# International Journal of Biotechnology and Bioengineering

**BioCore** Exploring Scientific Community

**Research Article**                                              **Open Access**

# Noise Tolerance Analysis for Reliable Analog and Digital Computation in Living Cells

Luna Rizik, Yaron Ram, Ramez Danial[*]

Biomedical Engineering Department, Technion – Israel Institute of Technology, Haifa, Israel

**Corresponding Author:** Ramez Danial, Biomedical Engineering Department, Technion – Israel Institute of Technology, Haifa, Israel. E-mail: ramizda@bm.technion.ac.il

**Abstract:**

Biomolecular computing, encompassing computations performed by molecules, proteins and DNA, is a central area of focus in Synthetic Biology research and development, which attempt to apply engineering design principles in living cells. Two major computation paradigms have been implemented so far in living cells - analog paradigm that computes with a continuous set of numbers and digital paradigm that computes with two-discreet set of numbers. Here, we analyze the biophysical and technological limits of large-scale gene networks created based on analog and digital computation in living cells. More specifically, we calculate the precision of analog systems and the noise margin of digital systems in living cells. We conclude that both systems are challenging to operate with low protein levels. To overcome this challenge, we show that analog systems should operate with a Hill coefficient smaller than 1 and digital systems should be buffered. Furthermore, an analytical description of a biophysical model recently developed for positive feedback linearization circuits and used in analog synthetic biology, is presented. Finally, we suggest new directions for engineering biological circuits capable of computation.

## Introduction:

Computation has become an integral part of our evolution and marks a significant landmark in modern technological revolutions. The first **abacus** "calculator" was invented before 2000 BC and was based on counting continuous numbers, a process known today as an analog computation. However, scaling the complexity of computation was only truly achieved in the last century, when the digital transistor that counts discrete values, was invented. Computation based on digital design is relatively straightforward, with clear ON and OFF states that can and provide reliable results and form the basis for screening. Furthermore, digital circuits, with tightly controlled physical parameters, can be simply assembled to form complex networks, with very low cross-talk between components. The evolution of digital computation mainly relies on shrinking the transistor dimensions, which have almost reached the fundamental physical limits of scaling laws, breaking Moore's law. In contrast to digital design, analog design computes with a continuous set of numbers, with each wire carrying many bits of information. In addition, it uses the powerful laws of physics, that are naturally embodied in analog transistors, to execute sophisticatedcomputational functions (e.g. addition, subtraction, multiplication, division, logarithms and power laws). The evolution of analog computation mainly relies on feedback loops to improve precision, attenuate noise and expand the working dynamic range [1].

The last decade witnessed major breakthroughs in biophysics and genomic technologies. Researchers have successfully applied biophysical models, by combining several genes to create basic biological networks with predictable behaviors in living cells [2,3,4]. At the same time, thanks to nanotechnology and biotechnology, significant advancements in genome DNA

engineering and assembly techniques have been achieved [5]. An outcome of these advancements is an extraordinary set of design rules and engineering tools that enable massive reprogramming of the DNA code in living organisms, including humans. This new technology, known as "synthetic biology" [6,7,8], attempts to translate engineering design principles to rational biological design [9,10], to achieve multi-signal integration and processing in living cells for diagnostic, therapeutic and biotechnological applications [11,12,13,14]. For example, living cells can be programmed to produce pharmaceutical compounds that are extremely challenging to synthesize using existing methods [15], microbiome bacteria can be programmed to detect and respond to changes in clinical homeostatis [16], and gene circuits can be engineered to identify and eliminate cancer cells [17]. These developments constitute a milestone that marks the beginning of new biomolecular computing technologies, based on nanoscale-level gene-circuits in living cells, that set an alternative limit to Moore's law.

Early efforts at biomolecular computing have used binding and unbinding reactions to represent the "ON/OFF" or "1/0" logic states. Consequently, proteins that bind to DNA or promoters and activate high levels of gene expression, represent the "1" logic state, while unbound, free proteins yield low levels of gene expression, and represent the "0" logic state. Many genetic circuits that mimic electronic digital circuits, have been constructed to perform Boolean logic gates [18,19,20], counter [21] and memory [22] devices in living cells. However, because signals in living cells are graded in their nature [23,26] and do not generally exist in only two possible states, digital paradigms are often an oversimplified means of describing signals in living cells. Thus, such representation can lead to errors in construction and implementation of genetic circuits and challenge gene-network scaling in living cells [24,25].

To date, engineered artificial logic gates in living cells have been proven difficult to scale due to cellular resource limitations, a lack of orthogonal genetic devices, high leakage levels of synthetic genetic devices and the absence of suitably sharp input-to-output transfer functions [24,25]. Recently, genetic circuits have been constructed based on analog design [26]. Such gene circuits take advantage of the complex operations already naturally present in living cells, to execute sophisticated computational functions. For example, analog genetic circuits exploit positive feedback loops to implement logarithmically linear sensing, addition, division [26] and negative feedback loops while performing square-root calculations to determine chemical concentrations [26]. Analog genetic circuits involve fewer components and resources, and execute more complex operations than their digital counterparts [23,26,27]. For example, an analog adder can be achieved by simply combining two parallel circuits, where each accepts different input molecules and produces common output molecules [26]. This lies in sharp contrast to digital adders, which sum two "1" binary numbers, and require another stage to hold the new bit "Carry out" ("10"). For instance, a 4-bit digital adder may require more than 30 synthetic parts to operate, and at the same

time, would place a substantial metabolic burden on a cell [23]. By analogy to electronics, noise in biological systems [28,29] can set the physical and technological limits of engineered analog-design large-scale gene networks based in living cells. For an in-depth analysis of the pros and cons of analog versus digital computation in living cells and electronics, readers are referred to excellent reviews on the subject [1, 27].

In the present article, we analyze the biophysical and technological limits of large-scale gene networks created based on analog and digital computation in living cells. The working dynamic range, noise margin, basal (leakage) level of biological parts, sharpness of input-to-output transfer functions and copy number of synthesized proteins/molecules are assessed. In the second part of this paper, we analyze analog computation in living cells. We close the work with suggestions for future directions for engineering computation functions in living cells.

## Accuracy of analog systems in living cells:

Figure 1a shows two computational elements in living cells; in the first one, the biochemical reaction occurs at the protein-DNA level. It includes an input protein signal (x) that binds to a promoter and activates transcriptional and translational processes to synthesize an output protein signal (z). In the second element, the biochemical reaction occurs at the chemical/protein-protein level. Both bio-computing elements can be described by a Michaelis–Menten enzyme–substrate binding reaction via a Hill function, given by:

$$z - z_0 = z_{ma} \frac{(x/K_d)^n}{1 + (x/K_d)^n} \qquad (1)$$

where, $K_d$ is a dissociation constant of a biochemical reaction ($K_d = K_-/K_I$), $z_0$ is the basal level of binding, $z_{max}$ is the maximum protein concentration achieved by the system, and $n$ is the Hill coefficient, describing cooperativity. Figures 1b and 1c describe the input-to-output transfer function of Equation 1, which includes two regions: an analog continuous mode and a digital mode. In the analog mode, the function can be described by a log-linear transduction ( $vs$ $y = \log(x/K_d)$, while in the digital mode, it can be viewed as two discreet values ("0" and "1"). Equation 1 can be approximated at $x = K_d$ or $(y=0)$, using Taylor series, as:

