Optimal parameter settings for information processing in gene regulatory networks

Dominique F. Chu [a],*, Nicolae Radu Zabet [a], Andrew N.W. Hone [b]

[a] *School of Computing, University of Kent, CT2 7NF, Canterbury, UK*
[b] *School of Mathematics, Statistics and Actuarial Science, University of Kent, CT2 7NF, Canterbury, UK*

ABSTRACT

Gene networks can often be interpreted as computational circuits. This article investigates the computational properties of gene regulatory networks defined in terms of the speed and the accuracy of the output of a gene network. It will be shown that there is no single optimal set of parameters, but instead, there is a trade-off between speed and accuracy. Using the trade-off it will also be shown how systems with various parameters can be ranked with respect to their computational efficiency. Numerical analysis suggests that the trade-off can be improved when the output gene is repressing itself, even though the accuracy or the speed of the auto-regulated system may be worse than the unregulated system.

© 2011 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Living systems depend crucially on their ability to adjust to changes of both internal or external conditions or combinations thereof. Examples are seasonal changes, nutrient availability, encounters with predators, prey or mating partners and the like. Reacting to such changes requires the organism to perform computations. Higher animals have specialised nervous systems to this end. However, even within unicellular organisms, such as bacteria, there are computational processes going on at a molecular level. Individual proteins have been described as performing computational functions (Bray, 1995). Another example of systems that can be considered as implementing computations are genes.

The idea that networks of genes can perform computational functions in the cell is now well established (see Fernando et al., 2009; Haynes et al., 2008; Ben-Hur and Siegelmann, 2004; Ziv et al., 2007). As a simple example consider the regulation of metabolic pathways in bacteria. The enzymes necessary to metabolize a specific nutrient will only be activated if the nutrient is present in the environment (see for example Chu et al., 2008), this metabolic control implements a basic if-then statement. In this case a bacterial cell takes as input the concentration of an external nutrient concentration, and "computes" an output, namely whether or not to turn on a metabolic pathway. This is probably the simplest exam-

ple of computing one could imagine. The next step in biological software sophistication is inducer exclusion (Narang, 2006; Narang and Pilyugin, 2007), i.e. the process whereby prokaryotic cells faced with several concurrently available nutrient sources exclusively take up the highest quality one. Beyond these examples, there are many more complicated instances of gene networks in cells where environmental information is integrated from a variety of sources to decide whether or not to turn on a gene (Mattick and Gagen, 2001).

The "computation" metaphor may not be useful in all contexts in theoretical biology, but in the case of gene networks it can be very enlightening. A network of genes that relays information from some input to some output can be parameterised in a number of ways. To see this consider the simplest model of the output y of a single gene that is regulated by some molecule with concentration x. This is a differential equation with a Hill functions as growth term:

$$\dot{y} = \alpha + \beta \frac{x^h}{K^h + x^h} - \mu_y y.$$

Here α and β are the leak rate and the maximal growth rate respectively, h and K are kinetic parameters of the Hill function, and x is the concentration of the regulator molecule. Depending on the particular parameters of the system, the gene activation function will be a concave ($h \leq 1$) or sigmoid ($h > 1$) function of the concentration of the input molecules; the sensitivity of the function to a particular range of input concentrations depends on K; α and β determine the output range of the gene. A gene network would be

---

* Corresponding author.
  E-mail address: D.F.Chu@kent.ac.uk (D.F. Chu).

modelled as a set of differential equations of the above type, where the network topology is determined by the dependencies between the equations.

For a cell, the particular choice of parameters is strongly constrained by a number of factors: (i) The kinetic parameters of each gene must be such that each gene can be regulated over a physiologically relevant range of the input x. Analogous considerations apply to the output range. (ii) The gene networks need to be able to process input sufficiently rapid, that is they need to change gene expression patterns, i.e. switch between steady states of protein concentrations, within a time that is appropriate given the biological context of the genes. Within the computational metaphor this can be thought of as the computational speed of a gene network. (iii) The parameters' impact on the metabolic cost must also be taken into account. Expressing proteins requires the cell to use ATP molecules; an expression rate β of a gene comes at a continuous cost per time unit that is proportional to β. This presumably exerts a strong adaptive pressure on organisms that compete for limited resources. (iv) Proteins are discrete units expressed and degraded stochastically. This means that even at steady state concentrations will be subject to random fluctuations (noise). Noise can be beneficial for the cell in some contexts, but is usually detrimental when the function of a gene network is to relay/process information. The fluctuations around the signal make it harder for the cell to tell apart two signals that are close together, hence reducing the ability of a gene network to distinguish between signals.

This uncertainty can be reduced if the cell averages the signals over a longer period of time. The intuition behind this is straightforward: Assume a series of data points and assume further that it is known that these data points are distributed around a mean value m_1 or m_2. One can now ask how many data points need to be examined before one can, with a given level of confidence, decide between the two mean values. The answer is: It depends! More precisely it depends on the level of noise that affects the data and the distance between m_1 and m_2. The closer they are together and the higher the noise the more sample points must be taken. Or, for a given level of noise, the more points are taken, the more confident one can be in determining the true mean value.

This is precisely the situation for a binary gene sensing its regulatory input. The more input data it collects, the more accurate the cell senses the activation state of the binary input, which implies that it will be more accurate in its output. However, collecting more input data takes time. What is more, mathematically it can be seen that it also entails that the gene switches slower between its two possible output levels. This manifests itself through a reduced computational speed of the gene. Hence, there is a trade-off between the accuracy of a cellular "computation" and its speed. This trade-off has been described n a recent contribution by Zabet and Chu (Zabet and Chu, 2009; Zabet et al., 2010), where it has been shown for the (near trivial) case of a single gene that noise and switching time cannot be separately optimised. They described a trade-off between the levels of noise of the output of a gene (which they call accuracy), the metabolic cost of expressing the gene, and the speed with which changes can be processed. Formally, this manifests itself through the relationship of the switching time and the output noise with the decay rate of the output of the gene; these are proportional and inversely proportional, respectively.

The question of noise and speed of gene networks has previously received considerable attention in the community. In particular the impact of local network topologies, so-called *network motifs* (Oltvai and Barabasi, 2002; Milo et al., 2002a,b; Ma et al., 2009), on speed and noise has been researched in considerable detail. One of the results of this research is an emerging consensus that negative auto-regulation decreases the noise and switching time of a gene

(Alon, 2006, 2007; Singh and Hespanha, 2009; Wang et al., 2010; Bruggeman et al., 2009). However, many of the relevant studies on networks consider a limited range of parameters only. Moreover, they often do not keep the metabolic cost of sets of parameters they compare fixed. Following Zabet and Chu (2009) this draws into question the comparison. Finally, most studies consider noise and switching speed separately, ignoring the trade-off between them.

Nearly all of the existing literature on stochastic fluctuations in gene networks focuses on noise at a particular steady state, in the sense that standard definitions of noise are normalised with respect to the steady state. This choice can be misleading when it is the aim to understand the computational limits of gene networks. Processing and relaying signals requires that the individual components are able to adequately distinguish between different input signals. The classic model in computer science is to consider only two possible states per component, e.g. "on" or "off." Translated into the realm of gene networks this would mean that each gene can be either turned on (high expression/protein abundance) or turned off (low expression/protein abundance). Within such a binary model noise manifests itself through an uncertainty about the activation state of a gene. Depending on the size of the fluctuation and the difference between the high state and the low state (the signal strength) the cell needs to average out fluctuations by "repeat measurements." The extent to which a given level of noise impairs the "distinguishability" of two signals does not primarily depend on the size of the fluctuations in relations to the steady state, but on the size of the fluctuations in relation to the difference between the two possible states, i.e. the signal strength rather than by the steady state concentration. Hence, the appropriate measure of noise in this context is the size of the fluctuation normalised by the mean signal strength, as discussed in the original article by Zabet and Chu. This seemingly minor modification, however, not only complicates the mathematics considerably, but also alters the optimality properties of the system.

In this article, we will ask the same question as Zabet and Chu (2009), but extend the analysis beyond their minimal network: Given a network of genes that is primarily computational in the above sense, what are the constraints on the choice of parameters? As it turns out, the answer to this question depends on the topology of the network. We therefore limit the scope of this article to networks small enough to allow some exact analysis, thus striking a balance between feasibility and practical relevance. The hope is that the insights gained from this (if not the results themselves) generalize and create the right intuition for later large scale numerical simulation studies.

