

## Author's Accepted Manuscript

An open-loop approach to calculate noise-induced transitions

Farzaneh Maleki, Attila Becskei



PII: S0022-5193(16)30422-2  
DOI: <http://dx.doi.org/10.1016/j.jtbi.2016.12.012>  
Reference: YJTBI8895

To appear in: *Journal of Theoretical Biology*

Received date: 25 April 2016  
Revised date: 3 November 2016  
Accepted date: 13 December 2016

Cite this article as: Farzaneh Maleki and Attila Becskei, An open-loop approach to calculate noise-induced transitions, *Journal of Theoretical Biology*, <http://dx.doi.org/10.1016/j.jtbi.2016.12.012>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain

## An open-loop approach to calculate noise-induced transitions

Farzaneh Maleki, Attila Becskei\*

Biozentrum, Computational and systems biology, University of Basel, 4056 Basel, Switzerland

\*Corresponding author. attila.becskei@unibas.ch

### Abstract

Bistability permits the co-existence of two distinct cell fates in a population of genetically identical cells. Noise induced transitions between two fates of a bistable system are difficult to calculate due to the intricate interplay between nonlinear dynamics and noise in bistable positive feedback loops. Here we opened multivariable feedback loops at the slowest variable to obtain the open-loop function and the fluctuations in the open-loop output. By the subsequent reclosing of the loop, we calculated the mean first passage time (MFPT) using the Fokker-Planck equation in good agreement with the exact stochastic simulation. When an external component interacts with a feedback component, it amplifies the extrinsic noise in the loop. Consequently, the open-loop function is shifted and the transition rates between the two states in the closed loop are increased. Despite this shift, the open-loop output reflects the system faithfully to predict the MFPT in the feedback loop. Therefore, the open-loop approach can help theoretical analysis. Furthermore, the measurement of the mean value, variance, and the reaction time-scale of the open-loop output permits the prediction of MFPT simply from experimental data, which underscores the practical value of the stochastic open-loop approach.

### Keywords

Gene regulatory network, linear noise approximation, diffusion, deviant effect, Stratonovich.

### Equation Section (Next)

## 1. Introduction

Transitions between two activity states represent a widespread phenomenon in biology (Biancalani et al., 2014; Pisarchik and Feudel, 2014; Thomas et al., 2014). Bistability in gene

regulatory network is the presence of two stable gene expression or signaling activity states, which commonly underlies alternative cell fates or stores cellular memory of past stimuli (Domingo-Sananes et al., 2015). Bistability in cells typically arises in genetic networks with a positive feedback loop (Kaufman et al., 2007), exemplified by a transcriptional factor (TF) that binds to its own promoter and activates its own expression. Positive feedback loops have been identified in a large number of networks, especially, in those that regulate metabolic utilization of alternative sugars and cell cycle transitions (Park et al., 2012; Pisarchik and Feudel, 2014; Verdugo et al., 2013). Bistable feedback loops may also influence the evolvability of genetic networks (Kuwahara and Soyer, 2012). The quantitative characterization of positive feedback loops can be facilitated by system reduction. An efficient method for system reduction is termed the open-loop approach (Angeli et al., 2004), which can be used to determine whether a feedback loop displays bistability. By opening the feedback loop, a feedback component is broken into a pair of input and output (Fig. 1, left). The mapping of the input to the output defines a one variable open-loop function, which contains all essential information on steady-state bistability without the need to analyze multiple variables. If the open-loop function is sigmoidal then the feedback loop has the potential to display bistability.

This deterministically formulated open-loop approach has important applications (Angeli et al., 2004; Majer et al., 2015; Rai et al., 2012). Experimentally, the open-loop response can be measured to determine whether a multiple-variable feedback gene network can display bistability, without knowing any details on the reactions in the feedback loop. In theoretical analysis, the open-loop approach can be used to determine the maximal bistable range of a parameter in a multidimensional parameter space.

In a deterministically defined bistable system, no transitions occur between two states. However, noise is a distinct feature of genetic networks and can induce transitions between the two states. Noise arises due to extrinsic fluctuations and the low copy number of molecules. Noise in bistable genetic networks plays an important role in transitions between cell fates and cellular adaptation to varying environmental conditions (Bednarz et al., 2014; Guerrero et al., 2016; Li and Wang, 2013; Park et al., 2012; Yuan et al., 2016). In the presence of noise, perturbations can make a bistable system to switch from one steady-state to the other. The average time it takes for a system to switch from one state to the other is the mean first passage time (MFPT), which is a defining characteristic of bistable systems in the presence of noise. The

MFPT is the inverse of the transition rate. The most common method to approximate the MFPT relies on the Fokker-Planck equation, which is typically used for one-dimensional systems. Since most biologically realistic feedback loops contain multiple variables, system reduction to a single variable is desirable.

To reduce the system and to approximate the MFPT, we aimed to formulate a stochastic version of the open-loop approach. For this purpose, we used truncations of the master equation, which describes chemical reactions stochastically. The linear noise approximation (LNA) (Kampen, 1992) is a leading order term truncation, and was employed to calculate the mean and variance of the open-loop output. The nonlinear Fokker-Planck equation (FPE) is a truncation after two terms, and was used to calculate the MFPT.

Using LNA, we calculated the fluctuations transmitted from all the system's components to the open-loop output (Fig. 1, right). Combining this with the open-loop function and the reaction time-scale of the output, we predicted the MFPT. The fluctuations in the nonlinear Fokker-Planck equation can be interpreted in the Ito and Stratonovich sense. The Ito integral implies that there is no dependence between the process (chemical reaction) and the fluctuation (Horsthemke and Lefever, 1984). This is a reasonable assumption for an idealized, white noise, which is typically applicable for intrinsic noise (Lei, 2009), which arises due to the low copy number of molecules. When extrinsic noise is present, the existence of a finite memory in the extrinsic noise has to be taken into account. Due to this memory, there is a dependence between the process and the random fluctuation. The Stratonovich integral takes into account such a correlation between the process and fluctuations. We applied the Stratonovich interpretation when the intensity of the extrinsic noise was significant. This improved the prediction of the MFPT.

**Equation Section (Next)**

## 2. Results

### 2.1 Characterization of bistability by the deterministic open-loop approach

The classical open-loop approach, which is formulated deterministically, consists of two steps: opening and reclosing the feedback loop (Fig. 1). To open the feedback loop, a component has to be broken into an input and output. The open-loop function,  $\eta = f(\omega)$ , assigns an output  $\eta$  to each

input  $\omega$ . Intersections of  $\eta = f(\omega)$  with the identity function  $f_i(\omega) = \omega$ , (black line, Fig. 2b), represent the steady states of the closed-loop system (Angeli et al., 2004). When there are three intersections, the system is bistable (blue curve, Fig. 2b). The upper and lower stable steady states are denoted by “ON” and “OFF” states.

First, we analyzed two versions of a simple feedback loop with one TF, which binds its own promoter cooperatively and activates its own expression. In the one-variable system (Fig. 2a), we neglect the encoding RNA, while we consider it explicitly in the two-variable version (Fig. 3a). Here we present the equations for the two-variable system ( $R$  for RNA and  $P$  for protein). The closed feedback system is described by the following ordinary differential equations (ODE):

$$\begin{aligned}\frac{d}{dt}R &= b + v_m \frac{P^n}{K_d^n + P^n} - \gamma_R R, \\ \frac{d}{dt}P &= \lambda R - \gamma_P P.\end{aligned}\tag{2.1}$$

$K_d$  denotes the equilibrium dissociation constant,  $b$  and  $v_m$  are the basal and maximal transcription rates and  $\lambda$  is the translation rate constant.  $\gamma_R$  and  $\gamma_P$  stand for the decay rate constants of the RNA and protein, respectively. The index of cooperativity, the Hill number, is denoted by  $n$ , and it was taken to be two for all calculations in order to generate a sigmoidal, ultrasensitive, open-loop function. In the example below, the parameters do not have physical dimension but their values are representative for yeast, or similar microorganism, when the time and concentration units are expressed in hours and nM (Bonde et al., 2014; Majer et al., 2015).

By opening the feedback loop at the level of protein, the resulting input,  $\omega$ , will be a stimulus that modulates the activity of the promoter. The protein  $P$  produced under the control of the promoter is the output, which is unable to bind to the promoter (Figs. 2a, 3a).

$$\begin{aligned}\frac{d}{dt}R &= b + v_m \frac{\omega^n}{K_d^n + \omega^n} - \gamma_R R, \\ \frac{d}{dt}P &= \lambda R - \gamma_P P.\end{aligned}\tag{2.2}$$

The open-loop function is given by

$$f(\omega) = \frac{\lambda}{\gamma_P \gamma_R} \left( b + v_m \frac{\omega^2}{K_d^2 + \omega^2} \right). \quad (2.3)$$

Applying the conditions of the fold bifurcation on the open-loop functions (Majer et al., 2015), we obtained the bistable range of the TF-DNA affinities,  $K_d$ , for the reclosed loop (Fig. 2a). The values for the translation and mRNA decay rates were chosen so that the open-loop functions of the one- and two-variable versions are identical. Consequently, the bistable ranges of  $K_d$  are identical in the feedback loops, which are plotted as the x-axis in Figs. 2d, 3c.

## 2.2 Calculation of drift and diffusion by the open-loop approach

To predict MFPT in the bistable range of the feedback loop system, two functions have to be specified for the reduced system: the drift and diffusion. The drift specifies the deterministic trend, which is zero when the reduced system is in the steady-state. The diffusion represents a random force which introduces fluctuations into a variable (Fig. 1).

