T_0}(M) \leftarrow K^N_{t=T_0}(M)'$;
    else
      if $exp(-(\Delta C/T)) > Random(0, 1)$ **then**
        $K^N_{t=T_0}(M) \leftarrow K^N_{t=T_0}(M)'$;
      end if
    end if
    $N \leftarrow (N - 1)$;
  end while
  Reset N
  $T_0 \leftarrow (CoolingRate \times T_0)$
end while

Approaching the target behaviour



Wu, Gao & Gilbert. CMSB 2010, Published by the ACM Digital Library

Brunel UNIVERSITY LONDON

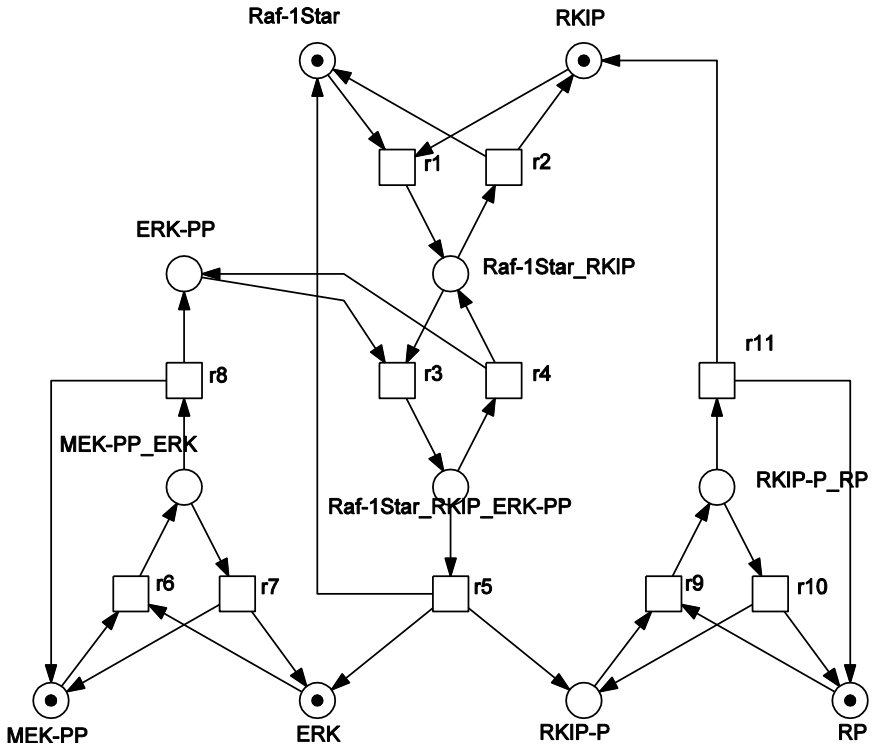


Case Study 1 : RKIP Pathway

Reference:

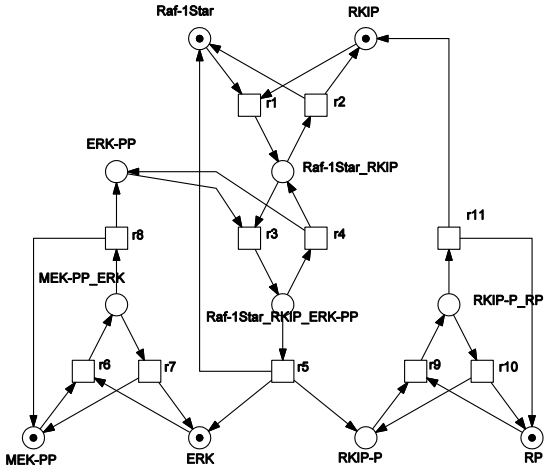
[1] Cho et al. (2003) Mathematical modeling of the influence of RKIP on the ERK signaling pathway.

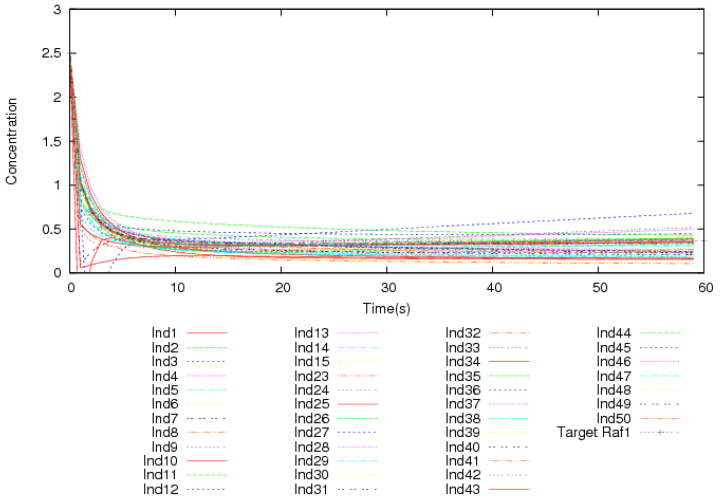
[2] Calder et al. (2004) Modelling the influence of RKIP on the ERK signalling pathway using the stochastic process algebra PEPA

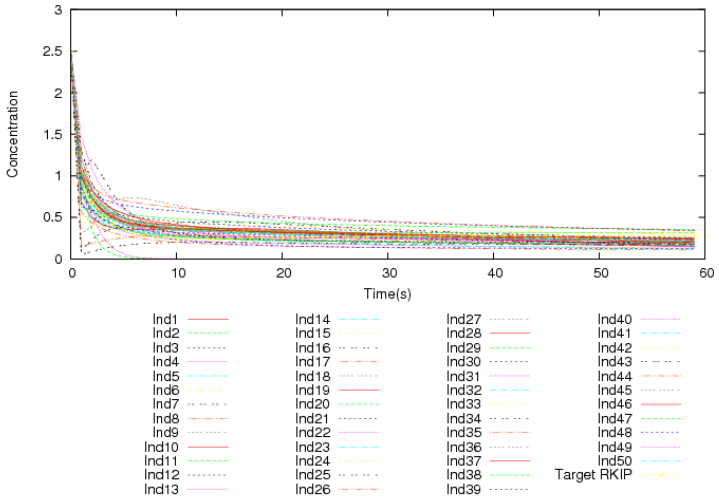


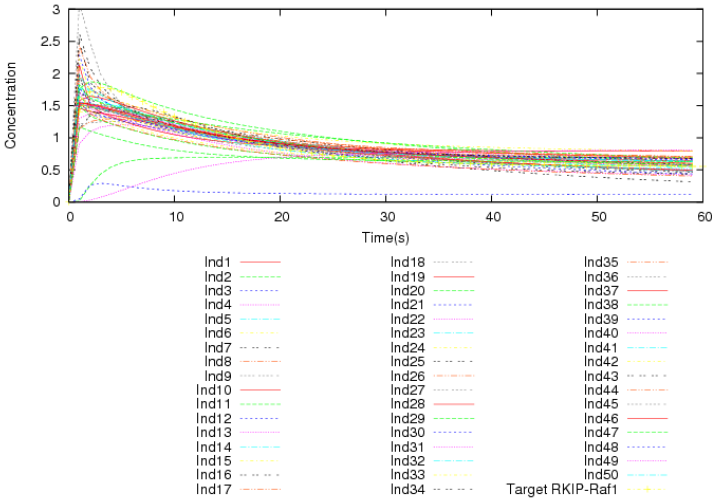
Brunel UNIVERSITY LONDON

Populations of solutions

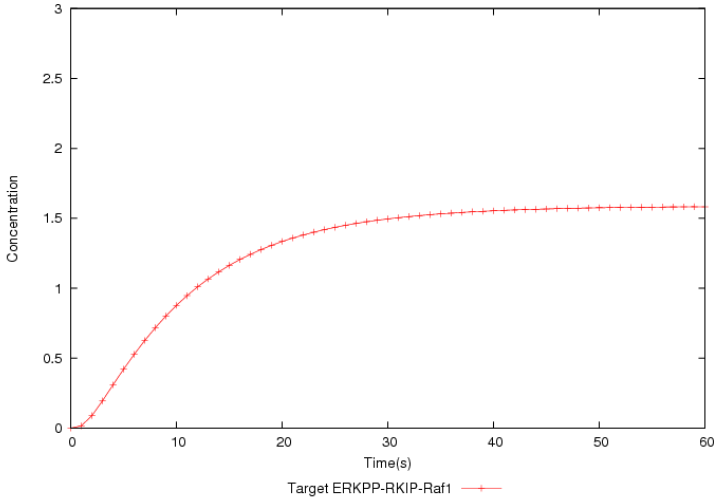




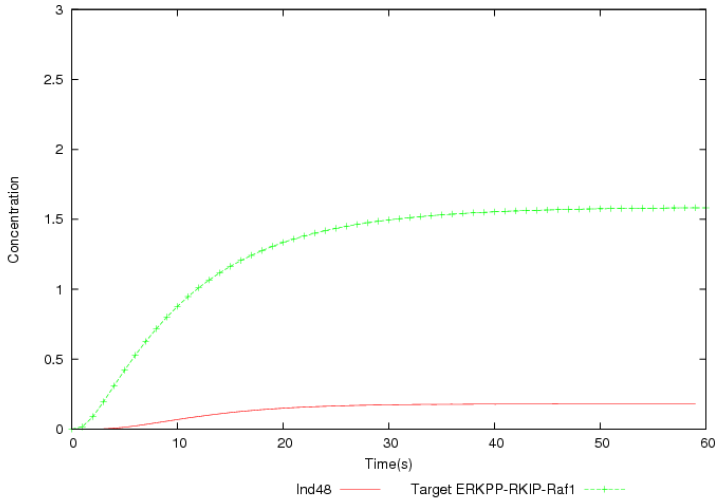




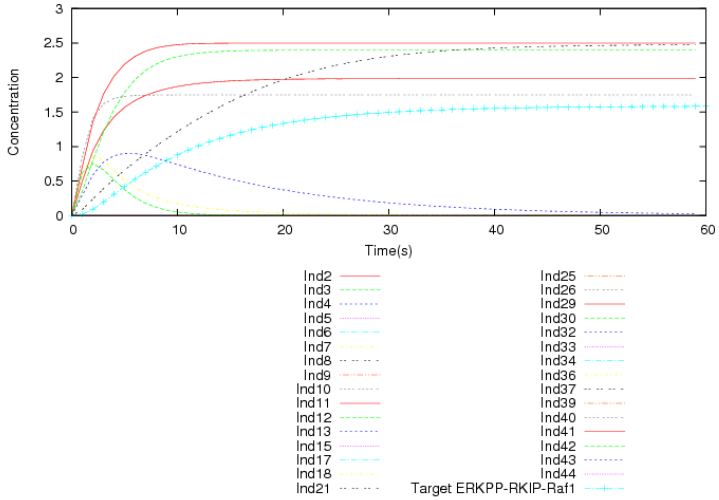
Brunel UNIVERSITY LONDON



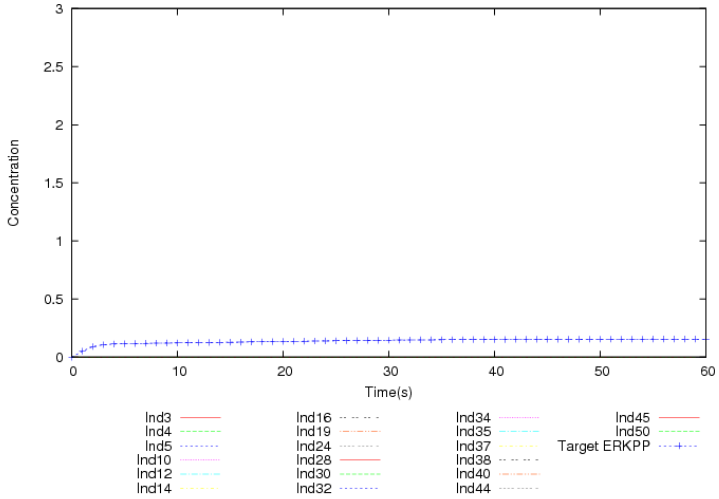
v9es_120216k_run1
(1000, 50, SES, ED+RP)

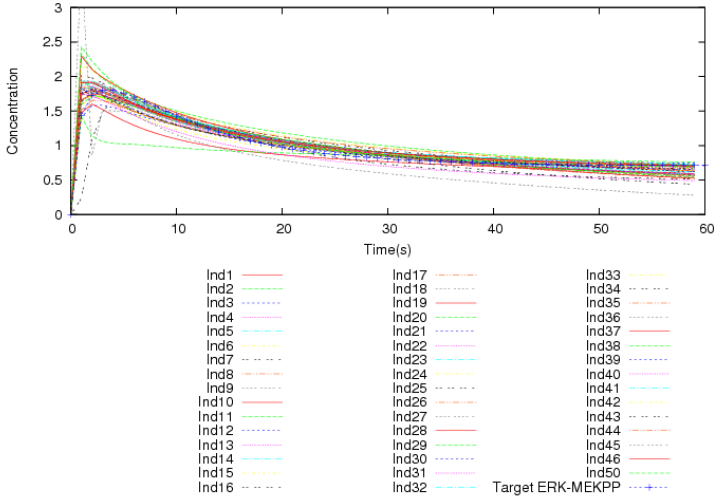


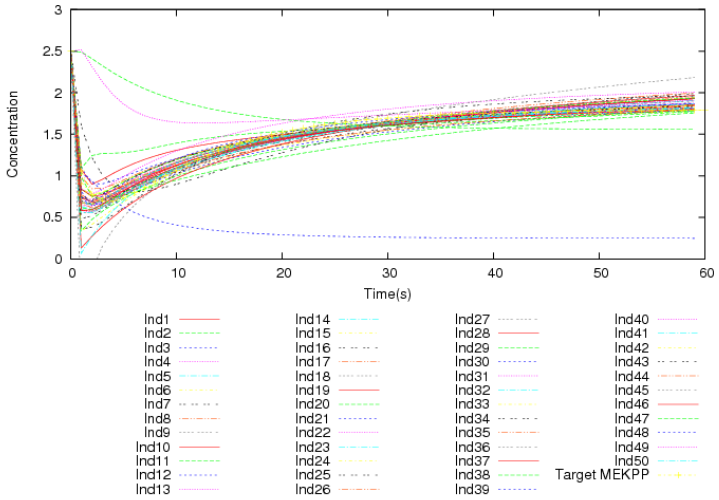
v9es_120216k_run3
(1000, 50, PES, ED)

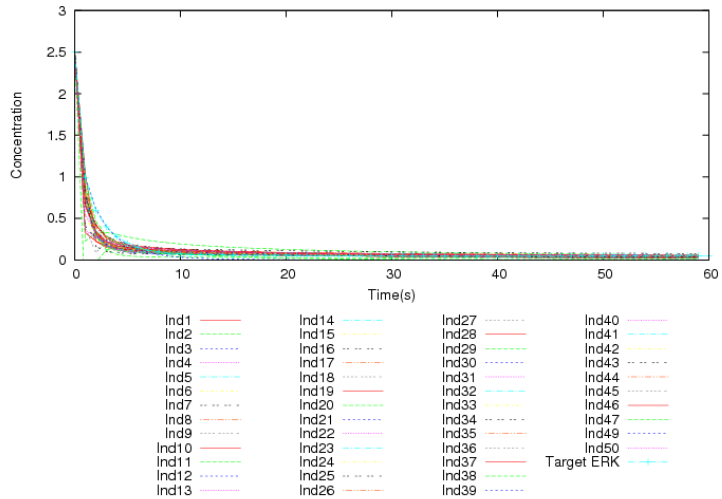


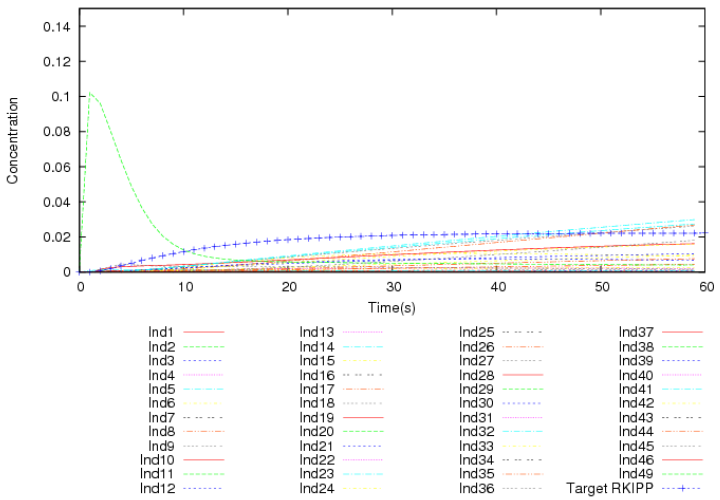
v9es_120209b_run1
(100, 45, SES, ED)

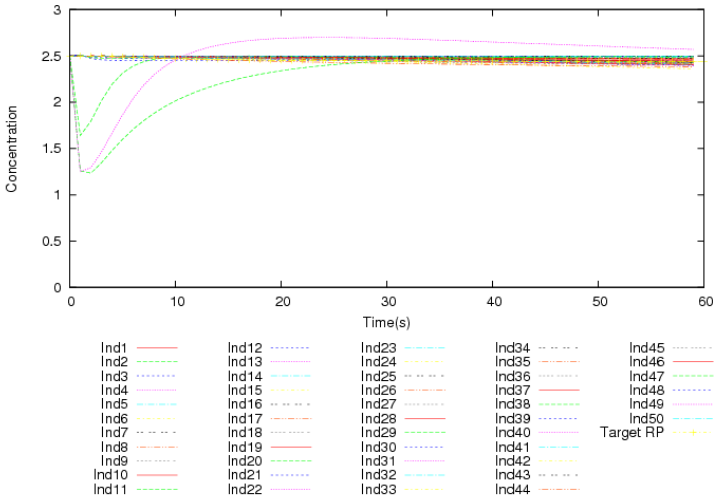

Concentration
Time(s)
Ind3
Ind4
Ind5
Ind10
Ind12
Ind14
Ind16
Ind19
Ind24
Ind28
Ind30
Ind32
Ind34
Ind35
Ind37
Ind38
Ind40
Ind44
Ind45
Ind50
Target ERKPP

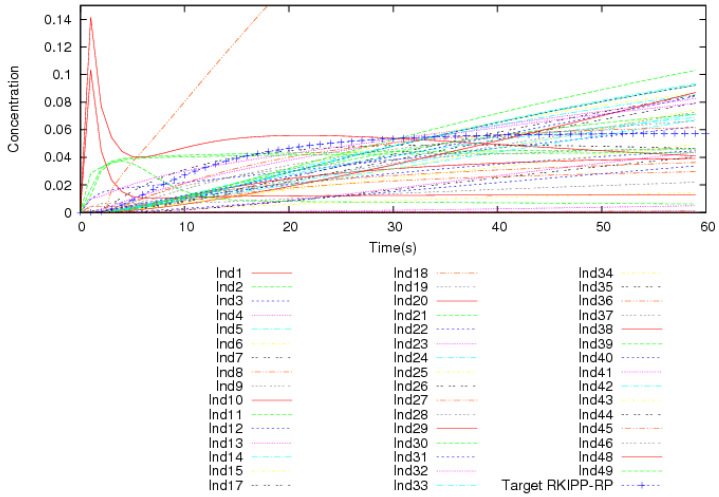

Concentration
Time(s)
Ind1
Ind2
Ind3
Ind4
Ind5
Ind6
Ind7
Ind8
Ind9
Ind10
Ind11
Ind12
Ind13
Ind14
Ind15
Ind16
Ind17
Ind18
Ind19
Ind20
Ind21
Ind22
Ind23
Ind24
Ind25
Ind26
Ind27
Ind28
Ind29
Ind30
Ind31
Ind32
Ind33
Ind34
Ind35
Ind36
Ind37
Ind38
Ind39
Ind40
Ind41
Ind42
Ind43
Ind44
Ind45
Ind46
Ind50
Target ERK-MEKPP

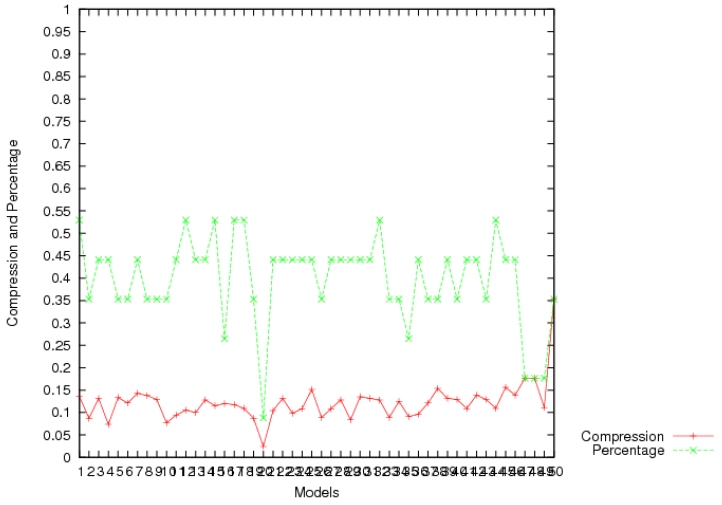

Concentration
Time(s)
Ind1
Ind2
Ind3
Ind4
Ind5
Ind6
Ind7
Ind8
Ind9
Ind10
Ind11
Ind12
Ind13
Ind14
Ind15
Ind16
Ind17
Ind18
Ind19
Ind20
Ind21
Ind22
Ind23
Ind24
Ind25
Ind26
Ind27
Ind28
Ind29
Ind30
Ind31
Ind32
Ind33
Ind34
Ind35
Ind36
Ind37
Ind38
Ind39
Ind40
Ind41
Ind42
Ind43
Ind44
Ind45
Ind46
Ind47
Ind48
Ind49
Ind50
Target MEKPP


Concentration
Time(s)
Ind1
Ind2
Ind3
Ind4
Ind5
Ind6
Ind7
Ind8
Ind9
Ind10
Ind11
Ind12
Ind13
Ind14
Ind15
Ind16
Ind17
Ind18
Ind19
Ind20
Ind21
Ind22
Ind23
Ind24
Ind25
Ind26
Ind27
Ind28
Ind29
Ind30
Ind31
Ind32
Ind33
Ind34
Ind35
Ind36
Ind37
Ind38
Ind39
Ind40
Ind41
Ind42
Ind43
Ind44
Ind45
Ind46
Ind47
Ind49
Ind50
Target ERK

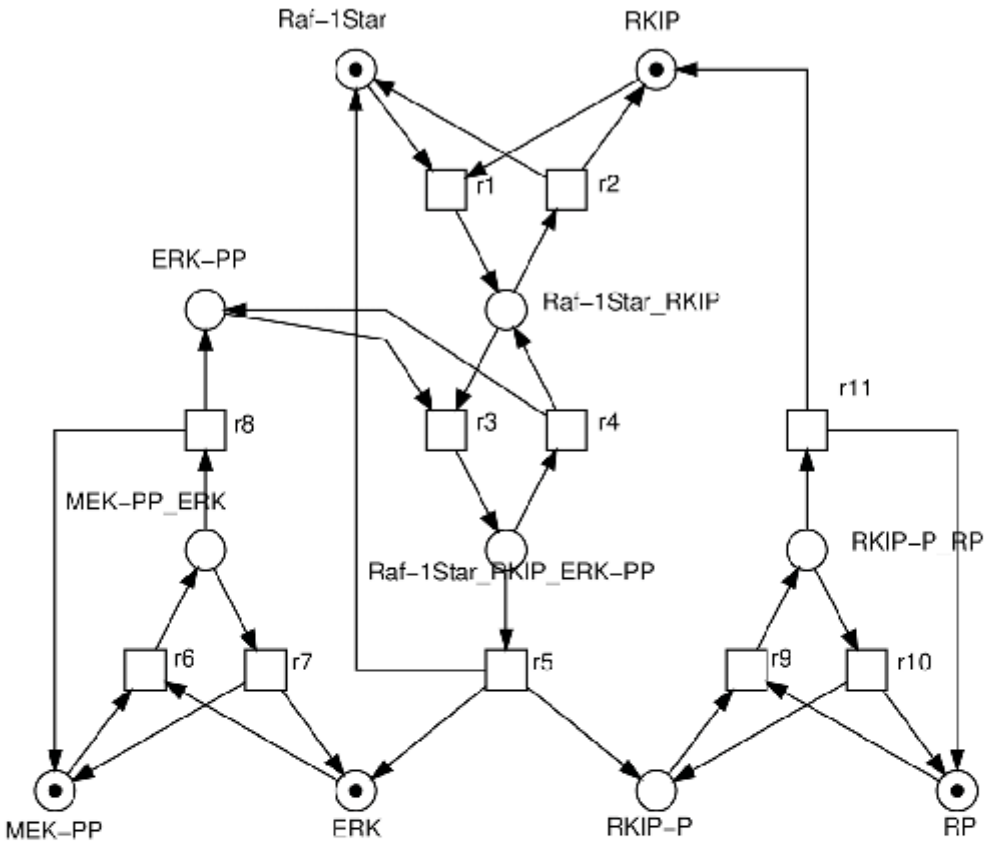

Concentration
Time(s)
Ind1
Ind2
Ind3
Ind4
Ind5
Ind6
Ind7
Ind8
Ind9
Ind10
Ind11
Ind12
Ind13
Ind14
Ind15
Ind16
Ind17
Ind18
Ind19
Ind20
Ind21
Ind22
Ind23
Ind24
Ind25
Ind26
Ind27
Ind28
Ind29
Ind30
Ind31
Ind32
Ind33
Ind34
Ind35
Ind36
Ind37
Ind38
Ind39
Ind40
Ind41
Ind42
Ind43
Ind44
Ind45
Ind46
Ind49
Target RKIPP


Concentration
Time(s)
Ind1
Ind2
Ind3
Ind4
Ind5
Ind6
Ind7
Ind8
Ind9
Ind10
Ind11
Ind12
Ind13
Ind14
Ind15
Ind16
Ind17
Ind18
Ind19
Ind20
Ind21
Ind22
Ind23
Ind24
Ind25
Ind26
Ind27
Ind28
Ind29
Ind30
Ind31
Ind32
Ind33
Ind34
Ind35
Ind36
Ind37
Ind38
Ind39
Ind40
Ind41
Ind42
Ind43
Ind44
Ind45
Ind46
Ind47
Ind48
Ind49
Ind50
Target RP


Concentration
Time(s)
Ind1
Ind2
Ind3
Ind4
Ind5
Ind6
Ind7
Ind8
Ind9
Ind10
Ind11
Ind12
Ind13
Ind14
Ind15
Ind17
Ind18
Ind19
Ind20
Ind21
Ind22
Ind23
Ind24
Ind25
Ind26
Ind27
Ind28
Ind29
Ind30
Ind31
Ind32
Ind33
Ind34
Ind35
Ind36
Ind37
Ind38
Ind39
Ind40
Ind41
Ind42
Ind43
Ind44
Ind45
Ind46
Ind48
Ind49
Target RKIPP-RP


Compression and Percentage
Models
Compression
Percentage

Target RKIP Pathway



Model Information:

Continuous places: 11
Continuous transitions: 11
Parameters(Reactions): 11
Edges(Arcs): 34

$$Compression = \frac{|Intersection|}{Max(|Target|, |Generated|)}$$

|Intersection| is the number of common (matched) arcs between generated and target models;

|Target| is the number of arcs in target pathway;

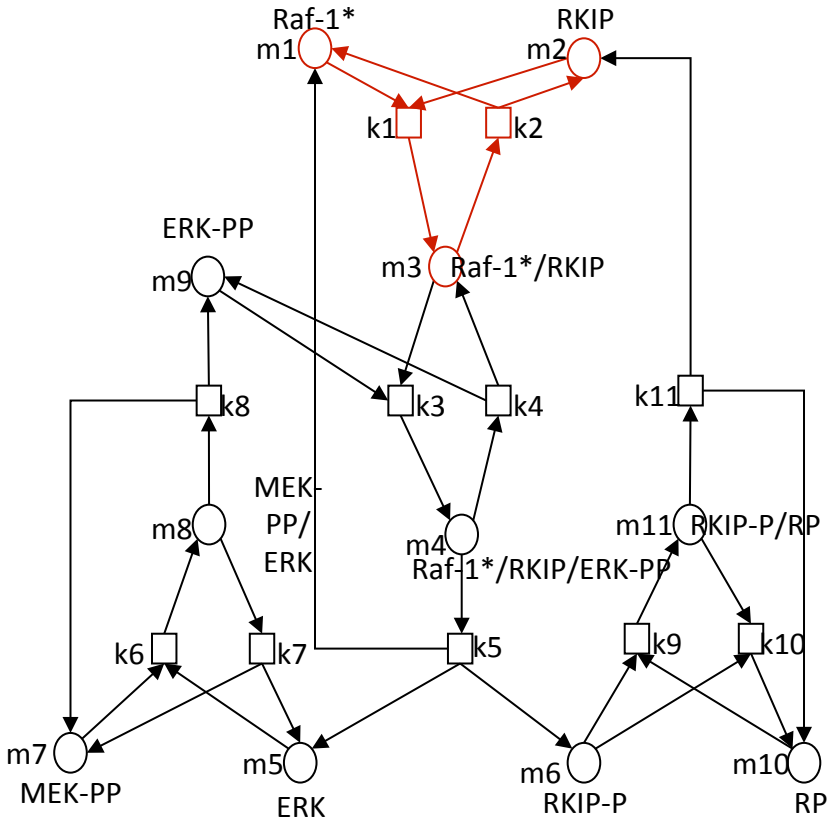
|Generated| is the number of arcs in generated model.

Goodness of generated topology

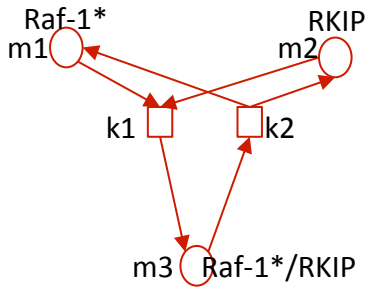
‘Small(er) is Beautiful’?

