COARSE-GRAINED DESCRIPTIONS OF A GENETIC CIRCUIT

D. Labavic[1], A. Garai[2], B. Waclaw[3], H. Nagel[4], H. Meyer-Ortmanns[1]

1: Jacobs University Bremen 2: UCL San Diego, 3: University of Edinburgh, 4:University of Leipzig

Abstract: We focus on a prototypic dynamical unit which consists of only two species interacting in a non-linear way. This unit may be regarded as a coarse-grained description of a genetic circuit . For comparison we first discuss the phase structure of this unit on a coarse-grained level, in a deterministic description. Then we turn to a fully stochastic description where we observe quasi-cycles for parameters that correspond to values deeply in the fixed-point regime in the deterministic limit [1].There we shall unravel the effect of demographic fluctuations and fluctuations in the reaction times. The power spectrum will show which source of stochastic behavior is dominant, in particular if the dynamics is very spiky. We compare analytic predictions with Gillespie simulations which come closest to experiments in vitro.

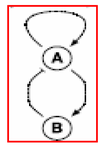
The genetic circuit has applications to all systems in which a self-activating species also activates its own repressor. Both interactions, the self-activation and the repression, need not necessarily be realized by direct links, but can amount to an effective description on a coarse scale with a different number of inter-mediate steps and different realizations of the very activation or repression. Intermediate steps, however, may introduce additional time scales. In our ongoing work [2] we therefore analyze the effect of competing time scales on the validity of the coarse-grained description. As it turns out, what is the appropriate model depends on the ratio of protein decay rates to binding/unbinding rates of transcription factors, leading to Different switching rates of genes to another expression level. This way the inherent time scales also determine the type of bifurcations which the circuit undergoes under variation of a certain parameter. For example, a whole regime with regular oscillations fades away if the gene states change over a time scale that is not short, but comparable to the lifetime of the involved proteins. This may give a hint on a possible origin for the malfunction of oscillatory systems like circadian clocks. Circadian clocks are just one example of systems that are supposed to be described by our genetic circuit.

Coupling these units to small network motifs, we know from our former results [3] that frustrated coupling can lead to multistable states and explain the observed multistability in synthetic gene circuits. So it is of much interest to analyze the effect of frustration in larger networks of these units.

I. Coarse-Grained Description of a Bistable Frustrated Unit

One bistable frustrated unit (S.Krishna, S. Semsey and M.H.Jensen, Phys.Biol.6 (2009))

$$\frac{dA}{dt} = \frac{\alpha}{1+(B/K)} \cdot \left(\frac{b+A^2}{1+A^2}\right) - A$$
$$\frac{dB}{dt} = \gamma(A-B),$$



A, B protein concentrations
γ ratio of half-life of A to that of B
K strength of the repression (of A by B)
α maximum rate of production of A (for full activation and no repression)
α b basal expression level of A

- may serve as basic building block in larger systems
- has its own rich phase structure
- has an intrinsic time scale (fast and slow variable)
- is "frustrated" on the basic level
- is realized in natural systems whenever bistable units are coupled to negative feedback loops, e.g. signalling system in the slime mold Dictyosthelium Discoideum, embryonic division control system, MAPK-cascade

I. On the coarse-grained level

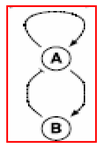
- in the deterministic realization, we analyze **the phase structure** as function of one control parameter to be characterized by excitable—oscillatory---excitable behavior
- in the stochastic realization, we search for **qualitatively new effects**: are there quasi-cycles or additional fixed points?
- We measure variances, autocorrelation functions and the power spectrum in order to **disentangle genuine limit cycles from quasi-cycles.**
- We observe large excursions in phase space (outside the perturbative regime).
- We emphasize the role of fluctuations in the reaction times for spiky dynamics.

II. For one possible realization on the fine-grained level

- We present results from Gillespie simulations of 6 master equations, and
- results of the derivation of the deterministic limit which depend on the kind of limit that is justified for "fast" genes, "slow" genes, and "ultraslow" genes.

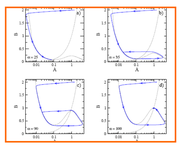
1. One bistable frustrated unit (S.Krishna, S. Semsey and M.H.Jensen, Phys.Biol.6 (2009)) **and its phase structure [3]**

$$\frac{dA}{dt} = \frac{\alpha}{1+(B/K)} \cdot \left(\frac{b+A^2}{1+A^2}\right) - A$$
$$\frac{dB}{dt} = \gamma(A-B),$$

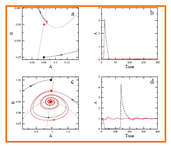


A, B protein concentrations
γ ratio of half-life of A to that of B
K strength of the repression (of B repressing A)
α maximum rate of production of A
α b basal rate

As function of α excitable, limit cycle, excitable behavior



zoom



2. Fully stochastic description of a bistable frustrated unit

Why at all?