$$z - z_0 = \frac{z_{max}}{2}(1 + \frac{n}{2}l(x/K_d)) \qquad (2)$$

Log-linear transduction, known as Weber's Law, is widely used in natural systems, such as audition, vision and cells [30], and offers advantages over linear-linear transduction. For example, small changes in the output of log-linear systems are proportional to small changes in the input signals divided by their intensity ($\Delta z \propto \Delta x/x$), demonstrating a memory element in the system. In contrast linear-linear systems show proportionality between small changes in the output and small changes in the input signals only $(\Delta z \propto \Delta x)$. The input dynamic range ($IDR$) in an analog mode is defined as (Figure 1b):

$$IDR \equiv y_{max} - y_{min} = (x_{max}/x_{min}) \qquad (3)$$

where $z(x=x_{max})-z_0 = 0.8\alpha$ and $z(x=x_{min})-z_0 = 0.2\alpha$ . Under there

definitions, the error between the log-linear analog function (Eq. 2) and a Hill function (Eq. 1) at the limits of the **IDR,** is less than 5%. By substituting **xmax and xmin** in equation 3, **IDR** is then given by:

$$IDR \approx 1.2/n \tag{4}$$

Equation 4 shows that by decreasing the Hill coefficient or the sharpness of the input-output transfer function of the binding reaction, one can increase the log-linear range. In natural biological systems, the Hill coefficient typically ranges between 1 and 4 [28], and then the **IDR** varies between 1 to 0.25 orders of magnitude. Recently, *Danial et al.* showed that by implementing a graded positive feedback loop in synthetic biological systems, one can increase the **IDR** by 4 orders of magnitudes [26].

Signals often originate from the transport of discrete random carriers in systems; in electronics, it is a drift/diffusion of electrons [1], in physics, it is the movement of photons and in biology, it is the diffusion of biochemical molecules and proteins [29,31]. Naturally, these signals propagate through networks with random fluctuations, which can be described by a Poisson process, generating shot noise that scales as the square-root of the molecular count [29]. Here, we analyze the design rules, determined by laws of cellular noise, which set the performance limits of analog and digital biological systems. Typically, there are two orthogonal sources of noise in any biological system [31,32]. The first source is the intrinsic noise, generated by the system itself, and the second source is the extrinsic noise, generated by random fluctuations in the input or another environmental parameter. A stochastic model for cellular intrinsic noise may be greater than Poisson process, with addition of burst size (**bint**) is given by [29]:

$$\sigma_{z,t} = \sqrt{(1+b_{int}) \cdot \bar{z}} \tag{5}$$

The burst size in a gene expression model is the average number of proteins synthesized ($\bar{z}$) per mRNA transcript. In a simple enzyme–substrate binding reaction, the cellular intrinsic noise is given by a Poisson process only. For simplicity, we assume that the system is operated at x= **Kd** and then, if **z=zmax/2** is substituted in equation 5, we get:

$$\sigma_{z,t} = \sqrt{(1+b_{int}) \cdot z_{max}2} \tag{6}$$

The gain of an analog system in a log-linear mode, amplifies random fluctuations in the input signal (Figure 1d). Then, the contribution of extrinsic noise (**σy**) on the output signal at **x=Kd** in a log-linear mode is expressed by:

$$\sigma_{z,t} = dz/dlog(x/Kd)|_{x=Kd} \cdot \sigma_y \tag{7}$$

The input y, is described by a log-linear function with input x, and therefore, the noise of y at **x=Kd** is a function of the noise of x (**σx**), and is given by:

$$\sigma_y = \sigma_x/x|_{x=Kd} \tag{8}$$

The input x, is a number of proteins or chemical molecules and thus, its noise (**σx**) can be described by a Poisson process, with addition of burst size (**bext**) ($\sigma_x = \sqrt{(1+b_{ext}) \cdot Kd}$). By substituting the last term of **σx** in Equation 8, we get:

$$\sigma_y = \sqrt{1+b_{ext}/Kd} \tag{9}$$

Equations 8 and 9 reveal that the noise in log-linear systems scales as the inverse of the square root of the molecular count, in contrast to linear-linear systems, where the noise scales as the square-root of the molecular count ($\sigma_x \propto \sqrt{x}$). It is simple to show that the gain of a log-linear system at *y=0* (or *x=Kd*) is equal to:

$$gain = dz/d(x/Kd)|_{x=Kd} = n \cdot z_{max}/4 \tag{10}$$

By substituting Equations 9 and 10 into Equation 7, we find that the contribution of the extrinsic input noise on the output signal is:

$$\sigma_{z,t} = n \cdot z_{max}/4 \cdot \sqrt{1+b_{ext}/Kd} \tag{11}$$

Because the intrinsic and extrinsic noise orthogonally contribute to the total noise of the system (**σz**) [29], the total noise can be given by:

$$\sigma_z = \sqrt{\sigma_{z,t}^2 + \sigma_{z,ext}^2} \tag{12}$$

If we substitute the values of intrinsic and extrinsic noise, i.e., Equations 6 and 11, respectively, into the last formula, we find that the total noise in the output of the analog signal in biochemical reactions is given by:

$$\sigma_z = \sqrt{z_{max}/2 \cdot [(1+b_{int}) + n^2 \cdot z_{max}/8Kd \cdot (1+b_{ext})]} \tag{13}$$

Any small change in the input **(Δy)**, within the **IDR** range, is amplified by the gain of the system and yields a change in the output (**Δz=gain· Δy**, Figure 1d, 1e). Biological systems have a log-linear transduction and therefore, the change of the output (**Δz**) as a response to change in the input (**Δy=Δx/x**) at **x=Kd**, is given by:

$$\Delta z = n z_{max}/4 \cdot \Delta y \tag{14}$$

For improved performance of analog systems, we require that changes in output are larger than the total noise of the system (**Δz>σz**) (Figure 1d). Thus, the minimum change in the input (**Δymin**) given by:

$$\Delta y_{min} = \sqrt{8 \cdot (1+b_{int})/n^2 \cdot z_{max} + (1+b_{ext})/Kd} \tag{15}$$

Equation 15 suggests that increasing the Hill coefficient (**n**), or the sharpness of input-to-output transfer function, of analog biological systems improves their performance. However, as we have shown in Equation 4, the **IDR** is reduced for high values of **n**, thereby affecting system performance (e.g., for a high value of **n**, the **IDR** can be smaller than the minimum change in the input (**Δymin**), thereby reducing the system's performance). Therefore, we define the precision of an analog system, which is equivalent to the signal-to-noise ratio, as the number of levels that the system can distinguish in the presence of noise. This can be calculated as the ratio of **IDR** (Equation 4) and minimum changes in the input (Equation 15):

$$\#N_{level} \equiv IDR/\Delta y_{min} \approx 1/\sqrt{8(1+b_{int})/z_{max} + (1+b_{ext})/Kd \cdot n^2} \tag{16}$$

Equation 16 represents the precision of analog systems in a log-linear mode, when we consider the contribution of extrinsic/intrinsic noise and the input dynamic range. The equation suggests that for high molecular counts or protein copies, the precision of the system will be enhanced. It also shows the contribution of extrinsic noise, which depends on a Hill coefficient, and the contribution of intrinsic noise, which is independent of a Hill coefficient. We now analyze Equation 16 under two different conditions, in accordance with the proposed basic bio-computing elements in living cells (Figure 1a). In a chemical/protein-protein reaction, the dissociation constant is often larger than the maximum protein copy number (**zmax<<Kd**), and therefore, Equation 16 can be approximated as:

$$\#N_{levl} \approx \sqrt{z_{max}/8}(1+b_{int}) \qquad (17)$$

In this case, the intrinsic noise can be viewed as the fluctuations in chemical/protein-protein binding and protein synthesis. While Equation 17 is only an approximation, it describes the precision of analog systems when the extrinsic noise is small. Under these conditions, the precision of the system is set only by the maximum protein copy number achieved by the system and by the intrinsic noise (Figure 2a), independent of **IDR**.

