Modelling biological timing and its robustness

1. Introduction

To perceive time is one of fundamental qualities of system of any purpose. People measure time in seconds, while planets measure time in years or higher. Time-scale is often associated with repeated behaviours, or oscillations. Many chemical or biological systems are known to oscillate for a long time, such as the oscillatory Belousov–Zhabotinsky reaction, glucose oscillation in cells, cAMP oscillations, Calcium oscillation. While timescales and oscillations have been heavily exploited in signaling and information theory[1], little analogy has been drawn in the scope of biology, despite the increasing amounts of examples, such as circadian rhythm, yeast cell-cycle, and synthetic oscillatory genetic circuits like repressilator[2]. Nevertheless, in absence of a uniting framework, important observations have been made with bifurcation theory and stochastic modeling.

Here we review two examples, the yeast cell-cycle network and the repressilator. Since the yeast cell-cycle is a complicated network, we do not address its particularities, but merely present it as an important observation made through bifurcation theory.

2. Cell-cycle as an example for deterministic model

We implemented “the generic cell cycle”(GCC) model from [3], and tested the cdc20- and the cdh- mutants.

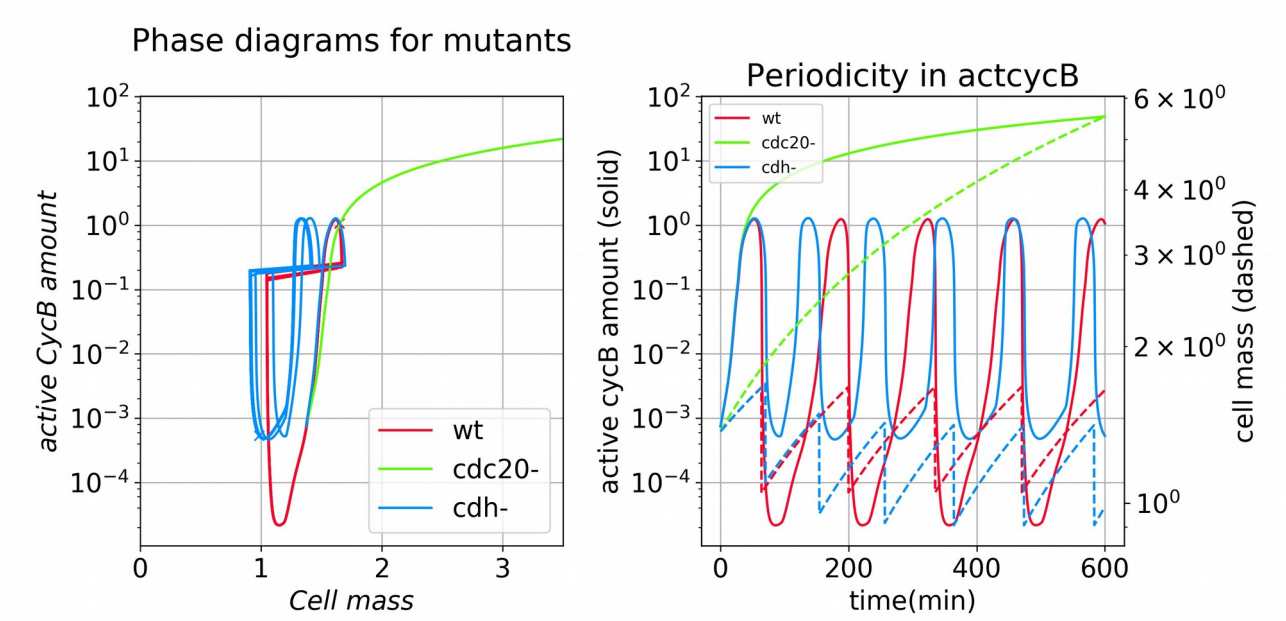


Figure 1: Testing the S. cerevisiae version of GCC model. Parameters: cdc20-:ks20p=0;ks20pp=0; cdh-:kah1p=0;kah1pp=0;

The qualitative behaviour captured includes (1) periodic changes in concentration of cyclins and cdk’s, well known for regulating cell-cycle progression. (2) phase progress in terms of changes in stability of steady-states/limit-cycles. e.g.: As the cell-size (modelled by cell mass) increases in S/G2, the previous steady state occupied by loses its stability in a saddle-node infinite-period bifurcation (SNIPER) to produce a stable limit cycle. The cell then proceed along the limit-cycle as a train follows a track, and divide according to the signal set by cyclins. It also captures the small cell-size at division of $cdh1^-$ mutant (figure 1), which is modelled by blocking the activation of cdh1 by cdk.

In experiments, cell-cycle arresting is a useful technique for achieving population synchronisation. For exmaple, alpha-pheromone is used to induce cell cycle arrest in yeast, which are then washed away after all cells reached the same stage. Quorum sensing(QS) is a yet another way to achieve synchronisation. Contrast to the blocking mechanism where cell-cycle is inhibited with existence of alpha-pheromone, the spike cycle in QS is inhibited with absence of an auto-inducer. Once auto-inducer reaches a certain threshold or alpha pheromone is diluted away, all cells suddenly start a program in unison. Since the equilibration of alpha-pheromone or introducer is limited only by diffusion, there is little spatial heterogeneity in the phases(the time when a program is started).

However, QS coupling of limit-cycle oscillators in bacteria is reported to produce significant spatio-temporal heterogeneity[4], on the timescale of the diffusion, where the macroscopic wave-speed V is proportional to the square-root of the diffusion constant of auto-inducer D ($V \sim D^{1/2}$), in line with the assertion that synchrony offered by QS is limited by the diffusion of auto-inducer, giving rise to phenomena observable only on a scale different from the typical cellular environment. Such link between scales serves as an excellent example of synthetic emergence. Reciprocally, we assert that a blocker like alpha-pheromone in an oscillatory background is also capable of inducing such emergence at diffusion level, though the alpha-pheromone itself may be constrained by long feedback time and other stochasticity. We intend to model such emergence in the future.

We note spontaneous synchronisation without distributing a blocking signal like alpha-pheromone is a long-standing topic in dynamical systems, and people have shown weak coupling is enough to result in synchronisation Kuramoto model [5]. In biology, heart cells synchronise in beats using some transient coupling[6].

