Robust network clocks: Design of genetic oscillators as a complex combinatorial optimization problem

Yasuaki Kobayashi,[1,*] Tatsuo Shibata,[2,3] Yoshiki Kuramoto,[4] and Alexander S. Mikhailov[1]

[1]*Department of Physical Chemistry, Fritz Haber Institute of the Max Planck Society, Faradayweg 4-6, D-14195 Berlin, Germany*
[2]*Laboratory for Physical Biology, RIKEN Center for Developmental Biology, Kobe 650-0047, Japan*
[3]*PRESTO, Japan Science and Technology Agency, 4-1-8 Honcho Kawaguchi, Saitama, Japan*
[4]*Research Institute for Mathematical Science, Kyoto University, Kyoto 606-8502, Japan*

Complex combinatorial optimization can be used to design network systems having desired dynamics and that are robust against structural perturbations. Here genetic networks exhibiting limit-cycle oscillations with prescribed periods and, furthermore, that are robust against the deletion of links and nodes or the application of noise are constructed. Large ensembles of robust genetic clocks with different periods could thus be obtained, and some of their statistical properties have been investigated. Similar methods can be used to design robust network oscillators of various origins.

PACS number(s): 87.18.Vf, 05.45.Xt, 87.18.Cf

A fundamental property of biological systems, including genetic networks, is that they are able to retain functional operation despite fluctuations and local damage [1]. As evidenced by knockout experiments, only 13% of all genes in yeast are essential [2]. The addition of a new regulatory link into the genetic network of *Escherichia coli* is, in 95% of cases, tolerated by the bacteria [3]. The circadian rhythm generator, a genetic clock of the cells, is able to maintain its period with high precision, despite variations of temperature and other external parameters [4]. After an oscillatory three-gene circuit (the repressilator) was proposed and experimentally implemented by Elowitz and Leibler [5], construction of genetic networks with prescribed dynamics has become a rapidly developing research field [6–8]. How do we design synthetic genetic networks with prescribed functions as robust as their real biological counterparts?

Already, N. Wiener has suggested that negative feedbacks can stabilize steady states [9]. Indeed, various feedbacks are present in genetic systems and reduce fluctuations in gene expression levels (see the review in Ref. [10]). For oscillator networks, robustness has been treated as an aspect of synchronization and global phase diffusion [11]. In addition to dynamical feedbacks, functional robustness can be increased through the use of special error-correcting network architectures [12]. Genetic networks with Boolean dynamics have been considered as being robust if their dynamics attractors, such as limit cycles, are not destroyed by the knockout of nodes or links [13].

Construction of dynamical networks with prescribed dynamics and improved robustness can be viewed as a problem of complex combinatorial optimization [14]. As demonstrated, e.g., for the traveling salesman problem, approximate optimization methods, such as simulated annealing, can be efficient for this class of systems [15]. Evolutionary optimization methods have been employed to design genetic networks with static Turing patterns [16], model signal

transduction networks [17], and oscillatory genetic networks with prescribed periods [18]. Networks of phase oscillators with improved synchronization properties have furthermore been obtained using stochastic Monte Carlo optimization with replica exchange [19].

Here we demonstrate how the simulated annealing method can be applied to design genetic networks that maintain their oscillation periods within a given tolerance window despite different perturbations, such as knockout of links or nodes and the introduction of distributed quenched noise.

We consider a genetic network with N genes. Some genes are regulated by other genes. All regulations are inhibitory, and we do not exclude the possibility of autoregulations. For simplicity, transcription and translation are treated as a single process, so that the only variables are gene expression levels [20]. The dynamic equations of the model are (see also [18])

$$\frac{du_i}{dt} = \frac{1}{1 + \left(\sum_{j=1}^{N} A_{ij} \phi_{ij} u_j \right)^n} - u_i, \qquad (1)$$

where u_i is the gene expression level of gene i ($1 \leqslant i \leqslant N$). The adjacency matrix A_{ij} determines the network structure ($A_{ij} = 1$ if a regulation from gene j to i exists, and otherwise, $A_{ij} = 0$). The equations are written in a dimensionless form, and the only remaining parameters are the regulation strength ϕ_{ij} and the Hill coefficient n. As the output signal, the expression level $u_1(t)$ of the first gene is chosen.

