Messenger RNA fluctuations and regulatory RNAs shape the dynamics of a negative feedback loop

María Rodríguez Martínez*
Department of Physics of Complex Systems, Weizmann Institute of Science, Rehovot, Israel and Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel

Jordi Soriano
Departament d'ECM, Facultat de Física, Universitat de Barcelona, Barcelona, Spain

Tsvi Tlusty[†,‡]
Department of Physics of Complex Systems, Weizmann Institute of Science, Rehovot, Israel

Yitzhak Pilpel[†,§] and Itay Furman
Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel
(Received 16 June 2009; revised manuscript received 30 November 2009; published 30 March 2010)

Single-cell experiments of simple regulatory networks can markedly differ from cell population experiments. Such differences arise from stochastic events in individual cells that are averaged out in cell populations. For instance, while individual cells may show sustained oscillations in the concentrations of some proteins, such oscillations may appear damped in the population average. In this paper we investigate the role of RNA stochastic fluctuations as a leading force to produce a sustained excitatory behavior at the single-cell level. As opposed to some previous models, we build a fully stochastic model of a negative feedback loop that explicitly takes into account the RNA stochastic dynamics. We find that messenger RNA random fluctuations can be amplified during translation and produce sustained pulses of protein expression. Motivated by the recent appreciation of the importance of noncoding regulatory RNAs in post-transcription regulation, we also consider the possibility that a regulatory RNA transcript could bind to the messenger RNA and repress translation. Our findings show that the regulatory transcript helps reducing gene expression variability both at the single-cell level and at the cell population level.

DOI: 10.1103/PhysRevE.81.031924 PACS number(s): 87.10.Mn, 87.18.Tt

I. INTRODUCTION

The growing interest in biological noise has led to many efforts to measure gene expression at the single-cell level [1–3], revealing a very distinct dynamics when compared to population cell experiments [4]. In two well-studied examples, the p53-mdm2 regulatory network and the NF-κB signaling pathway, sustained oscillations are observed in single cells following activation signals [5–7], while cell population experiments only show damped oscillations [8–10]. In both cases the core circuit consists of a negative feedback loop, one of the most common network designs, where the active transcription factor promotes the transcription of its own repressor.

Stable oscillations are not trivially generated in a single negative feedback loop [11]. A loop composed of only two agents does not oscillate for plausible macroscopic equations. Sustained oscillations require at least three agents, where the third one introduces a time delay that repeatedly causes the system to overshoot or undershoot above and below the steady state [12]. Some models achieve sustained oscillations by introducing *ad hoc* time delays to reproduce those that a system incurs when manufacturing the various molecular components [9,13]. The dynamics can also be enriched by considering combinations of negative and positive feedback loops [14–16], bistable switches [17], or by inheriting oscillatory signals from upstream regulators [10,13].

In this paper we show that the stochastic fluctuations in gene expression in a negative feedback loop can produce sustained pulses of protein expression. It has been suggested that protein fluctuations are driven by underlying messenger RNA (mRNA) fluctuations [2,3,18]. We show that the mRNA stochastic fluctuations can be amplified during translation and induce a sustained excitatory behavior characterized by a series of sustained anticorrelated pulses in the expression of the positive and negative regulators of the loop.

Noise-induced oscillations have already been found in other systems. Oscillations in a circadian clock consisting of a combination of a positive and a negative feedback loop are enhanced by the intrinsic biochemical noise [14]. Resonant amplification of the stochastic fluctuations can lead to cycling behavior in the Volterra system [19] and in self-regulatory genes [20]. Here, we show that a simple negative feedback loop consisting of an activator protein and its repressor is capable of producing protein pulses when the stochastic fluctuations of the mRNA are taken into account. This result does not rely on having a large number of mol-

*Present address: Center for Computational Biology and Bioinformatics, Columbia University, New York, USA.
†Corresponding author.
‡tsvi.tlusty@weizmann.ac.il
§pilpel@weizmann.ac.il

ecules or on the particular statistical properties of the noise, neither depends on upstream pulsating signals or couplings to additional loops.

Recently several experimental studies have shown that gene expression occurs in bursts of transcriptional activity [21–23]. These bursts are usually ascribed to random upstream events, such as chromatin remodeling or random promoter transitions. Here, we demonstrate that sustained pulses of protein expression can be produced merely by the stochastic nature of mRNA kinetics.

In view of the crucial significance of mRNA fluctuations, we asked ourselves how gene expression can be accurately regulated in the noisy cellular environment. Regulation of gene expression is a complex multilayered process that involves many different players. Since their discovery more than a decade ago, regulatory RNAs (termed "regRNAs" in this paper) have emerged as key regulators in virtually all the cellular processes studied to date. regRNAs are noncoding RNAs that regulate gene expression by base pairing to a partially or fully complementary mRNA target. Micro-RNAs (miRNA) [24] and antisense RNA [25] are two examples of regulatory RNAs.

In the mammalian genomes regRNAs often share the same transcriptional regulation as their targets, giving rise to a diversity of feed-forward loops [26,27]. Such pairs of target and regRNA transcripts are found to be coregulated, coexpressed, or inversely expressed more frequently than expected by chance, presumably due to sharing of common transcription factors [25,27–29]. In particular, some transcription factors have been found to bind to overlapping transcript pairs, thus potentially coupling the regulation of a coding gene and its regRNA [26]. It was suggested that one potential purpose of such a design is to filter transcription noise [25].

In this work we consider the case in which the positive regulator in the loop transcribes both mRNA and regRNA. The latter could be either a miRNA or an antisense. Additionally, we assume that the regulatory transcript binds to the mRNA and prevents its translation, but without promoting its degradation. We show that the presence of regRNA reduces the excitability of the system by increasing its capacity to buffer the noise.

II. MODELS

We consider three alternative designs of a negative feedback loop. The main loop is composed of a transcription factor (the positive regulator P) and its repressor (the negative regulator N). We assume that in response to some external cellular signal, P has been activated and is promoting the transcription of N. Two of the designs also model the transcript dynamics. In the three models, P is degraded at the post-translational level via protein-protein interaction with N. Schematic representations of the three models, as well as the individual chemical reactions, are shown in Fig. 1. The description of the biochemical parameters and their range of variation are summarized in Table I. The values adopted correspond to typical mammalian cells.