Bringing cell to synchrony also seems to improve the robustness of the individually sloppy oscillators. More precisely, we assert that phase drift phenomena introduced later is be drastically reduced in a QS-coupled population, and intend a simulation.

3. The deterministic and stochastic repressilator

Due to the limited length for this essay and the heavy bottom-level coding. Our efforts will focus on the introduction of stochasticity into deterministic models.

As indicated earlier, the deterministic model captures the qualitative asymptotic feature of a system. To simplify the analysis, we seek a simple model. Contrast to the complex aforementioned naturally evolved circuits, synthetic circuits are much simpler and more amenable to detailed analysis, whereas retaining the desirable functionality of natural circuits. Repressilator is one of the first artificial genetic circuit capable of asymptotic oscillation. It consist of three chained transcription inhibitors/repressors, namely LacI-LacO, TetR-TetO, and CI-CIO (inhibitor-promoter). We denote the loop as [CIO:LacI----| LacO:TetR----|TetO:CI----| CIO:…] (promoter:protein, inhibitor----|promoter). LacI inhibits the activity of LacO, which in turn controls TetR production. The three genes form an inhibitory loop (figure 2). An odd loop length ensures the loop can be coarse-grained to a delayed negative feedback from any species to itself.

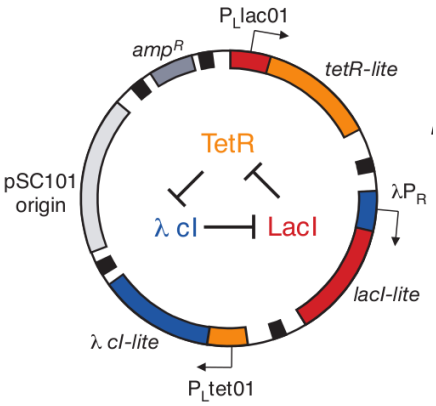


Figure 2: The original repressilator plasmid, where P_R is equivalent to CIO. Adapted from [7]

The system is abstracted into 6 variables in a minimal model, namely the mRNA number and the protein number of the three genes. Several assumptions have been made for simplicity. (1) The mRNA production rate is proportional to number of active promoters. (2) The protein production

rate is proportional to mRNA number. (3) Both mRNA degradation and protein degradation are Poisson processes, associated with different rate constants. (4) The inhibition of promoters by repressors are modelled by Hill functions, reflecting the cooperative inhibition. The governing equation can be found in the Appendix.

Using the parameters from [7], we reproduce a limit-cycle with a period of ~110min and an amplitude of ~3700 p_LacI (We use LacI protein number to measure amplitude), on the same order the reported 160±40min. To allow for modifcation regarding plasmid copy number and sponge promoters, we explicitly modelled the inhibition of the promoter with mass-action kinetics.

$$P + 2R \underset{k_{off}}{\overset{k_{on}}{\rightleftarrows}} P:2R$$

where $K^n = \frac{k_{off}}{k_{on}}$ (Note the off here refers to repressor binding on/coming off the promoter).
(Half-life of [P:2R]) = 318s, thus k_{off}=1/318s^{-1}, from [8]

Correspondingly, we re-manifest our assumption(1) and (4) to give rise to the following equations. The model is qualitatively unchanged with minor quantitative difference.

To illustrate the structure of the limit cycle, we fixed the m_LacI (mRNA of LacI) level to an arbitrary constant and record the steady state, effectively breaking the circular topology. There are two possible steady states, one at low m_LacI level and one at high m_LacI level. The first is associated with low CI and high TetR, the second associated with high CI and low TetR. The switching took place at m_LacI = $10^{0.4}$=2.5. Therefore we argue m_LacI level modulates the interplay between p_CI and p_TetR, effecively acting as a toggle-switch between a CI-dominating state and a TetR-dominating state.

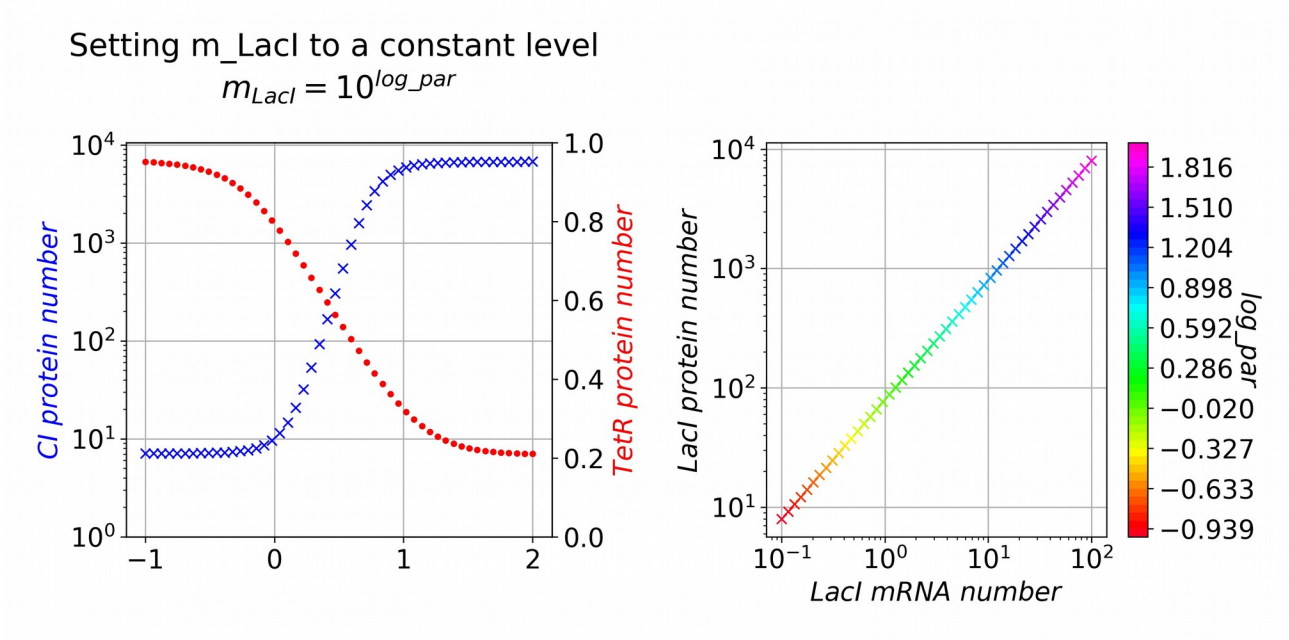


Figure 3: Steady states for any arbitrary m_LacI level. Note the sigmoid response curve

