POLITECNICO DI TORINO
Repository ISTITUZIONALE

Stochastic sequestration dynamics: A minimal model with extrinsic noise for bimodal distributions and competitors correlation

Original
Stochastic sequestration dynamics: A minimal model with extrinsic noise for bimodal distributions and competitors correlation / Del Giudice, M., Bosia, C., Grigolon, S., Bo, S.. - In: SCIENTIFIC REPORTS. - ISSN 2045-2322. - ELETTRONICO. - 8:1(2018), p. 10387. [10.1038/s41598-018-28647-9]

Availability:
This version is available at: 11583/2730096 since: 2019-04-04T18:44:45Z

Publisher:
Nature Publishing Group

Published
DOI:10.1038/s41598-018-28647-9

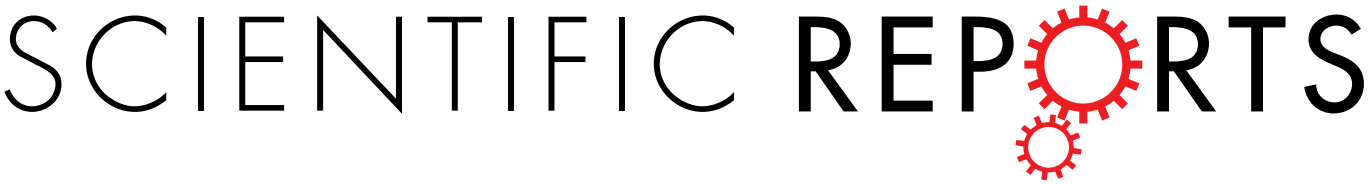
Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

27 November 2026



OPEN

Stochastic sequestration dynamics: a minimal model with extrinsic noise for bimodal distributions and competitors correlation

Received: 18 January 2018
Accepted: 21 June 2018
Published online: 10 July 2018

Marco Del Giudice[1,2], Carla Bosia[1,2], Silvia Grigolon[3] & Stefano Bo[4]

Many biological processes are known to be based on molecular sequestration. This kind of dynamics involves two types of molecular species, namely targets and sequestrants, that bind to form a complex. In the simple framework of mass-action law, key features of these systems appear to be threshold-like profiles of the amounts of free molecules as a function of the parameters determining their possible maximum abundance. However, biochemical processes are probabilistic and take place in stochastically fluctuating environments. How these different sources of noise affect the final outcome of the network is not completely characterised yet. In this paper we specifically investigate the effects induced by a source of extrinsic noise onto a minimal stochastic model of molecular sequestration. We analytically show how bimodal distributions of the targets can appear and characterise them as a result of noise filtering mediated by the threshold response. We then address the correlations between target species induced by the sequestrant and discuss how extrinsic noise can turn the negative correlation caused by competition into a positive one. Finally, we consider the more complex scenario of competitive inhibition for enzymatic kinetics and discuss the relevance of our findings with respect to applications.

Sequestration dynamics (also known as titrative dynamics) are ubiquitous in nature. Some of the best studied examples concern protein ubiquitination[1], growth factors signalling[2,3], gene-expression regulation at the transcriptional and post-transcriptional level, such as transcription factors sequestration[4–6], the interaction between RNA polymerase and its sigma factors in bacteria[7], mRNA-miRNA interaction in post-transcriptional gene regulation[8–13] and bacterial persistence[14]. The general scheme features one (or more) molecular species, the *sequestrant*, that bind to another molecular species, the *target*, to form a complex. In a simple mass-action law scenario, the system displays a threshold-like behaviour of the mean abundance of free targets as a function of the parameters determining its possible maximum amount (e.g. rate of synthesis for open systems or pool size for closed systems) as shown in Fig. 1a and in, *e.g.* refs[6,8,13,15]. This threshold-like behaviour follows from the action of the sequestrant. Intuitively, one can think about the following extreme scenarios. When the total number of the target is lower than the one of the sequestrant most of the target will tend to be bound. However, as the total number of targets begins to outnumber the sequestrant, the average of its free molecules starts to increase more markedly, often in a linear fashion. In the vicinity of the threshold, the average number of free targets moves from the mostly bound, slowly increasing regime, to the unbound, rapidly growing one. Therein, the system is defined to be *ultrasensitive*, meaning that, around the threshold, a small fold change in the amount of total target (input) leads to a larger fold change in the amount of free target (output), corresponding to a Hill coefficient larger than one[6].

On top of the average behaviour of the systems, one should consider stochastic effects since molecular interactions are known to be probabilistic and to take place in fluctuating environments. While the first source of noise, referred to as intrinsic in the following, has been fairly studied in the past, little is known about the effect of the second one, which we shall call extrinsic[16–18]. This kind of noise may be non-specific, i.e. affects in a similar way many components of a same network and it is often homogenous on a single-cell level but varies from one

[1]Department of Applied Science and Technology, Politecnico di Torino corso Duca degli Abruzzi 24, Turin, IT-10129, Italy. [2]Italian Institute for Genomic Medicine, via Nizza 52, I-10126, Torino, Italy. [3]The Francis Crick Institute, 1, Midland Road, London, NW1 1AT, United Kingdom. [4]Nordita, Royal Institute of Technology and Stockholm University, Roslagstullsbacken 23, SE-106 91, Stockholm, Sweden. Correspondence and requests for materials should be addressed to S.B. (email: stefano.bo@nordita.org)

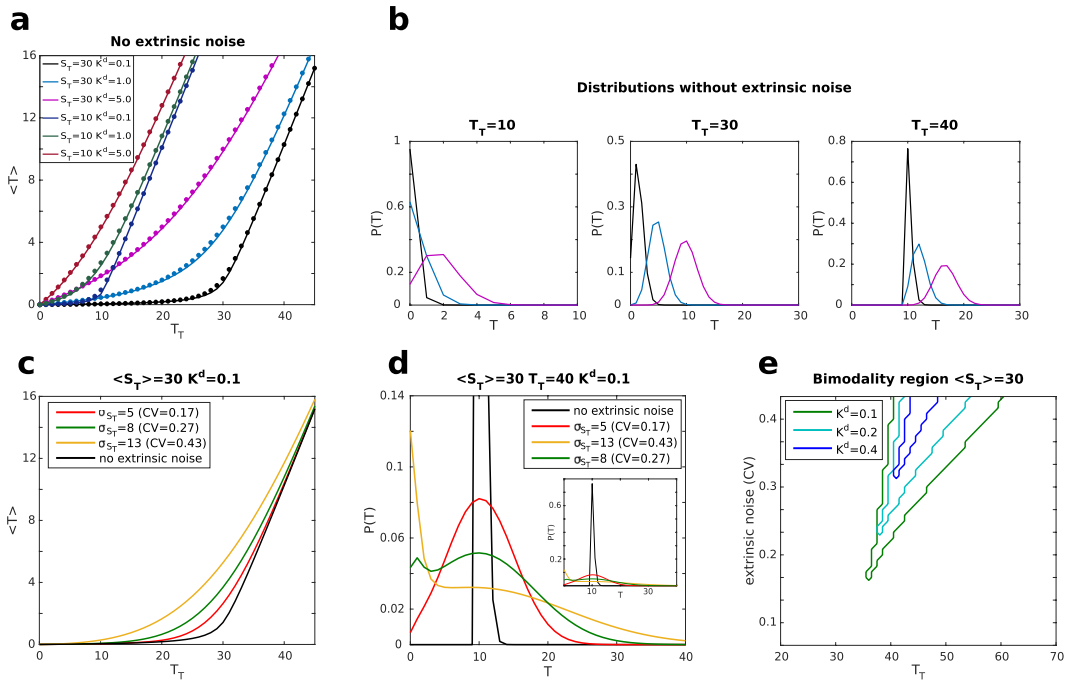


Figure 1. (**a**) $\langle T \rangle$ vs T_T for $S_T = 10$ and $S_T = 30$. Solid lines are the solution of the master eq. (12) while symbols the one of the rate eq. (6). K^d takes the values 0.1, 1.0, 5.0 (curves with lower K^d are below the ones with a higher one). (**b**) Examples of probability distributions from the solution of the master eq. (10). $S_T = 30$, K^d assumes the values: 5.0 (purple), 1.0 (cyan) and 0.1 (black). From left to right $T_T = 10$, 30, 40. (**c**) $\langle T \rangle$ vs T_T in presence of extrinsic noise on S_T as obtained from the solution of the master equation and the addition of extrinsic noise. For all the curves $K^d = 0.1$ and $\langle S_T \rangle = 30$. The black curve is the pure intrinsic noise case, while the other three correspond to a standard deviation of the Gaussian distribution of S_T equal to 5 (red), 8 (green) and 13 (yellow). (**d**) Examples of probability distribution of T in presence of extrinsic noise. The appearance of a bimodal distribution can be modulated by varying the extrinsic noise level. $\langle S_T \rangle = 30$, $T_T = 40$, $K^d = 0.1$, σ_{S_T} assumes the values: 5 (red), 8 (green) and 13 (yellow). (**e**) Plot of the bimodality region (presence of two distinct peaks) for different values of K^d as a function of T_T and extrinsic noise in units of the coefficient of variations ($CV = \sigma_{S_T}/\langle S_T \rangle$). Bimodal distributions are present for parameters inside the areas delimited by the plotted lines. A small K^d and a high noise level favour bimodality. $\langle S_T \rangle = 30$, K^d assumes the values: 0.1 (green), 0.2 (light blue) and 0.4 (blue). The size of the step along T_T is $\Delta T_T = 1$, while the size of the step for the extrinsic noise level is $\Delta \sigma_T = 0.25$ ($\Delta CV = 8 \cdot 10^{-3}$). For each point defined by these steps, the distribution $P(T)$ was computed analytically and the number of its maxima was evaluated. Note that the step-like features in the plot are due to the discreteness of T_T.

to another. Sources of extrinsic fluctuations can be variations in the amount of molecular machineries (*e.g.* RNA polymerase, ribosomes, mitochondria), gradients of temperature or chemicals' concentrations or, more widely, the coupling of the cell to the variability of the external environment. In the present contribution we resort to a minimal model of molecular sequestration and study the system in presence of both intrinsic and extrinsic noise[18,19]. This allows us to focus on two features of sequestration dynamics that have received relatively little attention in the past: the shape of the target probability distribution and the correlations amongst competing targets induced by the presence of the sequestrant. We analytically provide an exact solution to the intrinsic-noise case analysing both targets distributions and correlations in the case of one target and two different targets interacting with one single sequestrant. By adding an extrinsic source of noise, we show how the analytical distributions, previously found to be unimodal, can be reshaped into bimodal in a definite set of parameters. Bimodal distributions are often encountered in gene expression and other biomolecular assays data and are of particular interest as they may indicate the coexistence of two distinct phenotypes, with the modes of the distribution linked to different differentiation states or physiological conditions[20,21]. Inspired by the numerical results of ref.[15], we first analyse the role of ultrasensitivity close to the threshold as a possible tool to channel extrinsic fluctuations of the parameters of the system into bimodal distributions of free target amount. We find that these bimodal distributions are originated by stochastic effects only, the underlying deterministic system being indeed always monostable. Secondly, we show that the action of the sequestrant is not only responsible for the presence of the threshold which allows to channel noise, but also plays an important role in systems with multiple targets. Indeed, targets result to be effectively correlated by their competing interaction with the same sequestrant and this competition induces a negative correlation. Here we show that, if the amount of sequestrant fluctuates due to extrinsic noise, the above mentioned correlations can, from negative, become positive.