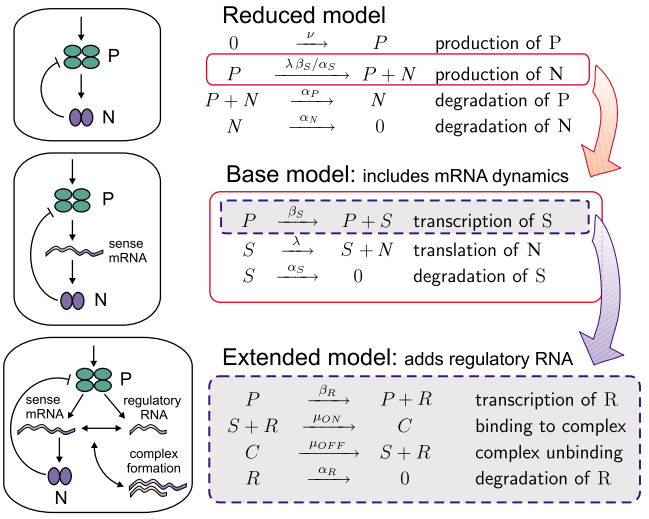

FIG. 1. (Color online) Schematic representation of the reduced, base, and extended models together with the chemical reactions. The reduced model, making up the simplest feedback loop, considers only the production and degradation of the proteins (P) and (N) and assumes that mRNA is in quasiequilibrium. The base model includes mRNA dynamics by expanding the chemical reaction of the production of N (solid red box). The additional chemical reactions describe the transcription of sense mRNA (S), the translation of N, and the degradation of S. Finally, the extended model puts into play regulatory RNAs, R (dashed blue box), by adding their transcription, degradation, and the formation and degradation of a sense-regRNA complex C.

A. Base model

The base model consists of the two protein regulators P and N, and an mRNA transcript S. (The latter will be interchangeably referred to as "transcript," "sense transcript," or "sense mRNA.") P promotes the transcription of S, which in

TABLE I. Symbols used as parameters of the models, their values or range of variation, and their description.

Symbol	Value (min^{-1})	Meaning
ν	200	Transcription factor (P) induction
β_S	0.05–10	mRNA (S) induction
β_R	0.01–20	Antisense (R) induction
λ	0.05–50	Protein (N) translation
μ_{ON}	0.001–10	Sense-regRNA binding
μ_{OFF}	0.001–10	Sense-regRNA unbinding
α_P	0.10	N-assisted decay of P
α_S	0.03	Sense autodegradation
α_R	0.03	regRNA autodegradation
α_N	0.10	Protein (N) autodegradation
n_S	3[a]	Hill's coeff., sense transcription
n_R	1[a]	Hill's coeff., regRNA transcription
k_S	500[a]	Threshold, sense transcription
k_R	300[a]	Threshold, regRNA transcription

[a]Dimensionless parameters.

turn encodes for the negative regulator N. The deterministic equations are

$$\frac{dP}{dt} = \nu - \alpha_P NP, \tag{1a}$$

$$\frac{dS}{dt} = \beta_S \frac{P^{n_S}}{P^{n_S} + k_S^{n_S}} - \alpha_S S, \tag{1b}$$

$$\frac{dN}{dt} = \lambda S - \alpha_N N. \tag{1c}$$

We assume that the cooperative binding of n_s molecules of P at the promoter site of N is required for efficient transcription, and we model the transcription rate with a Hill's function. The translation rate of N is proportional to the transcript level S. Finally, the rate of the N-mediated degradation of P is proportional to the concentration of both proteins.

B. Extended model: Incorporation of regulatory RNA

This model introduces a regRNA, denoted R, which targets S. P promotes the transcription of both S and R transcripts. R regulates S by base pairing to it, thus creating a hybrid S-R complex C. We assume that the complex molecule can unbind but not degrade, i.e., it returns to the system both complementary RNA transcripts.

The set of equations describing the model is given by

$$\frac{dP}{dt} = \nu - \alpha_P NP, \tag{2a}$$

$$\frac{dS}{dt} = \beta_S \frac{P^{n_S}}{P^{n_S} + k_S^{n_S}} - \mu_{ON} RS + \mu_{OFF} C - \alpha_S S, \tag{2b}$$

$$\frac{dR}{dt} = \beta_R \frac{P^{n_R}}{P^{n_R} + k_R^{n_R}} - \mu_{ON} RS + \mu_{OFF} C - \alpha_R R, \tag{2c}$$

$$\frac{dC}{dt} = \mu_{ON} RS - \mu_{OFF} C, \tag{2d}$$

$$\frac{dN}{dt} = \lambda S - \alpha_N N. \tag{2e}$$

The extended model contains 14 parameters described in Table I. In this work we focus on analyzing how the temporal behavior of the protein levels is influenced by the mRNA and regRNA transcription rates, the translation rate, and the S-R complex formation and destruction. Therefore, we systematically explored how the variation in β_S, β_R, λ, μ_{ON}, and μ_{OFF} affect the emergent system behavior. We maintained the remaining parameters constant throughout the simulations; random exploration showed that changes in their values did not affect the behavior of the system appreciably.

C. Reduced model: Neglected RNA kinetics

Finally, to isolate the effect of RNA fluctuations we consider a reduced model where the RNA transcripts are assumed to be in quasiequilibrium. From a molecular perspective, this model corresponds to the limit where the time scales associated with RNA transcription and degradation are much shorter than those associated with protein production and degradation. Thus, the transcript levels adjust rapidly to changes in P and N, and $\frac{dS}{dt} \approx 0$ in Eq. (1b). The equations become

$$\frac{dP}{dt} = \nu - \alpha_P NP, \tag{3a}$$

$$\frac{dN}{dt} = \frac{\lambda \beta_S}{\alpha_S} \frac{P^{n_S}}{P^{n_S} + k_S^{n_S}} - \alpha_N N. \tag{3b}$$

III. NUMERICAL ALGORITHM AND DATA ANALYSIS

We modeled the stochastic behavior at the single-cell level using the Gillespie algorithm [30]. To quantify the importance of the stochasticity we solved the deterministic equations and compared their trajectories with the stochastic dynamics for each set of chemical parameters in all three models. The deterministic equations always reached a steady state, which is the same in all three models for a given set of parameters. We chose this value as initial condition of the stochastic simulations in order to minimize the initial transient period. Each single stochastic simulation represents a possible cell realization, the equivalent of a single-cell experiment in this description. To simulate the behavior of a random distribution of cells, we computed 300 realizations with different random seeds for each set of chemical parameters. The typical length of each simulation was 2000 min.

For each molecular species X, where X stands for P, N, S, R, and C, we computed the average molecular count $\bar{X} = \langle X(t) \rangle_t$ (averaged across all simulations and over time) and the coefficient of variation, $\mathcal{V}$, defined as the ratio between the standard deviation and the average level,

$$\mathcal{V}(X) = \frac{\langle [X(t) - \bar{X}]^2 \rangle_t^{1/2}}{\bar{X}}. \tag{4}$$

$\mathcal{V}$ significantly increases with the amplitude of pulses of X; therefore, it provides a general signature of the stochasticity and strength of the pulsations for the three models.

To identify the pulsating dynamics in more detail, we computed the normalized autocorrelation function [31], given by

$$C(\tau) = \frac{\langle [X(t) - \bar{X}][X(t+\tau) - \bar{X}] \rangle_t}{\langle [X(t) - \bar{X}]^2 \rangle_t}. \tag{5}$$

The autocorrelation function for each model and parameters was averaged over 300 realizations of the initial conditions. Additionally, we compared the results of the autocorrelation with the power spectrum, defined as $S(k) = \langle F(k) F(-k) \rangle$, where $F(k)$ is the Fourier transform of $X(t)$ and $\langle \cdots \rangle$ denotes average over realizations.