The main results presented here are as follows: We prove a number of optimality properties of gene networks, including a lower bound on the noise of a gene network. Specifically, we show that the absolute fluctuations of negatively auto-regulated genes are always lower than those of comparable standard genes without negative auto-regulation, while this is no longer true when considering noise normalised by the signal strength. We also describe a methodology based on trade-off curves to compare two models of gene networks in terms of their computational efficiency. We could not find a single set of parameters for which negative auto-repression is inferior in terms of its computational properties.

2. Materials and Methods

2.1. The Model

In this article we consider three genes, G_x, G_y and G_z. The gene G_y is regulated by an input signal x which we assume to change instantaneously from a high state x_H to the low state x_L or *vice versa*. Similarly, G_z is regulated by the product of G_y.
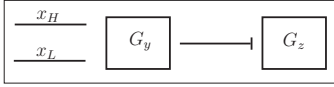
The output gene G_z thus takes one of two possible steady state values depending on the input x.

We model gene expression as a one-step process, thus ignoring the noise contribution from mRNA expression (McAdams and Arkin, 1997; Paulsson, 2004; Raj and van Oudenaarden, 2008). The amount of translational noise depends crucially on the number of proteins produced from a single mRNA transcript and the lifetime of the mRNA transcript in relation to the protein product. While translational noise can be dominating, for most genes it is true that the overall noise is mainly driven by the underlying transcription dynamics (Bar-Even et al., 2006). In terms of the dependence of noise and time on the parameters of the system, this is therefore the main level to consider.

We denote the concentration of gene G_z by z and analogously for G_y. In this article we will consider two regulation functions, the activator and the repressor. At steady state, if the input is x_H then the products of the genes take the values $z^* = z_H$ and $y^* = y_H$, in the case of the activator and $z^* = z_L$ and $y^* = y_H$ in the case of the repressor. The asterisk indicates the steady state value of the relevant concentration. We model the dynamics of the system as a set of two differential equations:

$$\dot{y} = \alpha_y + \omega g(x) - \mu_y y$$
$$\dot{z} = \alpha_z + \beta f(y) - \mu_z z \qquad (1)$$

The symbols α_y and α_z represent the leak expression of the gene in the case of complete de-activation. Non-vanishing leak expression is probably common in biological systems, but is also a source of inefficiency of the system (see Zabet and Chu, 2009); since it does not add any interesting dynamical effects, we will assume $\alpha_y = \alpha_z = 0$ throughout this contribution. Here the functions g and f are assumed to be Hill functions given by:

$$g(x) = \frac{x^{h_2}}{K_2^{h_2} + x^{h_2}}, \qquad f(y) = \frac{y^h}{K^h + y^h}$$

In the case of the repressor the Hill function is given by:

$$f(y) = \frac{K^h}{K^h + y^h}$$

Note that we always assume the gene G_y to be activated by its input. Following Zabet and Chu (2009) we define ζ, the cost of the activator network, as the sum of the maximum production rates of G_y and G_z:

$$\zeta = \omega g_H + \beta f_H.$$

Here $g_H \doteq g(x_H)$; f_H is defined analogously. In the case of the repressor we define the cost as $\zeta = \omega g_L + \beta f_H$, to reflect that at steady state the genes G_y and G_z are never simultaneously activated. In order to keep the total cost of the gene network constant, we assume that the production rate of z is given by

$$\beta = \frac{\zeta - \omega g_H}{f_H}.$$

The system (1) can be solved analytically for $y(t)$, but not for any subsequent gene products; however, $z(t)$ can be given explicitly in terms of a quadrature, namely

$$z(t) = e^{-\mu_z t} \left[\int_0^t e^{\mu_z s} \left(\alpha_z + \beta f(y(s)) \right) ds + C \right], \qquad (2)$$

where the constant C depends on initial conditions, and is zero if there is initially no product, $z(0) = 0$.

In the case of the negative auto-regulator (NAR), we model the system as follows:

$$\dot{z} = \beta f_z(y) \gamma R(z) - \mu_z z \qquad (3)$$

where the negative auto-regulation R is given by

$$R(z) = \frac{\bar{K}^{\bar{h}}}{\bar{K}^{\bar{h}} + z^{\bar{h}}}$$

and $\gamma \doteq R(z_H)^{-1}$.

When assessing the relevance of stochastic fluctuations for sub-cellular information processing, the common definition of noise (variance divided by the mean or square of the mean (Paulsson, 2004; Ozbudak et al., 2002)) may not always be appropriate. In essence, both of these measures relate the stochastic fluctuations at steady state to a (tacitly assumed) zero-concentration baseline. If defined in this way, noise is an indicator of the number of (statistically independent) measurements that are required in order to determine—within a given margin of error—the mean particle number in a cell. In many contexts this may be the relevant question to ask. In the context of information processing by genes it is not. If we assume binary genes, then
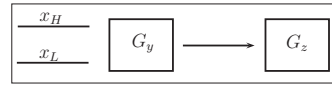
the absolute size of a signal S_1 may not be as important as the ability of the cell to distinguish S_1 from another signal S_2 in the presence of a given amount of stochastic fluctuations. This second signal S_2 may just be the absence of a signal, but this rarely works out to be a true zero baseline. Instead, "no signal" will typically mean "weak signal."

Therefore, in order to assess the impact of the stochastic fluctuations on the computational properties of gene regulatory networks, it is necessary to modify this standard definition of noise. Following Zabet and Chu (2009) we normalise the variance by the square of the signal strength (defined as $z_H - z_L$), rather than the signal itself, which is a measure of how the stochastic fluctuations limit an observer's ability to distinguish one signal from another one. To compute the noise we use Paulsson's version of van Kampen's linear noise approximation (van Kampen, 2007; Paulsson, 2004). For the standard system this yields the formula

$$\mathcal{N}_z = \underbrace{\frac{z_H}{(z_H - z_L)^2}}_{\text{intrinsic}} + \underbrace{\left[\overbrace{\frac{\beta f_H'}{z_H - z_L} \tau_z}^{\text{regulation factor}} \right]^2 \overbrace{\frac{\tau_y}{\tau_y + \tau_z}}^{\text{time factor}} \sigma_x^2}_{\text{extrinsic}} \qquad (4)$$

for the noise of z, while the noise formula for the NAR system evaluates to

$$\mathcal{N} = \frac{z_H}{(1 - \gamma f_H R_H' \mu_z^{-1})(z_H - z_L)^2} + \frac{1}{(z_H - z_L)^2} \frac{(\gamma \beta f_H' R_H)^2}{[\gamma \beta f_H R_H' - \mu_z][-\mu_y + \gamma \beta f_H R_H' - \mu_z]} \sigma_y^2 \qquad (5)$$

Here we abbreviated $f(y_H)$ by f_H, $df(y_H)/dy$ by f_H' and $1/\mu_z$ by τ_z.

We always evaluate the noise at the high state of G_z, because this state has the higher variance. The normalisation factor $z_H - z_L$ is the same hence the noise at the high state will be higher, and is therefore chosen as the indicator for the noise of the system as a whole. Note that Eqs. (4) and (5) are the formulas for the activator. In the case of G_z being repressed by G_y the terms f_H and f_H' need to be replaced by f_L and f_L'. This is because the high state z_H in G_z is achieved when $y^* = y_L$.

2.2. Switching Time

In the language of dynamical systems that we use to model our system of genes, the computational speed of a system is determined by the time to switch from one state to another. In the simplest possible case of a binary system that we consider here, there are only two transitions, namely from/to a high state z_H, and to/from a low state z_L. Specifically, we interpret the time to switch as the time of the output to reach a fraction of one steady state (say z_H) given that the system starts in the other possible state (which would be z_L in this case). In general, the transition time $T_d \left(\Theta_H \right)$, from the high to the low state, and the time $T_u \left(\Theta_H \right)$, from the low to the high states will not be equal, i.e. $T_u(\Theta_H) \neq T_d(\Theta_L)$. However, the maximum switching frequency is limited by $T(\Theta)$, the slowest state transition time in the system, which is therefore a good indicator of the computational speed of a network of genes:

$$T(\Theta) = \max \left(T_u \left(\Theta_H \right), T_d \left(\Theta_L \right) \right),$$

where $\Theta_L \doteq z_H - \theta(z_H - z_L)$ and $\Theta_H \doteq z_L + \theta(z_H - z_L)$.

Here the subscripts u and d abbreviate "up" and "down," signifying whether the output signal is rising (G_z turning on) or falling G_z (turning off). The optimum configuration is reached where the time to switch the system on equals the time to switch it off, i.e. $T_u(\Theta_H) = T_d(\Theta_L)$.