Suppose that our target dynamics is limit-cycle oscillations with the period T_0. We want to quantify the distance between the dynamics produced by a given network and the target. To evaluate the dynamics [18], we observe the signal over a fixed time interval, twice as large as the target period T_0, and determine its maximal and minimal values, $u_{\max}$ and $u_{\min}$, over this interval. The threshold level h is set as $h = (u_{\max} + u_{\min})/2$. We examine the output signal and determine all moments t_i when it crosses the threshold h from below. Thus, we also determine a set of time differences $\Delta_i = t_i - t_{i-1}$ between the crossing events. The average period is $T = \langle \Delta_i \rangle = (1/K) \sum_{i=1}^{K} \Delta_i$, and its variance is
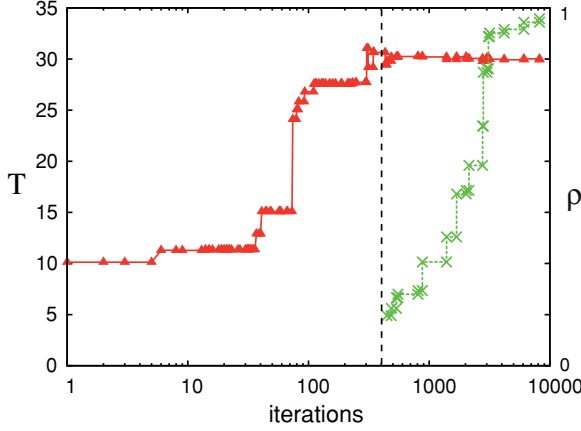
FIG. 1. (Color online) Example of the optimization process. The dependencies of the oscillation period (red triangles) and of the robustness against static noise (green crosses) on the iteration time are shown. The target period is $T_0 = 30$.

$\sigma^2 = \langle \Delta_i^2 \rangle - \langle \Delta_i \rangle^2$, where K is the total number of crossing events (we chose $K = 10$). The error function is introduced as

$$\epsilon = \frac{(T - T_0)^2}{T_0^2} + \frac{\sigma^2}{T^2}. \tag{2}$$

We shall say that a network is *functional* if its error ϵ does not exceed some threshold ϵ_{th}.

Robustness of a functional network is understood by us as its ability to retain the function despite structural perturbations. If only n different perturbations are possible, robustness ρ is defined as

$$\rho = \frac{1}{n} \sum_{i=1}^{n} \theta(\epsilon_{th} - \epsilon_i), \tag{3}$$

where ϵ_i is the error of a network under a perturbation i and $\theta(x) = 1$ for $x > 0$ and is zero otherwise. Generally, $\rho = \langle \theta(\epsilon_{th} - \epsilon_i) \rangle$, with statistical averaging over a perturbation ensemble.

The evolutionary optimization method represents a variant of simulated annealing and employs a sequence of network "mutations," each of which is followed by a decision to retain or to discard a change. A mutation consists of random relocation of a randomly picked link. At the first optimization stage, evolution is aimed at obtaining a functional network with the error below ϵ_{th}; this is done by using the algorithm described in [18]. Once a network with $\epsilon < \epsilon_{th}$ is found, the evolution is switched to the maximization of robustness. Now, for each network its robustness measures ρ and ρ' before and after a mutation are determined. A mutation is always accepted if $\rho' > \rho$. It is accepted with probability $p = \exp[(\rho' - \rho)/(1 - \rho)\mu]$ if $\rho' < \rho$. After each mutation, we also check whether a network is still functional; if $\epsilon' > \epsilon_{th}$, the evolution is changed back to the optimization with respect to the error ϵ. Note that this evolution method is essentially the stochastic Metropolis algorithm, where the role of energy is played by the network robustness ρ and the effective temperature is $(1 - \rho)\mu$. The temperature tends to zero as the solution with $\rho = 1$ is being approached, as assumed in the

simulated annealing method [15]. Both the number of nodes and the number of links are conserved in every evolution.

Each optimization involved up to 10 000 iteration steps. The iteration process was terminated before 10 000 when $\epsilon < \epsilon_{th}$ and $\rho = 1$ were achieved. The model parameters were $n = 3$ and $\phi = 100$. With such parameters, the three-gene repressilator circuit in our model had the oscillation period of $T_{rep} = 10.16$. The error threshold was set to $\epsilon_{th} = 0.0001$, so that the oscillation periods T were allowed to deviate up to 1% from the target period T_0. Alternatively, if $T = T_0$, the variance σ can be as large as $\sigma = 0.01 T_0$. The annealing temperature coefficient was chosen as $\mu = 0.01$. Networks with $N = 10$ genes and $M = 20$ regulatory links were considered.