- More realistic due to inherent stochasticity of various origin (finite number of species, biochemical reactions times, decay and birth processes happen in a stochastic way)

In general there may be **qualitative new effects** such as

- noise-induced ordering, noise-induced phase transitions, resonance phenomena
- new attractors arise: oscillations in space and time, or additional fixed points (pattern formation in ecological systems, Butler&Goldenfeld arXiv: 1011.0466, PRE (2009); for the brusselator see Boland, Galla& McKane J Stat Mech: Theory and Exp.(2008))

→ **quasi-cycles** in contrast to limit cycles

If the stochastic description goes along with a further zoom into the temporal resolution, there may be

→ **additional fixed points** as stable attractors (known from the toggle switch) (D.Schultz et al.PNAS(2008)H.Qian et al.PhysChemChemPhys (2009)).

Here: quasi-cycles, additional fixed points, new bifurcation scenarios

A= $\phi_A = N_A/N_0$ and $\phi_B = N_B/N_0$ =B

N0 parameterizes the system size, ranging from10 to 100000

Master equation:

$$\frac{\partial P(N_A,N_B)}{\partial t} = -(N_0 f(N_A/N_0, N_B/N_0) + N_A + \gamma N_A + \gamma N_B)P(N_A,N_B) \quad \text{loss}$$
$$+ (N_A+1)P(N_A+1,N_B)$$
$$+ N_0 f((N_A-1)/N_0, N_B/N_0)P(N_A-1,N_B)$$
$$+ \gamma(N_B+1)P(N_A,N_B+1) + \gamma N_A P(N_A,N_B-1). \quad (8) \quad \text{gain}$$

- van Kampen expansion in $N0^{-1/2}$ about the deterministic (infinite N0)-limit works away from the transition region and for short times, but also in the limit cycle regime.

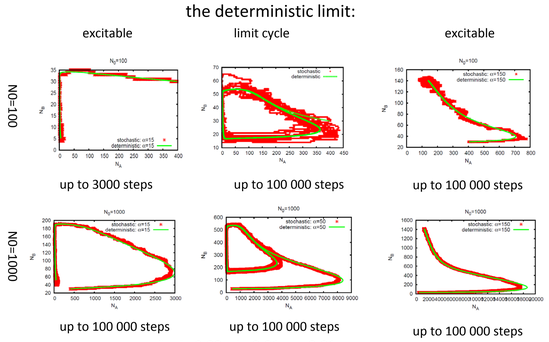
Corresponding reactions

production of A	$N_A \to N_A+1$	with rate $N_0 f(N_A/N_0, N_B/N_0)$
decay of A	$N_A \to N_A-1$	with rate N_A
production of B	$N_B \to N_B+1$	with rate γN_A
decay of B	$N_B \to N_B-1$	with rate γN_B

with $f(\phi_A,\phi_B) = \frac{\alpha}{1+\phi_B/K}\frac{b+\phi_A^2}{1+\phi_A^2}$.

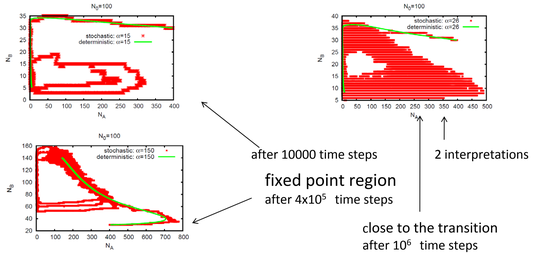
simulated with Gillespie simulations to obtain histograms

Gillespie trajectories show a similar phase structure as in the deterministic limit:



for k=0.02, b=0.01, γ=0.01

Quasi-cycles show up in Gillespie trajectories as "rare events" deeply in the former fixed point region:



for k=0.02, b=0.01, γ=0.01, N0=100

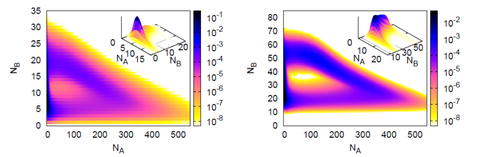


Figure 9. Density plots of the probability $P^*(N_A,N_B)$ in the stationary state, for $\alpha = 15$ (fixed-point regime, left) and $\alpha = 50$ (limit cycle, right), for $N_0 = 100$. Insets show peaks from small-N_A regions.