Analog systems can be alternative to their digital counterparts (1 bit of output precision) when operating with 4 to 8 levels of information (equivalent to 2-3 bits of output precision), which, based on our analysis (Figure 2a), can be achieved with 1000 proteins copies or molecular counts. In *Escherichia coli*, 1000 molecule counts is equal to a concentration of *1μM*, which is typically the levels of signaling proteins [33] (e.g., it was found that there are roughly 100 copies of EnvZ per cell and around 3500 copies of OmpR).

Protein-DNA biochemical reactions that involve transcription and translation processes, often operate with low protein copy numbers [33]. For such systems with both intrinsic and extrinsic noise sources, precision in a log-linear mode is described by Equation 16. For simplicity, we rearranged Equation 16 and assumed that the number of input and output protein copies are equal (**Kd=Zmax/2**) and the burst size for intrinsic and extrinsic noise is also equal (**bint=bext=b**):

$$\#N_{level} \approx \sqrt{z_{max}/8}(1+b)\cdot(1+n^2) \qquad (18)$$

The burst size relies on the translation rate, number of amino acids (aa) in the synthesized protein and on mRNA half time. Typically, in *Escherichia coli*, the translation rate ranges between 10-20 aa/sec, depending on growth conditions [33], and mRNA half time is around 3-5 min [33]. Therefore, the burst size in *Escherichia coli*, can range between 3-15. Figure 2b shows that, to achieve proper performance of analog systems based on protein-DNA biochemical reactions with 4-8 levels of information (2-3 bits of precision), the effective Hill coefficient should be smaller than one. The measured Hill coefficient in natural biological system is often higher than one, therefore, there are challenges in creating analog genetic circuits.

## Analog computation in living cells:

The first step toward implementation of synthetic analog computation in living cells, is to broaden the input dynamic range of genetic synthetic parts. Protein-DNA interactions typically have a narrow dynamic range, spanning 0.5 - 1 orders of magnitude. The input dynamic range of genetic parts is set by the cooperative binding of proteins to DNA and is often positive, with a Hill coefficient larger than one. This would mean that once one protein is bound to a DNA binding site, its affinity for other proteins increases. By contrast, a negative cooperative binding reaction has a Hill coefficient smaller than 1. *Dainal et al.* [26] implemented a positive feedback loop and decoy binding sites to shunt the proteins away from their target binding site, and achieved a Hill coefficient smaller than 1, with a very wide input dynamic range. Comprehensive biophysical and biochemical reaction models that fit their experimental results were presented [26]. In this article, we show a new analytical model that can explain the contribution of a shunt on an open loop and positive feedback loop. Figure 3a describes a transcription factor **x** (TF) that binds to **m** identical promoters. The **m-1** binding reactions act as a decoy or shunt pathway for the transcription factors. For simplicity, we assume that the Hill coefficients for all the promoters are equal to 1. The biochemical reaction model of this system is presented in Figure 3b and its solution in steady state is given by:

$$\{Pr_{b1}=x\cdot Prf_1 Kd \quad Prf_1=Pr-Pr_{b1} \quad x=x_T-m\cdot Pr_{b1} \qquad (19)$$

$$\rightarrow Prf_1 Pr = (x_T-m\cdot Prf_1)Kd / 1 + (x_T-m\cdot Prf_1)Kd/$$

where **Pr** is the total number of target promoters**, Prf** is the number of free target promoters, **Prb** is the number of target promoters occupied by transcription factors, **xT** is the total number of transcription factors and **Kd** is the dissociation constant of the binding reaction. Equation 19 can be viewed as a Michaelis–Menten (**MM**) model (equation 1) with a negative feedback (Figure 3c). The addition of decoy or shunt pathways increases the strength of the negative feedback loop and shifts the switch point of input-output transfer function to higher values (Figure 3d). If we fit the simulation results of Equation 19 to a Hill function, we find that the effective dissociation constant scales with the number of shunt reactions (Figure 3e **Kdeff=m·Kd**). For a very large number of shunt reactions or very strong negative feedback loops, Equation 19 can be approximated as a linear-linear function (insert of figure 3d), with a very weak signal and is given by:

$$z=z_{max}\cdot Prf_1+z_0 \qquad (20.1)$$

$$z=z_{max}\cdot x_T m\cdot Kd+z_0 \qquad (20.2)$$

Equation 20 represents the copy number of synthetized proteins in an open loop and shunt circuit. To amplify the weak signal of the open loop circuit, a positive feedback loop regulating only the target promoter, was included [26]. Figure 4a shows the new positive feedback loop and shunt circuit. Danial et al. used external inducers (e.g., AHL as a quorum sensing molecule or arabinose) to trigger the positive feedback loop and shunt. A simple model of the circuit is presented in Figure 4b and includes three elements:

(1) a linear circuit that demonstrates the contribution of shunt reactions (Equation 20), (2) a positive feedback loop, and (3) a multiplication operator. The inducer-transcription factor binding reaction is modeled by a multiplication between the transcription factor and Hill function *(xT=z\*f(In))*. By substituting the last expression into Equation 20, we can express the solution of a graded positive feedback loop and shunt circuit as:

$$z = z_0 / 1 - zmax/m \cdot Kd \cdot (In) \quad (21)$$

We can distinguish between two cases: (1) a very strong (*zmax/m·Kd>>1*) positive feedback loop, which yields a sharp input-output transfer function. In this case, the inducer-output protein transfer function is set by the transcription factor–promoter binding reaction and inducer-transcription factor binding reaction. The solution in this case is obtained by substituting *xT=z\*f(In)* into Equation 19 (Figure 4c). (2) A graded positive feedback (*zmax/m·Kd<<1*), which yields a log-linear transduction between input and output (Figure 4c). This can be achieved by increasing the number of shunted biochemical reactions, or by decreasing the binding efficiency of transcription factors to the promoter, or decreasing the translation/transcription rates of proteins affecting *zmax*. In this case, the inducer-output protein transfer function is set by the inducer-transcription factor binding reaction only and is given by:

$$z \approx z_0 \cdot (1 + z_{max}/m \cdot Kd \cdot f(In)) \quad (22$$

Figure 4c shows that the reduction of *zmax/m·Kd* broadens the input dynamic range of the positive feedback circuit. We can see that our analytical model fits (Equation 22) the exact model constructed based on biochemical reactions. The positive feedback loop and shunt circuit cannot widen the input dynamic range (*IDR*) more than the dynamic range of the inducer-transcription factor binding reaction (Figure 4d). The maximum signal that can be achieved in such a system is *z=z0·(1+zmax/m·Kd)*, and therefore, the addition of shunt biochemical reactions decreases the signal output. A simple explanation was provided by *Danial et al* [26], who suggest that the shunt creates several binding sites that delay the saturation of the transcription factor-binding site reaction at the target promoter. At the same time, as the inducer concentration increases, the positive feedback loop enables continuous production of just enough transcription factors.

