
Fighting with complexity of biological systems

Andrei Zinovyev

Institut Curie - INSERM U900 / Mines ParisTech
Computational Systems Biology of Cancer



Zurich, 2004

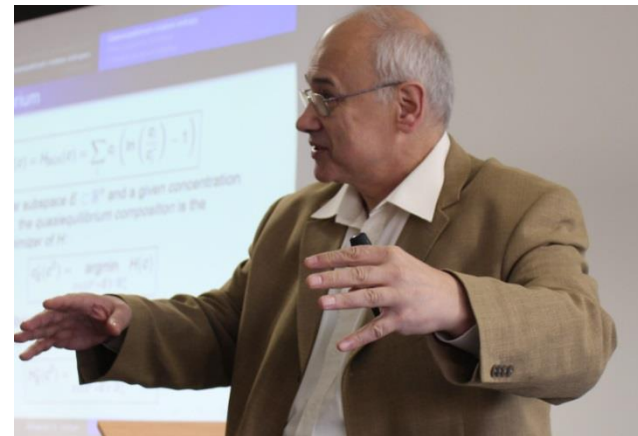
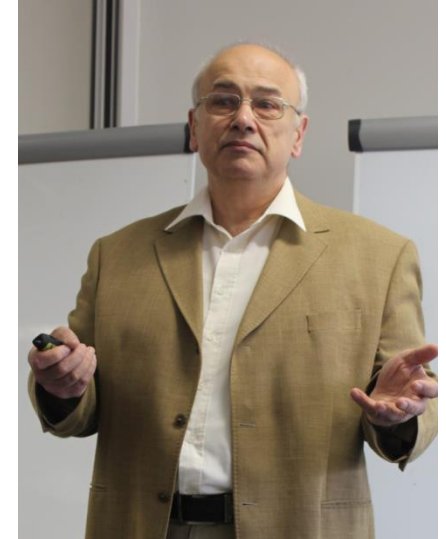


Ambleside, 2009

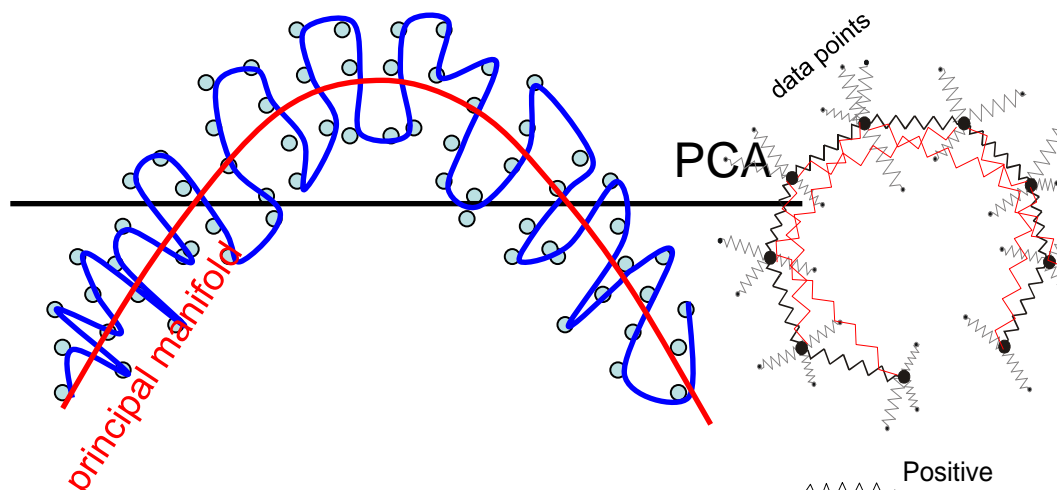


Leicester, 2014

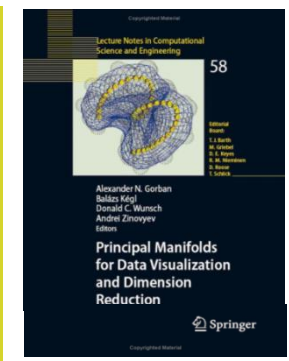
Scientific father



Non-linear principal manifolds and graphs

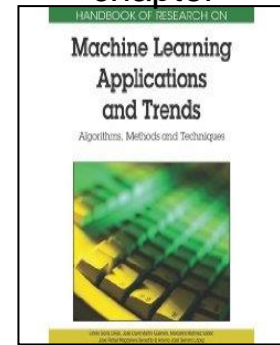


2000





2008

Principal graphs and manifolds chapter

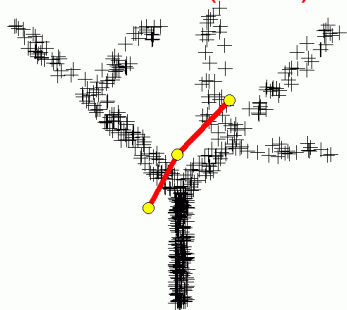


2010

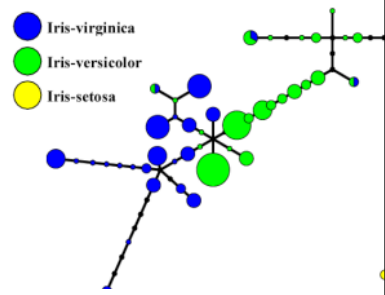
 Positive springs
 Negative springs

Two user-friendly software (VidaExpert and ViMiDa)

Principal harmonic dendrites (trees)



Metro maps data visualization



Data complexity measured by principal graphs

Andrei Zinovyev^{a,b,c,*}, Evgeny Mirkes^{d,e}

^a Institut Curie, rue d'Ulm 26, Paris, 75005, France

^b INSERM U900, Paris, France

^c Mines ParisTech, Fontainebleau, France

^d University of Leicester, University Road, Leicester, LE1 7RH, United Kingdom

^e Krasnoyarsk Institute of Railway Engineering, Krasnoyarsk-28, 660028, Russia

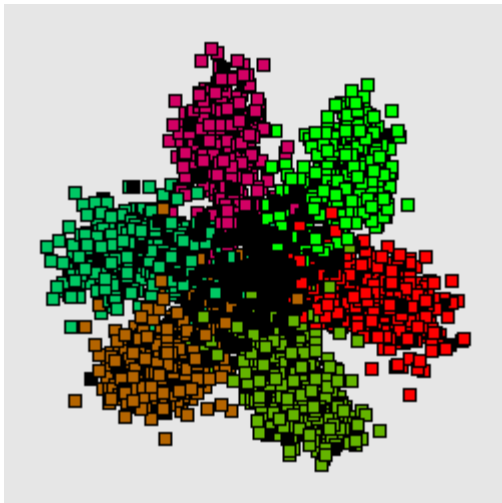
ARTICLE INFO

To our scientific father, Prof. Alexander N. Gorban, on his 60th birthday.

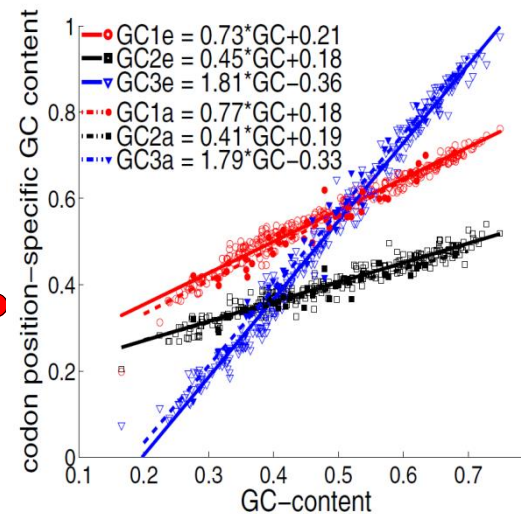
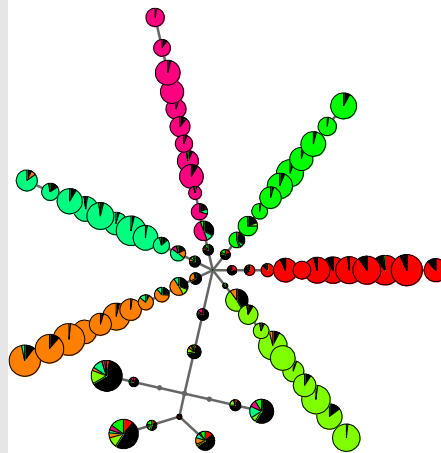
Keywords:

ABSTRACT

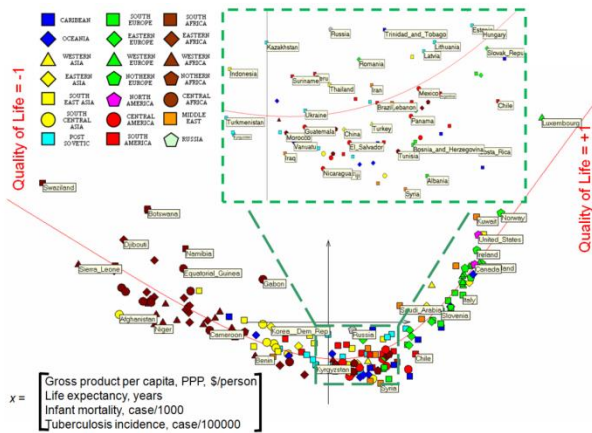
How to measure the complexity of a finite space? This is a non-trivial question which we suggest a set of data complexity measures. Principal cubic complexes datasets with non-trivial topologies. The ty



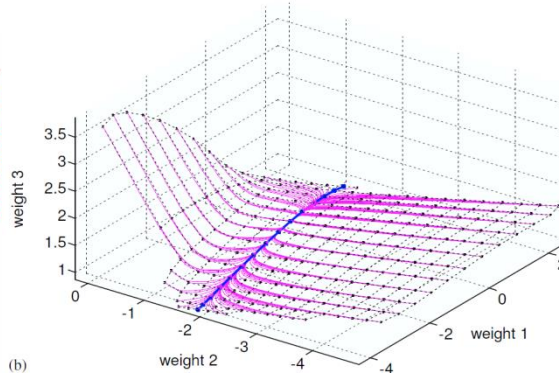
7-cluster structure of compact genomes
(Gorban et al, Physica A, 2007)



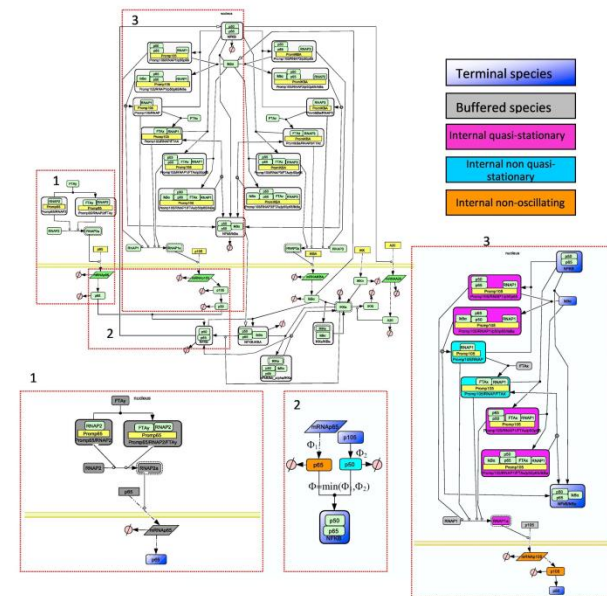
Mystery of 2 straight lines in bacterial genome statistics
(Gorban et al., Physica A, 2004)



Non-linear quality of life index
(Zinovyev and Gorban, Arxiv, 2010)



Method of invariant grids
(Gorban et al., Physica A, 2004)



Model reduction of biochemical networks
(Radulescu et al., BMC Sys Biol, 2008)

Fighting with complexity of biological systems

What is Complexity:

- 1) Big multidimensional data?
- 2) Difficulty with abstraction?
- 3) Non-linearity?
- 4) Emergence?
- 5) Limitation of human mind?

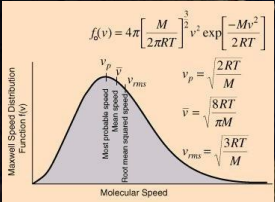
XXI century as the “century of complexity”

Stephen Hawking in his “millennium interview”:
“I think the next century will be the century of complexity”

Changing of era in science: from discovery of fundamental “scientific laws” to understanding/managing “complex systems”

“The complexity is recognized as the gap between the laws and the phenomena.”

