Design and Characterization of a Three-terminal Transcriptional Device Through Polymerase Per Second

Prasanna Amur Varadarajan and Domitilla Del Vecchio*, *Member, IEEE*

***Abstract*—In this paper, we provide an *in silico* input–output characterization of a three-terminal transcriptional device employing polymerase per second (PoPS) as input and output. The device is assembled from well-characterized parts of the bacteriophage λ switch transcriptional circuit. We draw the analogy between voltage and protein concentration and between current and PoPS to demonstrate that the characteristics of the three-terminal transcriptional device are qualitatively similar to those of a bipolar junction transistor (BJT). In particular, as it occurs in a BJT, the device can be tuned to operate either as a linear amplifier or as a switch. When the device operates as a linear amplifier, gains of twofolds can be obtained, which are considerably smaller than those obtained in a BJT (in which 100-fold amplification gains can be reached). This fact suggests that the parts extracted from natural transcriptional systems may be naturally designed mostly to process and store information as opposed to amplify signals.**

***Index Terms*—Electronic devices, gene regulation, stochastic simulation.**

I. Introduction

In the field of synthetic biology [1], [2], one aims at extending or modifying the behavior of organisms, and controlling them to perform new tasks. Through the *de novo* construction of simple elements and circuits, the field aims to foster an engineering discipline for obtaining new cell behaviors in a predictable and reliable fashion. One of the fundamental building blocks employed in synthetic biology is the process of transcriptional regulation, which is already found in natural transcriptional networks [3]. A transcriptional network is composed of a number of genes that express proteins that then act as transcription factors for other genes. The rate at which a gene is transcribed is controlled by the *promoter*, a regulatory region of DNA that precedes the gene. RNA polymerase is a molecule that binds a defined site (a specific DNA sequence) on the promoter. The quality of this site specifies the transcription rate of the gene (the sequence of the site determines the chemical affinity of RNA polymerase to the site). RNA polymerase acts on all of the genes. However, each transcription factor modulates the transcription rate of a set of target genes. Transcription factors affect the transcription rate by binding specific sites on the promoter region of the regulated genes. When bound, they change the probability per unit time that RNA polymerase binds the promoter region. Transcription factors thus affect the rate at which RNA polymerase initiates transcription. A transcription factor can act as a *repressor* when it prevents RNA polymerase from binding to the promoter site. A transcription factor acts as an *activator* if it facilitates the binding of RNA polymerase to the promoter.

Synthetic biomolecular circuits are fabricated typically in bacteria *Escherichia coli* or in yeast, by cutting and pasting together according to a desired sequence genes and promoter sites (natural and engineered). Since the expression of a gene is under the control of the upstream promoter region, one can this way create a desired circuit of activation and repression interactions among genes. Early examples of such transcriptional circuits include an activator–repressor system that can display toggle switch or clock behavior [4], a loop oscillator called the repressilator obtained by connecting three inverters in a ring topology [5], a toggle switch obtained connecting two inverters in a ring fashion [6], an autorepressed circuit [7], and an autoregulatory genetic module [8].

It has been suggested by several researchers in the biological sciences that suitable abstractions are needed in order to manage the complexity of biological circuit design [1]. In engineering, such abstractions often take the form of input–output descriptions of processes. The input and output signal carriers take the form of voltage and current in electrical and electronic systems, pressure and mass flow in hydraulic systems, and temperature and heat flow in thermodynamic systems. The specific choice of input and output of a component is of paramount importance when components are input–output connected together to form more complicated systems. In fact, the property of the interconnection will affect the behavior of the composed system. For a discussion on this fundamental topic, the reader is referred to [9]–[12]. What the right input–output description is for a biomolecular component is still subject of intense research [11]. We refer to a transcriptional component as the entire process, called gene expression, that transforms a gene into the corresponding protein [13]. It has been customary to experimentally characterize a transcriptional component by measuring the steady state amount of protein produced as a function of the concentration of transcription factor [14]. Within this type of characterization, we can view the amount of transcription factor as the input to the transcriptional component and the amount of protein produced as the output of the transcriptional

Manuscript received October 21, 2008; revised June 6, 2009. Current version published January 4, 2010. *Asterisk indicates corresponding author.*

P. A. Varadarajan is with the Department of Mechanical Engineering, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: prasn@umich.edu).

*D. Del Vecchio is with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: ddv@umich.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNB.2009.2028687

component. Therefore, protein concentration is implicitly chosen as the signal carrier.

It has been recently proposed that the standard signal carrier that should instead be considered for transcriptional components is the RNA polymerase per second, named polymerase per second (PoPS) [15]. The PoPS is defined as the number of times that an RNA polymerase molecule passes a specific point on the DNA per unit time. Thus, this quantity is conceptually similar to the current in a wire connecting two electronic components, to the mass flow in a hydraulic pipe, or to the heat flow between two objects at different temperatures. Currently, it is experimentally difficult to directly measure PoPS *in vivo*, and thus it is also hard to obtain a PoPS input–output characterization of a transcriptional component.[1] We thus perform *in silico* experiments in order to characterize the PoPS input–output response of a transcriptional device, in which a third terminal is employed for tuning the PoPS input–output response. To this end, we consider a detailed reaction model of the transcription and translation processes involved in gene expression, including transcription (translation) initiation, elongation, and termination. In particular, we incorporate the motion of the RNA polymerase molecule along the DNA for the transcription elongation process and the motion of the ribosome along the messenger RNA (mRNA) for the translation elongation process. Similar to other researchers [8], [16], we choose as genes and promoter sequences composing the transcriptional component, parts from the bacteriophage λ switch circuit [17], for which plenty of experimental data is available [18]–[20]. Since we would like to evaluate stochastic effects possibly due to small numbers of molecules, the Gillespie algorithm is employed to simulate the system model [21]. A number of numerical experiments are performed in order to compute the mean and standard deviation of the static device characteristics. By establishing a conceptual analogy between currents and PoPS and voltages and protein amounts, we demonstrate how the obtained device characteristics are qualitatively similar to those of a bipolar junction transistor (BJT). In particular, the input–output PoPS characteristics can be tuned through the third terminal in order to obtain either a linear amplification or a switch behavior in analogy to the behavior of the BJT [22]. However, while gains of 100 can be reached in a BJT, only a gain of twofold is reached in the three-terminal transcriptional device of this paper.