To further illustrate the usefulness of the deterministic model, we probe the response of the limit-cycle towards: (1) plasmid copy number (cpnum) (2) IPTG activation (IPTG) (3) TetR:TetO affinity (K_TetR) and sponge promoter (cpnum_TetO) (4)Protein degradation (k_dp) and mRNA degradation (k_dm). For each parameter set, we record the morphology of limit cycle in an 2D phase diagram, and note the corresponding amplitude and period. We note that these simulation ignore the stochastic elements in all variables and assume perfect certainty.

(1) The plasmid copy number manifest itself in sizing the promoter concentration, since promoter number is proportional to plasmid number. We resize the mRNA transcription rate from k_m to $k_m = k_{m\_base} \cdot cpnum$, where k_{m_base} is obtained by assuming cpnum=5 in the original model, where repressilator is hosted on pSC101 plasmid with a tightly controlled copy number of 5 per cell[7,9]. We note there exist mutants of pSC101 where copy number is higher, but may be less tightly regulated[10].

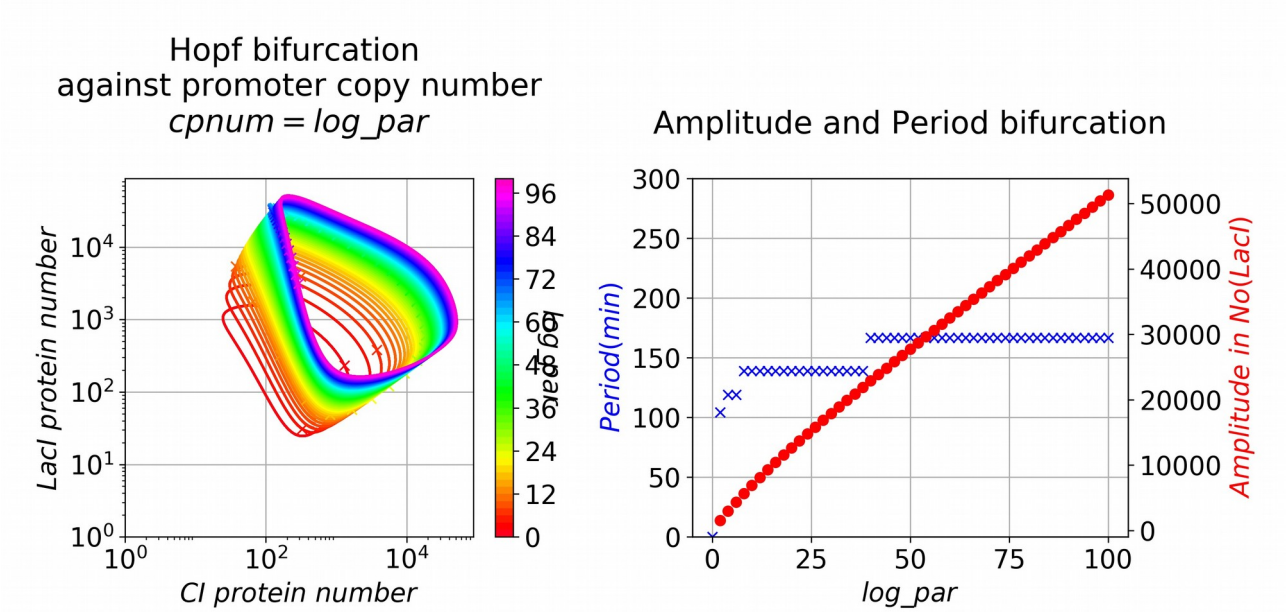


Figure 4: Bifurcation diagram for plasmid copy number. There is a presumably Hopf bifurcation at cpnum=0.

Upon change of cpnum between 0 and 96, the period remains mostly unchanged while the amplitude is approximately proportional to plasmid copy number (figure 4), which should be easy to observe unless masked. At a high copy number, the minimum protein number is also higher, and may reduce error due to random discrete protein events[7,11]. Stochastic modelling may be sought to further analyse its effect on robustness. We note LPT asserted that poorly regulated high-copy reporter plasmid caused heterogeneity in fluorescence, and resolved the issue by moving the reporter gene to pSC101 plasmid.

(2) IPTG is a common experimental reagent for inducing LacO transcription by inhibiting LacI's function. EQU4. In the model, the free p_LacI is reduced in a Hill function fashion.

$LacI_{eff} = \frac{LacI_{free}}{(1 + \frac{IPTG}{K_{IPTG}})^2}$, where K_{IPTG}= 6.0E6 au = 6mM [12]

