**R E S C I E N C E C**

[Re] A Multi-Functional Synthetic Gene Network

Gabriel Baiocchi de Sant'Anna[1, ID] and Mateus Favarin Costa[1]
[1]Universidade Federal de Santa Catarina (UFSC), Florianópolis, Brazil

Studies in the field of synthetic biology make constant additions to existing repositories of biological circuits, as well as to the theoretical understanding of their capabilities. Although many synthetic gene networks have been demonstrated able to perform computations using biomolecules, until recently the majority of such models were engineered to implement the functionality of single reusable circuit parts. Purcell et al.[1] have proposed a network capable of multiple functions, allowing for the selection of three different behaviours in a programmable fashion. This work provides an open-source reference implementation with which their *in silico* experiments are replicated and that can be reused by future studies.

1 Introduction

The field of biomolecular computing – and DNA-based computing in particular – has advanced remarkably over the last years [2]. There are numerous designs of biological parts which implement the behavior of digital logic gates [3], continuous-time systems [4], oscillators [5], memory components [6], asynchronous circuits [7] and so on. Such recent developments allow one to consider the possibility of exploting biologically derived materials and their aspects of massive parallelism and self-replication to build practical computing systems on biological *substrata* [8].

Amidst forward-engineered biochemical systems, genetic oscillators have been a focus of research [9] as they are required for the correct operation of synchronous sequential circuits and can also provide persistent periodic *stimulus* to other regulatory networks which may rely on them [10]. Genetic switches present another functionality specially useful [10] to digital logic: the ability to toggle between on or off states by either activating or repressing the expression of a certain gene makes them equivalent to a cellular memory unit [8]. One study has shown that combining an oscillator with a toggle switch under certain circumstances will result in the generation of a clock-like near square wave signal [11].

Purcell et al.[1] presented the *in silico* design of a novel genetic regulatory network which performs frequency division on an oscillating input. During experiments, that model was also discovered able to behave as a self-induced oscillator or toggle switch. We believe such multi-functionality may lead to more programmable reusable components in biological computing and this has led to the reproduction of simulations described in that paper on an open-source implementation of the aforementioned model.

2 Methods

The multi-functional synthetic gene network and its dynamics are wholly described in the original study. Supplementary material contains the full Ordinary Differential Equation (ODE) system under mass-action kinetics with Hill functions used to represent activation and repression of genetic promoters. The Quasi-Steady-State Assumption (QSSA) exploited to derive the reduced model considered in simulations is also provided, together with all reaction parameterisation and initial state of each experiment. These factors allowed for an easy replication of the model, even without direct reference to source code or any usage of the proprietary tools (MAPLE and MATLAB) originally employed.

We borrow the network's mathematical description from the original paper to replicate each of the hereby presented simulations. Unless otherwise stated, experimental results on the next section use initial conditions $R1 = R2 = 50nM$, $R3 = R4 = 0nM$ and reaction parameters are the same as given in the reference work. All numerical simulations were performed in Octave (version 5.1.0) using Euler's method with a fixed step size of 60 seconds (yielding an iteration count totaling around 6000 steps). The choice is justified based on the duration of experiments, the shortest of which take at least four days (virtual simulated time) in order to observe a couple of periods on the oscillating output of the frequency divider.

Every deterministic experiment was carried in two systems: one considering the full set of ODEs and another with the QSSA approximation that is used throughout that study. Figures show quantitative results in the complete model, but differences between that and the reduced one are highlighted. While stochastic experiments are only briefly mentioned on the main text, more details can be found inside supporting information documents. The Chemical Langevin Equations (CLEs) described therein were implemented with Gaussian noise being generated through the use of built-in Octave functions.

3 Results

3.1 Frequency Divider

Although the network is said to be a frequency multiplier, it actually performs frequency division as concentrations of repressor proteins oscillate in one half of the input frequency or, equivalently, with a twice as long period. This behaviour can be observed by feeding the model with a continuously oscillating input – this would be often the case considering existing genetic oscillators [9] – but also works with square waves.

With a sinusoidal input, output signals from the QSSA model have constant amplitude, whereas in the full model the first concentration peak of proteins R2 and R3 are higher than the following ones. This suggests mRNA reactions stabilize quicker in the first 10^5 seconds and thus the approximation is more accurate at those instants [12].

After that moment, however, the reduced model exhibits a persistent error in relation to the complete one. This can be observed by comparing concentration levels reached during oscillation peaks in Figures 2 and 3. In those simulations, while R1 and R4 concentrations in the full model reach *maxima* valued at 238.29 nM and 87.62 nM respectively, the reduced model goes up to 283.96 nM and 113.05 nM for each of these proteins. Peak values of R3 follow R1 closely and the same happens with R2 and R4.