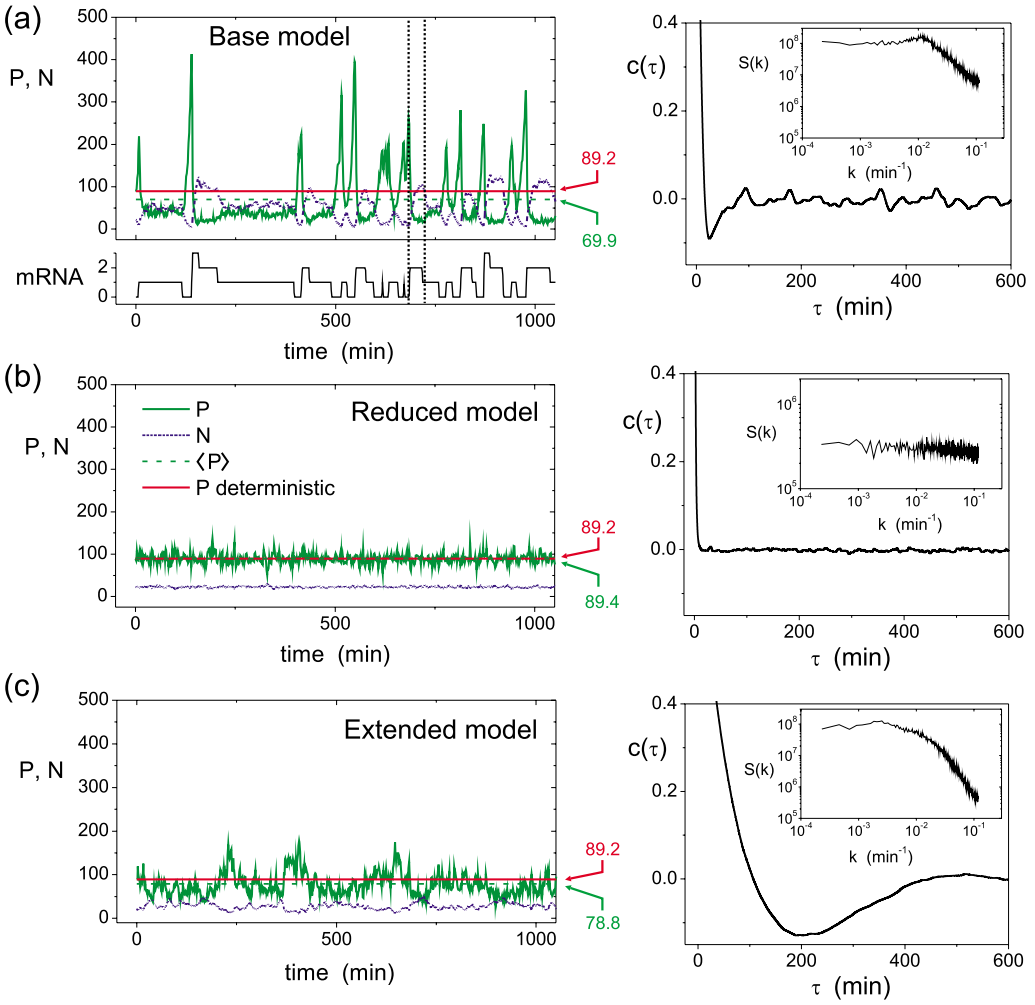

FIG. 2. (Color online) Examples of stochastic simulations in the (a) base, (b) reduced, and (c) extended models. Left panels: copy number of P (thick green line), N (dotted blue line), and mRNA (black line). The average P copy number (dashed green line) is compared to the deterministic trajectory of P (solid red line). For clarity, their average values are indicated on the right. The vertical dashed lines in (a) illustrate the correlation between pulses of N and mRNA. Right panels: detail of the corresponding autocorrelation functions $C(\tau)$ and power spectra $S(k)$ (inset), averaged over 300 runs.

IV. RESULTS

Figure 2 provides a snapshot of the stochastic behavior of the three models. For each model, an example of a typical simulation is shown on the left panels, while the corresponding correlation functions and power spectra (averaged over 300 runs) are depicted on the right panels. Overall, the base model is characterized by a clear excitatory behavior in the positive regulator P. This is confirmed by the peaked correlation function, which provides a characteristic period of the pulses. At the other extreme, the reduced model completely lacks pulses, and the correlation function rapidly decays to zero. Finally, the extended model shows a more complex behavior: the time evolution is characterized by a combination of short pulses and long-term fluctuations. The correlation function has a smooth oscillating shape that corresponds to the long-term fluctuations, and a characteristic period of the pulses cannot be identified. The detailed description of the results for each model, as well as an analysis of the conditions for the presence or lack of excitations, is provided next.

A. Stochastic fluctuations of RNAs can produce pulses of protein expression

We first considered the base network [Eqs. (1)], which contains the positive regulator P, the mRNA S, and the negative regulator N. Figure 2(a) shows the outcome of a typical stochastic simulation. For comparison, the P deterministic value is also shown. The stochastic and deterministic simulations exhibited remarkable differences. While the deterministic concentration was locked in a steady state, the stochastic integration showed a sustained excitatory behavior, where P and N were expressed in a series of anticorrelated pulses. The presence of regular pulses of protein expression in the base model is revealed by the autocorrelation function $C(\tau)$ of P, as shown in the right panel in Fig. 2(a). The autocorrelation presents a pulsating behavior that provides quantitative information about the dynamics of the system. The negative anticorrelation peak indicates the characteristic width of the pulses, around 25 min. The first positive peak provides the characteristic interval between consecutive pulses, around 100 min, and that is in agreement with the maximum in the power spectrum.