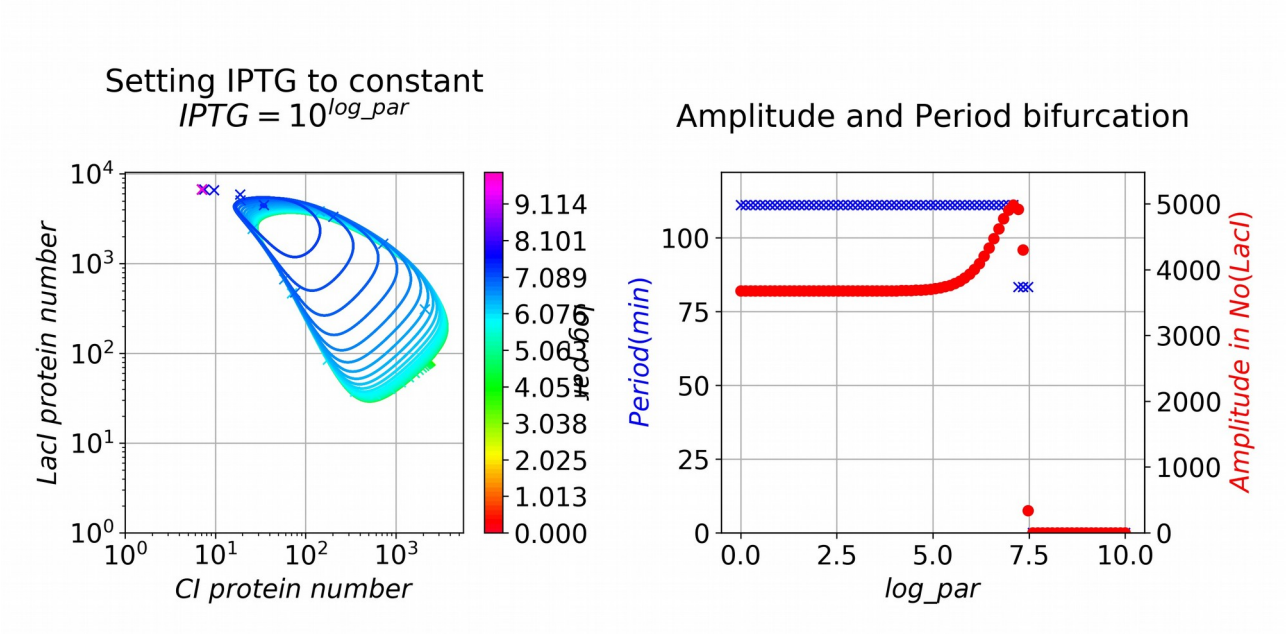


Figure 5:Bifurcation diagram for [IPTG]

As IPTG number increases, the minimum CI decrease. The limit cycle is suddenly lost near IPTG=$10^{7.5}$=~30 K_{IPTG}, with no change in period. We assert it is a Hopf bifurcation. After limit cycle is lost, the system rests at low CI, high TetR, high LacI steady state, as the LacI inhibition of LacO:TetR is abolished. This state is not visited by a normal limit cycle, and may cause heterogeneity when IPTG is removed.

In normal cell states, the LacI operon is controlled under allolactose and not under IPTG. The wild-type Lac operon activity is regulated by free LacI repressor, which is in turn regulated by the galactose level, whereas production of galactose from allolactose is regulated by β-galactosidase. In other words, the wt LacO operon is positively regulated by environmental lactose level, just as Lux operon is positively regulated by LuxI level. The difference is however that lactose is an energy source that is actively metabolised by bacteria and lactose has to be consumed to fuel the activation of Lac operon, whereas LuxI:LuxR fulfils the activation without consumption. How to modify the LacO operon to achieve quorum sensing will be an interesting open question. One possiblility is to engineer a membrane-bound version of beta-galactosidase.

(3) TetR:TetO affinity controls the stability of 2TetR:TetO inhibition complex. We assume the half-life of TetR:TetO is same as LacI:LacO, whereas production rate of TetR:TetO is higher, and hence a lower K_TetO give a higher k_on_TetO.

$k_{on\_TetO} = \frac{k_{off}}{{K_{TetO}}^n}$ where n = 2.1 is the hill coefficient.

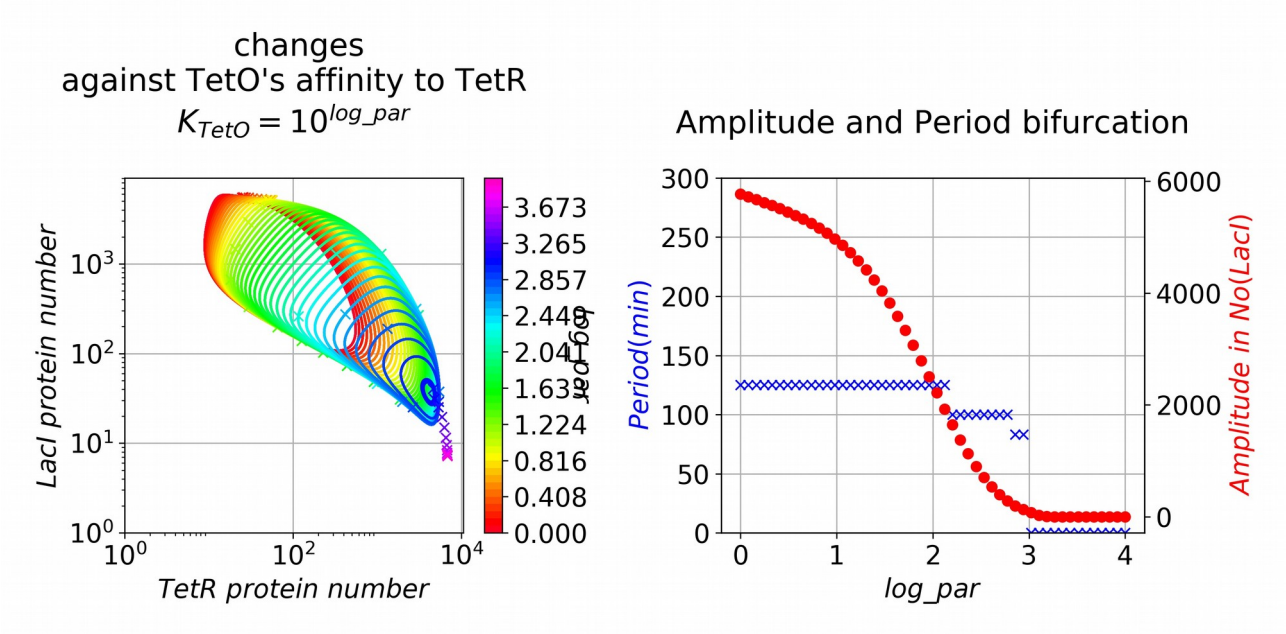


Figure 6: Bifurcation diagram for K_{TetO}

As K_{TetO} decreases, the amplitude become larger with little change in period. Importantly, at low K_{TetO}, the TetR number has to drop to several to allow progress into next phase (figure 6). As LPT asserted, this will amplify the stochasticity due to the breaking of mass action principle, and can be resolved by including more TetO promoter.

We modify TetO promoter number from $cpnum$ to $(cpnum + cpnum_{TetO})$, whereas only cpnum promoters are associated with CI production.