Synthetic analog parts that operate in a log-linear mode with a wide input dynamic range, can be simply integrated into more complex circuits for higher order functions [26]. For example, a genetic analog adder has been constructed in living cells by simply combining two analog synthetic parts (e.g., positive feedback loop and shunt) that each accept different input molecules and produce the same output molecules [26]. The addition operator was achieved by summing up the common diffusion fluxes of output molecules [26]. This operation is equivalent to Kirchhoff's current law in electronics. By contrast, a genetic digital adder cannot be constructed using the same principle that exploits a common

output signal, since every wire in digital design represents only a bit of information, and would require an additional stage to hold the carry out. For example, building a half 1 bit adder in bacteria requires 7 synthetic parts [34]. Analog computation presents an alternative to digital computation when the number of synthetic parts is limited. An analog subtractor can be constructed using the same principles applied for the analog adder [26]. The analog subtractor has two log-linear stages that produce common output proteins, one stage with a positive slope and another stage with a negative slope. *Danial et al.* [26] has used a LacI repressor to implement an analog stage with a negative slope.

## Noise margin of digital systems in living cells:

Figure 1c describes the input-to-output transfer function of Equation 1 in a digital mode, as log-log (*s=log (z/zmax) vs y=*log (*x/Kd*)). It can be viewed at two discrete levels (low and high):

$$s = \{ s_H \ y \geq y_{HSL} \ y \leq y_L \quad (23)$$

This device demonstrates a buffer logic gate operating in its extreme regions. This is exactly the opposite of its use as an analog device, where it operates in a log-linear region, at the middle of the transfer function (*x=Kd*). Digital logic gates utilize the gross nonlinearity exhibited by biochemical reactions in living cells. With these observations, the low-level output (*sL*) does not depend on the exact value of the input signal (*y*) as long as it does not exceed the low-level input (*yL*). Similarly, we observe that high-level output (*sH*) does not depend on the exact value of the input signal (*yL*), as long as its value does not fall below the high-level input (*yH*). When the input signal is higher than the low-level input and lower than the high-level input (*yL<y<yH*), the output increases and the logic gate enters its transition region, where the device can only act as an analog device. Similarly, we can define the logic levels for others logic gates. Ideal digital logic gates have a zero width of transition region and infinite sharpness of input-to-output transfer functions (very high Hill coefficient), operating in the middle of their transfer function at *x=Kd* (*yL=yH=0*), with maximum gain. However, the presence of intrinsic and extrinsic noise in biological and electronic systems limits the performance of ideal logic gates and drives a transition region with an input noise margin (*INM*) and output noise margin (*ONM*):

$$INM \equiv y_H - y_L = \log(x_H/x_L) \quad (24)$$

$$ONM \equiv s_H - s_L = \log(z_H/z_L)$$

The minimum *INM* will be set by the extrinsic noise of the input system and is given by:

$$INM \geq \sqrt{1 + b_{ext}/Kd} \quad (25)$$

In Equation 24, we assumed that the buffer logic gate operates at *x=Kd* (*INM>σY*). As we have shown, the transition region (or *IDR* in analog systems) is set by a Hill coefficient. For simplicity, we approximate the *INM≈1/n* (Equation 4). Then, Equation 25 can be given by:

$$n \leq \sqrt{Kd/1 + b_{ext}} \quad (26)$$

Equation 26 is demonstrated in Figure 5a, which shows that for a low level of input protein (or a low dissociation constant; for simplicity we assumed that **Kd=Zmax/2**), the digital logic gate should operate with a very high input noise margin and a low Hill coefficient. Under these conditions, the system has a graded behavior, acting as an analog system. Our analysis has shown that noise limits the performance of both analog and digital systems in living cells, rendering them extremely challenging to operate with low level of proteins. Alternatively, operating with a high level of input proteins can improve the performance of digital systems and reduce the **INM**. However, it increases the output noise margin. To quantify this insensitivity property, we consider the situation that often occurs in digital systems, where one buffer logic gate drives another buffer logic gate (Figure 5b). In this case, the digital cascade can only operate properly when the low-level output (**sL**) of the first stage is lower than the input level (**yL**) of the second stage and when the high-level output (**sH**) of the first stage is higher than the input level (**yH**) of the second stage. The output of the first stage often includes an intrinsic noise which sets the limits on the performance of the cascade (Figure 5b), and therefore we can write:

$$y_L \geq s_L + \sigma_{sL} \qquad (27.1)$$

$$y_H \leq s_H - \sigma_{sH} \qquad (27.2)$$

Because the relation between the output **s** and output **z** is described by a log-linear function, the noise of output **s** is given by $\sigma_s = \sigma_z/\bar{z}$ and $\sigma_s = \sqrt{(1+b_{int})}/\bar{z}$. Subtracting Equation 27.2 from Equation 27.1, i.e., substituting the noise of the output **s** and assumed that **zH >> zL** gives:

$$ONM \approx INM + \sqrt{1+b_{int}/z_{\bar{\bar{L}}}} \qquad (28)$$

To better understand Equation 28, we will cascade three identical buffer logic gates with **zL=10** and **bint=9** (Figure 5c). The resulting output noise margin of every layer **i** is larger than its input noise margin by one order of magnitude (**ONMi=INMi+1**), and the output noise margin of the last layer is larger than the input noise margin of the first layer by three orders of magnitudes (**ONM3=INM1+3**). For example, constructing a cascade of three logic layers, with an initial input noise margin of one order of magnitude, causes the last stage to have a very wide input dynamic range spanning 4 order of magnitudes (Figure 5c). Alternatively, we can increase the low-level output to 100 molecules, achieving ONM≈INM, however in this case, the high-level output is set to very high values. The basal level (**z0**) in synthetic biological parts, is often very large, and therefore, it sets the low-level output of digital systems (**zL≈z0**). If we substitute the **INM** in Equation 25 into Equation 28, we find other important relations:

$$ONM = \sqrt{1+b_{ext}/K_d} + \sqrt{1+b_{int}/z_{\bar{\bar{0}}}} \qquad (29.1)$$

$$ONM \approx 1/n + \sqrt{1+b_{int}/z_{\bar{\bar{0}}}} \qquad (29.2)$$

The last two equations quantify how the output noise margin of digital systems in living cells relates to intrinsic and extrinsic noise sources, Hill coefficient, basal level and molecule counts. Based on our analysis, in contrast to analog systems, the basal level is extremely important in determining the performance of digital systems (Figure 5d).

Digital computation in living cells have been widely used in synthetic biology and have been reviewed in several articles [7,8,35]. In this article, we briefly reviewed and discussed two key synthetic digital devices that were implemented in living cells. In the AND logic gate, output is only high if all inputs are high. The devices were constructed in bacteria [36,20], yeast [37] and mammalian cells [17], using a binding reaction between two synthetic parts regulated by input promoters. For example, Nissim et al. [20] constructed a system with two inputs that are duplicates of endogenous promoters that regulate the expression of a two-hybrid system, with one part fused to an activation domain, and the other to a binding domain. Together, they form a transcriptional complex that can bind a synthetic output promoter to express an output gene. By design, output is only generated if both endogenous promoters are active in the cell above a specific threshold. In the OR logic gate, output is high if at least one input is high. This device was constructed using two promoters that regulate the same gene [18]. Taking different approaches, several groups have constructed logic gates and memory using recombinase proteins [22].

