This is an accepted version of a paper published in Natural Computing. The final publication is available at Springer via http://dx.doi.org/10.1007/s11047-013-9383-8.

An adaptive genetic algorithm for parameter estimation of biological oscillator models to achieve target quantitative system response

Martin Stražar,* Miha Mraz, Nikolaj Zimic and Miha Moškon

Faculty of Computer and Information Science, University of Ljubljana, Tržaska cesta 25, Ljubljana, Slovenia
Phone: +(386)14768371; Fax: +(386)14264647

Email: Martin Stražar*- martin.strazar@gmail.com;

*Corresponding author

Abstract

Mathematical modeling has become an integral part of synthesizing gene regulatory networks. One of the common problems is the determination of parameters, which are a part of the model description. In the present work, we propose a customized genetic algorithm as a method to determine the parameters such that the underlying oscillatory system exhibits the target behavior. We propose a problem specific, adaptive fitness function evaluation and a method to quantify the effect of a single parameter on the system response. The properties of the algorithm are highlighted and confirmed on two test cases of synthetic biological oscillators.

Keywords: adaptive genetic algorithm, biological oscillator, gene regulatory networks, parameter estimation;

Introduction

In the recent years, notable improvement has been accomplished in understanding the component interactions in various living organisms, leading to improved and more accurate modeling techniques available in the fields of synthetic and systems biology (Alon, 2007).

A common biological subsystem is an oscillating regulatory network responsible for maintaining the circadian rhythm which in turn affects the sleep-wake cycle, thermogenesis, feeding, glucose and lipid metabolism, energy balance, regulation of peripheral body clocks, respiratory control, etc. Diseases such as diabetes,

Alzheimer's disease, sleeping disturbances and various metabolic diseases are known to be caused by disruption in the oscillatory mechanism (Turek et al, 2005; Morton et al, 2005; Wu et al, 2006). Understanding the behavior of oscillatory systems is essential upon approaching such problems.

The dynamics of the oscillatory systems are well understood, as well as the conditions for the emergence of the necessary unstable limit cycle (Novák and Tyson, 2008). Also, synthetic oscillators based on gene regulatory networks (GRNs) have been successfully realized both in prokaryotes (Elowitz and Leibler, 2000) as well as in mammalian cells (Tigges et al, 2009). A fundamental problem in modeling synthetic GRNs is the uncertainty in the values of the model parameters (Lillacci and Khammash, 2012, 2010) and the level of modeling complexity for a specific network. In ref. Scheper et al (1999) the authors show minimal required conditions for a biological oscillator model by treating various post-transcriptional processes as a single delay process. In order to achieve the targeted 24 hour period they partly explore the parameter space while restricting the parameters within predefined boundaries.

In the present study we extend the concept of exploring the parameter space towards achieving a targeted system response, which is characterized by the frequency and amplitude of the resulting oscillations. We employ a customized, adaptive genetic algorithm which can be applied to an arbitrary oscillator model, based on delayed differential equations (DDEs) and is able to find globally optimal parameters that cause responses which can range over several orders of magnitude for both amplitude and frequency.

Various methods and tools aimed at optimizing and seeking the desired behavior of gene regulatory networks have been proposed. General methods rely on having a pre-established library of modular components and searching for the optimal network topology that would produce the desired response. Implementations (Rodrigo et al, 2007; Rodrigo and Jaramillo, 2012) may employ a truth table for the desired function as an input and produce the optimal network topology as output using evolutionary techniques. Whether the resulting components would perform in a linear manner when wired together remains an open question due to observed non-modularity in living organisms (Hajimorad et al, 2011). (François and Hakim, 2004) perform a more specific study, where they define the characteristics of the desired response more precisely. Our approach is focused on a specific subgroup of networks in order to a) take advantage of the knowledge about the expected system's response and b) provide quantitative answers about network parameters such that they can be used to aid the design decision while building such networks.

Genetic algorithms have long been used to tackle analytically intractable optimization problems in multidimensional parameter space (Chambers, 2000). Recent developments in the field tend towards adaptive behavior of the algorithms. Examples include dynamic tuning of mutation rates, balancing the ratios between

mutations and crossover or adapting the fitness function depending on the current state of the population. In ref. Law and Szeto (2007), authors have demonstrated the advantages of mutation and crossover probabilities, proportional the fitness of an individual. Also, they have shown the importance of a single parameter's influence on the overall fitness can be quantified by calculating it's standard deviation. Summing up, less important parameters in less fit individuals tend to be subject to mutation with higher probability, while crossover is applied to fitter individuals more often. Using a similar method, the population of individuals in our case is replicated with a bias towards the fitter solutions.

The influence of each parameter in the resulting system response is quantified using the entropy function, an appropriate way to quantify determinism (Szendro et al, 2013). Repeatability is a measure of quantifying the determinism regarding an instance of evolution. It is defined as a probability that two independent evolutions will follow exactly the same path (Roy, 2009). By performing multiple runs of the developed genetic algorithm, we derive the statistics for each parameter and quantify it's influence on the resulting system's response.

Various choices of mutation probabilities are discussed in (Smith and Fogarty, 1996). Local maxima in the fitness landscapes are an emphasized problem of genetic algorithms. To be able to perform giant leaps in parameter space one should allow multiple alterations to be performed on candidate solutions. This can be achieved by allowing multiple mutations of parameters and/or usage of crossover. That way, a significantly larger number of possible evolution scenarios are available (Roy, 2009) and given enough time, the globally optimal solution will always take over the population (Szendro et al, 2013; Chambers, 2000). Allowing multiple mutations and crossover cause faster escape from the local maxima as well as opening the possibility of beneficial joint mutations which are otherwise deleterious (Weinreich and Chao, 2005). As we apply the algorithm to a specific family of problems, we assess the optimal choices for mutation probability and crossover method.