2.3. Stochastic Model

We compared the noise prediction from the linear noise approximation with a Markov chain model corresponding to the deterministic system. The Markov chain model was constructed using the PRISM probabilistic model checker (Kwiatkowska et al., 2001), using the following code:

```
        y : [miny..maxy] init miny;
        z : [minz..maxz] init minz;
[]    (y <maxy) & (x>0)
-> omeg*(pow(x,h2)/(pow(x,h2) + pow(K2,h2))): (y'=y+1) ;
[]    (y> miny) -> y*m: (y'=y-1);
[]    (z <maxz) & (y>0)
-> beta*(pow(y,h)/(pow(y,h) + pow(K,h))): (z'=z+1) ;
[]    (z> minz) -> z*m: (z'=z-1);
```

To save space, we suppressed the specification of the parameters here. The noise was computed using two reward structures namely (i) `true: z`; and (ii) `true: z*z;`. The relevant property query was then given by:

```
((R{2}=? [ S ]-(R{1}=? [ S ])*(R{1}=? [ S ])))/pow(zh-zl,2)
```

PRISM can compute the exact numerical values (up to machine precision) of the noise; this approach is thus not only computationally more efficient than noise estimates based on stochastic simulations using Gillespie's algorithm, but also more precise. The drawback of this approach is that it only works for relatively small models. In the present case, this works as long as the noise from the gene G_x is considered negligible, which was the case in Fig. 1(left). In the case of the positive auto-regulator in Fig. 1(right) this was no longer the case and we had to resort to approximations of the noise by simulation. The model used in this figure had an additional stochastic variable x representing the particle number of G_x with the following transition rates:

```
        x : [minx..maxx] init minx;
[]    (y <maxx) & (x>0)-> 10: (x'=x+1) ;
[]    (x> minx) -> x*m: (x'=x-1);
```

Here we did not duplicate the items already listed above.

3. Results

In this section we report how the computational properties of the three-gene system depend on its parameters. In all cases we assume that G_y is activated by G_x and we consider the cases where G_z is activated/repressed by G_y both for the standard system and the case of G_z repressing its own expression.

3.1. Linear Chain without Feedback—The Standard System

For a linear chain of genes there are two distinct trade-offs between noise and time. Firstly, at a fixed metabolic cost the overall noise of the output increases linearly with the decay rate (see equation (5) in SI 2), while the time to switch is inversely proportional to the decay rate (see SI 4.2). This is a direct generalization of the noise-time trade-off that has been identified previously for a single gene (Zabet and Chu, 2009).

Secondly, there are optimal values of the Hill threshold K for noise and switching time. In general, these will not be equal. Hence, the values of K in between the noise-optimum and the time-optimum realize a trade-off between noise and time, representing various possible combinations of noise and switching speed. In what follows, we describe these in more detail. For a fixed metabolic cost ζ, decay rate μ and Hill threshold K_2 both the switching time and the noise have an optimal parameter value for the Hill threshold K of the gene G_z. Conversely, the linear noise approximation predicts that for a fixed K there is an optimal value K_2 that minimizes the total noise of the output (see SI 3). When G_z is an activator, then as $K \to 0$ and K_2 is kept optimal with respect to K, Eq. (4) predicts the noise to approach

$$\lim_{K \to 0} \mathcal{N} = \left(\frac{x_H^{h \cdot h_2}}{x_L^{h \cdot h_2} - x_H^{h \cdot h_2}} \right)^2 \frac{\mu_z}{\zeta}. \tag{6}$$

Hence, the linear noise approximation does not predict an optimal value for the pair (K, K_2) in the case of the activator. We expect

the linear noise approximation to become inaccurate for extreme parameter values, i.e. close to $K = 0$. An exact numerical analysis of the full Markov chain model shows that the linear noise approximation holds well for even very low K but eventually breaks down before this minimum can be reached (see Fig. 1(left)).

The repressor case is somewhat different in that noise does not diverge for $K \to 0$, but approaches a finite value (see SI). Unlike the case of the activator, there is a global minimum for the noise in terms of the Hill coefficients K and K_2, but it is not possible to find a meaningful closed form solution for this global minimum.

Numerical analyses suggest that the time to switch is strongly dependent on the relative position of K between y_H and y_L. Intuitively, this can be understood as follows: In the case of an infinite Hill coefficient h the Hill threshold K defines the crucial limit for y below which there is no activation and above which the gene is fully activated. Hence, if the Hill threshold K is very low, then low levels of y will fully activate G_z, which entails that G_z will be activated soon after G_x has been activated. Hence, low K means rapid turn-on of G_z. Analogous reasoning applies when turning the activator off. The closer the Hill parameter to the activated state y_H the earlier the predecessor gene G_y falls to levels below the activation threshold, and hence higher values of K entail a faster switch-off. In the case of $h = \infty$ we therefore find that the time-optimal value of the Hill threshold is given by $K = (y_H - y_L)/2$.

If we relax the assumption of an infinite Hill coefficient, the same qualitative reasoning still holds, but the optimum K will not exactly coincide with the mid-point between the steady states of G_y. Fig. 1 illustrates the optimal K for a number of numerical examples. The graph records a normalised difference between the time required to switch the gene off and the time required to switch it on, i.e. $(T_d(K) - T_u(K))/T_d(K)$. The optimal K coincides with the respective times being equal, which is where the curves in the graph intersect the horizontal axis in Fig. 1. The case of the repressor gene is analogous.

As stated previously (and in Zabet and Chu, 2009), for each of the individual genes G_y and G_z the computing time crucially depends on the metabolic cost of the gene (assuming a fixed level of noise). In our present system the costs allocated to the individual genes can be altered as long as the boundary condition of a fixed total cost is respected, i.e. $\zeta = \beta f_H + \omega g_H$ remains constant. One could be tempted to conjecture that the computing time will depend on how cost is allocated to the individual genes. We show in SI 5 that this is not the case: The switching time is independent of the cost allocation. This rules out an allocation mediated trade-off between noise and time.

With respect to noise, the cost allocation does matter. In the limiting cases of all the cost being allocated to either G_y or G_z the noise goes to infinity which shows that there exists at least one minimum in between these extremes. It can also be shown that
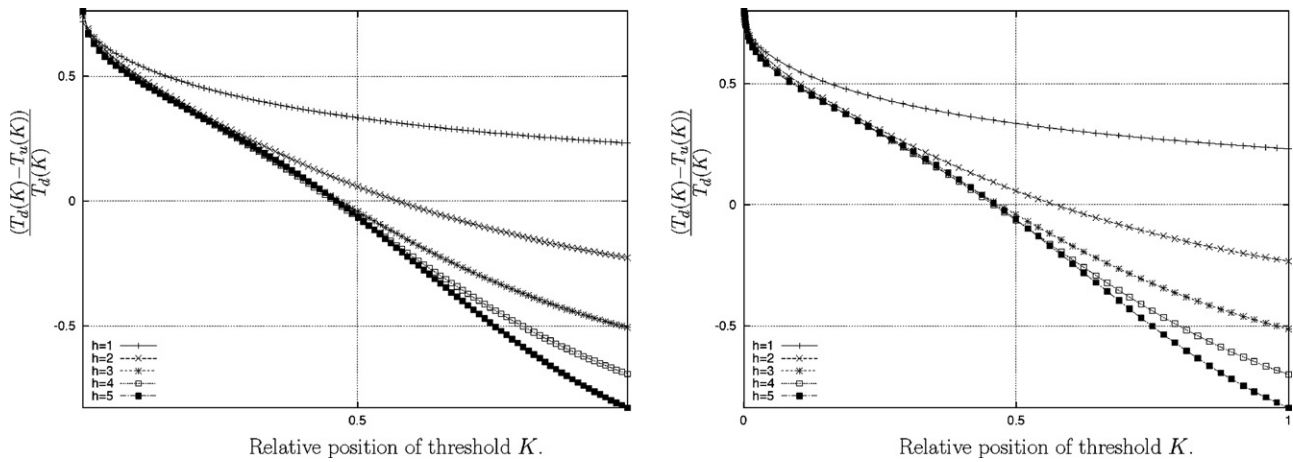
Fig. 1. Switching on and off. (left) The horizontal axis is the relative position of the threshold parameter. The vertical axis represents the difference between the time required to switch G_z on and off normalised by the time required to it switch on. The optimal parameter for K is where the line crosses $x = 0$. The various curves represent different Hill parameters h. The parameters used are as follows: $\omega = 2$, $h_2 = 2$, $x_H = 5$, $x_L = 0.1$, $\zeta = 4$. The value of the decay rate was set to $\mu = 2$ but note that each of the curves is invariant with respect to the decay rate. (right) Same as left but for the repressor.

any such minimum is independent of the decay rate μ, at least in the case of decay by dilution.

