Almost Linear Systems

- A null model for genetic regulatory networks

Rudolf Hanel

COMPLEX SYSTEMS RESEARCH GROUP
AT THE MEDICAL UNIVERSITY OF VIENNA

D. Stokic, R.Hanel and S.Thurner, *The window at the edge of chaos in a simple model of gene interaction networks* Physical Review E 77, 061917 (2008).

D. Stokic, R.Hanel and S.Thurner, *A fast and efficient gene-network reconstruction method from multiple over-expression experiments*, BMC Bioinformatics **10** (2009) 253.

R. Hanel, M. Pöchacker and S. Turner, *Living on the edge of chaos: minimally non-linear models of genetic regulatory dynamics*, Phil. Trans. Roy. Soc. xxx.

Abstract

Intracellular signalling in general or genetic regulation and enzyme kinetics in specific are processes known to be of highly non linear nature. Yet, the idea of parsimony guides evolution of theory along the way of the simplest available models in sufficient agreement with experimental observations. Almost Linear Systems, - i.e. systems that follow a linear dynamic for the concentration of agents under the constraint that agent concentrations must be non-negative -, can be understood as a null-model of observed genetic regulation processes like for instance - the recruitment of various agents into functional protein complexes. In this way Almost Linear Systems provide a starting-point for a systematic identification of crucial non-linear agent interactions that defy linearization or can be used to predict the dynamics of agents missing in considered assays.

Overview

- What is an Almost Nonlinear System?
- What are basic properties of ALS?
- Why and in which context are ALS interesting?

Genetic Regulation Networks

Who sees better?

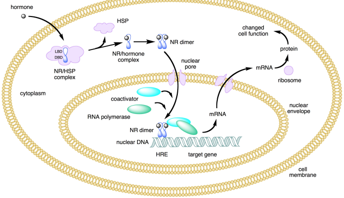
Boolean Goggles (fluorescence microscope based essays)

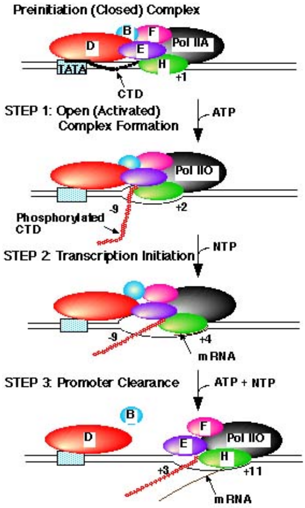
– or –

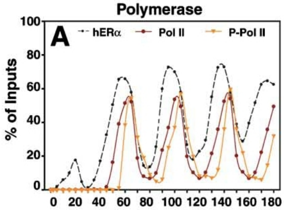
Differential Equations Goggles (ChIP: promotor-protein binding)

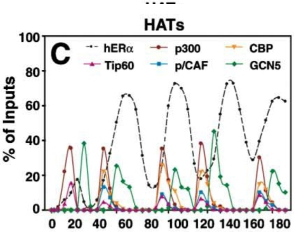
Introduction

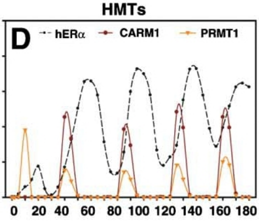
Example: Nuclear Receptor











1

[1] R. Metivier et al. Cell, **115** (2003) 751-736

Let us consider the following:

$$\dot{x}_i = S_{ij}\nu_j + J_i$$

- x_i ... concentration of substance i
  (mRNA, peptides/proteins, minerals, sugars, ...)
- ν_j ... process j,
  (e.g. proteins A and B form a complex C)
- J_i ... a flow of substances i into/out-of the system
- S_{ij} ... e.g. the matrix of stoichiometric coefficients
  (e.g. add one complex C and subtract one protein A and one protein B to/from the substrate)

Non linear systems ...

- usually require a large amount of parameters
- the exact form of the non-linear equations often is not exactly known

Possibility: Linearization of the Non-linear Equations ...

Problem with linear approximations: The bread and water problem, Positivity, ...

Note:

- $\nu_i = \nu_i(\mathbf{x})$ ... can be highly non-linear functions, depending on the set $\mathbf{x} = \{x_j\}$.
- $x_i \geq 0$ ... The system dynamics has to guarantee this, since concentrations x_i can not become negative.