To obtain the diffusion, we used linear noise approximation (LNA) (see Appendix 4.2.). It is important to emphasize that the LNA is applied on the open-loop system. For closed-loop bistable systems, LNA can be only used locally, around the steady-states and not globally.

The results of LNA were verified by the stochastic simulation of the chemical reaction network, for which we used the numerically exact stochastic simulation algorithm (SSA) (Gillespie, 1977). The steady-state mean values of all variables obtained by LNA are identical to their deterministic values (see Appendix 4.1 and 4.3). The mean and variance of the output calculated by LNA ((4.2) and (4.3)) in both one and two variable systems (Fig. 2c, 3b) matched the values obtained by SSA for the entire range of input values. This implies that the mean and variance of the open-loop output were approximated accurately by LNA, and we proceeded to calculate the diffusion.

The open-loop output is affected directly by intrinsic fluctuations and also indirectly by the fluctuations of all components in the loop that propagated to the output known as extrinsic fluctuations. Therefore, the total diffusion, which is the summation of intrinsic and all extrinsic diffusions, has to be re-defined for the single-variable open-loop output that reproduces the effect of fluctuations due to all components in the original system.

The open-loop function reduces a multivariable system to a one-variable function. Similarly, the diffusion matrix has to be reduced. The diffusion for the protein ( $B_{2,2}$ ) in (4.3), represents only the intrinsic diffusion for the protein. By reducing the system to a one-variable system ( $\bar{P}$ ), the total diffusion can be obtained using equation (4.5).

$$D_{total} = \underbrace{\frac{\lambda}{\gamma_R} \left( b + v_m \frac{\omega^2}{K_d^2 + \omega^2} \right)}_{D_{Intrinsic}} + \gamma_P \bar{P} + 2 \underbrace{\frac{\lambda^2}{\gamma_R^2 (\gamma_R + \gamma_P)} \left( b + v_m \frac{\omega^2}{K_d^2 + \omega^2} \right)}_{D_{Extrinsic}}. \quad (2.4)$$

In which  $D_{Intrinsic}$  and  $D_{Extrinsic}$  indicate the intrinsic and extrinsic diffusions respectively.

The total diffusion for the one-variable system (Fig. 2c) is identical to the intrinsic diffusion of the protein in (2.4).

The drift,  $H(\omega, \bar{P})$ , corresponds to the dynamics of protein in the reduced system obtained from (2.2):

$$H(\omega, \bar{P}) = \frac{\lambda}{\gamma_R} \left( b + v_m \frac{\omega^2}{K_d^2 + \omega^2} \right) - \gamma_P \bar{P}. \quad (2.5)$$

As discussed below (see section 2.6), the drift is a function of the mean open-loop output and the time-scale of the output (decay rate), (Fig. 1, right).

### 2.3 Calculation of MFPT upon reclosing the loop

We re-closed the loop, by replacing  $\omega$  by  $\bar{P}$  in (2.4) and (2.5). To predict the MFPT, we used the adjoint operator of the Fokker-Planck equation (FPE) (see (4.8)).

The MFPTs were calculated for the entire bistable range of  $K_d$ . The OFF - ON transitions become slower as by  $K_d$ , increases (Fig. 2d, 3c, left). The MFPTs were smaller for the two-variable system than for the one-variable system. In both systems, the bistable ranges of  $K_d$  are identical and the protein concentration varies between 5 and 500 nM. In the two-variable system, low RNA concentration (ranging between 0.5 and 50 nM) generates a stronger noise, and consequently, the stronger noise induces more frequent transitions. This also explains why the OFF - ON transitions are faster than the ON - OFF transitions (Fig. 2d, 3c, right).

To validate the predicted MFPTs we compared them to the MFPT obtained by the SSA. We simulated the system at least up to  $10^5$  hours, i.e. 11 years, which can be considered as an upper limit of the duration of a realistic experiment. We obtained a very good agreement between the calculated and simulated values. We have also checked the prediction of MFPT using Kramers formula (see (4.14)), which is a popular approximation of the FPE and can provide analytical insight on the determinants of the MFPT (Escudero and Kamenev, 2009; Hwang and Velazquez, 2013). The approximation was in good agreement with the SSA (Fig. 2d, 3c).

While opening can be performed at any variable (RNA or protein) for the deterministic characterization, this is not the case for stochastic open-loop approach. In the above two-variable system, the fast variable is the RNA. Therefore, it was appropriate to open the loop at the protein (see equation (2.2)).

#### 2.4 Opening of the loop at the slowest variable and the Stratonovich interpretations are necessary to accurate prediction of the MFPT

To see how the calculations can be extended to multi-variable feedback loops, we explored a double activator feedback system, in which two TFs activate each other's transcription. One of the TFs ( $P_1$ ) has a five-time faster decay rate than the other ( $P_2$ ) (Fig. 4a). The translation rate constants were adjusted so that the maximum concentrations for both TFs were equal. Since the RNAs encoding the two different TFs were explicitly modelled, we obtained a four-variable system.

$$\begin{aligned}
 \frac{d}{dt} R_1 &= b + v_m \frac{P_2^n}{K_d^n + P_2^n} - \gamma_{R_1} R_1, \\
 \frac{d}{dt} P_1 &= \lambda_1 R_1 - \gamma_{P_1} P_1, \\
 \frac{d}{dt} R_2 &= b + v_m \frac{P_1^n}{K_d^n + P_1^n} - \gamma_{R_2} R_2, \\
 \frac{d}{dt} P_2 &= \lambda_2 R_2 - \gamma_{P_2} P_2.
 \end{aligned} \tag{2.6}$$

New aspects have to be dealt with in order to calculate the MFPT for this system: the site of opening and the level of the extrinsic noise due to the large number of components in this loop, as evidenced by the multiple terms in the extrinsic diffusion (4.7).

To illustrate the importance of the choice where to open the loop, we first opened the feedback loop at the protein  $P_1$  with the fast decay rate. In this case, the predicted MFPT deviated from the SSA results (Fig. 4c). When we opened the feedback loop at  $P_2$ , the protein with the slow decay, the predictions improved significantly and were in good agreement with the SSA results. This confirms the expectation that the feedback loop has to be opened at the slowest variable. It is important to note that the simulated mean values of the open-loop output closely match the analytically calculated mean values for both openings (Fig. 4b).

The transitions in the two-TF system were generally slower in comparison to the feedback loop with a single TF (compare Figs. 3 and 4) because the four-variable system contains two promoters and the binding to both promoters is cooperative ( $n = 2$ ).

There is a small but consistent difference between the predicted and the exactly simulated MFPT values (Fig. 4c). We assumed that this discrepancy arises due to the larger number of variables between the input and output, which generates a more considerable extrinsic noise transmitted to the output. As mentioned in the introduction, the Stratonovich interpretation is more realistic to model the extrinsic noise than the Ito interpretation. So far, we have used the Ito interpretation to construct the FPE, which implies that the drift and diffusion are independent and the fluctuations are uncorrelated white noises (Appendix 4.3). The drift becomes dependent on extrinsic fluctuations due to the existence of nonlinearities. This dependency is shown mathematically by the Stratonovich interpretation and can shift the level of expectation values (Horsthemke and Lefever, 1984). The drift in the Stratonovich interpretation is extended from the drift in the Ito interpretation by the derivative of the external diffusion term with respect to the system variable (see also Appendix 4.3). If we open the feedback loop at  $P_i$ , the term of drift in the Stratonovich interpretation will be given by

$$H_S(\omega, P_i) = H_I(\omega, P_i) + \frac{1}{4} \frac{d}{dP_k} D_{Extrinsic}. \quad (2.7)$$

where  $H_I$  is the drift calculated by Ito interpretation while  $H_S$  represent the Stratonovich drift.

The values calculated with the Stratonovich interpretation were in a better agreement with the simulated values in comparison to the Ito-interpretation (Fig. 4c).

## 2.5. The effect of an external component with multiplicative effect on the feedback loop

With respect to stochastic control, it is particularly interesting to examine how extrinsic components affect the transitions in the feedback loop. For this purpose, we explicitly modelled a proteolytic enzyme that degrades the TF ( $P$  in (2.8)). Thus, the degradation rate of the TF is multiplied by the concentration of an enzyme ( $E$ ). The enzyme is modelled by a birth-death process (Fig. 5a) and displays a Poisson-distribution. The deterministic system is described by the following equations after the opening:

$$\begin{aligned} \frac{d}{dt} E &= \beta - \gamma_E E, \\ \frac{d}{dt} P &= b + v_m \frac{\omega^2}{K_d^2 + \omega^2} - \gamma_P EP. \end{aligned} \quad (2.8)$$

Despite having only two stochastic variables in the open-loop, the calculated MFPT, using the Ito interpretation, deviated considerably from the values obtained by SSA (Fig. 5c, orange curve). Furthermore, the mean value of the open-loop output deviated slightly but consistently from the SSA results (Fig. 5b). This is not fully surprising because it has been known that additional factors like non-zero cross-correlations between the fluctuations have to be taken into account in the presence of nonlinear reactions. A large number of studies have explored the effect of cross-correlation on the transitions (Fuliński and Telejko, 1991; Mei et al., 1999). Despite the strong effect of such correlations on the transitions, no rigorous methods have been proposed to derive such cross-correlations. The nonlinear interaction between the protein and enzyme indicates that the Stratonovich interpretation should be used. Such a nonlinear reaction results in a non-zero cross-correlation between the protein and enzyme, which adds an additional term to the total diffusion of  $\bar{P}$  in the reduced system obtained with uncorrelated noises (4.5).