Fig. 2 illustrates the noise-optimal cost distribution (and hence the optimal ω) as a function of the relative strength of the upstream noise. It indicates a strong relationship between the amount of the upstream noise relative to the total noise and the cost allocation. Intuitively, this can be understood as follows: Allocating more resource to G_y reduces the input noise to G_z by increasing the metabolic cost at G_y, but at the expense of increasing the intrinsic noise at G_z. Depending on which noise source is the dominant one, more or less cost should be allocated to G_y. The same reasoning holds in the case of the repressor, although the input noise tends to be low because it is generated by a low particle concentration y_L.

3.2. Negative Auto-regulation—The NAR System

Negative auto-regulation has been shown to attenuate noise and increase the switching time (Wang et al., 2010; Rosenfeld et al., 2002). Indeed, irrespective of its parameters, the variance of the
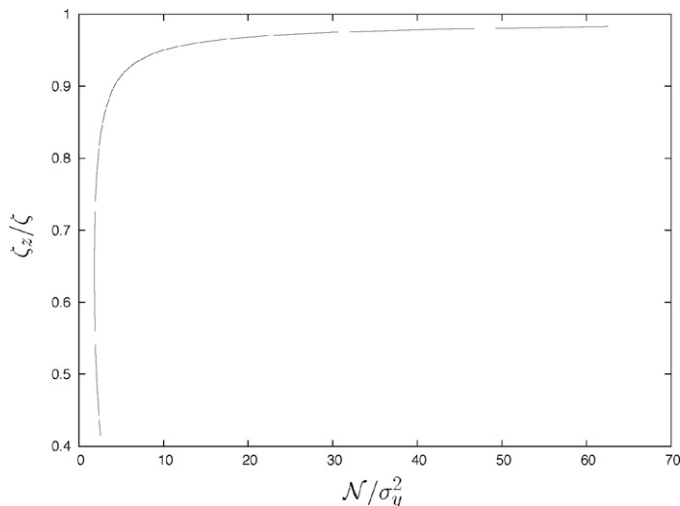

Fig. 2. Parametric plot of the optimal cost distribution (vertical) as a function of the ratio of the total noise to the input noise σ_y^2 (horizontal). The parameter varied along the curve is $K = [0 \ldots 20]$. The following parameters were used: $\omega = 2$, $h = h_2 = 2$, $x_H = 5$, $x_L = 0.1$, $\zeta = 4$, $K_2 = 0.04$. The value of the decay rate was set to $\mu = 2$ but note that each of the curves is invariant with respect to the decay rate and the cost.

NAR system is at worst equal to the noise of the corresponding standard system (see SI). Section SI 5 shows that in the NAR system the variance is lowest for $\bar{K} = 0$ and approaches the levels of the standard system for $K = \infty$. This is true for both the repressor and the activator. However, it should be noted that this does not necessarily translate into a lower noise. The signal $z_H - z_L$ in the NAR system is at best as strong as in the standard system (see SI 5). This means that there is the potential for the noise of the NAR system to be higher, even though the variance is lower. A full analytic treatment of the noise is hindered by the shape of the equation. In the case of the high state this is not a problem, because z_H is identical in the NAR system and in the corresponding standard system. However, the steady state equation for the low state z_L cannot be solved in general. We are therefore largely limited to doing numerical analysis.

Fig. 3 compares the noise of the NAR system and the standard system for some parameter values. The influence of $\bar{h}$ seems to be more pronounced in the case of the NAR repressor than for the activator. In both cases the Hill coefficient of the repression function does not seem to influence the noise qualitatively, in the sense that the range of parameters for which the standard system has better noise characteristics is roughly unchanged. For a particular set of parameters, Fig. 3 shows that the noise of the NAR activator is lower than the corresponding standard system. This contrasts with the case of the NAR repressor; the right hand side of Fig. 3 suggests that there is a significant proportion of the parameter space where the noise of the NAR repressor is higher than the corresponding standard system. Given that this is a particular example set of parameters, it is unclear to what extent this conclusion generalizes.

Figs. 4 and 5 compare the switching times of the NAR and the standard system. A somewhat different picture presents itself. While with respect to noise the NAR activator seemed to have shown more benign computational properties throughout the parameter space, when it comes to the switching time there is a sizeable area where the NAR activator is slower than the standard system. This is specifically true for lower values of the Hill threshold K, but the NAR activator is always faster once K is sufficiently increased. There is also an area for very low K where the NAR activator is faster. However, this is unlikely to be of biological significance, because this corresponds to G_z being only weakly controlled by G_y. The example also suggests that the Hill coefficient $\bar{h}$ correlates negatively with speed once it is greater than 2, although the difference does not seem to be big. Lower Hill coefficients $\bar{h}$ lead to NAR systems that are overall more similar to the standard system. This is expected and for $\bar{h} < 1$ this trend would continue. One can see from
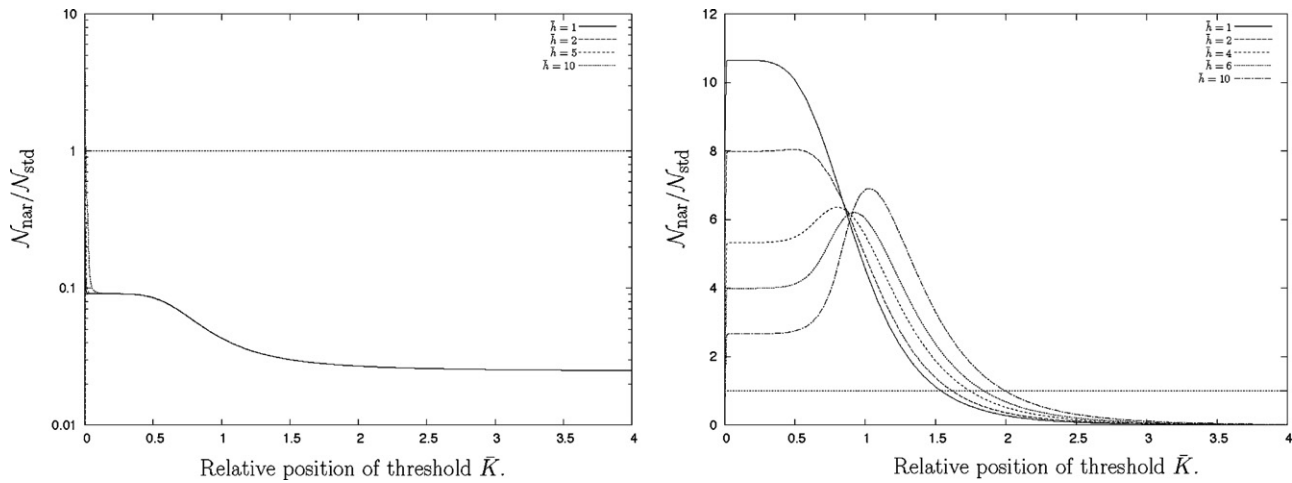
Fig. 3. The noise of the NAR system (activator and repressor). (left) Comparing the noise of the NAR system with the standard system. The graph shows the ratio of the noise of the standard system and the NAR system for different Hill thresholds of the repressor function $\bar{K}$ (x-axis) and Hill parameters $\bar{h}$. A value of 1 means that the systems have the same noise. A value <1 means that the standard system has higher noise. The parameters used are as follows: $\omega = 2$, $h = 3$, $h_2 = 2$, $x_H = 5$, $x_L = 0.1$, $\zeta = 4$, $\mu = 1$. The Hill constants K_2 and K have been kept at the midpoint between the high states of x and y, respectively. (right) Same as left but for the NAR repressor.
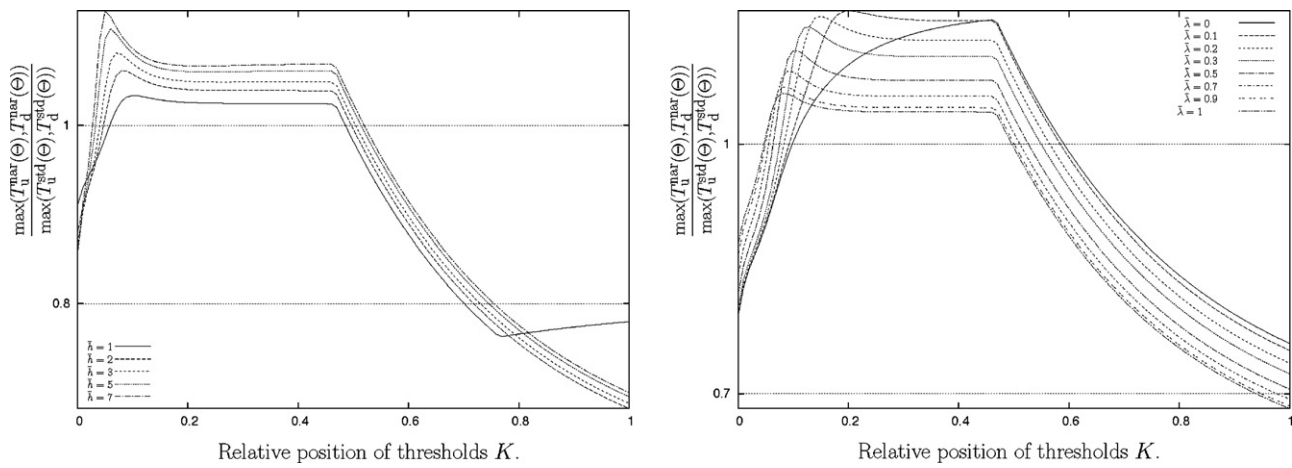