Results

Minimal Sequestration Model for noise channelling. *Sequestration defines the presence of a threshold.* Let us consider a simple model consisting of two species, namely S, the sequestrant, and T, the target. We shall view S as a titrating agent that binds T and sequesters it in the complex $\overline{TS}$. Here we simply assume a reversible dynamics, where the complex can in turn dissociate, releasing the two species in the environment for further interaction. This first order reaction network is described by the following equation:

$$T + S \underset{k_-}{\overset{k_+}{\rightleftharpoons}} \overline{TS}, \tag{1}$$

where k_+ and k_- are respectively the binding and unbinding rates. For simplicity, we shall assume in our model the total number of molecules of each species to be conserved, this assumption defining the following conservation laws:

$$S_T = S + \overline{TS} = const, \tag{2}$$

$$T_T = T + \overline{TS} = const. \tag{3}$$

In view of these conservation laws, the system has only one free variable, in the following assumed to be T. The average behaviour for this model can be studied by considering the associated rate equation for the concentration[6]. We start by writing the steady state solution of the rate equation for the case of two species that bind and unbind, with constant total concentration for each species. The rate equation for the concentration of T is:

$$\frac{d[T]}{dt} = -\hat{k}_+[T][S] + \hat{k}_-[\overline{TS}], \tag{4}$$

where $\hat{k}_+$ and $\hat{k}_-$ are the reaction rates for the concentrations. In particular $\hat{k}_- = k_-$ and $\hat{k}_+ = V_{sys}k_+$, with V_{sys} being the typical volume of the system in which reactions occur. By making use of the conservation laws for concentrations that trivially follow from (2) and (3) for constant volumes, the above equation can be easily solved at the steady state giving:

$$2[T] = [T_T] - [S_T] - \hat{K}^d + \sqrt{([T_T] - [S_T] - \hat{K}^d)^2 + 4\hat{K}^d[T_T]}, \tag{5}$$

where $\hat{K}^d \equiv \frac{k_-}{k_+V_{sys}}$ is the dissociation constant $K^d \equiv k_-/k_+$ rescaled by the volume. This deterministic approximation can be used to investigate the average behaviour of the number of free target molecules. By rescaling eq. (5) by the volume of the system we obtain:

$$2T = T_T - S_T - K^d + \sqrt{(T_T - S_T - K^d)^2 + 4K^dT_T}, \tag{6}$$

which is the steady-state deterministic solution for the amount of free target molecules.

As discussed in[6], the titrative interaction induces a threshold-like behaviour on the mean number of free molecules, upon the variation of their total amount (see Fig. 1a). The model considered here presents this feature with minimal ingredients, given by the binding and unbinding reactions between T and S and the conservation laws. Indeed, when the system is in the regime in which the total amount of target molecules is smaller than the sequestrant one, almost all of them are bound in complex and their mean free amount is close to zero. We name this region the *repressed regime*. Conversely, in the regime where the total amount of target molecules is larger than the sequestrant one, their mean number increases linearly with T_T, since the number of free molecules of the sequestrant is almost zero. We name this region the *unrepressed regime*. The position of the threshold is then located close to the equimolarity point, *i.e.* where $T_T \simeq S_T$. As a direct outcome of this behaviour, the system becomes ultrasensitive in proximity to the threshold. This means that around this point, a small fold-change variation in the total number of target molecules can result in a large fold-change of their mean free amount. In this specific framework, the steepness of the threshold is determined by the dissociation constant K^d. It follows that a small K^d implies a large affinity between the molecules, therefore leading to a sharper threshold. In the limit of infinitely large affinity between the sequestrant and the target, the system is continuous at the threshold but displays a discontinuous derivative. In this respect, the usage of the term *ultrasensitivity* depicts a scenario that differs from the ones typical of Goldbeter-Koshland, or cooperative Hill models[22] which, in the limit of infinite cooperation are discontinuous.

Full Master Equation Solution. Very often, in biochemical systems, the numbers of individual molecules at play is low. Fluctuations around the average behaviour (described by the rate equation) become then relevant and, to correctly characterise these systems, it is mandatory to take into account their stochastic nature[17]. In order to do that, we seek the explicit form of the probability distribution of the number of molecules in the system at a given time. Recalling that this model has only one independent variable, let us define $P(T, t)$ as the probability of observing T free molecules at time t. The dynamics of this probability distribution is Markovian and obeys the following chemical master equation[23,24]:

$$\frac{dP(T,\,t)}{dt} = k_+(T+1)(S+1)P(T+1,\ t) + k_-(\overline{TS}+1)P(T-1,\ t) - [k_+TS + k_-\overline{TS}]P(T,\,t). \tag{7}$$

Using the conservation laws, eqs (2) and (3), the master equation can be written as:

$$\begin{aligned}\frac{dP(T,\,t)}{dt} &= k_+(T+1)(S_T - T_T + T + 1)P(T+1,\,t)\\ &\quad + k_-(T_T - T + 1)P(T-1,\,t)\\ &\quad - [k_+T(S_T - T_T + T) + k_-(T_T - T)]P(T,\,t).\end{aligned} \tag{8}$$

We are mostly interested in the long-time behaviour of the system and we will then focus on the steady-state solution $P(T) = \lim_{t\to\infty} P(T,\,t)$. Since we have a single independent variable and all reactions are reversible, the detailed balance condition is satisfied. Then, at the steady state, equilibrium is reached and there are no probability flows between the states of the system. An important feature of the system is that the possible range of the number of free T depends both on the total number of available target molecules T_T and on the total number of sequestrant molecules S_T. Indeed, the minimum number of free T can either be $T_{min} = 0$ when $S_T \geq T_T$, or $T_{min} = T_T - S_T$ when $S_T < T_T$. To obtain the steady-state solution $P(T)$ to the master equation, one can recursively use the detailed balance condition

$$k_+(T+1)(S_T - T_T + T + 1)P(T+1) = k_-(T_T - T)P(T), \tag{9}$$

which leads to:

$$P(T) = \frac{1}{\mathcal{N}}(K^d)^T\binom{T_T}{T}\frac{1}{[S_T - (T_T - T)]!}, \tag{10}$$

where the normalisation factor is given by:

$$\mathcal{N} = \sum_{T=T_{min}}^{T_T}(K^d)^T\binom{T_T}{T}\frac{1}{[S_T - (T_T - T)]!}. \tag{11}$$

The shape of this distribution is mostly determined by the magnitude of the dissociation constant K^d and depends on whether the system is in the repressed ($T_T < S_T$) or unrepressed ($T_T > S_T$) regime. For low values of K^d the distribution of free targets T is mostly concentrated around the minimum value T_{min}, which, as mentioned before is 0 in the repressed regime and $T_T - S_T$ in the unrepressed one. Some examples of this probability distribution for different values of the dissociation constant and T_T are reported in Fig. 1b. For the explored parameters range, the target distribution was always unimodal. The various moments of the distributions can be obtained exactly by taking ensemble averages over the probability distribution $P(T)$ from eq. (10). The mean of T can be written in terms of Hypergeometric functions[25,26] as:

$$\begin{cases}\langle T\rangle = K^d\dfrac{T_T}{1 - T_T + S_T}\dfrac{{}_1F_1(1 - T_T,\,2 - T_T + S_T;\,-K^d)}{{}_1F_1(-T_T,\,1 - T_T + S_T;\,-K^d)} & T_T \leq S_T \quad (T_{min} = 0)\\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{if}\\ \langle T\rangle = (T_T - S_T)\dfrac{{}_1F_1(-S_T,\,T_T - S_T;\,-K^d)}{{}_1F_1(-S_T,\,1 + T_T - S_T;\,-K^d)} & T_T > S_T \quad (T_{min} = T_T - S_T)\end{cases} \tag{12}$$

as found in, *e.g.*, ref.[27]. Its behaviour is plotted in Fig. 1a confirming how the deterministic analysis of eq. (6) (and ref.[6]), which completely neglected fluctuations, gives an accurate description of the average number of free target molecules.

Extrinsic noise and bimodal distributions. In the previous section we focussed on the fluctuations originated by the discrete nature of molecules and the intrinsic randomness of their interactions. By introducing fluctuations in the total number of sequestrant molecules S_T, we here investigate the influence of extrinsic noise on the present system. Let us consider different copies of our system with different S_T, randomly assigned. This mimics, for instance, the scenario in heterogeneous cell populations randomly sampled. For the sake of simplicity we draw the amount of S_T in each system from a discretised Gaussian $P(S_T)$, restricted to positive values, with mean $\langle S_T\rangle$ and standard deviation σ_{S_T}.

The marginal probability distribution of free target molecules T (over the different values of S_T), $P(T)$, is now affected by the fluctuations in S_T and differs from the one given in eq. (10). Nonetheless, it can be expressed by making use of the law of total probability[28], i.e., via the superposition of the conditional probabilities $P(T|S_T)$ weighed by $P(S_T)$ as follows:

$$P(T) = \sum_{S_T=0}^{\infty}P(T|S_T)P(S_T), \tag{13}$$

where the expression of the conditional probability $P(T|S_T)$ coincides with the solution of the master equation obtained for a given S_T (Eq. (10)).