$$D_{total} = D_{Intrinsic} + \underbrace{D_{Extrinsic} + 2\alpha \sqrt{D_{Intrinsic} D_{Extrinsic}}}_{D_H}. \quad (2.9)$$

In which  $D_{Intrinsic}$  and  $D_{Extrinsic}$  indicate the intrinsic and extrinsic diffusions respectively. To specify the coefficient of the additional term ( $\alpha$ ), we decomposed it into a correlation coefficient between the enzyme and the TF and a normalization term:

$$\alpha = N_C \rho_{EP}. \quad (2.10)$$

The correlation coefficient for the enzyme and protein was calculated using LNA (4.2):

$$\rho_{EP} = \frac{C_{EP}}{\sqrt{C_{EE}C_{PP}}} = \frac{-\frac{\gamma_P E_s \bar{P}}{(\gamma_P E_s + \gamma_E)}}{\sqrt{\frac{1}{\gamma_P} \left( \frac{(\gamma_P \bar{P})^2 E_s}{(\gamma_P E_s + \gamma_E)} + \left( b + v_m \frac{\omega^2}{K_d^2 + \omega^2} \right) \right)}}. \quad (2.11)$$

The  $E_s$  represents the steady-state enzyme concentration. The normalized coefficient  $N_C$ , is obtained as the ratio between the extrinsic ( $\sigma_{EP}^2$ ) and intrinsic ( $\sigma_{P,Int}^2$ ) variances of protein normalized by the variances of the corresponding sources of fluctuation:

$$N_C = \frac{1}{\sqrt{2}} \sqrt{\frac{\frac{\sigma_{EP}^2}{\sigma_E^2}}{\frac{\sigma_{P,Int}^2}{\sigma_U^2}}} = \sqrt{\frac{(\gamma_P \bar{P})^2}{2(\gamma_P E_s + \gamma_E) \left( b + v_m \frac{\omega^2}{K_d^2 + \omega^2} \right)}}. \quad (2.12)$$

where  $C_{PP} = \sigma_{EP}^2 + \sigma_{P,Int}^2$  and  $C_{EE} = \sigma_E^2$ . Thus,  $N_C$ , and consequently the diffusion, becomes large when the normalized external fluctuation significantly exceeds the normalized internal fluctuation.  $\sigma_U = 1$  in our models since the order of fluctuation is  $1/\sqrt{\Omega}$  and we take  $\Omega = 1$  (see Appendix 4.5 Stochastic simulation algorithm).

Substituting (2.11) and (2.12) into (2.10), we obtain the total diffusion:

$$D_{total} = \underbrace{b + v_m \frac{\omega^2}{K_d^2 + \omega^2} + \gamma_P E_s \bar{P}}_{D_{Intrinsic}} + \underbrace{\frac{2(\gamma_P \bar{P})^2 E_s}{(\gamma_E + \gamma_P E_s)} - 2\rho_{EP} N_C \gamma_P \bar{P} \sqrt{\frac{2E_s}{(\gamma_E + \gamma_P E_s)} \left( b + v_m \frac{\omega^2}{K_d^2 + \omega^2} + \gamma_P E_s \bar{P} \right)}}_{D_H}. \quad (2.13)$$

Since the nonlinear interaction between protein and enzyme, the Stratonovich interpretation is applied to obtain the term of drift

$$H_S(\omega, \bar{P}) = H_I(\omega, \bar{P}) + \frac{1}{4} \frac{d}{d\bar{P}} D_H = b + v_m \frac{\omega^2}{K_d^2 + \omega^2} - \gamma_P \frac{\alpha}{\gamma_E} \bar{P} + \frac{1}{4} \frac{d}{d\bar{P}} D_H. \quad (2.14)$$

where  $H_I$  and  $H_S$  represent the Ito and Stratonovich drifts respectively. This noise-induced shift was in good agreement with the result of the stochastic simulation algorithm of the open-loop system (2.8) (Fig. 5b). Importantly, the increased variance in the open-loop output was also explained. Lastly, we verified the MFPT calculated with the diffusion term (2.13) that takes the

correlation into account. The prediction improved considerably, as revealed by comparison to the SSA (Fig. 5c). We also checked the goodness of prediction by varying the strength of the extrinsic noise introduced by the enzyme. By varying the birth-rate of the enzyme we explored a range of noise intensities of the enzyme,  $CV_E$ , up to the coefficient of variation of 0.5 (50%) (Fig. 6), which represents a strong noise in realistic biological systems (Newman et al., 2006). We calculated the mean and variance of the open-loop output at a particular value of the input in the Stratonovich interpretation. As the noise in the enzyme increased, the noise-induced shift in the mean and variance of the output also increased (Fig. 6a). The prediction by FPE taking into account the correlation was in good agreement with the SSA. The predicted MFPT decreased with increasing noise in the enzyme. The prediction was in good agreement with the simulations in the entire range of noise intensities (Fig. 6b). In summary, taking into account the noise-induced shift and correlations between the fluctuations improved the prediction substantially.

## 2.6. Experimental applications of the open-loop approach

The above results indicate that the drift and diffusion of the open-loop output faithfully reflect the stochastic system dynamics and can be used to calculate the MFPT even when the open-loop function is shifted by noise in an external component.

Can the stochastic open-loop approach be used to predict MFPT from experimentally measured open-loop response? Neither the drift nor the diffusion can be measured directly. The diffusion represents an abstract random force that introduces fluctuations into a system variable. However, the following relations provide link between them and measurable functions and parameters. The diffusion can be obtained from the measurable variance of the output (4.2), when the Jacobian of the output,  $J$ , is known.  $J = -\gamma_p$ .

$$D_{total}(\omega) = 2\gamma_p C_p(\omega, \eta) \quad (2.15)$$

and the drift can be obtained directly from the

$$H(\omega) = \gamma_p (f(\omega) - \eta). \quad (2.16)$$

The reaction time-scale of the output (Fig. 1, right panel) corresponds to the decay rate constant of the protein,  $J = -\gamma_p$ . It is the only physical parameter that has to be determined experimentally. Subsequently, the mean ( $f(\omega)$ ) and the variance ( $C_p(\omega)$ ) of the output as a function of input has

to be measured. Phenomenological functions can be then fitted to the input – output data without the need to identify specific physical parameters. Having determined the values of the parameters in the equations (2.15) and (2.16), we can proceed to calculate the MFPT.

To assess the practical feasibility of this approach, we simulated the mean and variance of the open-loop output derived from the above single-variable feedback loop (Fig. 7a), and added a normally distributed random variable to represent the measurement error. We have chosen two  $K_d$  values that are positioned close to the lower and higher end of the bistable range. We fitted then the appropriate functions to this simulated experimental data;  $f(\omega)$  was used to fit the simulated mean (4.15). To fit the variance  $C_p(\omega)$ , we used the first two terms in the function (2.13), without the correlation term. All the physical parameters were lumped into three parameters for the fitting ( $\alpha_1$  to  $\alpha_3$  in (4.16)). We assumed that the omission of the correlation terms will be compensated by the fitting of these three parameters. Indeed, we obtained a good fits for the mean and variance of the open-loop output (Fig. 7b). Using the fitted functions, we obtained the drift and diffusion, and calculated the MFPT using the Stratonovich interpretation (2.7). The MFPT predicted in this way was in good agreement with the expected values (Fig. 7c).

Next, we tested if the same equations can be used to predict the MFPT for a somewhat different system: a feedback loop consisting of two variables (RNA and protein) and the extrinsic noise is introduced by a ribosome, which translates the RNA into the protein. The ribosome itself is modeled by a birth-death process (Fig. 8a). The nonlinearities, i.e. the determinants of the bistability in the two systems (Fig. 7a and 8a), were identical. For this reason, we used the same functions for the fitting of the open-loop mean and variance. We expected that the fitting of the parameters in (4.16) will compensate the differences between the two systems. Indeed, we obtained a good fit to the open-loop functions simulated with the experimental error (Fig. 8b). The subsequently predicted MFPT was also in a good agreement with the expected value (Fig. 8c). This is important since it shows that relatively simple functions can be used to fit the open-loop mean and variance without the need to know the details of the system, which is a common impediment in realistic biological systems.

~~Equation Section (Next)~~

### 3. Discussion

The rate of transitions has two major determinants: the deterministic processes (drift) and random fluctuations (diffusion). They often originate in distinct biochemical processes. The drift is dominated by ultrasensitive reactions, such as cooperative binding and sequestration (Hsu et al., 2016b; Majer et al., 2015). Such reactions make the potential barrier between the two stable steady states higher, which slow down the transitions (Hsu et al., 2016a). On the other hand, noise is largely determined by copy number and extrinsic fluctuations (Rybakova et al., 2015; Vinuelas et al., 2013), which can also have dramatic effect on the MFPT (Horsthemke and Lefever, 1984; Hsu et al., 2012; Jaruszewicz et al., 2013). Thus, the Fokker-Planck equation (FPE) is in principle an ideal tool to calculate the MFPT since noise (and hence diffusion) and the deterministic drift are often measured separately.

The applicability of the FPE is however limited because it can be conveniently used only for one-variable systems, and, most feedback loops are multi-variable systems. For this reason, we employed linear the noise approximation (LNA), which has been extensively used to calculate noise in multi-variable systems (Paulsson, 2004). However, the LNA can be used to predict only features locally around the steady-states of the bistable system and not to predict global features (Scott, 2011). On the other hand, transitions are influenced by fluctuations throughout the passage between the two states and not only around the stable steady-states. The conversion of the closed feedback loop into an open loop obviates this problem since the resulting monostable system can be fully described by LNA, at the entire biologically or physically relevant range of open-loop input and output values.