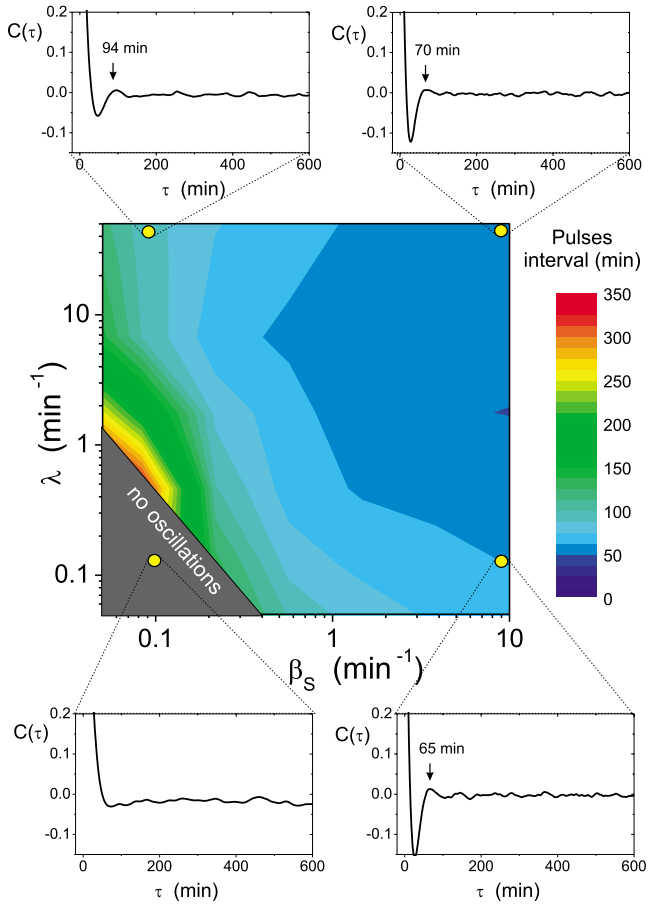

FIG. 3. (Color online) Color map of the characteristic P interpulse interval for the base model as a function of the transcription and translation coefficients (β_S and λ, respectively). A sustained excitatory behavior with a characteristic period appears once minimum translation and transcription rates are attained. The four surrounding plots, whose corresponding (β_S,λ) values are indicated with dots on the color map, illustrate the behavior of the autocorrelation function $C(\tau)$, averaged over 300 runs, for four extreme combinations of transcription and translation rates. Except for very low transcription and translation rates, the autocorrelation $C(\tau)$ shows peaks of correlation that identify the characteristic interpulse interval.

Minimal transcription and translation rates of the negative regulator N were required to identify an excitatory behavior with an observable pattern, as illustrated in Fig. 3. Below minimal values of ν and λ the amount of N was insufficient to implement an effective negative feedback: the pulses of protein expression became irregular, while the average interval between consecutive pulses increased. At extremely low transcription and translation rates the pulses disappeared altogether. Above these extreme conditions we observed that in general a characteristic interpulse period emerges in the base model for all tested sets of parameters (Fig. 3). The autocorrelation shows clear peaks that provide a characteristic period. The average period depends on the transcription and translation rates, and ranges between 30 and 200 min for realistic rates.

The origin of the pulses can be understood by observing the mRNA time trace. As shown in Fig. 2(a), the mRNA is

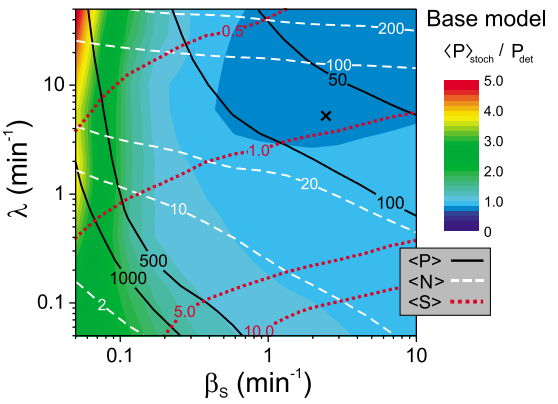

FIG. 4. (Color online) Ratio between the P stochastic level averaged over time and over 300 simulations and the P deterministic steady state in the base model, as a function of the transcription and translation rates. Contour lines show equal stochastic levels of P (solid black line), N (dashed white line), and mRNA (dotted red line). The black cross indicates the location of the example shown in Fig. 2(a).

transcribed in a series of micropulses that are subsequently amplified during translation and inherited by N. The strong correlation between the pulses of mRNA and N is evident in the figure.

We also observed a strong correlation between the mRNA interpulse time lags and the typical width of the pulses of P. As seen in Fig. 2(a), P only reached significant levels after a substantial decay in N prior to a new pulse of N. Therefore, the typical length of the P pulses depends on the decay time of N, which has been taken to be around 5 min [32]. The fast N degradation also explains the small width of P pulses compared to the typical time gap between consecutive pulses.

Finally, we note that the excitatory behavior of the network crucially depends on having a low number of mRNAs. In the base model, the average number of mRNA copies was below 5 for most of the explored parameter space (Figs. 4 and 5). This low average value is in agreement with measurements of transcripts copy number in mammalian cells, where it was found that many transcripts are present in less than one copy per cell on average [33].

1. Comparison with the reduced model: The importance of the mRNA stochastic fluctuations

To further test the impact of the mRNA stochastic dynamics, we considered an alternative circuit design, the reduced model [Eqs. (3)], where the mRNA was assumed to be in quasiequilibrium and therefore was not explicitly included in the circuit. Figure 2(b) shows an example of stochastic simulation run with the same biochemical parameters as the base model in Fig. 2(a). The protein dynamics was considerably different in both models. While the pulses were clearly present in the base model, no trace of pulses could be found in the reduced model. Similar conclusions are reached by comparing the autocorrelation and power spectrum in both models.

As opposed to the base model, intrinsic mRNA fluctuations were not allowed in the reduced model. Hence, the

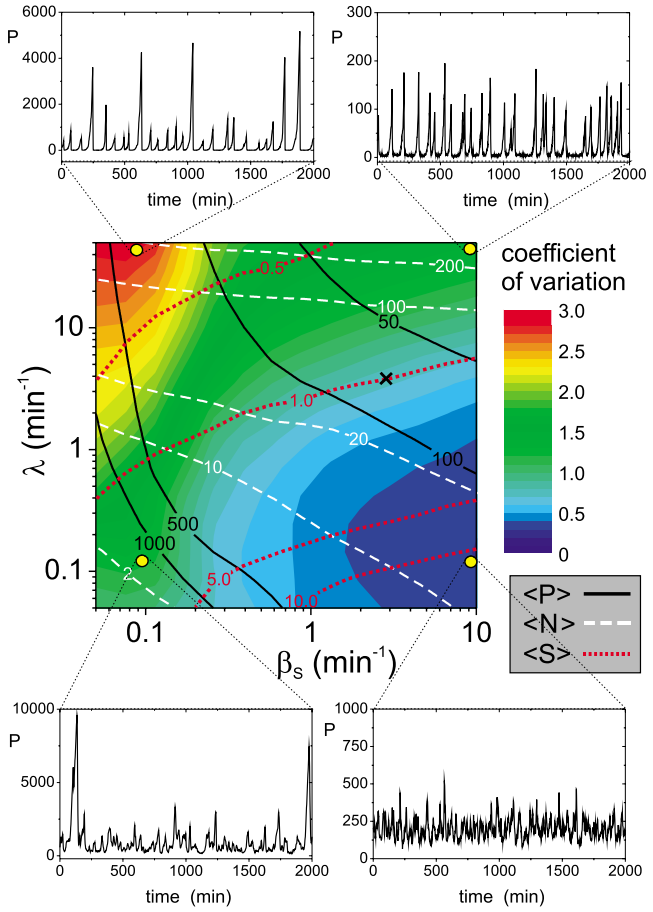

FIG. 5. (Color online) Color map of the coefficient of variation of P in the base model, as a function of the transcription and translation coefficients (β_S and λ, respectively). Contour lines show equal average stochastic levels of P (solid black line), N (dashed white line), and mRNA (dotted red line). P exhibits maximum coefficient of variation at low β_S and high λ. The black cross indicates the position of the example shown in Fig. 2(a). The four plots illustrate the $P(t)$ behavior for four different values of the coefficient of variation, showing that the pulsating behavior of P is maximum at high translation rates λ.

protein levels showed very uniform patterns with only small deviations around the deterministic steady state. The protein levels were also characterized by a very small coefficient of variation. The latter is in agreement with the experimental evidence that, when the contribution of external factor affecting cell-to-cell variability is subtracted, noise in protein expression is dominated by the stochastic production and destruction of mRNAs [2,3]. We stress that both networks were simulated with the same algorithm, kinetic parameters, and initial conditions.