Given this scheme, let us now study the consequences of the presence of the extrinsic noise in the system. A first quantity affected by this new source of noise is certainly the average of T. As the level of extrinsic noise is increased, the profile as a function of T_T becomes smoother with a less pronounced threshold, while the regions farther away from the threshold are not heavily affected (see Fig. 1c). Indeed, the effects of the extrinsic noise are stronger in the

vicinity of the threshold where the system is ultrasensitive. Therein, for a fixed T_T, small changes in the number of total sequestrants S_T can make the system transit form the repressed to the unrepressed regime, bearing significant consequences on the shape of the probability distribution. Indeed, stochastically sampling systems above and below the threshold can result in a bimodal distribution. The first peak of such distribution is narrow and located close to the origin $T=0$. It corresponds to the aggregated contribution of the repressed systems, *i.e.*, the ones that picked an S_T larger than T_T. The second peak is broader and corresponds to the superposition of the unrepressed systems, each of them with its own mean. We remark that the underlying deterministic system is not bistable and that the observed bimodality is due to the stochastic sampling of the repressed and unrepressed regimes across the threshold. This scenario shows how the threshold feature of the system effectively filters the variability introduced by the extrinsic noise, concentrating the contributions of all the systems below threshold. At this point, one may wonder whether any extrinsic noise would have the same effect on target distributions in any parameter range. The answer is no: bimodal distributions are present only close to the threshold and are favoured by steep thresholds (small K^d), which allow sampling between the unrepressed and repressed regimes (see Fig. 1d,e). Furthermore, not all distributions of extrinsic noise may induce bimodals for the target. To this aim, the extrinsic noise is required to have a peaked distribution, sufficiently broad to sample both below and above threshold. This becomes clear by examining the Gaussian case with small variances (see Fig. 1e) and the one of uniformly distributed extrinsic noise considered in the Supplementary Information online. In the latter case, the threshold behaviour concentrates the contribution of the systems below threshold, eventually giving rise to a repressed peak, but no mechanisms could induce the appearance of an unrepressed peak (see Fig. S1 in the SI).

In view of these results, the combination of the threshold-like response produced by the titrative interaction and a suitable extrinsic noise on the total amount of one of the species can be considered a general mechanism to achieve bimodality. This result is also interesting from a pure biological point of view. In biological systems, bimodal distributions of gene expression are particularly interesting as they may indicate the presence of two distinct physiological states. At the level of a cell population, this effect would therefore show heterogeneity in different cell gene expression, as observed in[16]. How this is normally achieved remains unknown but the mechanism here discussed represents a minimal way to obtain bimodal gene expression distributions in systems based on molecular sequestration and subject to extrinsic noise. It is worth noticing that the "static" source of extrinsic noise we consider here is a good approximation of the case of a system with a slow dynamically fluctuating sequestrant. In this case, the system has time to approximately reach a steady state before the amount of sequestrant changes considerably[15]. This timescale separation is thought to be present, for instance, in the mechanism of postranscriptional regulation by microRNAs (short non-coding segments of RNA)[10]. In these systems microRNA molecules bind and sequester coding messenger RNAs and inhibit their translation. The complex formation between microRNA and mRNA takes place on a scale that is much faster than the one of transcription and degradation of microRNA and mRNA. The total number of sequestrant molecules (and the one of target ones) then fluctuates on a slower scale than the sequestration dynamics (note that if one focusses on the complex dynamics without explicitly modelling transcription and degradation, the nature of the fluctuations on the number of the sequestrant molecules becomes extrinsic).

Sequestration as a tool to couple competing species. *Intrinsic noise.* We consider here a system composed of two molecular species (T_1 and T_2) competing for binding to a third molecule (S). The reaction network that defines the minimal model is described by eqs (14) and (15). Both T_1 and T_2 can bind to S respectively with rate k_+^1 and k_+^2 to form the complexes $\overline{T_1S}$ and $\overline{T_2S}$ which can then dissociate with rates k_-^1 and k_-^2, as follows:

$$T_1 + S \underset{k_-^1}{\overset{k_+^1}{\rightleftharpoons}} \overline{T_1S}, \tag{14}$$

$$T_2 + S \underset{k_-^2}{\overset{k_+^2}{\rightleftharpoons}} \overline{T_2S}. \tag{15}$$

T_1 can be seen to act as a sponge for the common resource S, preventing the binding with the competing species T_2 and vice versa. Also in this case, we assume the total amount of each species to be constant so that the following conservation laws hold:

$$T_{1T} = T_1 + \overline{T_1S} = const, \tag{16}$$

$$T_{2T} = T_2 + \overline{T_2S} = const \tag{17}$$

and

$$S_T = S + \overline{T_1S} + \overline{T_2S} = const. \tag{18}$$

These conservation laws limit the number of independent variables from 5 to 2. Since there are fewer independent variables than the number of conserved chemical species, the range of values that the number of free molecules of a given species can take is determined by the total amounts of the other species as well as detailed in Table 1. In particular, if the total number of molecules of a species is larger than the one of its titrant, there will always be some free molecule of that species (even if all the available titrant molecules are bound to it) so that its minimal value is greater than 0. When considering two targets, the range of values of free molecules of one target (*e.g.* T_1) is set by the number of free molecules of the sequestrant, which depends on the amount of complex

	MIN	MAX
S	$max\,(0, S_T - T_{1T} - T_{2T})$	S_T
T_1	$max\,(0, T_{1T} - S_T)$	T_{1T}
T_2	$max\,(0, T_{2T} - S_T)$	T_{2T}

Table 1. Range of allowed values for sequestrant and targets. The table shows the allowed maximum and minimum value for each species, only based on their relative total amounts.

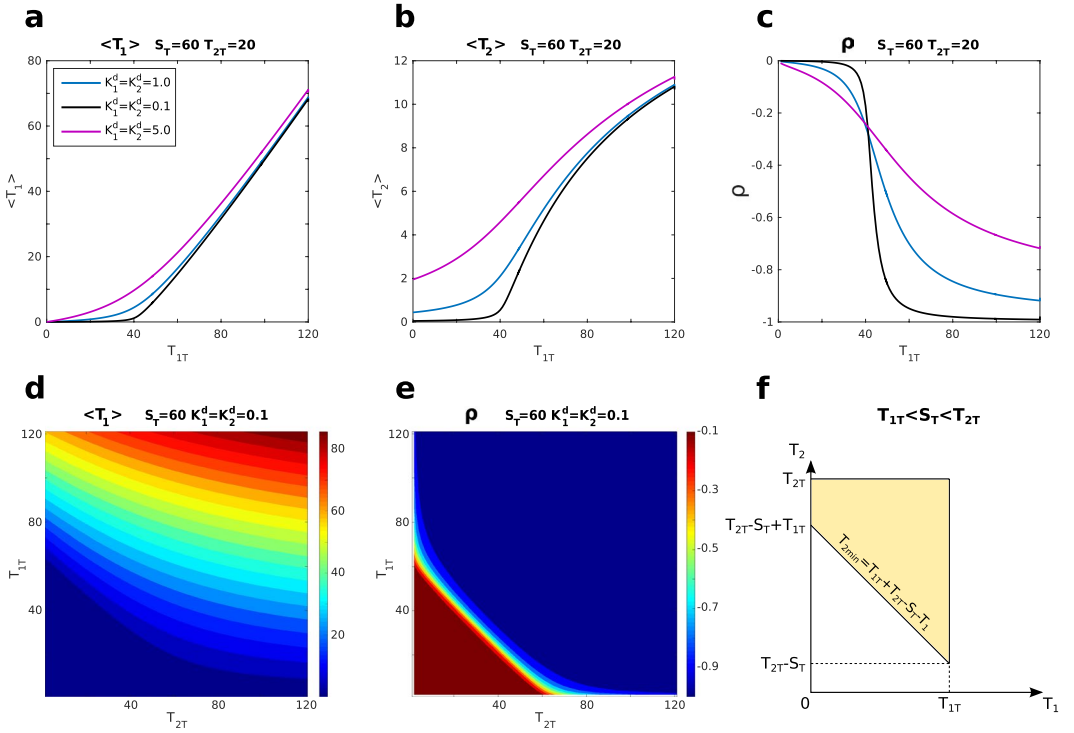


Figure 2. Two targets system. (**a**) Mean value of T_1. (**b**) Mean value of T_2. (**c**) Pearson correlation between T_1 and T_2 as a function of T_{1T} for a fixed value of T_{2T} and different values of K_1^d and K_2^d. The threshold behaviour of the means and the correlation drop are steeper for lower values of K^d. $S_T = 60$, $T_{2T} = 20$. K_1^d and K_2^d are always equal and assume the values: 5.0 (purple), 1.0 (cyan) and 0.1 (black). (**d**) Contour plot of the mean of T_1 as a function of T_{1T} and T_{2T}. $S_T = 60$, $K_1^d = K_2^d = 0.1$. (**e**) Contour plot of the Pearson correlation between T_1 and T_2 as a function of T_{1T} and T_{2T}. $S_T = 60$, $K_1^d = K_2^d = 0.1$. (**f**) Example of phase space for the case $T_{1T} < S_T < T_{2T}$.

between the second target and the sequestrant ($\overline{T_2 S}$). For instance, for a fixed amount of free target T_2 the number of free molecules of target 1 ranges between $max(0, T_{2T} + T_{1T} - S_T - T_2)$ and T_{1T}, as schematically shown in Fig. 2f. In the stochastic description, choosing T_1 and T_2 as independent variables, the model is described by eq. (19), which is the master equation governing the time evolution of the probability distribution of observing T_1 and T_2 free molecules at time t, *i.e.*:

$$\begin{aligned}\frac{dP(T_1, T_2, t)}{dt} &= (\overline{T_1 S} + 1)k_-^1 P(T_1 - 1, T_2) + (T_1 + 1)(S + 1)k_+^1 P(T_1 + 1, T_2) \\ &+ (\overline{T_2 S} + 1)k_-^2 P(T_1, T_2 - 1) + (T_2 + 1)(S + 1)k_+^2 P(T_1, T_2 + 1) \\ &- (\overline{T_1 S}k_-^1 + T_1 S k_+^1 + \overline{T_2 S}k_-^2 + T_2 S k_+^2)P(T_1, T_2).\end{aligned} \tag{19}$$