In the first part of this work, we predicted the MFPT using a stochastic open-loop approach. Three aspects had to be taken into account: opening at the slow variable, use of the Stratonovich interpretation and specifying the correlation of multiplicatively interacting components. Afterward, we have shown that simply fitting functions to the mean and variance of an experimentally realistic open-loop input and output data permits the prediction of MFPT.

We applied LNA to a multi-dimensional system in an open-loop setting and used a steady-state assumption to reduce the system to the slowest variable to obtain the drift and the total diffusion. Having the total diffusion with the drift, and reclosing the loop, we approximated the MFPT for feedback loops with multiple, serially arranged, components, in good agreement with the result of exact stochastic simulation of the reactions (Figs. 2, 3, 4).

The accurate approximation of the mean and variance in the open-loop output is an important but not sufficient condition for a good prediction of the MFPT. Only the opening at the slowest variable warrants an accurate prediction of MFPT. This is an important difference when loop opening is used for the deterministic characterization of the system; the site of the opening does not make a difference for the deterministic open-loop approach. The opening at the slowest variable of the four-variable system resulted in a good prediction of the MFPT (Fig. 4). It is also important to note that a good prediction is generally expected when the system can be separated into slow and fast variables with clearly separated time-scales.

When an external component interacts with a feedback component then the multiplicative effect of two variables introduces correlated fluctuations, which amplify the extrinsic noise. In this case, the system description with uncorrelated white noise in the Ito interpretation does not predict correctly the system behavior. This is not surprising because it has been known that for second-order reactions, the FPE with independent white noise is not accurate (Grima et al., 2011; Thomas et al., 2013). Different approaches have been introduced to resolve the discrepancy, exemplified by the colored noise approximations (Shahrezaei et al., 2008).

In this work, we included correlation between the two components and used the Stratonovich method. This change explained both the noise-induced shift of the open-loop function and the excess variance in the output. Furthermore, the predicted MFPTs were accurate at a broad range of realistic values of noise in the external component. The noise-induced shift, also known as a deviant effect, has been observed experimentally in promoter responses (Cai et al., 2008; Samoilov and Arkin, 2006). It was observed as a change in the mean value of the promoter response as the fluctuations were tuned. Thus, it is important that theoretical models properly describe the noise-induced shift. Importantly, even when the mean value of the output of the open-loop function is shifted by noise, the joint characteristics of the open-loop function, i.e. the drift and the diffusion, correctly reproduce the stochastic system dynamics to predict MFPT.

This has important practical implications. Firstly, it reveals that an experimentally opened feedback loop delivers the information for modelling faithfully, even when the open-loop function is shifted. Secondly, the loop opening provides an opportunity to predict MFPT from measuring only the characteristics of the output. Knowing the half-life of the protein, the mean and the variance of the output as a function of the input carries all the information to predict the MFPT. In principle, an appropriate function can be fitted to the measured open-loop mean and variance to obtain the drift and diffusion terms. Their ratio can be used to directly determine the MFPT, without the need to identify the physical parameter in the system. This provides an opportunity to predict MFPTs robustly using open-loop measurements.

There are in principle two different approaches to predict MFPT following experimental loop opening. If a feedback loop is opened at a component with a fast reaction time-scale, then the mean and the variance of the open-loop output is not informative. Consequently, the most important reactions steps between the input and output have to be characterized. For example, if the feedback loop is opened at the RNA, the turnover of which is typically faster than the turnover of a protein, then the time-scale of the protein decay and promoter noise have to be quantified. Afterward, the open-loop has to be modelled by a chemical master equations, typically by the exact numerical simulation (SSA), and ultimately the MFPT can be predicted (Hsu et al., 2016a).

On the other hand, the opening at the component with the slow-time scale (e.g. protein) permits the direct prediction of the MFPT simply by measuring and fitting the mean and variance of the open-loop output, as presented in this work. Performing time-series measurement of the output fluctuations may lead for further improvements in the prediction because it can be used to obtain the information about the nature of fluctuations and to refine the drift term (Erkal and Cecen, 2014; Pesce et al., 2013). Such an approach may be particularly useful when nonstationary deterministic processes, such as oscillations, and noise, jointly drive the transitions between the two states, which can be the case in the cell-cycle (Steuer, 2004). Therefore, we expect that further developments in the stochastic description of open-loop systems will further facilitate the understanding and prediction of transitions in bistable systems.

**Equation Section (Next)**

(4.1)

## Appendix.

### 4.1. The linear noise approximation (LNA)

The Fokker-Planck equation (FPE) is an approximation of the master equation for chemical reactions with nonlinear kinetics (Kampen, 1992). The basic assumption in FPE is to have high number of molecules  $N$  in a large volume  $\Omega$ . Linearization of the FPE results in LNA.

By repeating a stochastic process many times, we will observe some repeated averaged behaviors which are roughly in agreement by the deterministic description plus some fluctuations. This indicates the LNA assumption of the separation of the deterministic part ( $\bar{X}$ ) of a stochastic variable ( $X$ ) from the fluctuation ( $\varphi$ ), which is proportional to the square root of the system volume:

$$X = \bar{X} + \frac{\varphi}{\sqrt{\Omega}}, \quad (4.1)$$

Thus, the linear noise approximation (LNA) is the leading order term of the expansion of the master equation assuming small noise, which is equivalent to the linear FPE. The dynamics of the mean ( $X$ ) and covariance ( $C$ ) of the system is described by:

$$\begin{aligned} \frac{d}{dt} X &= F(X), \\ \frac{d}{dt} C &= J C + (J C)^T + B. \end{aligned} \quad (4.2)$$

where  $J = \frac{d}{dx} F(X)$  is the Jacobian matrix. The diffusion matrix  $B$  is given by

$$B = S \text{Diag}(V) S^T \quad (4.3)$$

where  $S$  is stoichiometric matrix and  $V$  is the reaction rate vector of the system,  $F(X) = S V$ .

### 4.2. Calculation of diffusion after system reduction to the slowest variable

In biological networks, some components have considerably shorter life times compared to the others. By eliminating the fast variables, we can focus on the dynamics of the slow variable, which reduces the complexity of the system (Thomas et al., 2012).

We open the system at the slowest variable in the multivariate system, and calculate the steady-state covariance for the output of the open-loop, i.e. the slowest variable ( $C_{ss}$ ) using the Lyapunov equation (4.2).

In the next step, we reduce the multivariate system to a system with a single variable, which has the slowest time scale. To determine the fluctuations (i.e. diffusion) that generate the same covariance of the slow variable in a reduced system  $\bar{C}_{ss}$ , we apply the steady-state assumption to the following reduced systems with a single slow variable:

$$\begin{aligned} \frac{d}{dt} \bar{X}_s &= F_s(\bar{X}_s), \\ \frac{d}{dt} \bar{C}_{ss} &= J_s \bar{C}_{ss} + \bar{C}_{ss} J_s^T + D_{total}, \end{aligned} \quad (4.4)$$

$$D_{total} = -2J_s \bar{C}_{ss}. \quad (4.5)$$

The slowest variable is taken as the intrinsic variable and the rest of the components are extrinsic variables which may induce some fluctuations to the intrinsic variable. Therefore for our purposes, the intrinsic variable is always the slowest variable.

For the 4-variable system (2.6), the opening at the protein  $P_i$  results in the following total diffusion:

$$D_{total} = 2\gamma_{P_i} \bar{C}_s. \quad (4.6)$$

For  $n = 2$ , and equal mRNA decay rates,  $\gamma_{R_1} = \gamma_{R_2} = \gamma_R$ , the above expression is

$$\begin{aligned}
D_{total}(\bar{P}_i) = & \underbrace{\frac{\lambda_i}{\gamma_R} \left[ b + v_m \frac{\left( \frac{\lambda_i}{\gamma_{P_k} \gamma_R} \omega \right)^2}{K_d^2 + \left( \frac{\lambda_i}{\gamma_{P_k} \gamma_R} \omega \right)^2} \right]}_{D_{Intrinsic}} - \gamma_{P_i} \bar{P}_i + \underbrace{\frac{2\lambda_i^2 \lambda_k^2 v_m \omega}{\gamma_{P_k} \gamma_R^2 (\gamma_{P_i} + \gamma_R)^2 (\gamma_{P_k} + \gamma_{P_i})} \left[ \frac{2\lambda_i K_d^2 v_m \omega}{\gamma_{P_k} \gamma_R \left( K_d^2 + \left( \frac{\lambda_i}{\gamma_{P_k} \gamma_R} \omega \right)^2 \right)^2} \right]}_{D_{Extrinsic(R_k \rightarrow P_i)}} + \\
& \underbrace{\frac{2\lambda_i^3 \omega}{\gamma_{P_k} \gamma_R^2 (\gamma_{P_i} + \gamma_R)^2 (\gamma_{P_k} + \gamma_{P_i})} \left[ \frac{2\lambda_i K_d^2 v_m \omega}{\gamma_{P_k} \gamma_R \left( K_d^2 + \left( \frac{\lambda_i}{\gamma_{P_k} \gamma_R} \omega \right)^2 \right)^2} \right]}_{D_{Extrinsic(P_k \rightarrow P_i)}} + \underbrace{\frac{2\lambda_i^2}{\gamma_R (\gamma_{P_k} + \gamma_{P_i})} \left[ b + v_m \frac{\left( \frac{\lambda_i}{\gamma_{P_k} \gamma_R} \omega \right)^2}{K_d^2 + \left( \frac{\lambda_i}{\gamma_{P_k} \gamma_R} \omega \right)^2} \right]}_{D_{Extrinsic(R_i \rightarrow P_i)}}
\end{aligned}
\tag{4.7}$$

### 4.3. Calculation of MFPT with FPE

The mean first passage time (MFPT) represents the average time for a random variable to switch from one state to another. The MFPT was calculated with the adjoint operator of the FPE (Scott, 2011):

$$H(x) \frac{d}{dx} T(x) + \frac{1}{2} D_{total}(x) \frac{\partial^2}{\partial x^2} T(x) = -1, \tag{4.8}$$

with the following boundary conditions

$$T(R) = 0, \quad \left. \frac{d}{dx} T(x) \right|_{x=R^*} = 0, \tag{4.9}$$

where  $R$  and  $R^*$  represent the absorbing and reflecting boundaries, respectively (reference is needed). Since the stochastic variable  $x$  is number of molecules, these boundaries are the identical.  $R = R^* = 0$  for the OFF – ON transitions, and  $R = R^* = \infty$  for the ON – OFF transitions.  $H(x)$  and  $D_{total}(x)$  represent the drift and diffusion respectively and  $T(x)$  is the MFPT for the system to reach the point  $x$ .