An additional difference between both models was given by the average of the stochastic simulations. The average, which provides an insight into cell population dynamics, converged toward the deterministic trajectory in the reduced model, but not in the base model. The deviation in the base model became especially important at low transcription rates.

Figure 4 shows the ratio of the stochastic P copy number (averaged over time and over 300 simulations) and the deterministic steady state. For transcription rates below 0.1 transcript/min, the deterministic description clearly underestimated the real circuit behavior. The fact that the divergence appeared at low transcription rates, and that it was absent in the reduced model, hints at the mRNA stochastic dynamics as the source of this divergence. This is another interesting example of a biological network presenting a deviation between the stochastic average and the deterministic equations. Such deviations are commonly expected in nonlinear systems and/or systems containing a small number of molecules, although more generic systems can also present large deviations [34].

2. Noise is minimized at high transcription and low translation rates

To achieve a given protein concentration an organism can adopt diverse strategies characterized by different transcription and translation rates. One strategy consists of producing a few mRNA transcripts and translating them efficiently. Alternatively, the organism can transcribe a larger number of transcripts and translate each one of them inefficiently. The former strategy is energetically favorable since a lower number of transcripts have to be produced; however, it was suggested that it may lead to a noisier pattern of protein expression [35].

In order to determine the influence of these two strategies on noise, we ran simulations for different transcription and translation rates in the base model. For each set of parameters we computed the coefficient of variation [Eq. (4)] as shown in Fig. 5. As a visual help we also plotted curves corresponding to equal average levels of P, N, and S. Examples of stochastic simulations at some extreme values are shown in the figure, as well.

Two axes of variation can be roughly identified. Along the main diagonal (from the bottom left to the upper right corner) the average level of N grows as the transcription and translation rates increase. The opposite behavior is seen for P, reflecting the negative feedback loop between both proteins. At the same time, the coefficient of variation of P remains practically constant along this axis. Along the other diagonal the coefficient of variation of P changes from a minimum at high transcription and low translation rates and reaches a maximum at low transcription and high translation rates. This behavior is seen for the coefficient of variation of N as well (data not shown). We also observed that higher levels of noise are obtained when the average number of mRNA transcripts is low, and vice versa (see the dotted red equi-$\langle S \rangle$ lines in Fig. 5), demonstrating the importance of the mRNA fluctuations to the excitatory behavior.

Experimental measurements of protein expression reveal that in many cases the variance in protein levels is roughly proportional to the mean. This trend is usually explained in terms of an increase in the transcription noise with the expression level [2,3]. Yet, some genes are observed to deviate from this scaling law.

Our results show a general agreement with the previous observation; at a fixed transcription rate, the noise increases with the expression level. However, our results also show that noise fundamentally depends on the interplay between

the transcription and translation rates. For instance, along the lines of constant protein levels the coefficient of variation reaches a maximum at low transcription and high translation rates. This suggests that the variability in protein expression depends not only on the average expression level, but also on the ratio of transcription versus translation efficiency. Additional factors such as nonlinear regulatory loops (as in our model system) may also lead to deviations from the scaling rule.

B. Regulatory RNA filter transcription noise

We want to address here the question of how the presence of a regRNA transcript, which targets the mRNA transcript S, modifies the excitatory behavior inherited from the mRNA stochastic fluctuations. Figure 2(c) shows a typical stochastic simulation of the extended model that includes a regRNA, run with the same biochemical parameters as the base model [Fig. 2(a)]. While pulses of protein expression were present in both networks, the extended model [Eqs. (2)] showed broader peaks with significantly lower amplitude relative to the average expression. A comparison of the autocorrelation in both models showed a weaker spiky behavior in the extended model. Similarly, while the power spectrum shared a similar trend with the base model, it lacked the clear maximum present in the latter one.

The deterministic levels of P (red line) are also shown for both models in Fig. 2. While none of the systems was correctly described by the deterministic solution, the departure from the deterministic dynamics was smaller in the model with regRNA. This suggests that, in the presence of a regRNA molecule, the individual cell's dynamics is more robust to transcript fluctuations, hence leading to a reduced cell-to-cell variability.

To verify that the reduced excitability in the extended model is due to the action of the regRNA, we tested the dynamics of the extended model at different regRNA transcription rates and compared it with the base model dynamics. If our hypothesis is correct, the two systems should show differences in their excitatory dynamics as the amount of regRNA increases. Simulations started at a minimum sense transcription rate of 0.1 transcript/min in order to have a sufficient amount of N for the effective repression of P. At low regRNA copy numbers, the coefficient of variation was maximized at low sense transcription rate, in agreement with the qualitative behavior of the base model [Fig. 6(a)]. However, at a high regRNA copy number (above 15 transcripts), the coefficient of variation became independent of the sense transcription rate.

The origin of this difference becomes clear when one observes the time traces of the different RNA transcripts in both models. Figure 6(b) shows the mRNA time trace in the base model, where no regRNA is present in the system. Clear mRNA micropulses with a typical length on the order of the mRNA half-life are observed. These micropulses are amplified during translation, producing the observed pulses of protein expression. An equivalent plot for the extended model [Fig. 6(c)] reveals a completely different dynamics. The regRNA serves as a capacitor: it sequesters sense mRNA when there are copies available and releases them back when there are none. Since the typical binding and unbinding rates are much faster than the RNA half-life, this sequester and release process is repeated many times during a typical mRNA micropulse, effectively erasing memory from any previous mRNA stochastic state. For completeness, Fig. 6(c) also shows the concentrations of sense, regRNA, and sense-regRNA complex. Notice also that the average mRNA concentration is significantly closer to the deterministic trajectory in the extended model, which evidences the reduction in the stochastic fluctuations in the network.

Based on these observations we conclude that the regRNA is able to filter out some of the stochastic fluctuations and induce a smoother protein expression pattern. Such a noise dampening capacity was predicted before in a circuit in which a shared transcription factor regulates a pair of sense and a highly expressed antisense transcript [25].

The filtering capability of the regRNA depends crucially on having fast coupling rates between both RNA strands. Figure 7(a) shows the coefficient of variation of P in the extended model as a function of the sense-regRNA binding and unbinding rates, and for two different values of the regRNA transcription rate.