Once again, detailed balance holds and therefore it is possible to derive the equilibrium solution. As presented in the Methods section, the relations obtained imposing detailed balance, together with the conservation laws, can be used to recursively derive the solution of the master equation, eq. (19), which reads:

$$P(T_1, T_2) = \frac{(K_1^d)^{T_1}(K_2^d)^{T_2}}{\mathcal{N}_2} \frac{\binom{T_{1T}}{T_1}\binom{T_{2T}}{T_2}}{[S_T + T_1 + T_2 - (T_{1T} + T_{2T})]!}, \tag{20}$$

in agreement with the grand canonical distribution for ideal particle mixtures[23] with the normalization factor given by:

$$\mathcal{N}_2 = \sum_{T_2=\max(0,T_{2T}-S_T)}^{T_{2T}} \sum_{T_1=\max(0,T_{2T}+T_{1T}-S_T-T_2)}^{T_{1T}} \frac{(K_1^d)^{T_1}(K_2^d)^{T_2}\binom{T_{1T}}{T_1}\binom{T_{2T}}{T_2}}{[S_T + T_1 + T_2 - (T_{1T} + T_{2T})]!}. \tag{21}$$

Thresholds and correlations. As shown in Fig. 2a, the average number of free molecules of target 1, $\langle T_1 \rangle$, displays a threshold behaviour when plotted as a function of its total number of molecules T_{1T}. The theoretical threshold is again located close to the equimolarity point, *i.e.* where $T_{1T} + T_{2T} \simeq S_T$. The average of the other target, $\langle T_2 \rangle$, shows instead a smooth sigmoidal profile with changing point in the vicinity of the threshold (Fig. 2b). The profiles of both averages are ultrasensitive around the threshold and the strength of the ultrasensitivity is controlled by the dissociation constants (see Fig. 2d). Small values of the dissociation constants mean steeper threshold responses and higher ultrasensitivity.

The common interaction between the targets and the sequestrant effectively correlates them (see e.g. refs[10,11,29]). When several molecules of a target are bound to the sequestrant, the propensity of binding for molecules of the second target is reduced. Additionally, when the total amount of molecules of the two targets globally exceeds the sequestrant, having a high number of one target molecules bound implies having fewer molecules of the other target in a complex. From this, it naturally follows that the two targets are negatively correlated via competition. To show this, we quantify such a correlation by means of the Pearson coefficient of the two targets, defined, as usual, as the ratio between the covariance and the product of the two standard deviations (σ_{T_1} and σ_{T_2}) as follows:

$$\rho \equiv \frac{\mathrm{cov}(T_1, T_2)}{\sigma_{T_1}\sigma_{T_2}}, \tag{22}$$

and investigate its dependence on the key parameters in the model. In a later section we will characterise the interaction of the two targets by means of mutual information[30] (as done in *e.g.* ref.[31] for a similar setup). In view of the applications of the present framework to biochemical systems, we focus on the dependence of the correlation on the targets abundances which are simpler to control in biological systems than the targets affinity for the sequestrant. Nonetheless, we report in the SI a detailed discussion of the dependence of the correlation on the dissociations constants. In Fig. 2c we plot the Pearson coefficient as a function of the total number of one of the targets (T_{1T}). We observe that the correlation starts close to zero for low target abundances and decreases close to the threshold $T_{1T}^* = S_T - T_{2T}$. For low dissociation constants, the Pearson coefficient displays a sigmoidal profile (red curve, Fig. 2c) with the drop located in proximity to the threshold. For high targets abundances, the correlation between the two targets saturates to a value close to -1 (the lower bound of the Pearson coefficient) and this behaviour is maintained also upon varying both targets abundances (Fig. 2e). The shape of the correlation as a function of the total number of one of the targets is affected in a different way by changes of the dissociation constants of the two targets. Keeping T_{2T} fixed, the steepness of the sigmoidal profile as a function of T_{1T} is mainly governed by the value of the dissociation constant of target 1, K_1^d, and increases with the decrease of the dissociation constant. Instead, K_2^d, the dissociation constant of target 2, whose abundance is not changed in the plot, affects the minimal value that the correlation asymptotically reaches as T_{1T} becomes large (see Fig. S3 in the SI). A lower dissociation constant K_2^d corresponds to a lower value of the minimal correlation (stronger negative correlation). However, this smoothens the correlation drop around the threshold and slightly influences its location. Figure 2e shows the behaviour of the Pearson coefficient when the abundances of both targets are independently changed. No global minimum is found and correlation is stronger (more negative) as the number of target molecules increases.

Extrinsic noise. To investigate the effects of external fluctuations, as done for the case of a single target species, we study the behaviour of the two targets system when the number of total sequestrant molecules is allowed to fluctuate between different realisations of the system. The probability distribution of S_T is again assumed to be a discretised Gaussian and the full probability distribution of the system is obtained as a weighed superposition by implementing the law of total probability. Having full analytical control on the system, the behaviour of the correlation in presence of extrinsic noise can be straightforwardly investigated and precisely quantified.

As a result, we observe that fluctuations in the amount of the sequestrant positively correlate the competing targets. This can be understood by considering the fact that the two target species interact with the same pool of sequestrant molecules. Then, fluctuations in the amount of the sequestrant S_T affect in the same way the two target species in each realisation of the extrinsic noise. For instance, if by an extrinsic fluctuation the amount of sequestrant is lower than the average, both target species will have a higher chance of being free from the sequestrant. This effectively induces a positive correlation between the targets, which can counterbalance the negative one due to competition and discussed in the previous section. The negative interference of these two opposite sources of correlation can result in basically uncorrelated systems. However, these two conflicting sources of correlation depend differently on the rates and the abundances. The positive correlation is felt the most in the proximity of the threshold. There, most of the sequestrant is bound and neither of the targets have many free molecules. Then, a change (say a decrease) in the sequestrant abundance directly reflects in a change (increase) in the number of free molecules of both targets. This is no longer the case when there is an excess of one target (e.g. large T_{1T}), for which the sequestrant is bound mostly to T_1 and a change in the sequestrant affects mostly the amount of free T_1 and less the amount of free T_2. When the sequestrant largely outnumbers the targets, changes in its abundance will have little impact on the targets, which will mostly be bound to the sequestrant anyway. From

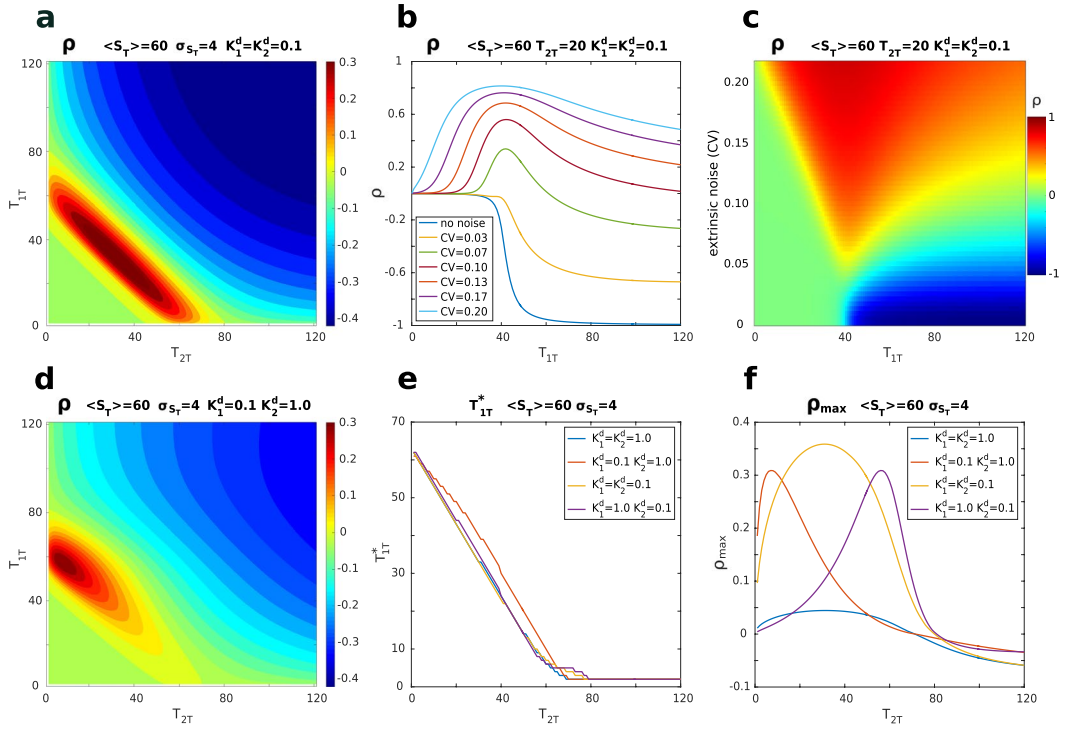


Figure 3. Correlation in presence of extrinsic noise. $\langle S_T \rangle = 60$, $\sigma_{S_T} = 4$ where not otherwise stated. (**a**,**d**) Contour plots of the correlation as a function of T_{1T} and T_{2T} for $K_1^d = K_2^d = 0.1$ (**a**) and $K_1^d = 0.1$, $K_2^d = 1.0$ (**d**). (**b**) Correlation as a function of T_{1T} for different levels of extrinsic noise. The blue line on the bottom corresponds to the pure intrinsic noise case, for the other lines σ_{S_T} assumes the values: 2, 4, 6, 8, 10, 12. $K_1^d = K_2^d = 0.1$, $T_{2T} = 20$. (**c**) Contour plot of the correlation as a function of T_{1T} and of the level of extrinsic noise. $K_1^d = K_2^d = 0.1$, $T_{2T} = 20$. The size of the step along T_{1T} is $\Delta T_{1T} = 1$, while the size of the step for the extrinsic noise level is $\Delta\sigma_T = 0.25$ ($\Delta CV = 4 \cdot 10^{-3}$). (**e**) Position (in terms of the value of T_{1T}) of the maximum of correlation, for a given T_{2T}, for different values of K_1^d and K_2^d. (**f**) Value of the maximum of correlation for a given T_{2T}, for different values of K_1^d and K_2^d.

the combination of these effects, correlation takes a non-trivial profile as a function of the abundances and can be positive for sizeable regions of the parameters space.