This offset distinguishing the two models happens because the QSSA used to derive the reduced ODE system considers a separation of time-scales between reactions which regulate mRNA production and those which describe protein translation. In the approximation, the former reactions are assumed to reach equilibrium instantaneously relative to the latter. Thus, the difference is a consequence of the fact that the original model maintains itself in a dynamic state that never actually reaches equilibrium [12], as it perpetually oscillates.
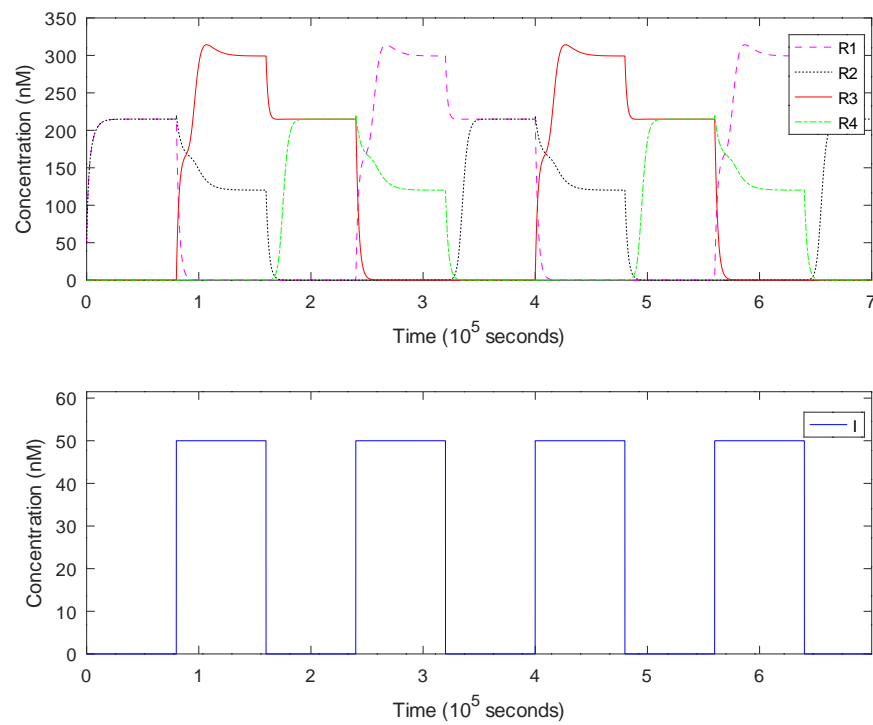
Figure 1. **Frequency division of clock-like input.** Input signal is a square wave with amplitude of $50\,\text{nM}$, period of 1.6×10^5 seconds and 50% duty cycle.
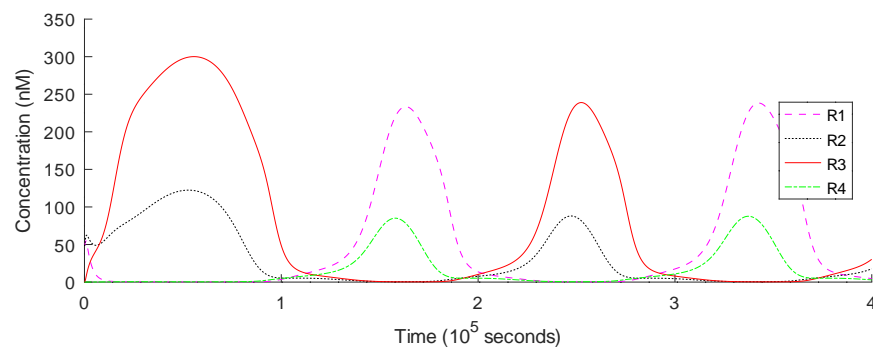
Figure 2. **Frequency division in the full model.** Sinusoidal input has an amplitude of $50\,\text{nM}$, minimum of $6\,\text{nM}$ and period is 0.9×10^5 seconds.
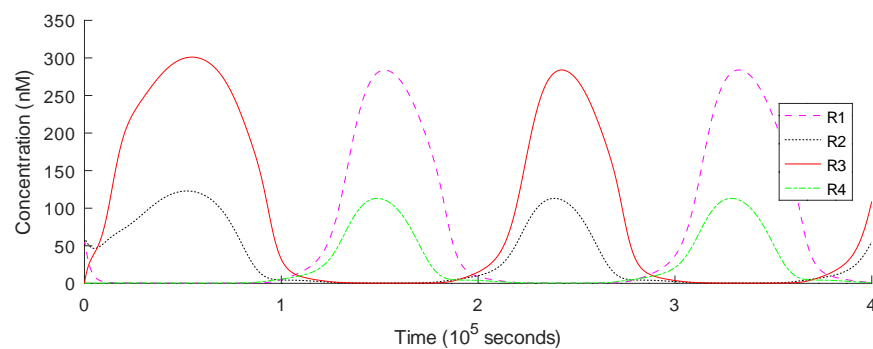
Figure 3. **Frequency division in the reduced model.** Input is the same as in Figure 2.