Adaptive behavior is further explored, using an adaptive fitness function. The performance gains caused by the usage of the population state dependent fitness evaluation are described in (Farmani and Wright, 2003). In addition, the authors introduce a concept of penalization of unfeasible solutions. We argue the performance gains of a problem specific, adaptive fitness function which also penalizes the unfeasible solutions. Evolutionary approaches towards tuning the response of biological circuits (Paladugu and Chickarmane, 2006; Szendro et al, 2013; Fang et al, 2009) tend to use generic fitness measures (e.g. non-linear least squares error), limiting the solution space. The problem specificity of the fitness function gives us the advantage of finding equivalently fit solutions in larger solution space, since it is not limited by absolute values of the

resulting system response.

The resulting set of parameter values guarantees the desired system response for a given model of a biological oscillator, offering a quantitative insight of the conditions that have to be met on the design level. The manipulation of parameter values of a gene regulatory is to some degree already achievable *in vivo*. Examples include engineering synthetic ribosome binding sites to influence translation efficiency (Salis et al, 2009), modifying the upstream promoter sequences to alter the transcription rate (Rhodius et al, 2012), using transcription factor proteins with different degree of cooperativity (Dill et al, 1993) or degradation tags to accelerate the protein degradation (Butz et al, 2011). Studies that rely on modifying the temporal behavior of rhythmic systems ought to make use of the quantitative knowledge of parameter values to achieve the desired effects (Mackey and Glass, 1977; Morton et al, 2005; Wu et al, 2006). A concept similar to our was also performed *in vivo* (Yokobayashi et al, 2002) in which the authors deliberately induce mutations during the gene cloning process to demonstrate the achievable change in protein-protein as well as protein-DNA interactions, which confirms the practical applicability of our approach.

Section Framework introduces the computational framework used in the study. Section Results presents results achieved on two different models of biological oscillators, which are discussed in Section Discussion.

Framework

Algorithm overview

A common way to model gene regulatory networks are ordinary differential equations (ODEs), which describe the change of a given species concentration in an infinitesimal interval of time. In general terms, the form of such description is

$$\frac{\delta S_i}{\delta t} = f(S_1(t), S_2(t), ..., S_n(t), p_1, p_2, ..., p_m) \tag{1}$$

where S_i is the observed chemical species, $S_1(t), S_2(t), ..., S_n(t)$ is the system state vector and $p_1, p_2, ..., p_m$ is a vector of parameters, which determine the system behavior. As processes in the cell span over multiple timescales (Scheper et al, 1999), the description can be altered to a system of delayed differential equations (DDEs):

$$\frac{\delta S_i}{\delta t} = f(S_1(t-\tau_1), S_2(t-\tau_2), ..., S_n(t-\tau_n), p_1, p_2, ..., p_m) \tag{2}$$

where the delays with which species affect the system are given by $\tau_1, \tau_2, ..., \tau_n$. The vector of parameters and the vector of time delays uniquely describe the response of a DDE system. A combination of their values

represent a candidate solution.

Fitness Evaluation

Considering the system response as a discrete time signal, we used it's standard deviation along the y-axis as an approximation to the amplitude A. The frequency f was in turn detected using the discrete fast Fourier transform (FFT). Furthermore, the *unfeasible* candidate solutions that reached a stable steady state could be detected. By acquiring the two values, the fitness of an individual candidate solution C_i was determined by computing the relative error ratio against targeted amplitude A and frequency f:

$$F(C_i) = \begin{cases} \Delta_a \cdot \Delta_f & \text{if } C_i \text{ is feasible,} \\ X + \Delta_a \cdot \Delta_f & \text{else;} \end{cases} \tag{3}$$

where $\Delta_a = \|(A_i - A)\|/A$ and $\Delta_f = \|(f_i - f)\|/f$ present obtained candidate solution response errors regarding the target amplitude and frequency values respectively. In case of an unfeasible solution an empirically derived large value X (empirically defined 10^9) was added to the resulting fitness in order to partition population on the feasibility of solutions. The approach agrees with the method proposed in ref. (Farmani and Wright, 2003), where unfeasible solutions are still kept in the population for potential adjustments. This turns out to be particularly important when feasible solutions are located in a very narrow parameter range and eliminating unfeasible solutions would cause the population to effectively decease.

An efficient approach in exploring the parameter space is to make large leaps in the earlier phases and afterwards performing local optimization. We further improved the algorithm by using an adaptive fitness function. In the beginning, candidate fitness is evaluated such that:

$$F(C_i) = \begin{cases} \Delta_f & \text{if } C_i \text{ is feasible,} \\ X + \Delta_f & \text{else;} \end{cases} \tag{4}$$

After the evolution finds a candidate solution with sufficiently small frequency error (empirically defined as 1 %), fitness evaluation method changes to Equation 3. Thus, by first adjusting the frequency and then perfuming local optimization considering both frequency and amplitude, the population proved to converge to the optimum value in a smaller number of steps. The approach mimics the simulated annealing algorithm (Vecchi, 1987) in a sense of performing large steps in the beginning and stems from the fact that some parameters mainly influence either amplitude or frequency while others influence both amplitude and frequency (see Section Results).