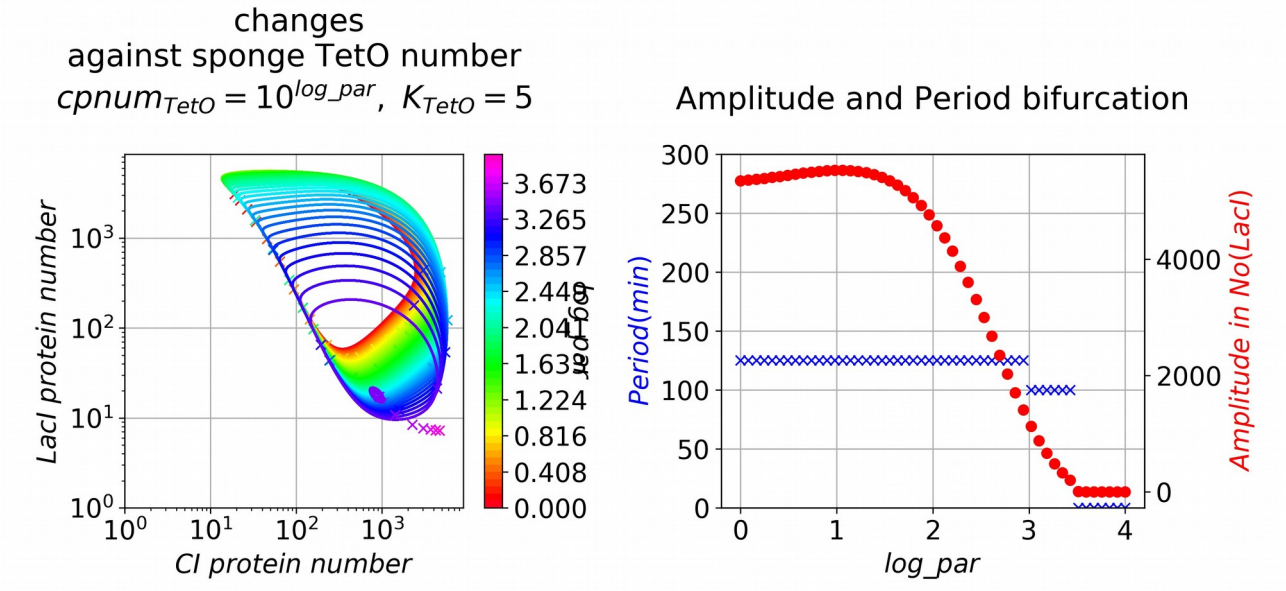


Figure 7: Bifurcation diagram for number of sponge TetO promoters

Note the bifurcation diagram for $cpnum_{TetO}$ is remarkably similar to that for K_{TetO}. In other words, low sponge promoter number and low K_{TetO} has similar phenomenological readouts. We continue to vary both parameters to produce a 2D bifurcation diagram, and argue there is certain

symmetry in the amplitude heatmap (figure 8) and an decrease in K_{TetO} can be compensated by an increased $cpnum_{TetO}$, which avoid low TetR number on phase transition and reduce stochasticity.

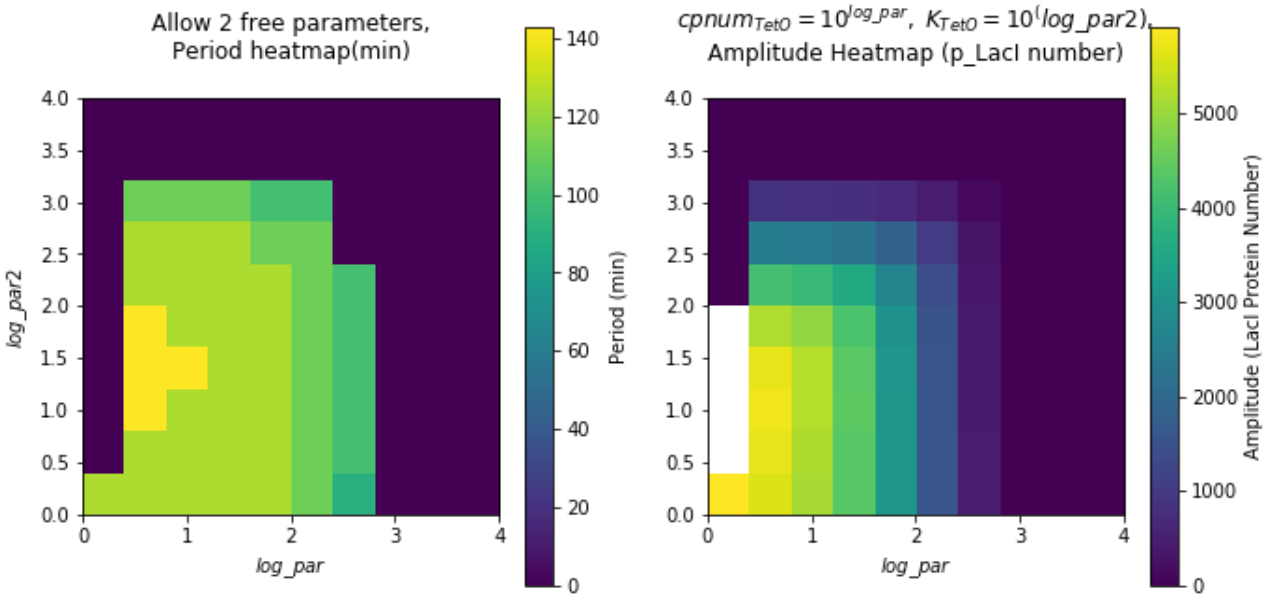
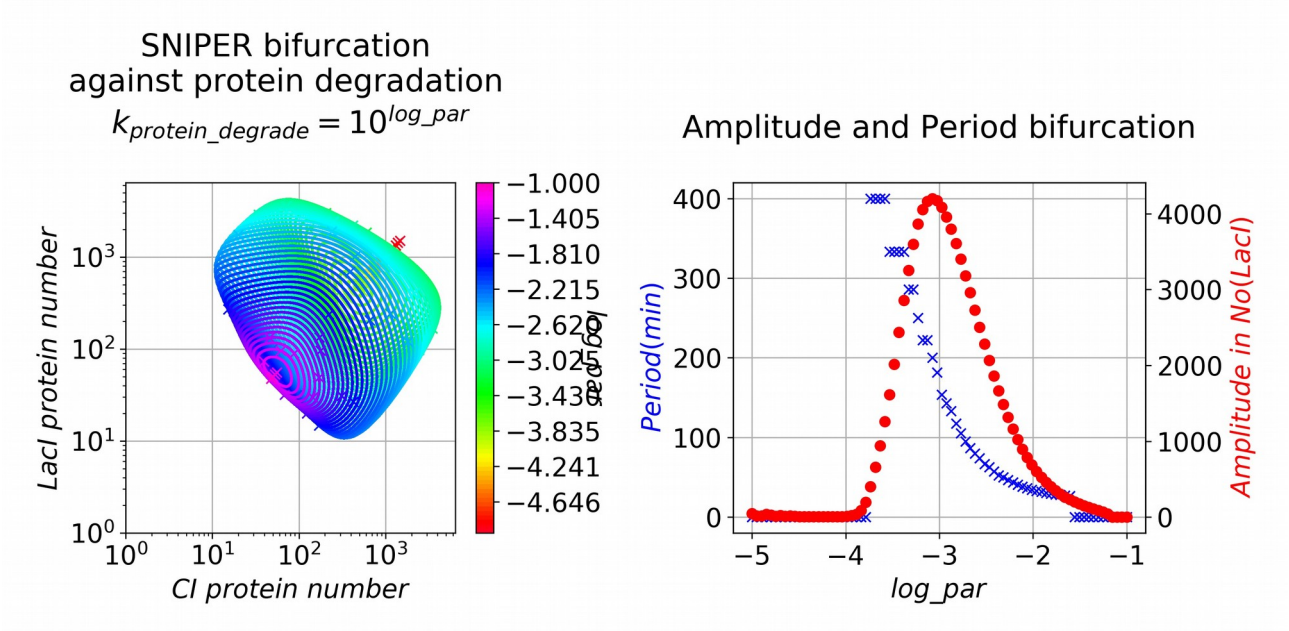