As mentioned in the original study, frequency division functionality can only be observed after a specific period threshold. We ran simulations over a range of input frequencies and found the period-doubling bifurcation to be located near values of 0.8×10^5 seconds (~ 22.22 hours) in both full and reduced models. Although this confirms the network's capability to interface with long-period oscillators, these results greatly diverge from those in the reference work, which state this threshold could be observed at input periods of approximately 0.275×10^5 seconds (~ 8 hours).
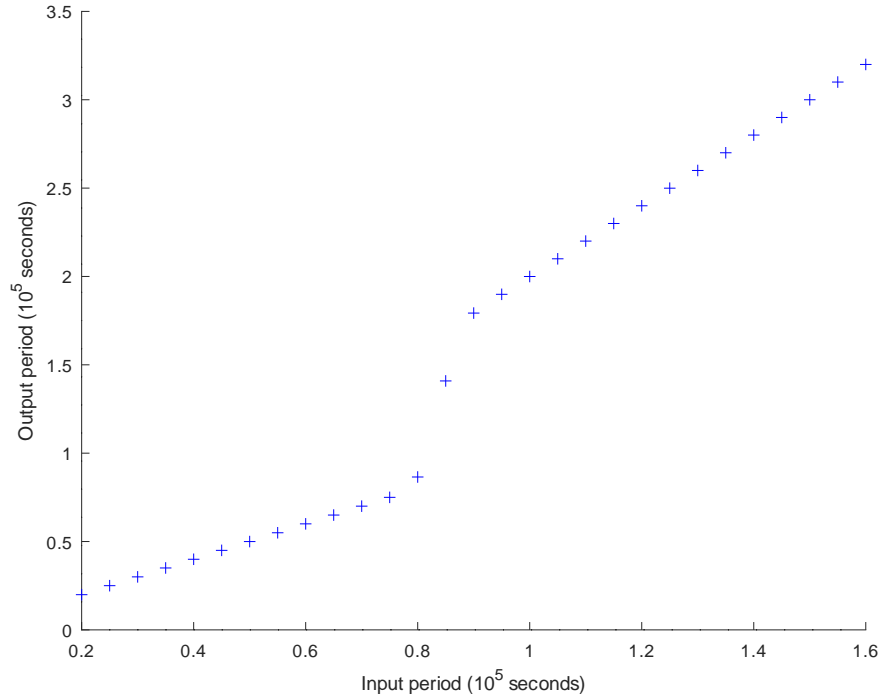

Figure 4. **Locating the period-doubling bifurcation.** Output period is detected by measuring the distance between R1 concentration peaks. Other than input frequency and simulation length (each run was configured to take as long as five times the input period), parameters are the same as in Figure 2.

3.2 Bifurcation Analysis

Purcell et al.[1] discovered the model's multiple extra functionalities by verifying different behaviours can be attained when input concentration is held constant at specific ranges. Experiments regarding the so-called bifurcation analysis were reproduced and Figures 5-8 show results under the complete model.