Note that we see always a donut-shape (due to quasi-cycles) and the variance of the probability distribution varies a lot along the cycle.

This is what we reproduced analytically in [1].

Power spectrum in the fixed-point regime with indications of small quasi-cycles

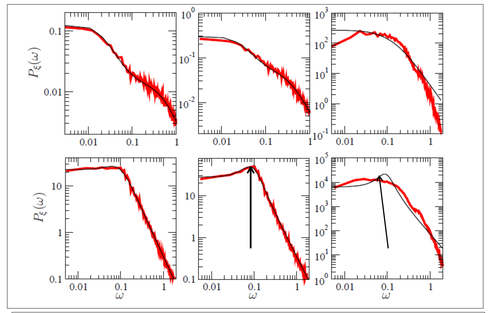


Figure 12. Comparison between the power spectrum $P_\xi(\omega)$ in the fixed point from Eq. (37) and obtained in simulations for $N_0 = 5000$ (top) and $N_0 = 10^5$ (bottom). Top left to right: $\alpha = 15, 20, 25$. Bottom left to right: $\alpha = 27, 28, 29$. see [1].

Power spectrum in the limit cycle phase and role of randomly distributed reaction times

We are interested in the stationary limit, but for a finite number of particles
$t \longrightarrow \infty$, $N_0 < \infty$
as realized in Gillespie simulations and biochemical reactions.

Observation: for long simulation times the width of the peaks first decreases and then stabilizes instead of shrinking to δ-peaks for all frequencies.

What to do instead? (B. Waclaw) Take into account that t is a random variable itself due to random reaction times dx= dt + W(t), x performing a Wiener process with

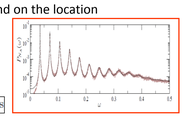
- taking into account that for fixed physical time the number of algorithmic steps or chemical reactions will fluctuate,
- neglecting fluctuations in N_A and N_B
- neglecting that the W(t) fluctuations in reaction time will depend on the location on the limit cycle, that is calculate the power spectrum as

$$P(x,t) = \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{-\frac{(x-\bar{x})^2}{2\sigma^2 t}}.$$

$$P_{N_A}(\omega) = N_0^2 \int_0^\infty dt \int_0^\infty dt' e^{i\omega(t-t')} \left\langle \phi_A(x(t))\phi_A(x(t')) \right\rangle_x$$

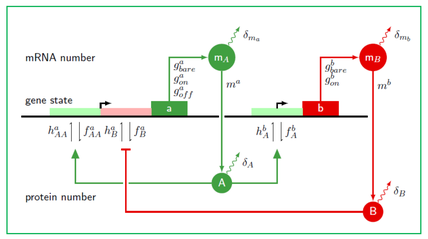
with the result of perfect agreement with Gillespie simulations

$N_0 = 4000, \gamma = 0.01$, and 2^{26} Gillespie steps averaged over 100 realisations



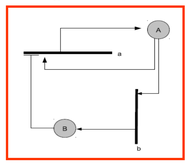
II. Fine-Grained Description: Zoom into the Genetic Circuit

From genes a, b ⟶ to mRNA m_A, m_B ⟶ to proteins A, B



h^a_{AA} = dimer binding rate of transcription factor A to promoter region of gene a
f^a_{AA} = corresponding unbinding rate etc.
δ_A, δ_B = decay rates of proteins A and B
$g^a_{bare,on,off}$ = production rate of mRNA of type A or B if gene a is in on, bare, off state
$g^b_{bare,on,}$ = production rate of mRNA of type A or B if gene b is in on or bare state

From genes to proteins



binding/unbinding reactions

$$\begin{cases} a_{bare}+2A \xrightarrow{h^a_{AA}/N_0^2} a_{on} \\ a_{on} \xrightarrow{f^a_{AA}} a_{bare}+2A \\ a_{bare}+B \xrightarrow{h^a_B/N_0^2} a_{off} \\ a_{off} \xrightarrow{f^a_B} a_{bare}+B \\ b_{bare}+A \xrightarrow{h^b_A/N_0} b_{on} \\ b_{on} \xrightarrow{f^b_A} b_{bare}+A \end{cases}$$

production and decay of proteins

$$\begin{cases} a_{on} \xrightarrow{g^a_{on}N_0} A+a_{on} \\ a_{bare} \xrightarrow{g^a_{bare}N_0} A+a_{bare} \\ a_{off} \xrightarrow{g^a_{off}N_0} A+a_{off} \\ A \xrightarrow{\delta^A} \phi \\ b_{on} \xrightarrow{g^b_{on}N_0} B+b_{on} \\ b_{bare} \xrightarrow{g^b_{bare}N_0} B+b_{bare} \\ B \xrightarrow{\delta^B} \phi \end{cases}$$