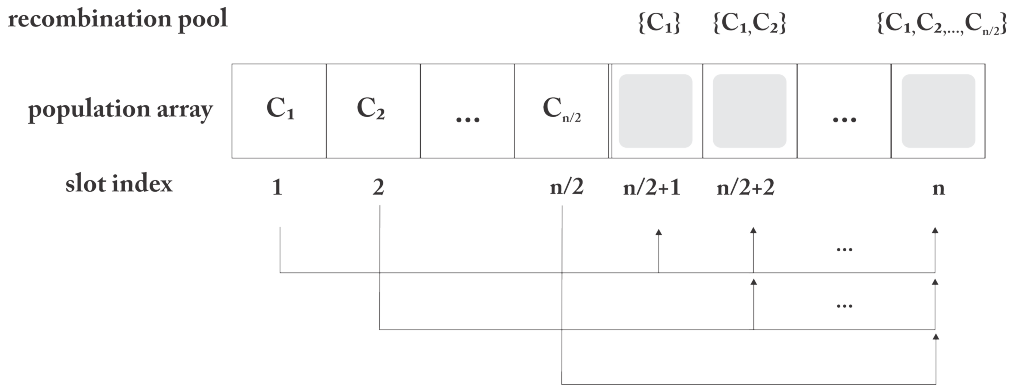

Figure 1: The proposed biased selection method. After the fitness of each individual candidate solution C_i is evaluated, the candidates are sorted and the bottom half (gray) is discarded. The resulting empty slots in the population array are then occupied by new candidates, which are derived by performing crossover of candidates from the recombination pool. The former is dynamic, in a sense that new candidate solutions are added for each slot in accordance with their fitness value. For example, the new candidate on index $n/2+1$ is a mutated best candidate, new candidate on index n/2+2 is a recombination from candidates C_1 and C_2, while at the last index n the resulting candidate is derived from a random pair of candidates from the whole upper half.

Selection and reproduction

The candidate solutions are afterwards sorted based on the ascending value of $F(C_i)$. Only the top half thus comprises the reproduction pool, while the lower half is discarded and slots are available for new candidate solutions. Random pairs of candidates are selected from the reproduction pool and their parameter vectors recombined into one with an equal probability for each component coming from either of parents (n-point crossover). This ensures the ability to make large steps in parameter space and avoid local maxima (Chambers, 2000).

To bias the selection towards fitter candidates, the reproduction pool grows dynamically during the crossover step for each unoccupied candidate solution slot. As the algorithm iterates through slots in previously discarded bottom half of the population, a candidate is added to the current reproduction pool when producing each new individual. In this way, fitter candidates will be in the reproduction pool sooner and their components will have higher probability to appear in the resulting new population. This modification appeared to produce better results and was based on the results obtained in (Law and Szeto, 2007) and is explained in more detail on Figure 1.

Mutations are applied to the parameter vector $p_{i,1}, p_{i,2}, ..., p_{i,l}$ of each candidate solution C_i with the probability proportional to the parameter vector length l (Back, 1992), such that

$$p_{i,j} = p_{i,j} + G(0,1) \tag{5}$$

where G stands for a normally distributed random variable with mean 0 and standard deviation 1. To ensure, that the quality of solutions monotonically increases, the current fittest candidate is never discarded and remains unchanged (elitism) in the next evolution cycle (Chambers, 2000).

Results

Oscillator model

The proposed approach was evaluated using the oscillator model, based on a negative feedback loop proposed in (Scheper et al, 1999) and schematically depicted in Figure 2a. It can be described with a DDE system:

$$\frac{\delta M_i}{\delta t} = \frac{r_m}{1 + (\frac{P(t)}{k})^h} - q_m \cdot M_i(t) \tag{6}$$

$$\frac{\delta M_i}{\delta t} = r_p \cdot M(t-\tau)^m + q_p \cdot P(t) \tag{7}$$

where the system state is composed of M as the mRNA concentration and P as the concentration of the repressor protein. System parameters that we have to determine are mRNA transcription rate r_M, protein translation rate r_P, mRNA degradation rate q_M, protein degradation rate q_P, scaling constant for the repressor protein k, mRNA non-linearity coefficient m, repressor protein cooperativity coefficient h and time delay between transcription and translation τ. To further exploit the *a priori* knowledge about the system, constraints on parameter values are employed. This helps us narrow down the search space and increase convergence speed. In ref. (Scheper et al, 1999), it is argued that described system can exhibit oscillatory behavior if and only if $\tau > 0$. On top of that, we constrained the degradation rates for both protein and mRNA such that $q_M, q_P < 1$ and non-linearity coefficients $m, h > 1$.

After reaching the unstable steady state, the system exhibits an oscillatory response. We performed a search for the target amplitude of 16 nM and target period of 24 h (frequency $0.042h^{-1}$) obtained in (Scheper et al, 1999) and found another equivalent solution in the parameter space (Figure 2b). The observations and conditions for oscillations observed in (Scheper et al, 1999; Novák and Tyson, 2008) apply to the resulting parameter vector as well, indicating there is more than one point in the solution space producing equivalent system response in terms of amplitude and frequency;

Running the evolutions for a range of amplitudes and frequencies spanning various orders of magnitude, the search method returned results with error values of the order 10^{-5} on average. To achieve optimal results, we adjusted and tested different properties of the proposed algorithm. Crossover of different parameter values proves beneficial for some problems while not for the others (Chambers, 2000). As the system