Suppose the dynamics has a fixed point $\dot{x}_i = 0$ at $\mathbf{x}^*$ for some current J^* such that one can simplify the non-linear dynamics by linearizing around $\mathbf{x}^*$.

$$\dot{x}_i = S_{ij} \left(\nu_j(\mathbf{x}^*) + \left. \frac{\partial \nu_j(\mathbf{x})}{\partial x_k} \right|_{\mathbf{x}=\mathbf{x}^*} (x_k - x_k^*) \right) + J_i$$

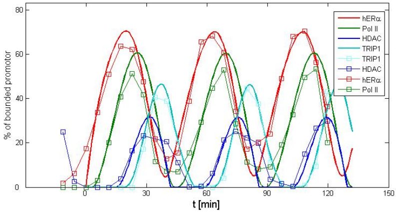
Rem: If $\mathbf{x}^*$ is a fixed point then $0 = S_{ij}\nu_j(\mathbf{x}^*) + J_i^*$ for all i.

Catalytic Equations
Linearization (2)

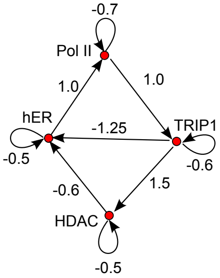
$$\dot{x}_i = \underbrace{S_{ij} \left. \frac{\partial \nu_j(\mathbf{x})}{\partial x_k} \right|_{\mathbf{x}=\mathbf{x}^*}}_{A_{ik}} (x_k - x_k^*) + \underbrace{J_i - J_i^*}_{\Delta J_i}$$

Not to forget ...

- concentrations have to be non negative ... $x_i \geq 0$
- due to possible low molecular concentrations stochastic influences may become important ... noise contributions ν_i



hERα
Pol II
HDAC
TRIP1
TRIP1
HDAC
hERα
Pol II
% of bounded promotor
t [min]
0
20
40
60
80
0
30
60
90
120
150



-0.7
Pol II
1.0
1.0
hER
-1.25
TRIP1
-0.5
-0.6
-0.6
1.5
HDAC
-0.5

- If $x_i > 0$ or if $x_i = 0$ and $\dot{x}_i = 0$ then:

$$\dot{x}_i = A_{ij}(x_j - x_j^*) + \Delta J_i + \nu_i$$

- If $x_i = 0$ and $\dot{x}_i < 0$ (as given by the linear equation) then:

$$\dot{x}_i = 0 \quad .$$

- x_i ... concentration levels of N molecular species i
- A_{ij} ... random adjacency matrix with average connectivity $\langle k \rangle$. Non-zero weights $A_{ij} \in N(0, \sigma_A)$
- $A_{ii} = -\sigma_A D$ ... decay rates are identical for all i
- ΔJ_i ... flow vector (difference to flow defining fixed-point; in the following set to $\Delta J = 0$)
- $\nu_i = \xi_i(t)(x_i - x_i^0) + \eta_i$ ... noise term with $\xi_i \in N(0, \sigma)$ and $\eta_i \in N(0, \bar{\sigma})$

The positivity condition $x_i \geq 0$ introduces a non-linearity:

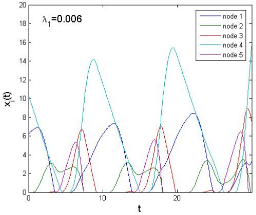
Dynamics can become chaotic:

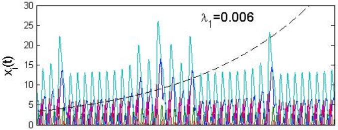
Idea:

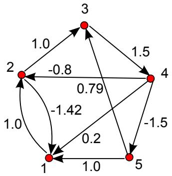
- $\lambda_1 < 0$ ... x_i converges to fixed point ($\lambda_1 \sim 0$)
- $0 \leq \lambda_1 < \varepsilon$ ... chaotic but non-exponentially growing x_i ($\lambda_1 \sim 0$)
- $\lambda_1 > \varepsilon$ ... exponentially growing x_i

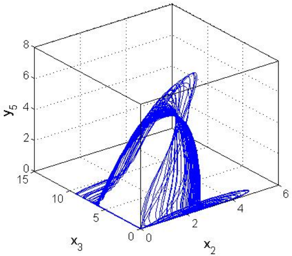
Chaos in ALS

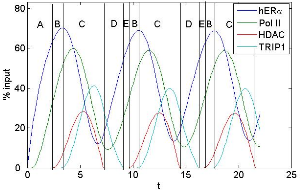
A weakly strange attractor of a 5-node network