The lowest coefficient of variation is reached in the panel with a higher regRNA transcription rate. Figure 7(a) also shows curves of equal amount of sense S, regRNA R, and S-R complex. In both panels the minimum coefficient of variation was reached as a trade-off between having fast binding and unbinding dynamics and having the highest number of regRNA molecules (magenta line). We conclude that the noise buffering capability needs both a sufficient number of regRNA transcripts and a fast kinetics to be efficient. This is clearly illustrated in Fig. 7(b), which shows examples of stochastic simulations for different combinations of binding and unbinding rates, and for two extreme values of the regRNA transcription rate β_R. Indeed, there is practically no buffering for low regRNA numbers (left panels), and the time evolution of both P and N shows a pulsating behavior even at large binding dynamics. On the contrary, for high regRNA numbers (right panels), such a pulsating behavior is observed only at very low binding dynamics to rapidly disappear as soon as the binding-unbinding kinetics, and therefore the buffering capacity, increases.

V. DISCUSSION

We have analyzed the response of three different network architectures of a feedback loop and we have observed remarkable differences among them that emerge from the RNA dynamical description. The fluctuations inherited from the transcription process can be amplified during translation and produce pulses of protein expression. The deterministic approach fails to describe such dynamics both at the single-cell level, where no trace of excitatory behavior is observed, and at the cell distribution level, where the deterministic steady state underestimates the protein levels.

Deviations between the stochastic average and the deterministic equations may have different origins [34]. In our case the difference emerges as a result of the nonlinear terms,

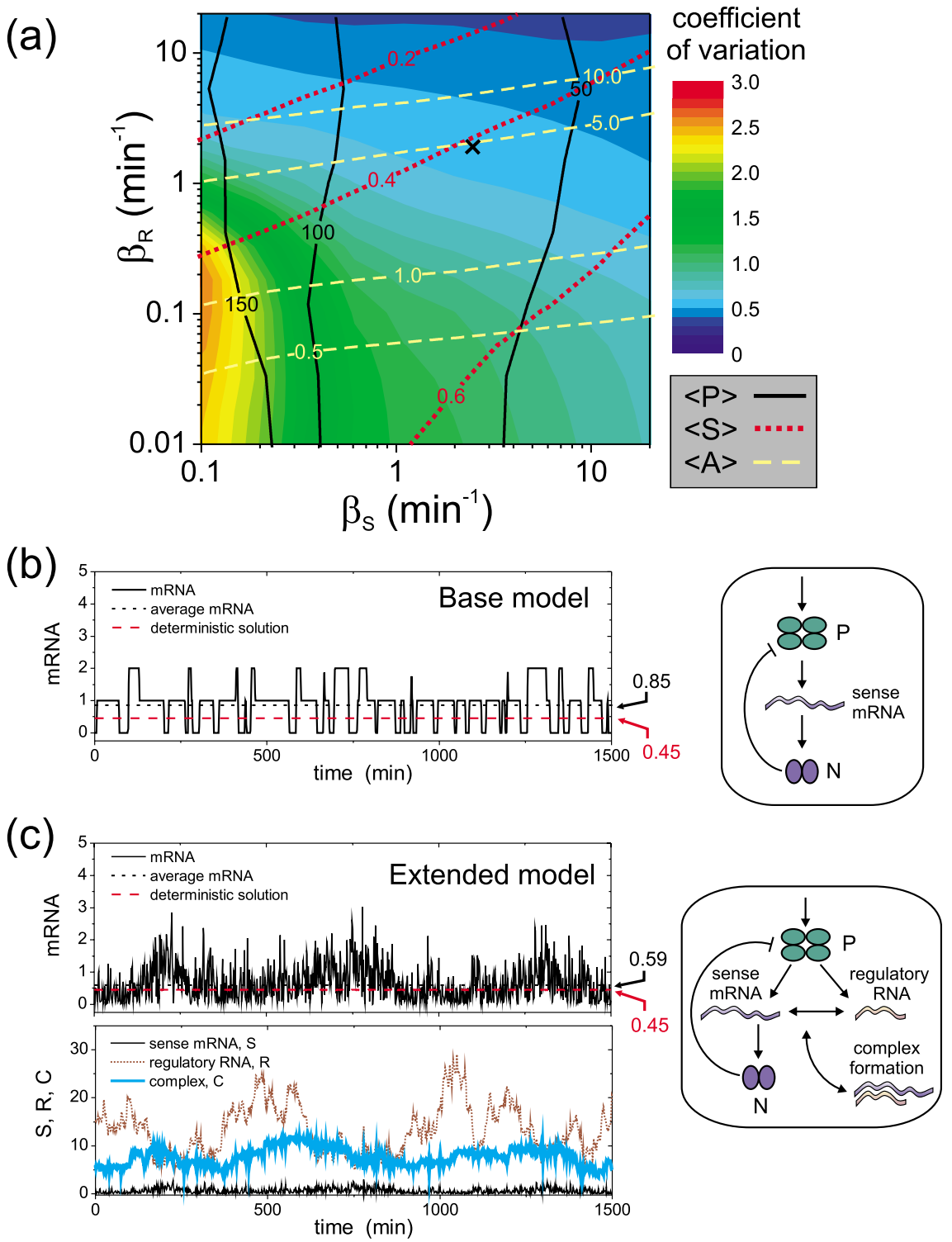

FIG. 6. (Color online) Buffering role of the regRNA in the extended model. (a) Color map of the coefficient of variation of P in the extended model, as a function of the sense and regRNA transcription coefficients (β_S and β_R, respectively), with $\lambda=10$. Contour lines show equal average levels of P (solid black line), mRNA (dotted red line), and regRNA (dashed yellow line). The coefficient of variation is maximized at low transcription rates. The black cross indicates the location of the example shown in Fig. 2(a). (b) Example of the mRNA variation (black) for a stochastic simulation of the base model with $\beta_S=2.4$ and $\lambda=5.0$. The dotted black and the dashed red lines show, respectively, the average mRNA concentration and the mRNA deterministic trajectory. (c) Top: mRNA concentration (black) for a stochastic simulation of the extended model with identical parameters and $\beta_R=2.0$. The mRNA average value is shown with the dotted black line, and the mRNA deterministic trajectory is indicated with the dashed red line. Bottom: concentrations of sense mRNA (S, thin black line), regRNA (R, dotted brown line), and S-R complex (C, thick cyan line). The right panels show schematic representations of the base and extended models.