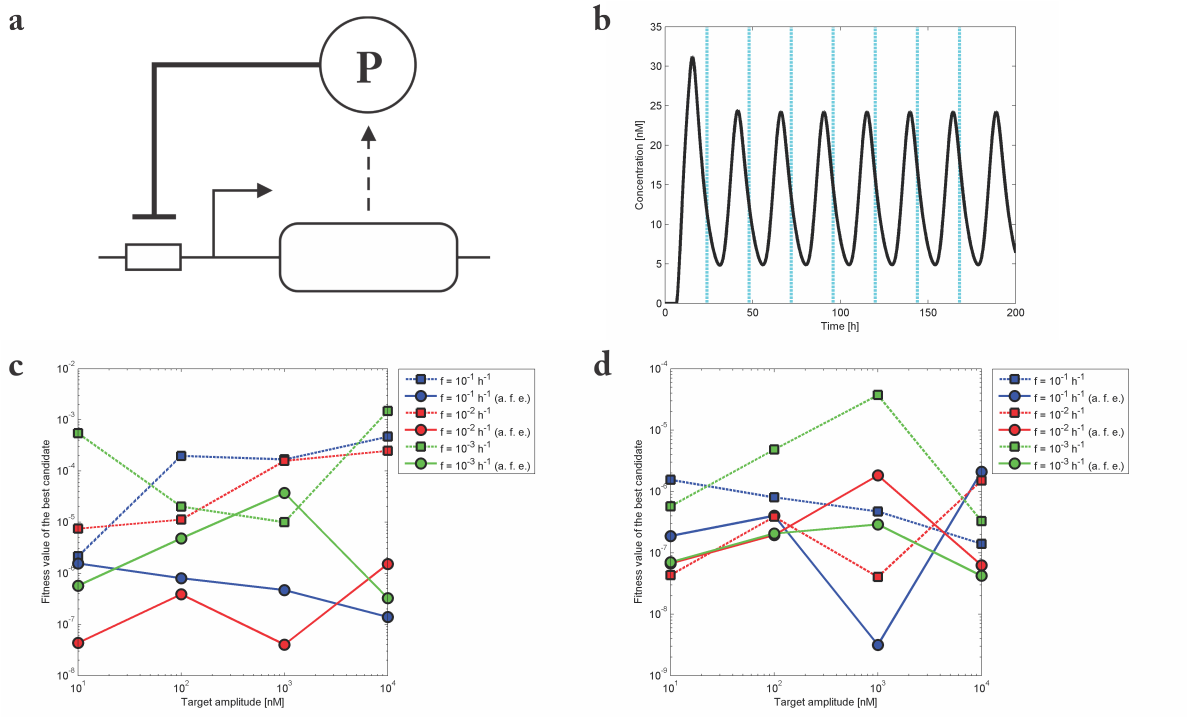

Figure 2: a) Schematic representation of a single repressor oscillator model; b) Simulation of optimal solution of parameter space search for target amplitude $A = 16\ nM$ and frequency $f = 0.042h^{-1}$. Parameters: $\tau = 12.906$, $r_M = 8.025$, $r_P = 15.884$, $q_M = 0.045$, $q_P = 0.067$, $m = 2.69$, $h = 2.083$ and $k = 9.88$. 24 hour periods are marked with a cyan dashed line. c) Comparison of parameters space investigation for amplitudes and frequencies of different orders of magnitude with (circles) and without crossover (squares); d) Comparison of parameters space investigation for amplitudes and frequencies of different orders of magnitude with (circles) and without adaptive fitness evaluation (squares).

response qualitatively depends on the ratios among parameter vector components (Novák and Tyson, 2008), conservation of fit subparts of parameter vectors proves beneficial. Figure 2c shows the comparison of evolution runs with and without using crossover. The fitness function value after running evolution runs with 150 population iterations was clearly lower by more than one order of magnitude for a large majority of evolution runs when using crossover, confirming the feasibility of it's use for our particular group of problems.

In a similar manner, we evaluated the use of the proposed adaptive fitness evaluation function explained in Section Framework. By first evaluating only the frequency of the system response and after the population exceeded a predetermined error threshold (in our case $< 1\%$), we started with the fitness evaluation function in Equation 4 until reaching the target frequency accuracy and only then proceeding with Equation 3. Thus evolution runs were able to achieve faster convergence. Consequently, the global error value is lower for evolutions using the adaptive error evaluation function, which is clearly seen on Figure 2d.

The hypothesis posed in Section Framework that some parameters influence either amplitude or frequency while others influence both was directly tested. Running various single oscillator simulations with gradually tuning two chosen parameters, we discover that both transcription rate r_M and translation rate r_P only affect the amplitude (Figure 4c), while none of them affects the resulting frequency (Figure 4d). This fact explains the advantage of the proposed adaptive fitness evaluation function, which offers the possibility to first find a solution with a satisfying frequency accuracy and subsequently tuning the parameters which influence amplitude without affecting the obtained frequency. This is an important result since it not only decreases the search time and reduces the fitness function dimensionality, but also improves the overall quality of the end solution.

To investigate whether the proposed algorithm is able to find a global optimum, we performed runs of 1000 independent evolutions with the same initial values (Figure 3). The resulting statistics for the final value of each individual parameter are modal distributions with a clearly distinguishable peak, indicating the algorithm ends the search near the global optimum with high probability.

As shown in ref. Law and Szeto (2007), the amount of information carried by an individual parameter is directly proportional to it's standard deviation across the population of candidate solutions. We calculate the entropy of parameter value distribution, similarly as proposed in ref. François and Hakim (2004) as a convenient way to quantify evolution determinism. For each individual parameter p_j a histogram with b bins is computed and it's entropy is calculated

$$H(p_j) = \sum_{k=1}^{b} c_{k,j} \cdot \log(c_{k,j}) \tag{8}$$

where $c_{k,j}$ presents the probability of parameter j having the value in the interval of the bin k.

As can be observed from the results obtained in Figure 3, the parameter that most notably influences the overall solution fitness (bears least entropy in value distribution) is time delay τ (consistent with findings in ref. (Scheper et al, 1999; Novák and Tyson, 2008)), while the influence of the scaling constant k is among the least noticable. Running the simulations by adjusting these two parameters while keeping the rest constant (at their calculated mean values) clearly shows the time delay τ influences both amplitude (Figure 4a) and frequency (Figure 4b), while the influence of k on both is negligible. Thus, the entropy proves an effective measure to quantify individual parameter importance for the measured response and the results can be applied in the design process of synthetic *in vivo* biological circuits.