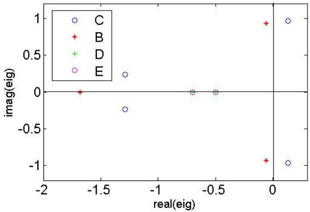








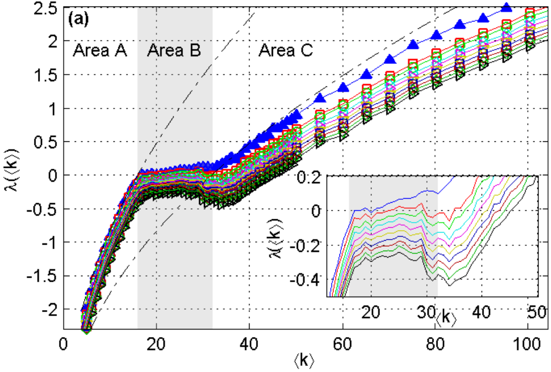
A
B
C
D
E
hERα
Pol II
HDAC
TRIP1
80
60
40
20
0
% input
0
5
10
15
20
25
t



C
B
D
E
1
0.5
0
-0.5
-1
imag(eig)
-2
-1.5
-1
-0.5
0
real(eig)

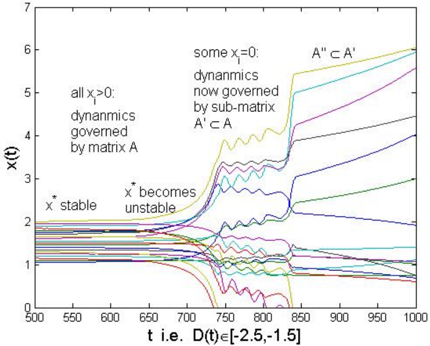
The edge of chaos

Lyapunov spectrum



Largest ten Lyapunov exponents (λ_p, $p = 1, \ldots, 10$) of the Lyapunov spectrum ($N = 500$). The two black dashed lines are theoretical curves – based on Girko's law – approximating $\lambda_1(\langle k \rangle)$ in the areas A and C. The intersection of these curves with the x-axis, $\lambda_1(\langle k \rangle) = 0$, estimate the beginning and end of the $\lambda_1(\langle k \rangle) \sim 0$ plateau (area B).

The edge of chaos
Range of decay rate



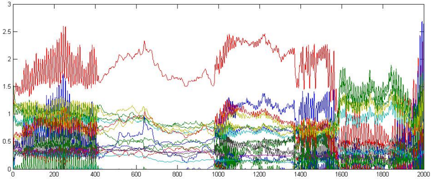
- While the decay rate rises some x_i become instable and are stopped at $x_i = 0$.
- The dynamics then is governed by the submatrix $A' \subset A$ which does not contain the rows and columns i for which $x_i = 0$.

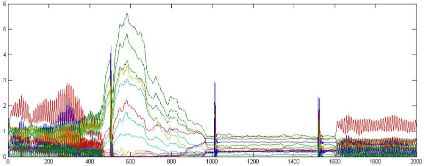
Conclusions

- ALS seem to provide a promising first approach to Genetic Regulation Networks.
- ALS possess an *inflated edge of chaos*.
- ALS are weakly chaotic.
- The dynamics of oscillating systems is governed by sequnces of submatrices $A_I \subset A$. A_I contains only indices $I \subset \{1, \ldots, n\}$.
- Stability can be understood by the alternating eigenvalue spectra of the A_I.

The End

FIN



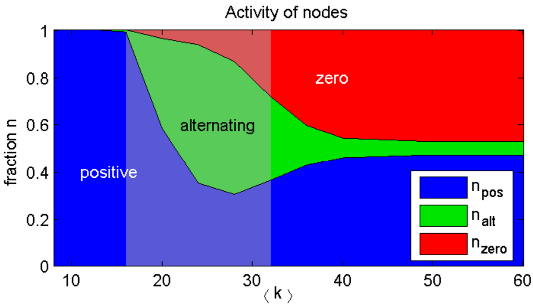


Typical time-serie; x-axis: time; y-axis: concentrations ($N = 30$, $\sigma = 0.005$ and 0.001, $dt = 0.1$, $\langle k \rangle = 5$, $D = 0.85$). Upper image: multi-stability; fluctuations occasionally causes mode-switching. Lower image: resilience of modes to random shocks; Shocks – just as fluctuations – switch modes by chance. Possibility: designing shocks for mode-switching.