- Compression: matched ‘arcs’ in target and generated (normalised 0..1)

Target



Generated



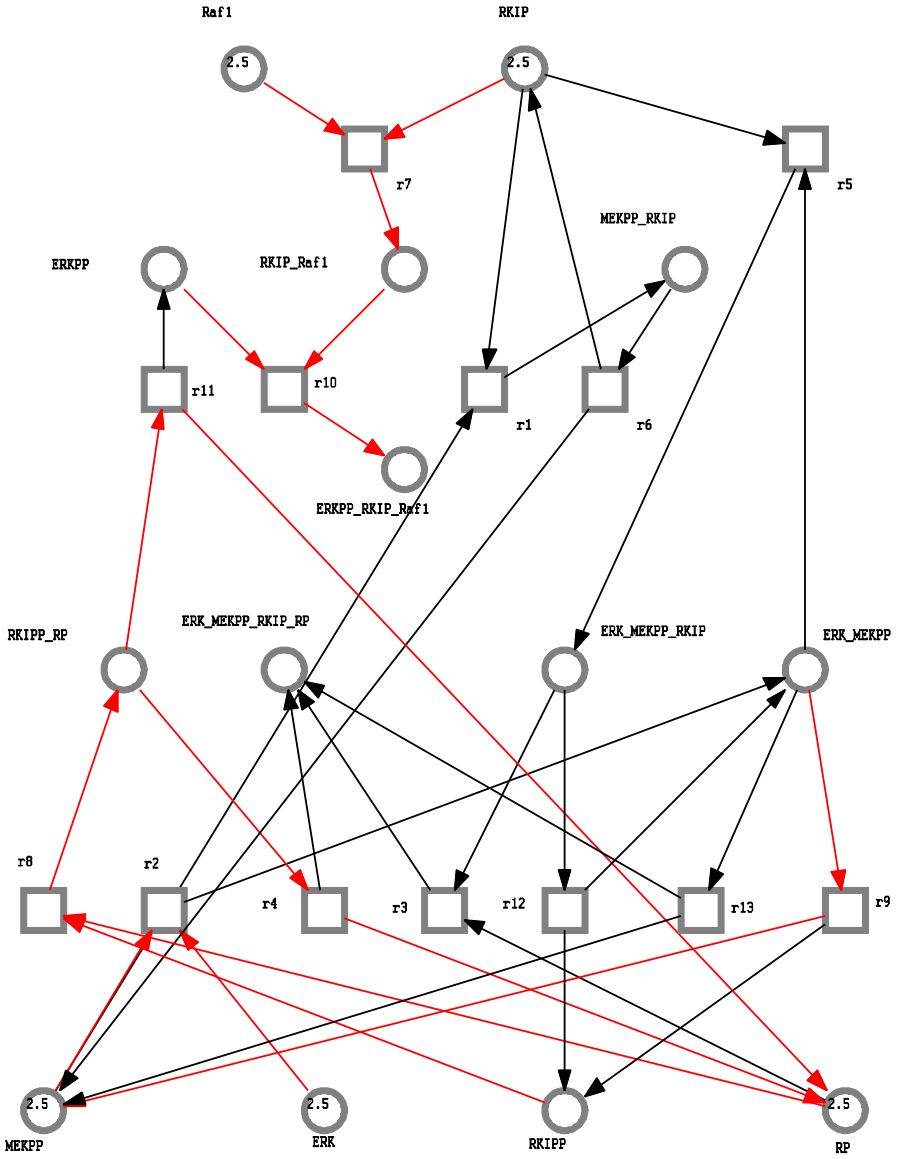
Matched arcs: ——

$$\text{Compression}_2 = \frac{|\text{Target} \cap \text{Generated}|}{Max(|\text{Target}|, |\text{Generated}|)}$$

$$= \frac{6}{Max(34,6)} = 0.176$$

Brunel UNIVERSITY LONDON

One Generated Model



Model Information:

Continuous places: 14
Continuous transitions: 13
Parameters(Reactions): 14
Edges(Arcs): 39
Matched arcs(in RED colour): 17

Compression = 17/Max(34, 39)
= 17/39
= 0.44

Brunel UNIVERSITY LONDON

Comparison between Generated and Target Models

Entities	$Topology_{Target}$	$Topology_{9th}$
Raf1	√	√
RKIP	√	√
Raf1|RKIP	√	√
Raf1|RKIP|ERKPP	√	
ERK	√	√
RKIPP	√	√
MEKPP	√	√
MEKPP|ERK	√	√
ERKPP	√	√
RP	√	√
RKIPP|RP	√	
ERKPP|Raf1		√
RKIPP|Raf1		√
ERK|Raf1		√
ERKPP|MEKPP		√
RKIPP|MEKPP		√
RKIPP|Raf1		√
ERK|RP		√
RKIP|RP		√
ERKPP|RP		√

Entities in target RKIP pathway

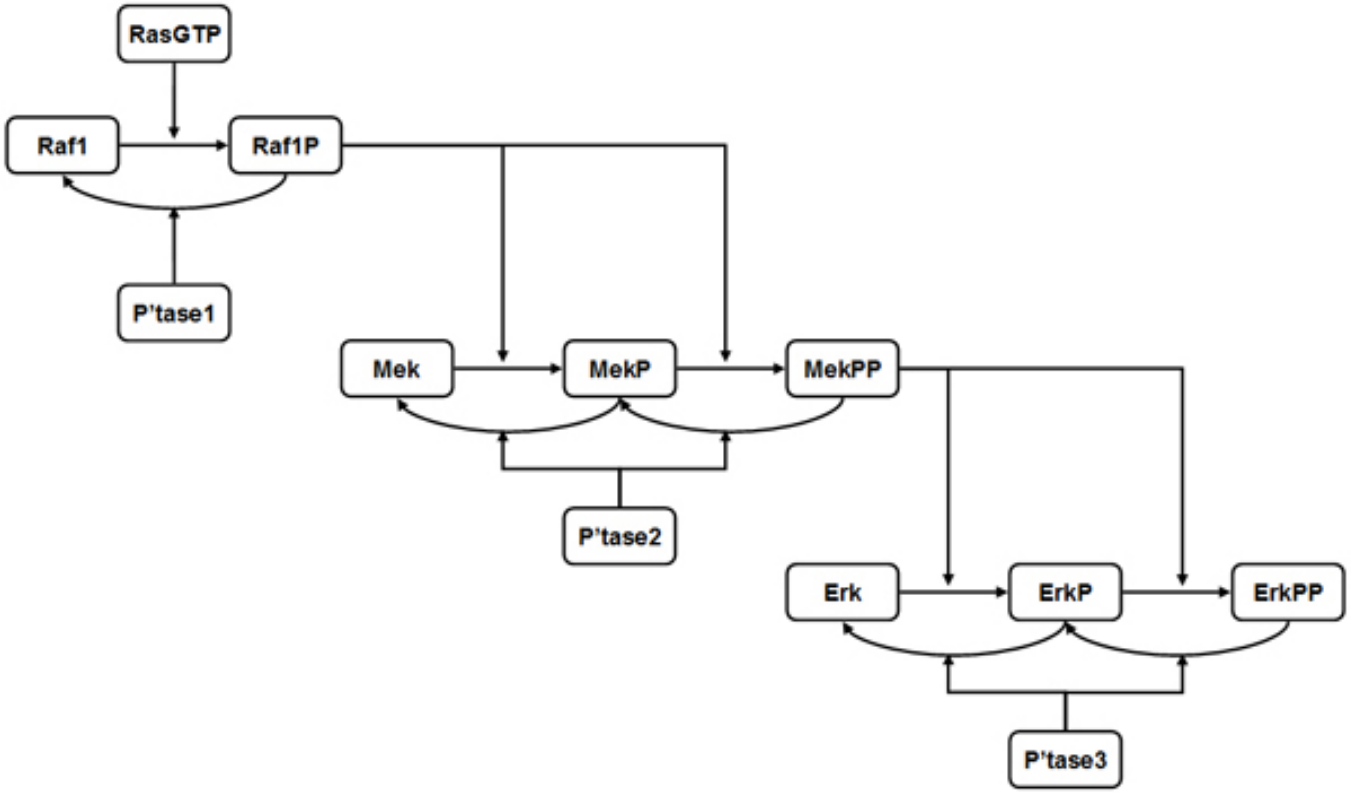
Entities in generated RKIP pathway

Wu, Gao & Gilbert. CMSB 2010, Published by the ACM Digital Library

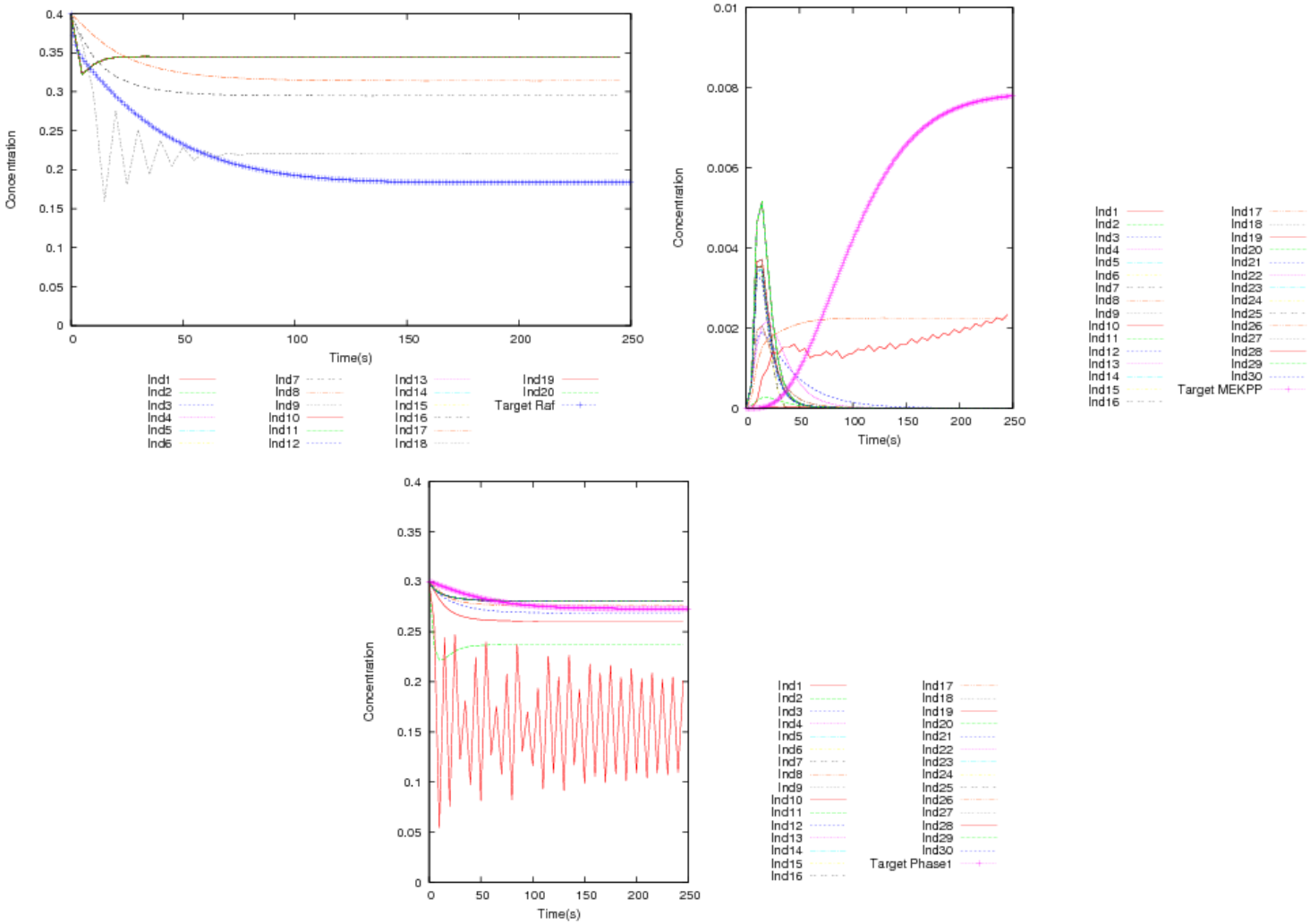
Case Study 2 : Levchenko2000

Reference:

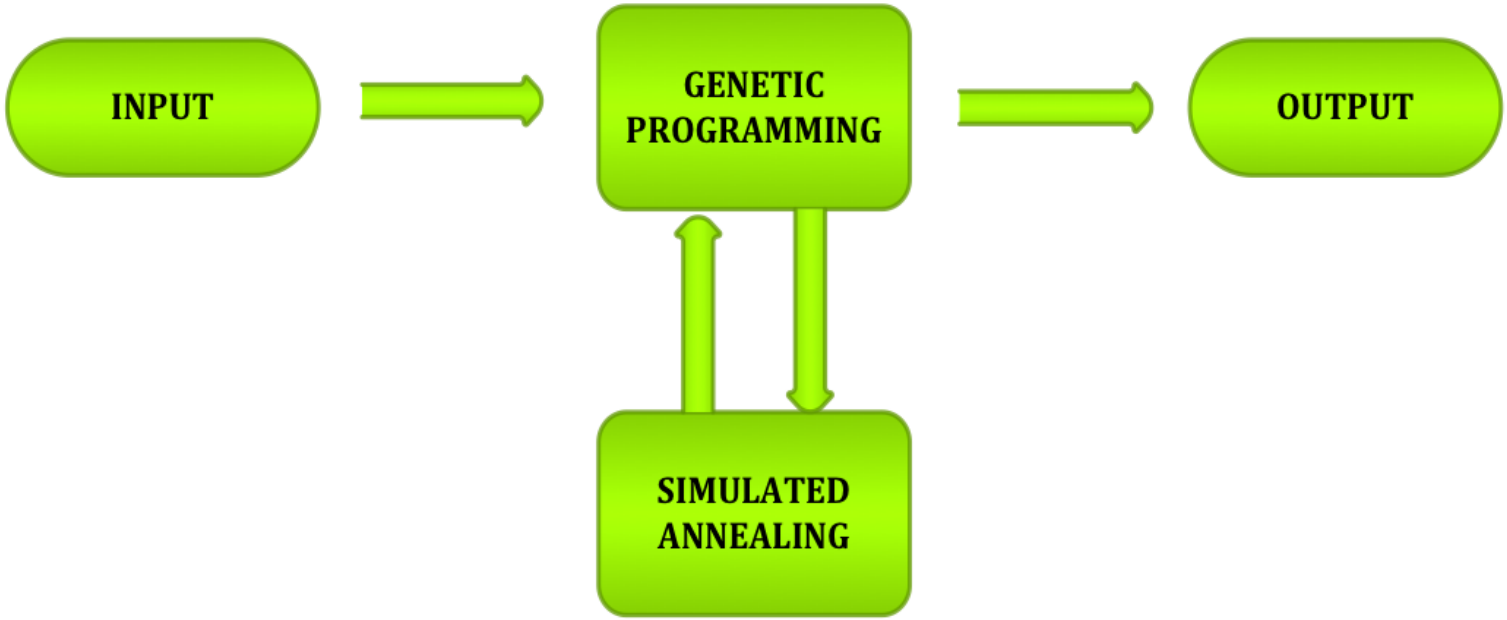
[1] Andre Levchenko et al. (2000) Scaffold proteins may biphasically affect the levels of mitogen-activated protein kinase signaling and reduce its threshold properties



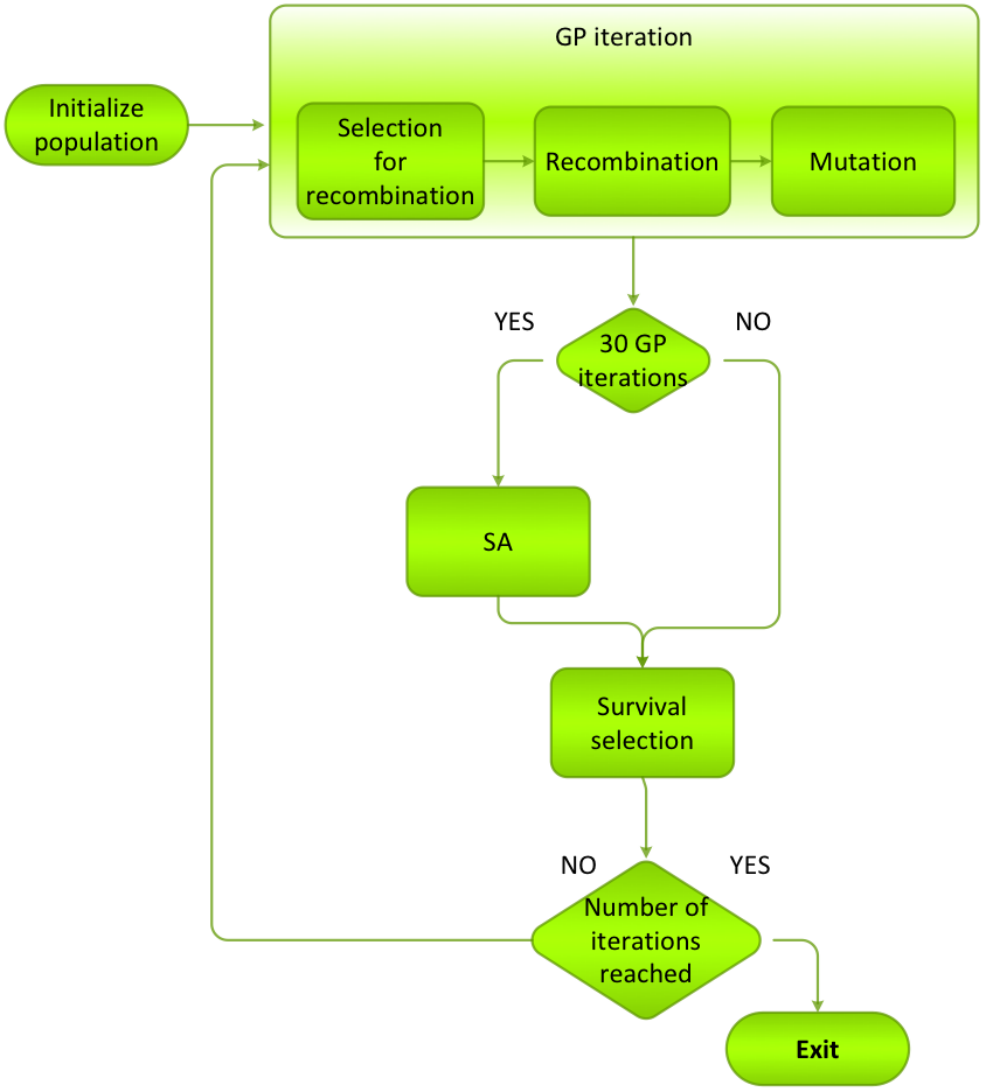
Brunel UNIVERSITY LONDON


Concentration
Time(s)
Ind1
Ind2
Ind3
Ind4
Ind5
Ind6
Ind7
Ind8
Ind9
Ind10
Ind11
Ind12
Ind13
Ind14
Ind15
Ind16
Ind17
Ind18
Ind19
Ind20
Target Raf
Target MEKPP
Target Phase1

Genetic programming – Simulated annealing



GP-SA integration



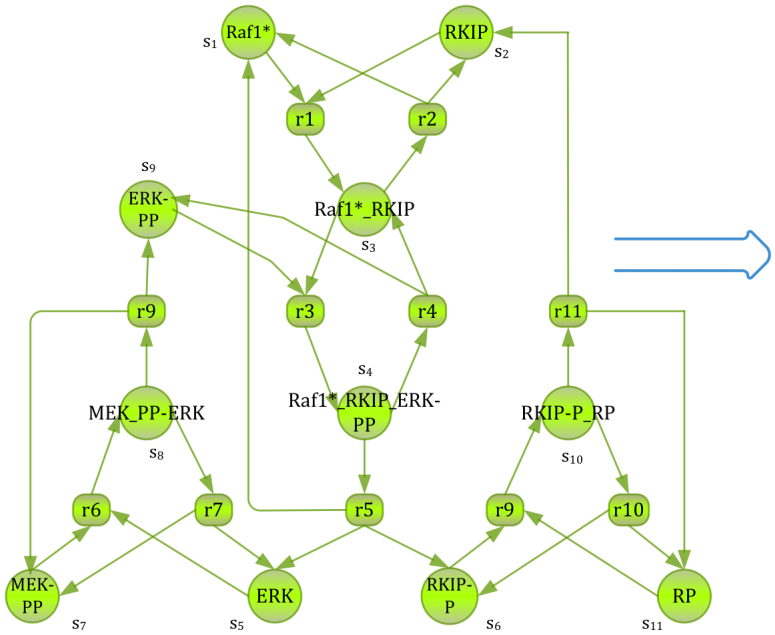
Brunel UNIVERSITY LONDON

GP elements

- Solution representation
- Fitness function
- Genetic operators (mutation and crossover)

Solution representation

- a network of reactions together with their kinetic rates.



Input species set	Output species set	Kinetic rates
{Raf1*, RKIP}	{Raf1*_RKIP}	r1 kinetic rate
{Raf1*_RKIP}	{Raf1*, RKIP}	r2 kinetic rate
{ERK-PP, Raf1*_RKIP}	{Raf1*_RKIP_ERK-PP}	r3 kinetic rate
{Raf1*_RKIP_ERK-PP}	{ERK-PP, Raf1*_RKIP}	r4 kinetic rate
{Raf1*_RKIP_ERK-PP}	{Raf1*, RKIP-PP, ERK}	r5 kinetic rate
{MEK-PP, ERK}	{MEK-PP_ERK}	r6 kinetic rate
{MEK-PP_ERK}	{MEK-PP, ERK}	r7 kinetic rate
{MEK-PP_ERK}	{MEK-PP, ERK-PP}	r8 kinetic rate
{RP, RKIP-P}	{RKIP-P_RP}	r9 kinetic rate
{RKIP-P_RP}	{RP, RKIP-P}	r10 kinetic rate
{RKIP-P_RP}	{RP, RKIP}	r11 kinetic rate

Solution validity

- a chromosome is valid if it contains every complex specie (composed of several simple species) in the output of at least one reaction

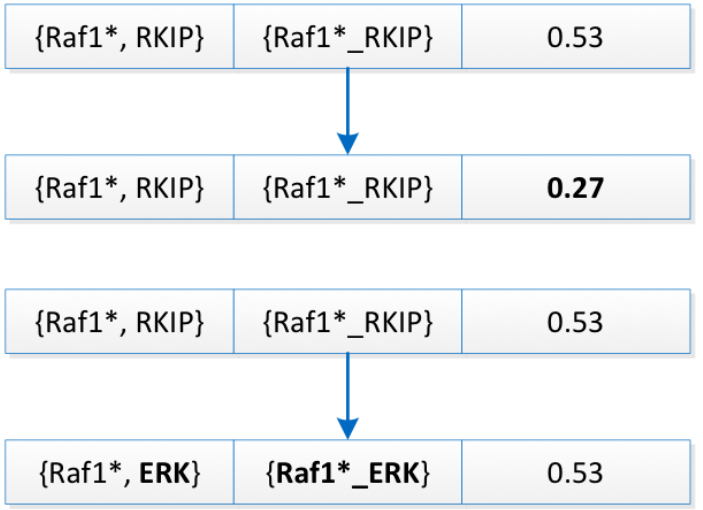
Fitness function

$$fitness = \sum_{t=0}^{m}\sum_{i=0}^{n}\left|s_i(t) - \text{target}_i(t)\right| - \sum_{t=0}^{m}\left(\sum_{i\in X} penalty\_missing_i + \sum_{j\in Y} penalty\_extra_j\right)$$

- *penalty_missing* is a constant representing penalty for missing specie from the current specie set
- *penalty_extra* is a constant representing penalty for extra specie in the current specie set

Mutation operators

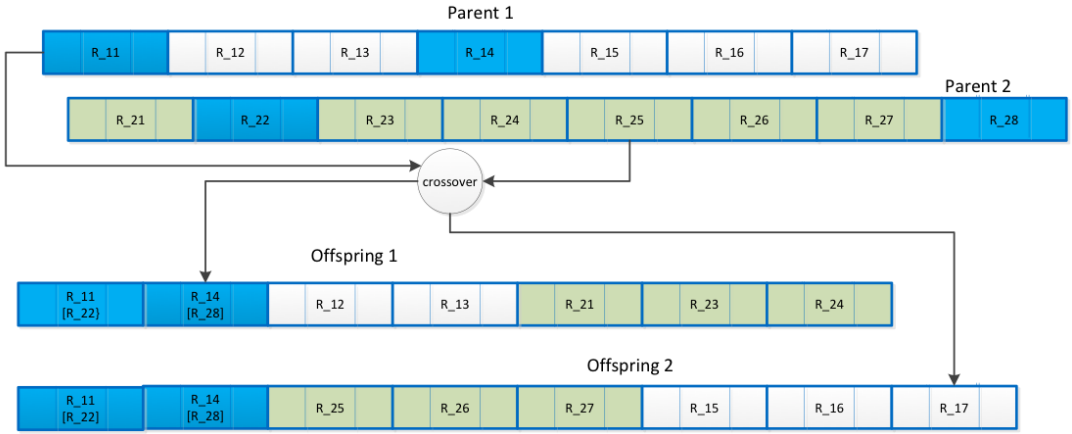
- Alteration of a kinetic rate



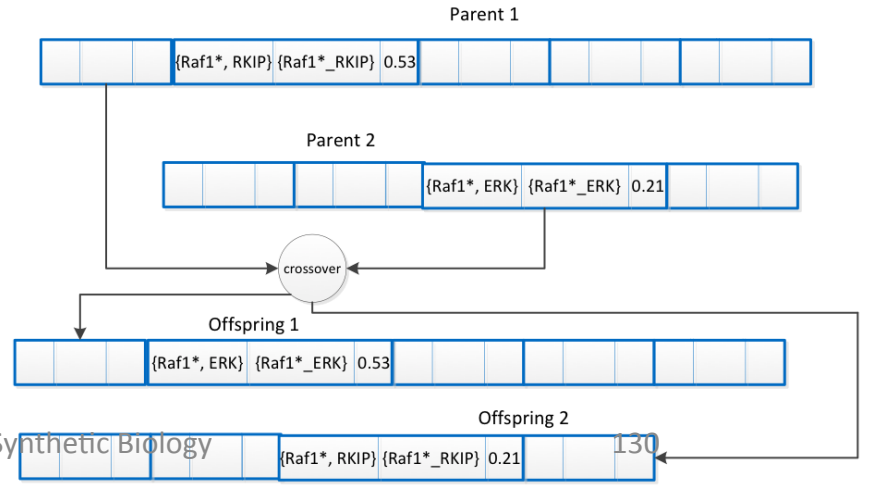
- Replacement of a species
- Insertion of a reaction
- Deletion of a reaction

Crossover operators

- Cut and splice



- Pick and replace



Now the rates…
(Parameter fitting)

1. Standard time-series distance-based

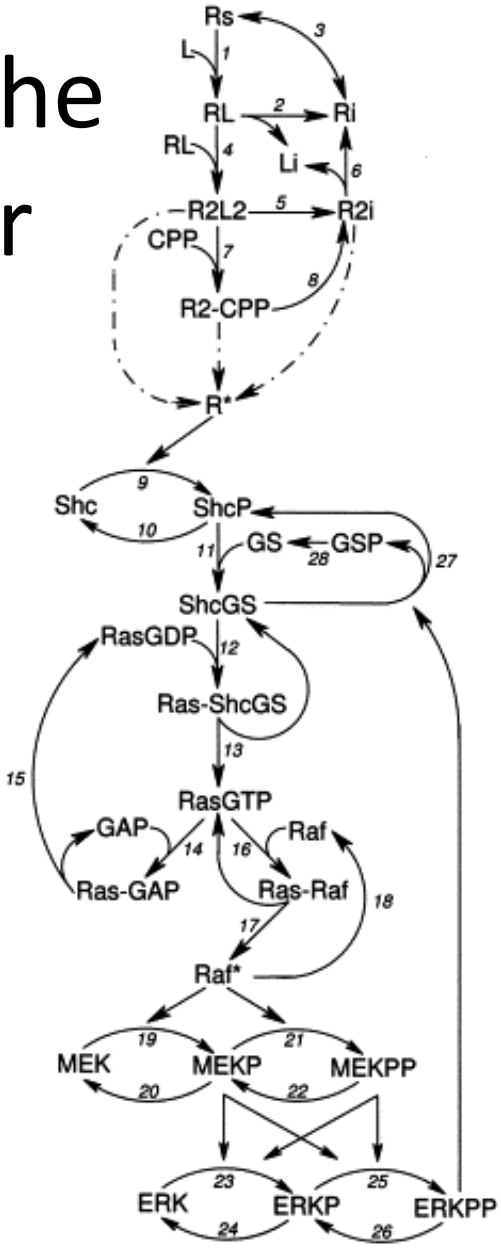
2. **Model checking: logical description of behaviours**

Parameter estimation using the Monte Carlo Model Checker

Continuous Brightman & Fell model:

- The EGF signal transduction pathway produces transient Ras, MEK and ERK activation whereas NGF stimulation produces sustained activation.

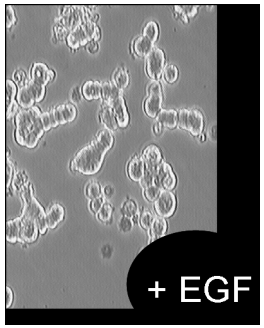
- Parameter V28 has the the highest probability of generating the desired behaviour, but requires 40-fold increase in value



Brightman & Fell, FEBS Lett 2000. "Differential feedback regulation of the MAPK cascade underlies the quantitative differences in EGF and NGF signalling in PC12 cells"

Brunel UNIVERSITY LONDON

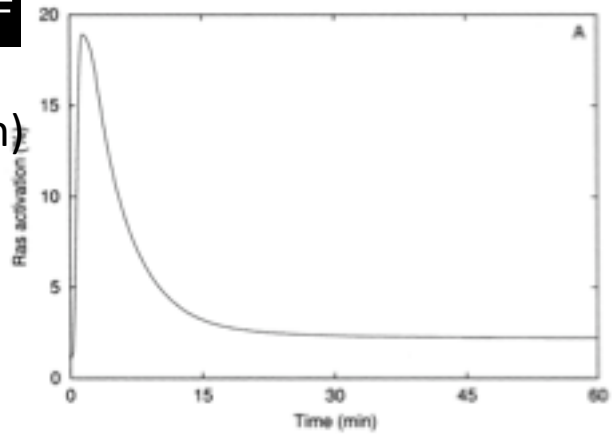
Response with EGF vs. NGF signal

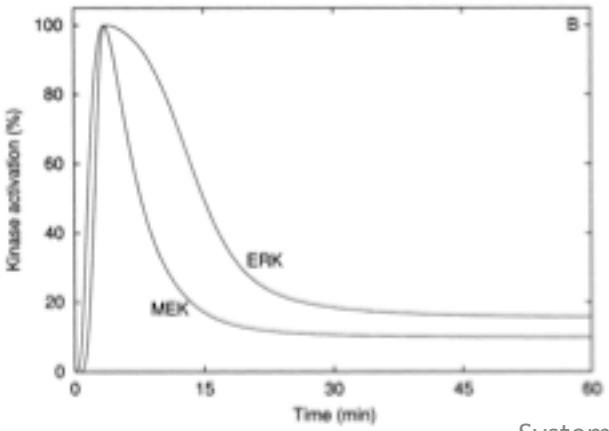


Proliferation
(cell division)
PC12 cells

EGF

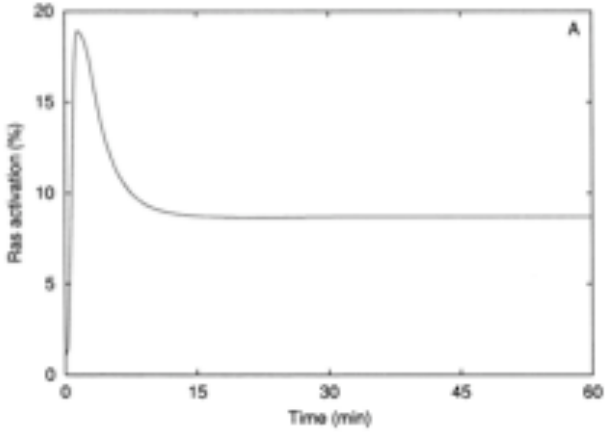
Transient activation of Ras, MEK and ERK

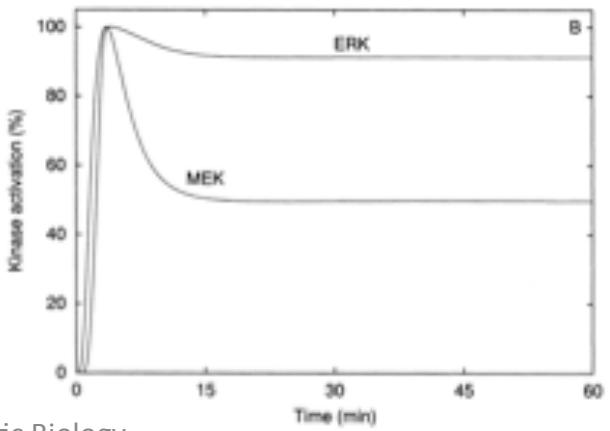


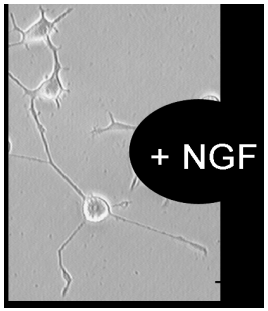


NGF

Sustained activation of Ras, MEK and ERK







Differentiation
(neurite
outgrowth)

Desired Behaviour in PLTLc

The desired (sustained) NGF behaviour of the pathway was written informally in the original paper.

We can formalise it in PLTLc as:

Sustained Ras: Active Ras peaks within 2 min to max 20% total Ras and stable between 5%..10%

$$P_{=?}\ [\ d(\text{active Ras}) > 0\ \ U\ (\ \text{time} \le 2\ \wedge\ \text{active Ras} \ge 0.15{*}\text{total Ras}$$
$$\wedge\ \text{active Ras} \le 0.2{*}\text{total Ras}\ \wedge\ (\ d(\text{active Ras}) < 0)$$
$$U\ (\ G(\ \text{active Ras} \ge 0.05{*}\text{total Ras}\ \wedge\ \text{active Ras} \le 0.10{*}\text{total Ras}\ )\ )\ )\ ]$$

Sustained MEK: Active MEK peaks in 2 to 5 min and is stable between 40%..50% of peak value

$$P_{=?}\ [\ d(\text{MEKPP}) > 0\ U\ (\ \text{time} \ge 2\ \wedge\ \text{time} \le 5\ \wedge\ d(\text{MEKPP}) < 0$$
$$U\ (\ G(\ \text{MEKPP} \ge 0.40{*}\max(\text{MEKPP})\ \wedge\ \text{MEKPP} \le 0.50{*}\max(\text{MEKPP})\ \ )\ )\ )\ ]$$

Sustained ERK: Active ERK peaks in 2 to 5 min and is stable between 85%..100% of peak value

$$P_{=?}\ [\ (\ d(\text{ERKPP}) > 0\ )\ U\ (\ \text{time} \ge 2\ \wedge\ \text{time} \le 5\ \wedge\ d(\text{ERKPP}) < 0$$
$$U\ (\ G(\ \text{ERKPP} \ge 0.85 * \max(\text{ERKPP})\ )\ )\ )\ ]$$

Robin Donaldson and David Gilbert (2008). A Model Checking Approach to the Parameter Estimation of Biochemical Pathways In proceedings CMSB 2008 (Computational Methods in Systems Biology). LNCS 5307/2008, pp269-287

Critical parameters

Critical parameters can produce sustained activation of Ras, MEK or ERK.

Used to give an idea which parameters to vary

Method:
- Vary the kinetic rate constant parameters in range ± 2 orders of magnitude from original value.

- Perform 1,000 simulations using different values for each parameter, linearly spaced in the range

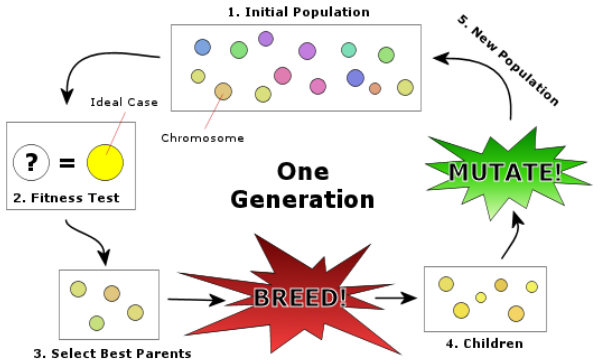
- The 'significance values' are the fraction of values in the range which give rise to sustained behaviour for each protein

- Found through model checking

parameter	sustained Ras	sustained MEK	sustained ERK
V_20	0.01	0.0	0.001
V_24	0.076	0.0	0.0
V_25	0.023	0.0	0.001
V_27	0.614	0.0	0.0
V_28	0.478	0.151	0.679
k1_14	0.0	0.0	0.778
k1_16	0.0	0.0	0.001
k1_18	0.001	0.0	0.807
k2_18	0.191	0.0	0.0
Km_20	0.001	0.0	0.797
kcat_21	0.001	0.0	0.688
kcat_23	0.001	0.0	0.186
Km_23	0.121	0.0	0.0
Km_25	0.001	0.0	0.157
kcat_26	0.0	0.0	0.001
Km_26	0.0	0.0	0.005

2000 models, 100 generations, 2.10^5 simulations/checks

Model construction



Starting with the EGF (transient) model, construct the NGF (sustained) model by varying the values of the ***critical*** kinetic rate constants.

I.e. minimise the distance of the model to sustained behaviour.

Could vary the initial concentrations, or topologies (ongoing research).