Repressilator model

Proposed approach was tested on the model with a larger number of interacting genes (Elowitz and Leibler, 2000) in order to investigate it's scalability. The general repressilator model is composed of an odd number of repressors (Figure 5a) such that

$$\frac{\delta M_x}{\delta t} = -\gamma \cdot M_x(t) + \frac{\alpha_1}{1 + (\frac{P_y(t)}{k})^h} + \alpha_0 \tag{9}$$

$$\frac{\delta P_x}{\delta t} = -\eta \cdot P_x(t) + \beta \cdot M_x(t - \tau) \tag{10}$$

where $x, y \, \epsilon \, 1, ..., N$ is the repressor protein index, M_x is the mRNA concentration of the corresponding repressor P_x. Parameters that have to be determined are mRNA degradation rate γ, maximal (unrepressed) mRNA production rate α_1, promoter leakage α_0, repressor scaling constant k, cooperativity coefficient h, protein degradation rate η, protein production rate β and transcription-translation time delay τ. In case of the repressilator with three repressors, the model is composed of 6 equations and 9 parameters. An example simulation run of the repressilator is shown on Figure 5b. The parameters found after 150 generations cause oscillating behaviour with shifted phases among the three repressor proteins as anticipated.

As the amplitude and frequency of the oscillations are mainly dependent on the stability of the protein (Elowitz and Leibler, 2000), we modified the model such that

$$\frac{\delta M_x}{\delta t} = -\gamma \cdot M_x(t) + \frac{\alpha_1}{1 + (\frac{P_y(t)}{k_x})^h} + \alpha_0 \tag{11}$$

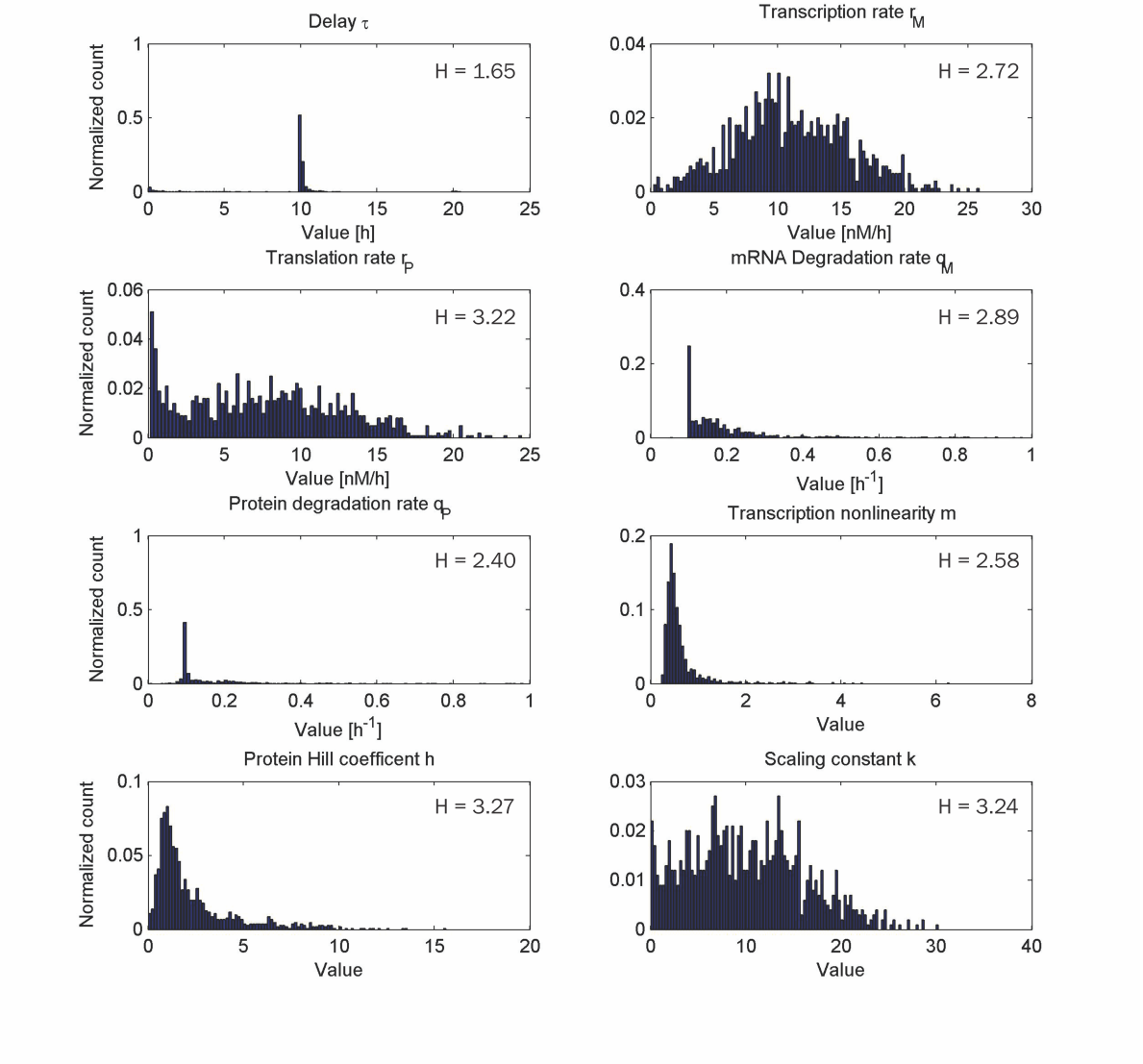

Figure 3: Distribution of individual parameter values for 1000 independent evolution runs with normalized counts and calculated entropies using Equation 8.