The definition of the drift depends in the nature of fluctuations. In the absence of correlation between the system variable and the fluctuations, the drift can be obtained based on Riemann integration and the method is known as Ito

$$H(X) = F(\bar{X}) \quad (4.10)$$

The Ito interpretation implies that noise is white, with a zero auto-correlation time. In real systems, fluctuations are time-dependent with non-zero auto-correlation time. Such time-dependence represents the correlation between the system variable and extrinsic fluctuations, which can change the expectation value of  $X(t)$ . From mathematical point of view, the absence or presence of such correlation can be imposed to the system by different integration methods. With the Lebesgue integration, the method is known as Stratonovich in which the drift is modified:

$$H(X) = F(\bar{X}) + \frac{1}{4} \frac{d}{dX} D_{Extrinsic} \quad (4.11)$$

In the present study, we adopt the suggestion to use the Ito interpretation for intrinsic noise and the Stratonovich interpretation for extrinsic noise (Horsthemke and Lefever, 1984; Lei, 2009). Thus, when the intensity of extrinsic noise becomes sufficiently large, we apply the Stratonovich interpretation.

#### 4.4. Calculation of MFPT by the Kramers' escape rate

Solving equation (4.8) with the boundary condition (4.9) results in a double integral:

$$T(x_h) - T(x_l) = \int_{x_l}^{x_h} \frac{1}{V(x)} \int_0^x \frac{V(x)}{D_{total}(x)} dx dx \quad (4.12)$$

, where

$$V(x) = - \int \frac{H(x)}{D_{total}(x)} dx, \quad (4.13)$$

is the effective potential function. The high (ON) and low (OFF) steady states are denoted by  $x_h$  and  $x_l$ , respectively. It is difficult to calculate (4.12). Kramers introduced a parabolic

approximation of  $V(x)$  around its minimum  $x_l$  (stable steady state) and maximum  $x_u$  (unstable steady state) to obtain the MFPT (Kramers, 1940):

$$T(x_l) - T(x_u) = \frac{2\pi}{\sqrt{|V''(x_l)V''(x_u)|}} e^{|V(x_l) - V(x_u)|}, \quad (4.14)$$

where  $V''(x)$  is the second derivative of the effective potential function with respect to the state variable.

#### 4.5. Stochastic simulation algorithm

The stochastic simulation algorithm (SSA) according to Gillespie was used to generate large ensemble realizations of stochastic events based on the master equations (Gillespie, 1977). We assumed a system size of  $\Omega = 1 \mu\text{m}^3$ , which permits a simple conversion of molecule copy number in a discrete system to concentrations in a continuous system since 1 molecule / cell corresponds to 1 nM concentration.  $\Omega = 1 \mu\text{m}^3$  is a realistic number for the yeast cell nucleus or for a bacterial cell.

For MFPT calculation from OFF to ON (ON to OFF) state, we set the initial condition at the OFF (ON) state, simulate for a period corresponding to the half-life of the slowest species in order to get a distribution for the initial condition. After setting such an initial condition with a proper probability distribution, we reset the time to zero and start counting the time until the system reaches the ON (OFF) state for the first time. Collecting first passage time from each iteration, we calculated their mean value to obtain the MFPT. MFPTs were calculated at least up to  $10^5$  hours, which corresponds to around 11 years.

#### 4.6. Fitted functions for mean and variance of the output

The following functions were used for the fitting of both open-loop systems (single variable and two-variable feedback loops).

$$f_{fitted} = b + v_m \frac{\omega^2}{K_d^2 + \omega^2}, \quad (4.15)$$

$$C_{P_{fitted}} = \alpha_1 (f(\omega) + \alpha_2 P) + \alpha_3 P^2 \Big|_{P=f(\omega)=f_{fitted}}, \quad (4.16)$$

For the variance (4.16), we used the first two terms in the function (2.13), omitting the correlation term.

## Acknowledgements

We thank V Jaquet for comments on the manuscript. This work was supported by a grant from the Swiss National Foundation (SNSF).

## References

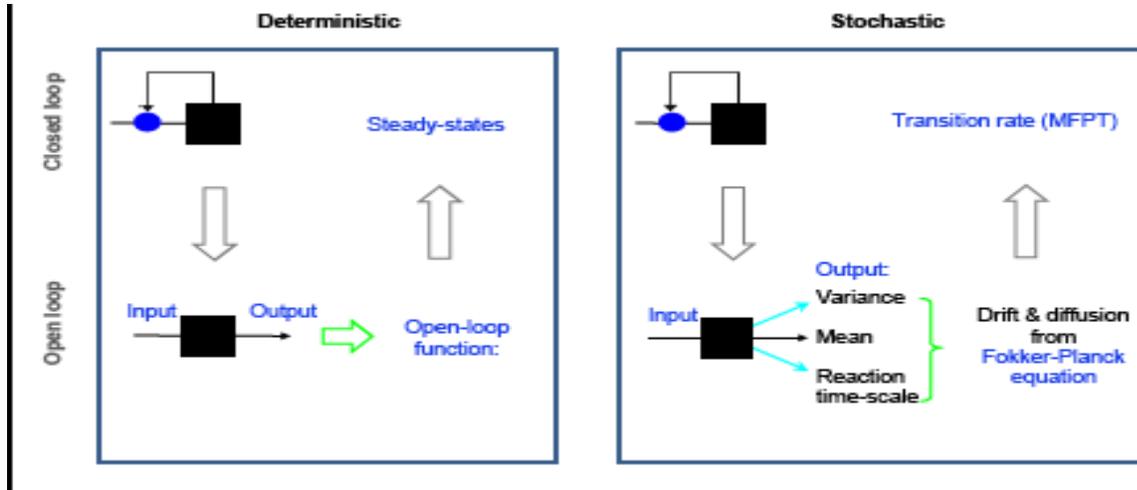
- Angeli, D., Ferrell, J. E., Jr., Sontag, E. D., 2004. Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems. *Proceedings of the National Academy of Sciences of the United States of America* 101, 1822-7, doi:10.1073/pnas.0308265100.
- Bednarz, M., Halliday, J. A., Herman, C., Golding, I., 2014. Revisiting bistability in the lysis/lysogeny circuit of bacteriophage lambda. *PLoS one* 9, e100876, doi:10.1371/journal.pone.0100876.
- Biancalani, T., Dyson, L., McKane, A. J., 2014. Noise-induced bistable states and their mean switching time in foraging colonies. *Physical review letters* 112, 038101, doi:10.1103/PhysRevLett.112.038101.
- Bonde, M. M., Voegeli, S., Baudrimont, A., Seraphin, B., Becskei, A., 2014. Quantification of pre-mRNA escape rate and synergy in splicing. *Nucleic acids research* 42, 12847-60, doi:10.1093/nar/gku1014.
- Cai, L., Dalal, C. K., Elowitz, M. B., 2008. Frequency-modulated nuclear localization bursts coordinate gene regulation. *Nature* 455, 485-90, doi:10.1038/nature07292.
- Domingo-Sananes, M. R., Szoor, B., Ferguson, M. A., Urbaniak, M. D., Matthews, K. R., 2015. Molecular control of irreversible bistability during trypanosome developmental commitment. *The Journal of cell biology* 211, 455-68, doi:10.1083/jcb.201506114.
- Erkal, C., Cecen, A. A., 2014. Empirical Fokker-Planck-based test of stationarity for time series. *Physical review. E, Statistical, nonlinear, and soft matter physics* 89, 062907, doi:10.1103/PhysRevE.89.062907.
- Escudero, C., Kamenev, A., 2009. Switching rates of multistep reactions. *Physical review. E, Statistical, nonlinear, and soft matter physics* 79, 041149, doi:10.1103/PhysRevE.79.041149.
- Fuliński, A., Telejko, T., 1991. On the effect of interference of additive and multiplicative noises. *Physics Letters A* 152, 11-14.
- Gillespie, D. T., 1977. Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry* 81, 2340-2361.
- Grima, R., Thomas, P., Straube, A. V., 2011. How accurate are the nonlinear chemical Fokker-Planck and chemical Langevin equations? *The Journal of chemical physics* 135, 084103, doi:10.1063/1.3625958.
- Guerrero, P., Byrne, H. M., Maini, P. K., Alarcon, T., 2016. From invasion to latency: intracellular noise and cell motility as key controls of the competition between resource-limited cellular populations. *Journal of mathematical biology* 72, 123-56, doi:10.1007/s00285-015-0883-2.
- Horsthemke, W., Lefever, R., 1984. Noise-induced transitions : theory and applications in physics, chemistry, and biology. Springer-Verlag, Berlin ; New York.