Genetic algorithm:

- Define a parameter space (±2 orders of magnitude for each parameter)
- Initialise a population of models randomly throughout parameter space
- For each generation of the algorithm:
  - Perform genetic operations on binary representations of the models' parameter values (crossover, mutation, reproduction)
  - Evaluate all models' fitness values using model checking – fitness is the distance to sustained behaviour
  - Probabilistically select models to survive to the next generation based on their fitness value

Fitness function using PLTLc

Probability:

- Can optimise the probability of a behaviour, which works fine on stochastic models.
  - On continuous models, the probability is boolean so not good in a fitness function – no gradient

Free variables:

- Can use the free variables in a PLTLc behaviour, works for continuous or stochastic models.
  - Can always get a numerical value for the fitness function, even in continuous models – good for search algorithm

- We specify the behaviour in PLTLc and at the same time characterise the 'tail' of the peak in a free variable.
- We have an idea of the desired behaviour of the tail and can calculate the distance, using the free variables, to give us a numerical value for the fitness function, whilst….
- the behaviour in PLTLc enforces a peak at the right position.

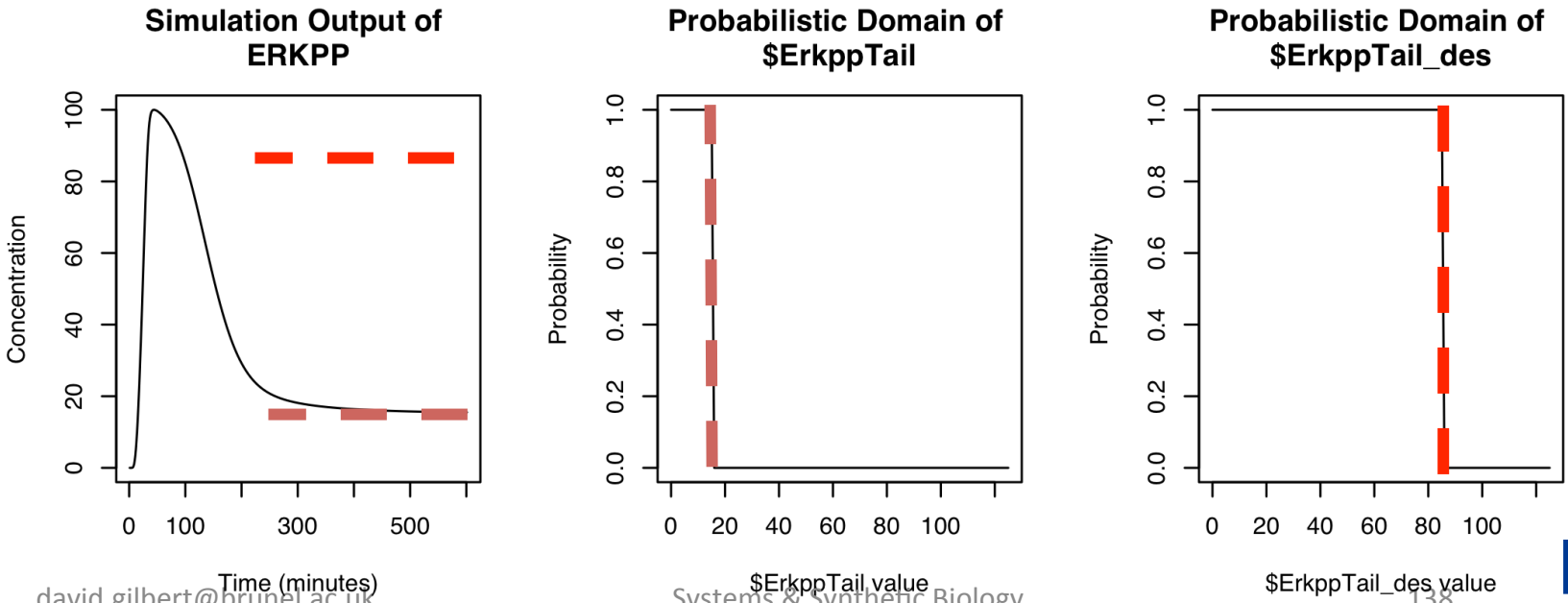
Fitness function using PLTLc

Fitness function for sustained ERK

Enforce a peak between time 2 and 5 and characterise the tail in $ErkppTail:

P=? [(d(ERKPP) > 0) U (time ≥ 2 ∧ time ≤ 5 ∧ d(ERKPP) < 0
∧ G(ERKPP ≥ $ErkppTail))]

- Fitness function is distance between tail and 85% of the peak height (sustained activation), calculated using the probabilistic domains:



Brunel UNIVERSITY LONDON

Distance metric

Uses Residual Sum of Squares

Over probabilistic domains in range m,n

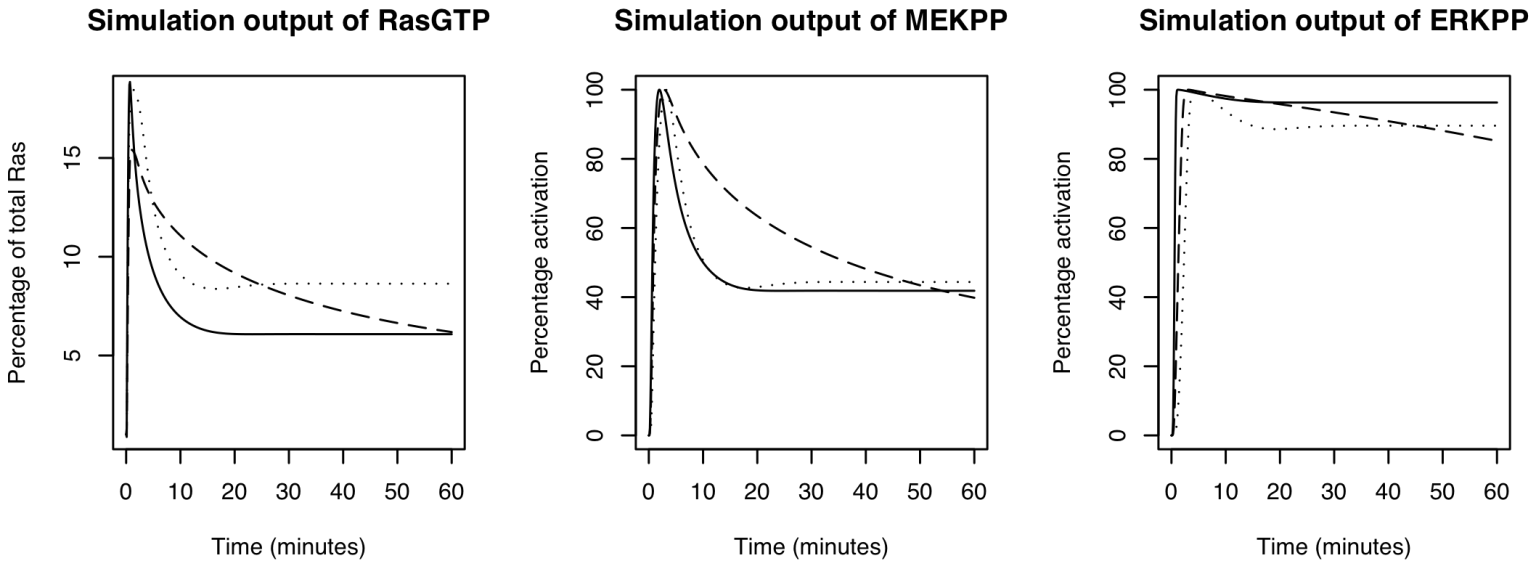
$$RSS(X,X',m,n) = \sum_{i=m}^{n} |\$X(i) - \$X'(i)|^2$$

Dist(Model,Model')=

(dist(Ras)+dist(MEKpp)+dist(ERKpp))/3

Results

- Built a fitness function for sustained Ras, MEK and ERK
- Ran the genetic algorithm with 100 generations with results:



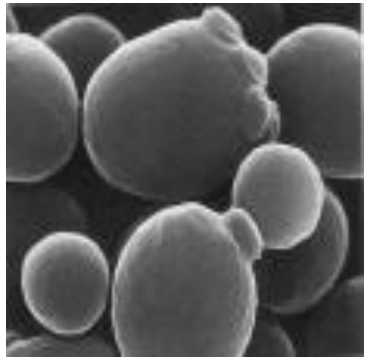
- Original model of the NGF signalling pathway varying V28 (dotted)
- Best model returned when varying the critical parameters (solid)
- Critical parameters without V28 (dashed).

The best model returned when varying the critical parameters only required a ***16-fold*** increase in V28 (compared with 40-fold in original paper)

Even possible to get similar behaviour **without** varying V28

Robin Donaldson and David Gilbert (2008). A Model Checking Approach to the Parameter Estimation of Biochemical Pathways In proceedings CMSB 2008 (Computational Methods in Systems Biology). LNCS 5307/2008, pp269-287

What about scaling up?

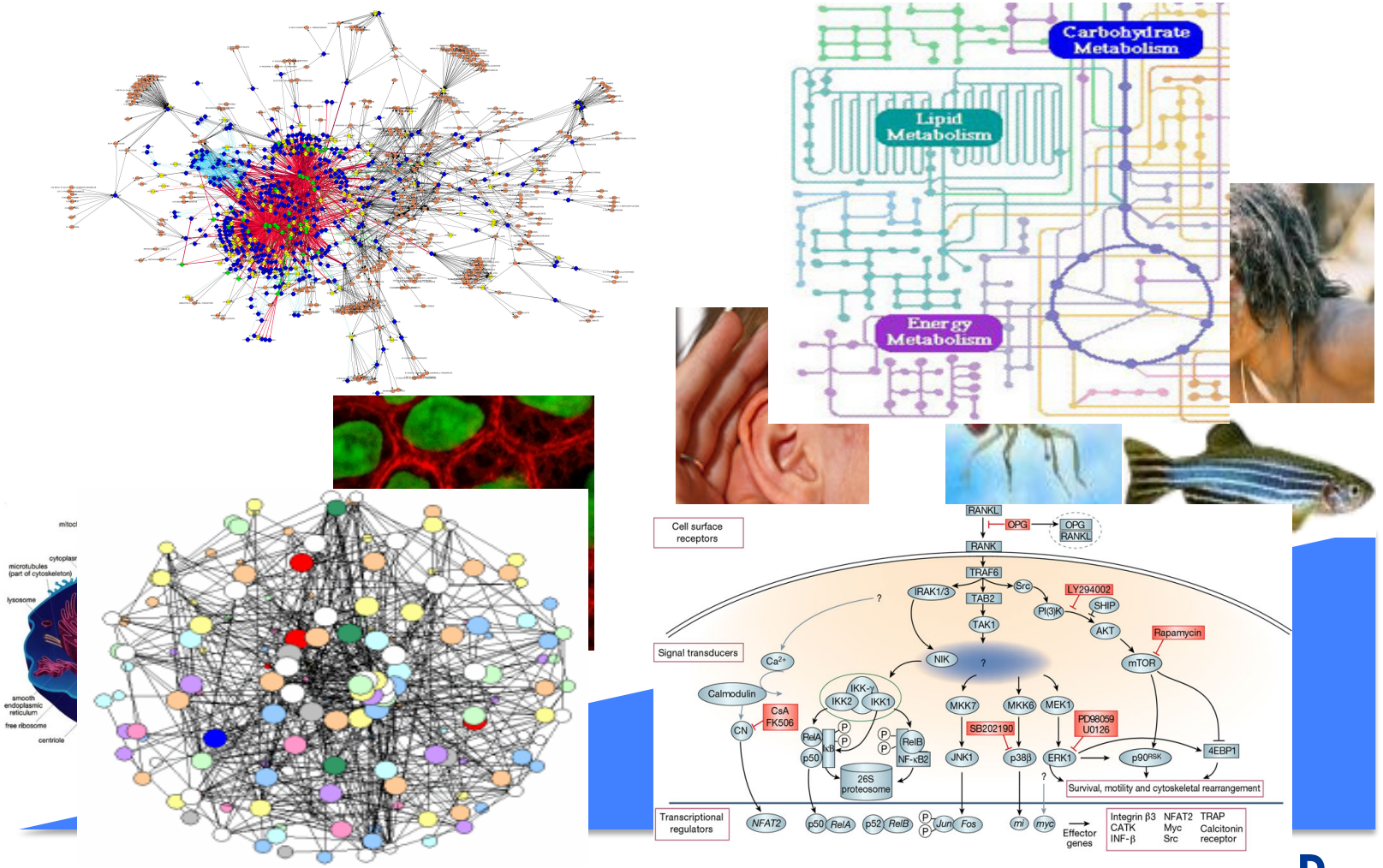







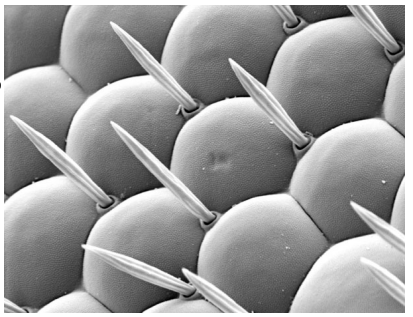
Brunel UNIVERSITY LONDON

Multiscale Modelling in Biology



Brunel UNIVERSITY LONDON

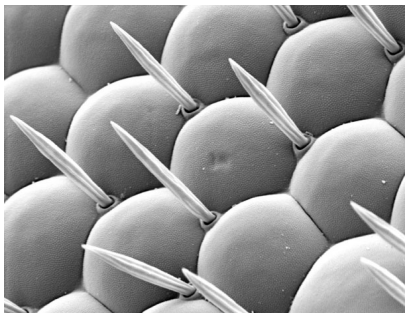
Multiscale Modelling Challenges



- ***Repetition*** – multiple components with similar definitions
- ***Variation*** – genetic mutants; random variants
- ***Organisation -*** regular / irregular patterns in 1, 2 or 3 dimensions
- ***Communication*** – short & long distance
- ***Hierarchical organisation*** – intra or inter cellular (tissues, organs, …)
- **Movement** – mobility (passive) & motility (active)

(Components could be molecules, organelles, cells, tissues, organs, organisms.)

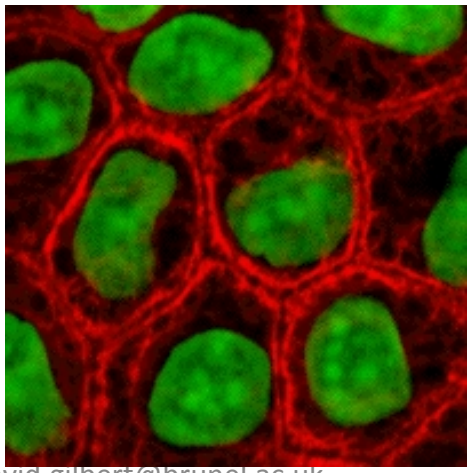
Multiscale Modelling Challenges (continued)



- **Replication** - reproduction
- **Deletion** – cell death
- **Irregular/semi-regular organisation** of components – for example a not-exact honeycomb grid.
- **Dynamic grid size** – for example alter size and/or topology of grid to model development. Also required for ability to insert/remove items.
- **Differentiation of components** - for example, differentiation of embryonic stem cells or immune cells makes a less specialized cell more specialized.
- **Pattern formation of components** - organizing a number of cells in appropriate one, two or three dimensional structures in space and time.

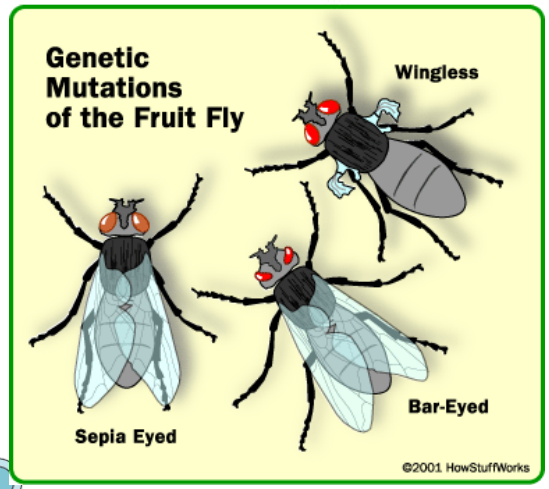
Repetition of individual components

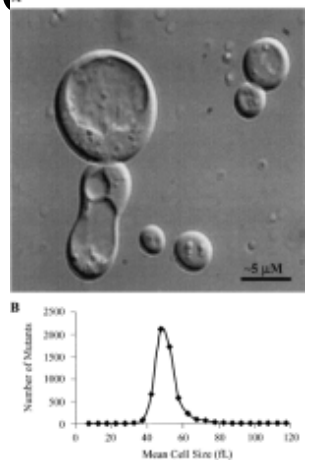
- Components within a cell (organelles etc)
- Multiple cells each of which having a similar definition
- Repeated tissue fragments
- Repeated organs (wings,…)
- Repeated individual organisms

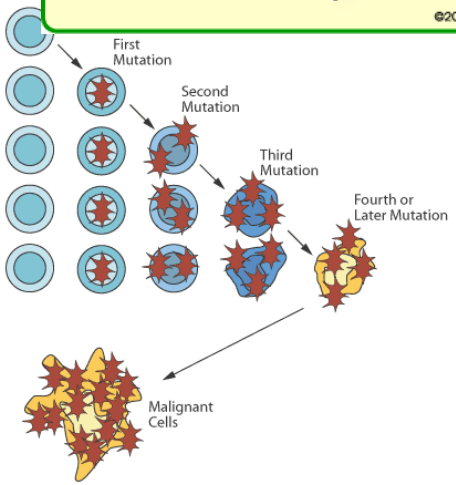


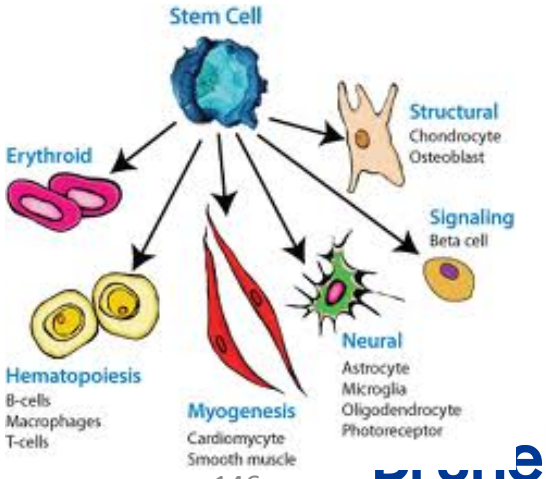
Variation

- Sets of similar components with defined variations
- Random mutation
- Genetic mutants
- Cancerous tissue
- Differentiation



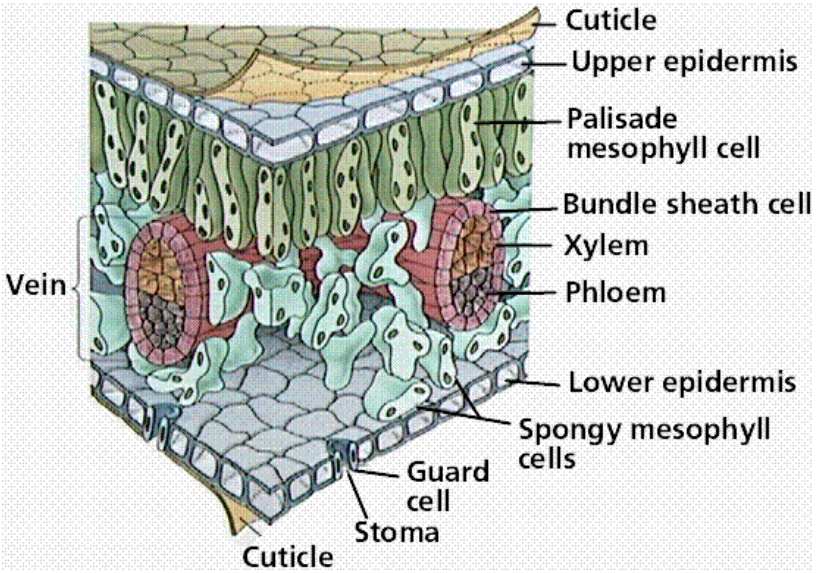




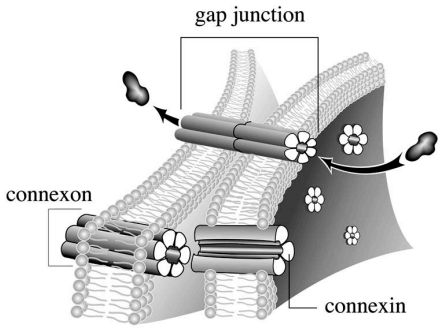


Spatial organisation

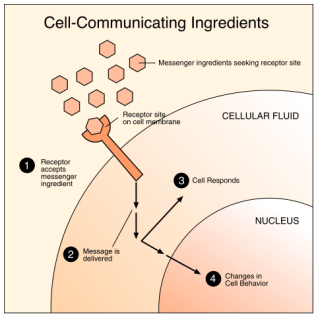
- Between cells
  - how they are organised into regular or irregular patterns over spatial networks in one, two or three dimensions.



Communication



- Between **immediate neighbours** (intracellular complexes)
- **Long-distance** (cytokines etc)



Further constraints:

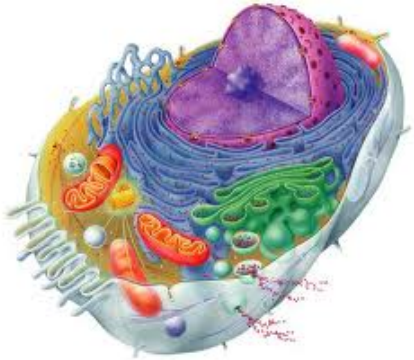
- Type of **relationship** between partners
- Type of component(s)
- History of component(s)
- **Position** of component(s) in spatial network.

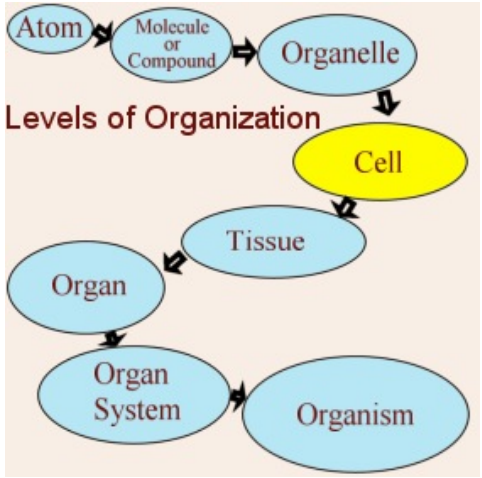
Brunel UNIVERSITY LONDON

Hierarchical organisation

Components containing repeated sub-components

- Cell containing s artm components.





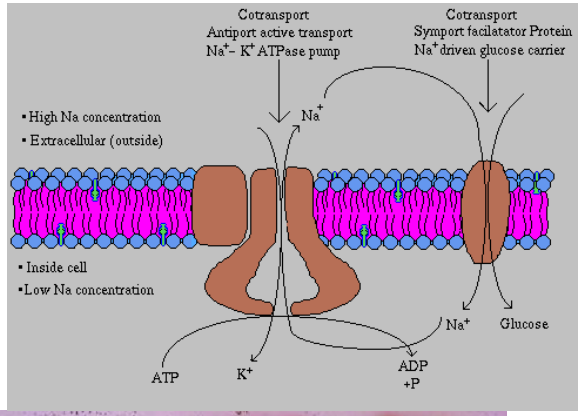
Enables the use of abstraction over level of detail used to describe components

Movement

- Mobility – passive movement.

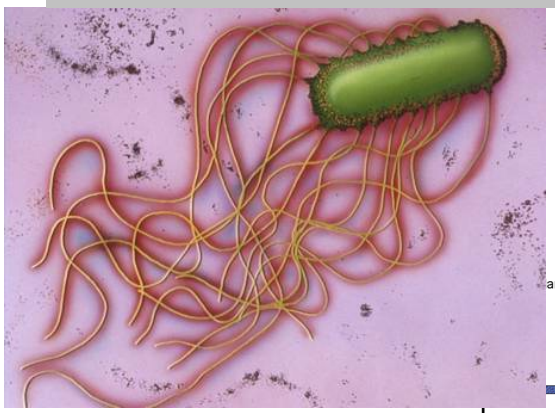
  Protein transport

  Sodium transport

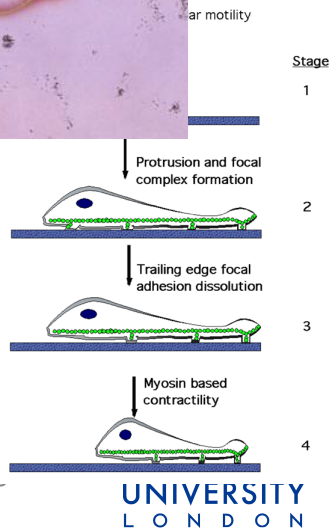


- Motility – active movement.

  Cells using organelles (flagellae)

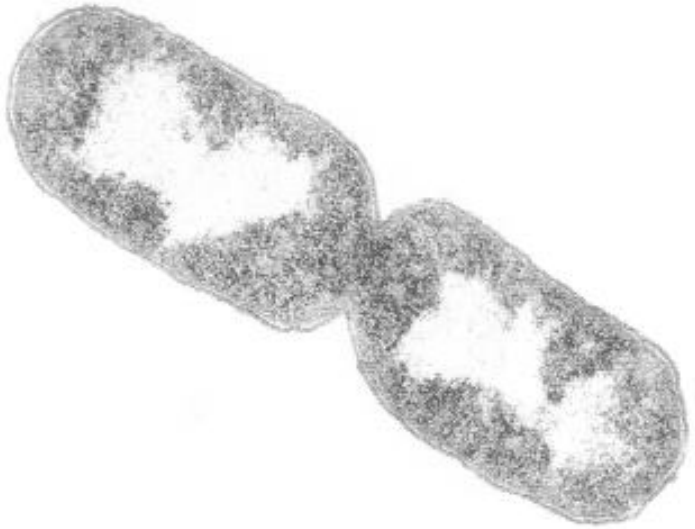


General cellular motility



UNIVERSITY LONDON

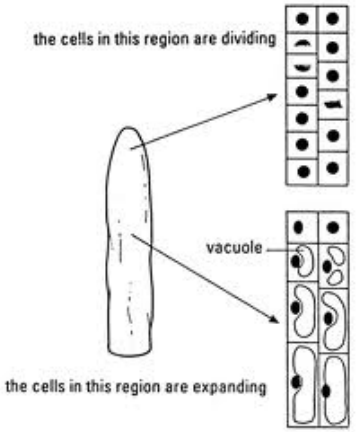
Replication



- E.g. cell division

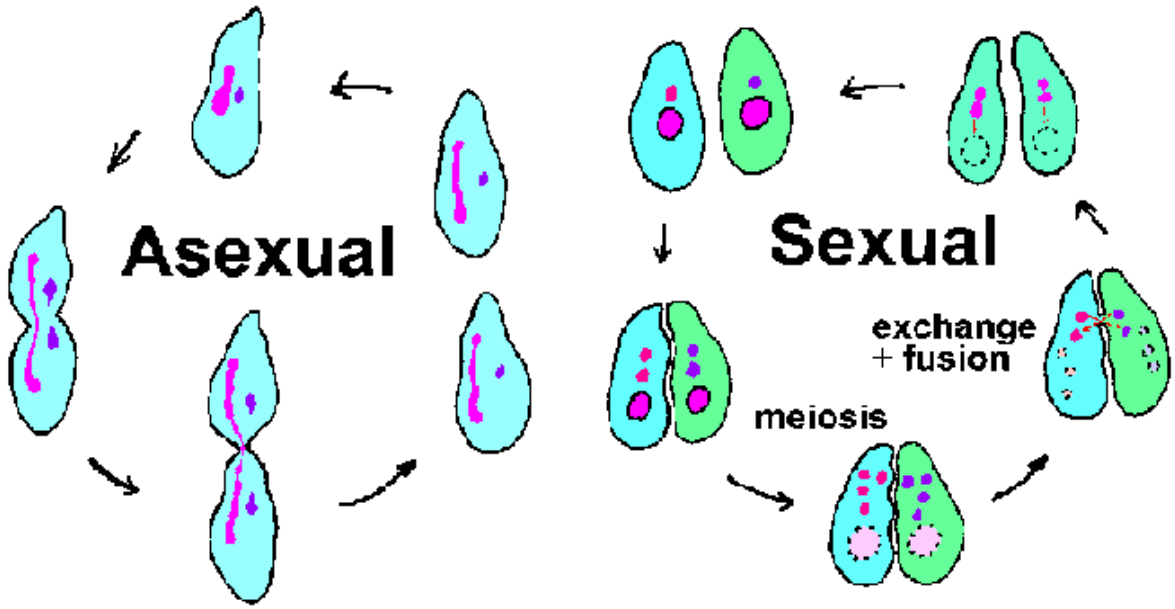
Can take into account:

- Mutation
- Spatial organisation / position

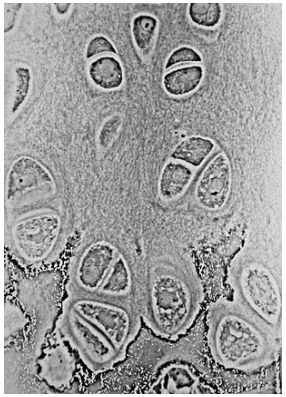


Exchange

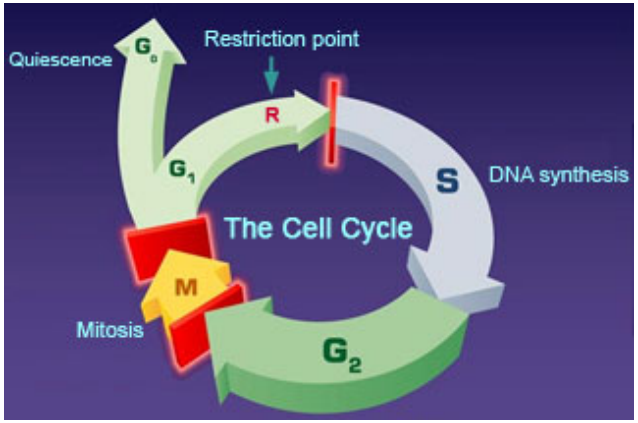
- Exchange of (genetic) information
- Sexual
- Asexual

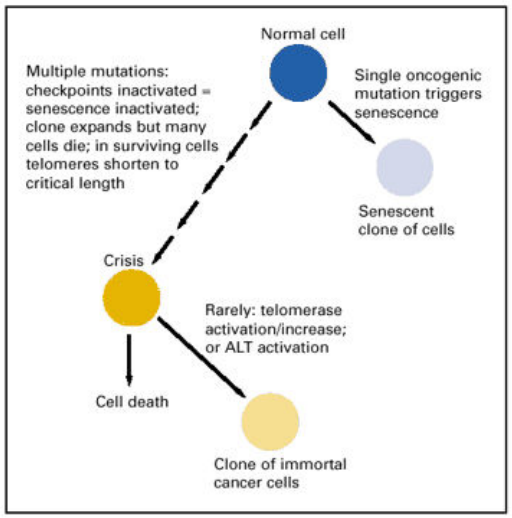


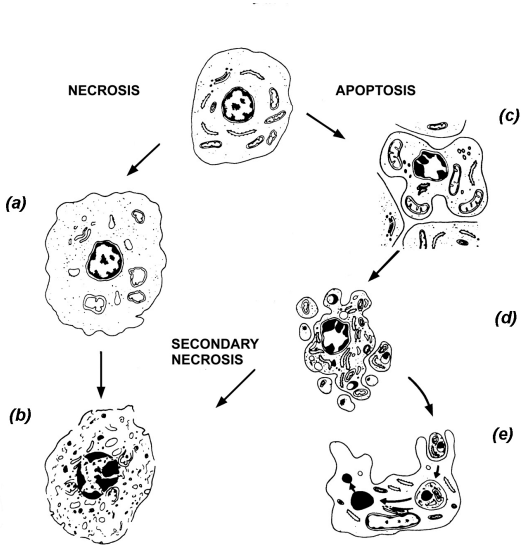
Death etc



- Cell death:
  - apoptosis (programmed), necrosis (traumatic)
- Quiescence
- Senility







Brunel UNIVERSITY LONDON

Coloured Petri Nets (CPN)

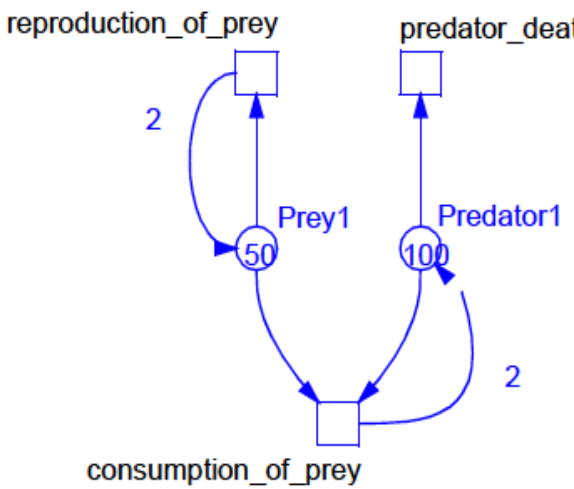
- Tokens distinguished via their colours.
- Each place gets a colour set, specifying the kind of tokens which can reside on the place.
- Each transition gets a guard, specifying which coloured tokens are required for firing.
- Each arc gets an arc expression specifying the kind of tokens flowing through it

- Allows for the discrimination of species (molecules, metabolites, proteins, secondary substances, genes, etc.).
- Colours can be used to distinguish between sub-populations of a species in different locations (cytosol, nucleus and so on).