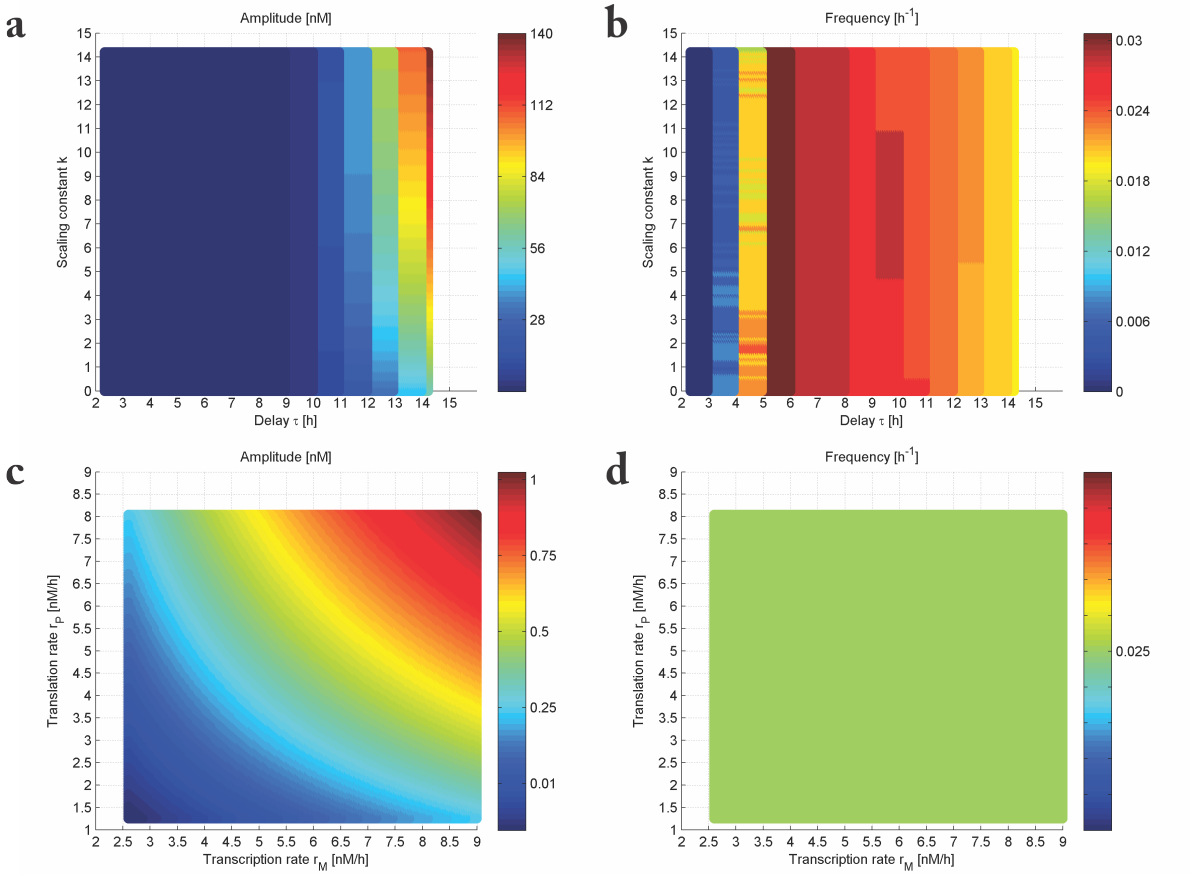

Figure 4: Effect on resulting amplitude and frequency of the single repressor oscillator model for tuning of the two parameters for a range $[\mu - \sigma, \mu + \sigma]$ where μ and σ are the mean and the standard deviations of parameter values obtained from a run of 1000 parallel simulations (Figure 3). Other parameters are kept constant at their mean values. a) Effect on amplitude by tuning parameters τ and k. b) Effect on frequency by tuning parameters τ and k. c) Effect on amplitude for parameters r_M and r_P. d) Effect on frequency for parameters r_M and r_P.

$$\frac{\delta P_x}{\delta t} = -\eta_x \cdot P_x(t) + \beta_x \cdot M_x(t-\tau) \tag{12}$$

Here the scaling constant k_x, protein degradation rates η_x and translation rates β_x unique for each protein in the circuit, which increases the number of parameters to 15. Results in Figure 5c and Figure 5d present relative error margins for end results of simulations for various orders of magnitude for targeted amplitudes and frequencies. The errors are below 0.01 % in majority of cases, which underlines the feasibility and the scalability of the approach. Also, by tuning separately a subgroup of parameters β_x, η_x, k_x while keeping other parameters equal for each of the three components of the network, we show that the frequency and amplitude response of the repressilator can be adjusted with the choice of interacting repressors' properties alone. This observation is constant with experimental studies (Elowitz and Leibler, 2000) while quantifying the observed information as well.