This paper is organized as follows. In Section II, we review the mechanisms of gene expression, which is the fundamental building block of our device. In Section III, we introduce a comparison between standard abstractions used in electrical engineering and recently proposed abstractions for biomolecular engineering. Section IV introduces the topology and specifications of the three-terminal transcriptional device. Its detailed model is described in Section V, while the *in silico* experiments and characteristics are discussed in Section VI.

[1]For a discussion on this issue, the reader is referred to `http://openwetware.org/wiki/Endy:Measuring_PoPS_on_a_plate_reader`.

II. Process of Gene Expression: Transcription and Translation

The term gene expression refers to the process by which genetic information is ultimately transformed into working proteins. The main steps are *transcription* from DNA to mRNA, *translation* from mRNA to linear amino acid sequences, and folding of these into functional proteins, but several intermediate editing steps usually take place as well. The DNA molecule is a double-stranded helix made of a sugar-phosphate backbone and nucleotide bases. Each strand carries the same information, which is encoded in a four-letter alphabet. The two strands are held together by hydrogen bonds between the bases, which give stability but can be broken-up for replication or transcription. Genes are substrings of the complete DNA sequence. Besides genes, there are regulatory regions and start/stop regions that help delimit genes. Transcription initiates when an RNA polymerase (RNAp) molecule, an enzyme, binds the transcription site on the promoter upstream of the gene and switches from a closed complex to an open complex separating the two DNA strands from each other. Then, the RNAp travels along the gene sequence through the process of transcription elongation. During this process, the RNAp reads one base pair at the time and adds one nucleotide to the growing nucleotide chain forming the mRNA. When the RNAp reaches the end of the gene, it falls off the DNA and the transcription process terminates with the release of one mRNA molecule. Protein synthesis from mRNA occurs through the process of translation. Ribosomes, which are large complexes of RNA and protein, bind a specific site on the mRNA, the RBS, which is upstream of the gene. It then starts moving along the mRNA sequence in the translation elongation process. During this process, the ribosome reads three base-pairs at a time (a codon) and for each read codon, it adds a new ammino acid to the ammino acid sequence that will later fold into the protein. Once the stop codon is reached, the ribosome falls off the mRNA and the protein is generated. Each gene can be transcribed into mRNA at the same time by several RNAp molecules concurrently traveling on it. Similarly, each mRNA strand can be translated into protein by several ribosomes before it degrades. The mRNA is quite unstable and is degraded by RNase, which is an enzyme that binds to a site on the mRNA, which is very close to the RBS. Once RNase binds to such a site, it starts degradation of the transcript.

III. Abstractions in Electrical Engineering and in Biomolecular Circuit Design

In order to design, fabricate, and analyze circuits, electrical engineering makes use of abstractions, which consider increasing levels of complexities: parts, devices, and systems. Parts lie at the low end of complexity and include objects such as resistors, capacitors, inductors, semiconductors, which have specific (and optimized) properties and functions. Devices are fabricated out of parts and include, for example, two-terminal devices such as diodes and three-terminal devices such as BJT or field effect transistors (FET). Devices have terminals through which they can be connected to other devices to form systems. The standard signals carried through the

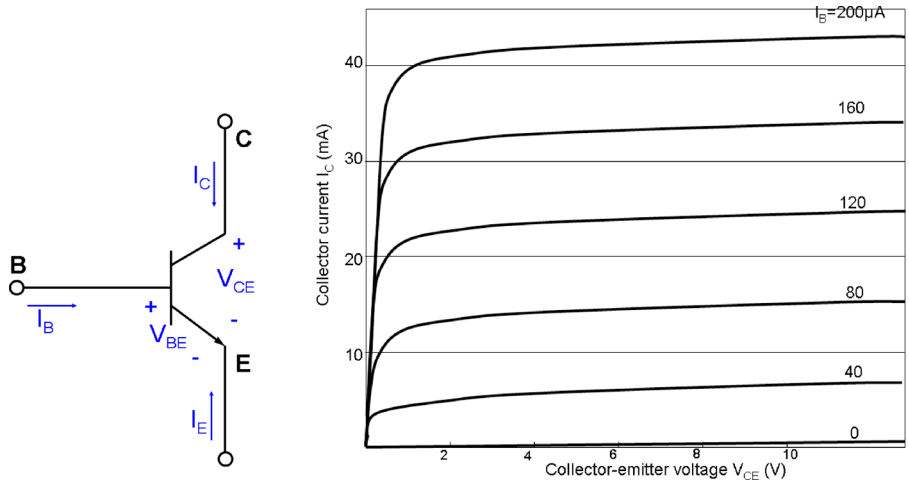

Fig. 1. Bipolar junction transistor (left) with its output characteristics (right).