## Summary:

Synthetic and Systems Biology have recently learned to exploit analog and digital genetic circuits for computation and decision making. In this work, we analyzed the precision of analog systems (Equation 18) and the noise margin of digital systems (Equation 29). We demonstrated that the performance of analog and digital systems in living cells is significantly impacted by extrinsic and intrinsic noise sources. We showed that both systems are challenging to operate with low protein levels and that both systems require optimization. For example, analog computation operates with Hill coefficients smaller than 1 and cascading of digital systems increases the input noise margin, conditions under which the digital system has a graded behavior acting as an analog system. We also have shown that, in contrast to analog systems, the basal level is extremely important in determining the performance of digital systems. Furthermore, we argue that, compared to digital design, analog computation is very efficient in its use of synthetic parts, however, embedded digital systems can operate reliably with low molecular counts. Therefore, biological systems that integrate both analog and digital circuits may provide an alternative strategy for scaling the complexities of computation in living cells [26], [27]. Although this design is widely used in electronics, in such contexts, it mostly aims to convert analog signals to a two-logic states and not to build efficient systems. Therefore, in our opinion, a hybrid analog-digital architecture in living cells should take a different approach than in electronics.
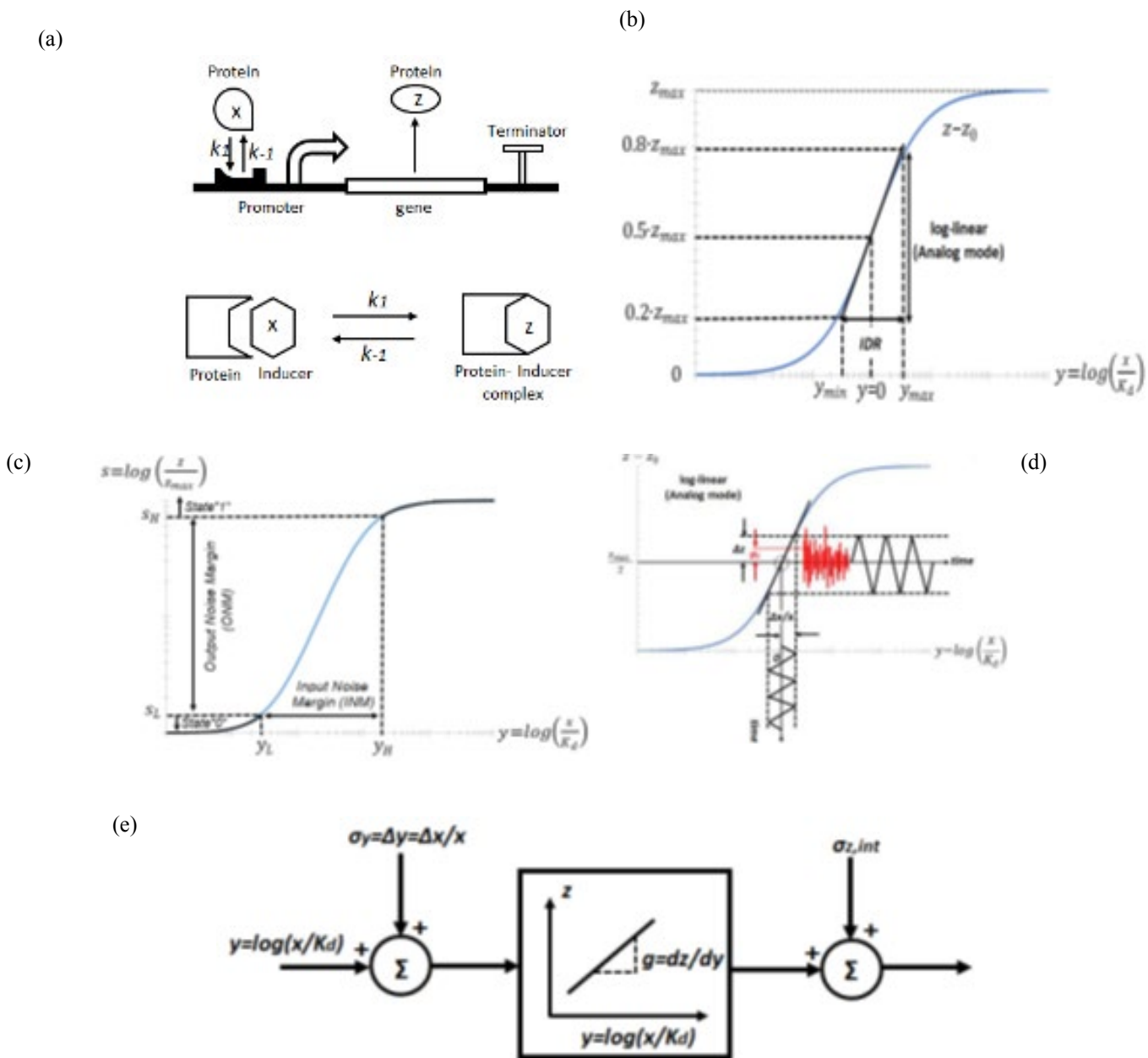
**Figure 1:** (a) Basic bio-computing elements in living cells, including a protein that binds to a promoter and an inducer that binds a promoter ($k1$ is the forward rate of the binding reaction, $k-1$ is the reverse rate of the binding (unbinding) reaction). (b) Analog mode: input-to-output transfer function of equation 1 (blue line) and log-linear function at y=0 (black line). (c) Digital mode: Equation 1 represents two logic states "0/1". (d) Noise analysis of log-linear analog systems, (e) Total noise is the sum of amplified extrinsic noise and intrinsic noise in a log-linear analog systems
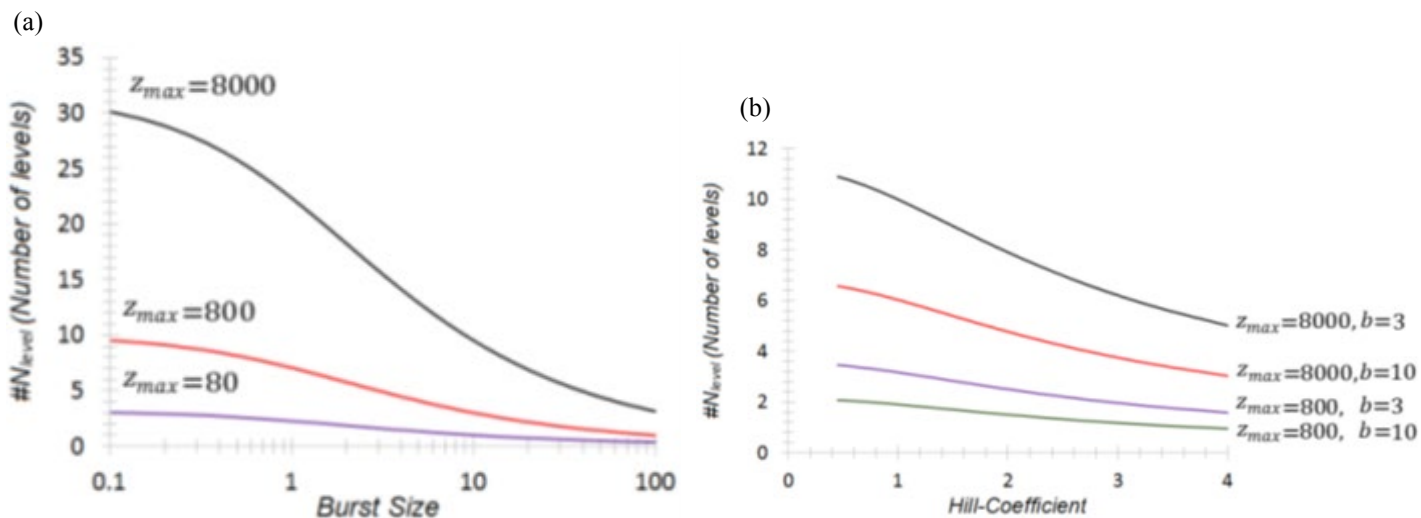
**Figure 2:** Noise Tolerance Analysis for reliable analog computation in living cells. (a) The figure shows the precision of analog systems in a log-linear mode, when we take into account the contribution of intrinsic noise and input dynamic range (Equation 17). (b) The figure shows the precision of analog systems in a log-linear mode, when we take into account the contribution of extrinsic/intrinsic noise and input dynamic range (Equation 18).