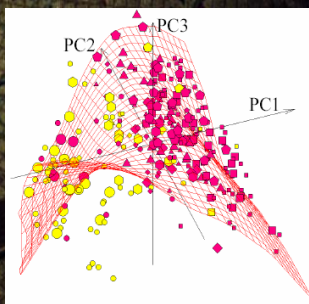
(Gorban, Grasping Complexity, *Comp&Math with Applications*, 2013)



Self-averaging complexity

Wild complexity

Reducible complexity



GORBAN'S MAP OF COMPLEXITY

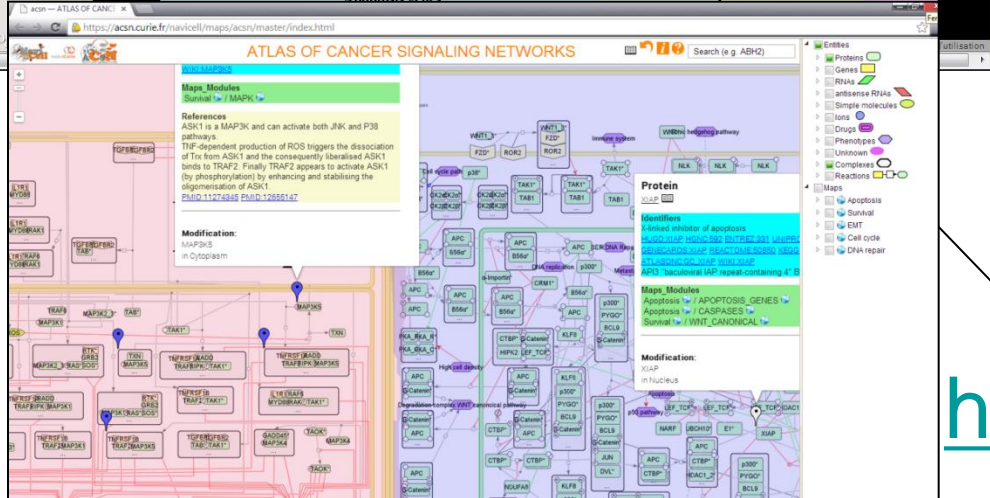
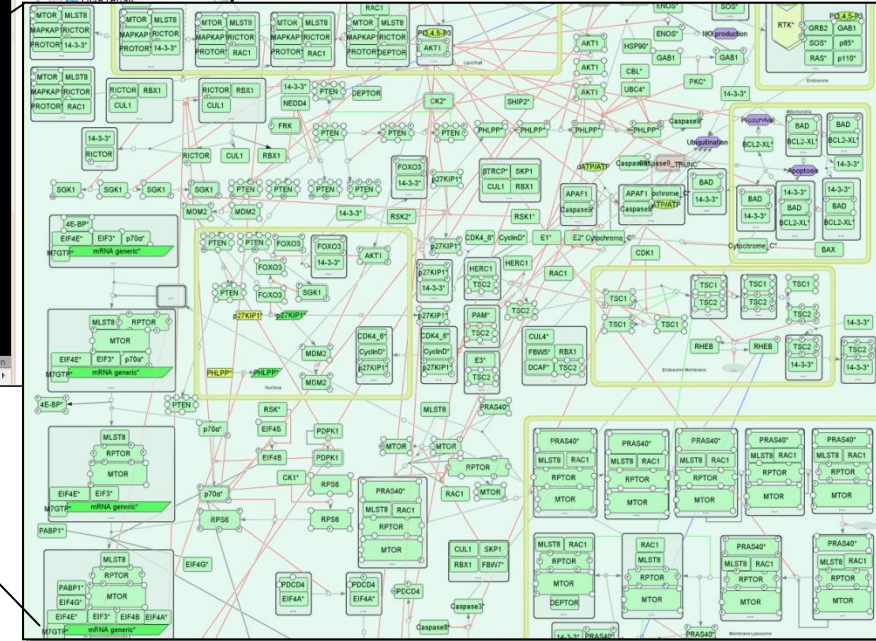
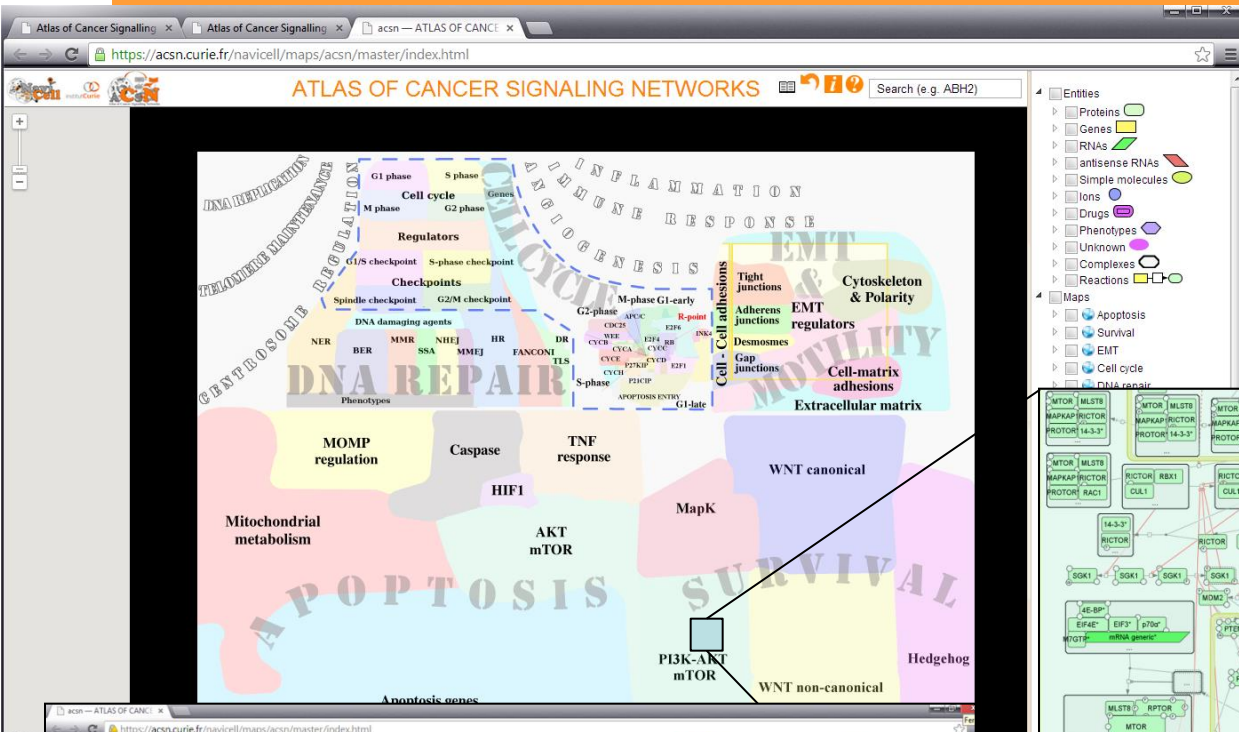
Gorban's map of complexity

- Imagine a set of points in multidimensional space
- Case of **reducible complexity**: there exists a low-dimensional object embedded in the space such that the most of points are located in its vicinity (*injection* of a low-dimensional object into multidimensional space).
- Case of **self-averaging complexity**: distributions of points *projected* into low dimensional spaces (screens) look very similar in almost all projections (concentration of measures phenomenon)
- **Wild complexity**: neither reducible nor self-averaging
- **Gorban's conjecture on brain functioning**: evolution pushes *brain functioning* to the Wild Complexity situation. If we observe that the brain functioning is reducible then it is not related to the main function of the brain but to something else (e.g., epilepsia crisis).
- Unlike brain, most of molecular mechanisms inside biological cells can be pushed by evolution to reducible complexity situation where they enjoy *robustness* and *controllability*.



Fighting with LARGE reaction networks: Atlas of Cancer Signaling Networks

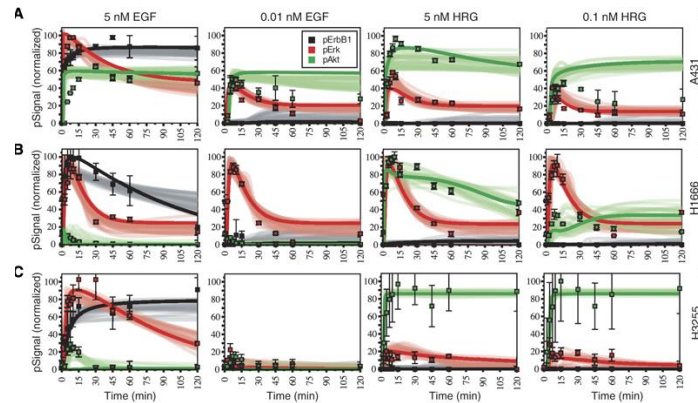
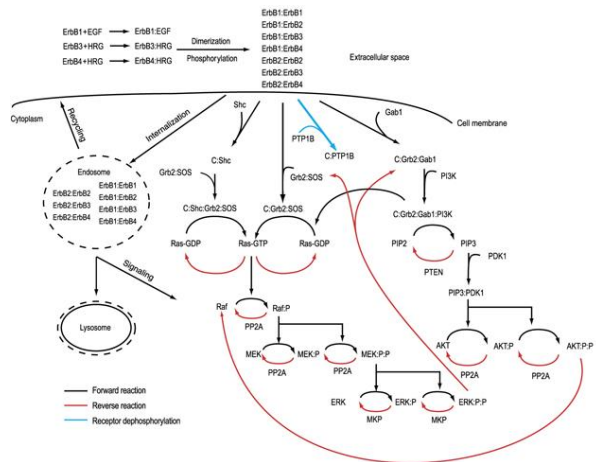
Google Maps
technology for
browsing LARGE
reaction network
(NaviCell tool,
Kuperstein et al, 2013)



<http://acsn.curie.fr>

Complex mathematical models of biological processes trained on experimental data, often have simple dynamics

From Chen et al., Input–output behavior of ErbB signaling pathways as revealed by a mass action model trained against dynamic data. 2009. *Molecular Systems Biology* 5:239.



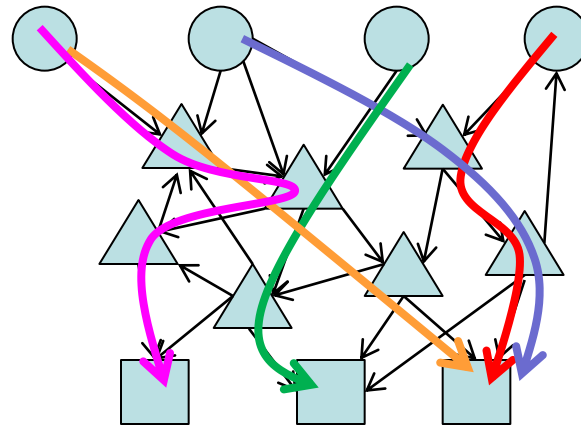
Model: 558 chemical species,
123 reactions, 478 equations

Data: Simple relaxation dynamics, which can be roughly characterized by the **steady state** value and the **relaxation time**

Why **biological** reaction networks are complex?

Hypothesis: biological networks are complex because they evolved to be able to respond to a variety of signals in a variety of conditions in a robust fashion, though each particular response can be simple and described by a small number of parameters

Then behavior of a complex network can be described as a “superimposition” of relatively simple models



Asymptotology of (bio)chemical reaction networks

(Gorban, Radulescu, Zinovyev, *Chem Eng Sci*, 2010)

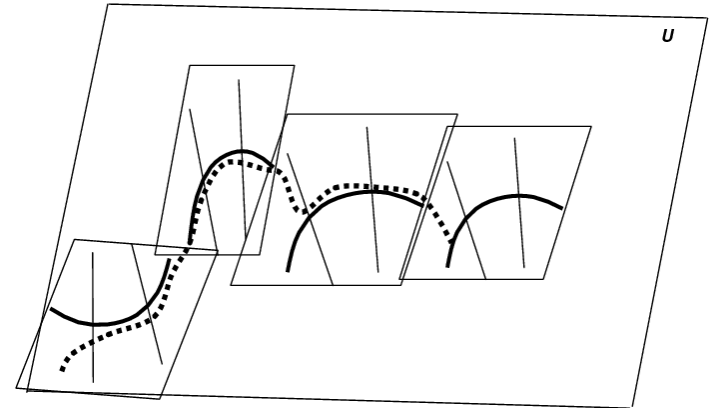
History of the term: accordingly to Martin Kruskal (1968), asymptotology is “*the science about the synthesis of simplicity and exactness by means of localization” or “the art of dealing with applied mathematical systems in limiting cases”*

Basic assumption: each particular behaviour of a biological system, represented by a mathematical model, corresponds to an asymptotic solution of this model, when some model quantities become small and can be neglected

Classical example: enzymatic catalysis $E+S \rightleftharpoons E:S \rightarrow E+P$, two asymptotic solutions: a quasi-steady one (Michaelis-Menten asymptotics) and a fast complex formation one (quasi-equilibrium asymptotics)

Asymptotology approach for dissecting complexity of mathematical models in biology

1) **For a given set of parameters,** decompose the complex dynamics into periods (epochs), each of which can be described in relatively simple terms (asymptotically)



2) **For a permissible set of parameters,** decompose the systems' phase space into regions, each of which is characterized by definite asymptotic properties



Five main tools for deriving asymptotics in chemical kinetics

- Quasi-equilibrium (fast reactions) approximation
- Quasy-steady state (fast intermediate chemical species) approximation
- Averaging (oscillating systems with small parameter)
- Lumping (aggregation of chemical species)
- **Rate limiting step and dominant systems**

Dominant systems and limitation in complex reaction networks

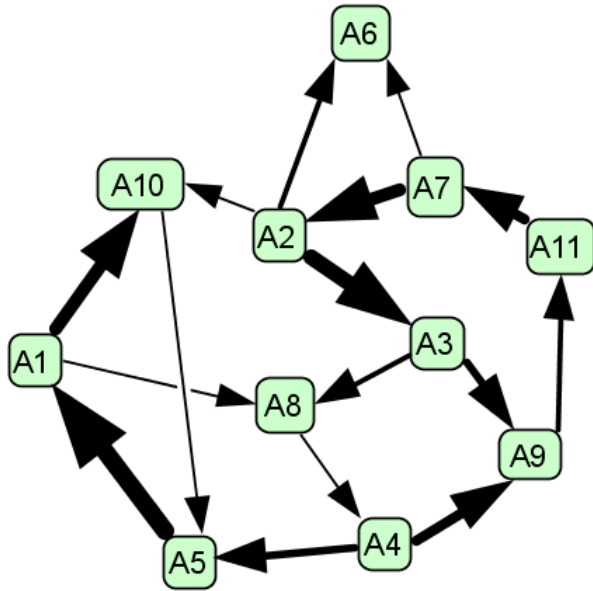
The steady-state rates of a **simple cycle** or a **linear chain** of reactions are determined by the limiting (slowest) reactions step (if the *slowest* reaction step can be well-defined). The relaxation dynamics is described by the second slowest step.