Three different kinds of structural perturbations were considered: deletion of a link or of a node (except for the output node $i = 1$) and application of static noise. The static noise, which does not change with time, was introduced through random variations of link weights: Every link weight ϕ_{ij} was perturbed as $\phi_{ij} = (1 + S\eta_{ij})\phi$, where S is the perturbation intensity and η_{ij} is a random number with a uniform distribution from the interval $[-1,1]$. In our simulations we chose $S = 0.1$. To determine the robustness of a network, $n = 100$ different static noise realizations were applied.

Figure 1 shows an example of an optimization process. After a functional network with the required period $T_0 = 30$ was found, optimization for robustness began, leading to the final network with the robustness $\rho = 1$, as compared with the robustness $\rho = 0.2$ of the initial functional network. Here and in all other simulations below, we have checked the values of T and σ separately and found that $\sigma = 0$ was always achieved; that is, the resulting dynamics always corresponded to periodic oscillations. In Fig. 2, examples of designed networks that are robust against noise and deletion of links or nodes are displayed. In Figs. 2(b) and 2(c), deletion of the outlined (red) links or nodes results in the loss of functionality.

By repeated optimizations, ensembles of networks with the same periods but that are robust against different kinds of perturbations were constructed. Thus, optimization efficiency and statistical properties of designed networks could be
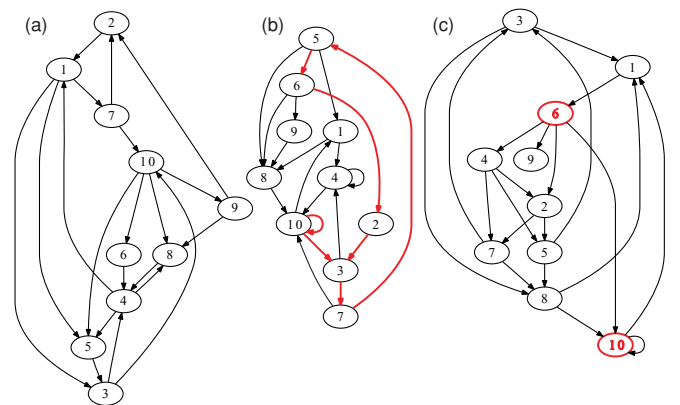

FIG. 2. (Color online) Examples of designed networks producing oscillations with period $T = 15$, optimized for robustness against (a) static random noise, (b) link deletions, and (c) node deletions. Thick red curves indicate the links and the nodes whose deletion brings a network outside of the tolerance window.
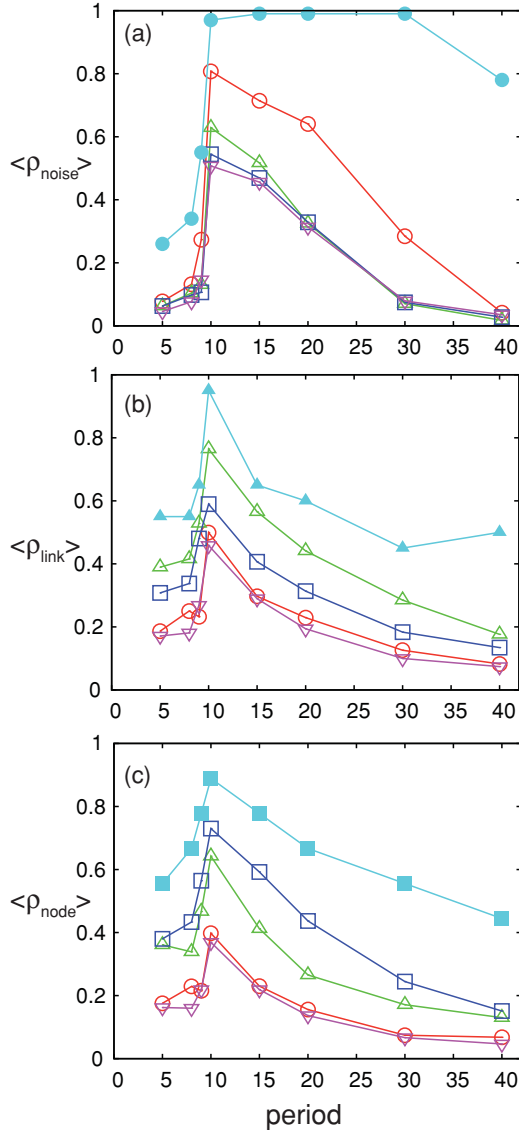
FIG. 3. (Color online) Average robustness against (a) noise, (b) link deletions, and (c) removal of nodes for the ensembles of networks optimized for noise (open red circles), deletion of links (open green triangles), and deletion of nodes (open blue squares). Additionally, the respective data for the control ensembles of the networks with different periods, which have not been optimized for any kind of robustness, are displayed (upsidedown pink triangles). The maximum robustness achieved in the optimizations against each kind of noise for a given oscillation period is also shown (solid symbols). Networks are of size $N = 10$ with $M = 20$ links and noise level $S = 0.1$. For each target period T_0, an ensemble of 100 designed networks is taken.

discussed, depending on the target period. Results of the statistical analysis are presented in Fig. 3.