Fig. 4. Comparing the switching time of the NAR (activator) system with the standard system. (left) The graph shows the ratio of the switching times of the NAR system and the standard system for different Hill coefficients of the NAR system ($\bar{h}$). A value of 1 means that the systems have the same switching time. A value <1 means that the standard system is slower. The parameters used are as follows: $\bar{K} = z_H$, $\omega = 2$, $h = 3$, $h_2 = 2$, $x_H = 5$, $x_L = 0.1$, $\zeta = 4$, $\mu = 1$. (right) Same, but $\bar{h} = 2$. The different plots correspond to different relative positions of $\bar{K} = \bar{\lambda} z_H$.
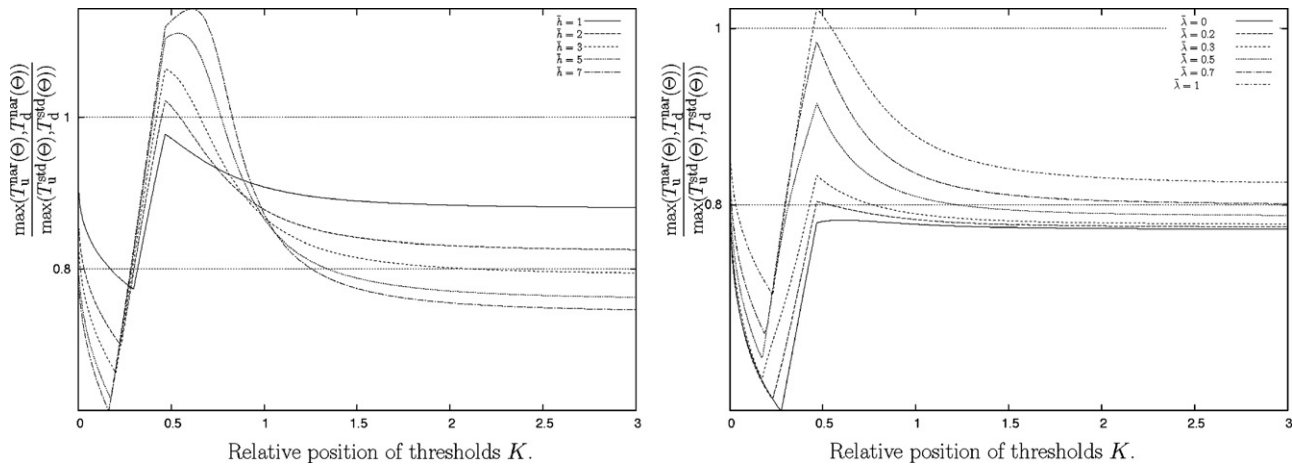

Fig. 5. Comparing the switching time of the NAR (repressor) system with the standard system. Parameters and interpretation are the same as in Fig. 4.
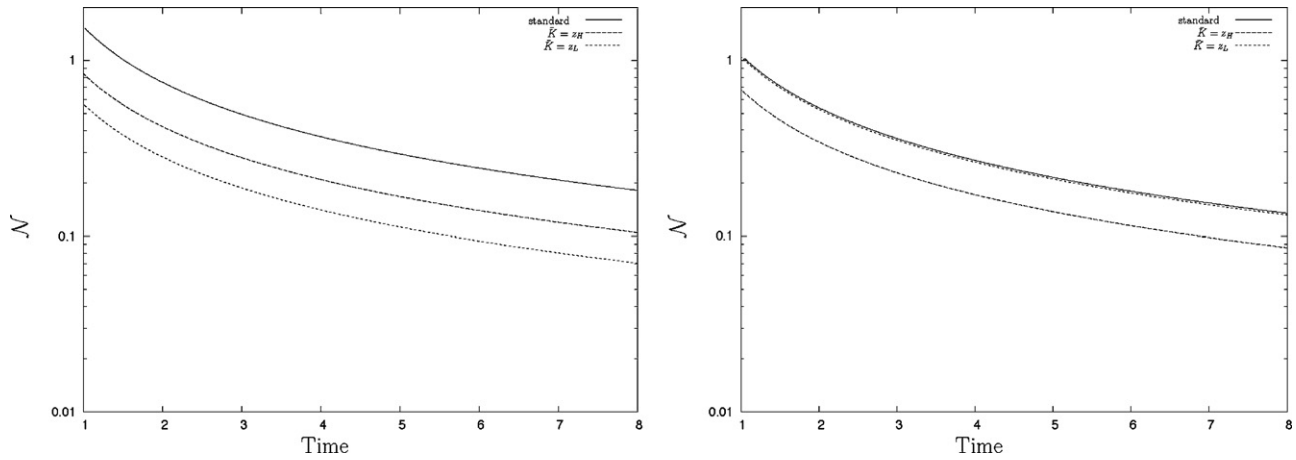
Fig. 6. Performance of the NAR system. (left) The μ-mediated noise-time trade-off of the standard system (solid line) compared to the NAR. The horizontal axis represents the switching time $T_{nar}(\Theta)$ and the vertical axis the noise $\mathcal{N}$. The parameters used are as follows: $\omega = 2$, $h = 3$, $h_2 = 2$, $\bar{h} = 2$, $x_H = 5$, $x_L = 0.1$, $\zeta = 4$, $K_2 = (1/2)x_H$, $K = 0.4y_H$. (right) Same as left but for the repressor and $K = 0.5y_H$.

the system of differential equations (3) that the regulation factor $R\gamma$ reduces to 1 for $\bar{h} = 0$ independent of $\bar{K}$; this means that in this limit the NAR system and the standard system are identical. The same is true for very large $\bar{K}$, when the contribution of $z(t)$ and z_H in $R(z)$ and γ respectively become negligible.

The example solutions shown in Fig. 5 indicate that the NAR repressor is faster than the standard system for most of the parameters, i.e. only a small proportion of each graph reaches beyond 1. It is difficult to obtain analytical results for the switching time, although in the limiting case of $\bar{K} = 0$ one can gain some insight. In this case the production rate in the differential equations is given by $J^+ = f_H z_H / z(t)$. This can be integrated to obtain a general solution corresponding to Eq. (2), namely

$$\bar{z}(T) = z_H \left(\frac{(1+\bar{h})}{z_H} \int_0^T f e^{\mu t (1+\bar{h})} dt + z_H \right)^{1/(1+\bar{h})} e^{-\mu T} \tag{7}$$

The trajectory of $\bar{z}(t)$ can be seen to be the scaled version of an un-repressed system with a decay rate $\mu'_z = \mu(1+\bar{h})$, whose solution is denoted $z'(t)$. The trajectory from the high state to the low state can then be written as follows:

$$\bar{z}(T) = z_H \left(\frac{z'(T)}{z_H} (1+\bar{h}) \right)^{1/(1+\bar{h})}. \tag{8}$$

Analogously, one can write the trajectory from the low state to the high state. The fact that the repressor is a scaled version of the un-repressed system, denoted by $z'(t)$, does not mean that $\bar{z}(t)$ and $z'(t)$ have the same switching times. The reason is that the points $\bar{\Theta}$ and Θ' where the respective systems reach a fraction θ of the distance between the high and the low state are not in direct correspondence. If we map the point $\bar{z}(T_{\bar{\Theta}}) = \bar{\Theta}$ onto the corresponding point on $z'(t)$ then we obtain

$$\bar{z}(T_{\bar{\Theta}}) = \left((1-\theta) + \frac{\bar{z}_L}{z_H} \theta \right)^{1/(1+\bar{h})}.$$

From Eq. (8) we can also calculate the low state of the system, that is

$$\bar{z}_L = z_H \left(\frac{z_L}{z_H} \right)^{1/(1+\bar{h})}. \tag{9}$$

The numerical examples so far show that the NAR system is sometimes noisier/slower than the standard system, this does not mean that the NAR system itself is "worse" than the standard system in terms of its computational properties. Fig. 6 compares two trade-off curves. The trade-off curves of the NAR activator

are clearly below the standard system, which means that across the parameters varied the NAR systems has computationally more benign properties/trade-offs, and similarly in the case of the repressor. It should be noted that some of the points in the trade-off curves correspond exactly to points in Figs. 3 and 4[1]: In the trade-off curves all points of the NAR system are below the standard system. At the same time, the activator corresponding to $K = 0.4y_H$ in Fig. 4 is clearly slower than the corresponding standard system; similarly, Fig. 3 shows that the noise of the NAR repressor is higher for $K = 0.5y_H$. Nonetheless, for both of these parameters the trade-off curve of the NAR system is below the standard system, which indicates an overall better computational performance than the corresponding standard system.