In **complex networks** there is usually no a single limiting reaction step. The steady state and relaxation dynamics is described by a **dominant system** (Gorban&Radulescu, 2007)

Dominant systems of **networks of monomolecular reactions** with well-separated kinetic constants often do not depend on exact values of kinetic parameters but on their ranking

In non-linear systems, there can be **several dominant systems**, and they can **change each other in time**: a complex systems “walks through” simpler subsystems.

Monomolecular networks with time separation can be solved without exact knowledge of kinetic rates



$$\frac{dC(t)}{dt} = KC(t) + K_0$$

$$c(t) = c^s + \sum_{k=1}^n r^k (l^k, c(0) - c^s) \exp(-\lambda_k t) \quad (2)$$

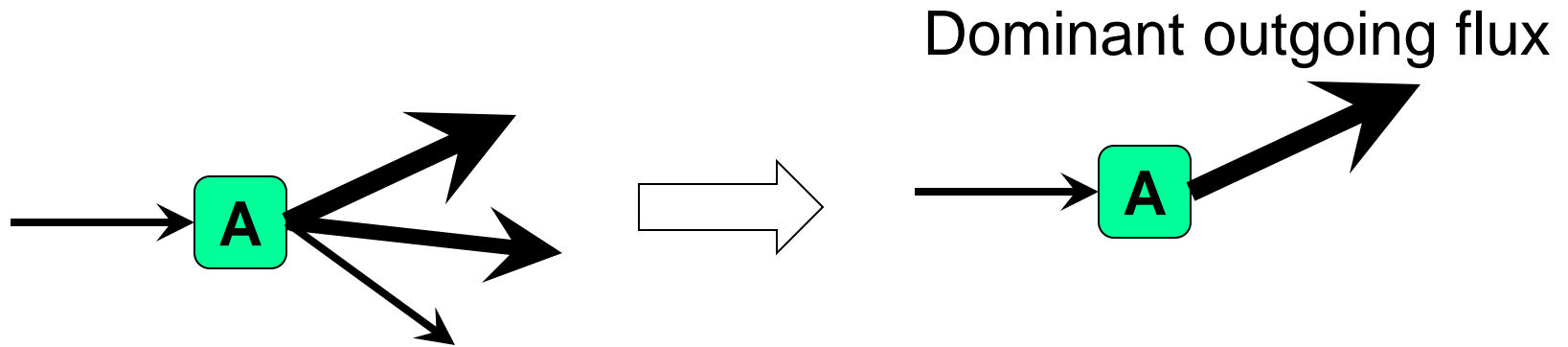
where λ_k , l^k , r^k , $k = 1, \dots, n$ are the eigenvalues, the left eigenvectors (vector-rows) and the right eigenvectors (vector-columns) of the matrix K , respectively, i.e.

$$K r^k = \lambda_k r^k, l^k K = \lambda_k l^k. \quad (3)$$

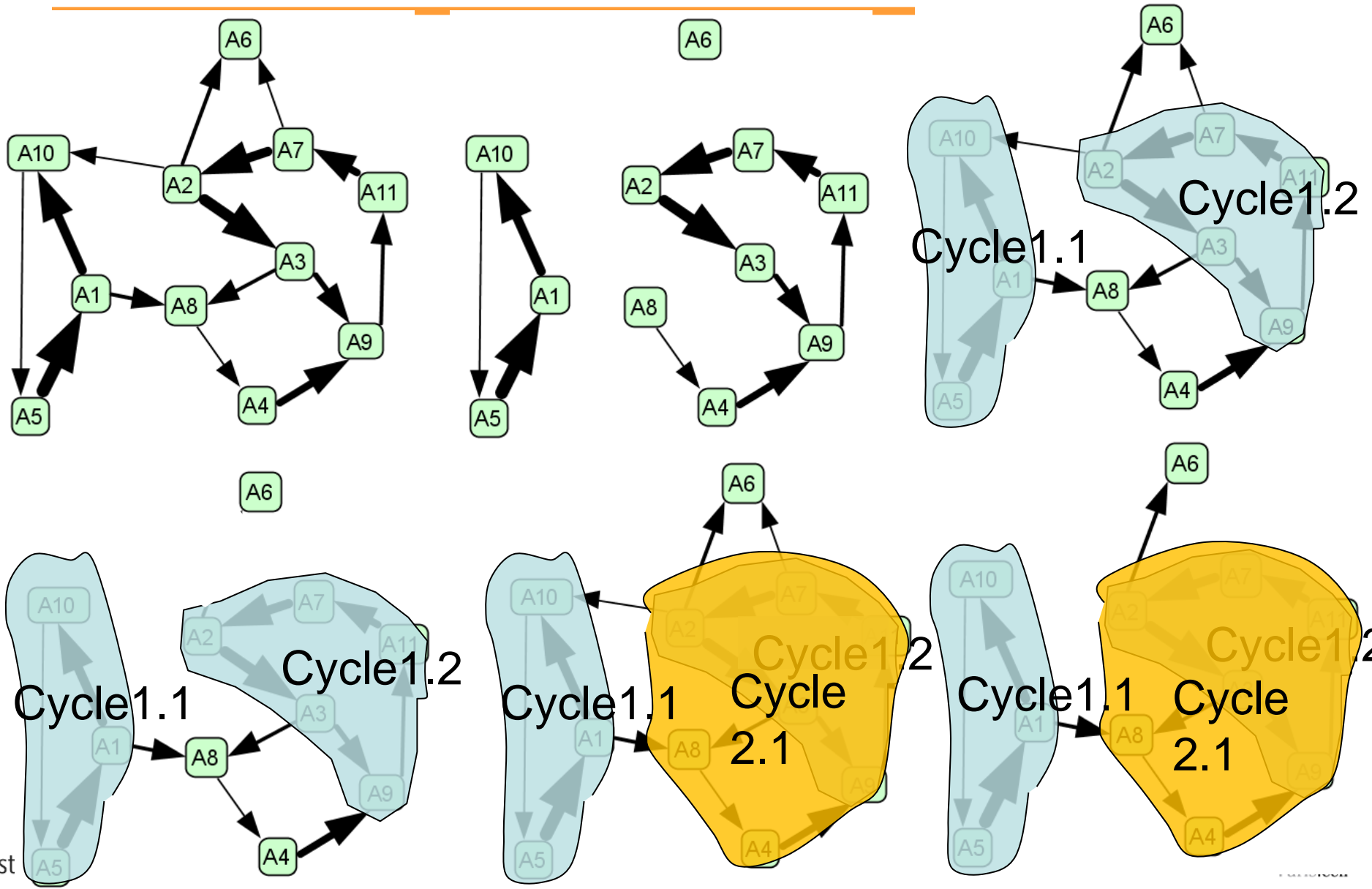
with the normalization $(l^i, r^j) = \delta_{ij}$, where δ_{ij} is Kronecker's delta.

For monomolecular systems with time separation, eigenvectors contain only -1, 0, 1 values. The pattern of these values can be derived from the network topology and the order of parameters (Gorban&Radulescu, 2007)

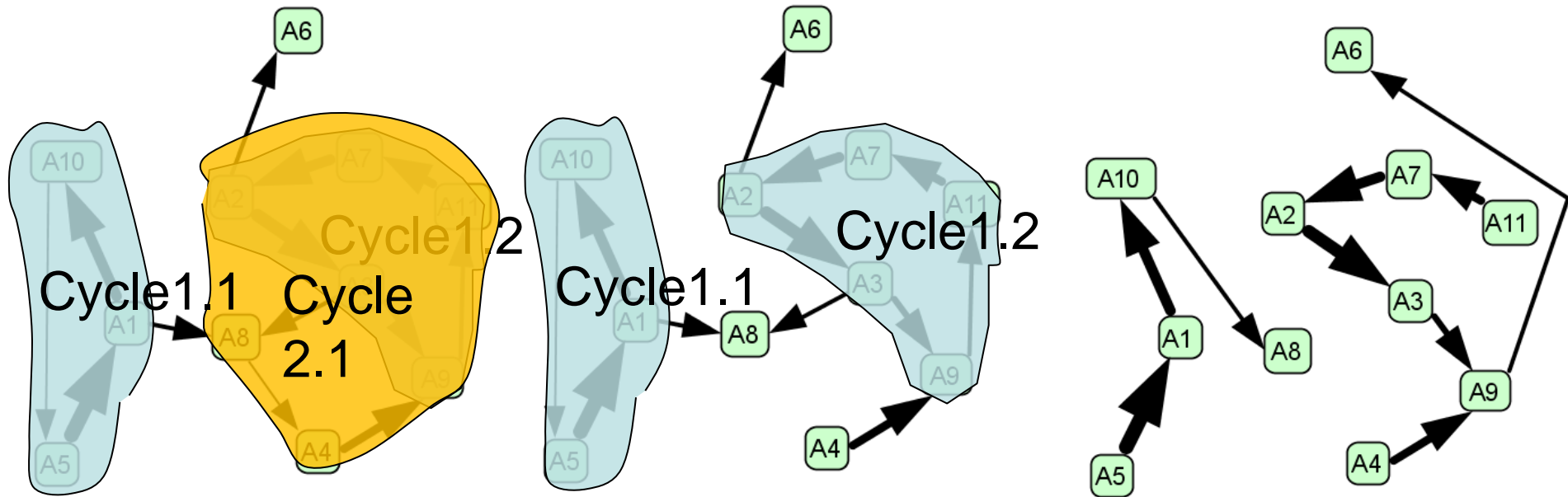
Elementary operation: neglect small outgoing fluxes in the reaction forks



Dominant dynamical system for linear networks: cycle gluing



Dominant dynamical system for linear networks: cycle “surgery”