TABLE I
PARTS, DEVICES, AND SYSTEMS

	Parts	Devices	Systems
Electrical circuits	resistor capacitor semiconductors	diodes BJT,FET	OPAMP amplifier
Bio-molecular circuits	gene promoter transcription factor	transcriptional component three-terminal devices	oscillator toggle switch

terminals are currents and voltages. Examples of popular electronic systems include the operational amplifier (OPAMP), current mirrors, inverting and non-inverting amplifiers, and oscillators. Such levels of abstractions have been proposed also for engineering biomolecular systems [1]. Biomolecular parts encode basic biological functions, such as genes, promoters, and transcription factors, which recognize specific DNA sequences upstream of a gene. Biomolecular devices are any combination of biomolecular parts that perform human-defined functions. One such example is provided by a transcriptional component that takes as input a transcription factor and produces as output another protein. Devices have terminals through which they can be connected to other devices, and which transmit signal carriers. It has been customary to use as signal carriers protein concentration. Endy [1] proposed to use PoPS as an alternative signal carrier at the device terminals. As PoPS represents the flow of RNAp molecules along the DNA, it is equivalent to a flux or a current. Such a flux is in turn induced by a transcription factor that binds to the promoter site upstream of a gene and alters the probability by which an RNAp molecule binds to the promoter and initiate transcription. Therefore, the amounts of transcription factor is directly responsible (possibly through a nonlinear function) of the value of PoPS in a way similar to how a voltage applied to a (nonlinear) resistor determines the current that flows in the resistor. Therefore, we view PoPS as playing the role of a current and protein amounts as playing the role of a voltage. Biomolecular systems include, for example, the oscillators of [4] and [5], the toggle switch of [6], the amplifier of [16], and the autoregulatory genetic module of [8]. A recent study on libraries of biomolecular parts, their combination to form devices, and the interconnection of devices to form systems is proposed in [23]. In this paper, we focus on the device abstraction level and in particular on the design of a three-terminal biomolecular device that has qualitatively the same functionality as a BJT.

IV. DEVICE TOPOLOGY AND PARTS

The BJT is one of the two principal semiconductor devices employed for signal amplification and switching. It has three terminals: the emitter, the collector, and the base, and it is represented in Fig. 1 (left). The configuration in which it is most employed is a common-emitter configuration, where the base current I_{B} and the voltage V_{CE} are the independent variables and the current I_{C} is with V_{BE} a dependent variable. The small base current I_{B} is employed as the input control quantity and V_{CE} is a parameter to tune the I_{B} to I_{C} response. Typical static characteristics are shown in Fig. 1 (right). These characteristics are divided into a linear region and a saturation region: the saturation region occurs for values of V_{CE} smaller than 0.3 V, while the linear region occurs for values of V_{CE} larger than 0.3 V. A BJT in the linear region has an almost linear I_{B} to I_{C} relationship, that is, $I_{\mathrm{C}} \approx \beta I_{\mathrm{B}}$, in which β is the current amplification gain and it is about 100. By virtue of these characteristics, the BJT can be employed as a current amplifier with amplification gain 100 when V_{CE} is in the linear operating region. For very small values of V_{CE}, the collector current becomes approximately independent of the base current and it reaches an almost constant saturation value as I_{B} increases. Because of these characteristics, the BJT can be properly connected to

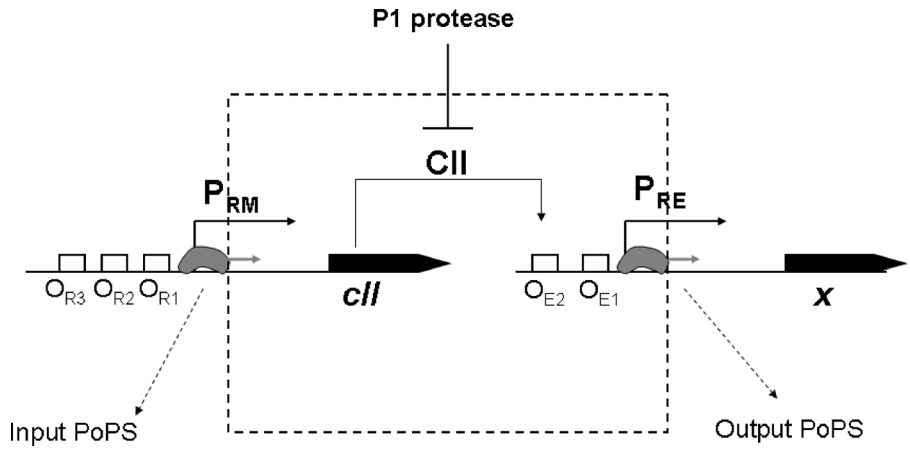

Fig. 2. Three-terminal transcriptional device.

optimally work either as a linear current amplifier or as a current ON-OFF switch [22].

The design of a transcriptional device qualitatively similar to a BJT focuses on reproducing input–output PoPS characteristics that can be either linear or switch-like depending on the value of the signal (protein amount) applied to a third terminal. The objective is thus to design a device that can be tuned to function as a linear amplifier or as a switch. The three-terminal transcriptional device we consider is shown in Fig. 2. It is composed of well characterized components from the λ-switch circuit [17] connected in a special way. Other authors have been investigating special ways of combining the parts from the λ-switch circuit to create synthetic biomolecular systems [8], [16]. In the new connection configuration of this paper, the cII gene is under the control of the $\mathrm{P_{RM}}$ promoter and the $\mathrm{P_{RE}}$ promoter is in turn activated by the CII protein. This protein is targeted for degradation by the P1 protease. The $\mathrm{P_{RM}}$ promoter has three operators at which CI or RNAp can bind. For transcription to start, RNAp must bind to at least one of the operators. If CI is bound to $\mathrm{O_{R2}}$, the probability that RNAp will bind to $\mathrm{P_{RM}}$ and initiate transcription is higher. Similarly, if CII is bound at $\mathrm{O_{E1}}$, the probability that RNAp will bind and start transcription is higher than is when CII is not at $\mathrm{O_{E1}}$. The device boundaries are shown in the same picture: for a fixed amount of protease P1, we view the RNAp flow along the DNA in the proximity of the $\mathrm{P_{RM}}$ promoter as the input to the device and the flow of the RNAp in the proximity of the $\mathrm{P_{RE}}$ promoter as the output of the device. We seek to characterize this device providing input–output PoPS characteristics when the amount of protease P1 is varied. Since the $\mathrm{P_{RE}}$ promoter has a (quite steep) sigmoidal activity response as a function of its transcription activator CII [18], we expect that if the amount of protease P1 is small, a small input PoPS at the $\mathrm{P_{RM}}$ promoter will cause immediate saturation of the PoPS at the $\mathrm{P_{RE}}$ promoter. This is because small values of input PoPS will cause large amounts of CII protein, which will in turn fully activate the $\mathrm{P_{RE}}$ promoter. If the amounts of protease P1 are large enough, we instead expect a more graded response between the input PoPS and the output PoPS. In fact, small amounts of input PoPS will translate to small amounts of protein CII, and thus the $\mathrm{P_{RE}}$ promoter will be more gradually activated as the input PoPS increases. To support this qualitative reasoning by quantitative data, we introduce a detailed reaction model for the three-terminal transcriptional device and perform numerical experiments to obtain the device PoPS characteristics.