In order to get a better impression of the behavior of the system across the parameter space we performed a Monte-Carlo sampling of the parameter space. Fig. 7 shows parameter sweeps for the behaviors of the NAR repressor and activator. Sampling different ω and K further corroborates the impression that there are significant parts of parameter space where the NAR repressor is noisier than the standard system, and significant parts where the NAR activator is slower than the standard system. The interpretation of these parameter sweeps is aided by comparison with random points taken at optimal K on the rhs of Fig. 7. Two main observations emerge from this. Firstly, the NAR repressor is both faster than the standard system and has lower noise when the Hill threshold parameter is optimised. Secondly, the activator has only lower noise, but is slower than the standard system at noise-optimal K.

Fig. 8 shows a parameter sweep showing the time optimal K for random parameters for the repressor. The result shows that for most points the NAR system is faster than the corresponding standard system.

4. Discussion

In order to survive, bio-systems need to perform computations on the state of their environment and their internal states. For bacteria one way to implement such computations are gene regulatory networks. These can switch genes or sets of genes on and off depending on environmental or internal conditions or rather concentrations of molecular species that indicate such conditions.

---

[1] It is not possible to know from the graph precisely which points in the trade-off curve map to the corresponding points in Figs. 3 and 4, because the dots in Fig. 6 are not labeled with respect to the free parameter μ.
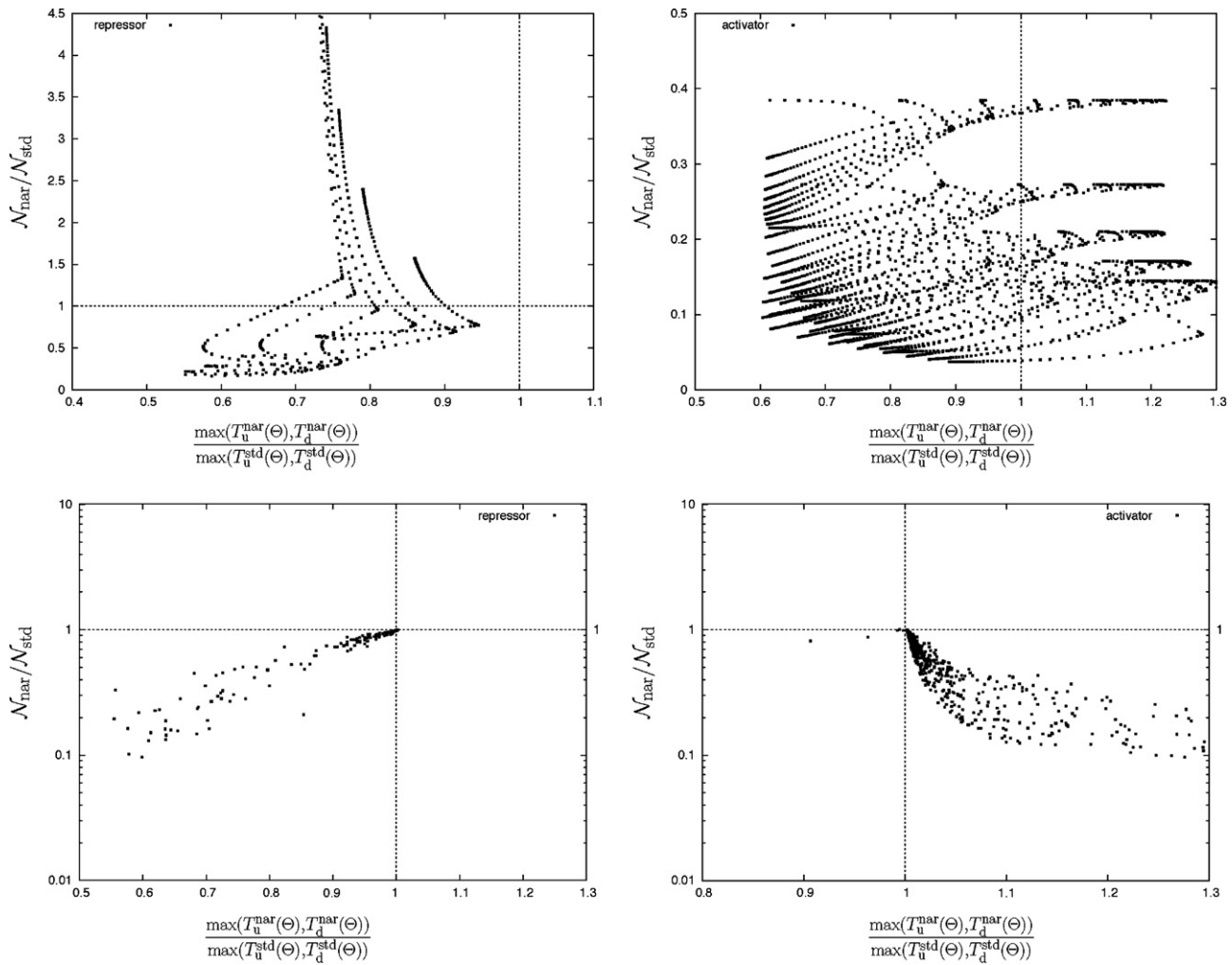
Fig. 7. Parameter sweep of the NAR system. (top) The graph shows noise-time trade-offs of the standard system compared to the NAR for both the positive regulator (lhs) and the negative regulator (rhs). This is a parametric plot of $\mathcal{N}_{nar}/\mathcal{N}_{std}$. The variables that are varied are $0 < \bar{K} < 2z_H$ and $0.1 \leq \omega \leq 4.1$ each for $\bar{h} = 1, 2, 3, 4, 5$. A value below 1 on the horizontal and vertical axis respectively means that the noise/switching time of the NAR system is lower. The parameters used are as follows: $\mu = 1, \omega = 2, h = 3, h_2 = 2, \bar{h} = 2, x_H = 5, x_L = 0.1, \zeta = 4, K_2 = (1/2)x_H$. (bottom) As top, but random parameter were chosen. For each set of parameters the optimal value of K was calculated and used to generate the point. This graph is in log-scale for better readability.
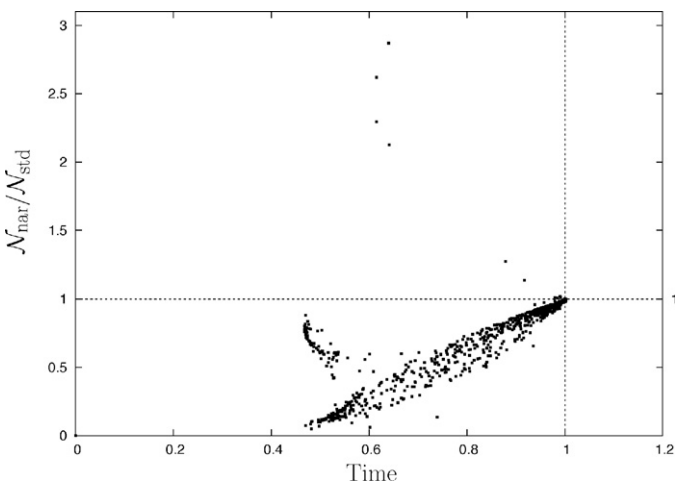

Fig. 8. Parameter sweep of the NAR repressor system. As in the bottom right graph in Fig. 7 but K is optimised for switching time.