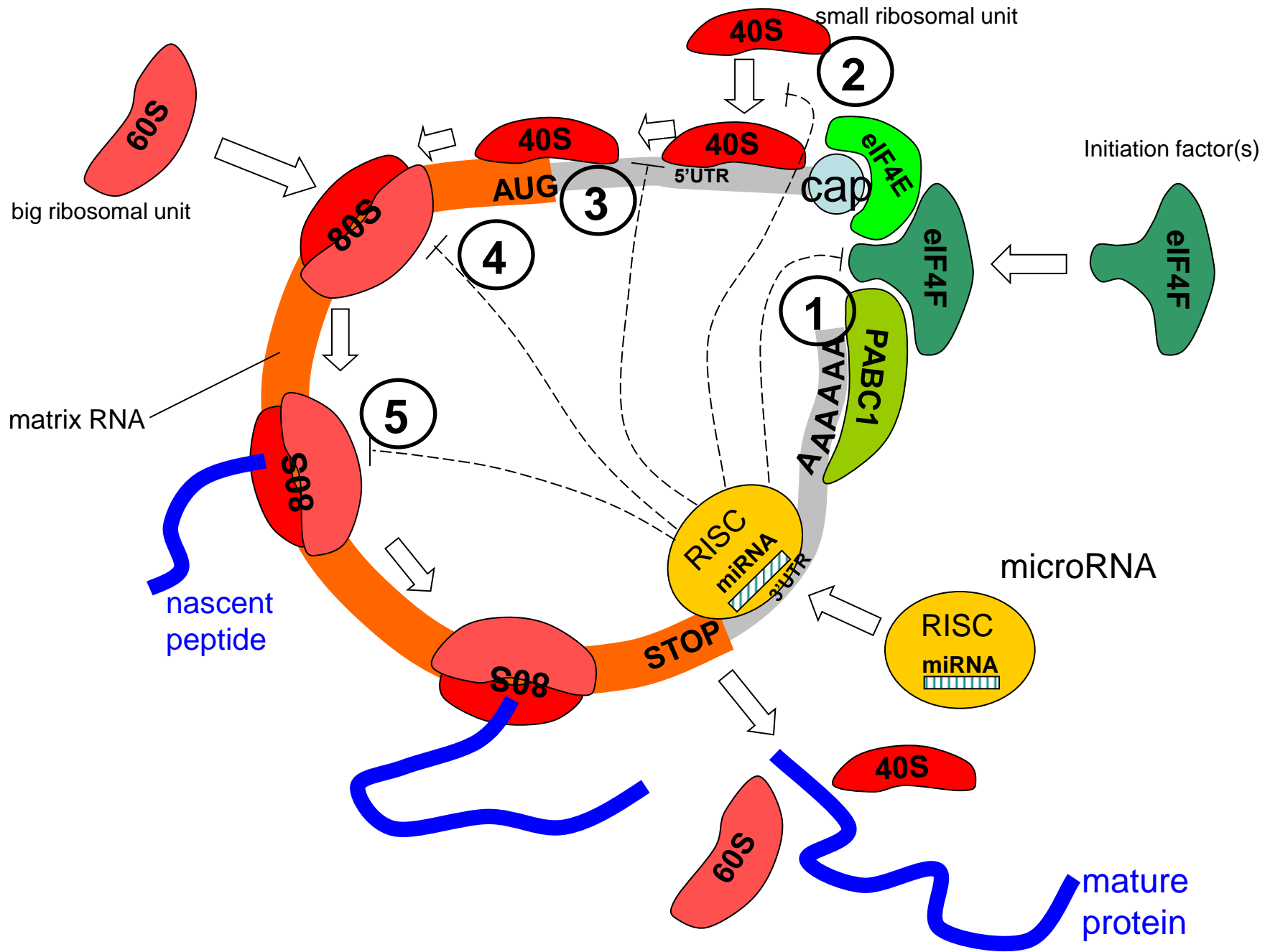


Acyclic dominant dynamical system approximating relaxation

«Simple» problem: modeling single protein synthesis and its regulation



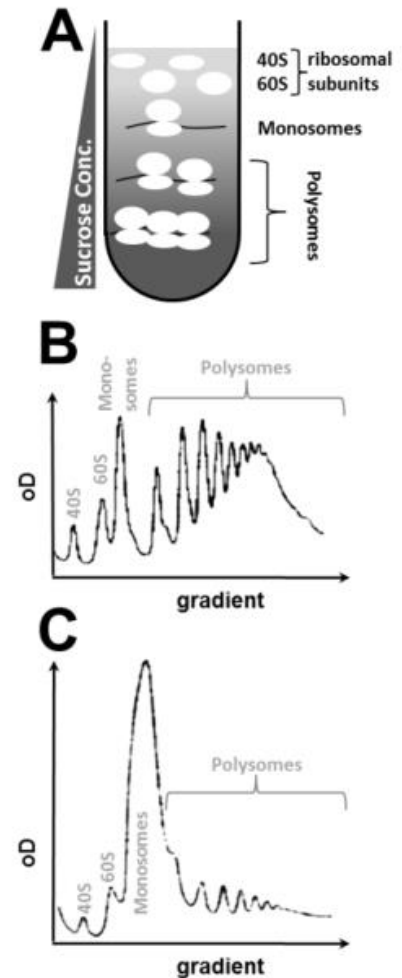
Protein synthesis and its regulation by microRNA



Observable quantities (macrovariables)

- Total amount of mRNA (**MT**)
- Amount of protein (**PR**)
- Polysomal profile (average number of ribosomes on a translated mRNA, **RB**)
- (May be) simple dynamics features such as the relaxation time

No **simple** mathematical model exists to describe these quantities altogether



Problem with the number of states

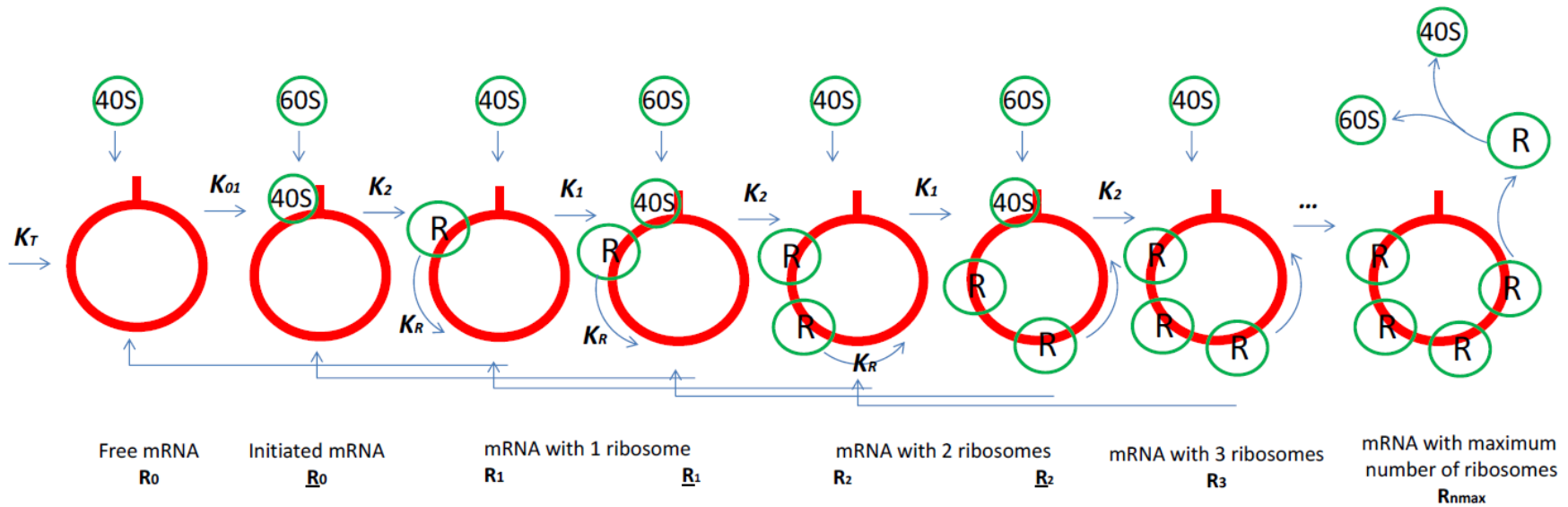


Figure 1. Schematic process of detailed translation representation. It requires $2 \times (n_{max} + 1)$ mRNA states.

Simplest basic model of transcription+translation+degradation (derived by lumping the species)

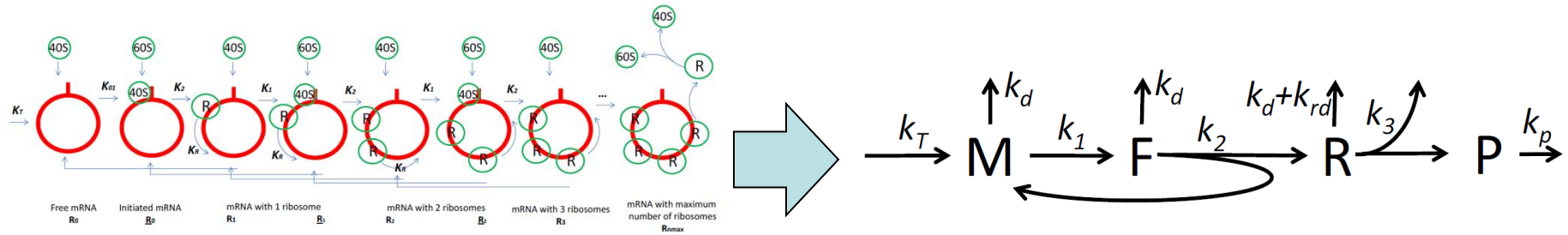


Figure 1. Schematic process of detailed translation representation. It requires $2 \times (nmax+1)$ mRNA states.

M – amount of mRNA with translation initiation site not occupied by assembling ribosome,

F – amount of mRNA with translation initiation site occupied by assembling ribosome,

R – amount of ribosomes sitting on mRNA synthesizing proteins,

P – amount of proteins.

In terms of \underline{R}_i and R_i variables, M and F represent the lumped values:

$$M = \sum_{i=0}^{nmax} R_i, \quad F = \sum_{i=0}^{nmax} \underline{R}_i \quad \text{and} \quad MT = M + F.$$

Simplest and basic model of transcription+translation+degradation (derived by lumping the species)

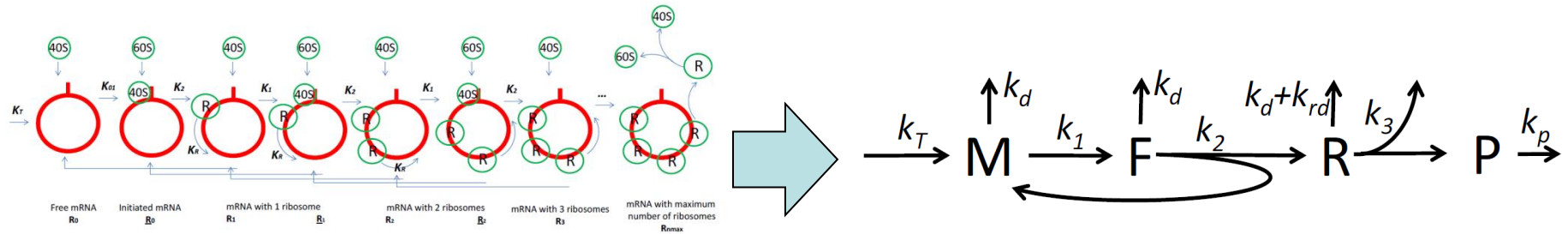
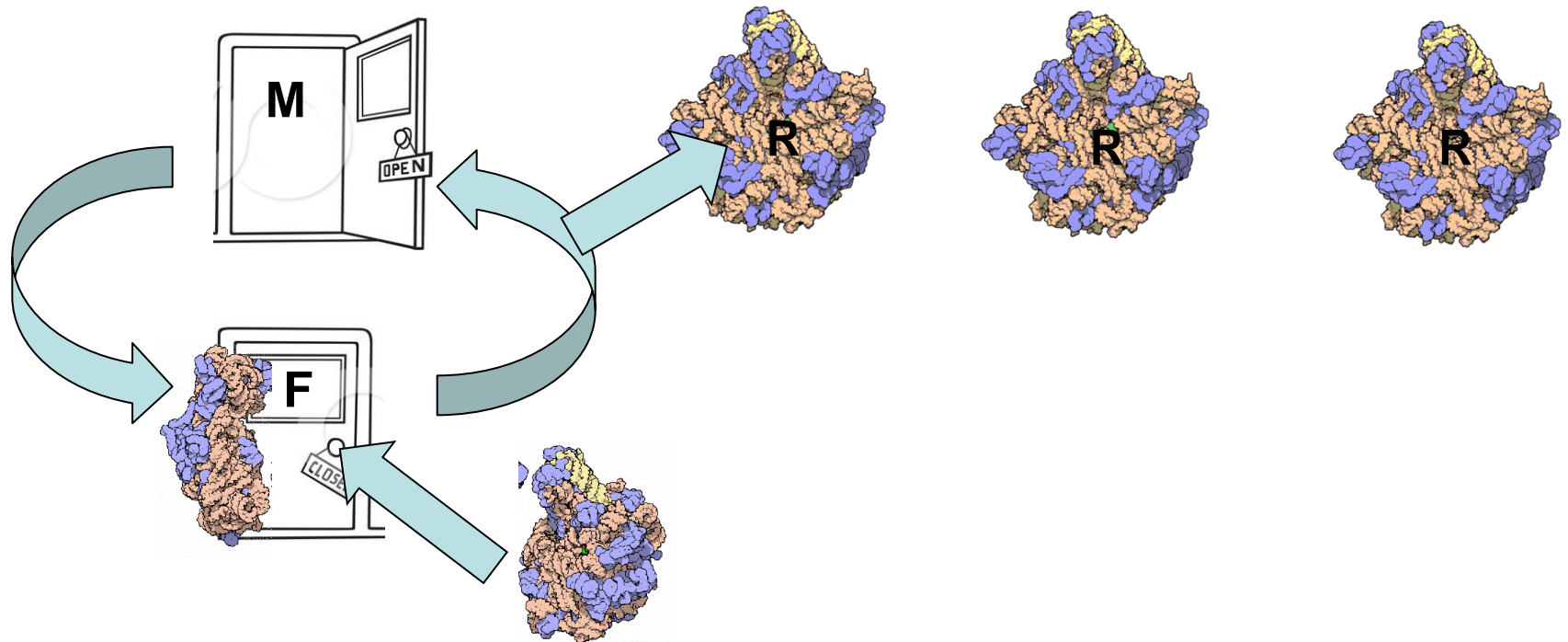


Figure 1. Schematic process of detailed translation representation. It requires $2 \times (n_{max} + 1)$ mRNA states.