V. Device Model and Integration

We employ a detailed reaction model for performing *in silico* experiments to determine the output characteristics of the three-terminal transcriptional device. For this sake we control the input PoPS by varying the amount of the CI protein, which acts as a transcription activator for the $\mathrm{P_{RM}}$ promoter after it has dimerized. The model that we consider thus includes the processes of transcription and translation of the genes involved in Fig. 2, the dimerization of the CI protein, and the degradation of the CII protein through its protease P1. The genetic mechanisms of transcription and translation are explicitly modeled by employing reaction models describing the motion of the RNAp along the DNA, the motion of the ribosome along the mRNA, and by considering mRNA degradation through RNase. The set of all reactions that we model are summarized in Table II along with the parameter values and the literature sources from which we obtained such values. Promoter operator sites are modeled using the statistical-thermodynamic model by Shea and Ackers [20] as such model and its parameters have been experimentally validated. This model describes the probability of each of the possible occupancy states of the operators. In order to take into account stochastic effects deriving from possibly low numbers of molecules, we employ a stochastic integration method based on the Gillespie algorithm [21]. The Gillespie algorithm relies on the Monte Carlo method to compute at each iteration the next reaction to occur in a set of possible reactions and the time interval after which it occurs. We describe next the details of the model and integration.

A. Model Description

1) Transcript Initiation: We employ the Shea and Ackers statistical-thermodynamic model [20] for computing the instantaneous probability of each of the possible occupancy states of the operator sites at both the $\mathrm{P_{RM}}$ and $\mathrm{P_{RE}}$ promoters. The occupancy states (or configurations) of the operator sites on the $\mathrm{P_{RE}}$ promoter are the same as those in [18]. The occupancy states of

TABLE II
KINETIC REACTIONS

Reaction	Parameter	Source
Dimerization		
2 CI $\underset{k_2}{\overset{k_1}{\leftrightarrows}}$ CI_2	$k_1 = 0.05\text{nM}^{-1}\text{s}^{-1}$, $k_2 = 0.5\text{s}^{-1}$	Arkin et al. [18]
Transcription initiation for P_{RM}		
RNAp+DNA $\overset{k}{\rightarrow}$ RNAp.DNA_1	$k = k_3 = 0.014\text{s}^{-1}$ (stimulated) $k = k_3 = 0.001s^{-1}$(basal)	Shea and Ackers [20]
Transcription initiation for P_{RE}		
RNAp+DNA $\overset{k}{\rightarrow}$ RNAp.DNA_1	$k = k_3' = 0.015\text{s}^{-1}$ (stimulated) $k = k_3' = 0.00004\text{s}^{-1}$(basal)	Arkin et al. [18]
Transcript elongation		
RNAp.DNA_n $\overset{k_4}{\rightarrow}$ RNAp.DNA_{n+1}	$k_4 = 30\text{nt s}^{-1}$	Arkin et al. [18]
RNAp.DNA_{80} $\overset{k_4}{\rightarrow}$ RNAp.DNA_{81}+RBS_1	$k_4 = 30\text{nt s}^{-1}$	–
Transcription termination		
RNAp.DNA_N $\overset{k_5}{\rightarrow}$ RNAp	$k_5 = 15\text{s}^{-1}$	–
Translation reactions		
Ribosome + RBS_i $\overset{k_6}{\rightarrow}$ RIB.RNA_1	$k_6 = 0.002\text{nM}^{-1}\text{ s}^{-1}$	–
$n \neq 80$ RIB.RNA_n $\overset{k_7}{\rightarrow}$RIB.$RNA_{n+1}$	$k_7 = 100\text{nt s}^{-1}$	–
RIB.RNA_{80} $\overset{k_7}{\rightarrow}$RIB.$RNA_{81}$+ RBS_i	$k_7 = 100\text{nt s}^{-1}$	McAdams and Arkin [19]
RIB.RNA_{294} $\overset{k_7}{\rightarrow}$ Ribosome + CII	$k_7 = 100\text{nt s}^{-1}$	
RNase + RBS_i $\overset{k_8}{\rightarrow}$ RNase	$k_8 = 0.2\text{nM}^{-1}\text{s}^{-1}$	Arkin et al. [18]
Degradation of CII		
CII + P1 $\underset{k_{10}}{\overset{k_9}{\leftrightarrows}}$ P1.CII	$k_9 = 0.01\text{nM}^{-1}\text{s}^{-1}$ $k_{10} = 0.01\text{s}^{-1}$	Arkin et al. [18]
P1.CII $\overset{k_{11}}{\rightarrow}$ P1	$k_{11} = 0.002\text{s}^{-1}$	Arkin et al. [18]