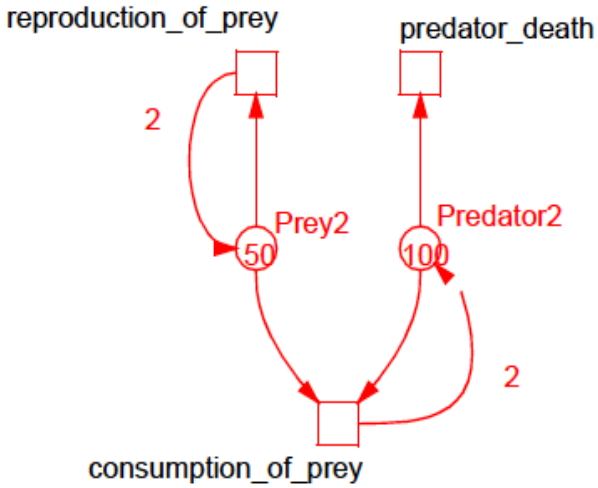
Repeated Petri Nets…

Example: Prey-Predator

sub-system1



sub-system2



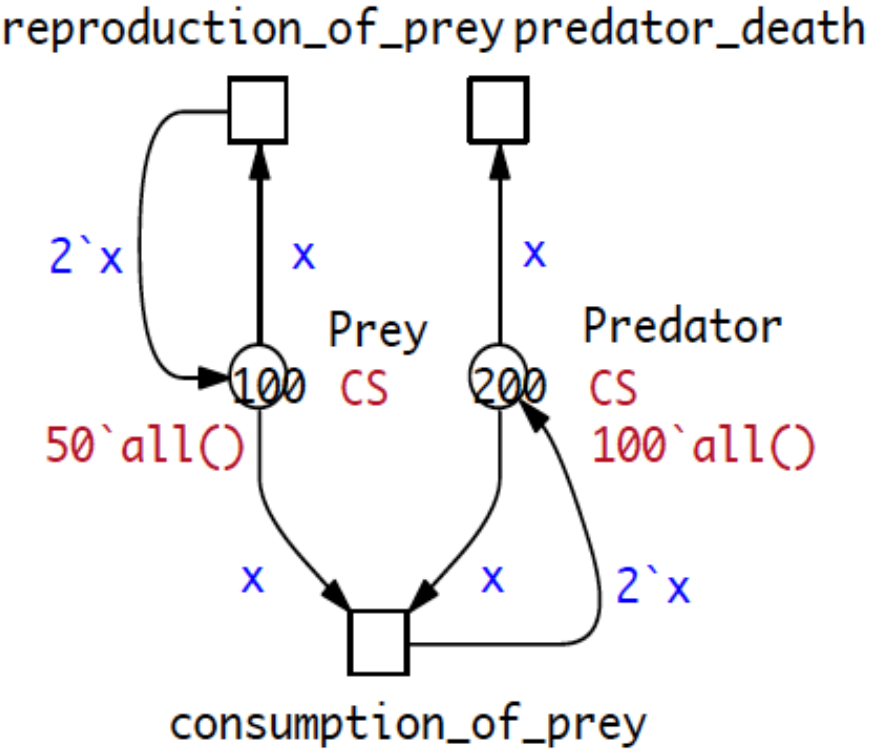
… folded to Coloured Petri Nets (CPN)

Example: Prey-Predator

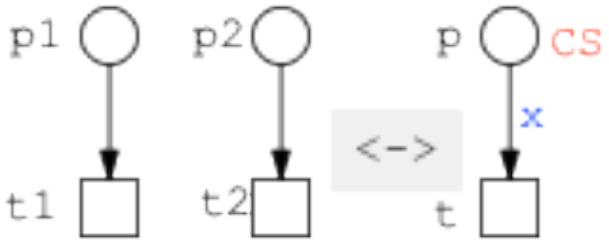
- Definitions

colourset CS = 1-2;

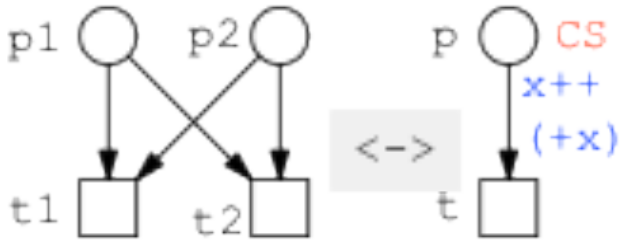
var x: CS;



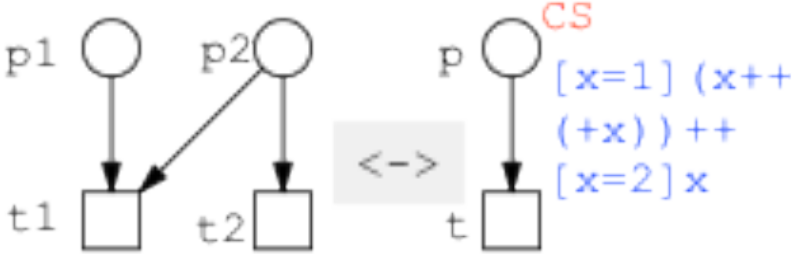
CPN Folding



(a)



(b)



(c)

```
Declarations:
colorset CS = int with 1,2;
variable x : CS;
```

(d)

++: multiset addition
(+x): successor
[x=2]: guard

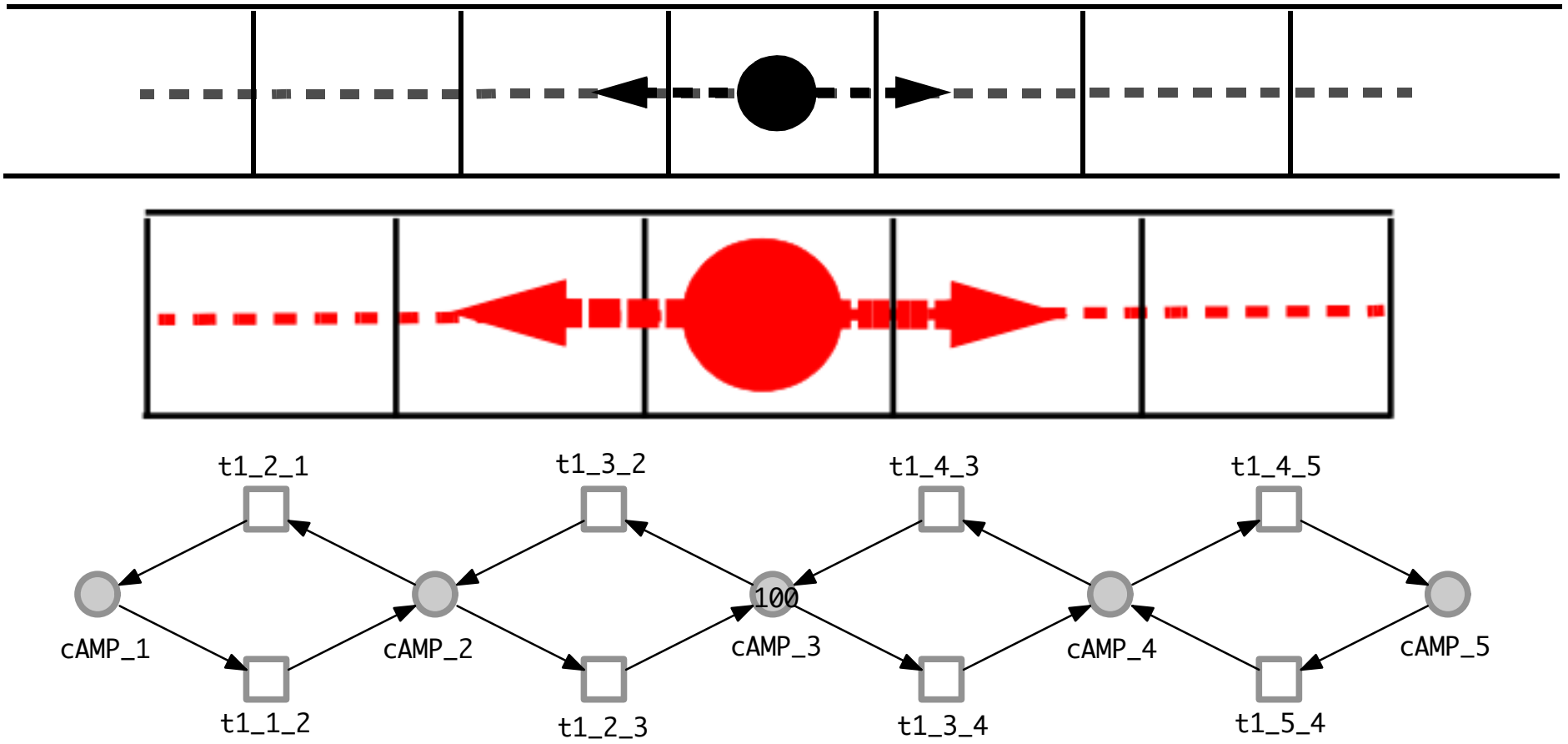
Brunel UNIVERSITY LONDON

Coloured Petri Nets (CPN)

Permits

- *Discrimination of species* (molecules, metabolites, proteins, secondary substances, genes, etc.).
- *Repeated* elements
- *Locality*: distinguish between sub-populations of a species in different locations (cytosol, nucleus and so on).

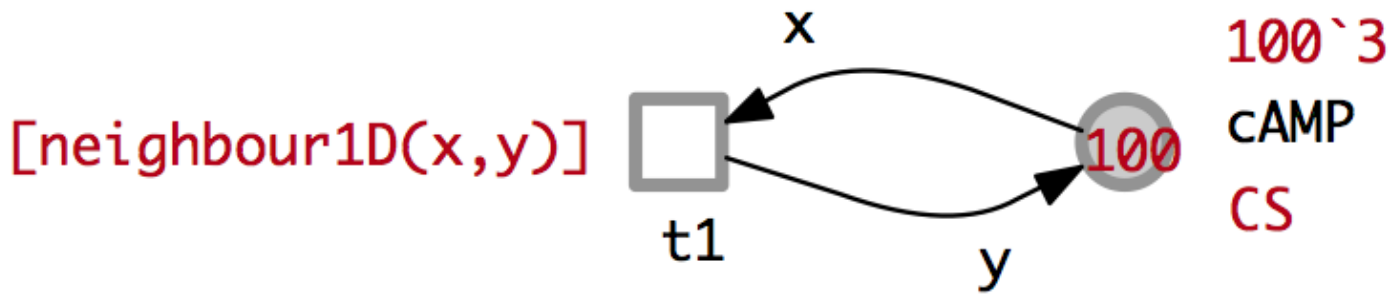
1-D diffusion



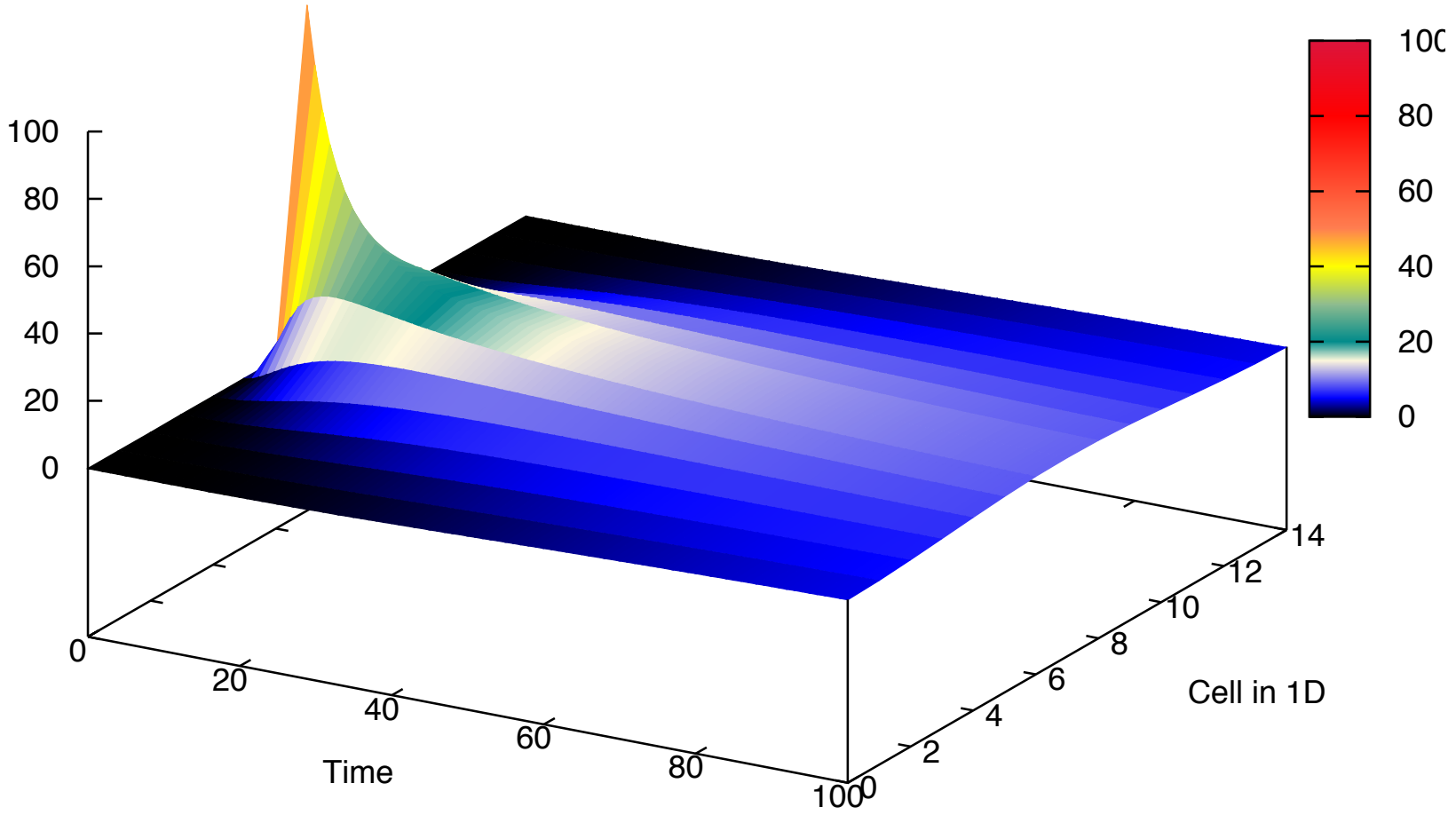
1D

- **definitions**

```
const D1 = 5;              // grid size
colorset CS = 1-D1;        // grid positions
var x,y : CS;

function neighbour1D (CS x,a) bool:
  // a is neighbour of x
  ( a=x-1 | a=x+1)  & (1<=a) & (a<=D1);
```

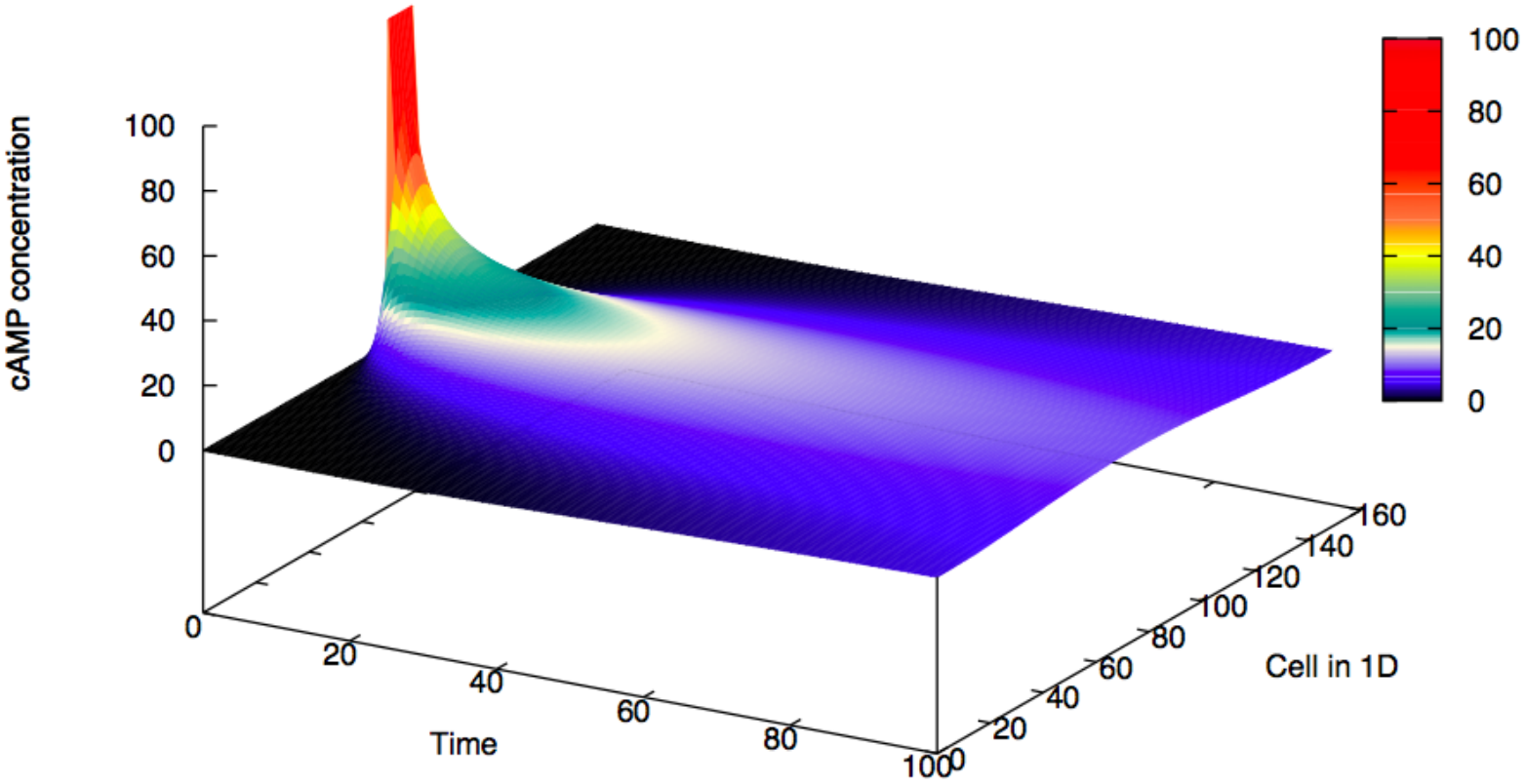


1D, 15 grid positions



Brunel UNIVERSITY LONDON

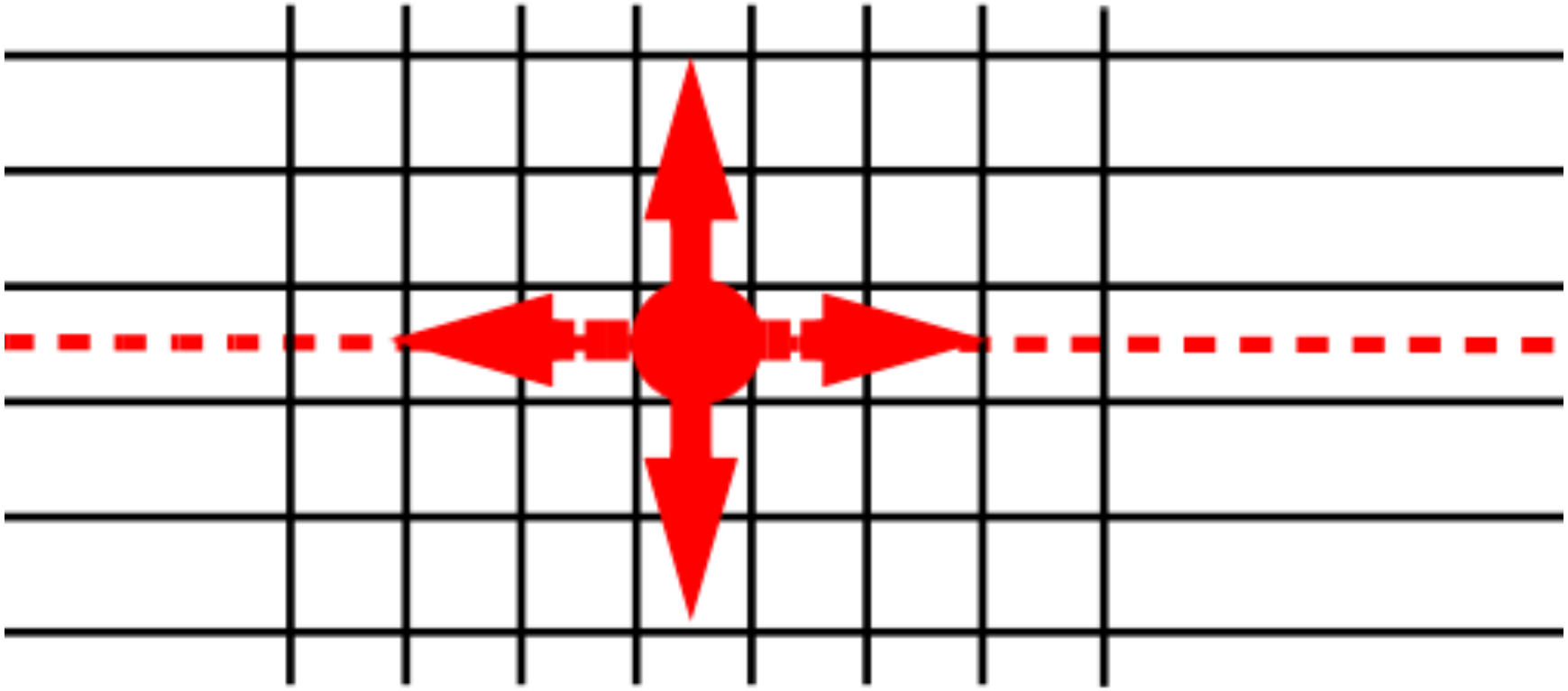
1D, 150 grid positions, scaling



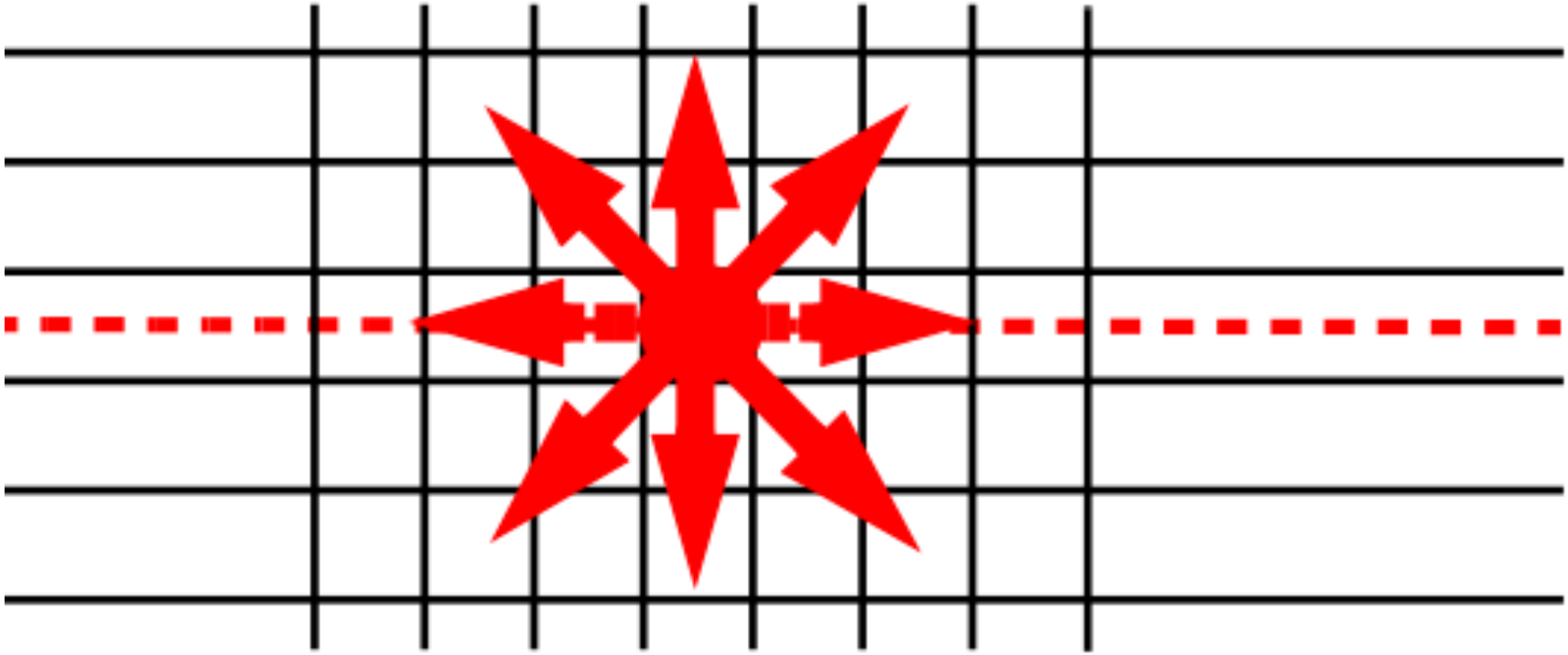
150 GRID POSITIONS, SCALING OF INITIAL MARKING AND RATES

Brunel UNIVERSITY LONDON

2-D (4)



2-D (8)



2D (8)

- **SCHEME**

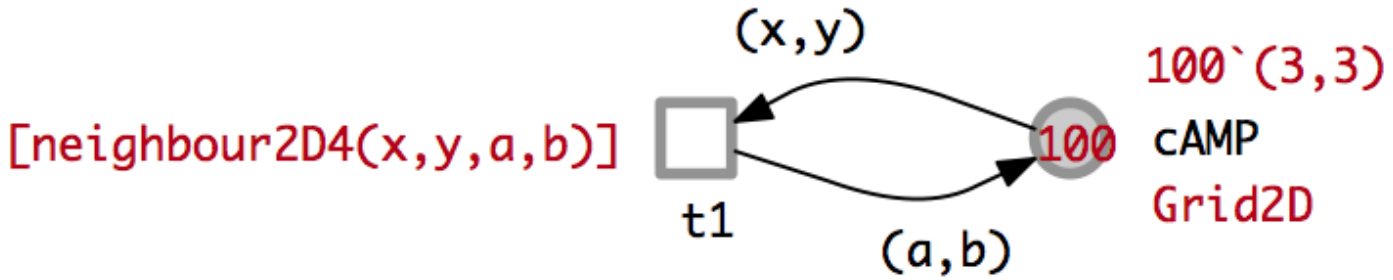
- **definitions**

```
const D1 = 5;                    // grid size first dimension
const D2 = 5;                    // grid size second dimension

colorset CD1 = 1-D1;             // row index
colorset CD2 = 1-D2;             // column index
colorset Grid2D = CD1 x CD2;     // 2D grid

var x, a : CD1;
var y, b : CD2;
```

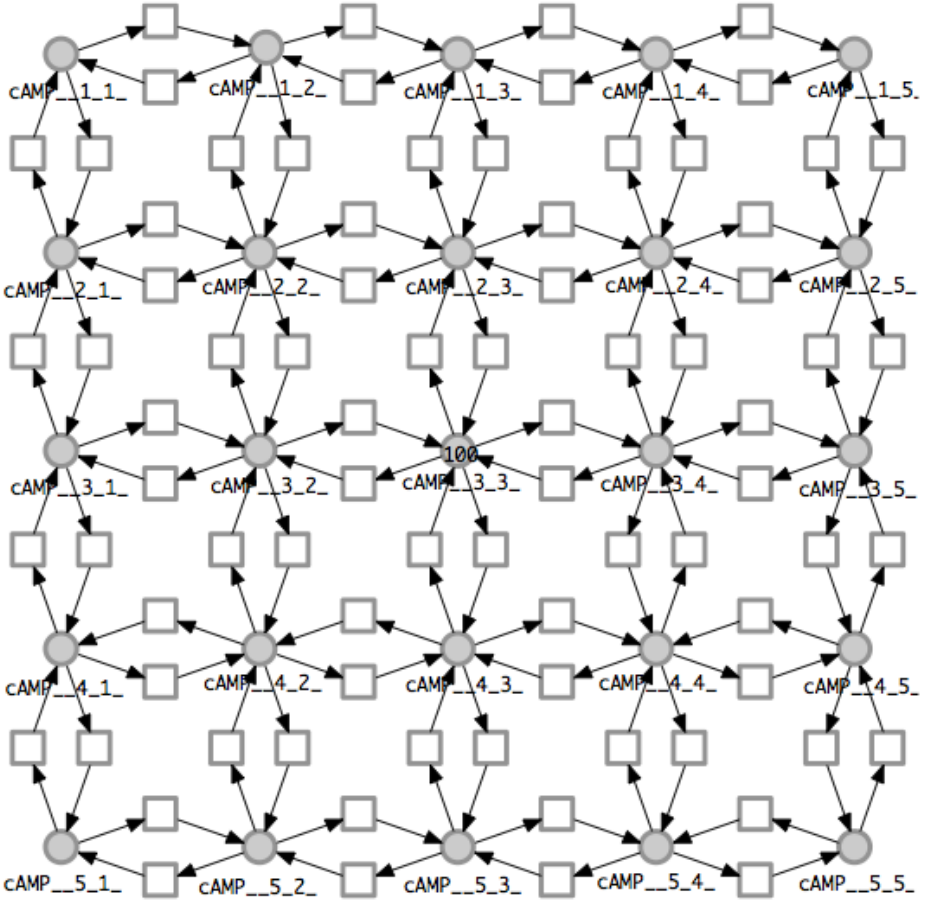
2D (4)

- **four neighbours**

```
function neighbour2D4 (CD1 x, CD2 y, CD1 a, CD2 b) bool:
  // (a,b) is one of the up to four neighbours of (x,y)
     (a=x & b=y-1) | (a=x & b=y+1)
  |  (b=y & a=x-1) | (b=y & a=x+1);
```



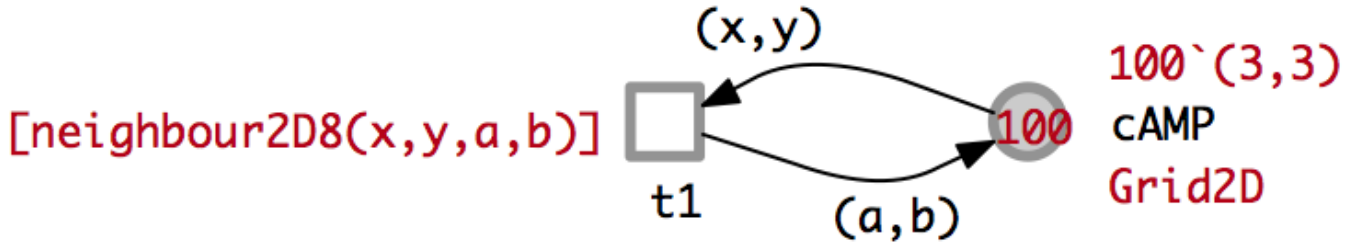
2D (4)



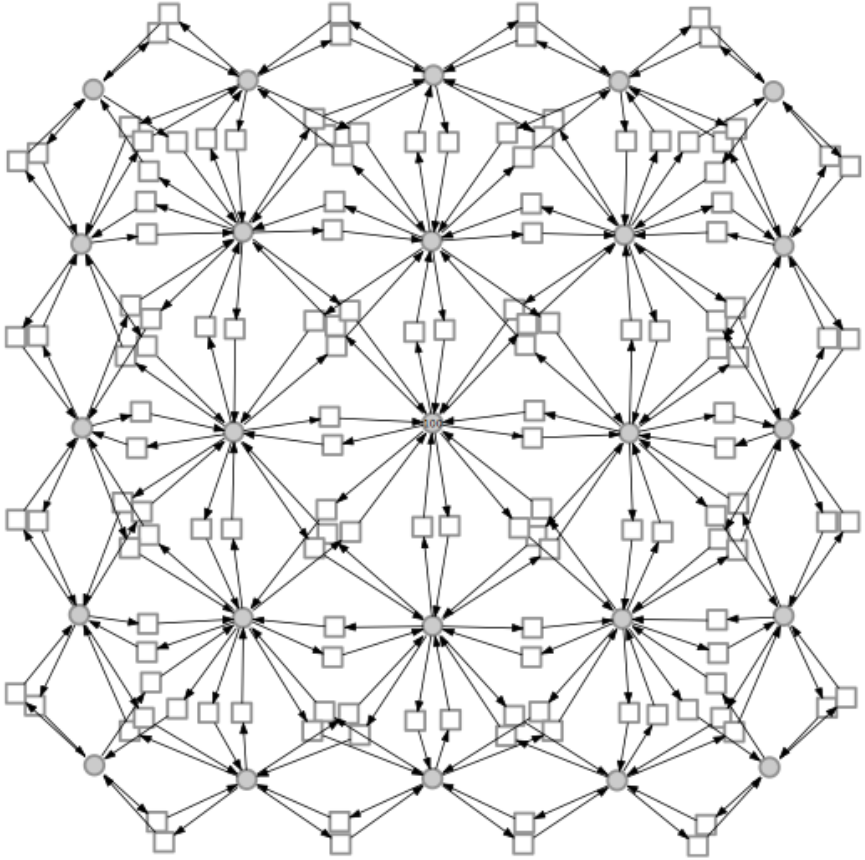
2D (8)