Simplest and basic model of transcription+translation+degradation (derived by lumping the species)

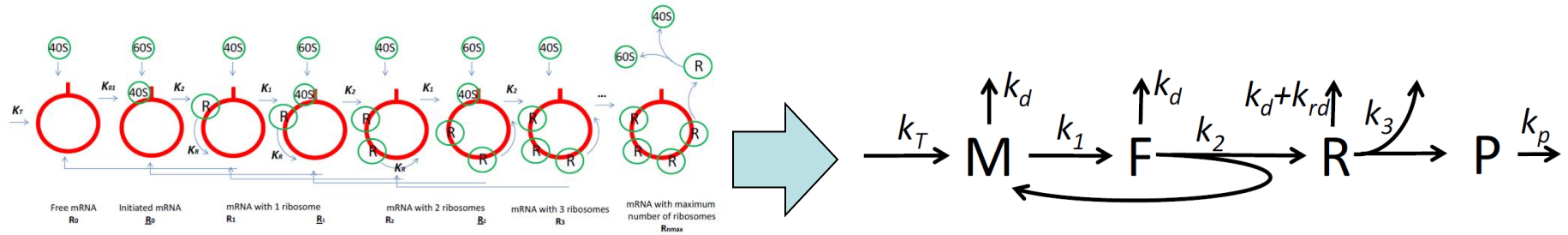
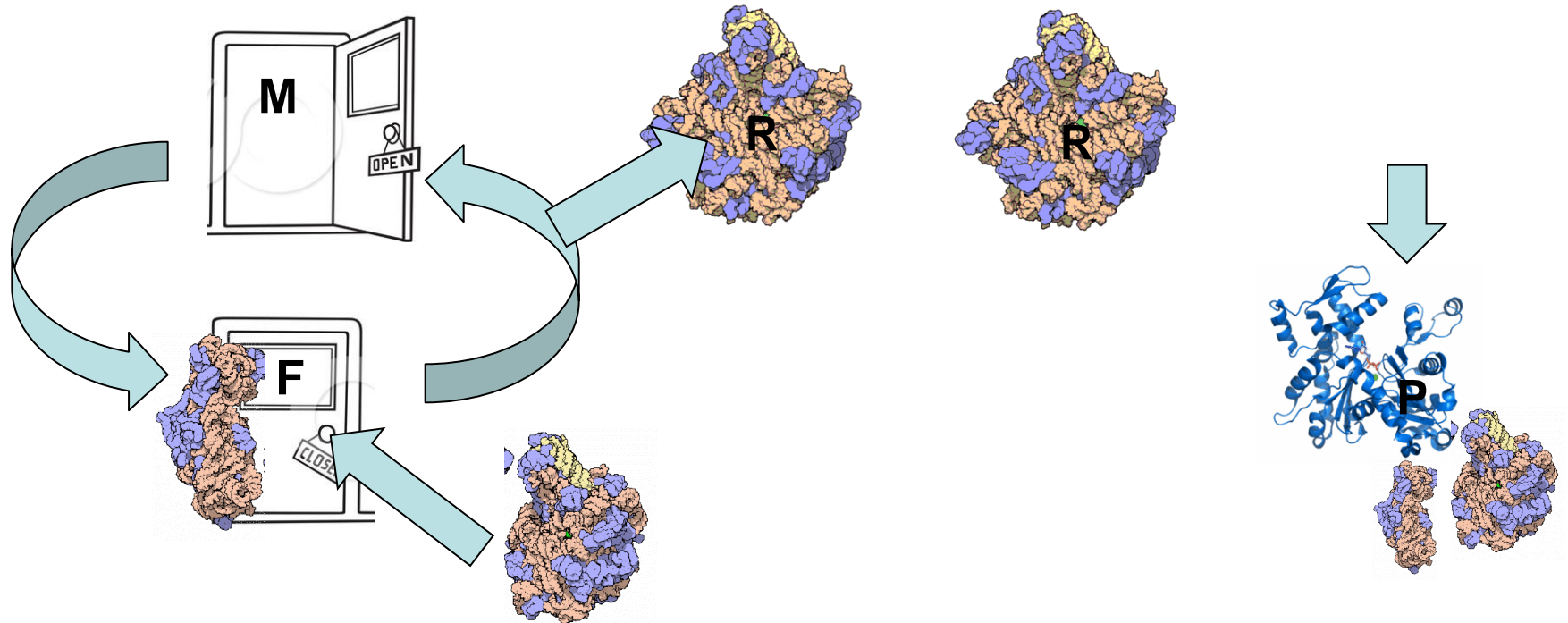


Figure 1. Schematic process of detailed translation representation. It requires $2 \times (n_{max} + 1)$ mRNA states.



Simplest and basic model of transcription+translation+degradation (derived by lumping the species)

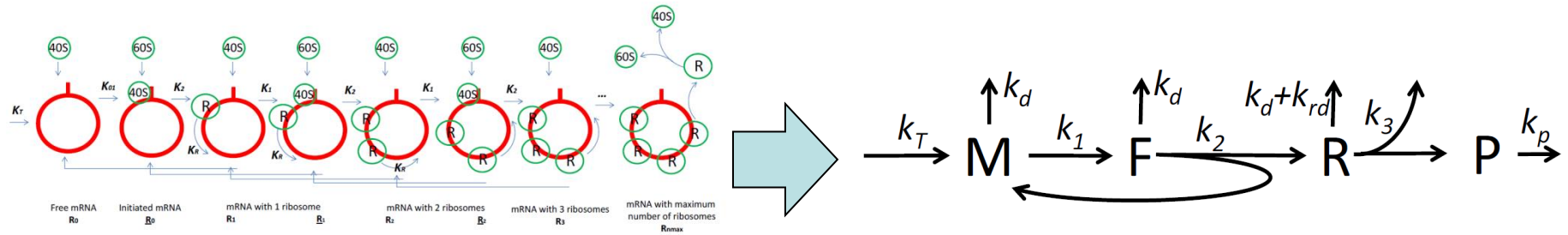
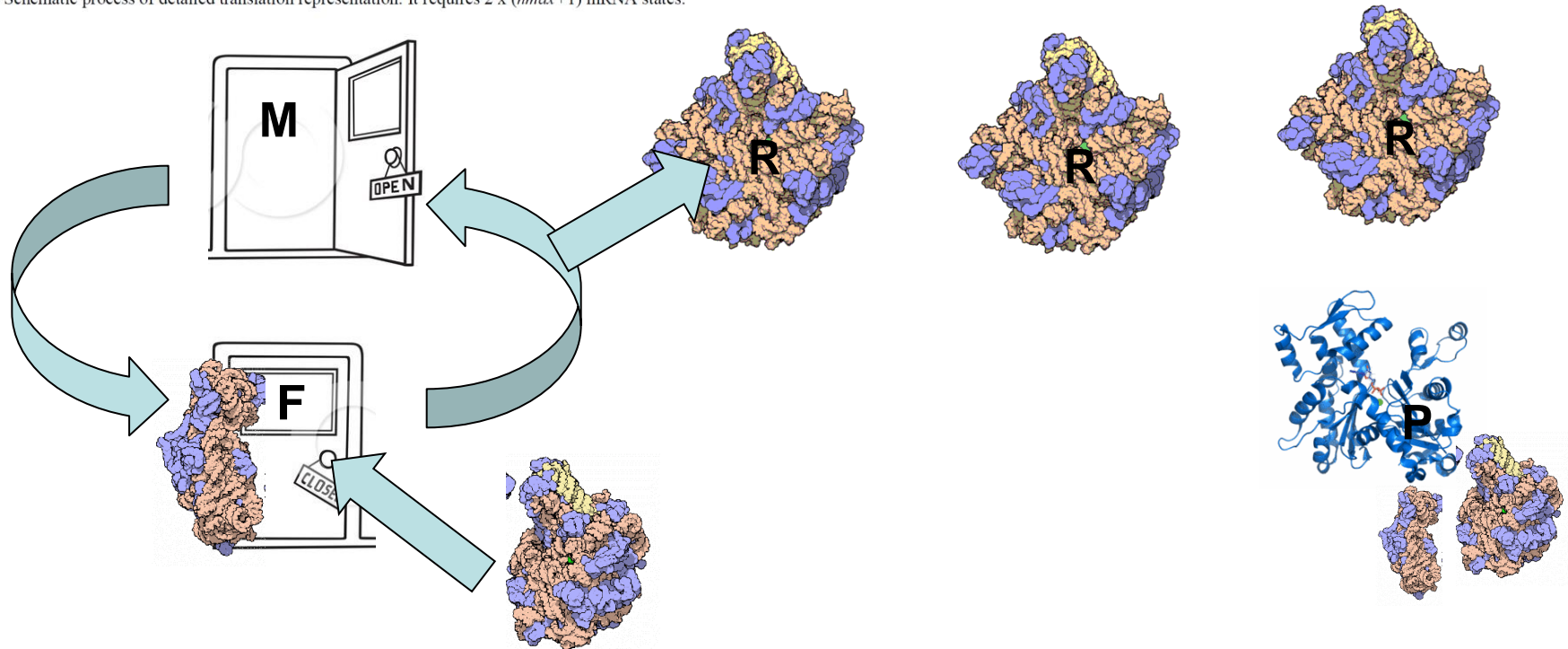


Figure 1. Schematic process of detailed translation representation. It requires $2 \times (n_{max} + 1)$ mRNA states.



Simplest and basic model of transcription+translation+degradation (derived by lumping the species)

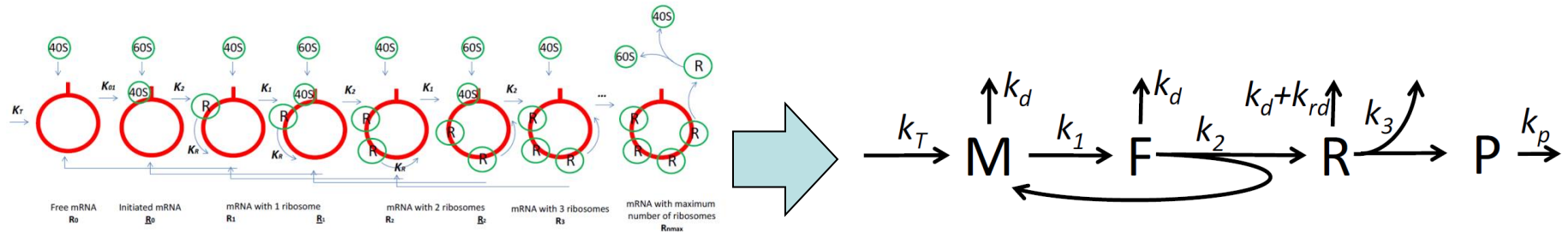
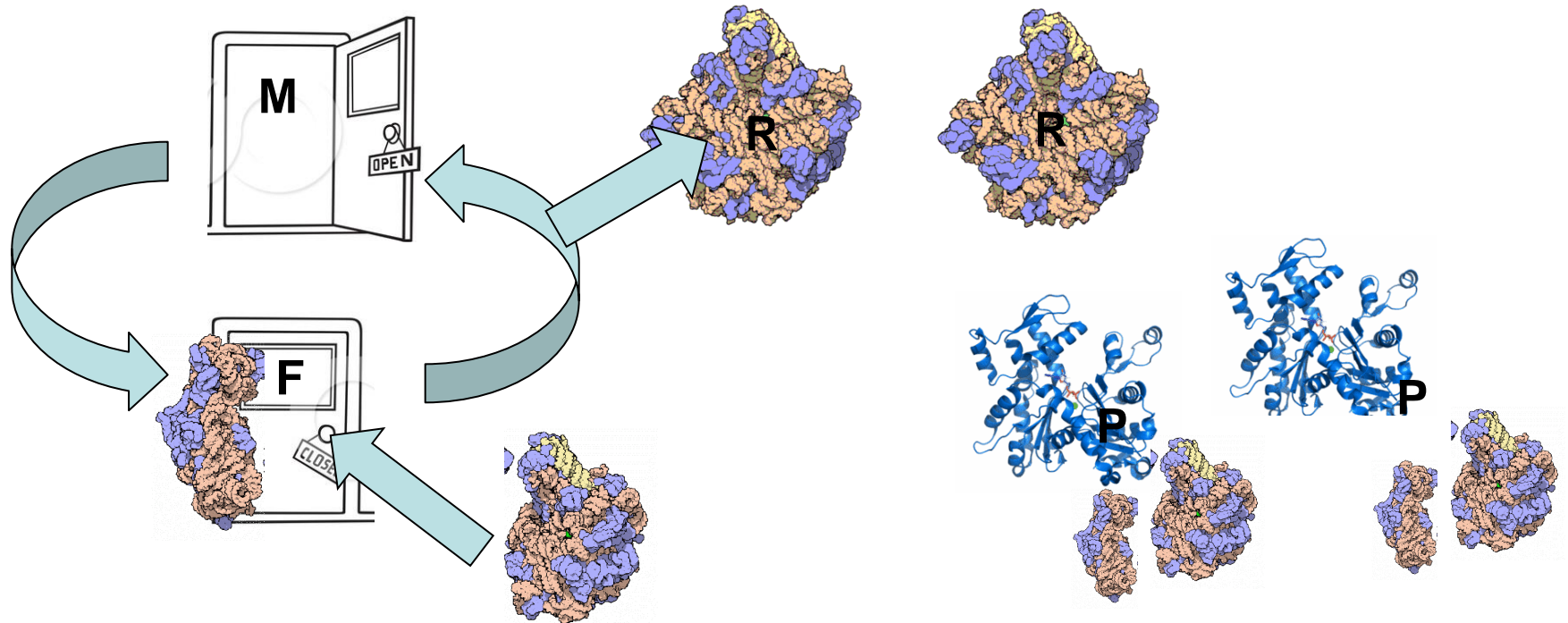


Figure 1. Schematic process of detailed translation representation. It requires $2 \times (n_{max} + 1)$ mRNA states.