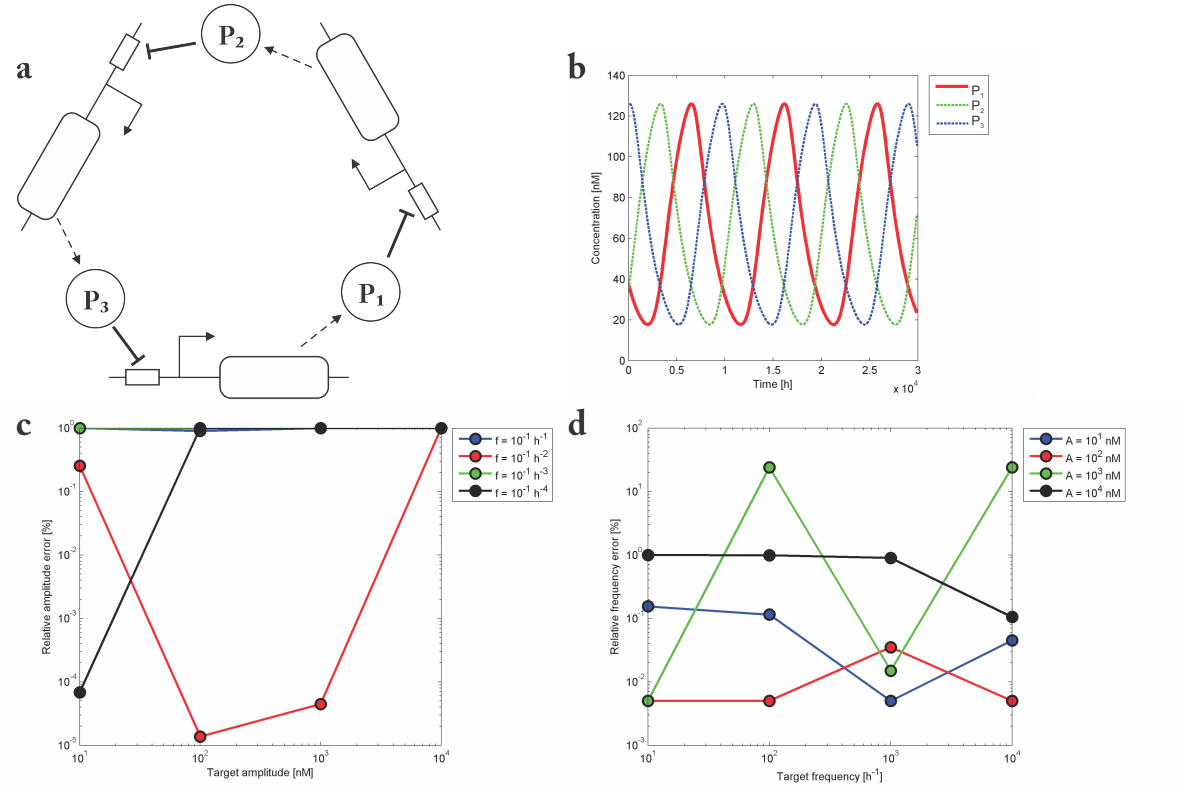

Figure 5: a) Schematic representation of the repressilator circuit b) Simulation of optimal solution of parameter space search for target amplitude $A = 100$ and frequency $f = 0.0001h^{-1}$. Parameters used: $\alpha_0 = 2.288^{-7}$, $\alpha_1 = 1.94$, $\beta = 0.0002$, $n = 4.199$, $\gamma = 0.0005$, $\eta = 0.007$, $k = 45.313$, $\tau = 0.09$. c) Comparison of parameters space investigation for amplitudes and frequencies of varying orders using Equations 9 and 10 (circles) or Equations 11 and 12 (squares). The y axis shows the relative amplitude error regarding the target amplitude. d) Comparison of parameters space investigation for amplitudes and frequencies of varying orders of magnitude Equations 9 and 10 (circles) or Equations 11 and 12 (squares). The y-axis shows the relative frequency error regarding the target frequency.

Discussion

We have developed a method to explore the parameter space and performed a thorough parameter analysis for biological oscillator circuit models with a predefined topology. By making use of *a priori* knowledge of the observable system dynamics, we were able to further narrow the parameter values while retaining the ability of finding the values that lead to the desired system response. The former is represented by a combination of oscillating amplitude and frequency which can be adjusted over a wide range of orders of magnitude. We efficiently applied digital signal processing methods to analyze the oscillating behavior to evaluate the candidate solutions.

The proposed genetic algorithm proved effective for two different types of models and wide range of input parameters which indirectly confirmed it's scalability. Comparing the evolution with and without using crossovers, we have concluded that it significantly improves the end result as a way of conserving fit subparts of parameter vectors. It reinforces the fact that the systems response is mainly dependent on ratios between different parameters. Hence the fitter combinations of a subgroup of parameters can be efficiently spread over candidate solutions pool (population). We included state-of-the-art developments in genetic algorithms by allowing multiple mutations, biased crossover towards fitter candidate solutions, partitioning the solutions on their feasibility and proposed a new approach of a problem dependent adaptive fitness evaluation, based on the effect of single parameters to the resulting system response.

Being able to perform giant leaps in parameter space by means of crossover and multiple mutations, the system successfully converges towards the optimal solution and avoids local maxima. Due to the use of elitism, convergence is always monotonic. By running various simulations with equal inputs, the simulations converged towards a vector with a clear, modal statistical distribution for each component (parameter). As the starting candidate solutions are purely random, we assert the vector of means is the global optimum with high probability. Each parameter was assigned an entropy value, which allowed us to determine and quantify the amount of information and the importance for a single parameter.

As we have shown the approach is applicable to oscillating systems with varying complexity, we turn towards the applicability of it's results. By assessing separately the tunable parameters used in the repressilator model while keeping the remaining untunable parameters constant, we conclude amplitude and frequency of oscillators can be adjusted only by altering the protein dependent parameters. Thus, the resulting vector of parameter values can be used to aid the decisions during design and experimentation of *in vivo* synthesis and analysis of biological circuits.

Acknowledgements

The research was supported by the scientific research programme Pervasive Computing (P2-0359) financed by Slovenian Reseach Agency in years from 2009 - 2012. Results presented here are in scope of PhD thesis that is being prepared by Martin Stražar.