The solid blue circles in Fig. 3(a) show the best robustness against noise ($S = 0.1$) that could be achieved. In the interval of oscillation periods from $T_0 = 10$ to $T_0 = 35$, networks with almost 100% robustness could be found. It was not, however, possible to construct robust networks with $T_0 < 10$, i.e., with the periods shorter than that of the three-gene repressilator. The open red circles in Fig. 3(a) show the average noise robustness achieved in the optimization runs. The optimization
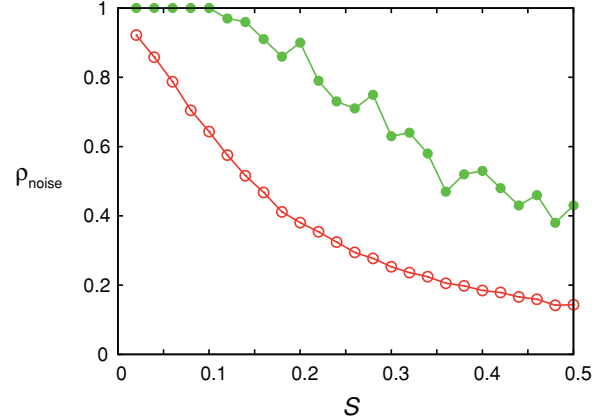

FIG. 4. (Color online) Robustness against noise with different noise intensities S for a designed network ensemble with $T = 20$. Evolutionary optimization is done with $S = 0.1$. Averages over an ensemble of 100 networks (open red circles) and maximum values (solid green circles) are shown.

efficiency deteriorates with the target period, and it becomes increasingly difficult to find a network with high robustness as T_0 is increased. The upsidedown pink triangles show the respective data for ensembles of functional networks without robustness optimization. We see that, on average, robustness against noise can be largely enhanced through architecture optimization.

Similar results are found when optimization against deletion of links or nodes is considered [Figs. 3(b) and 3(c)]. Both for the maximal values and the statistical averages, a substantial increase of robustness as compared to the control ensembles of networks without optimization could be reached. Note, however, that robust networks with short periods (less than 10) and with relatively large periods could not be found.

Networks optimized for one kind of perturbations are not generally optimal against the perturbations of different kinds. Figure 3(a) shows average robustness against noise for the networks optimized for link deletions (green triangles) and node deletions (blue squares). Their robustness is only slightly better than, or almost the same as, that of the control ensemble. Similar behavior is seen in the case of robustness against link deletions [Fig. 3(b)] and node deletions [Fig. 3(c)].

The networks robust against noise were designed above by applying noise with a fixed intensity $S = 0.1$. In Fig. 4, the maximum and average robustness of such networks under the increase of S is shown. As could be expected, robustness diminishes as the noise gets stronger. It is, however, remarkable that even when the regulation weights are randomly varied up to 25% ($S = 0.25$), the best designed networks can still maintain the oscillation periods with an accuracy of 1% in about 70% of noise realizations ($\rho_{noise} = 0.7$).

Are there any systematic statistical differences in the structural properties of networks robust against different kinds of perturbations? For three optimized network ensembles, optimized against noise, link deletions, and node deletions, and for the control ensemble of functional-only networks with period $T = 20$, we have computed the average clustering coefficients C and the characteristic path lengths l [21]: $C = 0.132 \pm 0.070$, 0.136 ± 0.069, 0.144 ± 0.067, and $0.136 \pm$

0.070, respectively, and $l = 2.14 \pm 0.16$, 2.28 ± 0.20, 2.28 ± 0.18, and 2.13 ± 0.14, respectively. The differences are small and lie almost within the statistical dispersion. Note also that all our networks have the same mean degree of $2M/N = 4$. Thus, classical properties, such as the mean degrees, path lengths, and clustering coefficients, are not enough to distinguish networks robust against various kinds of perturbations, and more refined statistical analysis, which will be undertaken in a separate study, is required.