The original study states simulations labeled $4c_1$ and $1b$ use the same initial conditions as $4c_2$ and $1a$ respectively. We believe these were typographical errors, as such settings would lead those pairs of experiments to the exact same results under deterministic semantics and this is not the exhibited behaviour. Instead, whereas reaction parameters and input levels are the same as in the original work, initial conditions used are $R1 = R2 = 0nM$ and $R3 = R4 = 50nM$ for simulations $1b$ and $4c_2$ and $R1 = R2 = 50nM$ and $R3 = R4 = 0nM$ for all others.
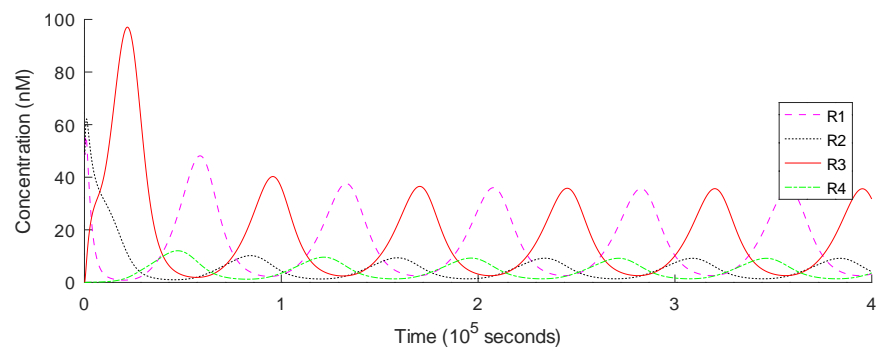
Once again, oscillatory amplitude differs between models in experiment 2, with R1 and R4 each peaking at 35.60 nM and 9.16 nM in the complete system (Figure 6) while going up to 30.08 nM and 8.18 nM in the QSSA reduction (not shown). Other experiments do not present such significant offset.
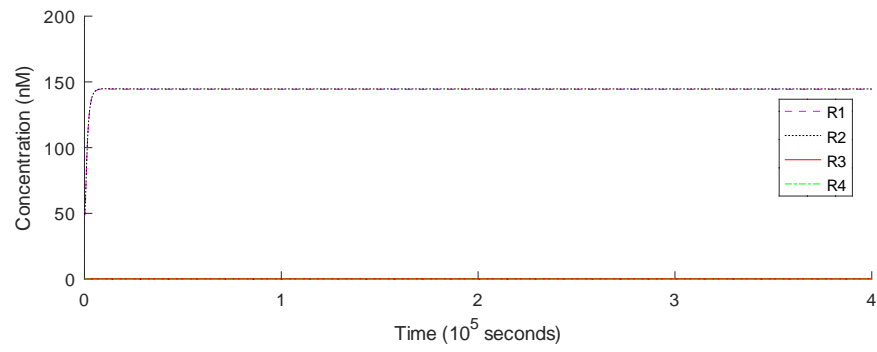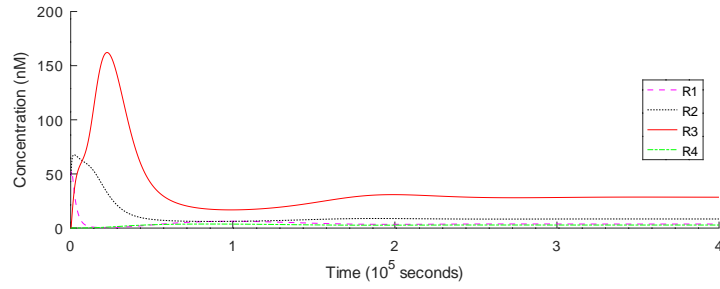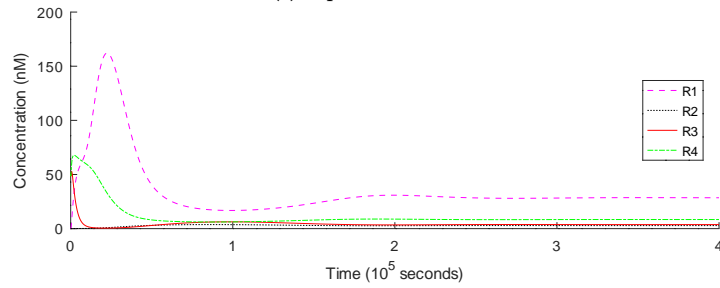
(a) Experiment $1a$.
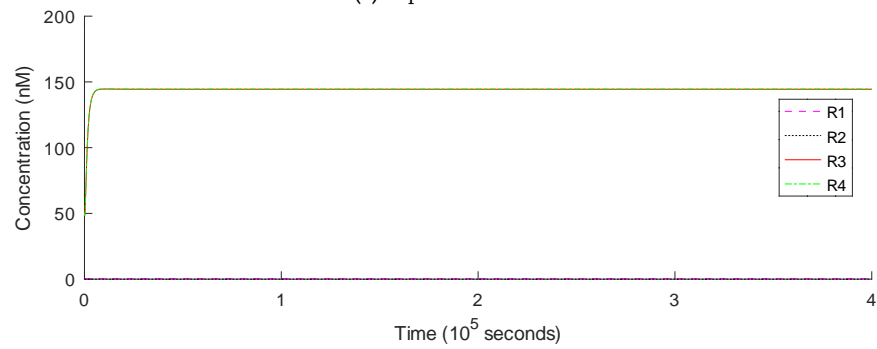
(b) Experiment $1b$.

Figure 5. Low concentration bistable behaviour. $I = 0.1nM$.

Figure 6. Experiment 2. $I = 5nM$.
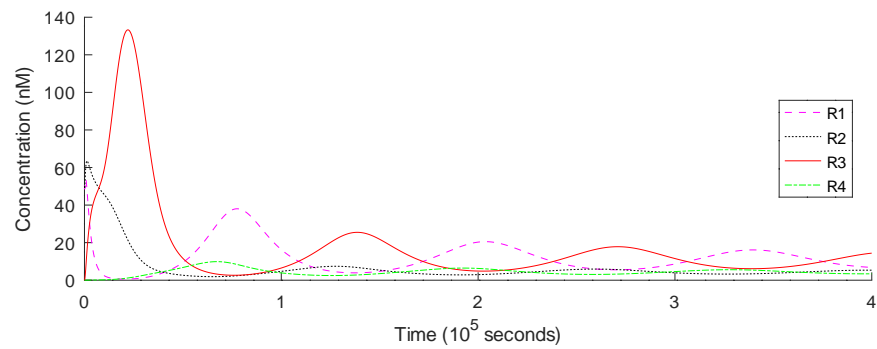
Figure 7. Experiment 3. $I = 7.5nM$. Oscillations are gradually damped and cease after about $12 \times 10^5 s$ elapsed time. This behaviour evidentiates a pitchfork bifurcation.