- Hsu, C., Jaquet, V., Maleki, F., Becskei, A., 2016a. Contribution of Bistability and Noise to Cell Fate Transitions Determined by Feedback Opening. *Journal of molecular biology*, doi:10.1016/j.jmb.2016.07.024.
- Hsu, C., Jaquet, V., Gencoglu, M., Becskei, A., 2016b. Protein Dimerization Generates Bistability in Positive Feedback Loops. *Cell reports* 16, 1204-10, doi:10.1016/j.celrep.2016.06.072.
- Hsu, C., Scherrer, S., Buetti-Dinh, A., Ratna, P., Pizzolato, J., Jaquet, V., Becskei, A., 2012. Stochastic signalling rewires the interaction map of a multiple feedback network during yeast evolution. *Nature communications* 3, 682, doi:10.1038/ncomms1687.
- Hwang, H. J., Velazquez, J. J. L., 2013. Bistable stochastic biochemical networks: highly specific systems with few chemicals. *Journal of Mathematical Chemistry* 51, 1343-1375, doi:10.1007/s10910-013-0150-y.
- Jaruszewicz, J., Zuk, P. J., Lipniacki, T., 2013. Type of noise defines global attractors in bistable molecular regulatory systems. *Journal of Theoretical Biology* 317, 140-51, doi:10.1016/j.jtbi.2012.10.004.
- Kampen, N. G. v., 1992. *Stochastic processes in physics and chemistry*. North-Holland, Amsterdam ; New York.
- Kaufman, M., Soule, C., Thomas, R., 2007. A new necessary condition on interaction graphs for multistationarity. *Journal of theoretical biology* 248, 675-85, doi:10.1016/j.jtbi.2007.06.016.
- Kramers, H. A., 1940. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* 7, 284-304.
- Kuwahara, H., Soyer, O. S., 2012. Bistability in feedback circuits as a byproduct of evolution of evolvability. *Molecular systems biology* 8, 564, doi:10.1038/msb.2011.98.
- Lei, J., 2009. Stochasticity in single gene expression with both intrinsic noise and fluctuation in kinetic parameters. *Journal of theoretical biology* 256, 485-92, doi:10.1016/j.jtbi.2008.10.028.
- Li, C., Wang, J., 2013. Quantifying Waddington landscapes and paths of non-adiabatic cell fate decisions for differentiation, reprogramming and transdifferentiation. *Journal of the Royal Society, Interface / the Royal Society* 10, 20130787, doi:10.1098/rsif.2013.0787.
- Majer, I., Hajihosseini, A., Becskei, A., 2015. Identification of optimal parameter combinations for the emergence of bistability. *Physical Biology* 12, 066011, doi:10.1088/1478-3975/12/6/066011.
- Mei, D., Xie, G., Cao, L., Wu, D., 1999. Mean first-passage time of a bistable kinetic model driven by cross-correlated noises. *Physical Review E* 59, 3880.
- Newman, J. R., Ghaemmaghami, S., Ihmels, J., Breslow, D. K., Noble, M., DeRisi, J. L., Weissman, J. S., 2006. Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature* 441, 840-6, doi:10.1038/nature04785.
- Park, B. O., Ahrends, R., Teruel, M. N., 2012. Consecutive positive feedback loops create a bistable switch that controls preadipocyte-to-adipocyte conversion. *Cell reports* 2, 976-90, doi:10.1016/j.celrep.2012.08.038.
- Paulsson, J., 2004. Summing up the noise in gene networks. *Nature* 427, 415-8, doi:10.1038/nature02257.
- Pesce, G., McDaniel, A., Hottovy, S., Wehr, J., Volpe, G., 2013. Stratonovich-to-Ito transition in noisy systems with multiplicative feedback. *Nature communications* 4, 2733, doi:10.1038/ncomms3733.
- Pisarchik, A. N., Feudel, U., 2014. Control of multistability. *Physics Reports-Review Section of Physics Letters* 540, 167-218, doi:DOI 10.1016/j.physrep.2014.02.007.
- Rai, N., Anand, R., Ramkumar, K., Sreenivasan, V., Dabholkar, S., Venkatesh, K. V., Thattai, M., 2012. Prediction by promoter logic in bacterial quorum sensing. *PLoS computational biology* 8, e1002361, doi:10.1371/journal.pcbi.1002361.
- Rybakova, K. N., Bruggeman, F. J., Tomaszewska, A., Mone, M. J., Carlberg, C., Westerhoff, H. V., 2015. Multiplex Eukaryotic Transcription (In)activation: Timing, Bursting and Cycling of a Ratchet Clock Mechanism. *PLoS computational biology* 11, e1004236, doi:10.1371/journal.pcbi.1004236.

- Samoilov, M. S., Arkin, A. P., 2006. Deviant effects in molecular reaction pathways. *Nature biotechnology* 24, 1235-40, doi:10.1038/nbt1253.
- Scott, M., 2011. *Applied stochastic processes in science and engineering*.
- Shahrezaei, V., Ollivier, J. F., Swain, P. S., 2008. Colored extrinsic fluctuations and stochastic gene expression. *Molecular Systems Biology* 4, 196, doi:10.1038/msb.2008.31.
- Steuer, R., 2004. Effects of stochasticity in models of the cell cycle: from quantized cycle times to noise-induced oscillations. *Journal of theoretical biology* 228, 293-301, doi:10.1016/j.jtbi.2004.01.012.
- Thomas, P., Straube, A. V., Grima, R., 2012. The slow-scale linear noise approximation: an accurate, reduced stochastic description of biochemical networks under timescale separation conditions. *BMC systems biology* 6, 39.
- Thomas, P., Matuschek, H., Grima, R., 2013. How reliable is the linear noise approximation of gene regulatory networks? *BMC Genomics* 14 Suppl 4, S5, doi:10.1186/1471-2164-14-S4-S5.
- Thomas, P., Popovic, N., Grima, R., 2014. Phenotypic switching in gene regulatory networks. *Proceedings of the National Academy of Sciences of the United States of America* 111, 6994-9, doi:10.1073/pnas.1400049111.
- Verdugo, A., Vinod, P. K., Tyson, J. J., Novak, B., 2013. Molecular mechanisms creating bistable switches at cell cycle transitions. *Open biology* 3, 120179, doi:10.1098/rsob.120179.
- Vinuelas, J., Kaneko, G., Coulon, A., Vallin, E., Morin, V., Mejia-Pous, C., Kupiec, J. J., Beslon, G., Gandrillon, O., 2013. Quantifying the contribution of chromatin dynamics to stochastic gene expression reveals long, locus-dependent periods between transcriptional bursts. *BMC biology* 11, 15, doi:10.1186/1741-7007-11-15.
- Yuan, L., Chan, G. C., Beeler, D., Janes, L., Spokes, K. C., Dharaneeswaran, H., Mojiri, A., Adams, W. J., Sciuto, T., Garcia-Cardena, G., Molema, G., Kang, P. M., Jahroudi, N., Marsden, P. A., Dvorak, A., Regan, E. R., Aird, W. C., 2016. A role of stochastic phenotype switching in generating mosaic endothelial cell heterogeneity. *Nature communications* 7, 10160, doi:10.1038/ncomms10160.

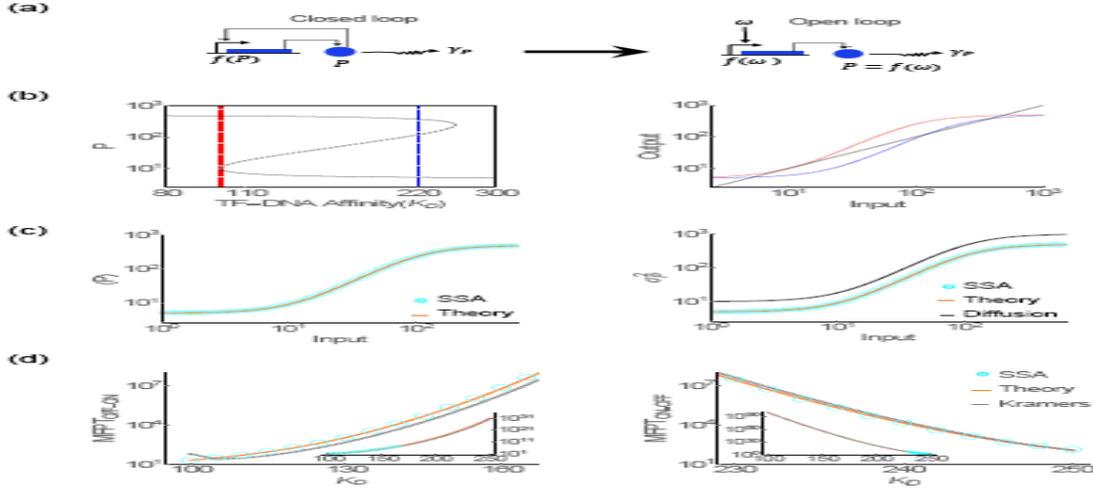
**Figure captions:**

**Fig. 1.** Principles of deterministic and stochastic modeling by loop opening. To open a (closed) feedback loop, a feedback component has to be broken into an input and output. In deterministic modeling (left), the open-loop function defines the number and value of steady-state levels for the feedback system. In stochastic modelling (right), the output is characterized by three characteristics: mean, variance and time-scale of the reaction that describes the output.