- **eight neighbours**

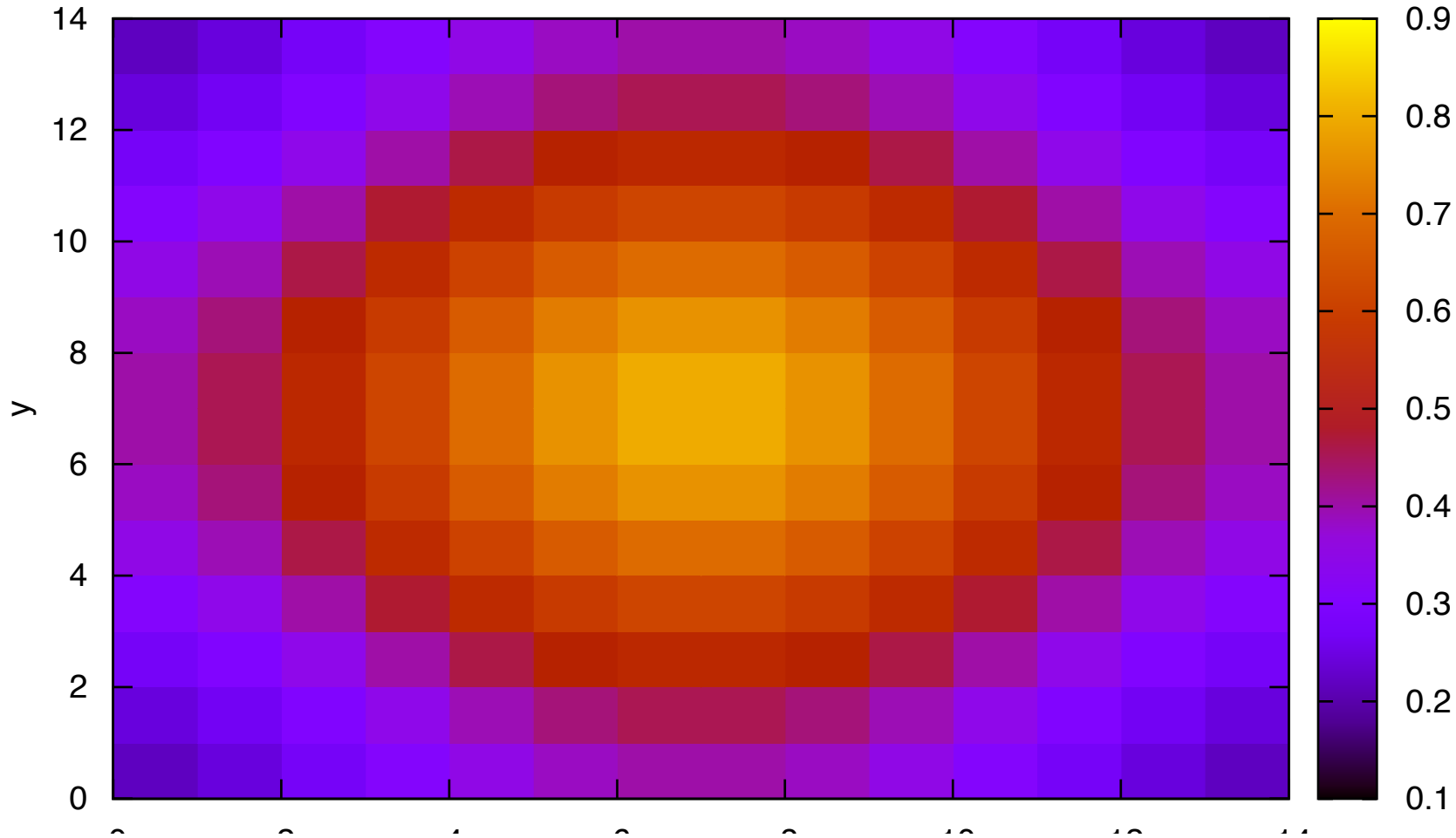
```
function neighbour2D8 (CD1 x, CD2 y, CD1 a, CD2 b) bool:
  // (a,b) is one of the up to eight neighbours of (x,y)
     (a=x-1 | a=x | a=x+1) & (b = y-1 | b=y | b=y+1)
  & (!(a=x & b=y))
  & (1<=a & a<=D1) & (1<=b & b<=D2);
```



2D (8)

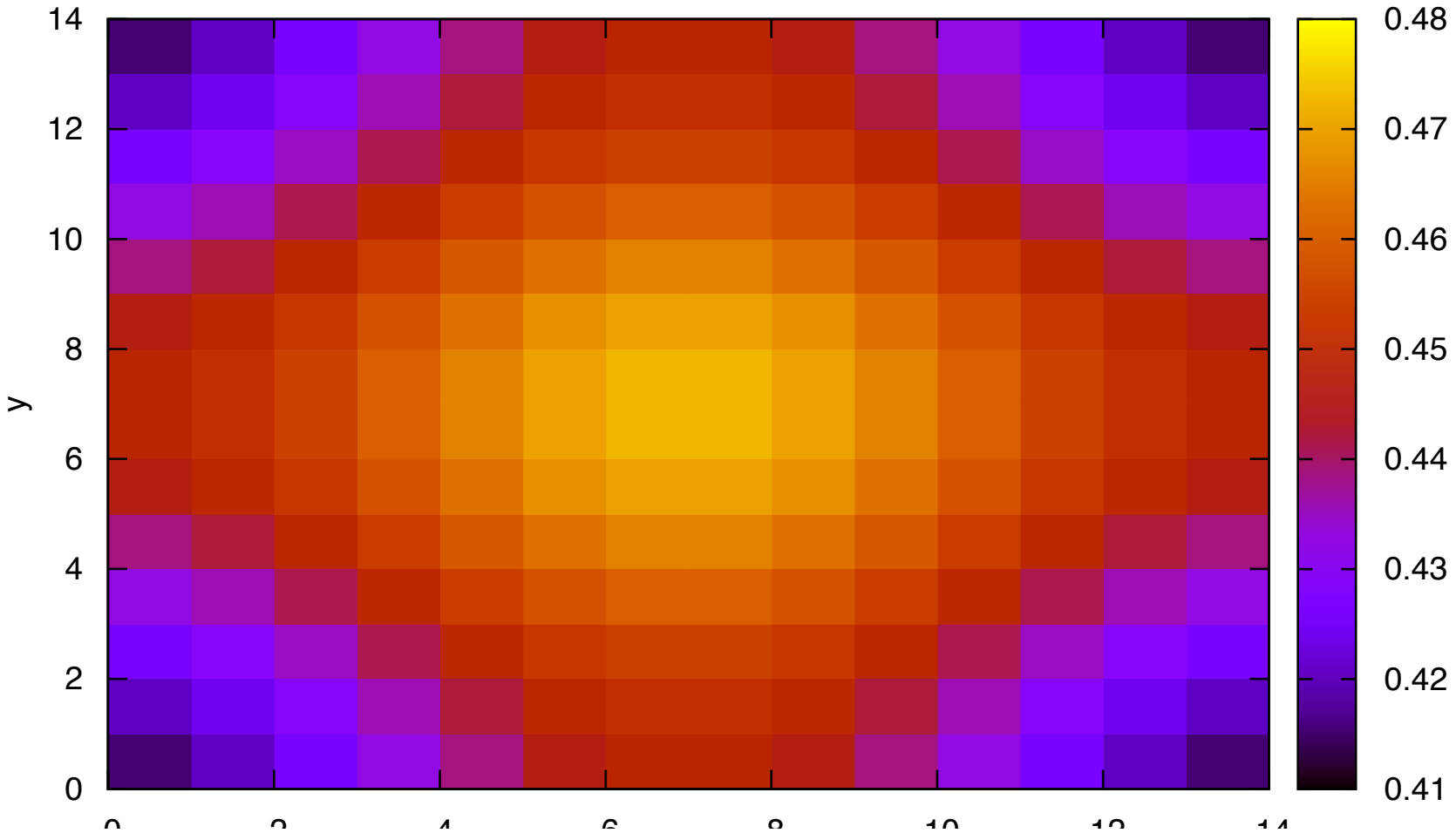


2D (4) gradient



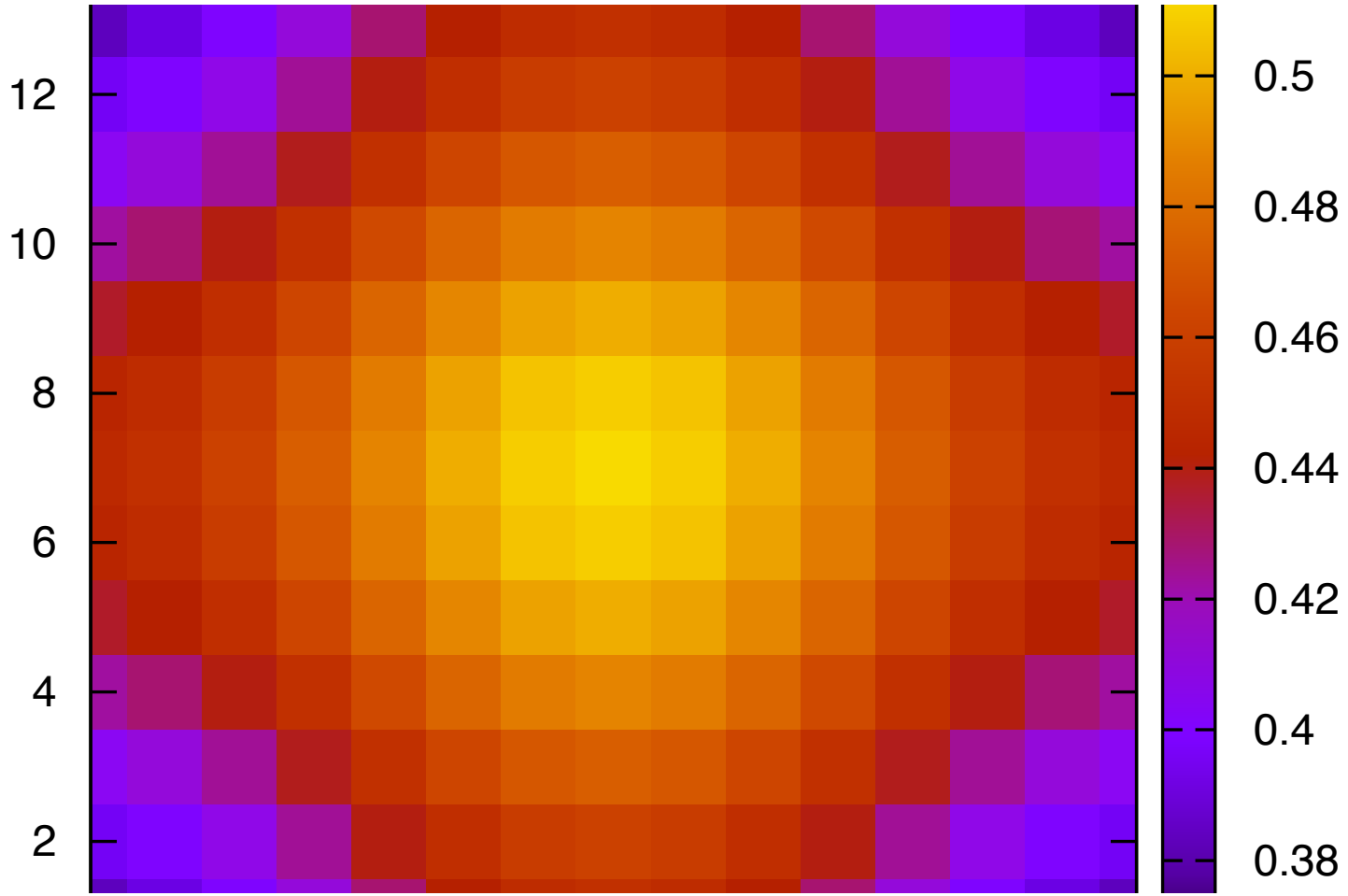
Brunel UNIVERSITY LONDON

2D (8) gradient

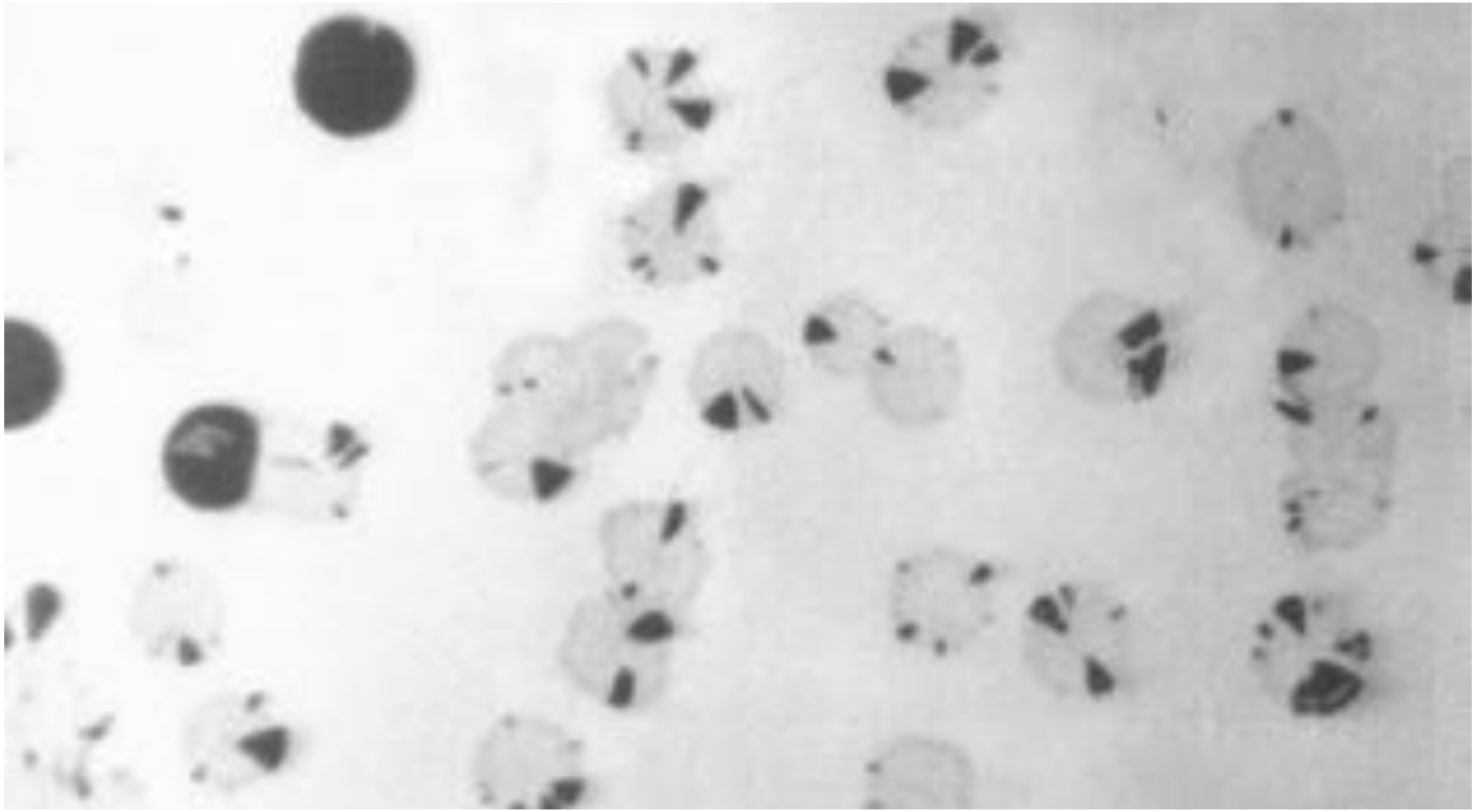


Brunel UNIVERSITY LONDON

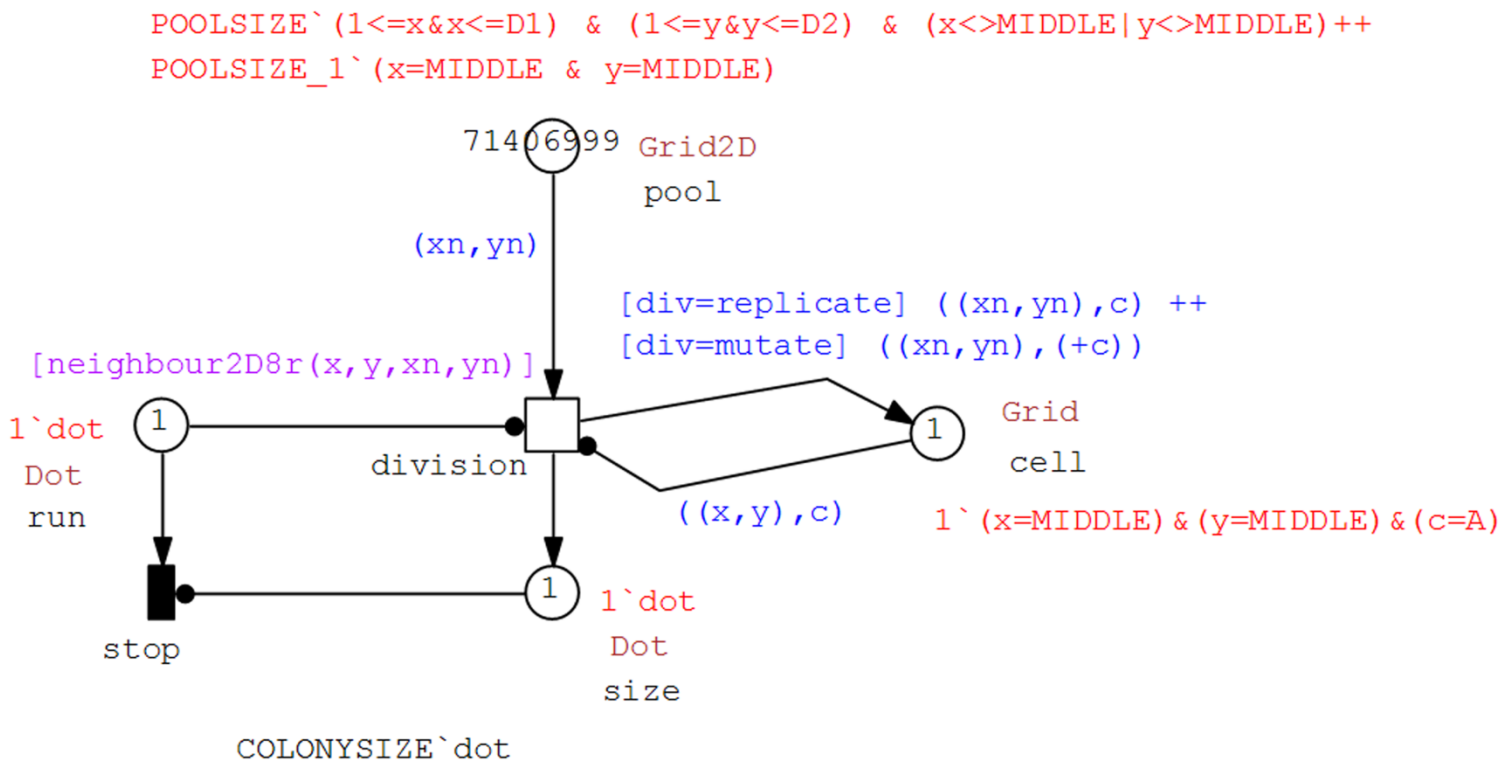
2D (8) gradient, higher resolutionb



Modelling phase variation in bacterial colonies (Monika Heiner, Ovidiu Parvu)



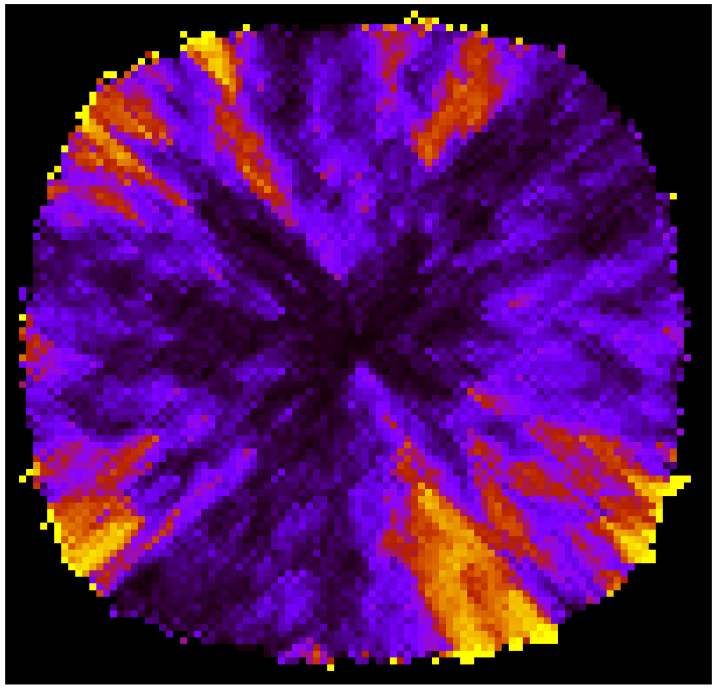
Models cont'd



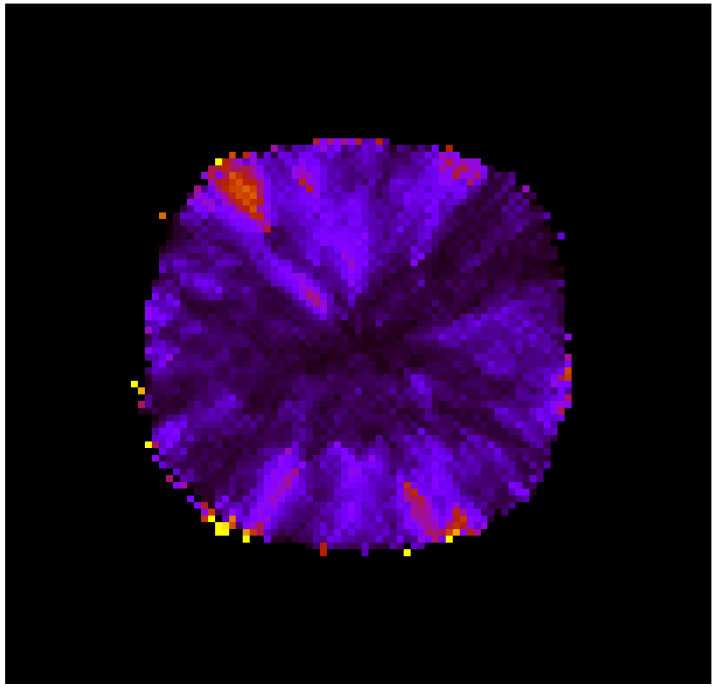
Rectangular

Parameter scanning

(Variable height)

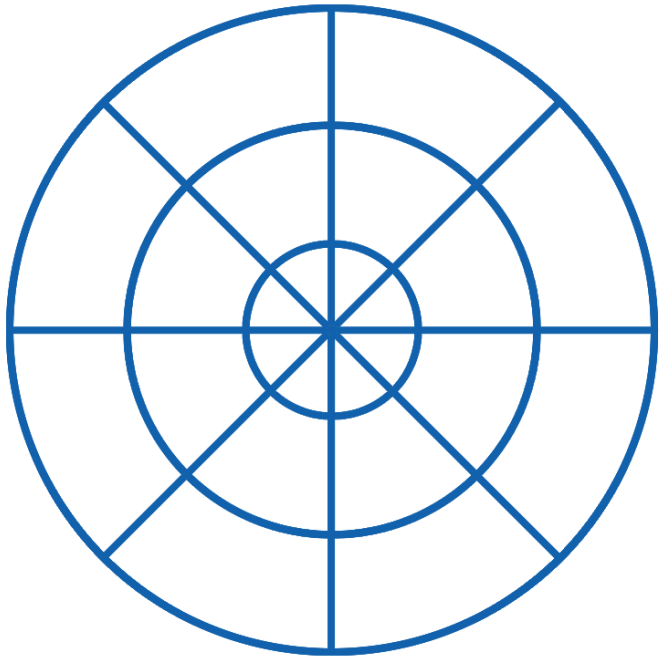


height = 12



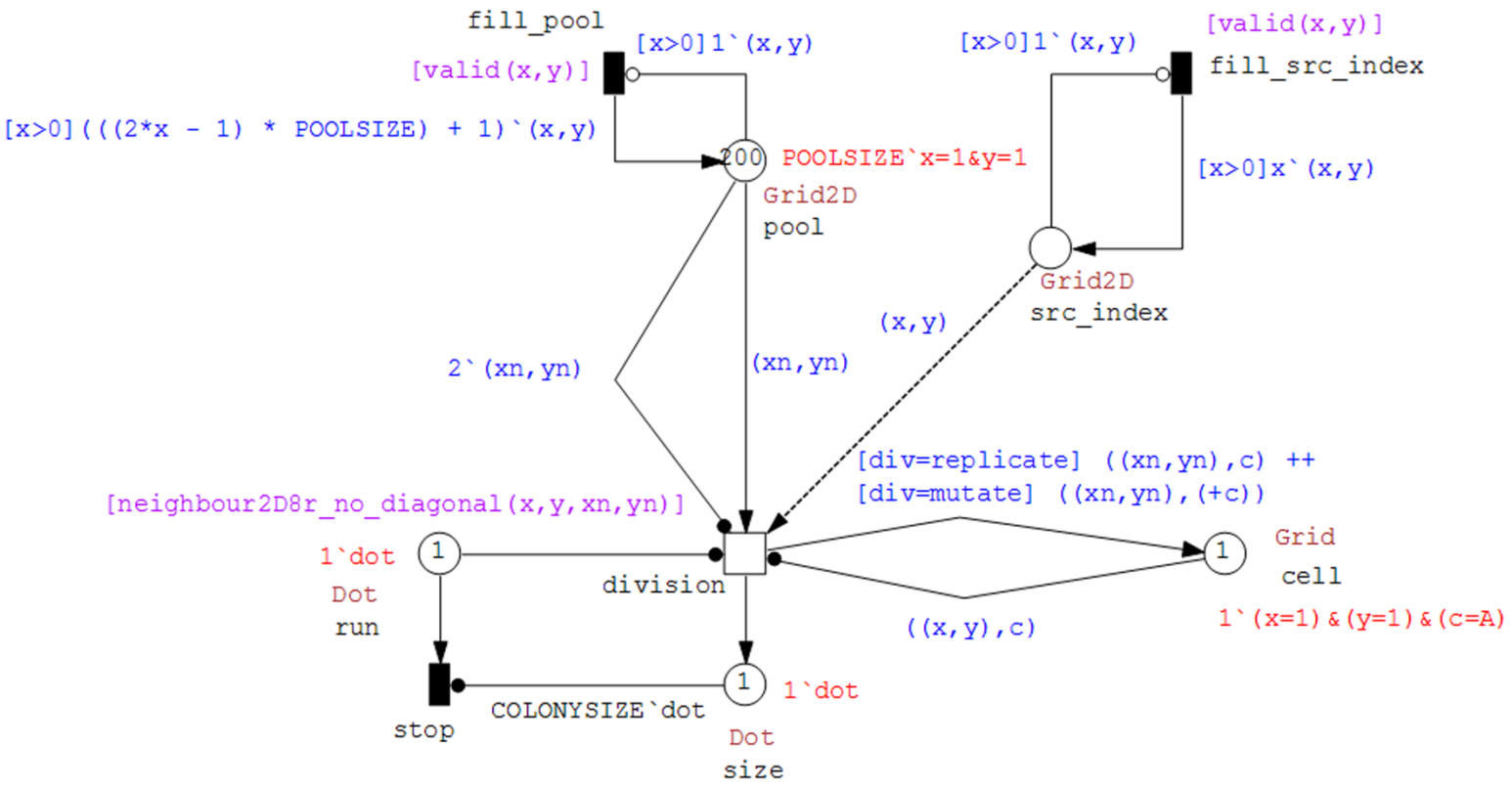
height = 29

Discretising the polar space



… obtaining ***N x M*** annular sectors.

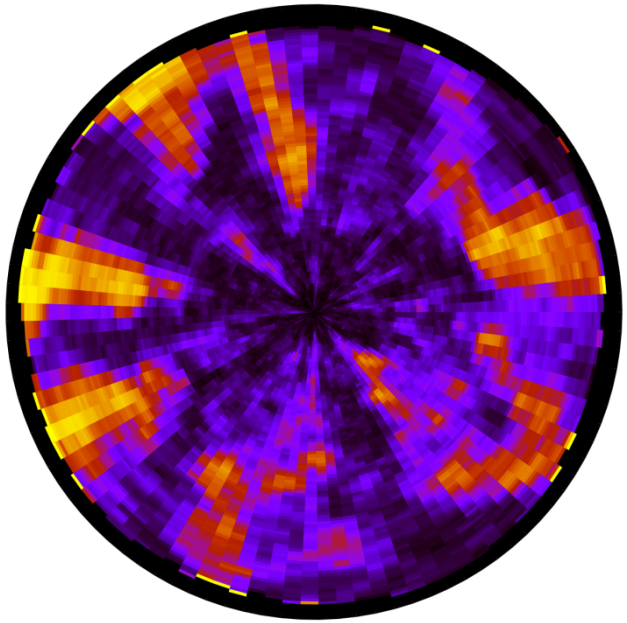
Models



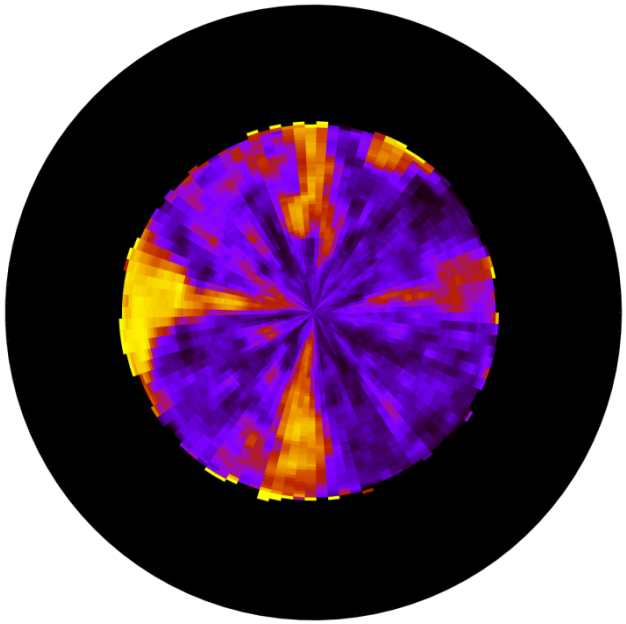
Circular

Parameter scanning

(Variable height)

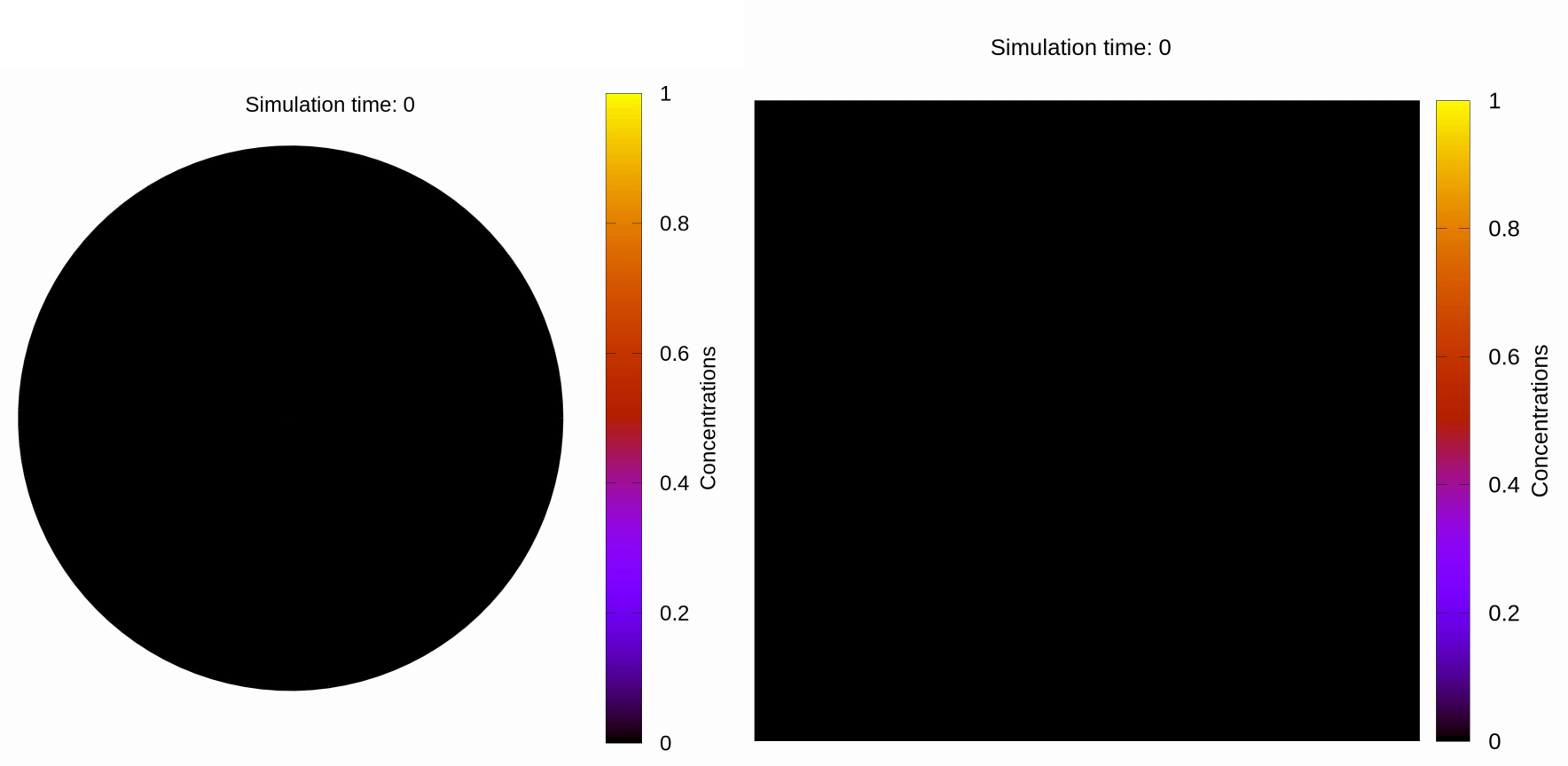


height = 12

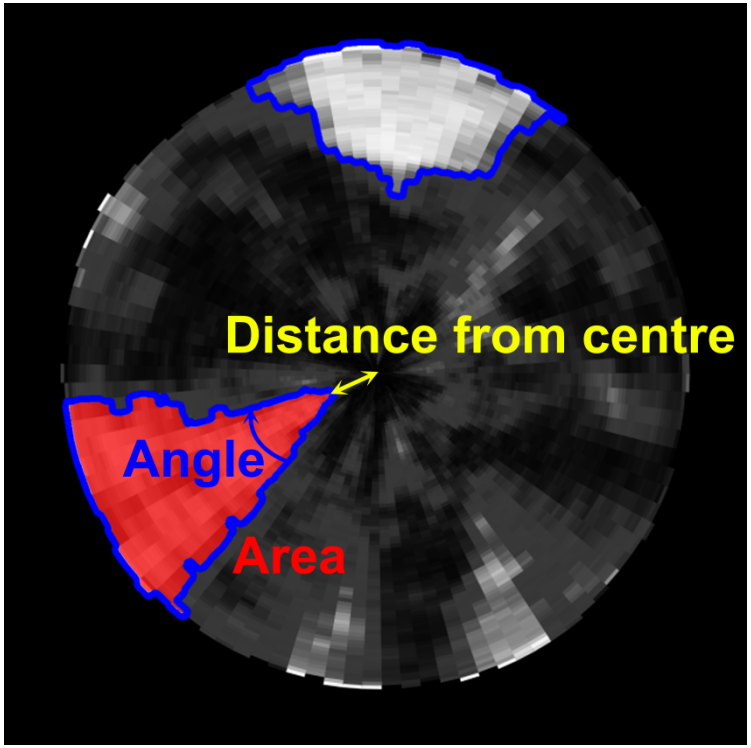


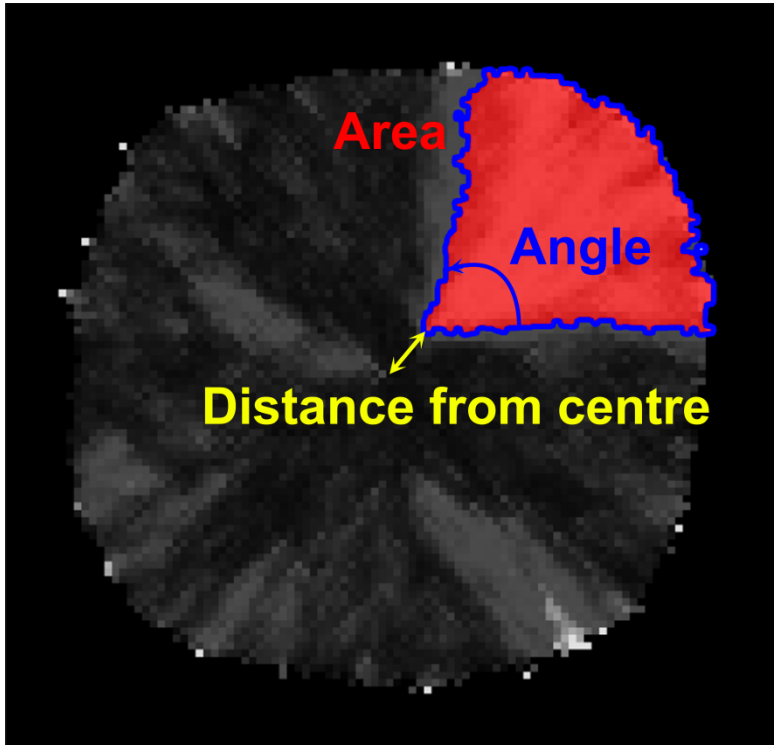
height = 29

Visualising results and “visual” model verification



Spatial analysis

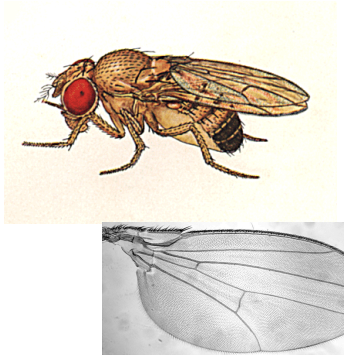


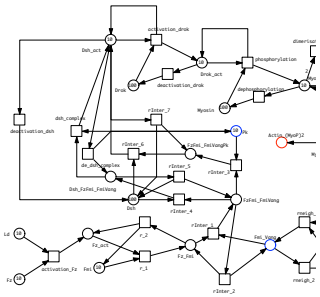


Phase variation simulations

- Rectangular
- Circular

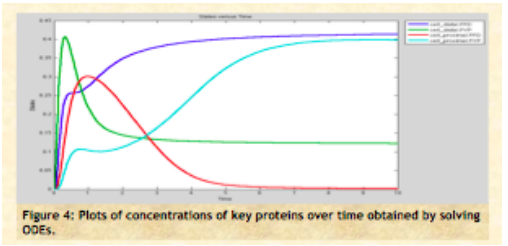
Multiscale: from signalling to organs





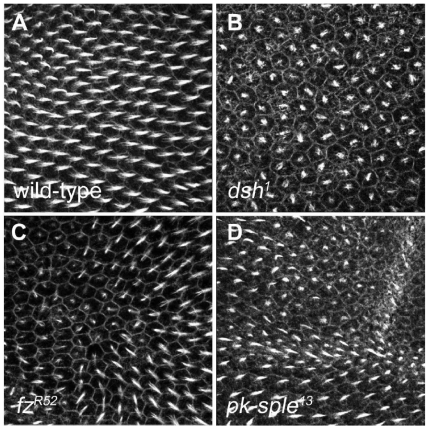
Petri nets (*coloured, hierarchical*)