Simplest and basic model of transcription+translation+degradation (derived by lumping the species)

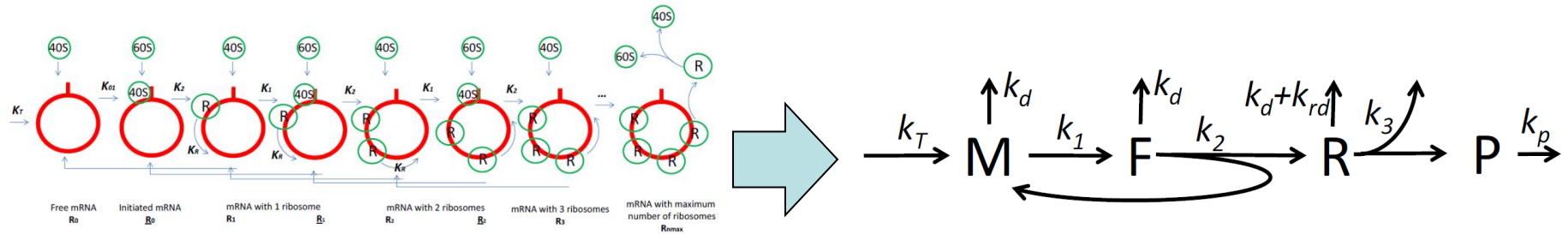
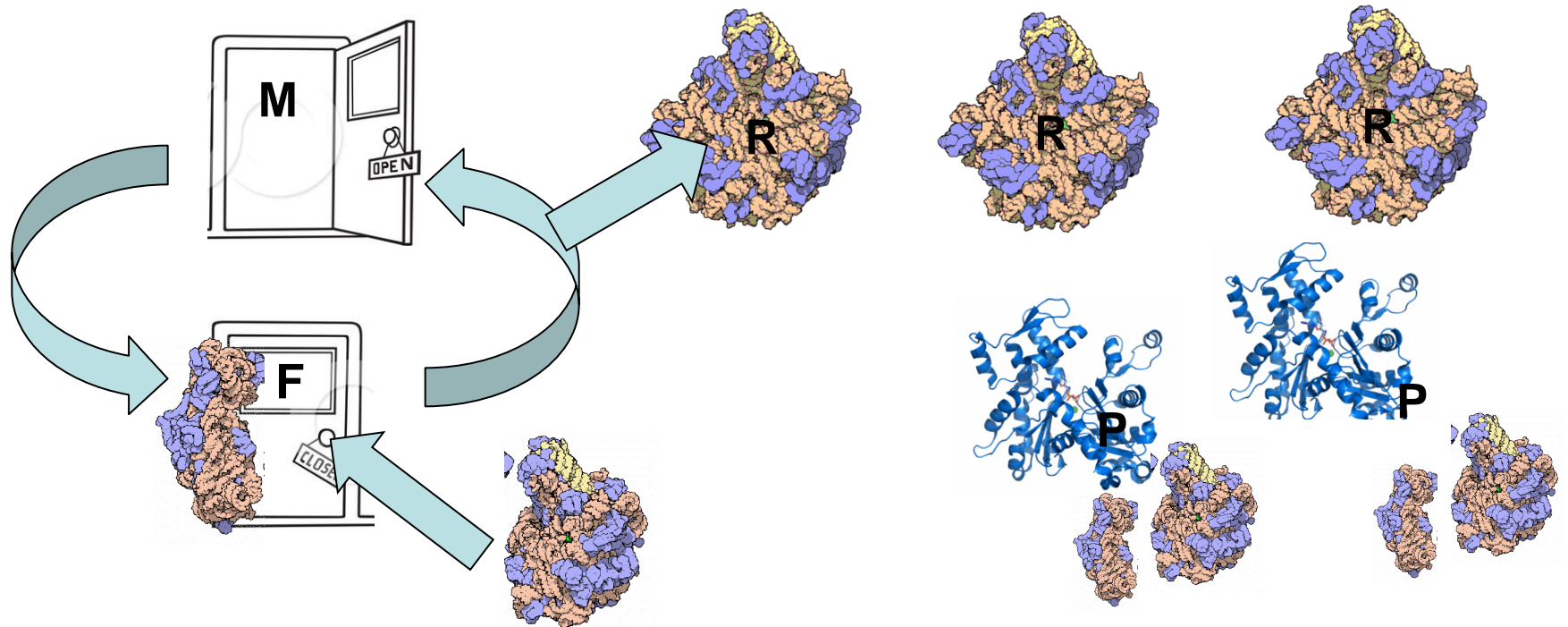
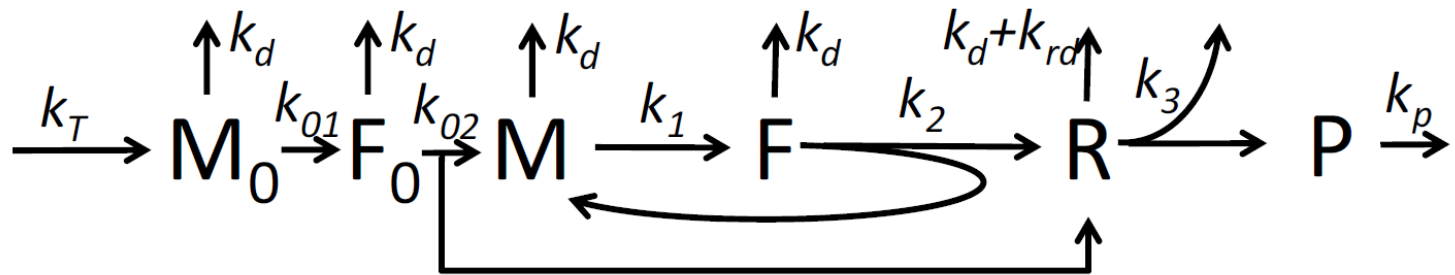
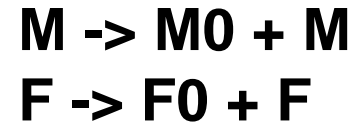
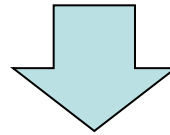
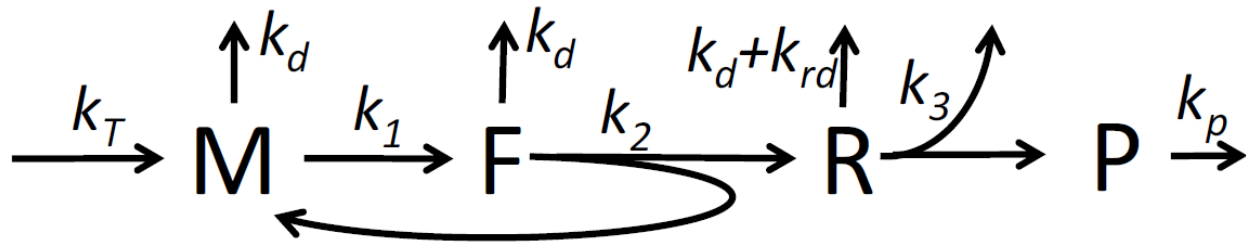


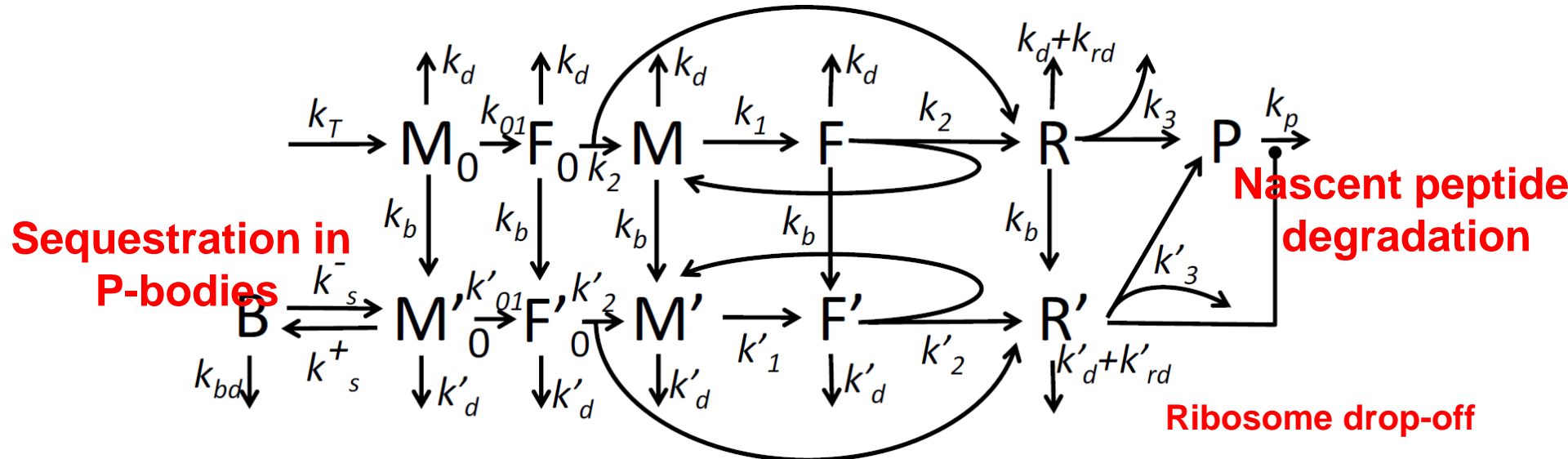
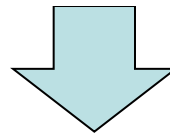
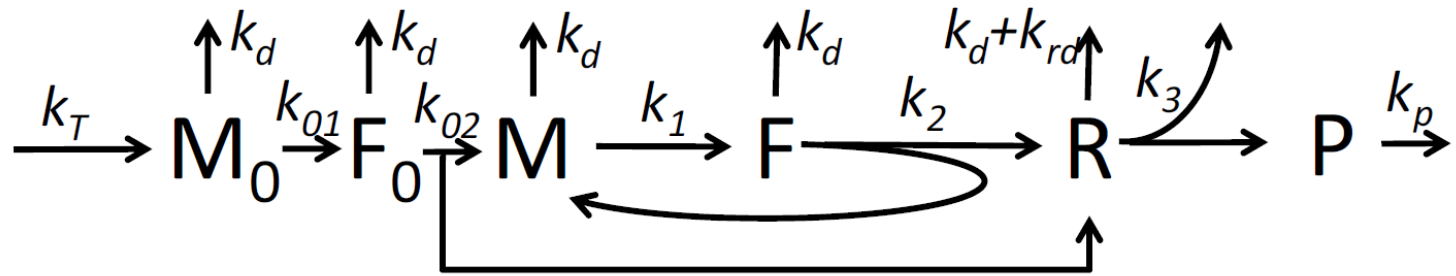
Figure 1. Schematic process of detailed translation representation. It requires $2 \times (n_{max} + 1)$ mRNA states.



Extension 1 of the model: representing explicitly the initiation stage



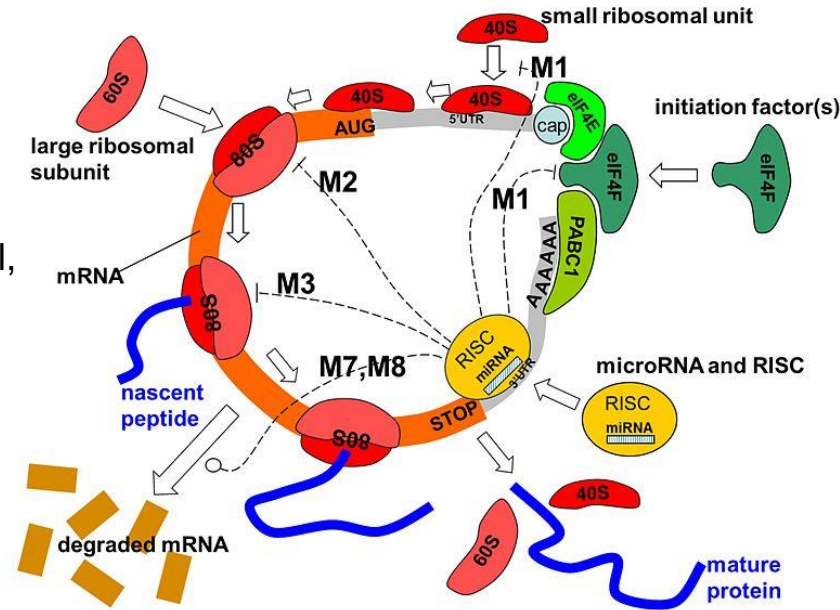
Extension 2 of the model: adding miRNA in the model



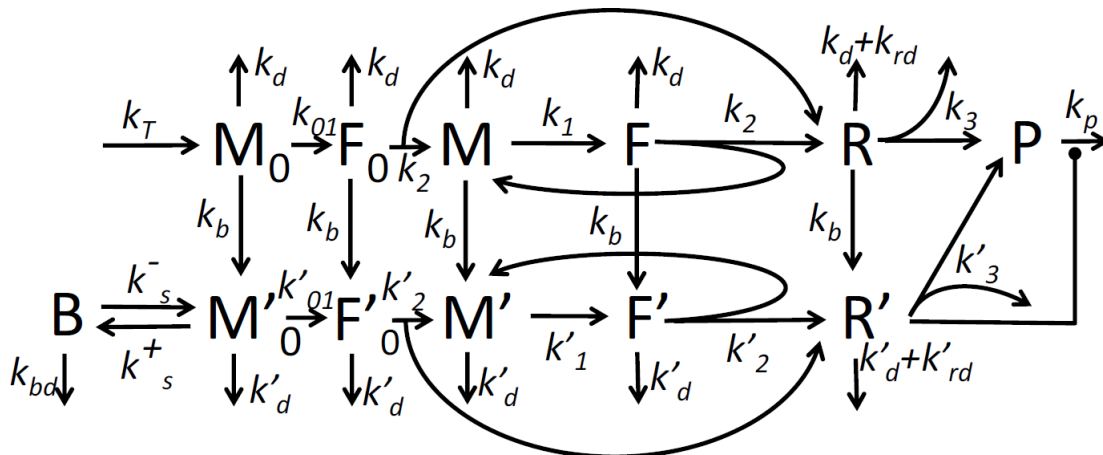
Mathematical model, describing 9 mechanisms of microRNA action on translation

Zinovyev A et al *BMC Systems Biology*, 2010

Morozova N, Zinovyev A. et al, *RNA*, 2012



9 distinct mechanisms of miRNA action



What is the “main” mechanism?