In this article we use a model of binary genes, that is each gene can be in one of two states only, "high" or "low." This assumption significantly simplifies the mathematics, but does not limit the generality of our conclusions. In genes that have more than two states, the same ideas as the one we present here apply. However, a detailed consideration of this case has to be left to future research.

We define a gene network as computing when it relays or processes changes of one and more input concentrations and regulates genes in response to this. In this article we exclusively consider binary genes, that is genes that have only two possible activation states. In gene networks that are optimised for computing, there are two distinct trade-offs between the time to compute and the accuracy of the computation, a trade-off mediated by the decay rate μ and one mediated by the Hill thresholds K. Mathematically, the former is the more fundamental one, in the sense that it is a consequence of how the noise and the switching time scale with the decay rate, namely $(\mathcal{N} \sim \mu)$ and $T(\Theta) \sim 1/\mu$, respectively. The biological interpretation of this μ-mediated trade-off is somewhat more difficult than the mathematics suggest. In bacterial cells, there is often no active break-down mechanism for proteins and concentrations diminish mainly through cell growth and division. Consequently, the decay rate will be related to the growth rate, which is time-varying and dependent on the nutrient supply.
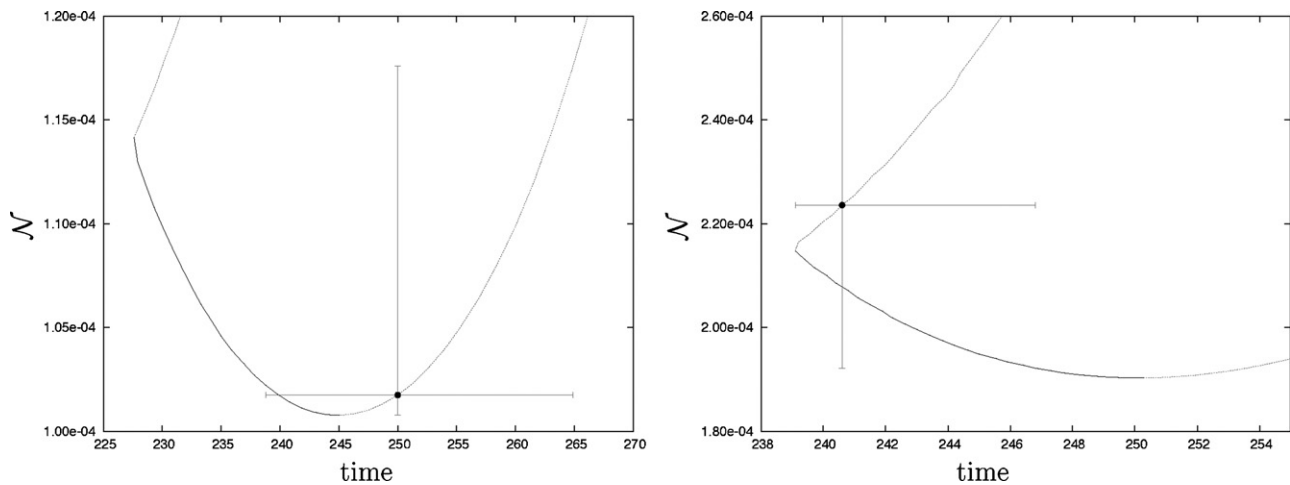
Fig. 9. Optimality analysis and biological experiments. This figure shows the K-mediated noise time trade-off for the pR promoter as reported in Rosenfeld et al. (2005). The trade-off is drawn in bold, and the sub-optimal points are in grey. We used the following set of parameters: $V = 1.5 \times 10^{-15}$ l, $\mu_x = \mu = \mu_{\text{dilution}} = \ln(2)/45 \, \text{min}^{-1}$, $\theta = 0.9$, $\alpha = 0 \, \mu\text{M min}^{-1}$, $x_H = 0.006 \, \mu\text{M}$ and $x_L = 0.140 \, \mu\text{M}$ (Rosenfeld et al., 2005). The error bars were produced assuming an error of 20% for $K \pm 20\%$. (left) Wild type as reported in Rosenfeld et al. (2005). (right) The mutant.

The growth rate and the availability of nutrients are only to a limited extent free parameters of the cell. The μ-mediated noise-time trade-off is therefore best thought of as a physical constraint on the cell rather than the result of an evolutionary strategy.

In the case of active particle breakdown the interpretation is not much clearer. The issue is complicated by the fact that active breakdown requires at least one additional species of particles to break down or inactivate the primary regulator. This comes itself at an additional metabolic cost and thus complicates the analysis. Moreover, the presence of the breakdown particle does not *per se* increase the switching speed because these breakdown particles need to be broken down themselves before a new signal can be built up. This may be circumvented by transporting particles to dedicated breakdown sites as it occurs in eukaryotic cells. Analyzing this case in terms of optimality considerations would go well beyond the scope of the present contribution.

Biologically, the K mediated trade-off is more rewarding to analyze, at least in the context of prokaryotic gene regulation. The Hill threshold K is related to the the binding and the un-binding rate constants of the transcription factor, and as such tunable over evolutionary time-scales. What makes the K-mediated trade-off interesting is that it provides insight into the adaptive pressures that shaped the parameters of the cell. When the parameters are known, then this can be used to understand what precisely the network is an adaptation to. When parameters are *not* known, then optimality considerations can be used to constrain the search space during parameter inference (e.g. during model fitting or via priors in Bayesian parameter estimation).

At present a problem are the high error rates which affect empirical estimates of kinetic parameters in gene regulatory networks. Fig. 9 shows a K-mediated trade-off curve for the *pR* promoter using parameters taken from Rosenfeld et al. (2005). They report errors for the value of the Hill threshold of about 20% and similarly high values for other parameters. Using the reported mean values, would locate the promoter just off the noise optimum; a mutant they reported in the same contribution would be just off the time optimum. Fig. 9 also shows error bars around the points. We assumed that only the Hill threshold is affected by an error and that all the other parameters are certain. The figure shows that even in this overly optimistic scenario it is not possible to meaningfully locate individual genes on the trade-off curve. Note that a proper treatment of errors would also have to note an uncertainty around the

trade-off curve itself. Future more accurate information about the parameters will make this type of analysis more rewarding.

In order to analyze the noise-time trade-offs we used the van Kampen–Paulsson linear noise approximation. A comparison with exact calculations of the same Markov chain shows that the predictions of this approximation are very accurate for most parameters. However, there is still some caution warranted when interpreting the results with respect to real organisms. One source of error are the simplifying assumptions we made in order to arrive at the dynamical model of the gene networks (i.e. Eqs. (1) and (3)).

One of these assumptions concerns the input (i.e. the product of the gene G_x) which we assumed to be flipping instantaneously between the high state x_H and the low state x_L, a condition that is clearly never fulfilled in nature. However, we have mainly concentrated on qualitative aspects of the dynamics, that is how observable features of the gene network depend on the parameters of the system and we are not trying to present a predictive model of a specific system. Relaxing the assumption of instantaneous input will not affect the qualitative relations we have established; the assumption is therefore of little consequence to our conclusions.

Another aspect that is clearly not taken into account in our model is the time delay due to the translation step which, amongst other things, will depend on the length of the protein. The length of the protein will also enter into the cost function thus complicating the analysis. For our current purposes these are second order effects that will be important in quantitative models of specific systems, but are of minor relevance for the conceptual understanding we are trying to develop here.

Of more direct impact for our models is the breakdown of the linear noise approximation for extreme parameters. The FDT approach predicts that there is a theoretical minimal value for the noise given by Eq. (6); this can be reached in the limit of a vanishing K (while keeping K_2 optimal). Fig. SI 1(left) indicates that the theoretical estimates for the noise become inaccurate for very low values of K which means that this global optimum will probably not be reproduced by real systems even though (at least for the parameters in the graph) the linear noise approximation is a good indicator for a wide range of parameters (see Fig. SI 2 and SI 1(left)). Add to this that the model itself is a simplification and we must conclude that the prediction of Eq. (6) may be far from biological reality in quantitative terms. On the other hand, the linear noise approximation

has successfully been applied to real biological systems and it is therefore reasonable to assume that the overall qualitative picture it paints is relevant.