Monika Heiner

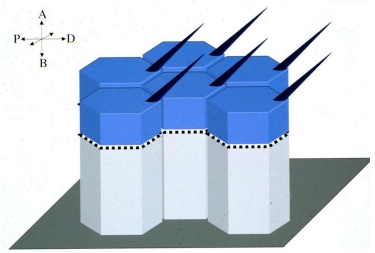


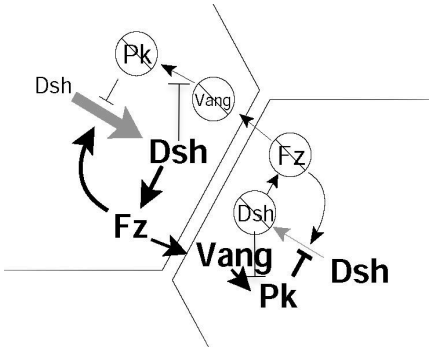
ODEs, stochastics

Planar Cell Polarity



→ **Design & genetically engineer 'patterns'!**





Pam Gao, David Tree

Brunel UNIVERSITY LONDON

Drosophila

What is great about flies:

- **Short life cycle** - 10 days in a single vial
- **Cheap** - small, simple food source
- **'Small' genome** (100 megabases)
  - Around 14,000 genes
  - Only 3 major chromosomes
- Good vertebrate model
  - 75% human genes
  - 60% human disease gene homologues

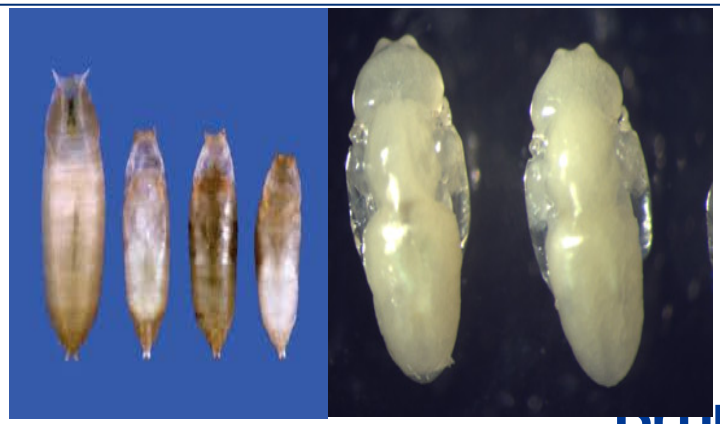
Disadvantages

•Cannot be frozen/stored **- require continuous culture**

- **No targeted mutagenesis** - e.g. homologous recombination
- **Able to fly**!

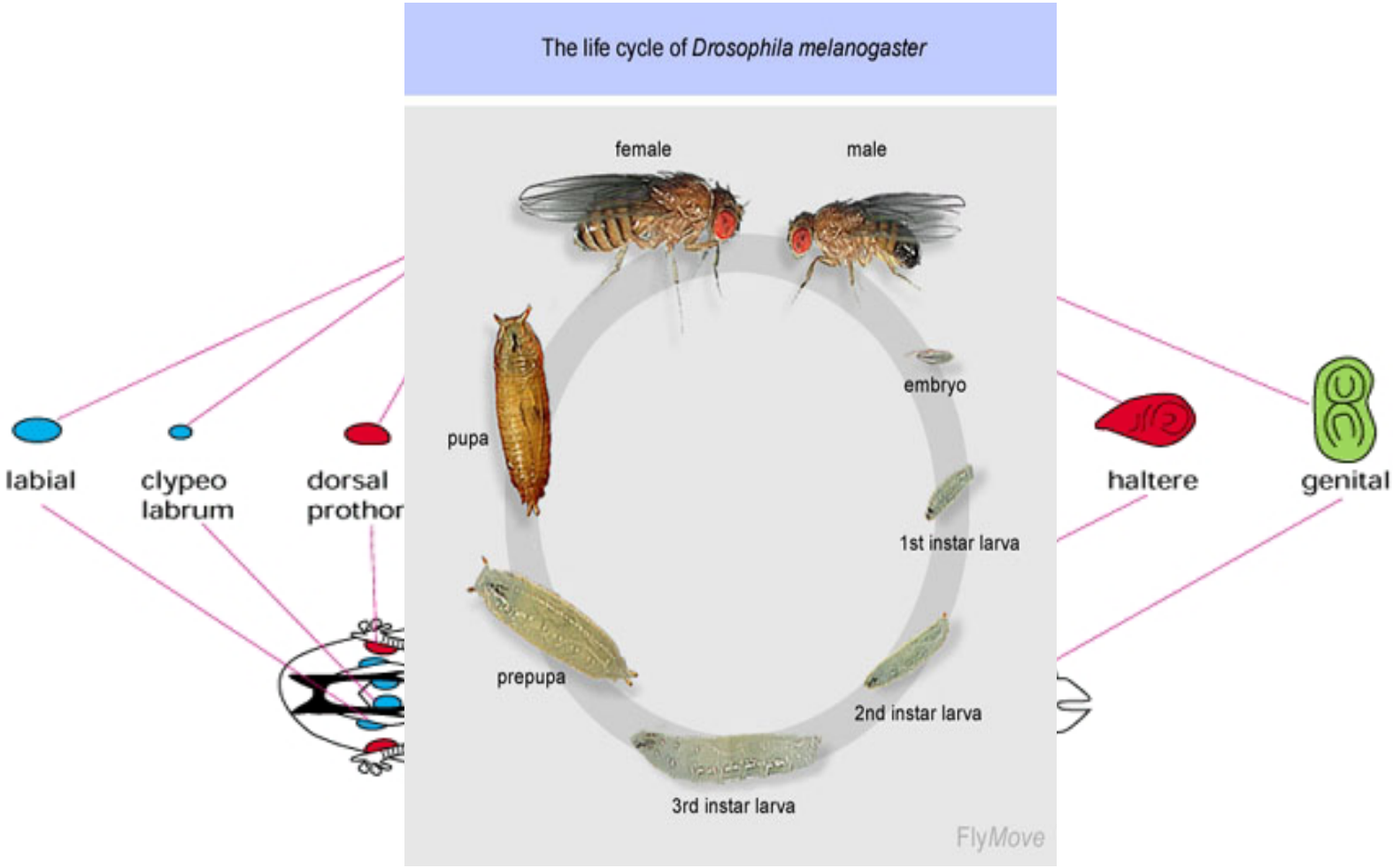
Dissecting fly wings to analyse wing hair patterns: knock-out and knock-in genes

Immuno-fluorescence imaging: imaging of a developing wing



Drosophila Development

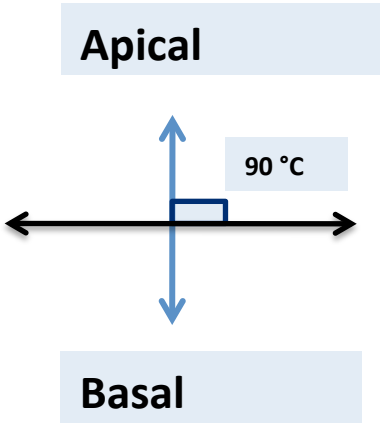
- **Patterning**: groups of cell aggregate and later become organs

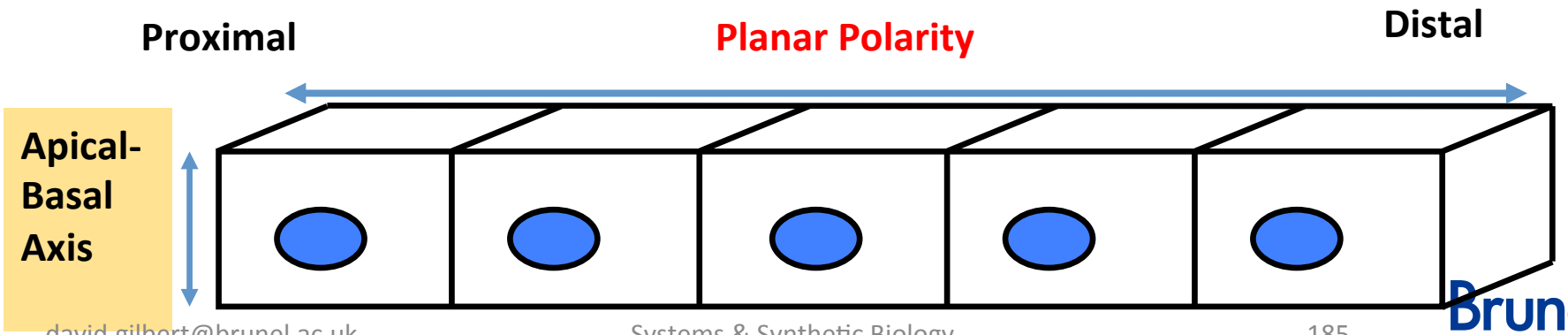


Planar Cell Polarity

Mechanism through which a number of tissue types determine the polarity of their cells perpendicular to their cellular apical-basal axis.

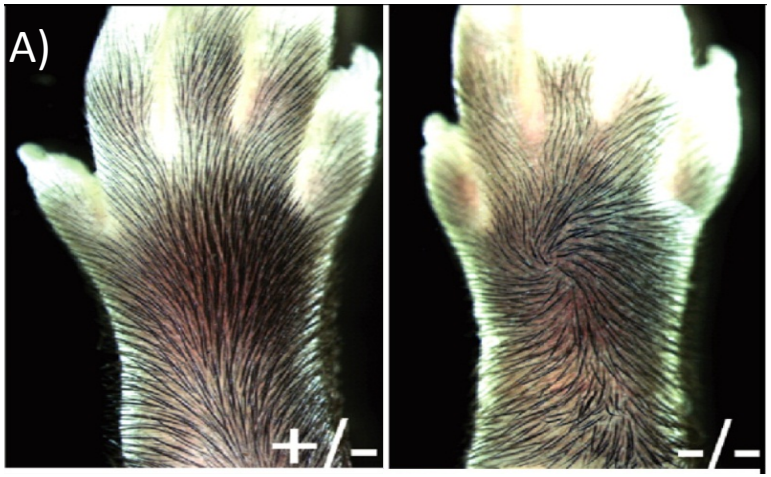
- Mammals: ear (sensory hair cell), eye (Equipotent R3/R4 cells) epithelia



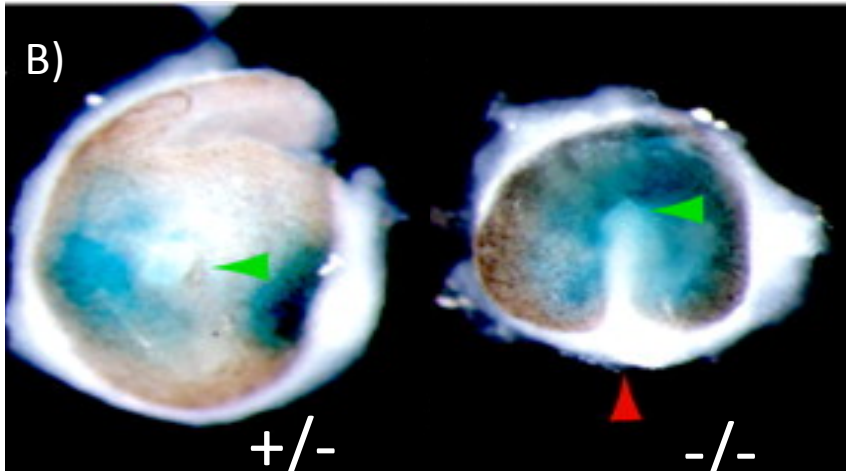


PCP: Vertebrate Phenotypes

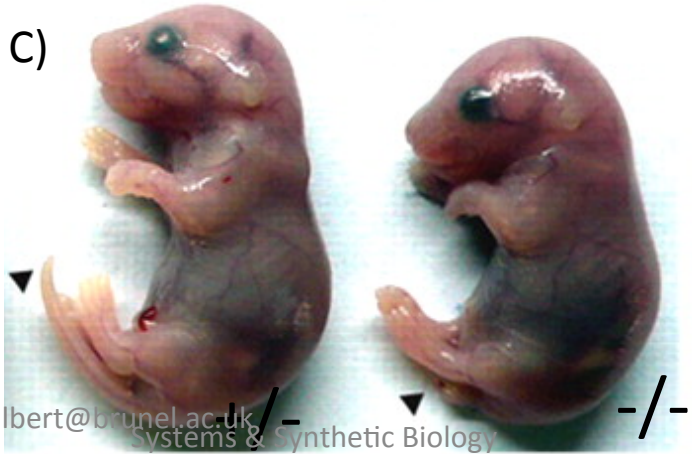
Fz6 Knockout: Hair Polarity



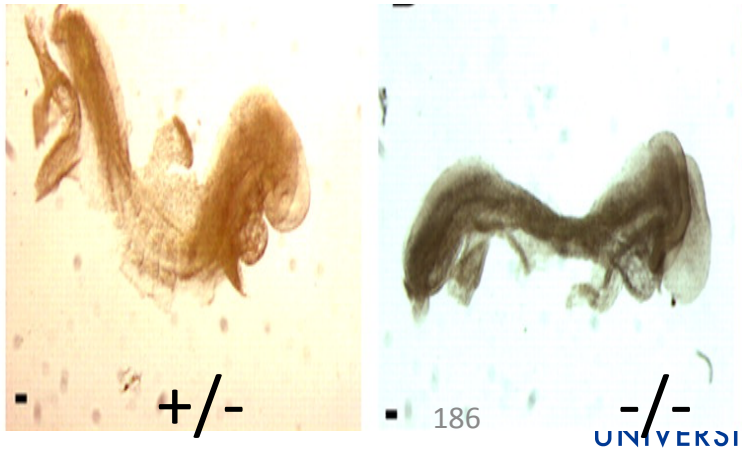
Fz5 Knockout: Coloboma



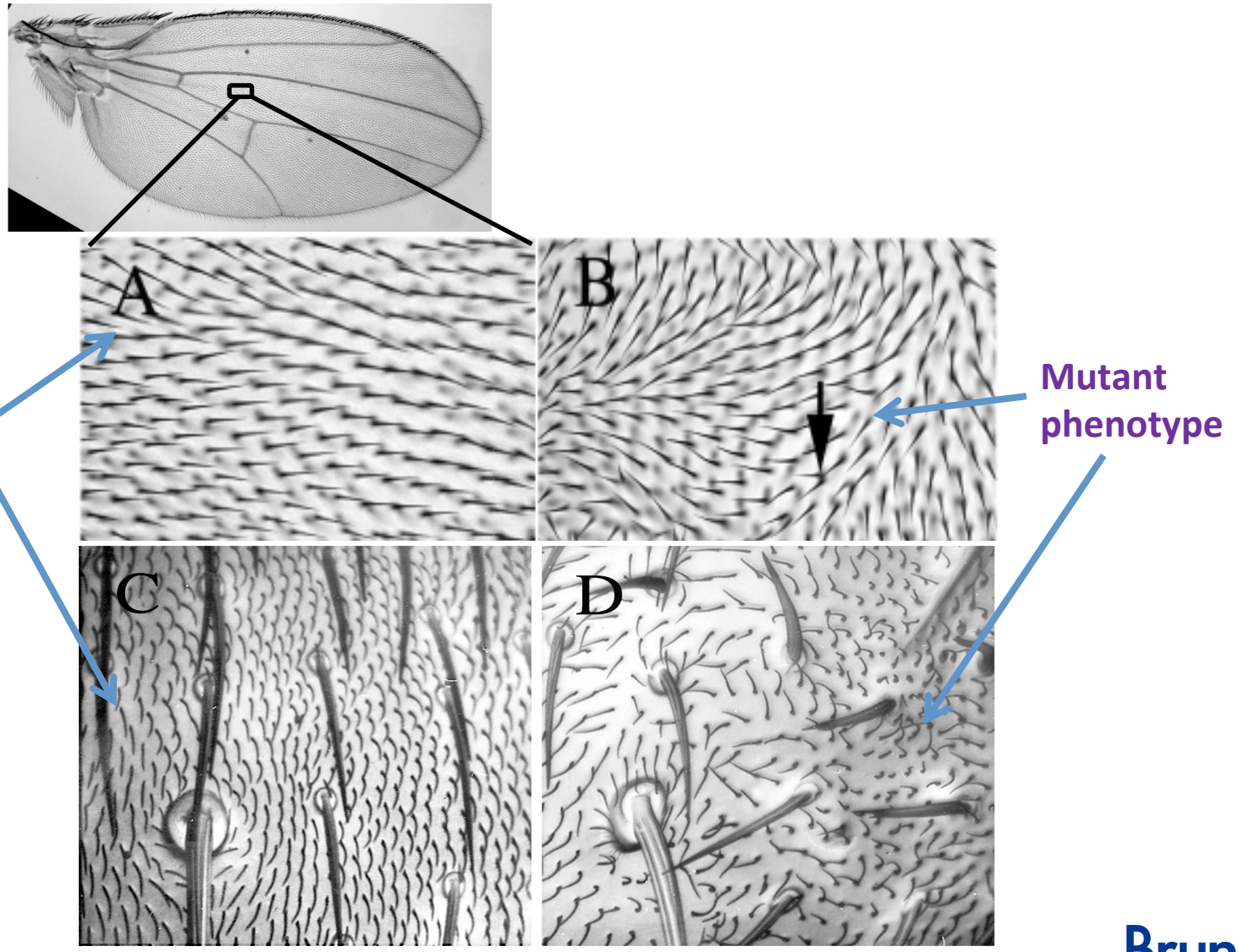
Dact1 Knockout: Urogenital Defects



Vangl2 Knockout: Impairs Cranial Neurulation

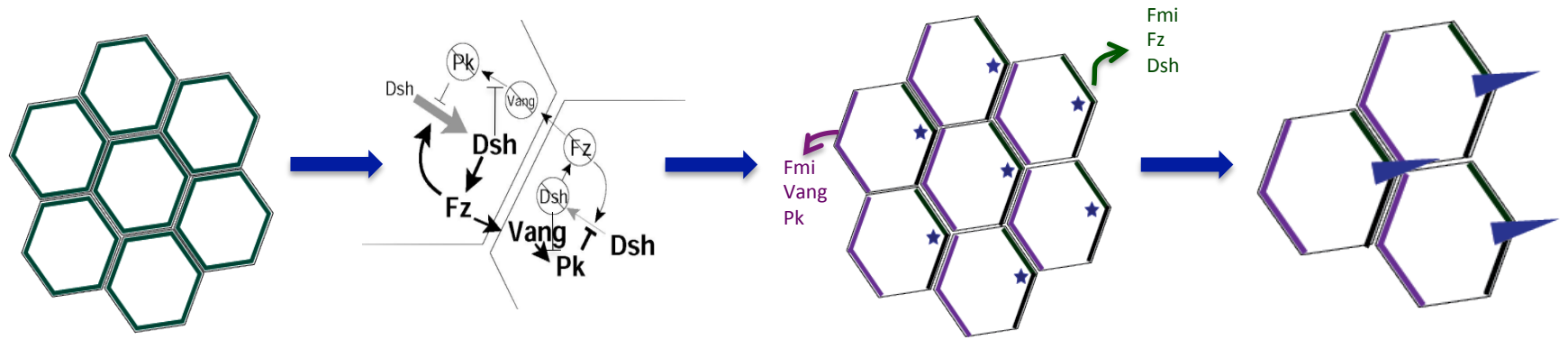


PCP: Drosophila Wing Phenotypes



Biological Model

- **PCP proteins** involved in the intercellular signalling: Flamingo (Fmi) , Frizzled (Fz), Dishevelled (Dsh), Prickle (Pk) and Van-Gogh (Vang) .
- **A core machinery** mediates a competition between the proximal and distal proteins: Frizzled (Fz), Dishevelled (Dsh), Prickle (Pk) and Van-Gogh (Vang). Flamingo (Fmi) localises at both distal and proximal edges.
- **Feedback loops**: cells tent to align cell polarity as asymmetric distribution.

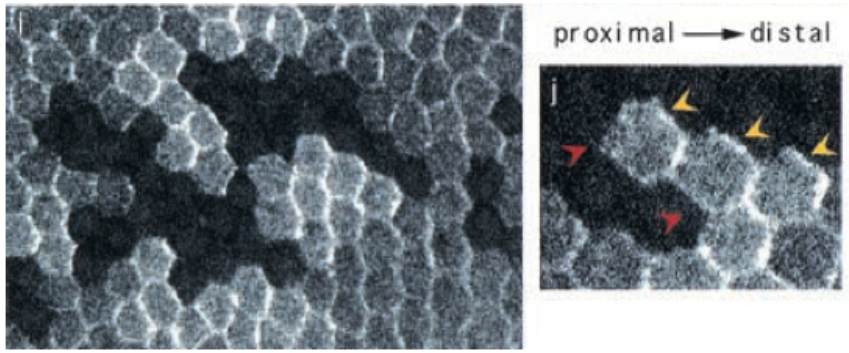


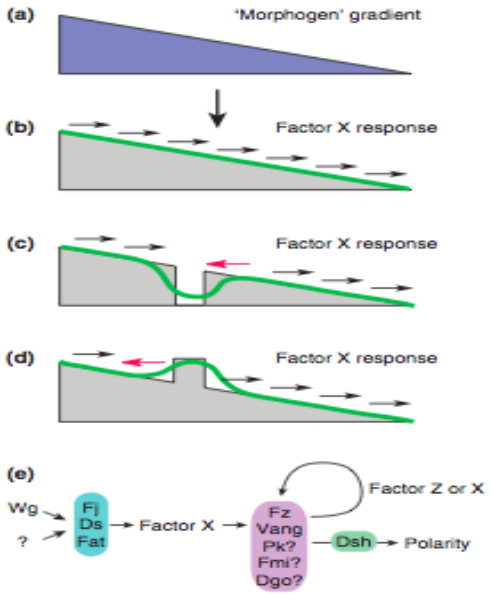
Symmetric distribution | Polarising signalling | Asymmetric distribution | Prehair formation

Hypotheses in PCP

- **Morphogen Factor X gradient**
  - Tissue-wide
  - Within cell

• J.D. Axelrod, C.J. Tomlin (2011). Modeling the control of planar cell polarity. WIREs Systems Biology and Medicine, p865





- **Microtubules**

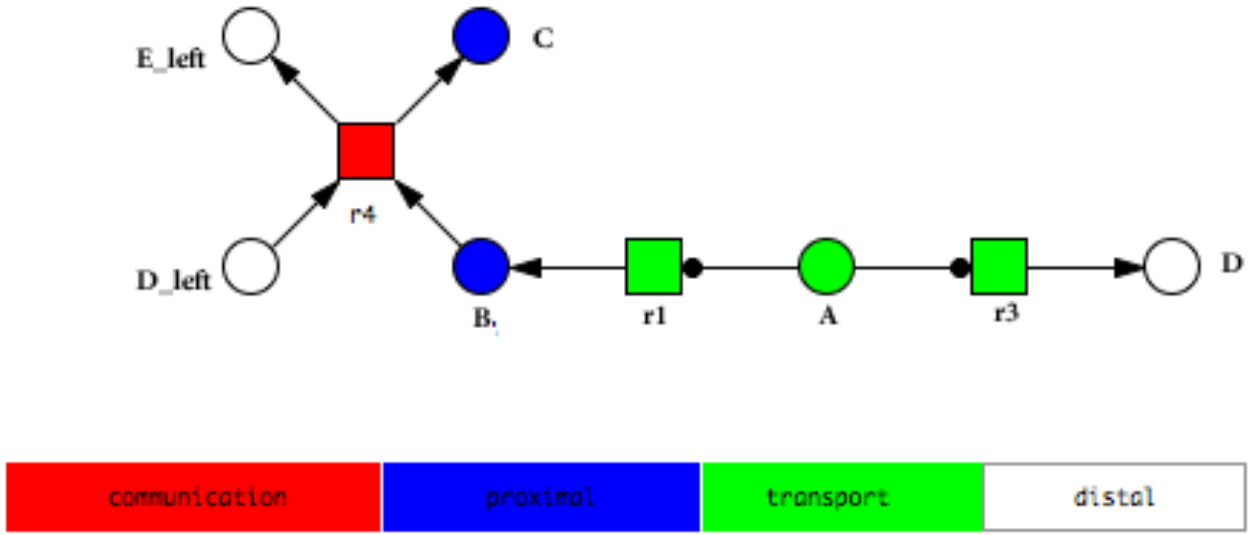
Directional transport of proteins

Issues

- Petri Nets and ODEs do not have specific facilities to create our multiscale descriptions of PCP.
  - Hard to scale up
  - Difficult to describe variants (mutations)

- Coloured Petri nets (CPN)
  - methodology for engineering robust multiscale models

Single cell
Abstract Level

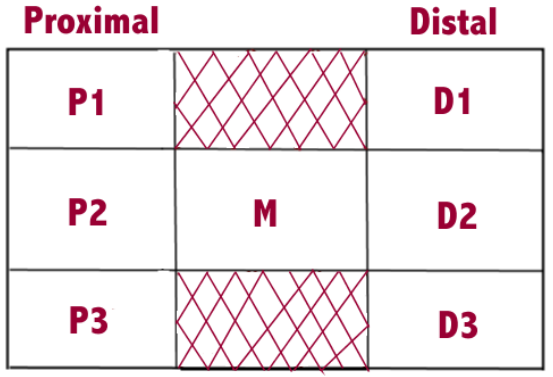


- Four spatial regions as labelled *(Labelled colours are **not** CPN colour sets)*
- D_left & E_left: two molecular species (places) from the left-hand side neighbouring cell(s)

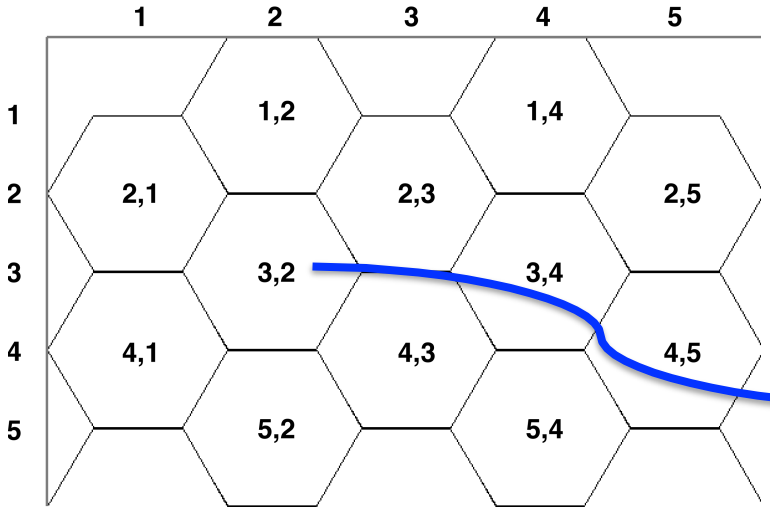
Hierarchical Organisation

- Hierarchically coloured

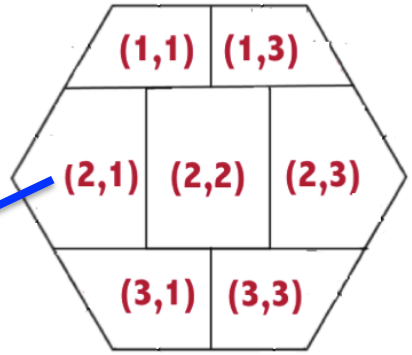
Intracellular compartments



Tissue (Cells)



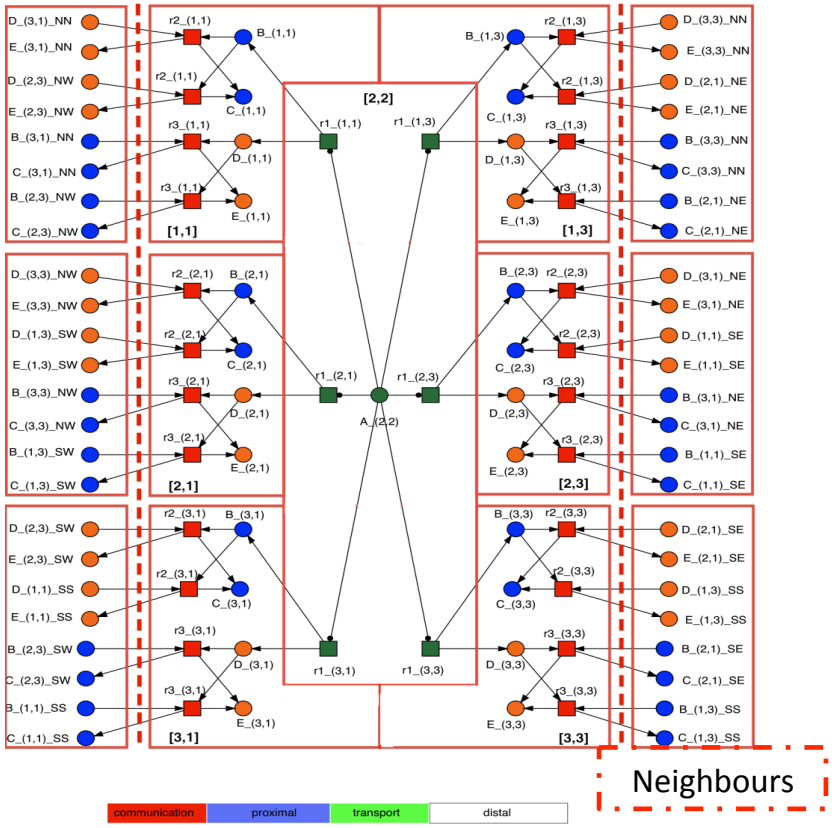
Cell: (3,2)
Compartment (2,1)



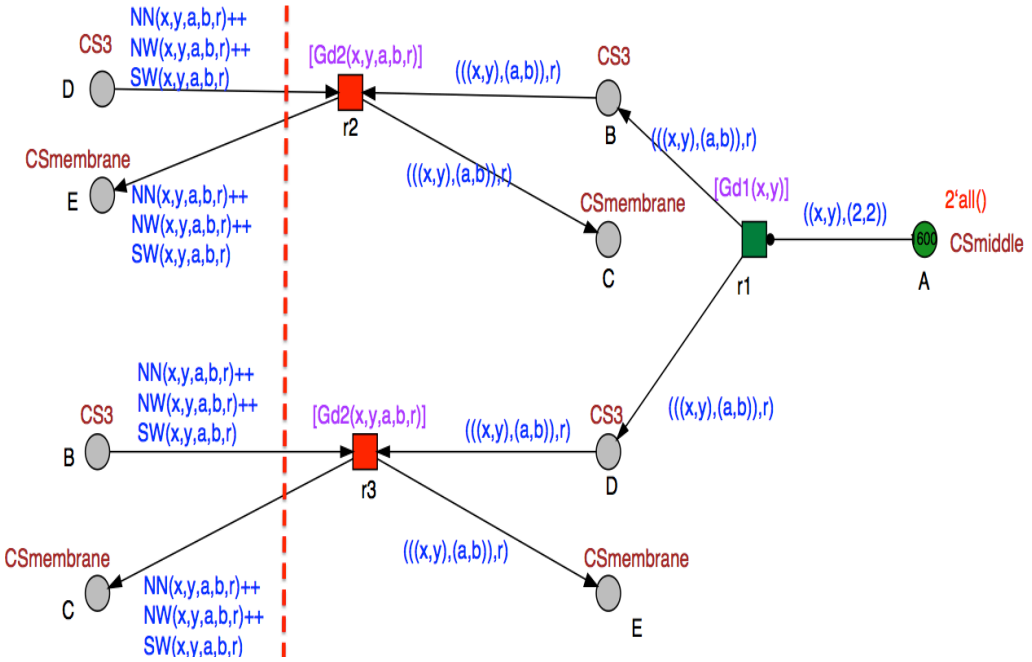
Colourset = {…, {((3,2)(1,1)), ((3,2)(2,1)), ((3,2)(3,1)),……((3,2)(3,3))}, …

Brunel UNIVERSITY LONDON

PN Model for Single Cell
Abstract level



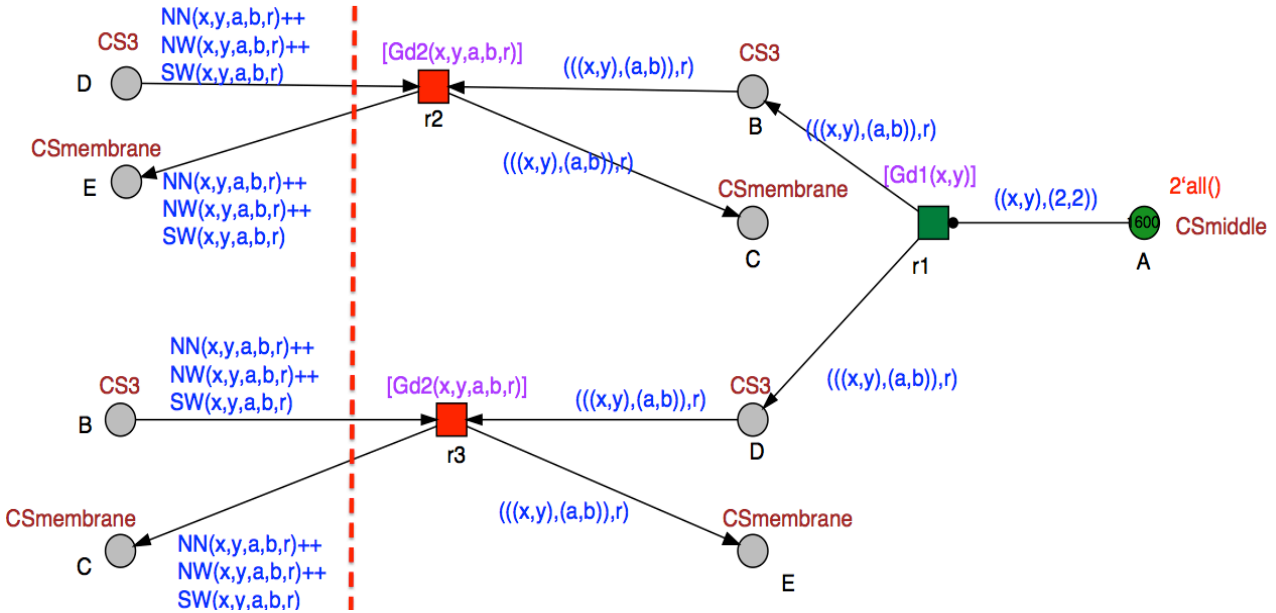
HCPN Model for Tissue
(Hierarchically Coloured Petri Net)
Abstract level



- 4 spatial regions: communication, proximal, transport and distal.
- Seven virtual compartments ((1, 1), (2, 1),..., (3,3)), denoted by ((x,y), (a,b)).
- Each place or transition belongs to a specific compartment.
- NN, NW and SW denote three left neighbours of the current cell.