References

Alon U (2007) An introduction to systems biology: design principles of biological circuits. Chapman & Hall/CRC, New York, NY

Back T (1992) Self-adaptation in genetic algorithms. Towards a Practice of Autonomous Systems Proceedings of the First European Conference on Artificial Life (1992), MIT Press pp 263–271

Butz M, Neuenschwander M, Kast P, Hilvert D (2011) An N-terminal protein degradation tag enables robust selection of highly active enzymes. Biochemistry 50(40):8594–602

Chambers LD (2000) The Practical Handbook of Genetic Algorithms: Applications. Chapman and Hall/CRC, Perth

Dill Ka, Fiebig KM, Chan HS (1993) Cooperativity in protein-folding kinetics. Proceedings of the National Academy of Sciences of the United States of America 90(5):1942–6

Elowitz MB, Leibler S (2000) A synthetic oscillatory network of transcriptional regulators. Nature 403(6767):335–8

Fang F, Ni BJ, Yu HQ (2009) Estimating the kinetic parameters of activated sludge storage using weighted non-linear least-squares and accelerating genetic algorithm. Water research 43(10):2595–604

Farmani R, Wright J (2003) Self-adaptive fitness formulation for constrained optimization. IEEE Transactions on Evolutionary Computation 7(5):445–455

François P, Hakim V (2004) Design of genetic networks with specified functions by evolution in silico. Proceedings of the National Academy of Sciences of the United States of America 101(2):580–5

Hajimorad M, Gray PR, Keasling JD (2011) A framework and model system to investigate linear system behavior in Escherichia coli. Journal of biological engineering 5(1):3

Law N, Szeto K (2007) Adaptive genetic algorithm with mutation and crossover matrices. In: Proceedings of the 20th International Joint Conference on Artificial Intelligence, Hyderabad, India, i, pp 2330–2333

Lillacci G, Khammash M (2010) Parameter estimation and model selection in computational biology. PLoS computational biology 6(3):696

Lillacci G, Khammash M (2012) A distribution matching method for parameter estimation and model selection in computational biology. International Journal of Robust and Nonlinear Control

Mackey M, Glass L (1977) Oscillation and chaos in physiological control systems. Science 197(4300):287–289

Morton aJ, Wood NI, Hastings MH, Hurelbrink C, Barker Ra, Maywood ES (2005) Disintegration of the sleep-wake cycle and circadian timing in Huntington's disease. The Journal of neuroscience : the official journal of the Society for Neuroscience 25(1):157–63

Novák B, Tyson JJ (2008) Design principles of biochemical oscillators. Nature reviews Molecular cell biology 9(12):981–91

Paladugu S, Chickarmane V (2006) In silico evolution of functional modules in biochemical networks. Systems biology 153(4):223–236

Rhodius Va, Mutalik VK, Gross Ca (2012) Predicting the strength of UP-elements and full-length E. coli σE promoters. Nucleic acids research 40(7):2907–24

Rodrigo G, Jaramillo A (2012) AutoBioCAD: Full Biodesign Automation of Genetic Circuits. ACS Synthetic Biology

Rodrigo G, Carrera J, Jaramillo A (2007) Genetdes: automatic design of transcriptional networks. Bioinformatics (Oxford, England) 23(14):1857–8

Roy SW (2009) Probing evolutionary repeatability: neutral and double changes and the predictability of evolutionary adaptation. PloS one 4(2):e4500

Salis HM, Mirsky Ea, Voigt Ca (2009) Automated design of synthetic ribosome binding sites to control protein expression. Nature biotechnology 27(10):946–50

Scheper T, Klinkenberg D, Pennartz C, van Pelt J (1999) A mathematical model for the intracellular circadian rhythm generator. The Journal of neuroscience : the official journal of the Society for Neuroscience 19(1):40–7

Smith J, Fogarty T (1996) Self adaptation of mutation rates in a steady state genetic algorithm. In: Evolutionary Computation, 1996., Proceedings of IEEE International Conference on, pp 318–323

Szendro IG, Franke J, de Visser JAGM, Krug J (2013) Predictability of evolution depends nonmonotonically on population size. Proceedings of the National Academy of Sciences of the United States of America 110(2):571–6

Tigges M, Marquez-Lago TT, Stelling J, Fussenegger M (2009) A tunable synthetic mammalian oscillator. Nature 457(7227):309–12, DOI 10.1038/nature07616, URL http://www.ncbi.nlm.nih.gov/pubmed/19148099

Turek FW, Joshu C, Kohsaka A, Lin E, Ivanova G, McDearmon E, Laposky A, Losee-Olson S, Easton A, Jensen DR, Eckel RH, Takahashi JS, Bass J (2005) Obesity and metabolic syndrome in circadian Clock mutant mice. Science (New York, NY) 308(5724):1043–5

Vecchi SKCDGMP (1987) Optimization by Simulated Annealing, vol 220. Science, New Series, Vol. 220, No. 4598. (May 13, 1983), pp. 671-680.

Weinreich DM, Chao L (2005) Rapid evolutionary escape by large populations from local fitness peaks is likely in nature. Evolution; international journal of organic evolution 59(6):1175–82

Wu YH, Fischer DF, Kalsbeek A, Garidou-Boof ML, van der Vliet J, van Heijningen C, Liu RY, Zhou JN, Swaab DF (2006) Pineal clock gene oscillation is disturbed in Alzheimer's disease, due to functional disconnection from the "master clock". FASEB journal : official publication of the Federation of American Societies for Experimental Biology 20(11):1874–6

Yokobayashi Y, Weiss R, Arnold FH (2002) Directed evolution of a genetic circuit. Proceedings of the National Academy of Sciences of the United States of America 99(26):16,587–91