Our study, using simple model networks, reveals that robustness can easily emerge in the process of evolutionary optimization and, probably, could also have easily arisen in biological evolution. The plasticity of genetic networks is astonishing. In a small genetic circuit with only 10 genes and 20 regulations, mere rewiring of connections yields thousands of limit-cycle oscillators with largely varying periods. Genetic oscillators, which keep their period with an accuracy of 1% under deletion of links or knockout of nodes, could be obtained in large numbers. Robust genetic clocks retaining the oscillation period with this accuracy even when regulation weights were randomly varied up to $\pm 20\%$ were constructed.

The architectures of robust network clocks strongly depend on the kind of structural perturbations against which they need to be resistant. The networks robust against distributed noise are typically not robust against deletion of nodes or links. Apparently, different mechanisms of error correction are being used by the networks to ensure the stability of their function despite structural perturbations. The mechanisms of robustness need also to be further investigated. Optimization for robustness breaks down for periods shorter than the period

$T_{\text{rep}} = 10.16$ of the repressilator circuit. Repressilators may play a special role in robust oscillatory circuits. The designed networks often included such three-gene ring structures as elementary motifs, and some synchronization phenomena could have been involved. Note, however, that all interactions had equal strength, and therefore, the networks did not represent ensembles of weakly coupled oscillators, where classical synchronization analysis is applicable.

In a model study, we have demonstrated that genetic networks with a degree of robustness comparable to their real biological counterparts can be easily designed by evolutionary optimization methods. Remarkably, high levels of robustness can be already achieved in networks with relatively small numbers of genes and with only inhibitory interactions. It would be interesting to apply similar methods for the construction of synthetic genetic networks that may be experimentally implemented. Moreover, our analysis can provide a reference frame for better understanding of natural genetic networks, which are responsible, e.g., for circadian rhythm generation. Finally, it should be mentioned that, although only genetic networks have been considered here, the design method is general and can be used to construct robust functional dynamical networks of various origins. While robustness against only static damage or structural perturbations has been investigated, similar methods can be used to improve robustness of dynamical networks with respect to the dynamical noise as well.

[1] M. A. Savageau, Nature (London) **229**, 542 (1971).

[2] D. Deutscher *et al.*, Nat. Genet. **38**, 993 (2006).

[3] M. Isalan *et al.*, Nature (London) **452**, 840 (2008).

[4] J. C. Dunlap, J. J. Loros, and P. J. DeCoursey, *Chronobiology: Biological Timekeeping* (Sinauer, Sunderland, MA, 2004).

[5] M. B. Elowitz and S. Leibler, Nature (London) **403**, 335 (2000).

[6] J. Stricker *et al.*, Nature (London) **456**, 516 (2008).

[7] M. Tigges *et al.*, Nature (London) **457**, 309 (2009).

[8] T. Danino *et al.*, Nature (London) **463**, 326 (2010).

[9] N. Wiener, *Cybernetics: Or Control and Communication in the Animal and the Machine* (MIT Press, Cambridge, MA, 1948).

[10] C. V. Rao, D. M. Wolf, and A. P. Arkin, Nature (London) **420**, 231 (2002).

[11] H. Kori, Y. Kawamura, H. Nakao, K. Arai, and Y. Kuramoto, Phys. Rev. E **80**, 036207 (2009).

[12] P. Kaluza, M. Vingron, and A. S. Mikhailov, Chaos **18**, 026113 (2008).

[13] S. Bornholdt and K. Sneppen, Proc. R. Soc. London, Ser. B **267**, 2281 (2000).

[14] M. R. Garey and D. S. Johnson, *Computers and Intractability* (Freeman, San Francisco, 1979).

[15] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, Science **220**, 671 (1983).

[16] K. Fujimoto, S. Ishihara, and K. Kaneko, PLoS ONE **3**, e2772 (2007).

[17] P. Kaluza, M. Ipsen, M. Vingron, and A. S. Mikhailov, Phys. Rev. E **75**, 015101 (2007); P. Kaluza and A. S. Mikhailov, Europhys. Lett. **79**, 48001 (2007).

[18] Y. Kobayashi, T. Shibata, Y. Kuramoto, and A. S. Mikhailov, Eur. Phys. J. B **76**, 167 (2010).

[19] T. Yanagita and A. S. Mikhailov, Phys. Rev. E **81**, 056204 (2010).

[20] S. Ishihara, K. Fujimoto, and T. Shibata, Genes Cells **10**, 1025 (2005).

[21] The characteristic path length is calculated as the inverse of the average l_{ij}^{-1}, where l_{ij} is the shortest path between nodes i and j. If the path from i to j does not exist, we set $l_{if}^{-1} = 0$.