(e)
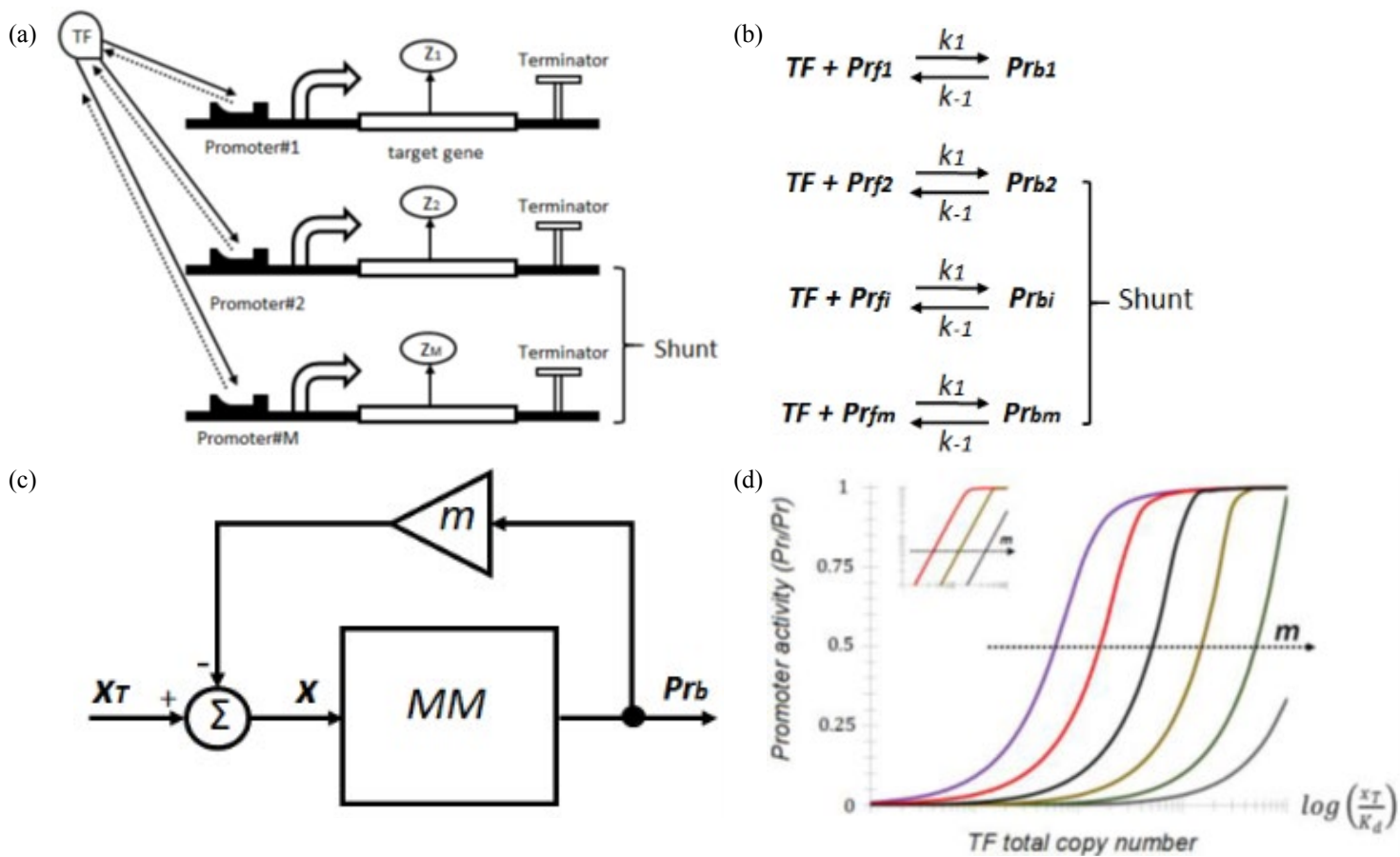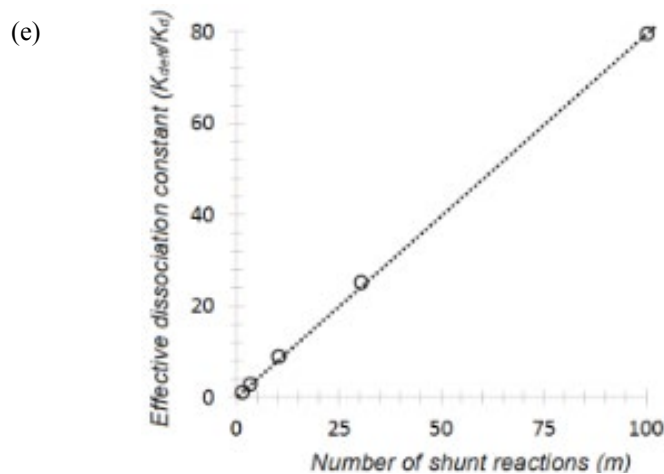