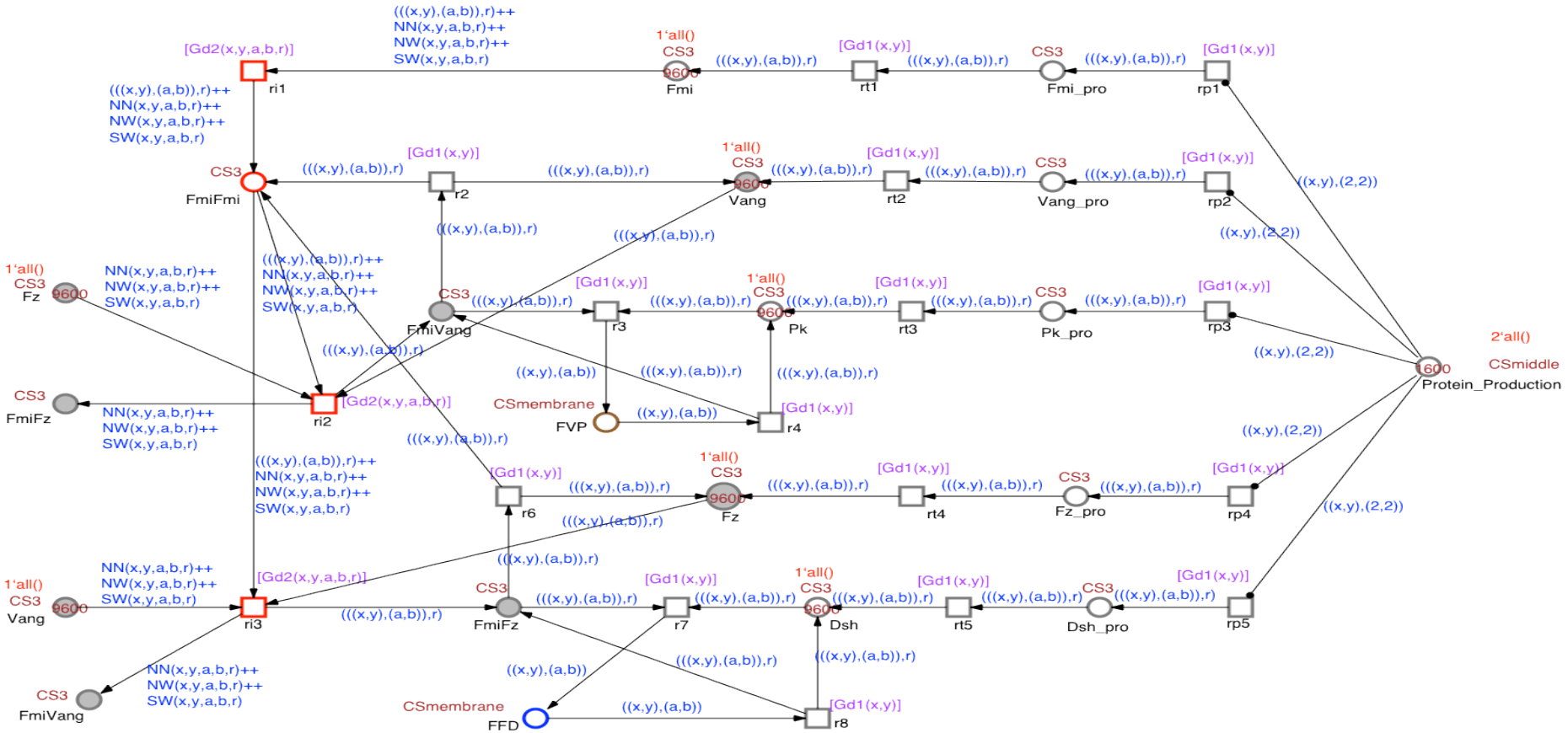
Brunel UNIVERSITY LONDON

CPN Model
Abstract Level



- 4 spatial regions: communication, proximal, transport and distal.
- Seven virtual compartments ((1, 1), (2, 1),..., (3,3)).
- Each place or transition belongs to a specific compartment.
- NW and SW denote two left neighbours of the current cell.

HCPN Model for Tissue
(Hierarchically Coloured Petri Net)
Detailed level



Brunel UNIVERSITY LONDON

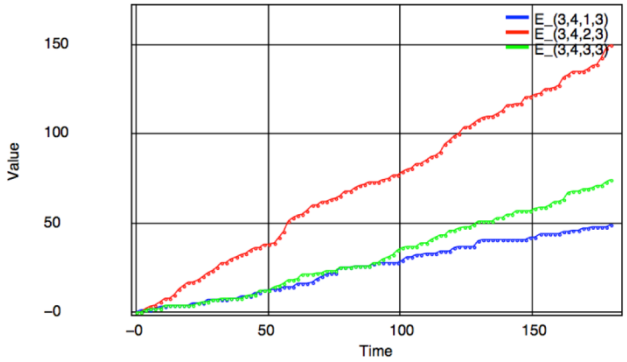
Declarations

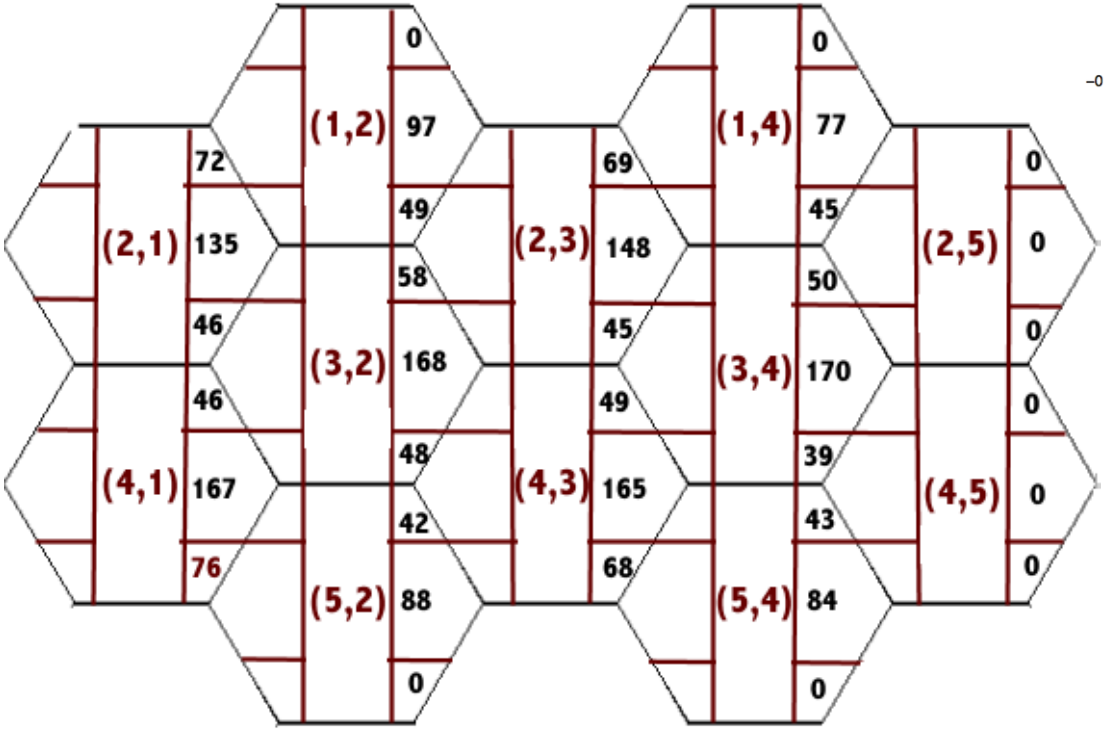
Declaration

Constant M = int with 5 ;
Constant N = int with 5;
Constant C = int with 3;
Constant R = int with 3;
colourset Row = int with $1 - M$;
colourset $Column$ = int with $1 - N$;
colourset $ComR$ = int with $1 - R$;
colourset $ComC$ = int with $1 - C$;
colourset $CSr4$ = enum with $c5, c6\_1, c6\_1, c7$;
colourset $CS1$ = product with $Row \times Column$;
colourset $CS2 = CS1$ with $x\%2 = 1\&y\%2 = 0|x\%2 = 0\&y\%2 = 1$;
colourset CS = product with $Row \times Column \times ComR \times ComC$;
colourset $CS4 = CS3$ with $x\%2 = 1\&y\%2 = 0|x\%2 = 0\&y\%2 = 1$;
colourset $CSdistal = CS4$ with $b = 3$;
colourset $CSproximal = CS4$ with $b = 1$;
colourset $CSmiddle = CS4$ with $b = 2$;
Variable $x : Row$;
Variable $y : Column$;
Variable $a : ComR$;
Variable $b : ComC$;
Variable $r4 : CSr4$;
Function $CSproximal\ NW(Row\ x, Column\ y, ComR\ a, ComC\ b)$;
Function $CSproximal\ SW(Row\ x, Column\ y, ComR\ a, ComC\ b)$;

Analysis Results

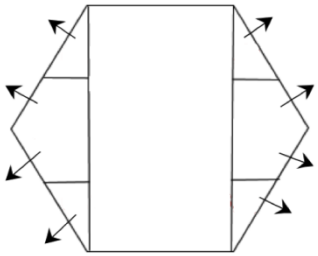
Stochastic simulation results





• Q. Gao, F. Liu, D. Tree & D. Gilbert. A Multi-cell Modelling Using Coloured Petri Nets Applied to Planar Cell Polarity. **BioPPN 2011**, Newcastle, UK.

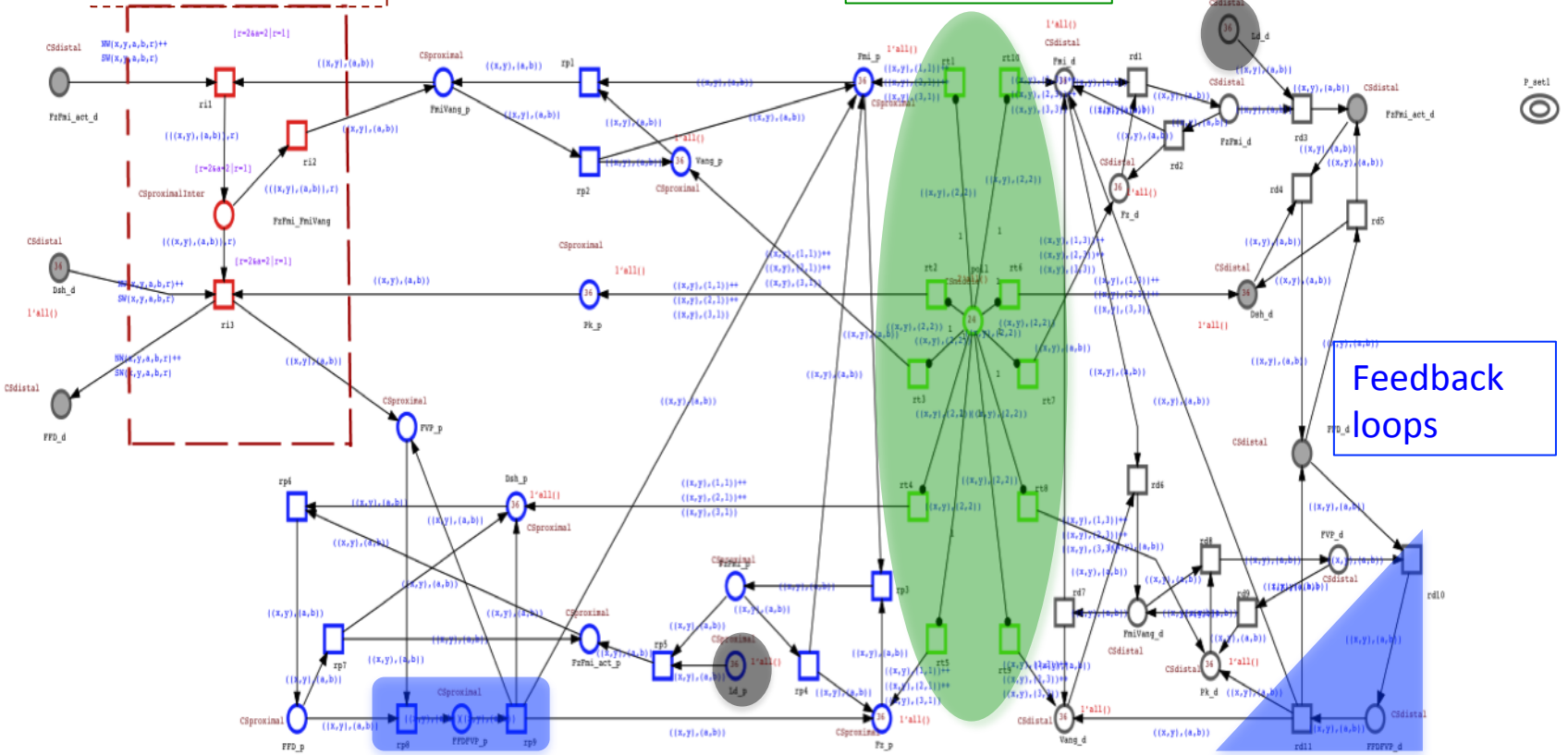
Initial CPN Model
PCP Signalling



Intercellular Communication

Microtubules

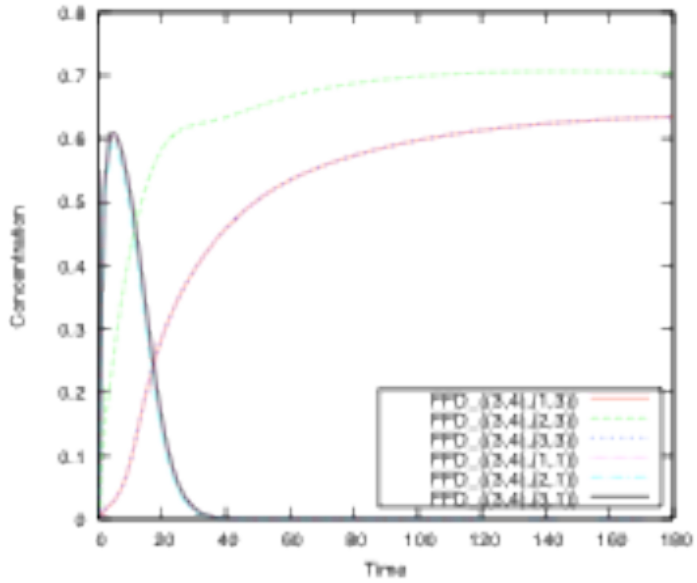
Factor X



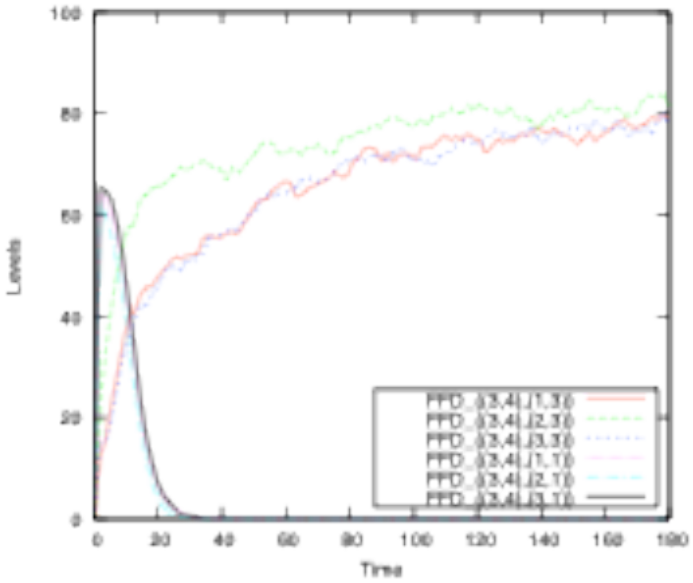
Feedback loops

Simulations

Continuous simulation

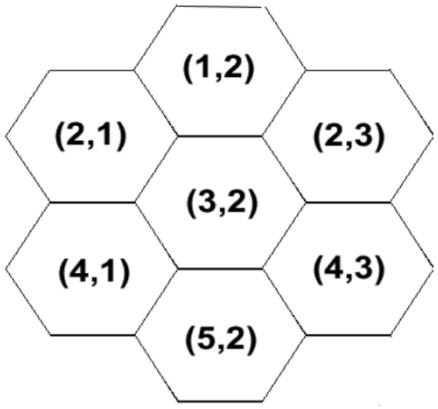


Stochastic simulation (average of 10 runs)



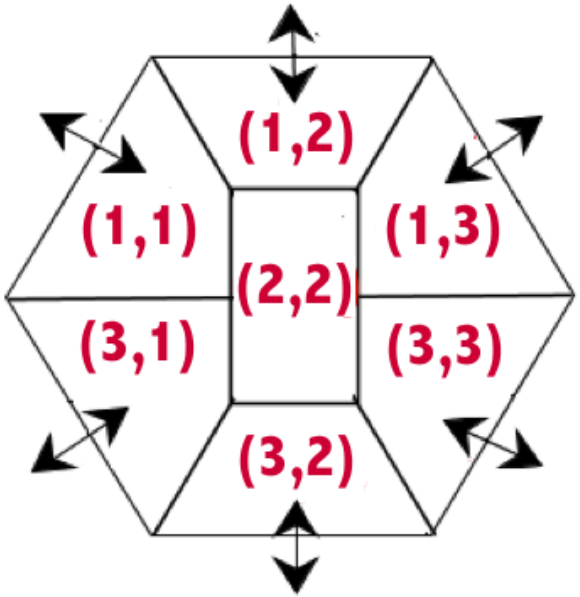
FFD accumulates at the distal edge of the cell rather than the proximal edge at the end of signalling.

Different Compartmentalisation

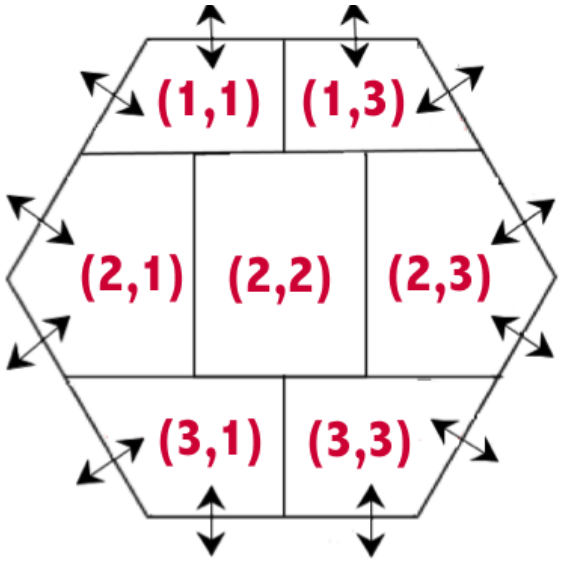


Symmetric Models

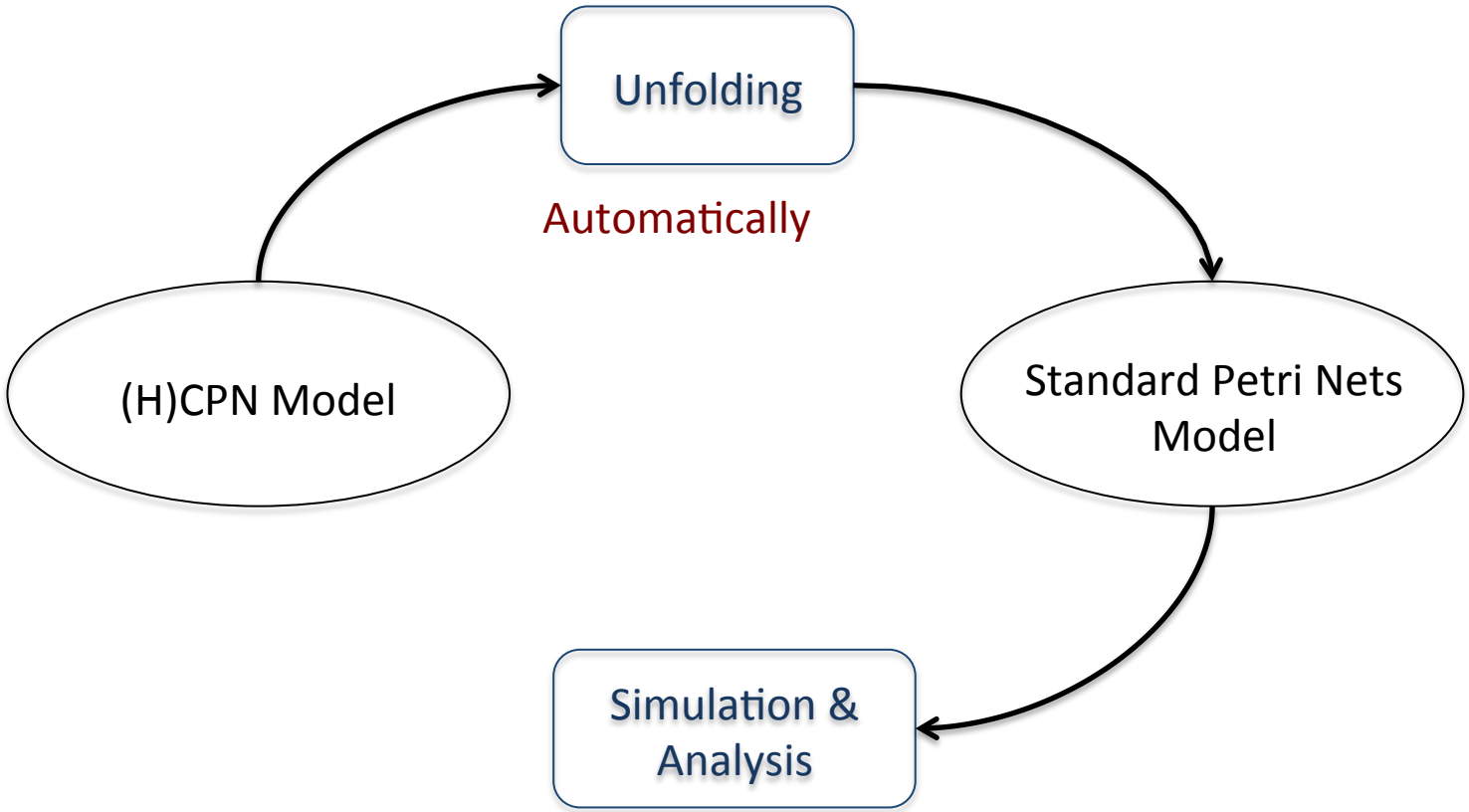
1:1



1:2



Unfolding, Simulation & Analysis (Snoopy)



Some Statistics

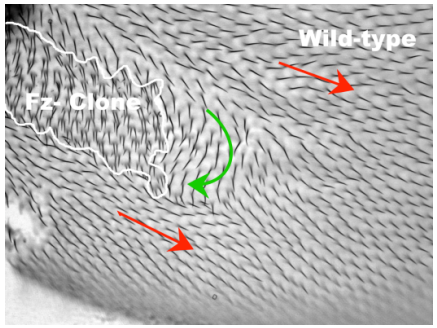
Unbiased PCP model size and runtime[a] for unfolding and continuous simulation over 1000 time units.

Size				Unfolding runtime (seconds)		Simulation runtime (seconds)
Grid($M \times N$)	Cells	Places	Transitions	Before optimisation	After optimisation	
5×5	12	2,028	2,802	3.195	1.154	3.145
10×10	50	8,450	11,826	9.714	2.613	14.618
15×15	112	18,928	26,622	22.771	4.495	42.586
20×20	200	33,800	47,646	44.818	9.231	88.886
40×40	800	135,200	191,286	280.598	83.162	371.647
40×40^b	800	164,000	229,686	329.384	120.186	7,399.544

[a] performed on a Mac Quad-core Intel Xeon, CPU 2× 2.26GHz, memory (DDR 3) 8 GB; [b] for the biased model BFXWt.

Constraint solver used for optimisation – enables larger size tissue to be simulated.

Modelling Mutant Clones

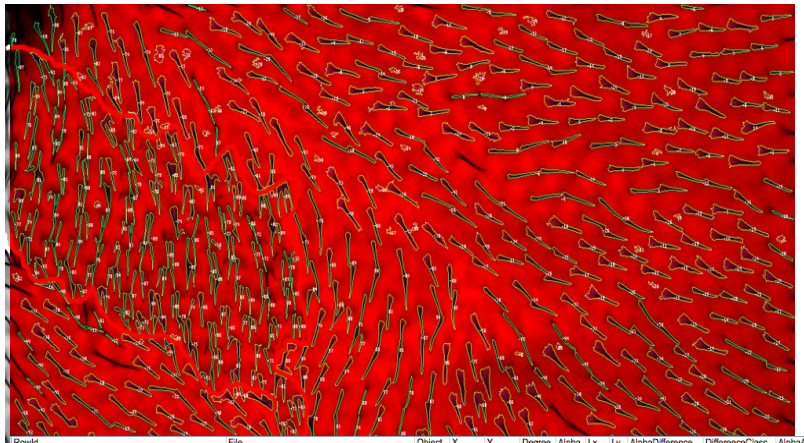


- Knock-out: cell clones in which a certain gene is knocked out are induced in the tissue (Biological experiments)
  - no corresponding protein produced.
- Petri nets: set the protein concentration zero
- CPN (repeat, with variations)
  - Big enough patch (tissue): 800 cells
  - Size / shape of clone: 80 cells (10% of the patch) in a circle-like shape

Models

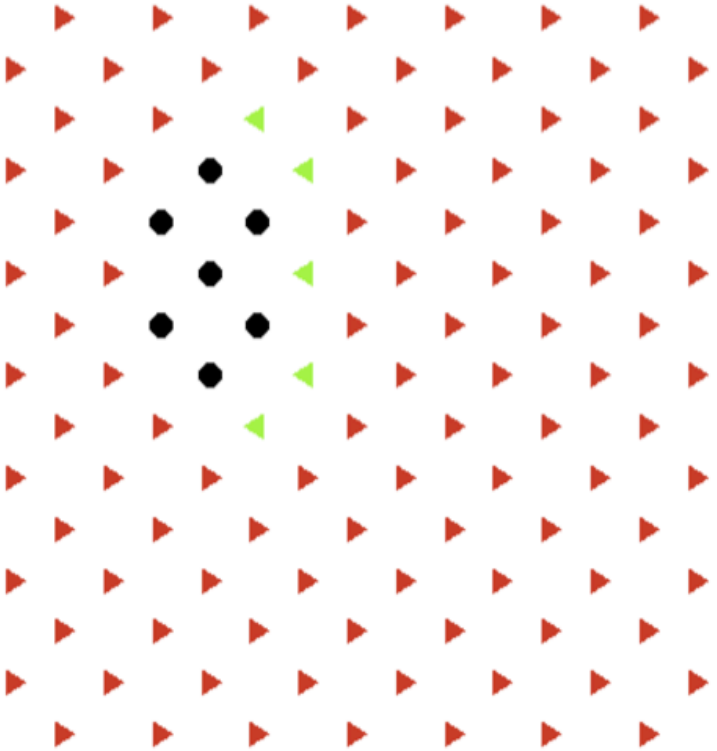
- Unbiased model: U [Wt]
  - wild type, no bias
- Biased models: B F [X] [T] [Wt]
  - Wild type
  - Always include feedback loops, and optionally either or both
    - Factor X
    - Directional transport
- Mutant models: B F [X] [T] [Fz|Vang]
  - One mutant clone induced to a biased model