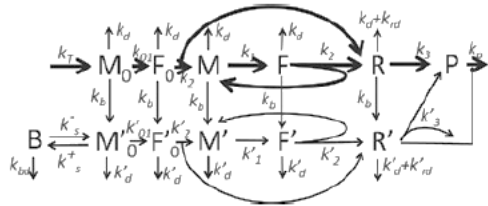
What does it mean?

Dominant paths and miRNA mechanisms

Dominant path

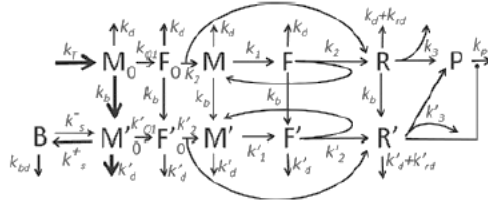
Biological interpretation

Corresponding miRNA-mediated translation repression mechanism(s)



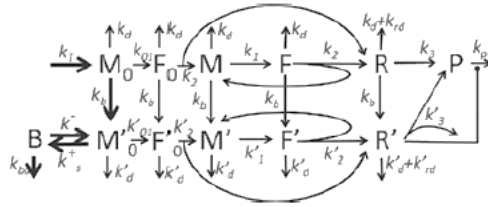
M_0F_0MFRP
Normal translation with negligible effect of miRNA

None



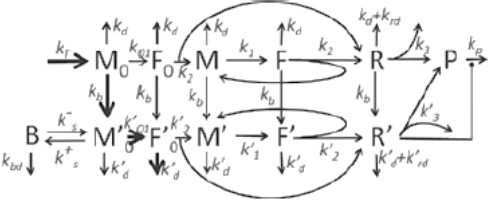
$M_0M'_0$
The dominant effect is degradation of mRNA by miRNA.

M1: Cap inhibition
M7: Decay
M8: Cleavage



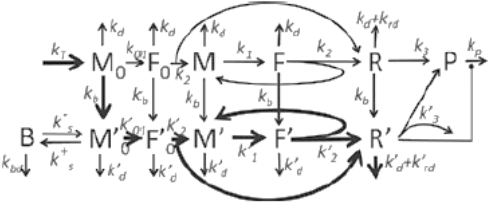
$M_0M'_0B$
mRNA is captured in P-bodies.

M6: Sequestration of mRNA in P-Bodies



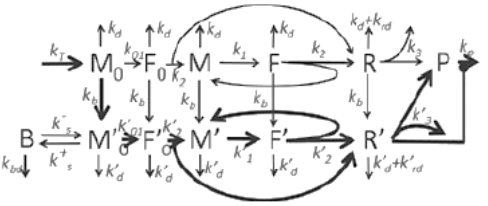
$M_0M'_0F'_0$
mRNA translation is stuck after initiation, before the assembly of the ribosome.

M2: 60S subunit joining inhibition



$M_0M'_0F'_0M'F'R'$
mRNA is stuck with ribosomes on it and destroyed, or mRNA translation is prematurely aborted.

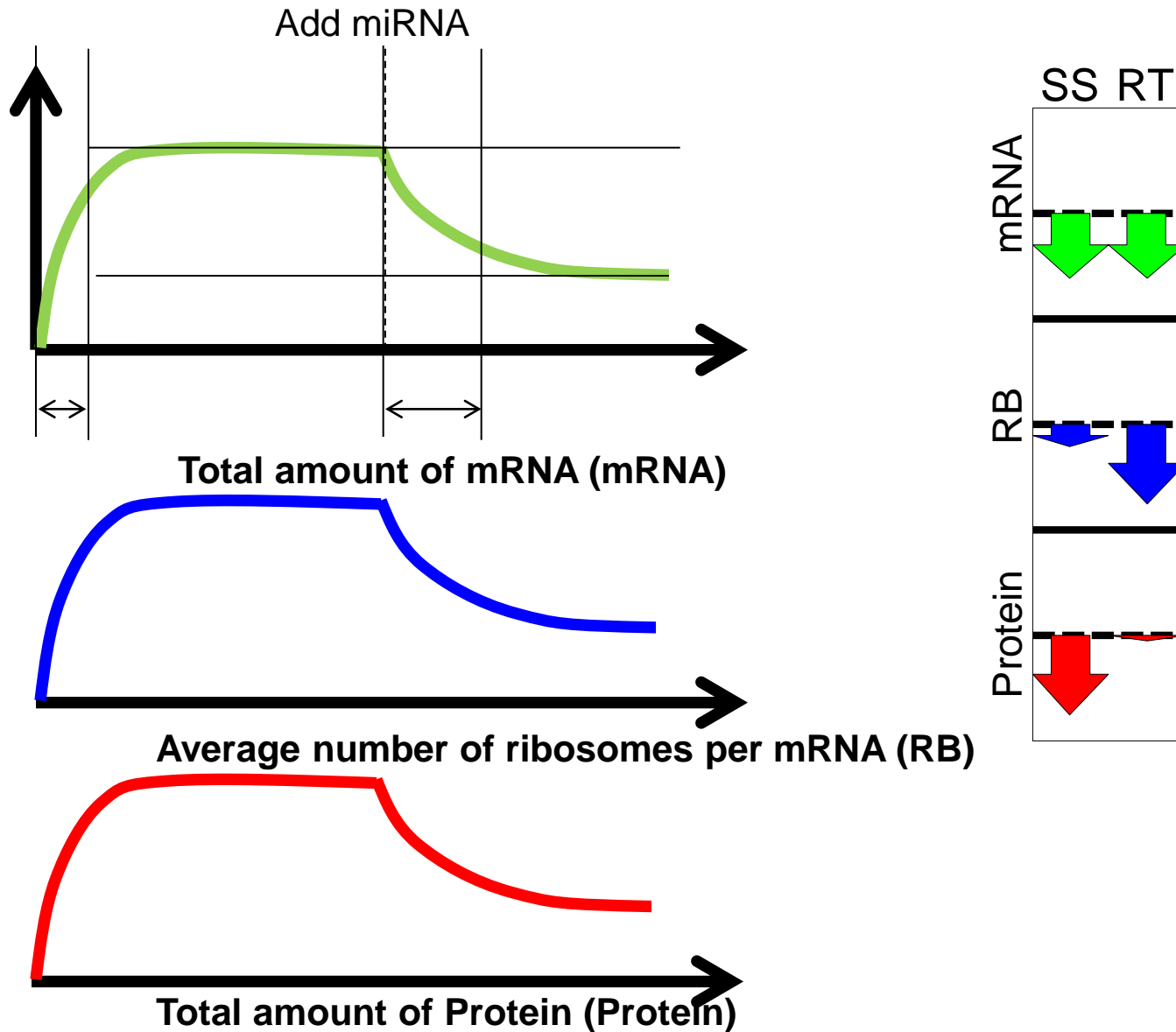
M3: Elongation inhibition
M4: Ribosome drop-off



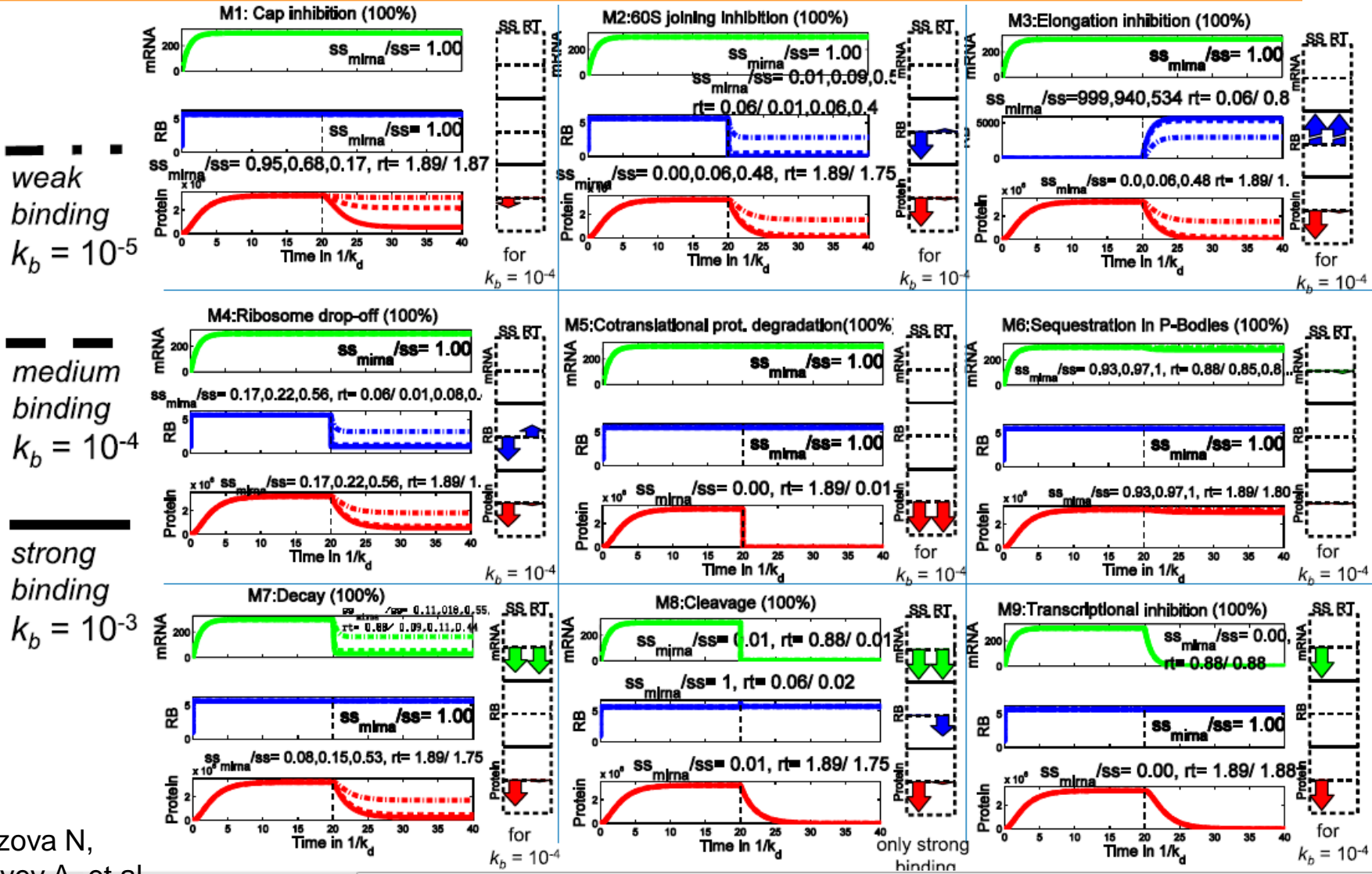
$M_0M'_0F'_0M'F'R'P$
Protein synthesis in the presence of miRNA with low mRNA degradation

M1: Cap inhibition
M2: 60S subunit joining inhibition
M3: Elongation inhibition
M5: Cotranslational protein degradation mechanisms