**Figure 3:** (a) Open loop and shunt circuit: a transcription factor binds to m identical promoters, (b) biochemical reaction model of an open loop and shunt circuit, (c) a schematic model includes Michaelis–Menten and a negative feedback for the open loop and shunt circuit. (d) Simulation results show the contribution of shunt biochemical reactions on the activity of the target promoter. (e) The simulation results show the contribution of the shunt biochemical reactions on the effective dissociation constant.
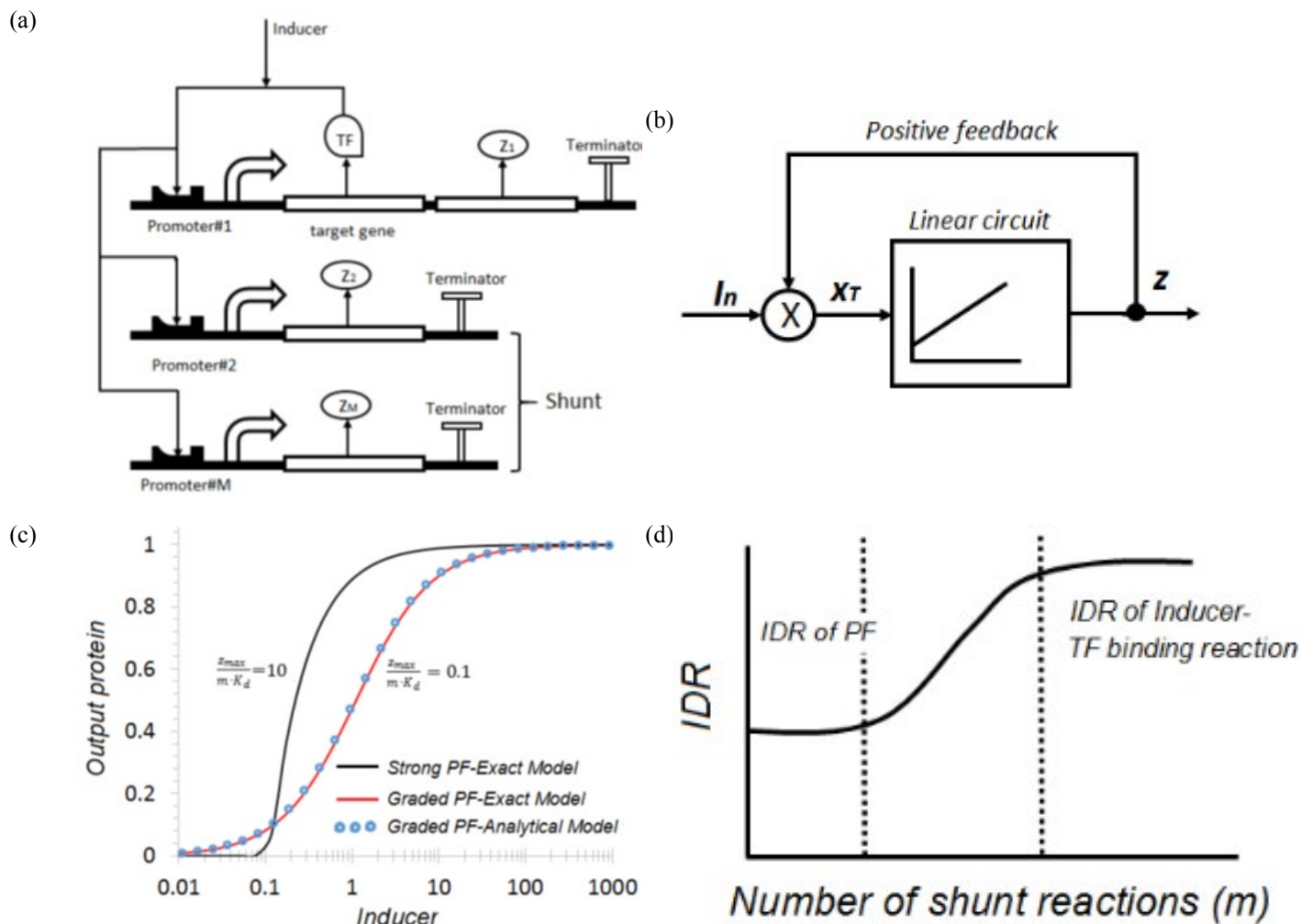


**Figure 4:** (a) Positive feedback loop and shunt circuit: the transcription factor is produced by its own promoter and binds to m identical promoters. (b) A schematic model of positive feedback and shunt circuit includes a linear part. (c) Simulation and analytical results showing a graded positive feedback loop. (d) A schematic model for the input dynamic range of positive feedback shows the contribution of shunt biochemical reactions
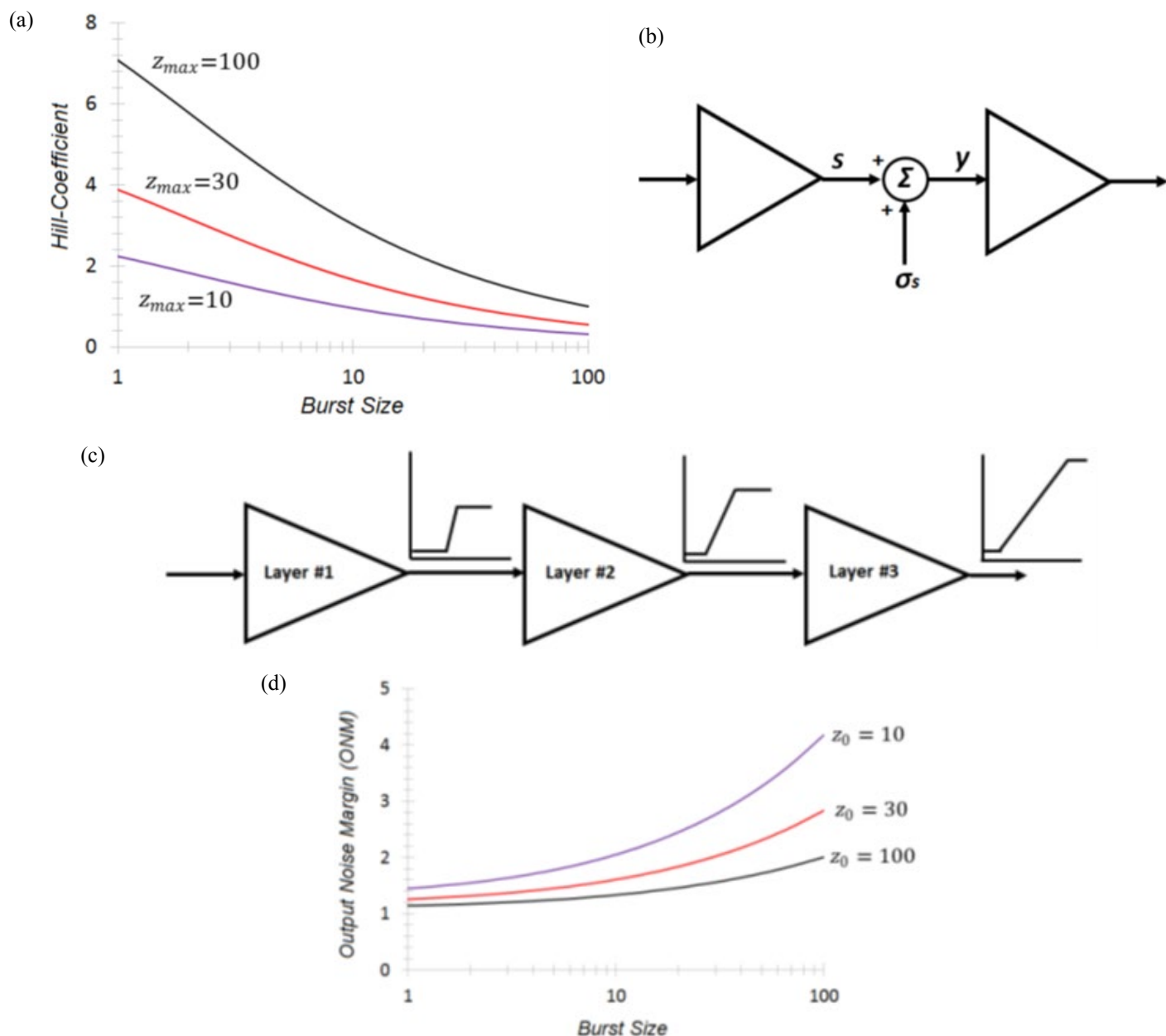
Figure 5: (a) Input noise margin analysis for digital circuits in living cells (Equation 26). (b) Contribution of intrinsic noise to the output noise margin in digital circuits in living cells. (c) Contribution of noise margin on cascading digital circuits in living cells (d) Output noise margin analysis for digital circuits in living cells (Equation 29).