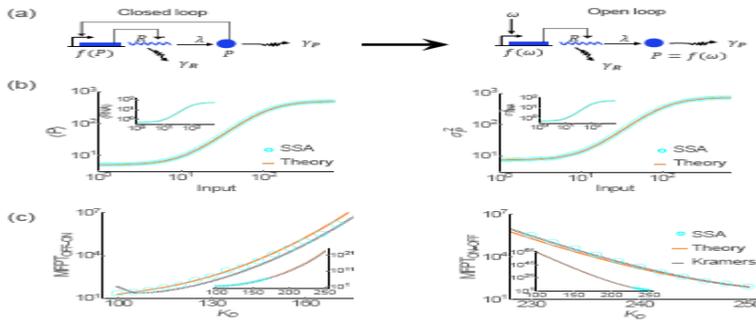


**Fig. 2.** Deterministic and stochastic modeling of the single-variable loop. (a) The scheme of the reactions in closed-loop (left) which has been opened at protein (right). In the closed-loop, protein regulates the promoter cooperatively, described by a Hill-function  $h(P)$  with a Hill-coefficient,  $n=2$ . (b) Open and closed-loop behavior of a positive feedback loop with the following fixed parameters  $v_m = 500, b = \frac{v_m}{100}, \gamma_p = 1, n = 2$ . For the deterministic models (left), the open-loop response is shown (left down) for  $K_D = 100$  (red) and  $K_D = 220$  (blue). The corresponding steady-state values in the closed-loop system are indicated by dashed lines of the same color (left top). For the stochastic modeling of the MFPT (right), not only the (mean) steady-state value of the output is needed but also the noise (diffusion or variance) and reaction time-scale. (c) The mean (left) and the variance (right) of the open-loop output with respect to the input (at  $K_D = 110$ ) obtained by stochastic simulation (cyan circle) and theory (orange solid). The model corresponds to that shown in (a, b). The total diffusion (black solid) is linearly proportional to the variance of output. (d) Mean first passage time (MFPT) from the OFF to the ON state (left) and to the ON to OFF state (right) have been calculated by stochastic simulation

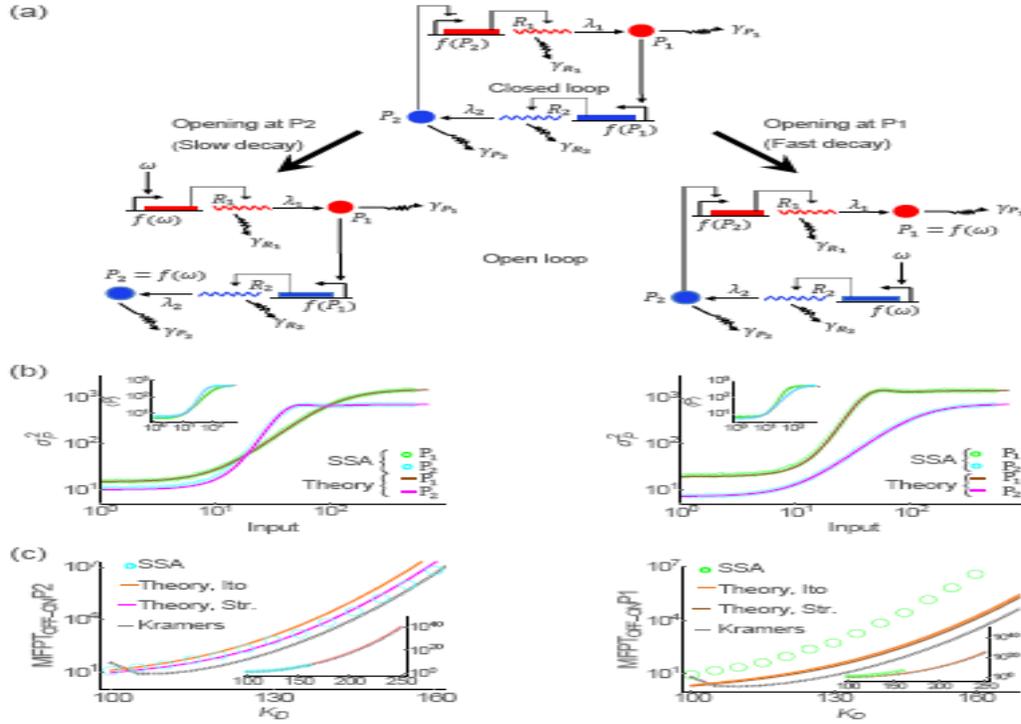
algorithm up to  $10^6$ . MFPT for the bistable range of the TF-DNA affinity ( $K_D$ ) is shown in the inset, (see Calculation of MFPT with FPE in the Appendix). The good agreement between the stochastic simulation (cyan circle) and theory (orange rectangle) indicates the accuracy of the method. With the drift and diffusion terms, MFPT was also calculated using the Kramers method, as well (gray dashed).



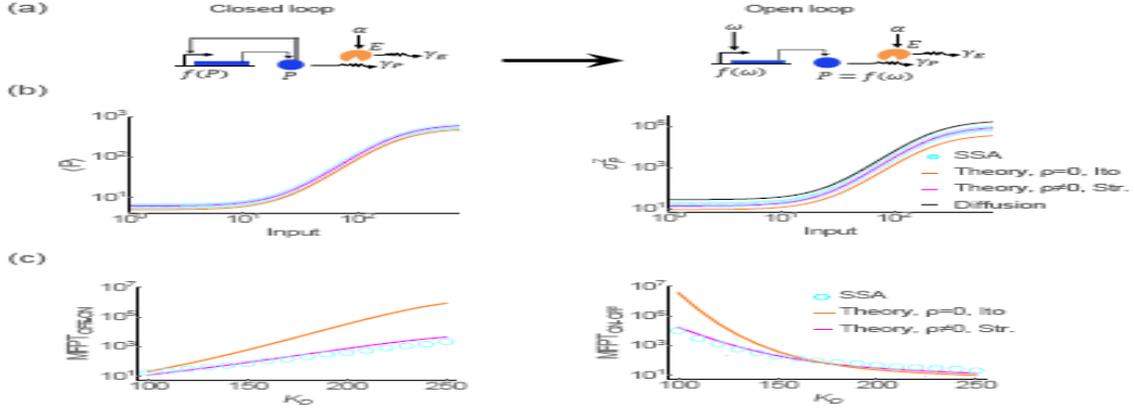
**Fig. 3.** Prediction of MFPT in the two-variable feedback comprising RNA and protein. (a) The scheme of reactions in the closed loop (left), which has been opened at the protein (right). The following parameters had fixed values:  $v_m = 1000, b = \frac{v_m}{100}, \gamma_{RNA} = 20, \gamma_p = 1, \lambda = 10, n = 2$ . (b) The mean (left) and the variance (right) of the output with respect to the input (at  $K_D = 110$ ) calculated with stochastic simulation (cyan circle) and theoretically (orange solid). Mean and variance of RNA values are shown in the inset. (c) MFPT has been calculated by stochastic simulation algorithm up to  $10^6$ . MFPT for the bistable range of TF-DNA affinity ( $K_D$ ) is shown in the inset. MFPT was also calculated using the Kramers method (gray dashed).



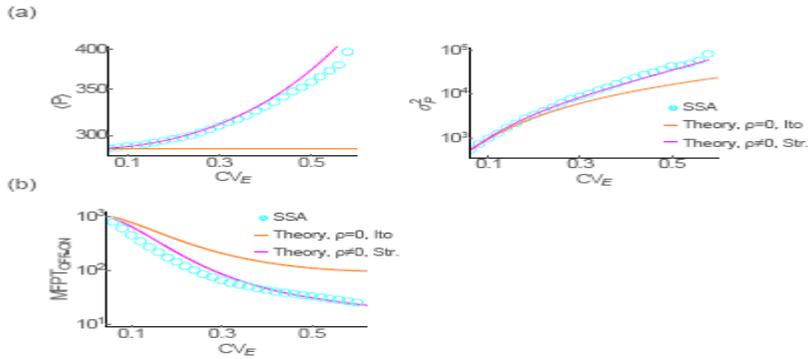
**Fig. 4.** Stochastic modelling of the closed and open loops of a four-variable feedback comprising two RNA and two protein species. (a) The scheme of reactions in the closed loop (top), which has been opened either at  $P_1$  with the shorter half-life (right) or at  $P_2$  with longer half-life (left). In the closed-loop,  $P_2$  regulates the transcription of  $RNA_1$  and  $P_1$  regulates the transcription of  $RNA_2$  cooperatively, with a Hill-coefficient of  $n=2$  at each promoter. The following parameters had fixed values:  $v_m = 1000, b = \frac{v_m}{100}, \gamma_{RNA1} = \gamma_{RNA2} = 20, \gamma_{P1} = 5, \gamma_{P2} = 1, \lambda_1 = 50, \lambda_2 = 10$ . (b) The mean and variance were calculated by stochastic simulation for  $P_1$  (green circle) and  $P_2$  (cyan circle) and by theory (brown & magenta solid, respectively) at  $K_D = 110$ . The feedback loop has been opened at  $P_1$  (right) or  $P_2$  (left). Mean values of both proteins are shown in the inset. (c) MFPT from OFF to the ON state has been calculated for  $P_1$  (right) and  $P_2$  (left) by stochastic simulation algorithm up to  $10^6$ . MFPT for the bistable range of TF-DNA affinity ( $K_D$ ) is shown in the inset. A good agreement between the SSA and theory for the opening at  $P_2$  (left) but not for the opening at  $P_1$  (right) indicates that only variables that are fast than the input and output can be considered to be in steady-state, in order to calculate the total diffusion. It can be seen that using Stratonovich interpretation (magenta) may not have a considerable effect on the open-loop result but improves the MFPT results compared to those from Ito interpretation (orange solid curve). MFPT was also calculated using the Kramers method (gray dashed).