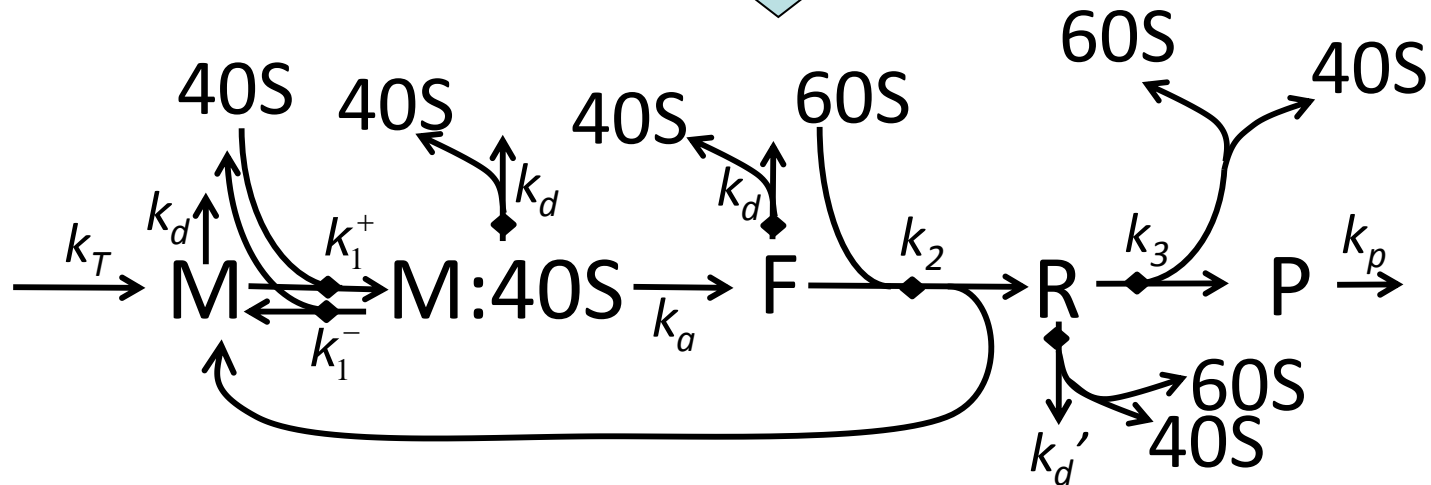
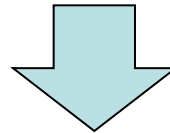
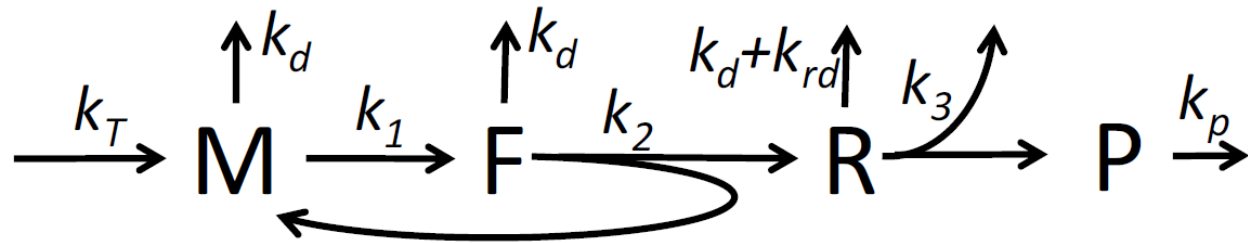
Notion of kinetic signature



Kinetic signatures of miRNA action

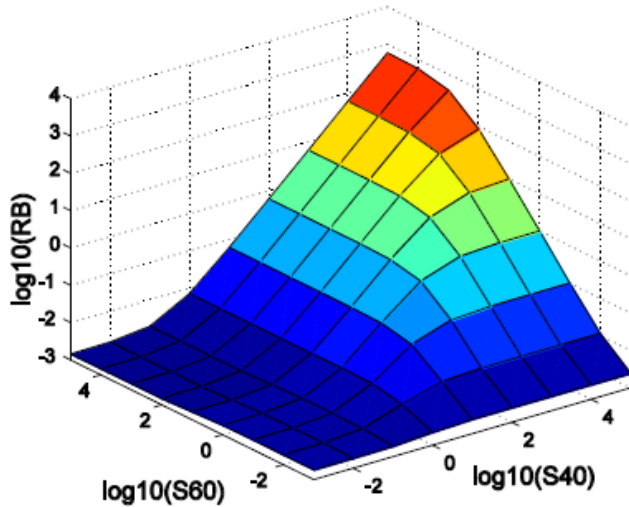
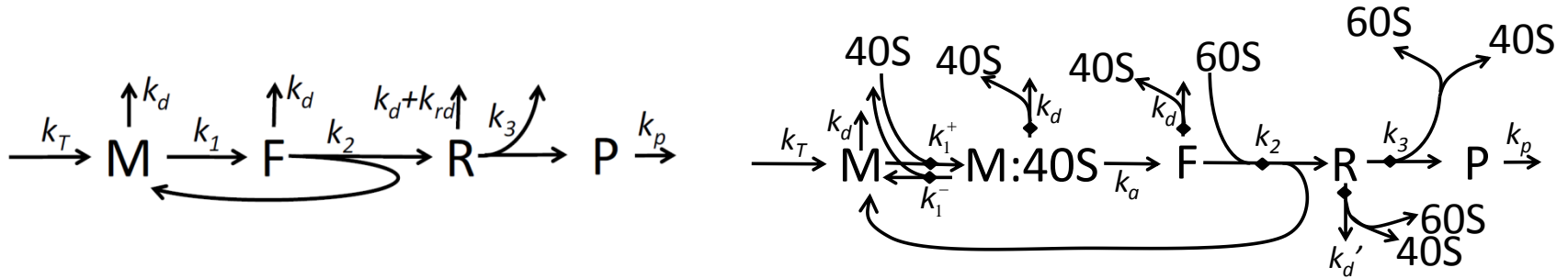


Problem of infinite polysome size

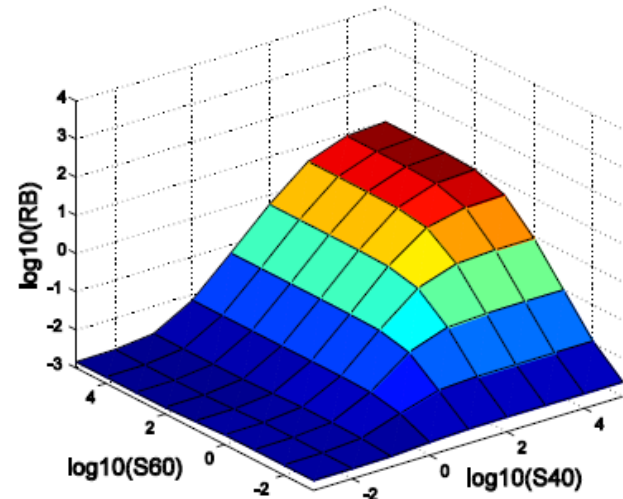


$$RB = R/(M+F)$$

average number of ribosomes per mRNA



f10')



f11')

Figure 7. Number of translating ribosomes per mRNA (RB) in $f10'$ and $f11'$ models of translation as a function of concentrations of small ($S40$) and large ($S60$) ribosomal subunits for fixed concentrations of the initiation factors.

Other possible model extensions

More explicit representation of translation termination or elongation, description of ribosome stalling phenomenon.

More detailed representation of the mRNA initiation process.

Description of phenomena connected with uneven distribution of ribosomes along mRNA, such as described in recent literature on explicit studies of ribosome positioning on mRNAs (Ingolia et al, 2009).

Other possible model extensions

Explicit modeling of the mRNA codon usage.

Mean-field models of the ribosomes' interaction: The simplest method to include the interaction of ribosomes in the lumped model is a dependence of the ribosome drop-off constant k_{rd} on the average concentration θ of the ribosomes per initiated molecule of mRNA: $k_{rd}=k_{rd}(\theta)$. For example, for the scheme presented in Figure 3 it may be $k_{rd}(\theta)=a/(b-\theta)$, where $\theta=R/(M+F)$.

Mean-field models of how the property of the mRNA (such as its stability) might change depending on the state and also on the history of mRNA. For example, one can imagine a (very hypothetical) version of mRNA kinetics with “mRNA aging” such as each new round of translation makes mRNA more fragile and prone to destruction. Or in opposite, mRNA can become more stable with ribosomes sitting on it.

Other possible model extensions (current work)

Modeling distribution of model parameters, leading to existence of **population of mRNAs** with different speeds of different steps of translation

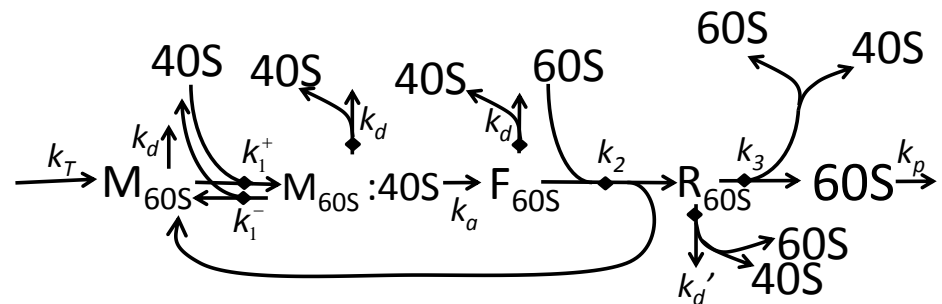
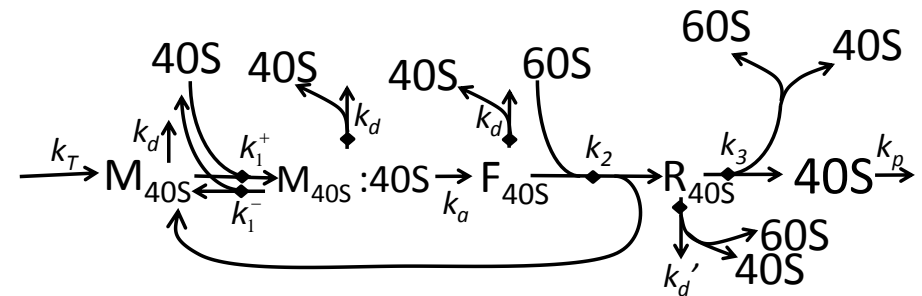
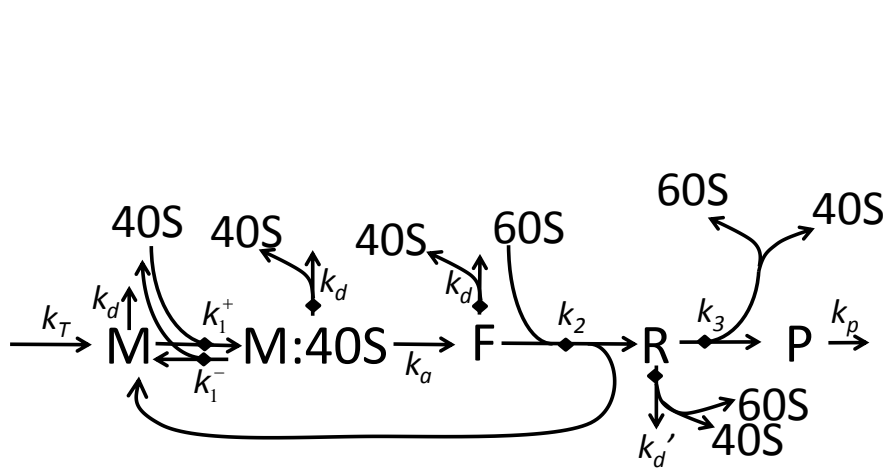
Explicit modeling of **competition of various protein syntheses processes for resources** (ribosomal subunits and initiation factors). The most interesting is to include in this picture the **production of the resources themselves** (transcription, translation, degradation), which will introduce complex global regulatory feedbacks

Global cellular protein synthesis model

P – total pool of proteins in a cell except ribosomal units

40S – total pool of small ribosomal units

60S – total pool of large ribosomal units



Continuity of life

- There exists a minimal pool size of ribosomes needed for sustaining translation of cellular proteins
- Life can not be created (translated) from scratch (isolated genome), this minimal ribosomal pool size should pre-exist

Acknowledgements

Alexander Gorban,
Department of Mathematics, University of Leicester, UK



Nadezhda Morozova, CNRS, France



Ovidiu Radulescu, Univ.of Montpellier



Nora Nonne, CNRS, France



Annick Harel-Belan, CNRS, France



Emmanuel Barillot, Institut Curie, France

Several references

- A. Gorban, B. Kegl, D. Wunsch, A. Zinovyev (Eds.), [Principal Manifolds for Data Visualisation and Dimension Reduction](#), *Lecture Notes in Computational Science and Engineering*, Vol. 58, Springer, Berlin – Heidelberg – New York, 2007. (ISBN 978-3-540-73749-0)
- A.N. Gorban, A. Y. Zinovyev, [Principal Graphs and Manifolds](#), Chapter 2 in: Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques, Emilio Soria Olivas et al. (eds), IGI Global, Hershey, PA, USA, 2009, pp. 28-59 (and also [arXiv:0809.0490v1](#) [cs.LG])
- A.N. Gorban, I.V. Karlin, A.Y. Zinovyev, [Constructive methods of invariant manifolds for kinetic problems](#) Phys. Rep., 396, 2004, 197-403
- A. Zinovyev, N. Morozova, A.N. Gorban, and A. Harel-Belan, [Mathematical Modeling of microRNA-Mediated Mechanisms of Translation Repression](#), *Advances in Experimental Medicine and Biology* Vol. 774, Springer, 2013, pp. 189-224.
- N. Morozova, A. Zinovyev, N. Nonne, L.-L. Pritchard, A.N. Gorban, and A. Harel-Bellan, [Kinetic signatures of microRNA modes of action](#), *RNA*, Vol. 18, No. 9 (2012) 1635-1655,
- A. N. Gorban, A. Zinovyev, [Principal manifolds and graphs in practice: from molecular biology to dynamical systems](#), *International Journal of Neural Systems*, Vol. 20, No. 3 (2010) 219–232.
- A. N. Gorban, O. Radulescu, A. Y. Zinovyev, [Asymptotology of chemical reaction networks](#), *Chemical Engineering Science* 65 (2010) 2310–2324
- A. Zinovyev, [Dealing with complexity of biological systems: from data to models](#) (habilitation thesis) , [arXiv:1404.1626](#) [q-bio.QM]
- L. Calzone, A. Gelay, A. Zinovyev, F. Radvanyi, E. Barillot, [A comprehensive modular map of molecular interactions in RB/E2F pathway](#), *Molecular Systems Biology*, March 2008.