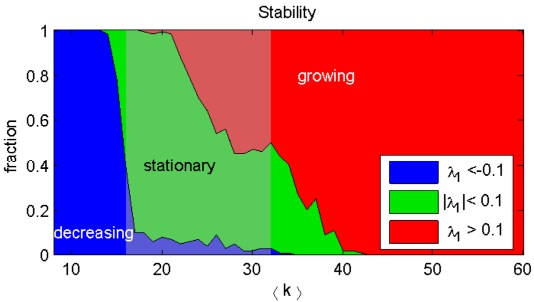
Dynamic Topology

The full and the active network

- In a network of size N there are N_{zero} nodes that always have $x_i = 0$.
- The active network only consists of N_{on} (active) nodes with non-zero x_i. *Active links* are links between active nodes.
- $N_{\text{on}} = N_{\text{pos}} + N_{\text{alt}}$
- $N = N_{\text{pos}} + N_{\text{alt}} + N_{\text{zero}}$



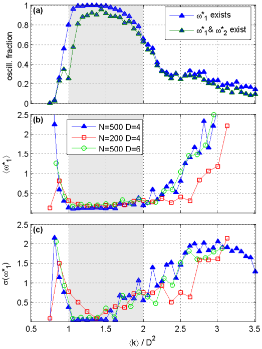
Average fractions of x_i: positive (n_{pos}), zero (n_{zero}), or alternating (n_{alt}). Averages are taken over 1000 realizations, time interval, $[500,\ 1000]$, $N = 500$, $D = 4$, $\sigma = \bar{\sigma} = 0$, $\sigma_A = 1$, $x^0 = 1000$.



Fraction of realizations which lead to exponentially growing ($\lambda_1 > 0.1$), decaying ($\lambda_1 < -0.1$) and stable time series ($|\lambda_1| \leq 0.1$) computed from 100 realizations, $N = 500$, $D = 4$, time interval, $[200, 1000]$, $\sigma = \bar{\sigma} = 0.1$, $\sigma_A = 1$, and x^0=1000.

Oscillatory Dynamics

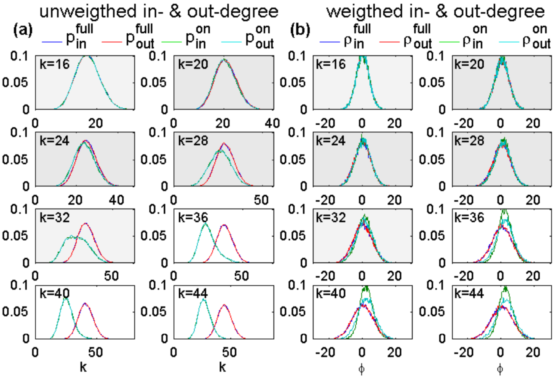
Frequencies



(a) Probability of finding oscillating realizations: existing fundamental frequency ω_1^* (blue) both, existing ω_1^* and ω_2^* (green). (b) Average ω_1^* as a function of $\langle k\rangle$. (c) Standard deviation of ω_1^*. $N = 500$, time interval, $[1000\ ,\ 3000]$, $D = 4$, $\sigma = \bar{\sigma} = 0$, $\sigma_A = 1$, $x^0 = 1000$. In (b) and (c) $N = 500$, $D = 6$, (green circles) and $N = 200$, $D = 4$ (red squares) are shown for comparison.

Active Network

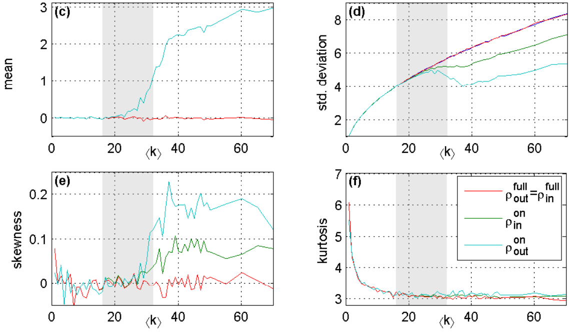
Degree distributions



(a) Unweighted in- and out-degree distributions of the active regulatory sub-network for various $\langle k \rangle$. Active in- and out-degree, $p^{\mathrm{on}}_{\mathrm{in/out}}(k)$, are practically indistinguishable. (b) Weighted in- and out-degree distributions. In- and out-weight distributions, $\rho^{\mathrm{on}}_{\mathrm{in/out}}(\phi)$, of active weights are clearly distinguishable. $\phi = \sum A_{ij}$ and the sum runs over i or j for in- and out-weight distribution, respectively.

Active Network

Degree distributions



(c) Mean, (d) standard deviation, (e) skewness and (f) kurtosis of the in/out-weight distributions. Differences between in- and out-weight distributions are found in the standard deviation and the skewness. Averages are taken over 50 realizations, $N = 500$, time interval, $[500,\ 1300]$, $D = 4$, $\sigma = \bar{\sigma} = 0.1$, $\sigma_A = 1$, $x^0 = 1000$.

Conclusions

- Stable (non-exponentially growing) dynamics dominant in plateau region and coincides with dominating number of alternating nodes
- Fundamental frequency almost certainly exists in plateau region
- Symmetry breaking in the in- and out- weight distributions exist. Yet, effect is not strong enough to explain topological differences of in- and out- degree distributions in real genetic regulatory networks alone

The End

FIN