(a) Experiment $4c_1$.

(b) Experiment $4c_2$.

Figure 8. High concentration bistable behaviour. $I = 10nM$.

3.3 Self-induced Oscillator

We verified the network's oscillatory dynamics in the region between saddle-node and Hopf bifurcations by measuring its output period for each given input concentration level. Results illustrated in Figure 9 describe a relation with similar behaviour and the same near-vertical increase in period as input approaches the lower bistability region.
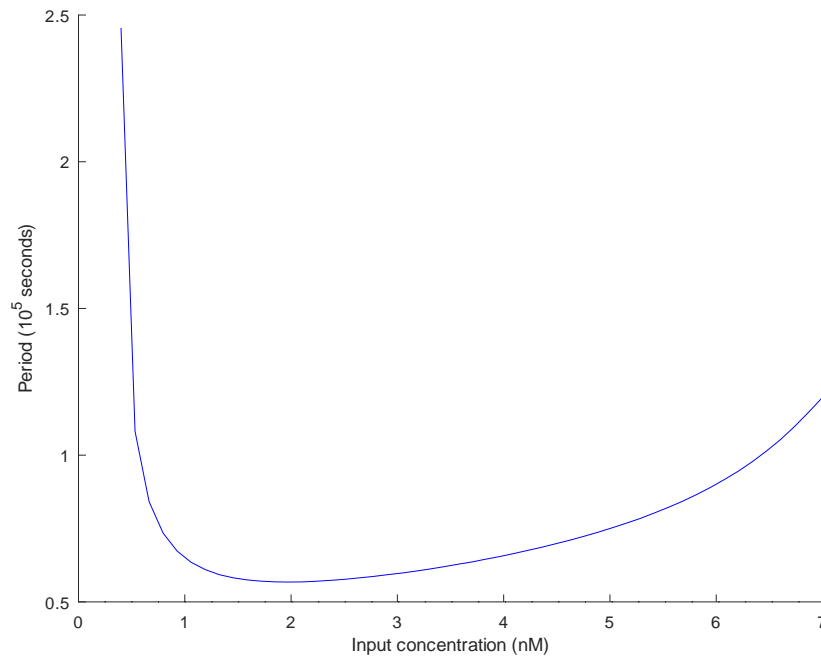

Figure 9. **Analysing oscillatory dynamics.** Simulation configuration is the same as in Figure 6 except for input levels, which are held in the range $[0.4nM, 7nM]$. Output period is detected by measuring the distance between R1 concentration peaks.

Toggle Switch

As revealed during bifurcation analysis, the network exhibits bistability when input concentration is held outside the oscillatory range, that is, at levels lower than 0.4 nM or greater than 7 nM. Figures 10-13 illustrate simulations with the system being used as a toggle switch which is "triggered" by varying binding affinity of particular repressors as described in the reference work. This set of experiments were exactly reproduced and no significant difference is found between reduced and full models.
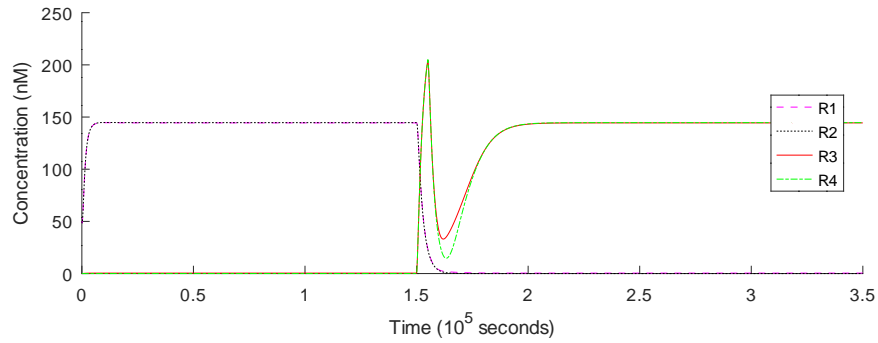

Figure 10. Toggle switch behaviour at low concentration. Switching active state from R1 & R2 to R3 & R4 at $I = 0.1nM$ by lowering binding affinity of proteins R1 and R2 between 1.5×10^5 and 1.55×10^5 seconds. Initially, $R1 = R2 = 50nM$ and $R3 = R4 = 0nM$.
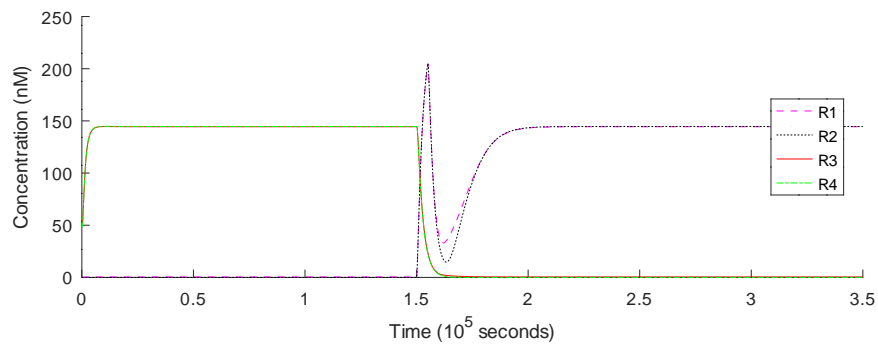