For a more quantitative discussion, we first focus on the correlation profile as a function of the total abundance of target 1, T_{1T}, keeping T_{2T} fixed. As shown in Fig. 3b,c, a low level of fluctuations on the number of sequestrant molecules opposes the negative correlation induced by competition. When the level of extrinsic noise is increased, the positive correlation induced by the sequestrant fluctuations can counterbalance competition, leading to a correlation close to zero. A further increase of extrinsic noise eventually induces a positive correlation between the targets. Because of the different dependence of the sources of correlation on the system features, the resulting profile displays a correlation maximum in the vicinity of the threshold, where the system is more sensitive to fluctuations of the sequestrant amount. This can also be observed by analysing the two-dimensional (varying both target abundances) profile of correlation shown in Fig. 3a,d. If we focus on slices with a fixed number of T_{2T} we see that for larger values of T_{2T} the maximum correlation is reached for lower total amounts of the T_{1T}, for which the system is still in the vicinity of the threshold. This is investigated in Fig. 3e where the abundance of target 1, for which the maximum correlation is attained (T_{1T}^*), is plotted as a function of the abundance of the second target T_{2T}. When the dissociation constants of the two targets are equal, such position of the maximum decreases linearly (with slope -1) with the increase of the abundance of the second target until the overall number of target molecules ($T_{1T} + T_{2T}$) exceeds that of the sequestrant and eventually saturates at low levels. This is more evident for systems with a strong affinity for the sequestrant where the threshold is sharp. The overall trend is preserved when the two targets have different affinities for the sequestrant but the linear decrease is less pronounced when the second target (T_2) has a lower affinity with the sequestrant. When the overall number of targets starts exceeding that of the sequestrant, the correlation decreases and the profile of the location of the maximum depends on the specific dissociation constants. For systems with higher levels of extrinsic noise, such profiles beyond the threshold may be non monotonic, as shown in Fig. S6e in the SI. The value of the maximum of correlation for a given amount of the second target: $\rho_{max}(T_{2T}) \equiv \rho(T_{1T}^*, T_{2T})$ is plotted in Fig. 3f. Such plot can be used to characterise the location and the value of the global maximum of correlation, which is seen when considering the two-dimensional perspective (varying both targets abundances) reported in Fig. 3a,d. Indeed, the maximum of $\rho_{max}(T_{2T})$ corresponds to the global maximum. As expected, it is located in proximity of the region of equimolarity between targets and sequestrant ($T_{1T} + T_{2T} \simeq S_T$). The maximizing individual target abundances are set by

their respective affinity with the sequestrant: the higher the affinity of a target, the higher its abundance at the maximum. In other words, as shown in Fig. 3f, when the affinities are equal, $K_1^d = K_2^d$, the peak is reached for $T_{1T} \simeq T_{2T} \simeq S_T/2$ and when $K_1^d < K_2^d$ the peak moves towards higher values of T_{1T} and smaller values of T_{2T}. The value of the global maximum is also controlled by the dissociation constants, and takes higher values for lower dissociation constants (higher affinities). Qualitatively similar behaviours are recovered by varying the average amount of sequestrant S_T and the intensity of the noise. We report this in the SI.

Mutual Information. Having access to the explicit expression of the joint probability allows us to characterise the coupling between the competing targets beyond the Pearson coefficient and to study the mutual information between the competing targets. In this setting, mutual information $I(T_1, T_2)$ quantifies the amount of information, measured in bits, that could be obtained on one target by measuring the other target. Its definition[30] reads:

$$I(T_1, T_2) = \sum_{T_1, T_2} P(T_1, T_2) \log \frac{P(T_1, T_2)}{P(T_1)P(T_2)}, \tag{23}$$

where $P(T_1, T_2)$ is the joint probability distribution and $P(T_1) = \sum_{T_2} P(T_1, T_2)$ the marginal one. Mutual information is never negative, which implies that both positive and negative correlations result in positive mutual information (there is no distinction between correlated and anti-correlated systems) and, for discrete probabilities, is bounded by the logarithm of the number of states accessible to the system. In Fig. 4a we explore the profile of the mutual information between the targets upon varying both targets' abundances and in Fig. 4b upon variation of T_{1T} (in analogy with Fig. 2c,e). The mutual information profile starts at zero and begins to increase in the vicinity of the threshold. As for the Pearson correlation, when plotted as a function of T_{1T}, the dissociation constant K_1^d governs the steepness of the profile, while K_2^d mainly controls the maximum value of mutual information that can be achieved, (see the SI for additional plots). In Fig. 4d–f we show the effect of extrinsic noise on mutual information, in analogy with Fig. 3. The profiles are qualitatively similar to the ones of the Pearson coefficient and offer the same insights. However, mutual information captures further non-linear behaviour in the interaction between two targets. This becomes more evident if we convert correlation in units of $-\frac{1}{2}\log[1 - \rho^2]$ where ρ is Pearson correlation coefficient, which corresponds to the mutual information that two jointly Gaussian variables of correlation ρ would have. This quantity is plotted as full lines in Fig. 4. For the parameters investigated in the previous sections, mutual information takes larger values than the one associated with a Gaussian approximation with the measured Pearson coefficient, showing that the co-dependence between the targets goes beyond the correlation captured by the Pearson coefficient. There is one notable difference between the Pearson correlation and the mutual information, which is seen in the absence of extrinsic noise when one of the target is in large excess (this is not noticeable for the parameters explored in the previous sections) and shown in Fig. 4c. Fixing the abundance of one target and varying the other one (as done in Fig. 4a) correlation is monotonic and reaches a plateau, whereas mutual information displays a maximum when the target starts being in excess. This maximum becomes more peaked as the difference in magnitude of the two dissociation constants increases. The plateau in the Pearson coefficient for an excess target arises from the combined trends of the covariance and the standard deviations. The covariance, similarly to mutual information, is non monotonic and displays a peak. Its decrease after the peak is compensated by the decrease in the standard deviations of the two targets whose marginal distributions get narrower as the system gets more saturated. This discrepancy between the mutual information and the correlation profile can be traced back to the marked non-Gaussianity of the joint target distribution past saturation. In this region the distribution gets strongly peaked, a scenario which has been related to instances of mutual information lower than the one predicted by Gaussian approximations (see *e.g.* ref.[32]). Finally, it is important to note how mutual information (and correlations) are strongly damped by the presence of extrinsic noise as a consequence of the negative interference between the negative correlation due to competition and the positive one induced by fluctuations in the sequestrant.

Extrinsic noise effects on competitive inhibition kinetics. Let us now explore how the key features we have identified in the minimal model carry over to more complex settings. To this aim, let us consider the case of competitive inhibition in an enzymatic reaction. The system is made of an enzyme that, when free (T_F), can bind to a substrate (which we assume to be at fixed concentration c_s) and form the active enzyme T_A from which the product is made. The free enzyme can also bind to an inhibitor (S) that prevents it from combining with the substrate and becoming active. In the language of the previous sections, the inhibitor plays the role of the sequestrant and the enzyme that of the target. The reaction network that describes this system is defined by

$$T_F + S \underset{k_-}{\overset{k_+}{\rightleftharpoons}} \overline{TS} \tag{24}$$

$$\text{Substrate} + T_F \underset{k_r}{\overset{k_f c_s}{\rightleftharpoons}} T_A \rightarrow \text{Product} + T_F \tag{25}$$

We consider the case in which both substrate and product concentrations are large so that their fluctuations are negligible and focus on the stochastic dynamics of the enzyme (target) and the inhibitor (sequestrant). Their total amounts are conserved, defining the following conservation laws:

$$T_T = T_F + T_A + \overline{TS} = const, \tag{26}$$

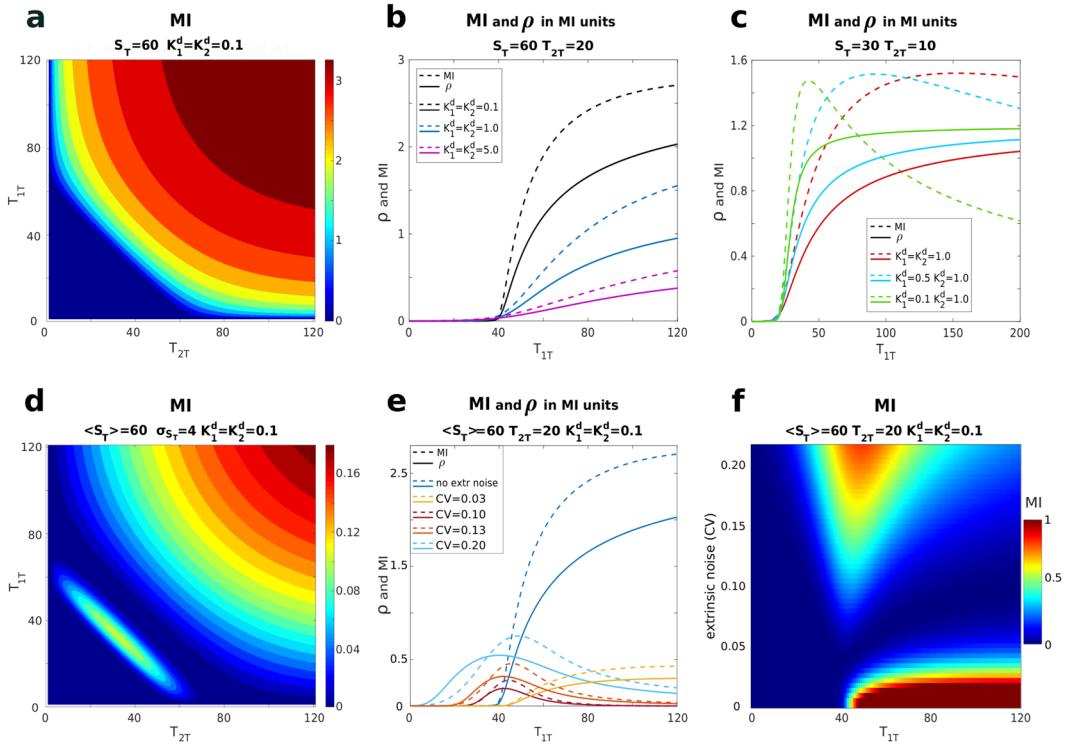


Figure 4. Mutual information (bits) (**a**–**c**) pure intrinsic noise case. (**d**–**f**) Mutual information in presence of extrinsic noise. (**a**) Contour plot of the mutual information between T_1 and T_2 as a function of T_{1T} and T_{2T}. The parameters are as in Fig. 2e: $S_T = 60$, $K_1^d = K_2^d = 0.1$. (**b**) The dashed lines refer to mutual information between T_1 and T_2 as a function of T_{1T} for a fixed value of T_{2T} and different values of K_1^d and K_2^d. The solid lines are the Pearson correlation expressed in units of $-\frac{1}{2}\log[1-\rho^2]$ as discussed in the text. The values of the parameters correspond to the ones of Fig. 2d: $S_T = 60$, $T_{2T} = 20$. K_1^d and K_2^d are always equal and assume the values: 5.0 (purple), 1.0 (cyan) and 0.1 (black). (**c**) Mutual information and correlation past saturation. Again the dashed lines refer to mutual information and the solid ones are the Pearson correlation expressed in units of $-\frac{1}{2}\log[1-\rho^2]$ the values of the parameters are now different, with $S_T = 30$, $T_{2T} = 10$ and the dissociation constants are reported in the inset. (**d**) Contour plot of the mutual information as a function of T_{1T} and T_{2T} for $K_1^d = K_2^d = 0.1$ in presence of extrinsic noise with $\langle S_T \rangle = 60$ and $\sigma_{S_T} = 4$ (same parameters as Fig. 3a). (**e**) Mutual information and correlation as a function of T_{1T} for different levels of extrinsic noise (see the legend). $K_1^d = K_2^d = 0.1$, $T_{2T} = 20$ (same parameters as Fig. 3b) (**f**) Contour plot of the mutual information as a function of T_{1T} and of the level of extrinsic noise. $K_1^d = K_2^d = 0.1$, $T_{2T} = 20$.

$$S_T = S + \overline{TS} = const. \tag{27}$$

As a consequence of the conservation laws, the number of free stochastic variables for this system is reduced to 2. Here, we focus on the active enzyme T_A from which the product is formed and the inhibitor-bound enzyme $\overline{TS}$. For the sake of clarity we consider the case of quasi-equilibrium dynamics in which the product formation reaction is much slower than the others. However, the general case can be addressed with minor modifications since the master equation's structure and solution is left basically unchanged (see SI). As in the cases considered in the previous sections (and shown in the SI) the steady-state distribution can be analytically derived and the effect of fluctuations in the number of the sequestrant (inhibitor) molecules investigated. We again model the fluctuations by a discretised Gaussian distribution. For comparison with the minimal case, we plot the average behaviour of the active enzyme T_A and the inhibited one $\overline{TS}$ as a function of the total number of enzymes (targets) in Fig. 5a. Both quantities present the threshold-like profile typical of molecular sequestration. As for the minimal case, extrinsic noise affects the shape of the probability distributions and Fig. 5b shows the occurrence of bimodality for both the inhibited complex (sequestrant -target) and the active enzyme (bound to the substrate). However, it is important to note that, due to the presence of an additional chemical reaction, with its inherent stochasticity, the probability distributions are smoother with respect to the minimal model. As a consequence, stronger affinities (lower dissociation constants) are required to ensure robust bimodal phenotypes, especially for the active enzyme (see Fig. 5c,d). In analogy with the minimal model, the distributions can be tuned from being bimodal to unimodal and vice versa by modulating the level of extrinsic noise. The effects of the extrinsic noise level and of the dissociation constant ($K_d = k_-/k_+$) on bimodality are in qualitative agreement with the minimal model. Indeed, for both T_A and $\overline{TS}$, as extrinsic noise is increased, the range of bimodality over the values of T_T becomes wider. Similarly, a low value of the dissociation constant, i.e. a steeper threshold, favours the presence of bimodality (Fig. 5c,d).

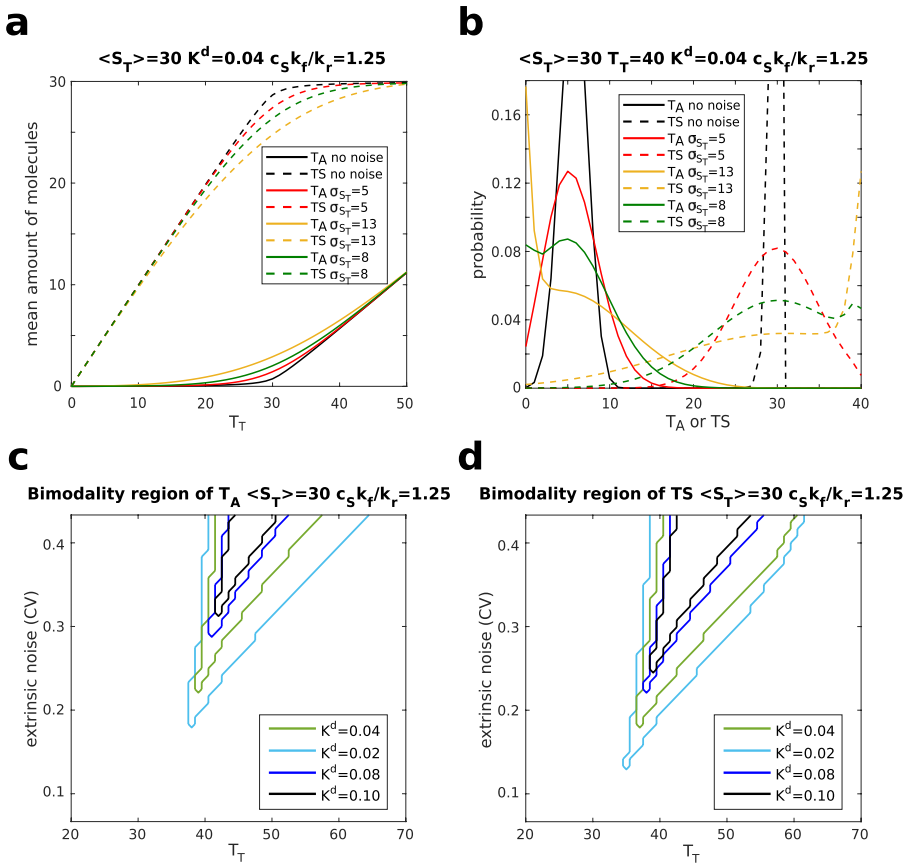


Figure 5. Extrinsic noise effects on competitive inhibition kinetics. (**a**) $\langle T_A \rangle$ (solid lines) and $\langle \overline{TS} \rangle$ (dashed lines) vs T_T for different levels of extrinsic noise. For all the curves $K^d = 0.04$, $c_S k_f/k_r = 1.25$ and $\langle S_T \rangle = 30$. The black curve is the pure intrinsic noise case, while the other three correspond to a standard deviation of the Gaussian distribution of S_T equal to 5 (red), 8 (green) and 13 (yellow). (**b**) Examples of probability distribution of T_A (solid lines) and $\overline{TS}$ (dashed lines) in presence of extrinsic noise. $\langle S_T \rangle = 30$, $T_T = 40$, $K^d = 0.04$, $c_S k_f/k_r = 1.25$, σ_{S_T} assumes the values: 5 (red), 8 (green) and 13 (yellow). (**c**,**d**) Plots of the bimodality region of the marginal distributions of T_A (**c**) and $\overline{TS}$ (**d**) for different values of K^d as a function of T_T and extrinsic noise. Bimodal distributions are present for parameters inside the areas delimited by the plotted lines. $\langle S_T \rangle = 30$, $c_S k_f/k_r = 1.25$, K^d assumes the values: 0.02 (light blue), 0.04 (green), 0.08 (blue) and 0.10 (black). The size of the step along T_T is $\Delta T_T = 1$, while the size of the step for the extrinsic noise level is $\Delta\sigma_T = 0.25$ ($\Delta CV = 8{\cdot}10^{-3}$). For each point defined by these steps, the distribution $P(T)$ was computed analytically and the number of its maxima was evaluated. Note that the step-like features in the plot are due to the discreteness of T_T.

Correlations. In the previous sections we showed that two target species competing for the same sequestrant can become positively correlated under the effect of extrinsic noise on the sequestrant. We now analyse the dynamics of two species of enzymes (targets) inhibited by the same sequestrant. The corresponding reaction network is described by the following reactions:

$$T_{1F} + S \underset{k_-^1}{\overset{k_+^1}{\rightleftharpoons}} \overline{T_1 S} \tag{28}$$

$$\text{Substrate} + T_{1F} \underset{k_r^1}{\overset{k_f^1 c_s}{\rightleftharpoons}} T_{1A} \rightarrow \text{Product} + T_{1F} \tag{29}$$

$$T_{2F} + S \underset{k_-^2}{\overset{k_+^2}{\rightleftharpoons}} \overline{T_2 S} \tag{30}$$

$$\text{Substrate} + T_{2F} \underset{k_r^2}{\overset{k_f^2 c_s}{\rightleftharpoons}} T_{2A} \rightarrow \text{Product} + T_{2F}. \tag{31}$$