Figure 8: Two-parameter bifurcation, x = log(sponge promoter number), y = log(K_TetO)

4) Protein degradation (k_{dp}) and mRNA degradation (k_{dm}). Varying k_{dp} and k_{dm} is straightfoward.



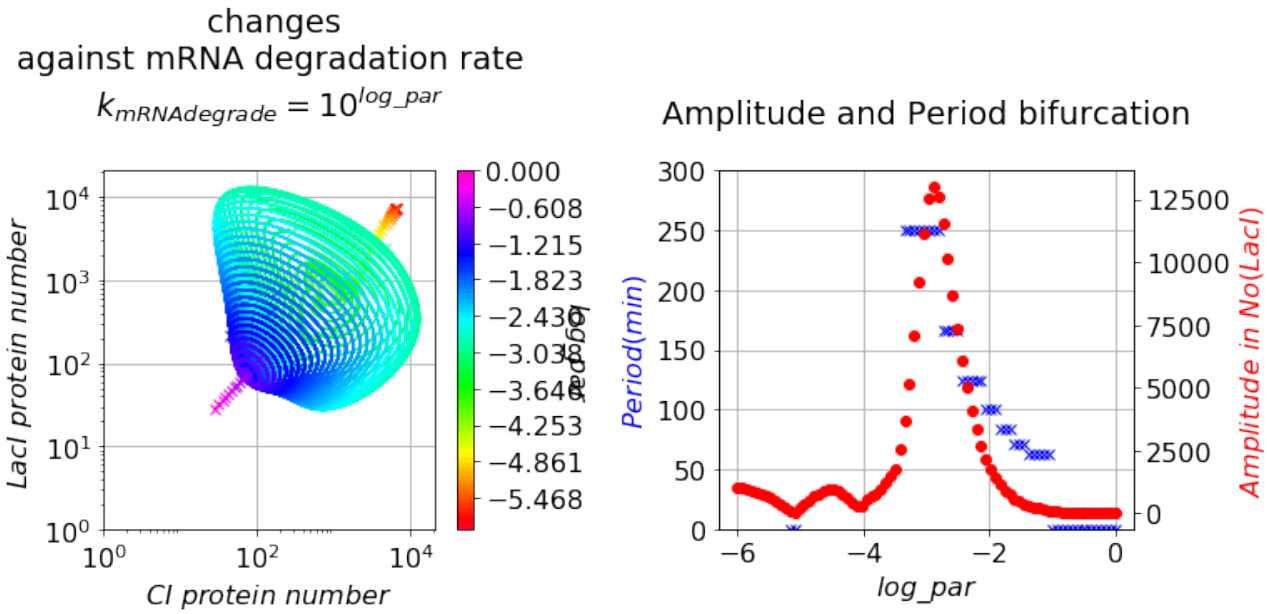


Figure 9: Bifurcation diagrams for protein degradation rate (top) and for mRNA degradation rate (bottom)

Interestingly, the amplitude does not change monotonically as degradation rate changes, but takes the shape of a spike (figure 8). At high degradation rate, the period is extremely small and close to zero. As degradation slows down, both amplitude and period increase. The amplitude peaks at a critical degradation rate (k_{crit}) and start to decreases afterwards. The left and right of k_{crit} may correspond to the relaxation and harmonic regimes correspondingly, as postulated in [11].

As shown above, deterministic models suggest how the macroscopic properties like period and amplitude may change as the underlying parameter varies. In reality, however, such smooth change of parameter is rarely possible. For a genetic mutant like cdh$^-$, the rate constant of some interaction like cdh activation is set to zero directly, essentially representing a discrete jump in parametric space. In biology, one is faced with heterogeneity and discreteness, that is uncrackable with bifurcation theory that involves intrinsic smoothness. That said, one can introduce heterogeneity into a deterministic model by replacing a single point in parametric space with a probability distribution, and the evolution of this PDF can be modelled with a deterministic operator. In Gillespie algorithm, one repeatedly sample the underlying PDF and apply a stochastic operator to estimate the trajectory of the PDF, though this process may well be approximated with stochastic calculus.