Figure 11. Toggle switch behaviour at low concentration. Switching active state from R3 & R4 to R1 & R2 at $I = 0.1nM$ by lowering binding affinity of proteins R3 and R4 between 1.5×10^5 and 1.55×10^5 seconds. Initially, $R1 = R2 = 0nM$ and $R3 = R4 = 50nM$.

Figure 12. Toggle switch behaviour at high concentration. Switching active state from R2 & R3 to R1 & R4 at $I = 50nM$ by lowering binding affinity of proteins R1 and R2 between 1.5×10^5 and 1.55×10^5 seconds. Initially, $R1 = R2 = 50nM$ and $R3 = R4 = 0nM$.
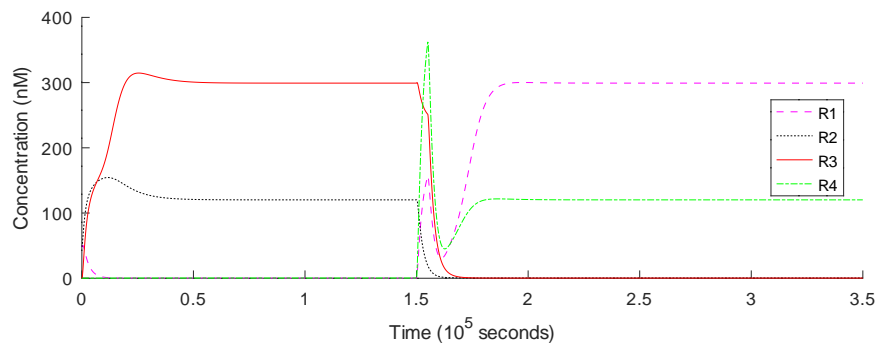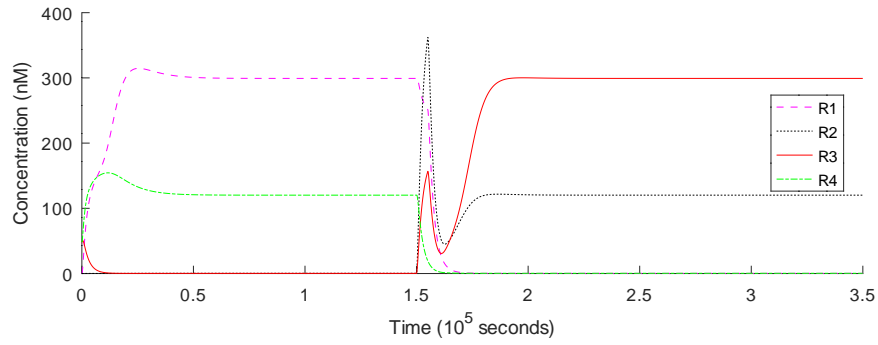
Figure 13. Toggle switch behaviour at high concentration. Switching active state from R1 & R4 to R2 & R3 at $I = 50nM$ by lowering binding affinity of proteins R3 and R4 between 1.5×10^5 and 1.55×10^5 seconds. Initially, $R1 = R2 = 0nM$ and $R3 = R4 = 50nM$.

3.5 Stochastic Simulations

We implemented the CLEs shown in the related supporting information document provided by Purcell et al.[1], but it was not possible to verify similar results without modifications to the noise-inducing function. While original authors state the usage of Gaussian noise with zero mean and variance of 1, simply employing Octave's `randn` function – which provides such a distribution [13] – proved being too noisy as stochastic fluctuations began dominating the model's behaviour. It was found through trial and error that downscaling random numbers by a factor of $s \in [\frac{1}{100}, \frac{1}{10}]$ (and consequently scaling variance by s^2) would bring results closer to what is seen in the reference work. In order to increase reproducibility, stochastic simulations are performed with a hardcoded random seed for the noise-generating function.
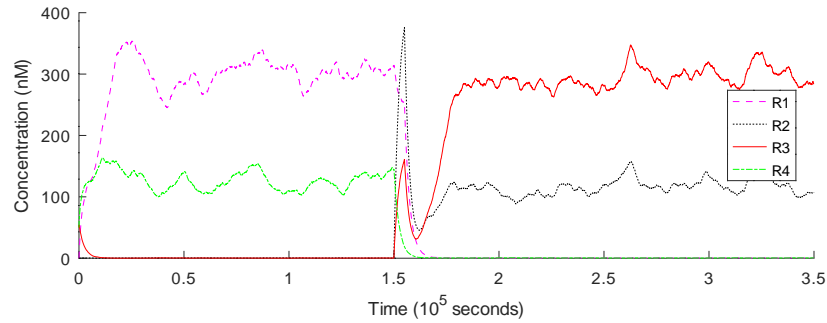
(a) Initially, $R1 = R2 = 0nM$ and $R3 = R4 = 50nM$.
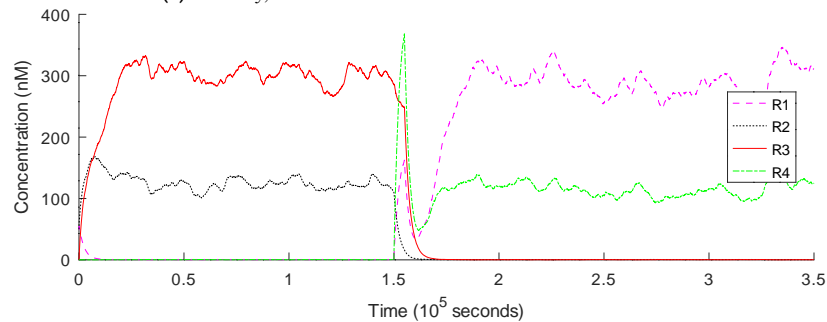
(b) Initially, $R1 = R2 = 50nM$ and $R3 = R4 = 0nM$.