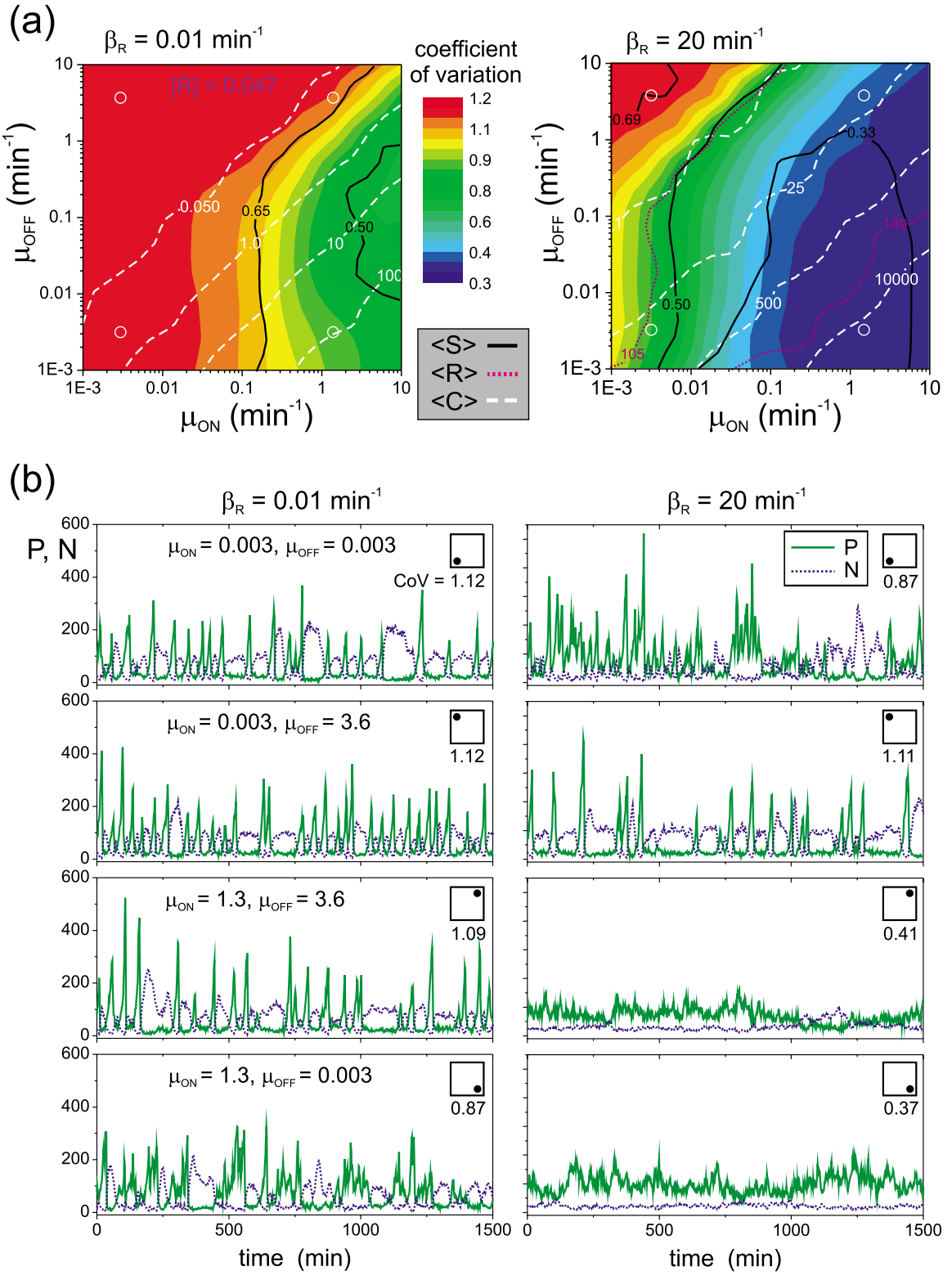

FIG. 7. (Color online) Effect of the binding-unbinding rates on the filtering capability of the regRNA. (a) Color maps of the coefficient of variation of P as a function of the binding and unbinding coefficients (μ_{ON} and μ_{OFF}, respectively) of the sense and regRNA in the extended model. Data are shown for two different regRNA transcription rates β_R. Sense transcription and translation rates are maintained constant. The contour lines show equal average stochastic levels of sense molecules (S, solid black line), regRNA (R, dotted pink line), and complex (C, dashed white line). For β_R=0.01 min^{-1} the number of regRNA is almost constant and only the average value is shown. (b) Examples of P (thick green line) and N (dotted blue line) stochastic simulations for two different values of regRNA transcription rate β_R, and for four different combinations of complex binding (μ_{ON}) and unbinding rate (μ_{OFF}). The sketch in the top right of each panel indicates the location of the examples shown in (a). The value underneath each sketch shows the corresponding coefficient of variation.

which induce correlations that are not present in the deterministic framework. These correlations become particularly important at low mRNA levels, where the largest deviation is observed. Indeed, several works in the literature point at the possibility of having very few transcripts per cell, in some cases, less than 1 on average [36]. At such low copy numbers, our findings stress the importance of using stochastic models to accurately describe the network dynamics.

The excitatory behavior was observed only after simulating the circuit dynamics with the Gillespie algorithm [30]. There are alternative methods to model stochastic fluctuations. The Langevin approach, for instance [37], adds a small stochastic term to the continuous deterministic equations to account for noise. This approach is not suitable in our system because of the strong RNA fluctuations that prevent the characterization of a smooth continuous background. An alternative approach is based on the linear noise approximation of the master equation [38]. It assumes that the system contains a large number of particles and models noise as a continuous linear Gaussian perturbation. This linearized description has been applied to processes involving two molecular species [39]. Opposed to the linear approaches, the Gillespie algorithm provides a simple yet powerful method to obtain stochastic and dynamical solutions compatible with the full master equation. It does not require external assumptions on the noise, the number of molecules and species, or the dynamical regime of the system.

The protein pulses originate in the RNAs stochastic fluctuations. This result is supported by the observed correlation between the typical protein pulse length or frequency and the mRNA fluctuations, as shown in Fig. 2. Additional evidence is provided by the fact that when the RNA intrinsic randomness is neglected, the stochastic behavior converges toward the deterministic concentration. These results are in agreement with experimental evidences that mRNA fluctuations are a fundamental source of noise in protein expression [2,3].

To produce a given average protein concentration the cellular machinery may choose among different transcription and translation rates. Minimized variability among cell population is obtained with a combination of high transcription and low translation rates, as shown in the case of an autoregulated gene in steady state using a linearized stochastic model [18]. Our results generalize this observation. We have analyzed a dynamical stochastic negative feedback loop and found, similarly, that noise is minimized at high transcription and low translation rates, as shown in Fig. 5.

The presence of a noncoding regRNA may help buffering the mRNA fluctuations while allowing the cell to maintain low rates of transcription. In this work we have considered a regRNA that binds the sense transcript and sequesters it from the cellular environment, thus preventing its translation. A regRNA molecule sequesters mRNA when there are some coding transcripts available in the medium and releases them back when there are none. If this process is repeated sufficient number of times during a typical mRNA half-life, the memory of previous states inherited from the transcription process can be erased and, therefore, noise can be partially buffered. In this way, regRNA contributes to reduce the temporal variability at the single-cell level and, consequently, also the cell-to-cell variability. A requisite for this mechanism to work efficiently is fast binding and unbinding (compared to the typical RNA lifetime) between the regulatory RNA and its target mRNA, as shown in Fig. 6. As a result of this buffering, the extended model is less excitable. The circuit still shows pulses, yet compared to the base model where there is no regRNA, the pulses are broader, of smaller amplitude, and they appear at a lower frequency.