For problems as complex as the repressilator, not amenable to analytical methods, however, simulations become the only way to abstract information from the model. Here we apply stochastic simulation to analyse the phase drift phenomena observed in the repressilator[11]. Instead of assuming a constant rate of phase change, a random phase change rate is assumed and modelled as a Poisson process, given that the phase is prohibited to go backwards in oscillation. The implementation is implicit in Gillespie algorithm, where mass action kinetics is replaced with multiple Poisson processes, the sum of which emerge to be a Poisson process in the phase space in the vicinity of the deterministic limit-cycle.

The downside of stochastic modelling is that enormous sampling is required before the trajectory of PDF converges. To reduce the computational cost, we control step-size by pooling random events together until the time-elapsed reach a specified amount, and apply pooled changes simultaneously, based on the assumption that the system is approximately the same during this step-size. This approximation becomes invalid when a single discrete change corresponds to a big leap in the

phase, which might be true during relaxation[11]. In practice, we observe As long as the step size is much smaller than the period of limit-cycle, the simulation result does not change much.

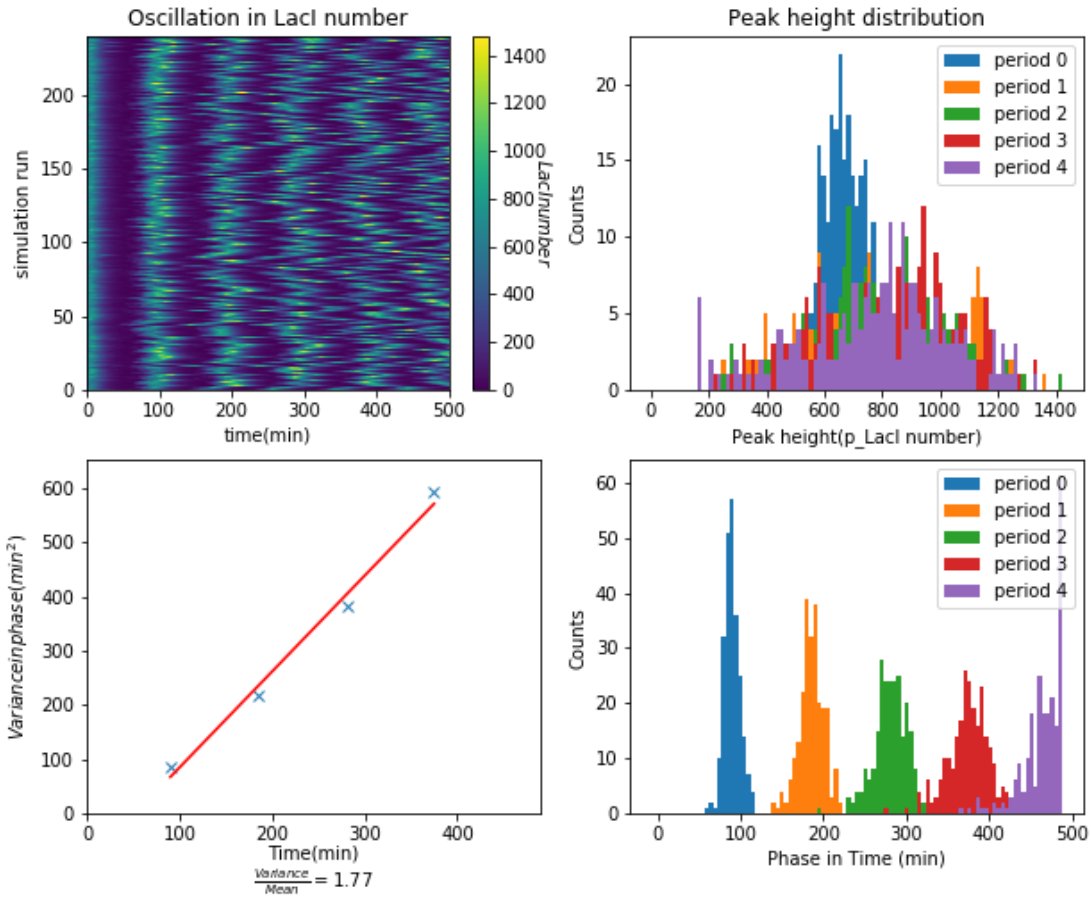


Figure 10:Stochastic modelling at k_dm=0.00577s^{-1} Top-Left:p_LacI levels form different simulation runs. Top-right: Amplitude PDF at timepoints Bottom-Left: Variance-Time plot. Bottom-right: Phase PDF at timepoints

We first simulated the modified model at original parameters and got a period of ~100min. We then record the time-point and amplitude at which p_LacI peaked and plotted accordingly. The simulation result is in line with the report that variance in phase increases linearly in time [11] (figure 9), at a rate of ~10min. The amplitude distribution experiences a flattening process, presumably due to the simulation being started a point outside the limiting distribution.

We then test whether the variance increasing rate(VIR) is dependent on the model parameters by setting mRNA degradtion to 0.0010s^{-1}. The results show a higher VIR at 22.32min, indicating a faster phase drifting. At this parameter, many points also lagged 1 period behind the majority, causing the period PDF to become bimodal, presumably increasing the variance. Whether a changed VIR can be experimentally observed, is an open question yet to test.