TABLE III
OPERATOR STATES FOR THE P_{RM} PROMOTER

Promoter state				Total Free Energy
s	O_R3	O_R2	O_R1	ΔG_s (Kcal)
1	O	O	O	0.0
2	O	O	R	-11.7
3	O	R	O	-10.1
4	R	O	O	-10.1
4*	RNAp	O	O	-11.5
5*	O	RNAp	—	-12.5
6	O	R	R	-23.7
7	R	O	R	-21.8
8	R	R	O	-21.6
9*	R	RNAp	—	-22.6
10*	RNAp	R	O	-21.6
11*	RNAp	O	R	-23.2
12	R	R	R	-33.8
13	RNAp	R	R	-35.2

the operator sites on the P_{RM} promoter must instead be adapted from those in [20] because in our system there is no Cro protein, which acts as a transcription factor for the P_{RM} promoter in the natural λ-switch circuit. Such occupancy states, adapted from [20] removing all those states in which Cro appears, are listed in Table III. The "O" denotes free operator, "R" denotes the CI dimer, and "—" denotes that the RNAp on the nearby site covers also that site. The asterisks denote transcriptionally active states, that is, operator configurations from which transcription can start. For each of the operator states, Shea and Ackers [20] formulated the expression of its probability as a function of protein concentration according to principles of thermodynamics. The probability of an operator configuration s is given by

$$< p_s >= \frac{\exp(-\Delta G_s/RT)[\text{R}]^i[\text{RNAp}]^k}{\Sigma_{s,i,k}\exp(-\Delta G_s/RT)[\text{R}]^i[\text{RNAp}]^k} \quad (1)$$

in which ΔG_s is the Gibbs free energy of the s operator state, R is the gas constant, T is the absolute temperature, [R], [RNAp] are the concentrations of unbound transcription factor (CI dimer for P_{RM} and CII for P_{RE}) and of RNAp. The probability of each promoter configuration is then used in the stochastic formulation of kinetics defined by Gillespie [21]. Assuming a rapid equilibrium of the binding and unbinding of proteins to the DNA (typically much faster than transcription and translation processes [3]), the binding state of the promoter is modeled by randomly choosing the promoter state at each instant using the probabilities given by the partition functions (1). If the promoter state selected is one from which transcription can occur then the transcript initiation reaction is included in the list of possible reactions of the next Monte Carlo step according to Gillespie algorithm. We have two different initiation rates for the promoters

corresponding to the stimulated and basal rates as considered in the Shea and Ackers model [20]. For P_{RM}, we select the stimulated initiation rate for all configurations in which RNAp is at P_{RM} and CI dimer is at O_{R2}, while we select the basal rate for those configurations in which RNAp is at P_{RM} and no CI dimer is at O_{R2}. Similarly for P_{RE}, we select the basal rate when CII is not at O_{E1} and RNAp is at P_{RE}, while we select the stimulated rate when RNAp is at P_{RE} and CII is at O_{E1}.

2) Transcript Elongation and Termination: The motion of the RNA polymerase along the DNA is modeled as a sequence of one step nucleotide reactions, in which each reaction is assumed to be unidirectional. Also, we assume that once the RNA polymerase is bound to the DNA, it just moves along the DNA and does not fall off before it reaches the termination region. In this model, we assume that at the termination site the RNA polymerase has probability one of detaching from the DNA. More sophisticated models [18] consider a non-unit probability of detaching.

3) Translation: Translation can start when a RBS becomes available at the beginning of the elongating mRNA strand, so that a ribosome can bind. We assume that an RBS becomes available after 80 nucleotides have been transcribed as according to [19] a ribosome can bind as soon as 80 nucleotides are available behind the transcribing RNA polymerase. The elongation process is modeled as described above and termination occurs when the ribosome reaches the end of the mRNA strand. The RNase E is responsible for mRNA degradation. It recognizes a site on the mRNA that is very close to the RBS, so that when a ribosome occludes the RBS, RNase E cannot bind to start degradation of mRNA. We, therefore, model as in [19] the degradation of mRNA by RNase as a competition between ribosome and RNase for the RBS, and we assume that once a ribosome has started translating, it concludes the process with the production of a protein. Once the ribosome reaches the end of the mRNA strand and a protein is produced, it falls off the mRNA.

Additionally, we model dimerization of the CI protein and degradation of the CII protein through binding to the P1 protease. For the *cII* gene, $N = 294$ nucleotides, while for the x gene we assumed for reducing the computation time that it is 50 nucleotides long. This is not relevant for our calculations as we are interested in the rates of RNA polymerase on the cII gene and on the x gene. For simulating the system employing Gillespie algorithm, all reactions rates k_i are substituted by reaction probabilities c_i. These are related to the rate constants k_i by powers of the volume V of the system (the *E. coli* cell in our case) with exponents as established in [21] on the basis of the reaction. The average volume of an *E. coli* cell fluctuates with cell growth [19]. In this paper, we take an average volume of 1.67×10^{-15} L. Such a volume can be approximately written as $V = 1$ (nM)$^{-1}$ by writing 1 L as one mole divided by 1 M and by employing the Avogadro number NA $= 6 \times 10^{23}$ in place of the mole [25]. This way, the actual numerical values of the reaction probabilities c_i in the Gillespie algorithm are identical to the numerical values of the reaction rates k_i as long as these are expressed in nM. This way, concentrations measured in nM directly correspond to numbers of molecules of the respective species in an *E. coli* cell.

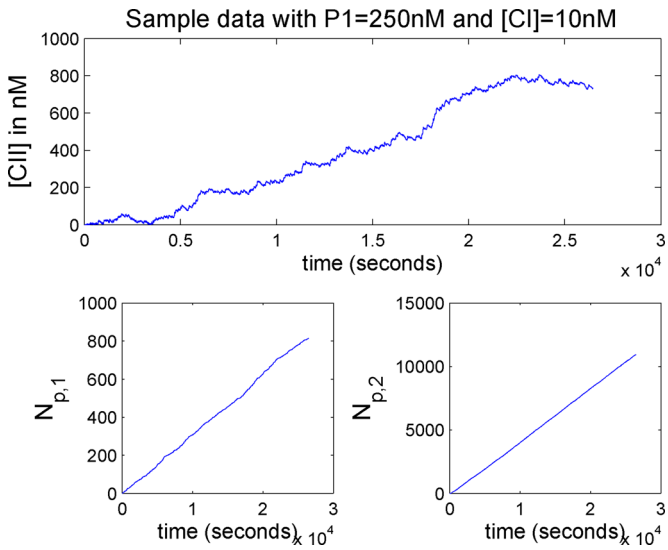

Fig. 3. Sample time data. On the y-axis the units are in number of molecules.