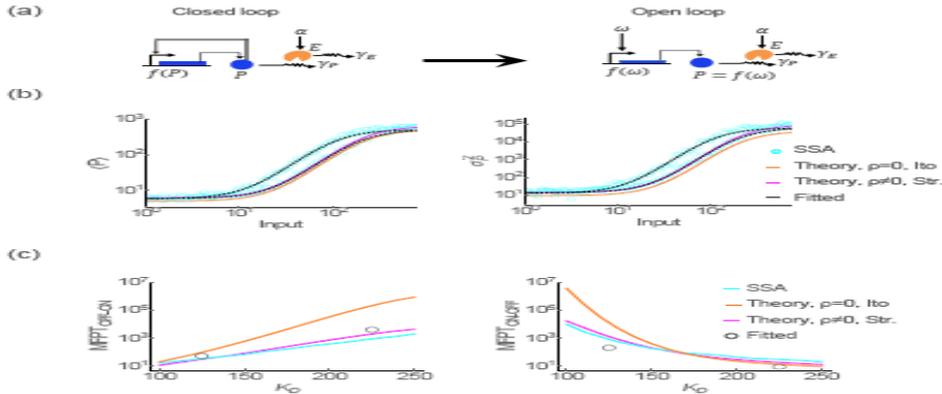
**Fig. 5.** Stochastic modelling of the closed and open loops of a one-variable feedback loop interacting with an external component. (a) The scheme of the reactions in the closed-loop (left), which has been opened at the protein (right).  $v_m = 500, b = \frac{v_m}{100}, \gamma_p = 0.2$  and  $n = 2$ . The protein is degraded by an enzyme, which displays a Poisson distribution arising due to a birth-death process, with  $\beta = 1.25, \gamma_E = 0.25$ . (b) The open-loop response (left) and the variance of the output with respect to the input (right) is shown for  $K_d = 125$  calculated by stochastic simulation (cyan circle) and theory when the random forces that induce fluctuations are uncorrelated (orange solid) or correlated (magenta solid). For uncorrelated fluctuation, the Ito interpretation was used whereas the Stratonovich interpretation was used for cross-correlated noises. (c) MFPT from the OFF to the ON state (left) and from the ON to OFF state (right) was calculated for the bistable range of the TF-DNA affinity ( $K_D$ ).



**Fig. 6.** Open and closed loop behavior in response to variation of the external noise. The same model was studied as in Fig. 4, with the following differences. The production rate of the enzyme was tuned to adjust the noise,  $\beta = 0.25/CV_E^2$ , where  $CV_E$  is the coefficient of the variation of the steady-state enzyme distribution. To keep the decay of protein at a constant rate, the decay rate constant was compensating the variations in the enzyme concentration:  $\gamma_P = CV_E^2$ . (a) The open-loop response (left) and the variance of the output (right) with respect to  $CV_E$ , as the extrinsic source of fluctuation for  $input = 125$ ,  $K_D = 130$ . The stochastic simulation results (cyan circle) match that of theory with correlated noises (solid magenta). The total diffusion is shown for the correlated noise (solid black). (b) MFPT for protein from the OFF to ON state is shown as  $CV_E$  was varied. The good agreement between the results of stochastic simulation (cyan circle) and theory (magenta rectangle) reflects the importance of the correlation between noises. The parameters are fixed at  $v_m = 500$ ,  $b = \frac{v_m}{100}$ ,  $\gamma_P = CV_E^2$ ,  $\gamma_E = 0.25$ ,  $\beta = 1/4CV_E^2$ ,  $K_D = 130$ ,  $n = 2$ .

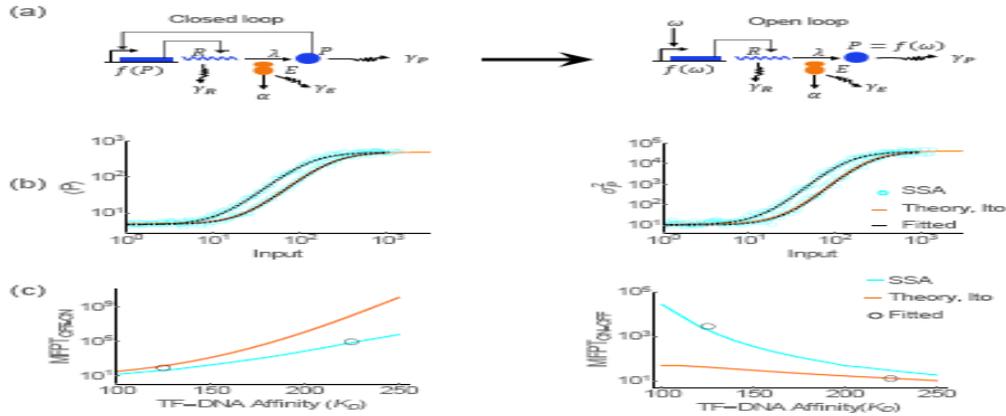


**Fig. 7.** Prediction of the MFPT based on the fitted open-loop function. (a) The scheme of reactions in a single-variable closed-loop (left, as in Fig. 5), which has been opened at the protein (right). (b) For the open-loop response, the mean (left) and the variance of the output (right) are shown for  $K_d = 125$  and  $K_d = 225$ . To emulate an experimentally realistic open-loop response (cyan circle), a normally distributed random variable ( $mean = 0$ ;  $variance = 0.1 \cdot \overline{output}$  (mean of the output) and  $variance = 0.2 \cdot Var(output)$ ), which represents the experimental error, was added to the simulated value (SSA). For comparison, the theory values are shown (as in Fig. 5) for one of the responses ( $K_d = 225$ ) (orange and magenta solid). The black dashed curves are the function fitted to mean value and the variance of the output in the open-loop system, with the following values ( $\pm$  standard error of the fitting): ( $K_d = 125$ ):  $n = 2 \pm 0.05$ ,  $K_d = 113 \pm 5.2$ ,  $v_m = 500 \pm 17.8$ ,  $b = 5.9 \pm 0.15$ ;  $\alpha_1 = 0.17 \pm 0.024$ ,  $\alpha_2 = 4.95 \pm 0$  and,  $\alpha_3 = 0.258 \pm 0.01$ ; ( $K_d = 225$ )  $n = 2 \pm 0.05$ ,  $K_d = 201 \pm 9.5$ ,  $v_m = 500 \pm 20.2$ ,  $b = 5.7 \pm 0.11$ ;  $\alpha_1 = 0.11 \pm 0.017$ ,  $\alpha_2 = 5.15 \pm 0$  and  $\alpha_3 = 0.283 \pm 0.009$ . (c) MFPT from the OFF to the ON state (left) and from the ON to OFF state (right) was calculated for the bistable range of the TF-DNA affinity ( $K_D$ ). The black circles represent MFPT which is calculated using the drift and diffusion obtained from equations (2.13) and (2.14) by having the open-loop function and the variance from (4.15) and (4.16) fitted to the open-loop data.



**Fig. 8.** Prediction of the MFPT based on the fitted function for the open-loop obtained from a two-variable feedback loop interacting with an external component. (a) The scheme of the reactions in the closed-loop (left), which has been opened at the protein (right). (b) For the open-loop response, the mean (left) and the variance of the output (right) are shown for  $K_d = 125$  and  $K_d = 225$ . To emulate an experimentally realistic open-loop response (cyan circle), a

normally distributed random variable ( $mean = 0$ ;  $variance = 0.1 \cdot \overline{output}$  and  $variance = 0.2 \cdot Var(output)$ ), which represents the experimental error, was added to the simulated value (SSA). For comparison, the theory values are shown (as in Fig. 5) for one of the responses ( $K_d = 225$ ) (orange and magenta solid). The black dashed curves are the function fitted to mean value and the variance of the output in the open-loop system with the following values ( $\pm$  standard error of the fitting: ( $K_d=125$ ):  $n = 2 \pm 0.04$ ,  $K_d = 130 \pm 4.9$ ,  $v_m = 498 \pm 14.2$ ,  $b = 4.9 \pm 0.11$ ;  $\alpha_1 = 0.56 \pm 0.029$ ,  $\alpha_2 = 0.15 \pm 0.005$  and,  $\alpha_3 = 1.94 \pm 0.005$ ; ( $K_d=225$ )  $n = 1.92 \pm 0.04$ ,  $K_d = 235 \pm 9.3$ ,  $v_m = 499 \pm 15.2$ ,  $b = 4.6 \pm 0.09$ ;  $\alpha_1 = 0.55 \pm 0.028$ ,  $\alpha_2 = 0.153 \pm 0.005$  and  $\alpha_3 = 1.84 \pm 0.005$ . (c) MFPT from the OFF to the ON state (left) and from the ON to OFF state (right) was calculated for the bistable range of the TF-DNA affinity ( $K_D$ ). The black circles represent MFPT which is calculated using the drift and diffusion obtained from equations (2.13) and (2.14) by fitting the equations (4.15) and (4.16) fitted to the open-loop data. The parameters are fixed at  $v_m = 500$ ,  $b = \frac{v_m}{100}$ ,  $\gamma_p = 0.2$ ,  $\alpha = 1.25$ ,  $\gamma_E = 0.25$ ,  $n = 2$ .



## Highlights

- A stochastic open-loop approach is formulated to predict transition rates in bistable feedback loops.
- The open-loop output is characterized by three basic properties: the mean and variance of the output and its time-scale.
- The three basic properties predict the transition rates without the need to characterize the internal components of the feedback.
- The prediction performs well even when a stochastic deviant effect shifts the mean value of the open-loop output.