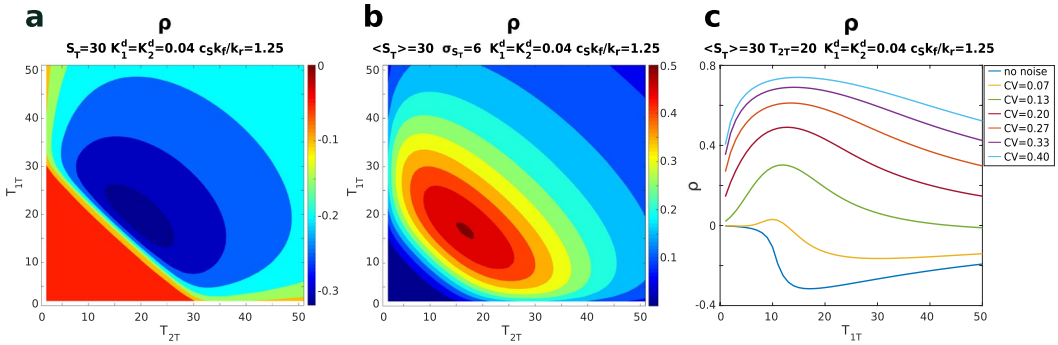


Figure 6. Correlations for competitive inhibition. (**a**) Pure intrinsic noise case: contour plot of the Pearson correlation between the active enzymes T_{1A} and T_{2A} as a function of the total number of enzymes T_{1T} and T_{2T}. $S_T = 30$, $K_1^d = K_2^d = 0.04, c_S k_f^1/k_r^1 = c_S k_f^2/k_r^2 = 1.25$. (**b**) Extrinsic noise: contour plot of the Pearson correlation between T_{1A} and T_{2A} as a function of T_{1T} and T_{2T}. $\langle S_T \rangle = 30, \sigma_{S_T} = 6, K_1^d = K_2^d = 0.04, c_S k_f^1/k_r^1 = c_S k_f^2/k_r^2 = 1.25$. (**c**) Correlation between the active enzymes as a function of T_{1T} for different levels of extrinsic noise. The blue line on the bottom corresponds to the pure intrinsic noise case, for the other lines σ_{S_T} assumes the values: 2, 4, 6, 8, 10, 12. $\langle S_T \rangle = 30$, $T_{2T} = 20, K_1^d = K_2^d = 0.04, c_S k_f^1/k_r^1 = c_S k_f^2/k_r^2 = 1.25$.

Both the unbound targets, T_{1F} and T_{2F}, can be activated (into T_{1A} and T_{2A}) by binding to a substrate with rates proportional to the substrate concentration and to their intrinsic activation rates k_f^1 and k_f^2. For the sake of simplicity, we assume that the substrate is the same for both the enzymes. The active targets T_{1A} and T_{2A} can be deactivated with rates k_r^1 and k_r^2 respectively. The unbound, inactive targets can be sequestered by the inhibitor molecule S into the inactive complexes $\overline{T_1 S}$ and $\overline{T_2 S}$, the rates of these reactions are k_+^1 and k_+^2. Finally, the complexes $\overline{T_1 S}$ and $\overline{T_2 S}$ can dissociate with rates k_-^1 and k_-^2 respectively. We focus on the correlations that are induced on the active enzymes by the competitive interactions with the inhibitor. In the absence of noise, the two enzymes are negatively correlated, especially in proximity to the threshold (see Fig. 6a). The activation reaction introduces an additional source of stochasticity. As a consequence, the correlations between the active enzymes (T_{1a} and T_{2a}) are generally lower than the ones between the sequestered, inactive enzymes ($\overline{T_1 S}$ and $\overline{T_2 S}$) and of the ones of the minimal model. This additional layer of stochasticity also affects the profile of the Pearson correlation coefficient, which now displays a minimum around the threshold instead of the plateau shown for the minimal model (see Fig. 2c,e). This can be intuitively understood considering the fact that, in the minimal model, the plateau arose because the standard deviation of the targets roughly followed the behaviour of the covariance. The additional stochasticity related to the activation dynamics is not affected much by the sequestration dynamics, thus the standard deviation keeps growing with the total number of targets not compensating the slow-down of the covariance, giving rise to the minimum.

Let us now turn our attention to the case in which extrinsic noise introduces fluctuations in the number of sequestrant molecules. As in the minimal model, such fluctuations can induce positive correlations between the active target, especially around the threshold (see Fig. 6b,c). Again, since the interaction with the sequestrant is diluted by the presence of an additional stochastic reaction, the correlation tends to be generally lower.

Discussion

Sequestration models are good approximations of many processes in nature, especially in biological systems. Although fairly studied, this kind of models still show novel features when their stochasticity is addressed. In this contribution we have considered how different sources of noise can affect the properties of a minimal model with sequestration dynamics. We started by analysing again the deterministic version of this system, where, as expected[6], two species with fixed abundances, here called sequestrant and target, mutually inhibiting each other, give rise to a threshold response. We then moved to a stochastic version taking into account intrinsic noise due to the probabilistic nature of the interactions in the system. Via an exact analytical solution we checked that the threshold behaviour is maintained for the average values of the targets in this context as well. We then made use of the derived set-up to investigate the effects introduced by an extrinsic source of noise, which may be due to fluctuating environments, be they global external surroundings or simply other cell molecular networks interacting with the one under consideration. We have shown that extrinsic fluctuations in the total number of one of the inhibitors (the sequestrant), in combination with the threshold profile may result in a bimodal distribution of the other one (the target). This means that, close to the threshold, the extrinsic variability is channelled by the nonlinear sequestration mechanism into two main outputs corresponding to a repressed and an unrepressed state. This result identifies two minimal ingredients giving rise to bimodal profiles without the need of other regulatory links, i.e., (i) titrative (sequestration) interactions and (ii) extrinsic source of noise.

We have then extended the model to the case of two species competitively binding to a third one. Once again, taking advantage of the simplicity of the minimal model we have derived the analytical solution of the master equations to quantitatively address another key feature of the sequestration dynamics: the correlation induced by the competition for binding to the same sequestrant. We have shown that, without extrinsic fluctuations on the total abundance of the sequestrant, a significant negative correlation is present when the total amount of targets outnumbers the sequestrant. Together with this expected effect of the competition, we have identified non-trivial

trends of the correlation in presence of an extrinsic source of noise. Fluctuations in the total sequestrant amount affect in a similar way the free amounts of both targets. As a result, despite the competitive dynamics of the targets, positive correlations are induced by the extrinsic fluctuations of the common sequestrant.

These findings, obtained in a highly simplified description, have highlighted some key features that are likely to occur in more detailed models describing specific biochemical settings. To probe this, we have investigated the more complex case of competitive inhibition in enzymatic kinetics. This system features an additional reaction downstream of the sequestration interaction, which, with its stochasticity, partially dilutes the coupling between competitors, smoothening the distribution profiles and partially hindering the appearance of bimodality. However, we found that, with the due adjustments, extrinsic noise on the inhibitor can induce bimodal distributions and positive correlations between the enzymes. This suggests that the possible constructive role of extrinsic noise we have highlighted in the minimal model may be at play in more complex systems of biological relevance. Models of these systems would be more complicated, including different intermediate steps in the reactions of mediation between sequestrant and target or using saturation functions for production rates. Nonetheless, whenever sequestration dynamics are present, we expect our findings to be of relevance for such complex systems as well, where they may serve as guidelines to identify the most relevant parameters affecting correlations and the presence of bimodal distributions (as for instance in the cases discussed in refs[1–14]). A system for which our results will be of particular importance is the post-transcriptional regulation achieved by short non-coding RNAs interacting with messenger transcripts in eukaryotic cells[10,13,15].

Methods

Solution of the two-targets model. As presented in the main text, from eqs (14) and (15) with the conservation laws (16–18) one can write the master equation that describes the dynamics of the system:

$$\begin{aligned}\frac{dP(T_1,S,t)}{dt} &= (\overline{T_1S}+1)k_-^1P(T_1-1,S-1)+(T_1+1)(S+1)k_+^1P(T_1+1,S+1)\\ &+(\overline{T_2S}+1)k_-^2P(T_1,S-1)+(T_2+1)(S+1)k_+^2P(T_1,S+1)\\ &-(\overline{T_1S}k_-^1+T_1Sk_+^1+\overline{T_2S}k_-^2+T_2Sk_+^2)P(T_1,S).\end{aligned} \tag{32}$$

Making use again of the conservation laws (16–18), the master equation can be written in terms of T_1 and S only:

$$\begin{aligned}\frac{dP(T_1,S,t)}{dt} &= (T_{1T}-T_1+1)k_-^1P(T_1-1,S-1)\\ &+(T_1+1)(S+1)k_+^1P(T_1+1,S+1)\\ &+(S_T-S-T_{1T}+T_1+1)k_-^2P(T_1,S-1)\\ &+(T_{2T}-S_T+S+T_{1T}-T_1+1)(S+1)k_+^2P(T_1,S+1)\\ &-[(T_{1T}-T_1)k_-^1+T_1Sk_+^1+(S_T-S-T_{1T}+T_1)k_-^2\\ &+(T_{2T}-S_T+S+T_{1T}-T_1)Sk_+^2]P(T_1,S).\end{aligned} \tag{33}$$

In order to find its steady-state solution, we first notice that detailed balance holds. By using detailed balance, the following relations for the steady-state probability are obtained:

$$k_-^1(T_{1T}-T_1)P(T_1,S)=k_+^1(T_1+1)(S+1)P(T_1+1,S+1), \tag{34}$$

$$k_-^2(S_T-T_{1T}-S+T_1)P(T_1,S)=k_+^2(S+1)(T_{1T}+T_{2T}-S_T+S-T_1+1)P(T_1,S+1). \tag{35}$$

Dividing eq. (35) by $P(T_1)$, we can write a recursive relation for the probability of having S free molecules conditioned on the fact that a generic number T_1 of molecules are free:

$$P(S+1|T_1)=\frac{k_-^2}{k_+^2}\frac{S_T-T_{1T}-S+T_1}{(S+1)(T_{1T}+T_{2T}-S_T+S-T_1+1)}P(S|T_1). \tag{36}$$

In this system, as discussed in the main text (and shown for an example in Fig. 2f), the minimal number of free molecules of S can either be $S_{min}=0$ if $S_T\leq T_{1T}+T_{2T}-T_1$, or $S_{min}=S_T-T_{1T}-T_{2T}+T_1$ when $S_T>T_{1T}+T_{2T}-T_1$. Equation (36) can be used to recursively write the expression of $P(S|T_1)$ starting from the probability $P(S_{min}|T_1)$:

$$\begin{aligned}P(S|T_1) &= \left(\frac{k_-^2}{k_+^2}\right)^{S-S_{min}}\frac{(S_T-T_{1T}+T_1-S_{min})!}{(S_T-T_{1T}+T_1-S)!}\frac{S_{min}!}{S!}\\ &\cdot\frac{(S_{min}+T_{1T}+T_{2T}-S_T-T_1)!}{(S+T_{1T}+T_{2T}-S_T-T_1)!}P(S_{min}|T_1).\end{aligned} \tag{37}$$