In all simulations, the number of RNA polymerase molecules is fixed at 30 and that of the ribosome molecules is fixed at 500 as in [18]. The concentration of CI is varied from 0 to 5 in steps of 1 and the protease P1 concentration is varied from 50 to 200 in steps of 10. All other concentrations are initially set to 0.

VI. Results and Discussion

A. Calculation of PoPS

The calculation of PoPS is performed by first counting the number of RNAp molecules that have transited at time t from nucleotide n to nucleotide $(n + 1)$. Then, the steady state value of PoPS is obtained by differentiating such a number with respect to time. For both the cII and x genes, we consider $n = 26$ to be the nucleotide at which the count of RNA polymerase molecules per unit time is performed: we keep track at every time step of the number of RNA polymerase molecules that have transited from the 26th nucleotide to the 27th nucleotide. Let, thus $N_p(t)$ represent the current cumulative number of RNAp molecules that have advanced from the 26th to the 27th nucleotide. We formally define the instantaneous PoPS as

$$\text{PoPS}(t) = \frac{d}{dt} N_p(t)$$

which is identical (apart from the absence of a minus sign) to the definition of electrical current if the number of RNAp is substituted by the number of electrons. Fig. 3 shows a sample of time plots obtained for the input $N_p(t)$ at P_{RM}, for the protein CII amount, and for the output $N_p(t)$ at P_{RE}. All simulations are run for 2 000 000 iterations and it is observed that the slope of $N_p(t)$ reaches an almost constant value after 10 000 iterations. The average steady state slope of $N_p(t)$ for each simulation is thus computed by taking ten intervals within the iterations between the 10 000th and the 2 000 000th and by computing for each interval the ratio between the difference of the N_p values at the ends of the interval and the difference of the corresponding

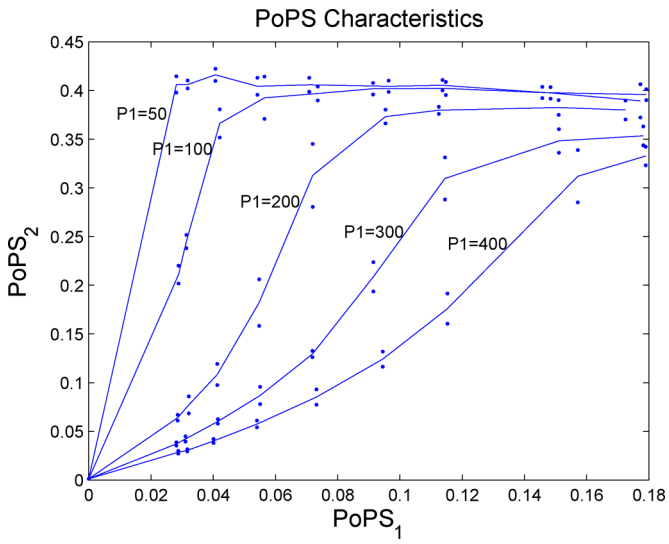

Fig. 4. PoPS Characteristics of the transcriptional device as the amounts of P1 are changed. The solid lines denote the sample mean over ten data samples, while the dots denote the standard deviation.

times at the ends of the interval. The average of the ten slopes so obtained is taken as the steady state PoPS value.

One of our objectives is to compute the input–output amplification gain of the device. Correctly establishing such an amplification on the basis of protein input and output amounts is much more difficult than establishing it on the basis of PoPS. First of all, input and output proteins are in general different from each other, so it is not necessarily clear what one means by amplification. This would have the same meaning as, for example, saying that three apples are amplified into six oranges. Second, different protein species have different decay rates. So, to correctly compute an amplification gain, one needs to normalize the protein amounts by their decay rates as performed for example in [16]. PoPS provides a natural way to measuring input–output responses of a device as the input and output species are the same and are involved in the same molecular processes.

B. PoPS Characteristics

The values of the input PoPS (PoPS_1) are controlled by varying the amounts of the CI protein from 0 to 20 nM. We consider values in steps of 50 nM of the amount of protease P1 of CII from 50 to 400 nM. For each combination of the CI and protease P1 amounts, we run ten experiments. For each time-run the steady state value of PoPS is computed as described in the previous section. Then, the sample mean and standard deviation of the steady state values of PoPS_1 and PoPS_2 are computed for each combination of CI and protease values using the ten data samples. The resulting characteristics of the input–output PoPS transcriptional device at the steady state are shown in Fig. 4. For any fixed amount of protease P1, the system is characterized by an input–output PoPS characteristic. If one lets the protease vary, one obtains a three-terminal device. Such a device can be controlled through the protease P1 in such a way to have an almost linear input–output PoPS characteristic with a linear gain of up to twofolds for high levels of protease. For low levels of protease, the system reaches a switch-like behavior. This device can thus be tuned to work as a linear amplifier or as a switch. In Fig. 4, the value of PoPS_2 in correspondence to $\mathrm{PoPS}_1 = 0$ is obtained from the basal initiation rate at $\mathrm{P_{RE}}$, and it is measured in correspondence of a zero amount of CII in the system. In Fig. 5, we compare the protein amount/PoPS plots of the three-terminal device of Fig. 2 to those of a BJT. This comparison is performed by viewing the input PoPS (PoPS_1) as playing a similar role as the base current I_{B}, the output PoPS (PoPS_2) as playing a similar role as the collector current I_{C}, and the protease P1 amount as playing a similar role as the collector-emitter voltage V_{CE}. In both devices characteristics, we can identify a saturation region and a linear region. In the saturation region, the collector current is almost independent of the base current in the bipolar junction transistor. Similarly, in such a region the output PoPS (PoPS_2) is almost independent of the input PoPS (PoPS_1) for the transcriptional three-terminal device. The saturation region for the three-terminal transcriptional device corresponds to lower levels of protease P1. In such a case, in fact the steady state concentration of such a protein can be very large even for small values of PoPS_1, thus saturating the P_{RE} promoter. Therefore, the values of PoPS_2 are always very high almost independently of the values of PoPS_1. As a consequence, for a zero value of PoPS_1, the value of PoPS_2 is at its small basal level. Then a minimal increase of PoPS_1 produces a switch of PoPS_2 to the highest values. This is better observed in Fig. 4 in correspondence to the plots for $\mathrm{P1} = 50$ nM, 100 nM. In the linear region, the relationships between the base current I_{B} and the collector current I_{C} for the BJT is almost linear. Similarly, the polymerase per second PoPS_1 and the polymerase per second PoPS_2 for the three-terminal transcriptional device is also linear for high values of the protease concentration P1. In fact, for higher values of protease concentration P1, the concentration of the CII protein corresponding to a value of the polymerase per second PoPS_1 will be lower. Therefore the P_{RE} promoter will not be saturated by protein CII. The linear gain between PoPS_1 and PoPS_2 can be more easily seen in Fig. 4 and reaches a value of twofolds. So, while the qualitative characteristics of the three-terminal transcriptional device are similar to those of the BJT, they largely differ in the value of the linear amplification gain that can be reached, which is about 100-fold for the BJT.