A prime example of an oscillating negative feedback loop is provided by the p53-mdm2 regulatory network. This system has been shown to oscillate at the single and cell population levels (sustained oscillations versus damped oscillations, respectively) [8,9]. The maintenance and shape of the oscillations have been linked to two upstream signaling kinases as well as to an additional negative feedback loop [10]. Interestingly however, a single nucleotide polymorphism (SNP309) found in the mdm2 promoter that results in higher levels of mdm2 mRNA and protein [40] has been shown to disrupt the oscillations of p53 and mdm2 proteins [41]. This finding supports the idea that RNAs could play an important role in the oscillatory dynamics in this network, as a high number of mdm2 mRNAs would minimize the importance of the intrinsic stochastic fluctuations, and thus attenuate the pulses of protein expression.

While this work has considered a negative feedback loop, the mechanism that we have described is more general and applies to a wide variety of regulatory networks. We have shown that the RNA dynamics is a fundamental source of intrinsic noise, suggesting that a realistic description of genetic networks requires the stochastic modeling of the transcription stages of protein expression.

ACKNOWLEDGMENTS

We thank Arnold Levine, Yael Aylon, Yolanda Espinosa-Parrilla, Michal Lapidot, and Reut Shalgi for interesting discussions. We especially thank Gil Hornung and Orna Dahan for a critical reading of the manuscript and for raising many interesting observations. We acknowledge the support of the Clore Center for Biological Physics, the Center for Complexity Science, the Minerva Foundation, the European Research Council (ERC) Ideas Grant, and the EC FP7 funding (ONCOMIRS, Grant Agreement No. 201102).

[1] M. Elowitz, A. Levine, E. Siggia, and P. Swain, Science **297**, 1183 (2002).
[2] A. Bar-Even, J. Paulsson, N. Maheshri, M. Carmi, E. O'Shea, Y. Pilpel, and N. Barkai, Nat. Genet. **38**, 636 (2006).
[3] J. Newman, S. Ghaemmaghami, J. Ihmels, D. Breslow, M. Noble, J. DeRisi, and J. Weissman, Nature (London) **441**, 840 (2006).
[4] A. Raj and A. van Oudenaarden, Annu. Rev. Biophys. **38**, 255 (2009).
[5] R. Lev Bar-Or, R. Maya, L. Segel, U. Alon, A. Levine, and M. Oren, Proc. Natl. Acad. Sci. U.S.A. **97**, 11250 (2000).
[6] G. Tiana, M. Jensen, and K. Sneppen, Eur. Phys. J. B **29**, 135 (2002).
[7] D. Nelson *et al.*, Science **306**, 704 (2004).
[8] G. Lahav, N. Rosenfeld, A. Sigal, N. Geva-Zatorsky, A. Levine, M. Elowitz, and U. Alon, Nat. Genet. **36**, 147 (2004).
[9] N. Geva-Zatorsky *et al.*, Mol. Syst. Biol. **2**, 2006.0033 (2006).
[10] E. Batchelor, C. Mock, I. Bhan, A. Loewer, and G. Lahav, Mol. Cell **30**, 277 (2008).
[11] J. Griffith, J. Theor. Biol. **20**, 202 (1968).
[12] M. Elowitz and S. Leibler, Nature (London) **403**, 335 (2000).
[13] L. Ma, J. Wagner, J. Rice, W. Hu, A. Levine, and G. Stolovitzky, Proc. Natl. Acad. Sci. U.S.A. **102**, 14266 (2005).
[14] J. Vilar, H. Kueh, N. Barkai, and S. Leibler, Proc. Natl. Acad. Sci. U.S.A. **99**, 5988 (2002).
[15] A. Ciliberto, B. Novak, and J. Tyson, Cell Cycle **4**, 488 (2005).
[16] T. Zhang, P. Brazhnik, and J. Tyson, Cell Cycle **6**, 85 (2007).
[17] P. François and V. Hakim, Phys. Rev. E **72**, 031908 (2005).
[18] M. Thattai and A. van Oudenaarden, Proc. Natl. Acad. Sci. U.S.A. **98**, 8614 (2001).
[19] A. J. McKane and T. J. Newman, Phys. Rev. Lett. **94**, 218102 (2005).
[20] A. McKane, J. Nagy, T. Newman, and M. Stefanini, J. Stat. Phys. **128**, 165 (2007).
[21] I. Golding, J. Paulsson, S. Zawilski, and E. Cox, Cell **123**, 1025 (2005).
[22] J. Chubb, T. Trcek, S. Shenoy, and R. Singer, Curr. Biol. **16**, 1018 (2006).
[23] A. Raj, C. Peskin, D. Tranchina, D. Vargas, and S. Tyagi, PLoS Biol. **4**, e309 (2006).
[24] D. Bartel, Cell **116**, 281 (2004).
[25] M. Lapidot and Y. Pilpel, EMBO Rep. **7**, 1216 (2006).
[26] S. Cawley *et al.*, Cell **116**, 499 (2004).
[27] R. Shalgi, D. Lieber, M. Oren, and Y. Pilpel, PLOS Comput. Biol. **3**, e131 (2007).
[28] J. Chen, M. Sun, L. Hurst, G. Carmichael, and J. Rowley, Trends Genet. **21**, 326 (2005).
[29] S. Katayama *et al.*, Science **309**, 1564 (2005).
[30] D. Gillespie, J. Phys. Chem. **81**, 2340 (1977).
[31] A. Loinger and O. Biham, Phys. Rev. E **76**, 051917 (2007).
[32] J. Stommel and G. Wahl, EMBO J. **23**, 1547 (2004).
[33] M. Carter, A. Sharov, V. VanBuren, D. Dudekula, C. Carmack, C. Nelson, and M. Ko, Genome Biol. **6**, R61 (2005).
[34] M. Samoilov and A. Arkin, Nat. Biotechnol. **24**, 1235 (2006).
[35] H. Fraser, A. Hirsh, G. Giaever, J. Kumm, and M. Eisen, PLoS Biol. **2**, e137 (2004).
[36] F. Holstege, E. Jennings, J. Wyrick, T. Lee, C. Hengartner, M. Green, T. Golub, E. Lander, and R. Young, Cell **95**, 717 (1998).
[37] C. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*, Springer Series in Synergetics, 3rd ed. (Springer-Verlag, Berlin, 2004).
[38] N. Van Kampen, *Stochastic Processes in Physics and Chemistry*, 3rd ed. (North-Holland, Amsterdam, 2007).
[39] J. Paulsson, Nature (London) **427**, 415 (2004).
[40] G. Bond *et al.*, Cell **119**, 591 (2004).
[41] W. Hu, Z. Feng, L. Ma, J. Wagner, J. Rice, G. Stolovitzky, and A. Levine, Cancer Res. **67**, 2757 (2007).