Writing $P(T_1,S)=P(S|T_1)P(T_1)$, we can insert eq. (37) in the first relation of detailed balance, eq. (34). Defining $S_{min}\equiv S^*_{min}+T_1$, the obtained relation can be explicited for the marginal probability $P(T_1+1)$ as follows:

$$P(T_1+1)=\frac{k_-^1}{k_+^1}\frac{T_{1T}-T_1}{(T_1+1)(S^*_{min}+T_1+1)}\frac{P(S^*_{min}+T_1|T_1)}{P(S^*_{min}+T_1+1|T_1+1)}P(T_1). \tag{38}$$

Recursively applying eq. (38), we can write the analytical expression of the marginal probability $P(T_1)$:

$$P(T_1)=\left(\frac{k_-^1}{k_+^1}\right)^{T_1-T_{1min}}\frac{(T_{1T}-T_{1min})!T_{1min}!(S^*_{min}+T_{1min})!}{(T_{1T}-T_1)!T_1!(S^*_{min}+T_1)!}\frac{P(S^*_{min}+T_{1min}|T_{1min})}{P(S^*_{min}+T_1|T_1)}P(T_{1min}), \tag{39}$$

where $P(T_{min})$ is the probability of having the minimal number allowed of free T_1. T_{min} is defined by the total amount of S and T_1. According to Table 1, the two possible cases can either be $T_{1min}=0$ when $S_T\geq T_{1T}$, which means that in principle all the molecules of T_1 can be bound to molecules of S, or $T_{1min}=T_{1T}-S_T$ when $S_T<T_{1T}$, meaning that even if all the molecules of S are bound to molecules of T_1, T_{1min} molecules are free. When $T_{1min}>0$ we have that $S^*_{min}+T_{1min}=S_T-T_{1T}-T_{2T}+T_{1T}-S_T=-T_{2T}\leq 0$, then the conditional probability at the numerator in eq. (39) is $P(0|T_{1min})$. Conversely, if $T_{1min}=0$, the conditional probability can either be $P(0|T_{1min})$ or $P(S^*_{min}|T_{1min})$, depending on the sign of S^*_{min}. The probability $P(T_1, S)$ can be reconstructed by using the definition of joint probability $P(T_1, S)=P(S|T_1)P(T_1)$, which leads to the following expression:

$$P(T_1,S)=\left(\frac{k_-^1}{k_+^1}\right)^{T_1}\left(\frac{k_-^2}{k_+^2}\right)^{S-T_1}\frac{P(T_{1min},S^*_{min}+T_{1min})}{(S_T-T_{1T}+T_1-S)!S!(T_{1T}+T_{2T}-S_T-T_1+S)!(T_{1T}-T_1)!T_1!}, \tag{40}$$

with $P(T_{1min}, S^*_{min}+T_{1min})$ given by the normalization of the probability distribution:

$$\begin{aligned}P(T_{1min},S^*_{min}+T_{1min})^{-1} &= \sum_{T_1=T_{1min}}^{T_{1T}}\sum_{S=S^*_{min}+T_{1min}}^{S_T}\left(\frac{k_-^1}{k_+^1}\right)^{T_1}\left(\frac{k_-^2}{k_+^2}\right)^{S-T_1}\\ &\cdot\frac{1}{(S_T-T_{1T}+T_1-S)!S!(T_{1T}-T_1)!T_1!}\\ &\cdot\frac{1}{(T_{1T}+T_{2T}-S_T-T_1+S)!}.\end{aligned} \tag{41}$$

Also for the joint probability if $T_{1min}>0$, then $S^*_{min}+T_{1min}\equiv 0$, while when $T_{1min}=0$, the value of S^*_{min} is given by its definition $S^*_{min}\equiv \max(S_T-T_{1T}-T_{2T}, 0)$.

References

1. Pickart, C. M. & Eddins, M. J. Ubiquitin: structures, functions, mechanisms. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* **1695**, 55–72 (2004).
2. Klein, D. E., Nappi, V. M., Reeves, G. T., Shvartsman, S. Y. & Lemmon, M. A. Argos inhibits epidermal growth factor receptor signalling by ligand sequestration. *Nature* **430**, 1040–1044 (2004).
3. Yanagita, M. BMP antagonists: Their roles in development and involvement in pathophysiology. *Cytokine and Growth Factor Reviews* **16**, 309–317 (2005).
4. Van Doren, M., Ellis, H. M. & Posakony, J. W. The Drosophila extramacrochaetae protein antagonizes sequence-specific DNA binding by daughterless/achaete-scute protein complexes. *Development* **113**, 245–255 (1991).
5. Benezra, R., Davis, R. L., Lockshon, D., Turner, D. L. & Weintraub, H. The protein id: A negative regulator of helix-loop-helix dna binding proteins. *Cell* **61**, 49–59 (1990).
6. Buchler, N. E. & Louis, M. Molecular titration and ultrasensitivity in regulatory networks. *Journal of molecular biology* **384**, 1106–1119 (2008).
7. Grigorova, I. L., Phleger, N. J., Mutalik, V. K. & Gross, C. A. Insights into transcriptional regulation and σ competition from an equilibrium model of rna polymerase binding to dna. *Proceedings of the National Academy of Sciences* **103**, 5332–5337 (2006).
8. Mukherji, S. *et al*. MicroRNAs can generate thresholds in target gene expression. *Nat Genet* **43**(9), 854–859 (2011).
9. Ala, U. *et al*. Integrated transcriptional and competitive endogenous RNA networks are cross-regulated in permissive molecular environments. *Proc Natl Acad Sci USA* **110**(18), 7154–7159 (2013).
10. Bosia, C., Pagnani, A. & Zecchina, R. Modelling competing endogenous RNA networks. *PLoS One* **8**(6), e66609 (2013).
11. Figliuzzi, M., Marinari, E. & De Martino, A. MicroRNAs as a selective channel of communication between competing RNAs: a steady-state theory. *Biophysical Journal* **104**(5), 1203–1213 (2013).
12. Lai, X., Wolkenhauer, O. & Vera, J. Understanding microRNA-mediated gene regulatory networks through mathematical modelling. *Nucleic acids research* **44**, 6019–6035 (2016).
13. Bosia, C. *et al*. Rnas competing for microRNAs mutually influence their fluctuations in a highly non-linear microRNA-dependent manner in single cells. *Genome Biology* **18**, 37, https://doi.org/10.1186/s13059-017-1162-x (2017).
14. Rotem, E. *et al*. Regulation of phenotypic variability by a threshold-based mechanism underlies bacterial persistence. *Proc. Natl. Acad. Sci.* **107**, 12541 (2010).
15. Del Giudice, M., Bo, S., Grigolon, S. & Bosia, C. On the role of extrinsic noise in microRNA-mediated bimodal gene expression. *Plos Computational Biology* **14**(4), e1006063 (2018).
16. Swain, P., Elowitz, M. & Siggia, E. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proceedings of the National Academy of Sciences of the United States of America* **99**(20), 12795–12800 (2002).
17. Elowitz, M., Levine, A., Siggia, E. & Swain, P. Stochastic gene expression in a single cell. *Science* **297**(5584), 1183–1186 (2002).
18. Tsimring, L. S. Noise in biology. *Reports on Progress in Physics* **77**, 026601 (2014).
19. Munsky, B., Neuert, G. & Van Oudenaarden, A. Using gene expression noise to understand gene regulation. *Science* **336**, 183–187 (2012).
20. Bessarabova, M. *et al*. Bimodal gene expression patterns in breast cancer. *BMC Genomics* **11**(Suppl I), S8 (2010).
21. Shalek, A. K. *et al*. Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. *Nature* **498**, 236–240 (2013).

22. Goldbeter, A. & Koshland, D. E. Sensitivity amplification in biochemical systems. *Quarterly reviews of biophysics* **15.3**, 555–591 (1982).
23. Van Kampen, N. G. *Stochastic processes in physics and chemistry*, vol. 1 (Elsevier, 1992).
24. Gillespie, D. T. A rigorous derivation of the chemical master equation. *Physica A: Statistical Mechanics and its Applications* **188**, 404–425 (1992).
25. Abramowitz, M. & Stegun, I. *Handbook of Mathematical Functions*. (Dover Books on Mathematics, 1965).
26. Gradshteyn, I. & Ryzhik, I. *Table of Integrals, Series and Products*. (Academic Press, 1994).
27. McQuarrie, D. A. Stochastic approach to chemical kinetics. *Journal of applied probability* **4**, 413–478 (1967).
28. Thomas, P., Popovic, N. & Grima, R. Phenotypic switching in gene regulatory networks. *Proceedings of the National Academy of Sciences* **111**, 6994–6999, https://doi.org/10.1073/pnas.1400049111 (2014).
29. Martirosyan, A., Marsili, M. & De Martino, A. Translating ceRNA susceptibilities into correlation functions. *Biophysical Journal* **113**, 206–213 (2017).
30. Cover, T. M. and Thomas, J. A. *Elements of information theory* (John Wiley & Sons, 2012).
31. Martirosyan, A., De Martino, A., Pagnani, A. & Marinari, E. ceRNA crosstalk stabilizes protein expression and affects the correlation pattern of interacting proteins. *Scientific Reports* **7**, 43673 (2017).
32. Foster, D. & Grassberger, P. Lower bounds on mutual information. *Phys. Rev. E* **83**, 010101 (2011).

Acknowledgements

S. B. acknowledges financial support from the Swedish Science Council (Vetenskapsrådet) under the Grant No. 638-2013-9243. S. G. acknowledges the Francis Crick Institute which receives its core funding from Cancer Research UK (FC001317), the UK Medical Research Council (FC001317) and the Wellcome Trust (FC001317). M. D. G. and C. B. wish to thank Nordita for hospitality during part of this work, which was partially supported by the Swedish Science Council (Vetenskapsrådet) under the Grant No. 621-2012-2982.

Author Contributions

All the authors intellectually conceived the study. S.B. and C.B. planned the study. M.D.G. and S.B. performed the analysis and obtained the results. All the authors drafted and reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-28647-9.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2018

Reproduced with permission of copyright owner. Further reproduction prohibited without permission.