From the PoPS characteristics of Fig. 4, one deduces what key parameters regulate the linear amplification gain. If the maximum reachable PoPS_1 (about 0.18 in the figure) is high enough to attain the maximum reachable PoPS_2 (about 0.4 in the figure) for small values of protease P1, then the maximal achievable linear gain will be approximately given by $(\text{Maximal } \mathrm{PoPS}_2)/(\text{Maximal } \mathrm{PoPS}_1)$. As a consequence, to increase the linear gain of the device, it is necessary to choose promoters such that the output range of the one controlling PoPS_2 is much larger than the output range of the one controlling PoPS_1. These ranges are experimentally available for a number of promoters. For example, Ellis *et al.* [23] constructed a library of TetR-regulated promoters and LacI-regulated

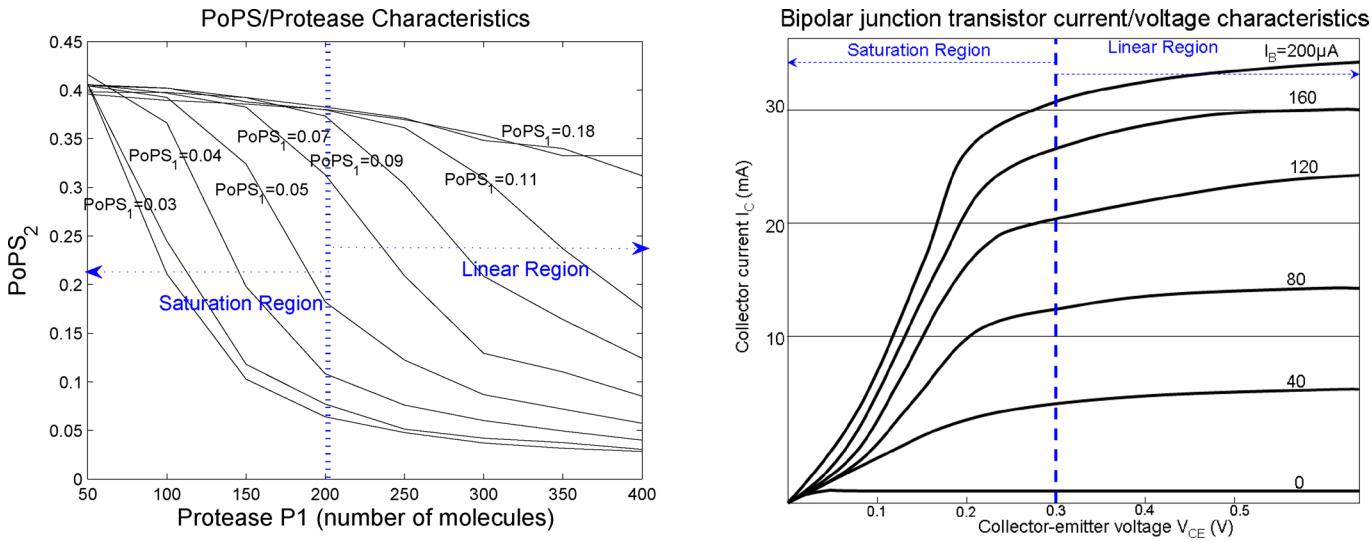

Fig. 5. Three-terminal device characteristics on the left compared to those of a bipolar junction transistor (BJT) on the right (taken from [22]). In both cases, depending on the values of the voltage/protein concentration at one of the terminals, we have a saturation region and a linear region. For a BJT, the saturation region usually starts around 0.3 V, while the operation value of the collector–emitter voltage in the linear region is about 0.7 V.

promoters and determined output ranges for each one of them. By their characterization data, one could combine promoters for controlling $PoPS_1$ and $PoPS_2$ in order to achieve approximately a linear gain of 30-fold, which is much larger than what we obtained here using promoters from the λ switch system. In summary, to achieve high linear gains, the promoter controlling $PoPS_2$ should be highly sensitive to its transcription factor (so that the maximal $PoPS_1$ is enough to achieve the maximal $PoPS_2$ for small protease values) and it should have an output range that is much higher than that of the promoter controlling $PoPS_1$.

These relatively low linear amplification gains obtained by combining biomolecular parts from transcriptional networks suggest that transcriptional networks have been designed mostly to process and store information as opposed to amplifying signals. The task of signal amplification in biomolecular systems may be exclusively covered by other devices such as phosphorylation/dephosphorylation loops and by systems obtained cascading them [24]. These systems transmit and amplify signals traveling from the cell membrane down to the nucleus, where transcriptional networks then process the received information. It has been theoretically shown that larger gains in transcriptional networks are needed to attenuate the effects of retroactivity at interconnections arising from downstream loads. Attenuating retroactivity effects enforces, in turn, network modularity [12]. Gains of two or tenfolds may be large enough to accomplish this task for the amount of downstream load that is naturally encountered in transcriptional networks. These gain values may instead be not enough to enforce modularity in synthetically constructed gene networks, in which one can arbitrarily increase the amounts of downstream load to any device or system. In such a case, it will be necessary to recur to parts extracted from systems that enable higher amplification gains such as, for example, signaling systems.