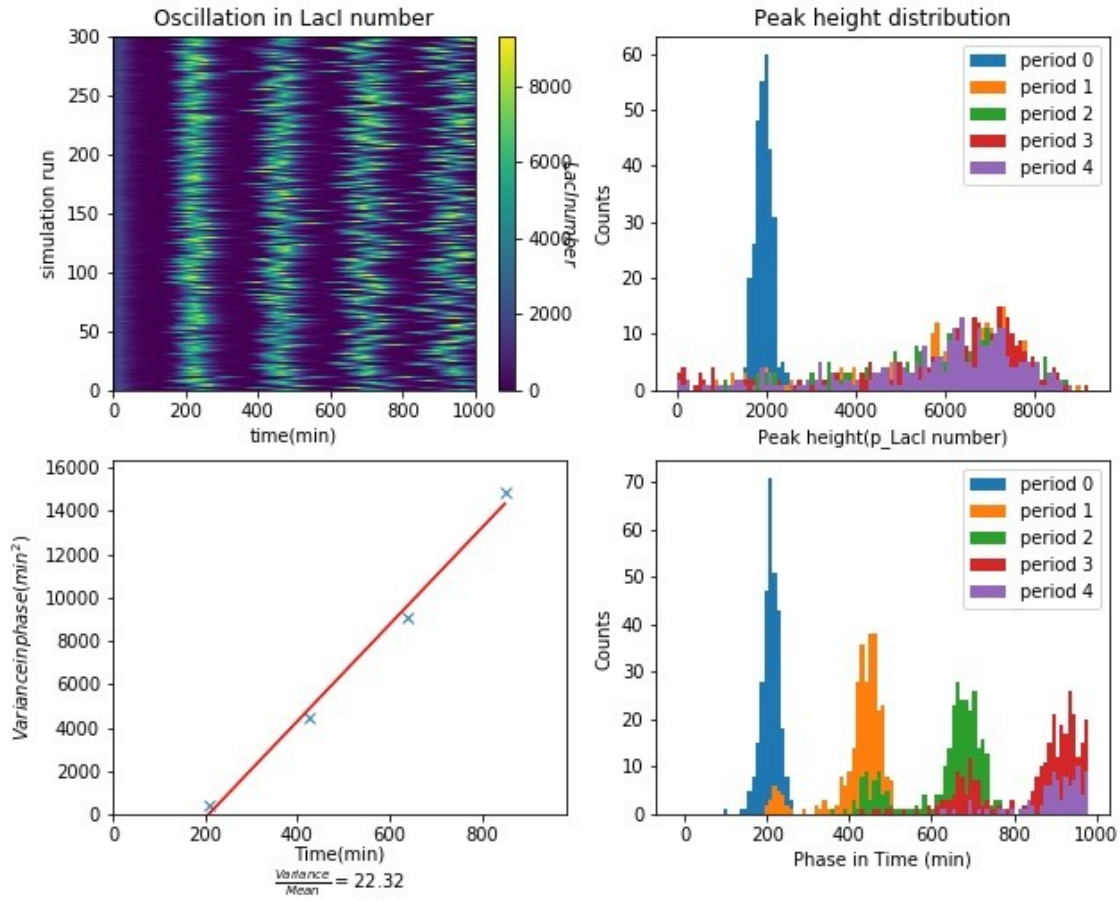


Figure 11: Stochastic simulation at k_dm=0.0010. Top-Left:p_LacI levels form different simulation runs. Top-right: Amplitude PDF at timepoints Bottom-Left: Variance-Time plot. Bottom-right: Phase PDF at timepoints

4.Concluding Remarks:

Modelling is a powerful approach to consolidate our understanding of biological phenomena. Deterministic modelling in terms of ODE formulation and bifurcation has been successfully applied to predict the existence of limit cycles. However these predictions can easily be masked by the intrinsic heterogeneity in biological objects. Such variations, in one way, reflects incomplete knowledge about the underlying mechanism, and are presumably eliminated once a complete mechanistic picture is obtained.

However, such development is not always possible and often impractical for technical reasons and exponentially increasing complexity. Incorporation of these uncertainties into a stochastic model circumvent this need for mechanistic information, and give verifiable predictions about the heterogeneity in system. In the phase drifting example, experiments have identified (Laurent Potvin-Trottier 2016) plasmid copy number, repressor affinities, protein degradation rate and cellular stress as important for robustness, and it is possible to reproduce these qualities in stochastic modelling. One should note, however, it is not reproducing experimental observation but mechanistic simplification that qualify a model for insightful modelling.

As it happens, the birth of synthetic biology offers researchers more controllable systems for more detailed tests, and finally eliminating “the unknown parameter” problem as encountered in modelling of cell cycle. We believe stochastic modelling will replace deterministic modelling in the

nearly future and fuel a new round of discoveries in synthetic biology and thus in design principles of life.

Future directions have been indicated elsewhere in the essay will not be explicitly addressed.

Appendix:
1) 3D diagrams for deterministic one-parameter bifurcations.
2) Equations
3) Codes to reproduce the figures

References:

1. Nyquist–Shannon_sampling_theorem. Available at: https://en.wikipedia.org/wiki/Nyquist–Shannon_sampling_theorem.
2. Novak, B. & Tyson, J. J. Design Principles of Biochemical Oscillator. **37,** 872–880 (2009).
3. Csikász-Nagy, A., Battogtokh, D., Chen, K. C., Novák, B. & Tyson, J. J. Analysis of a generic model of eukaryotic cell-cycle regulation. *Biophys. J.* **90,** 4361–79 (2006).
4. Tsimring, L. *et al.* A synchronized quorum of genetic clocks. *Nature* **463,** 326–330 (2010).
5. Acebrón, J. A., Bonilla, L. L., Vicente, C. J. P., Ritort, F. & Spigler, R. The Kuramoto model: A simple paradigm for synchronization phenomena. *Rev. Mod. Phys.* **77,** 137–185 (2005).
6. Strogatz, S. H. & Hall, K. Ny 14853 1. *Notes* 2–4 (1997). doi:0-7803-3728-W97
7. Elowitz, M. B. & Leibler, S. A synthetic oscillatory network of transcriptional regulators. *Nature* **403,** 335–338 (2000).
8. Guschlbauer Wilhelm, ed. *DNA-ligand Interactions: From Drugs to Proteins*. (Plenum Press, 1986).
9. Vocke, C. & Bastia, D. Primary structure of the essential replicon of the plasmid pSC101. *Proc. Natl. Acad. Sci. U. S. A.* **80,** 6557–6561 (1983).
10. Xia, G. X. *et al.* A copy-number mutant of plasmid pSC101. *Mol. Microbiol.* **5,** 631–640 (1991).
11. Nathan, L. P., Glenn, D. L. & Johan, V. Synchronous long-term oscillations in a synthetic gene circuit. *Nature* 1–4 (2016). doi:10.1038/nature19841
12. Kuhlman, T., Zhang, Z., Saier, M. H. & Hwa, T. Combinatorial transcriptional control of the lactose operon of Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 6043–8 (2007).