Figure 14. Effect of noise on switching function. The trigger is held between 1.5×10^5 and 1.55×10^5 seconds at $I = 50nM$. Random seed is set to 73544911520192 and fluctuations are scaled by $\frac{1}{55}$.
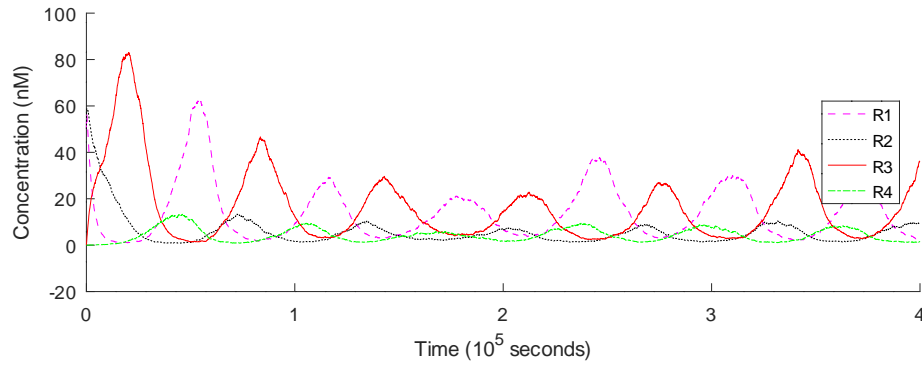
Figure 15. **Effect of noise on oscillator functionality.** Input is set to 5 nM, initial repressor state is $R1 = R2 = 50nM$, $R3 = R4 = 0nM$ and noise is scaled by $\frac{1}{55}$. Random seed: **73544911520192**.
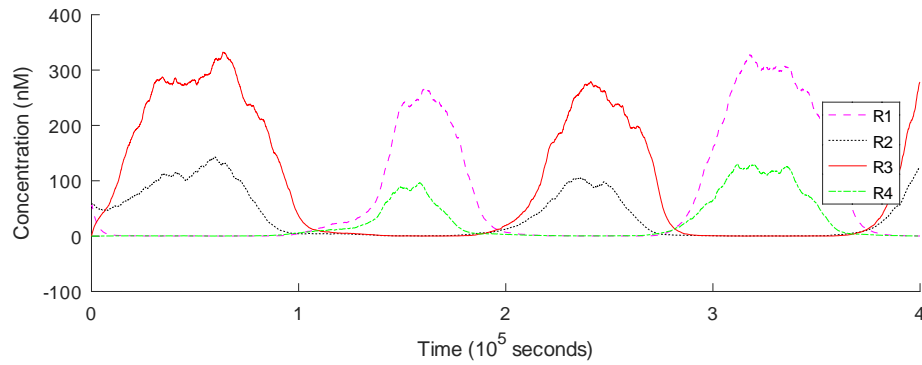

Figure 16. **Usual frequency division behaviour.** Noise is scaled by a factor of $\frac{1}{55}$ and the random seed used is **73544911520192**. Other parameters are the same as in Figure 2.
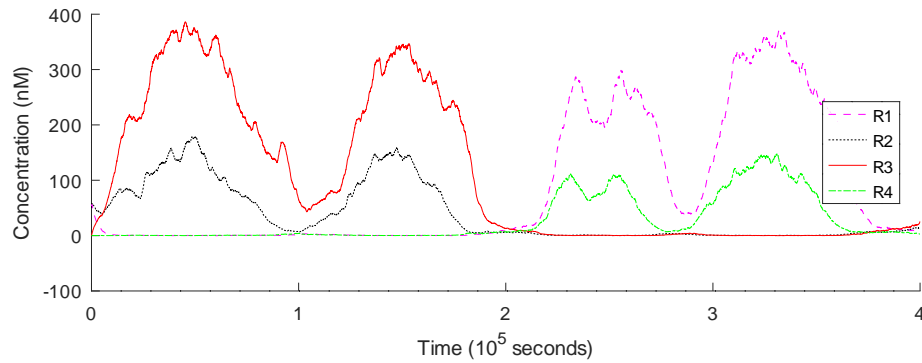
Figure 17. **Irregular oscillations caused by stochastic fluctuations.** Parameters are the same as in Figure 16, but with the random seed set to **0.8977812586558097** and noise scaling to $\frac{1}{26}$.