Stochastic description in terms of 6 master equations summarized as

$$\frac{dP_{ij}(N_A,N_B;t)}{dt} = -(g^a_0 N_0 + \delta^A N_A + g^b_0 N_0 + \delta^B N_B)P_{ij}(N_A,N_B)$$
$$+ g^a_0 N_0 P_{ij}(N_A-1,N_B) + \delta^A(N_A+1)P_{ij}(N_A+1,N_B)$$
$$+ g^b_0 N_0 P_{ij}(N_A,N_B-1) + \delta^B(N_B+1)P_{ij}(N_A,N_B+1)$$
$$+ h^a_A(i)\frac{N_A-1}{N_0}P_{bare,j}(N_A,N_B) - f^a_A(i)P_{on,j}(N_A,N_B)$$
$$+ h^a_B(i)\frac{N_B}{N_0}P_{bare,j}(N_A,N_B) - f^a_B(i)P_{off,j}(N_A,N_B)$$
$$+ h^b_A(j)\frac{N_A}{N_0}P_{i,bare}(N_A,N_B) - f^b_A(j)P_{i,on}(N_A,N_B)$$

with $P_{i,j}(N_A,N_B,t)$ = prob to find N_p proteins of type P if gene a is in state i and gene b in state j

$$\frac{d\langle(N_A)\,a_{on}\rangle}{dt} = g^a_{on}N_0 a_{on} + h^a_{AA}\frac{\langle N_A\rangle^2}{N_0^2}a_{bare} - f^a_{AA}\langle N_A\rangle a_{on} - \delta^A\langle N_A\rangle a_{on} \quad (1)$$

$$\frac{d\langle(N_A)\,a_{off}\rangle}{dt} = g^a_{off}N_0 a_{off} + h^a_B\langle N_A\rangle\frac{\langle N_B\rangle^2}{N_0^2}a_{bare} - f^a_A\langle N_A\rangle a_{off} - \delta^A\langle N_A\rangle a_{off} \quad (2)$$

$$\frac{d\langle(N_A)\,a_{bare}\rangle}{dt} = g^a_{bare}N_0 a_{bare} - \delta^A\langle N_A\rangle a_{bare} - h^a_{AA}\frac{\langle N_A\rangle^2}{N_0^2}a_{bare} + f^a_{AA}\langle N_A\rangle a_{on}$$
$$- h^a_B\langle N_A\rangle\frac{\langle N_B\rangle^2}{N_0^2}a_{bare} + f^a_A\langle N_A\rangle a_{off} \quad (3)$$

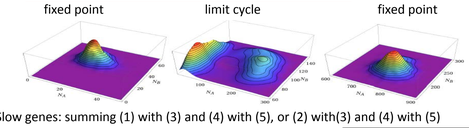
$$\frac{d\langle(N_B)\,b_{on}\rangle}{dt} = g^b_{on}N_0 b_{on} - \delta^B\langle N_B\rangle b_{on} + h^b_A\frac{\langle N_A\rangle}{N_0}\langle N_B\rangle b_{bare} - f^b_A\langle N_B\rangle b_{on} \quad (4)$$

$$\frac{d\langle(N_B)\,b_{bare}\rangle}{dt} = g^b_{bare}N_0 b_{bare} - \delta^B\langle N_B\rangle b_{bare} - h^b_A\frac{\langle N_A\rangle}{N_0}\langle N_B\rangle b_{bare} + f^b_A\langle N_B\rangle b_{on} \quad (5)$$

Fast genes: summing (1)-(3) and (4)-(5) with stationary values for the genes ⟶

$$\frac{d\langle N_A\rangle}{dt} = N_0\frac{g^a_{bare} + g^a_{on}x^a_{AA}\frac{\langle N_A\rangle^2}{N_0^2} + g^a_{off}x^a_B\frac{\langle N_B\rangle^2}{N_0^2}}{1+x^a_{AA}\frac{\langle N_A\rangle^2}{N_0^2}+x^a_B\frac{\langle N_B\rangle^2}{N_0^2}} - \delta^A_p\langle N_A\rangle$$
$$\frac{d\langle N_B\rangle}{dt} = N_0\frac{g^b_{bare} + g^b_{on}x^b_A\frac{\langle N_A\rangle}{N_0}}{1+x^b_A\frac{\langle N_A\rangle}{N_0}} - \delta^B_p\langle N_B\rangle \quad \text{where } x^m_n = \frac{h^m_n}{f^m_n}.$$

fixed point | limit cycle | fixed point



Slow genes: summing (1) with (3) and (4) with (5), or (2) with(3) and (4) with (5)