In the literature, there is a general consensus developing that self-repressing genes have more benign noise and time (Rosenfeld et al., 2002; Wang et al., 2010) properties than a standard gene with the same parameters. Indeed, in bacteria estimates for the proportion of NAR genes among all genes range from 40 to 60% (Rosenfeld et al., 2002; Alon, 2007). However, there is no obvious metabolic cost in negative auto-regulation, which begs the question why not all genes are auto-repressors. Part of the answer is that the idea of gene networks as computing input is not always true. Sometimes genes fulfill functions other than information processing. In those the main adaptive pressures may not come from noise and/or switching speed, but from other considerations. There may be systems that need to switch fast in one direction only. In this case the trade-off curves would have to be drawn very differently, thus altering the conclusions reached here. Finally, in our analysis we did not consider motifs other than the standard system and negative auto-regulation. For example, in some circumstances Feed-Forward Loop network motifs (Mangan and Alon, 2003) may be more suitable than negative auto-regulation.

The linear noise approximation predicts that the variance is always lower in the NAR system than in the standard system. This entails that the noise, if taken against a zero-concentration baseline is lower. However, if the baseline signal corresponds to a non-vanishing concentration, then the noise of the NAR system may be higher than in the standard system. In the case of the NAR repressor we even found this to be the typical behavior for the range of parameters we considered. On the other hand, our Monte-Carlo parameter sweeps indicate that the NAR-activator is typically less noisy than the corresponding standard system, although there are significant parts of parameter space where the auto-regulated system is slower than the standard system.

Biologically the typical behavior of a system may not be as relevant as the optimal behavior to which adaptive pressures drove the system. This is particularly true in the present case where the typical behavior of the system is very different from the optimal behavior: Both at its noise-optimal K and its time-optimal K the repressor is faster and less noisy than the standard system (see Figs. 7 and 8). In the case of the activator the case is not as clear cut. For all but a few example parameters we generated, at noise-optimal K the NAR system has lower noise than the standard system, but the switching time is higher.

From our analysis it is clear that the computational efficiency of gene networks cannot be indicated by a single variable only, but must be assessed across the parameter space using trade-off curves. While we found that sometimes the negative regulator has indeed higher noise than the corresponding standard system, we could not find a single example where the NAR system did not lead to an overall better noise-time trade-off than the corresponding standard system. This leads us to the conclusion that the NAR system has a better computational performance than the standard system. More generally, however, this suggests that trade-off curves can be very useful tool when exploring optimal parameter ranges for bio-systems.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.biosystems.2011.01.006.

References

Alon, U., 2006. An Introduction to Systems Biology: Design Principles of Biological Circuits. Chapman & Hall.

Alon, U., 2007. Network motifs: theory and experimental approaches. Nature Review Genetics 8 (June (6)), 450–461, http://dx.doi.org/10.1038/nrg2102.

Bar-Even, A., Paulsson, J., Maheshri, N., Carmi, M., O'Shea, E., Pilpel, Y., Barkai, N., 2006. Noise in protein expression scales with natural protein abundance. Nature Genetics 38 (June (6)), 636–643, http://dx.doi.org/10.1038/ng1807.

Ben-Hur, A., Siegelmann, H., 2004. Computation in gene networks. Chaos 14 (March (1)), 145–151, http://dx.doi.org/10.1063/1.1633371.

Bray, D., 1995. Protein molecules as computational elements in living cells. Nature 376, 307–313.

Bruggeman, F., Bluethgen, N., Westerhoff, H., 2009. Noise management by molecular networks. PLoS Computational Biology 5 (September (9)), e1000506, http://dx.doi.org/10.1371/journal.pcbi.1000506.

Chu, D., Roobol, J., Blomfield, I., 2008. A theoretical interpretation of the transient sialic acid toxicity of a *nanR* mutant of *Escherichia coli*. Journal of Molecular Biology 375, 875–889.

Fernando, C., Liekens, A., Bingle, L., Beck, C., Lenser, T., Stekel, D., Rowe, J., 2009. Molecular circuits for associative learning in single-celled organisms. Journal of the Royal Society Interface 6 (May (34)), 463–469, http://dx.doi.org/10.1098/rsif.2008.0344.

Haynes, K., Broderick, M., Brown, A., Butner, T., Dickson, J., Harde, W., Heard, L., Jessen, E., Malloy, K., Ogden, B., Rosemond, S., Simpson, S., Zwack, E., Campbell, A., Eckdahl, T., Heyer, L., Poet, J., 2008. Engineering bacteria to solve the burnt pancake problem. Journal of Biological Engineering 2 (1), 8, http://www.jbioleng.org/content/2/1/8.

Kwiatkowska, M., Norman, G., Parker, D., 2001. PRISM: Probabilistic symbolic model checker. In: Kemper, P. (Ed.), Proc. Tools Session of Aachen 2001 International Multiconference on Measurement, Modelling and Evaluation of Computer-Communication Systems, pp. 7–12, available as Technical Report 760/2001, University of Dortmund, September.

Ma, W., Trusina, A., El-Samad, H., Lim, W.A., Tang, C., 2009. Defining network topologies that can achieve biochemical adaptation. Cell 138 (4), 760–773.

Mangan, S., Alon, U., 2003. Structure and function of the feed-forward loop network motif. Proceedings of the National Academy of Sciences of the United States of America 100 (October (21)), 11980–11985, http://dx.doi.org/10.1073/pnas.2133841100.

Mattick, J., Gagen, M., 2001. The evolution of controlled multitasked gene networks: the role of introns and other noncoding rnas in the development of complex organisms. Molecular Biology and Evolution 18 (September (9)), 1611–1630.

McAdams, H., Arkin, A., 1997. Stochastic mechanisms in gene expression. Proceedings of the National Academy of Sciences of the United States of America 94 (February (3)), 814–819.

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, S., Alon, U., 2002a. Network motifs in the transcriptional regulation network of *Escherichia coli*. Nature Genetics 31, 64–68.

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, S., Alon, U., 2002b. Network motifs: simple building blocks of complex networks. Science 298, 824–827.

Narang, A., 2006. Comparative analysis of some models of gene regulation in mixed-substrate microbial growth. Journal of Theoretical Biology 242 (September (2)), 489–501.

Narang, A., Pilyugin, S., 2007. Bacterial gene regulation in diauxic and non-diauxic growth. Journal of Theoretical Biology 244 (January (2)), 326–348, http://dx.doi.org/10.1016/j.jtbi.2006.08.007.

Oltvai, Z., Barabasi, A., 2002. Life's complexity pyramid. Science 298, 763–764.

Ozbudak, E., Thattai, M., Kurtser, I., Grossman, A., van Oudenaarden, A., 2002. Regulation of noise in the expression of a single gene. Nature Genetics 31 (May (1)), 69–73, http://dx.doi.org/10.1038/ng869.

Paulsson, J., 2004. Summing up the noise in gene networks. Nature 427 (January (6973)), 415–418, http://dx.doi.org/10.1038/nature02257.

Raj, A., van Oudenaarden, A., 2008. Nature, nurture, or chance: stochastic gene expression and its consequences. Cell 135 (October (2)), 216–226, http://dx.doi.org/10.1016/j.cell.2008.09.050.

Rosenfeld, N., Elowitz, M., Alon, U., 2002. Negative autoregulation speeds the response times of transcription networks. Journal of Molecular Biology 323 (November (5)), 785–793.

Rosenfeld, N., Young, J., Alon, U., Swain, P., Elowitz, M., 2005. Gene regulation at the single-cell level. Science 307 (March (5717)), 1962–1965, http://dx.doi.org/10.1126/science.1106914.

Singh, A., Hespanha, J., 2009. Optimal feedback strength for noise suppression in auto-regulatory gene networks. Biophysical Journal 96 (May (10)), 4013–4023.

van Kampen, N., 2007. Stochastic Processes in Physics and Chemistry, third edition. Elsevier, Amsterdam.

Wang, L., Xin, J., Nie, Q., 2010. A critical quantity for noise attenuation in feedback systems. PLoS Computational Biology 6 (4), e1000764, http://dx.doi.org/10.1371/journal.pcbi.1000764.

Zabet, N., Chu, D., 2009. Computational limits to binary genes. Journal of the Royal Society Interface, 10, http://www.cs.kent.ac.uk/pubs/2009/2969, December.

Zabet, N., Hone, A., Chu, D., 2010. Design principles of transcriptional logic circuits. In: Artificial Life XII Proceedings of the Twelfth International Conference on the Synthesis and Simulation of Living Systems,. MIT Press, August, pp. 186–193, http://www.cs.kent.ac.uk/pubs/2010/3036.

Ziv, E., Nemenman, I., Wiggins, C., 2007. Optimal signal processing in small stochastic biochemical networks. PLoS One 2 (10), e1077, http://dx.doi.org/10.1371/journal.pone.0001077.