### References:

1. Sarpeshkar, R. Ultra Low Power Bioelectronics: Fundamentals, Biomedical Applications, and Bio-Inspired Systems (CambridgeUniv. Press, 2010).

2. Elowitz MB, Leibler S. 2000 A synthetic oscillatory network of transcriptional regulators. Nature 403, 335–338. (doi:10.1038/35002125)

3. Gardner, T.S., Cantor, C.R., and Collins, J.J. (2000) Construction of a genetic toggle switch in Escherichia coli. Nature, 403,339–342.

4. Hooshangi, S., Thiberge, S., and Weiss, R. (2005) Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. Proc. Natl Acad. Sci. USA, 102, 3581–3586.

5. Carlson, R, Biology is Technology: the promise peril and new business of engineering life (Harvard University Press, 2010).

6. Lu, T.K., Khalil, A.S., and Collins, J.J. (2009) Next-generation synthetic gene networks. Nat. Biotechnol., 27, 1139–1150.

7. B. Jusiak, R. Daniel, F. Farzadfard, L. Nissim, O. Purcell, J. Rubens, and T. K. Lu, "Synthetic Gene Networks", Encyclopedia of Molecular Cell Biology and Molecular Medicine, October 20, 2014

8. Purnick, P. E. M., and Weiss, R., (June 2009). The second wave of synthetic biology: from modules to systems. Nature Reviews Molecular Cell Biology, DOI:10.1038/nrm2698

9. Chen YY, Galloway KE, and Smolke CD. Synthetic biology: advancing biological frontiers by building synthetic systems. Genome Biol. 13, 240 (2012) (doi:10.1186/gb-2012-13-2-240).

10. Baker D, Church G, Collins J, Endy D, Jacobson J, Keasling J, Modrich P, Smolke C, and Weiss R. Engineering life: building a FAB for biology. Sci. Am. 294, 44–51 (2006) (doi:10.1038/scientificamerican0606-44).

11. Ro DK, Paradise EM, Ouellet M, Fisher KJ, Newman KL, et al. 2006. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. Nature 440:940–43

12. JW, Kerns SJ, Shaket LA, Siraj, L, Collins JJ, Way JC, and Silver PA. Programmable bacteria detect and record an environmental signal in the mammalian gut. Proc Natl Acad Sci, 111:4838-43 (2014).

13. Z Xie, Wroblewska L, Prochazka L, Weiss R, and Benenson, Y. Multi-input RNAi-based logic circuit for identification of specific cancer cells. Science, 333, 1307-11 (2011)

14. Weber, W. and M. Fussenegger, Emerging biomedical applications of synthetic biology. Nat Rev Genet, 13(1): p. 21-35 (2012).

15. Ro DK, Paradise EM, Ouellet M, Fisher KJ, Newman KL, et al. 2006. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. Nature 440:940–43

16. JW, Kerns SJ, Shaket LA, Siraj, L, Collins JJ, Way JC, and Silver PA. Programmable bacteria detect and record an environmental signal in the mammalian gut. Proc Natl Acad Sci, 111:4838-43 (2014).

17. Nissim, L. and Bar-Ziv, R.H. (2010) A tunable dual-promoter integrator for targeting of cancer cells. Mol. Syst. Biol., 6, 444.

18. Tamsir A, Tabor J.J. and Voigt C.A. Robust multicellular computing using genetically encoded NOR gates and chemical 'wires'. Nature 469, 212–215 (2011).

19. Bonnet J, Yin P, Ortiz M.E, Subsoontorn P and Endy D, Amplifying genetic logic gates, Science 340, 559-603 (2013) ( doi: 10.1126/science.1232758).

20. Moon T.S, Lou C, Tamsir A, Stanton B.C, and Voigt C.A, Genetic programs constructed from layered logic gates in single cells, Natrue, 491, 249-253 (2012).

21. Ajo-Franklin C.M, Drubin D.A, Eskin J.A, Gee E.P.S, Landgraf D, Phillips I and Silver P.A, Rational design of memory in eukaryotic cells. Genes Dev. 21, 2271–2276 (2007).

22. Siuti P, Yazbek J, and Lu T.K, Synthetic circuits integrating logic and memory in living cells, Nat. Biotechnol., 31, 448-452 (2013).

23. Sauro H.M and Kim K, Synthetic biology: it's an analog world. Nature 497, 572–573 (2013) (doi:10.1038/nature12246).

24. Arkin A.P and Cardinale S, Contextualizing context for synthetic biology–identifying causes of failure of synthetic biological systems, Biotechnology journal 7,7,856-866 (2012).

25. Brophy J.A.N and Voigt C.A, Principles of genetic circuits design, Nat Methods, 101: 508-520 (2014).

26. Daniel R, Rubens J.R, Sarpeshkar R and Lu T.K, Synthetic analog computation in living cells. Nature 497, 619–623 (2013) (doi:10.1038/nature12148)

27. Sarpeshkar R, Analog Synthetic Biology, Phil. Trans. R. Soc. A 2014 372 (2014).

28. Uri A. 2007 An introduction to systems biology: design principles of biological circuits. Boca Raton, FL: Chapman & Hall/CRC.

29. Ozbudak EM, Thattai M, Kurtser I, Grossman AD, van Oudenaarden A. 2002 Regulation of noise in the expression of a single gene. Nat. Genet. 31, 69–73. (doi:10.1038/ng869)

30. Ferrell JE. 2009 Signaling motifs and Weber's law. Mol. Cell 36, 724–727. (doi:10.1016/j.molcel.2009.11.032

31. Elowitz MB, Levine AJ, Siggia ED, Swain PS. Stochastic Gene Expression in a Single Cell Science. 2002 Aug 16;297(5584):1183-6.

32. Swain PS, Elowitz MB, Siggia ED. Intrinsic and Extrinsic Contributions to Stochasticity in Gene Expression Proc Natl Acad Sci U S A. 2002 Oct 1;99(20):12795-800

33. Ron Milo and Rob Phillips, Cell Biology by Numbers. http://book.bionumbers.org/

34. Wong A, Wang H, Poh Ch, Kitne RI, Layering genetic circuits to build a single, bacterial half adder. BMC Biology DOI: 10.1186/s12915-015-0146-0 (2015).

35. O. Purcell and T. K. Lu, "Synthetic Analog and Digital Circuits for Cellular Computation and Memory", Current Opinion in Biotechnology, vol. 29, pp. 146-155, October 2014.

36. Anderson, JC, Voigt, CA, & Arkin, AP. (2007). Environmental signal integration by a modular AND gate. Molecular Systems Biology, 3: 133.

37. Nissim L, Beatus T, Bar-Ziv R., An autonomous system for identifying and governing a cell's state in yeast, Phys Biol. 2007 Aug 16;4(3):154-63.