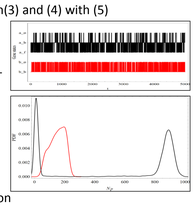
1w3 4w5

$$\begin{cases} \frac{d\langle N_A\rangle}{dt} = N_0\frac{g^a_{bare}+g^a_{on}x^a_{AA}\frac{\langle N_A\rangle^2}{N_0^2}}{1+x^a_{AA}\frac{\langle N_A\rangle^2}{N_0^2}} - \delta^A_p\langle N_A\rangle \\ \frac{d\langle N_B\rangle}{dt} = N_0\frac{g^b_{bare}+g^b_{on}x^b_A\frac{\langle N_A\rangle}{N_0}}{1+x^b_A\frac{\langle N_A\rangle}{N_0}} - \delta^B_p\langle N_B\rangle \end{cases}$$

2w3 4w5

$$\begin{cases} \frac{d\langle N_A\rangle}{dt} = N_0\frac{g^a_{bare}+g^a_{off}x^a_B\frac{\langle N_B\rangle^2}{N_0^2}}{1+x^a_B\frac{\langle N_B\rangle^2}{N_0^2}} - \delta^A_p\langle N_A\rangle \\ \frac{d\langle N_B\rangle}{dt} = N_0\frac{g^b_{bare}+g^b_{on}x^b_A\frac{\langle N_A\rangle}{N_0}}{1+x^b_A\frac{\langle N_A\rangle}{N_0}} - \delta^B_p\langle N_B\rangle \end{cases}$$

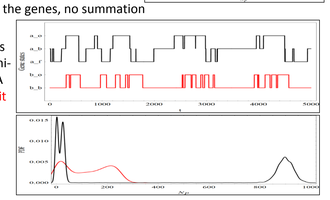
2x1 fixed points in the deterministic limit for NA and NB, no limit cycle.



Ultraslow genes: no steady states for the genes, no summation

$$\frac{d\langle N_A\rangle}{dt} = g^a_{on}N_0 - \delta^A\langle N_A\rangle$$
$$\frac{d\langle N_A\rangle}{dt} = g^a_{bare}N_0 - \delta^A\langle N_A\rangle$$
$$\frac{d\langle N_A\rangle}{dt} = g^a_{off}N_0 - \delta^A\langle N_A\rangle$$
$$\frac{d\langle N_B\rangle}{dt} = g^b_{on}N_0 - \delta^A\langle N_B\rangle$$
$$\frac{d\langle N_B\rangle}{dt} = g^b_{bare}N_0 - \delta^A\langle N_B\rangle$$

3x2 fixed points in the deterministic limit for NA and NB, no limit cycle.



Summary:

The bistable frustrated unit is a challenging motif with an intrinsic time scale. We see quasi-cycles, the more the larger the fluctuations, the smaller N0 and the larger γ.

Stochastic case for a single unit in the coarse-grained description:

- (Large) fluctuations induce quasi-cycles even deeply in the former fixed point region.
- Quasi-cycles can be disentangled from limit cycles via the autocorrelation function.
- Quasi-cycles make the BFU even more flexible so that no fine-tuning is needed for having oscillations, but are both oscillations really equivalent also when these units are coupled?
- Large excursions in phase space cannot be treated within the van Kampen expansion
- Random reaction times are the most important source of stochasticity.

Challenge: Identify quasi-cycles in natural oscillatory genetic systems. What do they serve for? What is the "normal" mode of performance?

The effective deterministic models along with their bifurcation patterns (I) depend on the inherent time scales on the fine-grained level. This is seen when these models are derived via taking appropriate averages over processes, which are fast as compared to those in the main focus. Here additional fixed points show up and regular limit cycles disappear. So coarse-grained descriptions should be derived with care.

[1] A. Garai, B. Waclaw, H. Nagel and H. Meyer-Ortmanns, J. Stat. Mech. (2012) P01009, 28 pages. [2] D. Labavic, H. Nagel, W. Janke, H. Meyer-Ortmanns, in preparation. [3] P. Kaluza, and H. Meyer-Ortmanns: Chaos (2010) 20:043111, & Virtual Journal of Biological Physics Research -- November 15, 20 (10) (2010).