VII. Conclusion

In this work, we have defined the notion of PoPS in order to characterize the input–output response of a three-terminal transcriptional device. By making the analogy between PoPS and current and protein concentration and voltage, we showed how the characteristics of the three-terminal transcriptional device are qualitatively similar to those of a bipolar junction transistor (BJT). In particular, just like a BJT, the three-terminal transcriptional device can be tuned to function as a linear amplifier or as a switch. While the current amplification gain in a BJT is about 100-folds, we obtained a PoPS amplification gain of up to twofold for the three-terminal device. While significantly smaller than the amplification gain of a BJT, the PoPS gain may be enough for the proper functioning of natural transcriptional networks.

References

[1] D. Endy, "Foundations for engineering biology," *Nature*, vol. 438, no. 24, pp. 449–452, 2005.
[2] E. Andrianantoandro, S. Basu, D. K. Karig, and R. Weiss, "Synthetic biology: New engineering rules for an emerging discipline," *Mol. Syst. Biol.*, vol. 2, pp. 1–14, 2006.
[3] U. Alon, *An Introduction to Systems Biology. Design Principles of Biological Circuits*. London, U.K.: Chapman & Hall, 2007.
[4] M. R. Atkinson, M. A. Savageau, J. T. Meyers, and A. J. Ninfa, "Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in *Escherichia coli*," *Cell*, vol. 113, pp. 597–607, 2003.
[5] M. B. Elowitz and S. Liebler, "A synthetic oscillatory network of transcriptional regulators," *Nature*, vol. 403, pp. 339–342, 2000.
[6] T. S. Gardner, C. R. Cantor, and J. J. Collins, "Construction of the genetic toggle switch in *Escherichia Coli*," *Nature*, vol. 403, pp. 339–342, 2000.
[7] A. Becskei and L. Serrano, "Engineering stability in gene networks by autoregulation," *Nature*, vol. 405, pp. 590–593, 2000.
[8] F. J. Isaacs, J. Hasty, C. R. Cantor, and J. J. Collins, "Prediction and measurement of an autoregulatory genetic module," *PNAS*, vol. 100, no. 13, pp. 7714–7719, 2003.
[9] J. Willems, "Behaviors, latent variables, and interconnections," *Syst., Control Inf.*, vol. 43, pp. 453–464, 1999.

[10] J. Saez-Rodriguez, A. Kremling, and E. Gilles, "Dissecting the puzzle of life: Modularization of signal transduction networks," *Comput. Chem. Eng.*, vol. 29, pp. 619–629, 2005.
[11] J. Saez-Rodriguez, A. Kremling, H. Conzelmann, K. Bettenbrock, and E. D. Gilles, "Modular analysis of signal transduction networks," *IEEE Control Syst. Mag.*, vol. 24, no. 4, pp. 35–52, Aug. 2004.
[12] D. D. Vecchio, A. J. Ninfa, and E. D. Sontag, "Modular cell biology: Retroactivity and insulation," *Nature/EMBO Mol. Syst. Biol.*, vol. 4, pp. 161–176, 2008.
[13] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Molecular Biology of the Cell*. New York: Garland, 2002.
[14] N. Rosenfeld, J. W. Young, U. Alon, P. S. Swain, and M. B. Elowitz, "Gene regulation at the single-cell level," *Science*, vol. 307, pp. 1962–1965, 2005.
[15] D. Baker, G. Church, J. Collins, D. Endy, J. Jacobson, J. Keasling, P. Modrich, C. Smolke, and R. Weiss, "Engineering life: Building a FAB for biology," *Sci. Amer.*, vol. 294, pp. 44–51, Jun. 2006.
[16] G. D. Rubertis and S. W. Davies, "A genetic circuit amplifier: Design and simulation," *IEEE Trans. Nanobiosci.*, vol. 2, no. 4, pp. 239–246, Dec. 2003.
[17] M. Ptashne, *Genetic Switch: Phage Lambda Revisited*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 2004.
[18] A. Arkin, J. Ross, and H. H. McAdams, "Stochastic kinetic analysis of developmental pathway bifurcation in phage λ-infected *Escherichia coli* cells," *Genetics*, vol. 149, pp. 1633–1648, 1998.
[19] H. H. McAdams and A. Arkin, "Stochastic mechanisms in gene expression," *PNAS*, vol. 94, pp. 814–819, 1997.
[20] M. A. Shea and G. K. Ackers, "The O_r control system of bacteriophage lambda a physical-chemical model for gene regulation," *J. Mol. Biol.*, vol. 181, pp. 211–230, 1985.
[21] D. T. Gillespie, "Exact stochastic simulation of coupled chemical reactions," *J. Phys. Chem.*, vol. 81, pp. 2340–2361, 1977.
[22] D. L. Schilling and C. Belove, *Electronic Circuits: Discrete and Integrated*. New York: McGraw-Hill, 1968.
[23] T. Ellis, X. Wang, and J. J. Collins, "Diversity-based, model-guided construction of synthetic gene networks with predicted functions," *Nature Biotechnol.*, vol. 27, no. 5, pp. 465–471, 2009.
[24] B. N. Kholodenko, "Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades," *Eur. J. Biochem.*, vol. 267, pp. 1583–1588, 2000.
[25] R. Bundschuh, F. Hayot, and C. Jayaprakash, "Fluctuations and slow variables in genetic networks," *Biophys. J.* pp. 1606–1615, 2003.

Authors' photographs and biographies not available at the time of publication.