Mutated tissue
Experimental vs In-silico



RowId	File	Object	X	Y	Degree	Alpha	Lx	Ly	AlphaDifference	DifferenceClass	AlphaAverage
dsh1-1001005_012.TIF_1	dsh1-1001005_012.TIF	1	23	1104	105	3.40	15	1	0.27	5	3.12
dsh1-1001005_012.TIF_2	dsh1-1001005_012.TIF	2	23	1469	28	2.06	15	1	1.08	2	3.12
dsh1-1001005_012.TIF_3	dsh1-1001005_012.TIF	3	24	822	113	3.55	15	1	0.43	5	3.12
dsh1-1001005_012.TIF_4	dsh1-1001005_012.TIF	4	25	537	158	4.33	15	1	1.21	2	3.12
dsh1-1001005_012.TIF_5	dsh1-1001005_012.TIF	5	25	1602	46	2.37	15	1	0.75	3	3.12
dsh1-1001005_012.TIF_6	dsh1-1001005_012.TIF	6	26	1479	29	2.08	15	1	1.04	2	3.12
dsh1-1001005_012.TIF_7	dsh1-1001005_012.TIF	7	28	1489	20	1.92	15	1	1.21	2	3.12
dsh1-1001005_012.TIF_8	dsh1-1001005_012.TIF	8	28	1545	177	4.65	15	1	1.53	0	3.12
dsh1-1001005_012.TIF_9	dsh1-1001005_012.TIF	9	30	1517	67	2.73	15	1	0.39	5	3.12
dsh1-1001005_012.TIF_10	dsh1-1001005_012.TIF	10	31	1635	11	1.76	15	1	1.36	1	3.12
dsh1-1001005_012.TIF_11	dsh1-1001005_012.TIF	11	32	1619	36	2.19	15	1	0.93	3	3.12
dsh1-1001005_012.TIF_12	dsh1-1001005_012.TIF	12	33	1426	59	2.60	15	1	0.52	4	3.12
dsh1-1001005_012.TIF_13	dsh1-1001005_012.TIF	13	35	1389	40	2.27	15	1	0.85	3	3.12
dsh1-1001005_012.TIF_14	dsh1-1001005_012.TIF	14	35	1563	82	3.01	15	1	0.12	6	3.12
dsh1-1001005_012.TIF_15	dsh1-1001005_012.TIF	15	35	1777	15	1.84	15	1	1.28	1	3.12
dsh1-1001005_012.TIF_16	dsh1-1001005_012.TIF	16	37	835	149	4.17	15	1	1.05	2	3.12
dsh1-1001005_012.TIF_17	dsh1-1001005_012.TIF	17	37	1446	68	2.75	15	1	0.37	5	3.12
dsh1-1001005_012.TIF_18	dsh1-1001005_012.TIF	18	37	1535	15	1.84	15	1	1.28	1	3.12
dsh1-1001005_012.TIF_19	dsh1-1001005_012.TIF	19	39	805	153	4.25	15	1	1.12	2	3.12
dsh1-1001005_012.TIF_20	dsh1-1001005_012.TIF	20	41	1464	50	2.45	15	1	0.68	4	3.12
dsh1-1001005_012.TIF_21	dsh1-1001005_012.TIF	21	44	992	123	3.72	15	1	0.60	4	3.12
dsh1-1001005_012.TIF_22	dsh1-1001005_012.TIF	22	44	1554	16	1.86	15	1	1.27	1	3.12
dsh1-1001005_012.TIF_23	dsh1-1001005_012.TIF	23	44	1608	19	1.90	15	1	1.23	1	3.12
dsh1-1001005_012.TIF_24	dsh1-1001005_012.TIF	24	46	1376	50	2.45	15	1	0.68	4	3.12
dsh1-1001005_012.TIF_25	dsh1-1001005_012.TIF	25	46	1484	53	2.50	15	1	0.62	4	3.12
dsh1-1001005_012.TIF_26	dsh1-1001005_012.TIF	26	47	1087	87	3.08	15	1	0.04	6	3.12
dsh1-1001005_012.TIF_27	dsh1-1001005_012.TIF	27	47	1626	178	4.67	15	1	1.55	0	3.12
dsh1-1001005_012.TIF_28	dsh1-1001005_012.TIF	28	49	846	111	3.51	15	1	0.39	5	3.12
dsh1-1001005_012.TIF_29	dsh1-1001005_012.TIF	29	49	1524	18	1.88	15	1	1.24	1	3.12
dsh1-1001005_012.TIF_30	dsh1-1001005_012.TIF	30	49	1645	40	2.27	15	1	0.86	3	3.12
dsh1-1001005_012.TIF_31	dsh1-1001005_012.TIF	31	50	1502	36	2.20	15	1	0.93	3	3.12
dsh1-1001005_012.TIF_32	dsh1-1001005_012.TIF	32	51	818	144	4.09	15	1	0.97	2	3.12
dsh1-1001005_012.TIF_33	dsh1-1001005_012.TIF	33	51	1164	26	2.02	15	1	1.11	2	3.12
dsh1-1001005_012.TIF_34	dsh1-1001005_012.TIF	34	51	1787	16	1.86	15	1	1.27	1	3.12
dsh1-1001005_012.TIF_35	dsh1-1001005_012.TIF	35	53	1009	59	2.60	15	1	0.52	4	3.12
dsh1-1001005_012.TIF_36	dsh1-1001005_012.TIF	36	53	1145	51	2.47	15	1	0.66	4	3.12
dsh1-1001005_012.TIF_37	dsh1-1001005_012.TIF	37	53	1431	75	2.89	15	1	0.24	6	3.12
dsh1-1001005_012.TIF_38	dsh1-1001005_012.TIF	38	53	1667	10	1.74	15	1	1.38	1	3.12
dsh1-1001005_012.TIF_39	dsh1-1001005_012.TIF	39	53	1683	10	1.74	15	1	1.38	1	3.12
dsh1-1001005_012.TIF_40	dsh1-1001005_012.TIF	40	54	1573	22	1.96	15	1	1.17	2	3.12
dsh1-1001005_012.TIF_41	dsh1-1001005_012.TIF	41	55	922	106	3.42	15	1	0.29	5	3.12
dsh1-1001005_012.TIF_42	dsh1-1001005_012.TIF	42	55	1472	60	2.62	15	1	0.51	4	3.12
dsh1-1001005_012.TIF_43	dsh1-1001005_012.TIF	43	57	1393	39	2.25	15	1	0.87	3	3.12
dsh1-1001005_012.TIF_44	dsh1-1001005_012.TIF	44	57	1451	69	2.77	15	1	0.35	5	3.12

FFD at distal vs FFD at proximal over Tissue

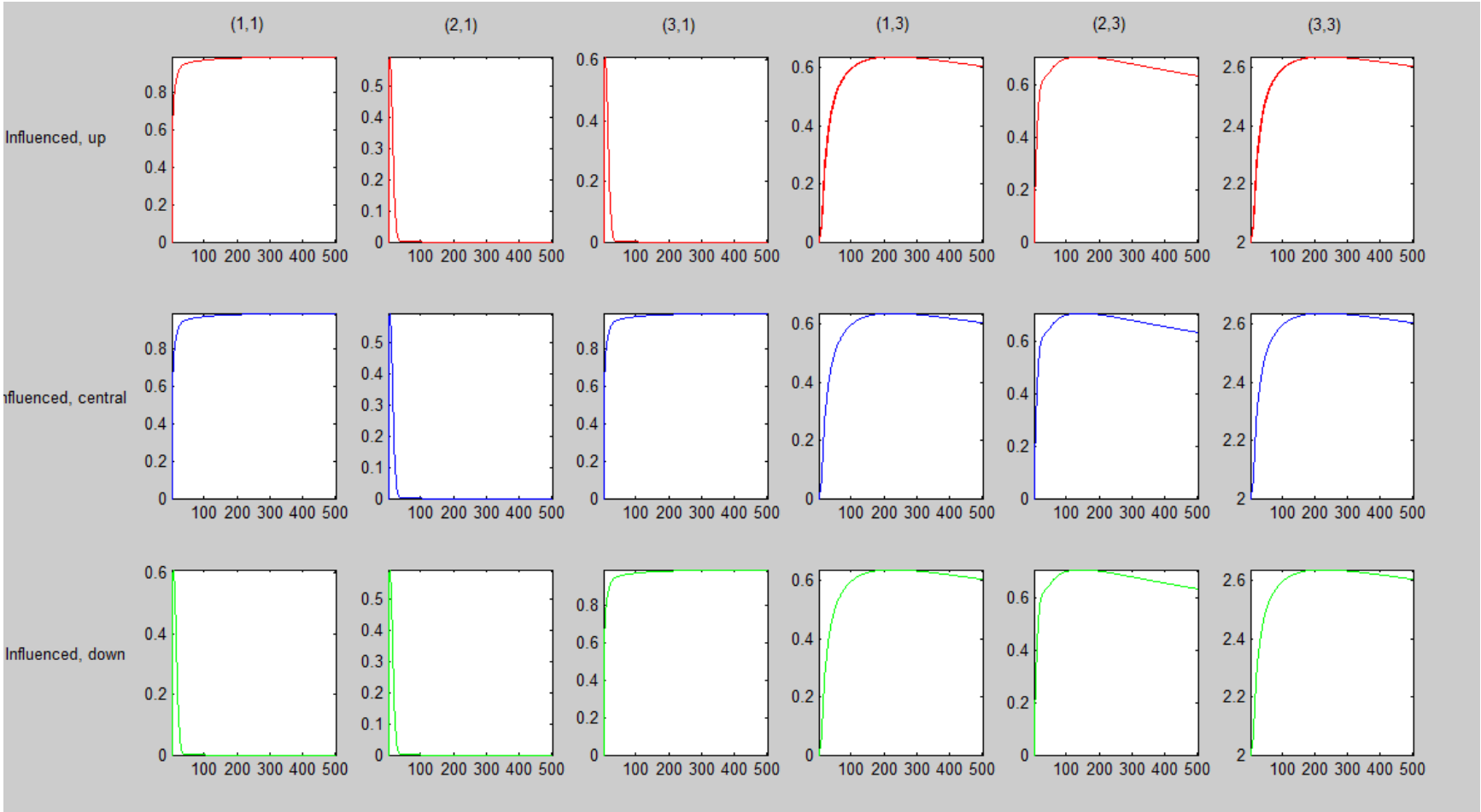


• Q. Gao, F. Liu, D. Gilbert, M. Heiner & D. Tree. **CMSB 2011**, Paris, France.

Analysis & Visualisation

- Clustering
  - DBScan
  - Hierarchical clustering
  - K-means
  - SOMs
- Model checking

Clustering of behaviours

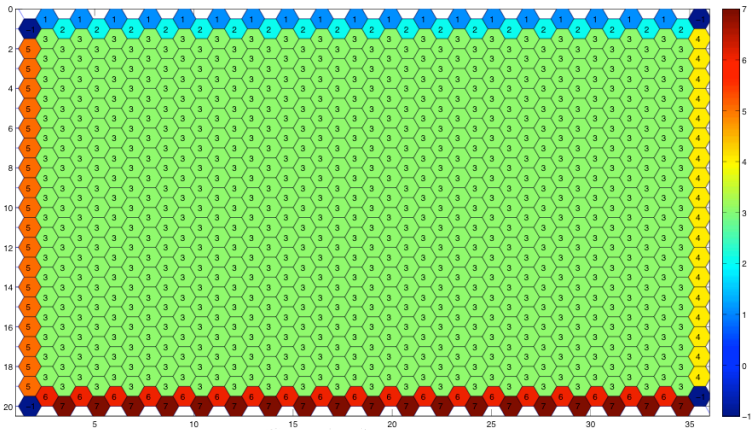


Brunel UNIVERSITY LONDON

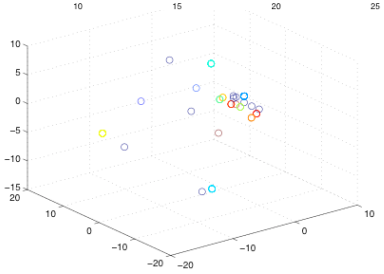
Clustering

- DBScan with Principal Component Analysis (PCA)

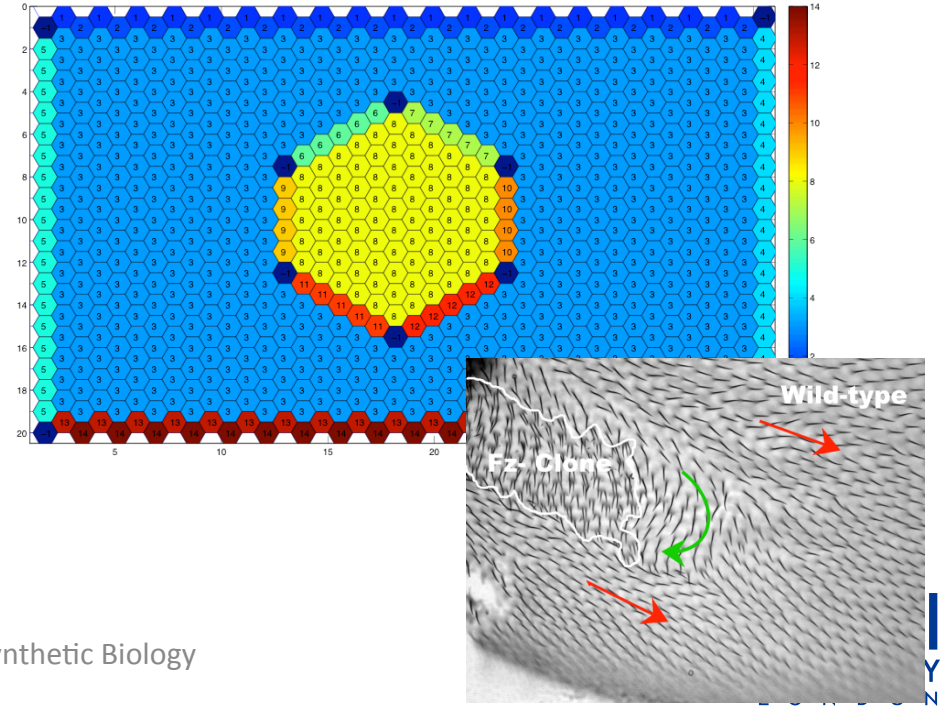
Unbiased model:
Grid 40*40 (800 cells)



Feature selection: PCA

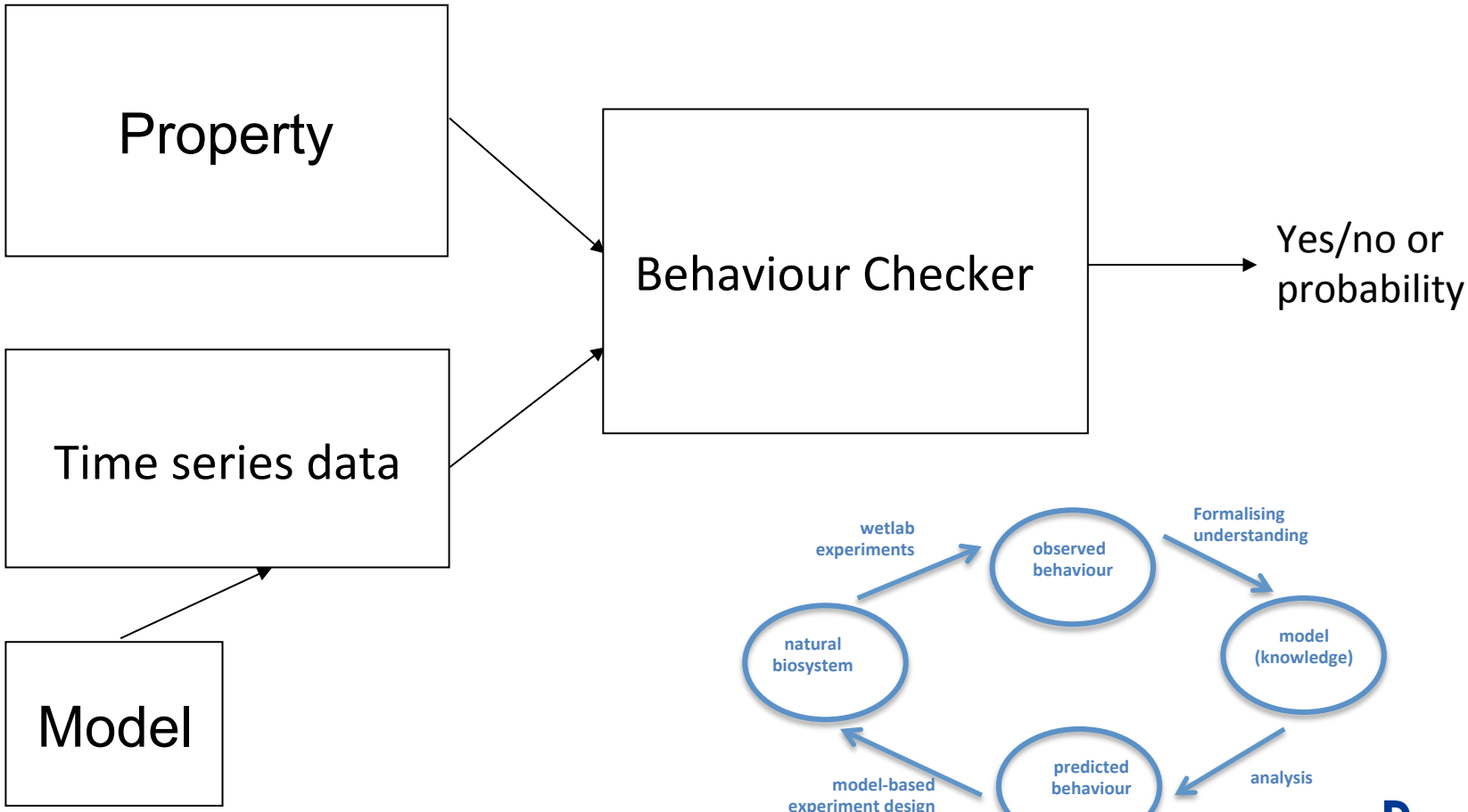


Fz- mutant clone model:
A patch of mutated cells lacking Frizzled (Fz) in a wild-type background



Simulation-based Model Checking

- PLTL using MC2 [Donaldson&Gilbert CMSB 2008]



Brunel UNIVERSITY LONDON

PLTL Language

- Behaviours to be checked against a model is expressed in temporal logic

- Probabilistic logic called Probabilistic Linear-time Temporal Logic (PLTL) – MC2 [Donaldson&Gilbert CMSB 2008]

- Main PLTL operators:

  G (P) – P always happens

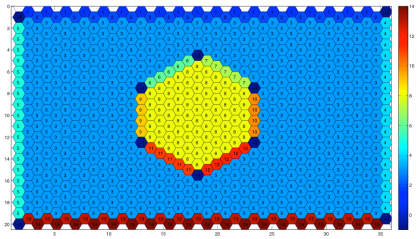
  F (P) – P happens at some time

  X (P) – P happens in the next time point

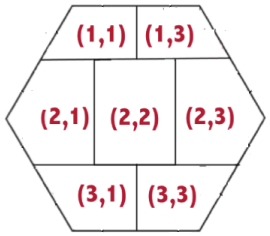
  (P1) U (P2) – P1 happens until P2 happens

  P1 { P2 } – P1 happens from the first time P2 happens

  time > ε – After a time point



Model Checking Primary data



Fz- mutant clone model

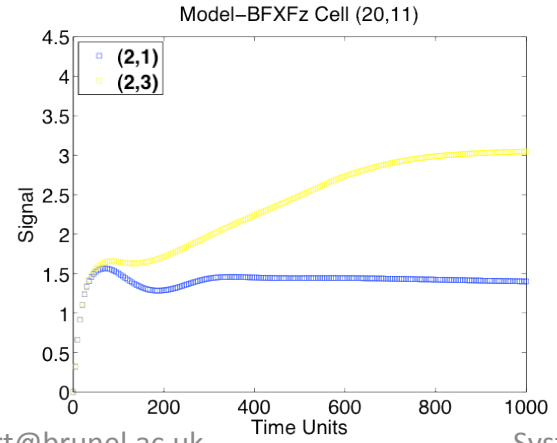
Unlike in the wild-type cells, for the **cells distally neighbouring to the Fz- clone** the concentration of FFD in the middle distal compartment is always lower than that of the middle proximal compartment:

P=? [time > 0 → G(D2 < P2)]

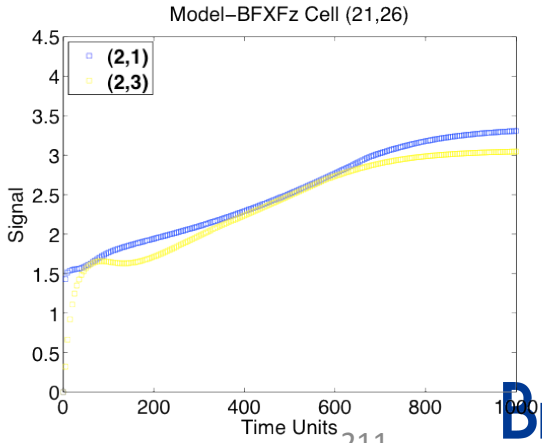
Moreover, the trace of D2 exhibits a peak followed by a trough, which is not true for P2:

P=?[F(d(D2) > 0 ∧ F(d(D2) < 0 ∧ F(d(D2) > 0)))]

Wild-type



Distally neighbouring to the clone

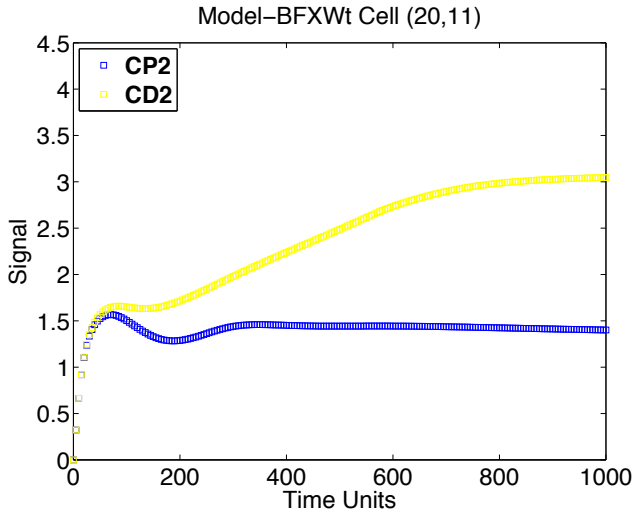


Brunel UNIVERSITY LONDON

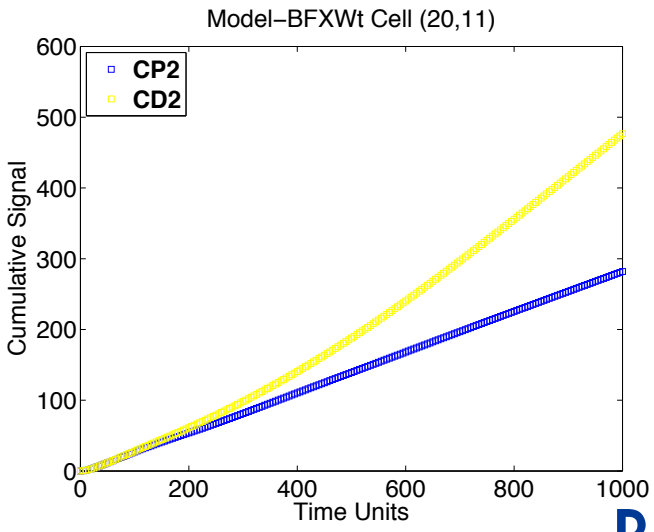
Big idea – check cumulative signal!

- Cumulative signal: time-series of accumulated concentrations of FFD (secondary data)
- Why?

  - The localisation of PCP signalling at any given time point is the result of the cumulative effect of the sum over the signalling events until that point.



Primary data



Secondary data

Brunel UNIVERSITY LONDON

Model Checking

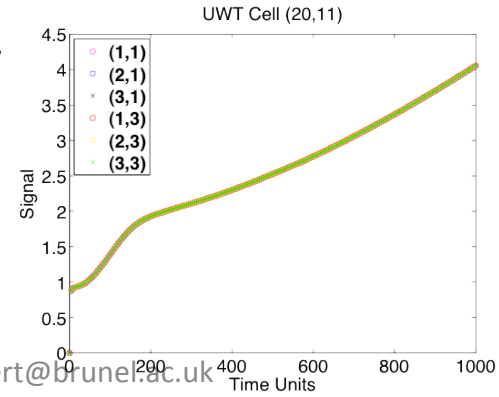
Unbiased model: reproduce the lack of polarisation, causing proteins to be symmetrically distributed along the cell membrane.

Query:

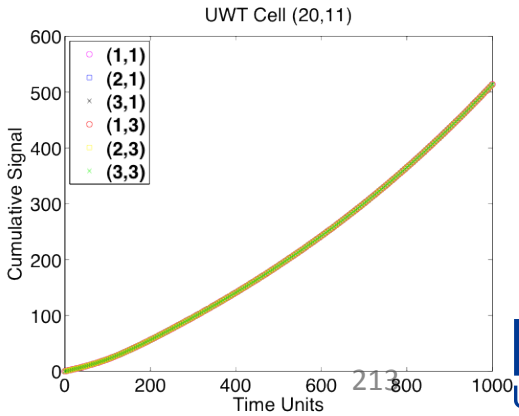
$$P=?[G(P1 = P2 \pm\delta \wedge P1 = P3 \pm\delta \wedge P1 = D1 \pm\delta \wedge P1 = D2 \pm\delta \wedge P1 = D3 \pm\delta)]$$

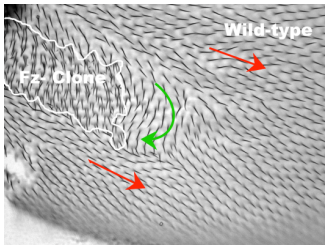
Where P1, P2, P3, D1, D2 and D3 denote the concentration of FFD in the proximal and distal compartments. Note that $P1 = P2 \pm\delta$ is a short-hand for $P1 < P2 + \delta \wedge P1 > P2 - \delta$.
Primary data set $\delta = 0.003$
Secondary data set $\delta = 0.04$ due to the cumulative nature of variations

Primary Data



Secondary Data

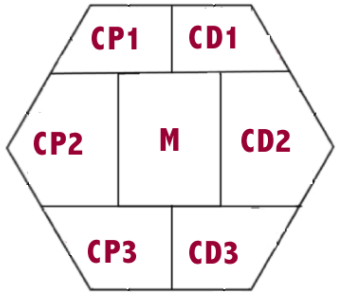




Model Checking

Secondary data

Fz- mutant clone model



Wild type cells in the tissue (i.e. away from the clone area).
After short initial period: Always middle distal cumulative[FFD] greater than middle proximal cumulative[FFD]

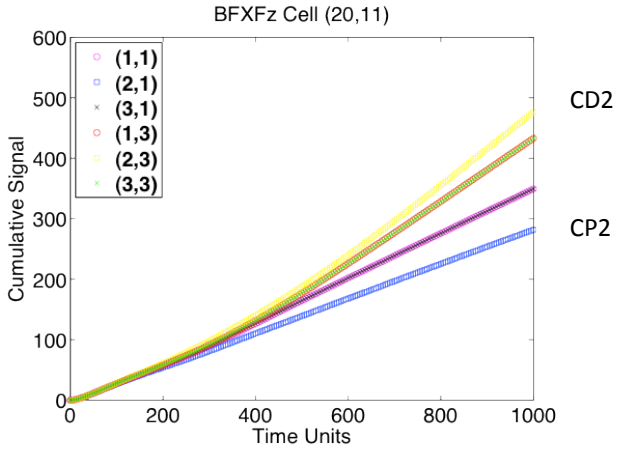
P=? [time > ε → G(CD2 > CP2)]

Wild type cells distally neighbouring to clone in the tissue
After short initial period: Always middle distal cumulative[FFD] less than middle proximal cumulative[FFD]

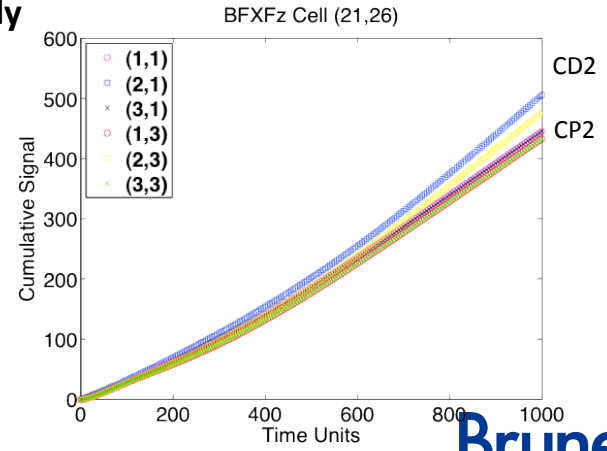
P=? [time > ε → G(CD2 < CP2)]

Hairs grow normally in wild-type, but disturbed in WT distally near clone, influence from the clone

Wild-type



Wild-type distally neighbouring to clone



Brunel UNIVERSITY LONDON

Acknowledgements – BioModel Engineering

Brunel

- Pam Gao
- Ovidiu Parvu
- David Tree
- Zujian Wu

Glasgow

- Robin Donaldson (MC2)
- Richard Orton
- Rainer Breitling

Cottbus

- Monika Heiner

Snoopy: www-dssz.informatik.tu-cottbus.de/DSSZ/Software/Snoopy

MC2 model checker: www.brc.dcs.gla.ac.uk/software/mc2

Brunel Centre for Systems & Synthetic Biology

Open positions at Brunel:

- Contact

david.gilbert@brunel.ac.uk

Brunel UNIVERSITY LONDON

References

Basic introduction

- (Basic)
  * W Marwan, C Rohr and M Heiner:
  Petri nets in Snoopy: A unifying framework for the graphical display, computational modelling, and simulation of bacterial regulatory networks. In Methods in Molecular Biology - Bacterial Molecular Networks, (Jv Helden, A Toussaint and D Thieffry, Eds.), Humana Press, pages 409-437, 2012
- (Advanced)
  ** MA Blätke, M Heiner, and W Marwan (2011),
  Tutorial - Petri Nets in Systems Biology. Otto von Guericke University Magdeburg, Magdeburg Centre for Systems Biology
- (Advanced)
  ** F Liu, M Heiner and C Rohr (2012):
  Manual for Colored Petri Nets in Snoopy, Technical report 02-12, Brandenburg University of Technology Cottbus, Department of Computer Science

Basic framework

- (For a beginner:)
  * Monika Heiner, Robin Donaldson and David Gilbert (2010): **Petri Nets for Systems Biology**; In
  Symbolic Systems Biology: Theory and Methods, (MS Iyengar, Ed.), Jones & Bartlett Learning, LCC, pages 61-97. [Preprint]
- (Formal definitions, more advanced examples:)
  ** Monika Heiner, David Gilbert, and Robin Donaldson (2008),
  Petri Nets for Systems and Synthetic Biology. In M Bernardo, P Degano, and G Zavattaro (Eds.): Formal Methods for Systems Biology SFM 2008, Springer LNCS 5016, pp. 215-264, 2008. [Preprint]
- (Application to biological assays)
  *** Monika Heiner, Sebastian Lehrack, David Gilbert and Wolfgang Marwan (2009),
  Extended Stochastic Petri Nets for Model-Based Design of Wetlab Experiments in C Priami, RJ Back, I Petre (eds.): Computational models for cell processes; Trans. on Computational Systems Biology XI, Springer LNCS 5750, pp. 138-163. [Preprint]

References

BioModel engineering

- (Simple introduction)
  * Rainer Breitling, Robin Donaldson, David Gilbert, Monika Heiner (2010):
  Biomodel Engineering - From Structure to Behavior; (position paper) in C Priami et al. (eds.): Trans. on Computational Systems Biology XII, Springer LNCS 5945, pp. 1-12. [Preprint].
- (More detailed treatment)
  ** Rainer Breitling, David Gilbert, Monika Heiner, Richard Orton (2008).
  A structured approach for the engineering of biochemical network models, illustrated for signalling pathways. Briefings in Bioinformatics, 2008 9(5):404-42, doi:10.1093/bib/bbn026 [preprint]
- (Case study)
  ** Zujian Wu, Shengxian Wang and David Gilbert. (2012). **A Hybrid Approach to Piece-wise Modelling of Biochemical Systems**
  PPSN 2012 (12th International Conference on Parallel Problem Solving From Nature - September 1-5, 2012 Taormina, Italy) [Preprint]
- (Case study)
  ** Robin Donaldson and David Gilbert (2008).
  A Model Checking Approach to the Parameter Estimation of Biochemical Pathways. In proceedings CMSB 2008 (Computational Methods in Systems Biology). Heiner, Monika; Uhrmacher, Adelinde M. (Eds.) LNCS 5307/2008, pp269-287. ISBN: 978-3-540-88561-0. [Preprint]

References

Multiscale modelling

- (Overview)
  * David Gilbert and Monika Heiner (2011), **Petri nets for multiscale Systems Biology**. Project summary (website).
- (Case study)
  ** Q. Gao, D. Gilbert, M. Heiner, F. Liu, D. Maccagnola and D. Tree, **Multiscale Modelling and Analysis of Planar Cell Polarity in the Drosophila Wing**, IEEE/ACM Transactions on Computational Biology and Bioinformatics, to appear. [Preprint]

Synthetic biology

- (iGEM example, simple treatment)
  * X.Gu, M.Trybiło, S.Ramsay, M.Jensen, R.Fulton, S. Rosser and D. Gilbert (2010),
  Engineering a novel self-powering electrochemical biosensor, in Journal for Systems and Synthetic Biology, Springer, Special Issue:Top-down and bottom-up approaches to the design of biological systems - papers arising from the Second European Conference on Synthetic Biology 1:4, 203-214, DOI: 10.1007/s11693-010-9063.
- (iGEM example, detailed treatment)
  ** David Gilbert, Monika Heiner, Susan Rosser, Rachael Fulton, Xu Gu and Maciej Trybiło (2008),
  A Case Study in Model-driven Synthetic Biology. In Biologically Inspired Cooperative Computing: BICC 2008. IFIP Springer ISSN 1571-5736 (Print) 1861-2288 (Online). Volume 268/2008. DOI 10.1007/978-0-387-09655-1. ISBN 978-0-387-09654-4. pp 163-175. [Preprint]

References

Further reading

- *** M Heiner and D Gilbert:
  How Might Petri Nets Enhance Your Systems Biology Toolkit; In Proc. PETRI NETS 2011, LNCS 6709/2011, Springer, pages 17-37, DOI: 10.1007/978-3-642-21834-7_2. [Preprint]
- *** D. Gilbert, M. Heiner, R. Breitling, R. Orton (2010).
  Computational modelling of kinase signalling cascades in MAP Kinase Signaling Protocols, Rony Seger (Ed), Methods in Molecular Biology, 2010, Volume 661, Part 4, 369-384, DOI: 10.1007/978-1-60761-795-2_22. [PubMed]

Wider reading

- (Overview)
  * David Gilbert, Alfonso Jaramillo, Natalio Krasnogor and Victor de Lorenzo (2010),
  Synthetic biology gains momentum in Europe, in Journal for Systems and Synthetic Biology, Springer, Special Issue:Top-down and bottom-up approaches to the design of biological systems - papers arising from the Second European Conference on Synthetic Biology 1:4, pp 145-147, DOI: 10.1007/s11693-010-9065-0
- (Link to volume with many examples of different modelling methodologies / examples)
  *** Corrado Priami, Rainer Breitling, David Gilbert, Monika Heiner, Adelinde M. Uhrmacher (Eds.) (2010).
  Transactions on Computational Systems Biology XII: Special Issue on Modeling Methodologies. Springer LNCS 5945, ISSN 0302-9743 (Print), 1611-3349 (Online).