We verified that among the three functionalities, frequency division possesses less robustness to intrinsic noise. During the stage where input is applied and proteins are all at low concentrations, random fluctuations may cause unintended oscillations as a pair of repressors rise in concentration and prevent transcription of the other two. An example of such irregular oscillatory behaviour is given in Figure 17.

4 Conclusion

All operation modes in the multi-functional synthetic gene network were verified and to a large degree the results presented in [1] were replicated. It should be emphasised that the original work stands as an example of reproducible science, with detailed descriptions of model derivation, parameter decisions and extra experiments inside openly available supplementary material. Simplifications made using the QSSA to reduce the full ODE system proved being a good approximation as even though some small quantitative differences were observed, there was no impact on the network's qualitative dynamics.

However, stochastic simulations would not provide the expected results without a significant reduction to noise variance. In fact, Gaussian noise seems capable of undermining frequency division functionality to some degree. Additionally, the period-doubling bifurcation was empirically found to be at a much lower frequency than what is stated in the reference study. Although this may have been a typographical mistake, it is a critical piece of information which can be used to know whether or not the model would work with existing or future oscillators.

We highlight the idea that as synthetic networks grow in complexity and size, multi-functionality may arise more frequently and become difficult to avoid. While this can bring the benefits of reusable programmable components, it might also end up becoming a nuisance if models start behaving unexpectedly under the influence of certain inputs. At last, we believe further works should focus on a deeper exploration of parameter space. The motivation for this is double: it may allow for the characterisation of *in vivo* implementation designs with reusable biological parts and will potentially verify the full set of specifications for this programmable genetic component and its multiple operating configurations.

References

1. O. Purcell, M. di Bernardo, C. S. Grierson, and N. J. Savery. "A Multi-Functional Synthetic Gene Network: A Frequency Multiplier, Oscillator and Switch." In: **PLOS ONE** 6.2 (Feb. 2011), pp. 1–12.
2. D. E. Cameron, C. J. Bashor, and J. J. Collins. "A brief history of synthetic biology." In: **Nature Reviews Microbiology** 12 (5 Apr. 2014), pp. 381–390.
3. A. Goni-Moreno and M. Amos. "A reconfigurable NAND/NOR genetic logic gate." In: **BMC Systems Biology** 6 (Sept. 2012), p. 126.
4. S. A. Salehi, H. Jiang, M. D. Riedel, and K. K. Parhi. "Molecular Sensing and Computing Systems." In: **IEEE Transactions on Molecular, Biological and Multi-Scale Communications** 1.3 (Sept. 2015), pp. 249–264.
5. M. B. Elowitz and S. Leibler. "A synthetic oscillatory network of transcriptional regulators." In: **Nature** 403 (Jan. 2000), pp. 335–338.
6. C. Lin, T. Kuo, and Y. Chen. "Implementation of a genetic logic circuit: bio-register." In: **Systems and Synthetic Biology** 9.1 (Dec. 2015), pp. 43–48.
7. L. Cardelli, M. Kwiatkowska, and M. Whitby. "Chemical reaction network designs for asynchronous logic circuits." In: **Natural Computing** 17.1 (Mar. 2018), pp. 109–130.
8. G. H. Moe-Behrens. "The Biological Microprocessor, or How to Build a Computer with Biological Parts." In: **Computational and Structural Biotechnology Journal** 7.8 (2013), e201304003.
9. T. Mahajan and K. Rai. "A novel optogenetically tunable frequency modulating oscillator." In: **PLOS ONE** 13.2 (Feb. 2018), pp. 1–29.
10. A. Khalil and J. Collins. "Synthetic Biology: Applications Come of Age." In: **Nature reviews. Genetics** 11 (May 2010), pp. 367–379.
11. C. Lin, P. Chen, and Y. Cheng. "Synthesising gene clock with toggle switch and oscillator." In: **IET Systems Biology** 9.3 (2015), pp. 88–94.
12. B. P. Ingalls. **Mathematical Modeling in Systems Biology: An Introduction**. The MIT Press, Jan. 2013.
13. Octave Forge Community. **Octave randn Documentation**. URL: https://octave.sourceforge.io/octave/function/randn.html (visited on 11/11/2019).