Molecular switch architecture drives response properties

Khem Raj Ghusinga[a,b,c], Roger D. Jones[d], Alan M. Jones[a,c,†], and Timothy C. Elston[b,c,†]

[a]Department of Biology, University of North Carolina at Chapel Hill, NC, USA
[b]Department of Pharmacology, University of North Carolina at Chapel Hill, NC, USA
[c]Computational Medicine Program, University of North Carolina at Chapel Hill, NC, USA
[d]Center for Complex Systems and Enterprises, Stevens Institute of Technology, Hoboken, NJ, USA
[†]Correspondence: alan_jones@bio.unc.edu, timothy_elston@med.unc.edu

Abstract

A wide range of signaling pathways consist of molecular switches, which are proteins that transition between two states—*on* and *off*. Signal transduction occurs when a stimulus activates its cognate receptor, which then turns on a downstream switch by one of the following mechanisms: activation, where the transition rate from the *off* state to the *on* state increases; derepression, where the transition rate from the *on* state to the *off* state decreases; and concerted, where both activation and derepression operate simultaneously. Here we use mathematical modeling to compare and contrast these signaling mechanisms in terms of their dose response curves, response times, and abilities to process upstream fluctuations. Our analysis elucidates two general principles. First, activation makes the switch response more sensitive than the receptor response whereas derepression does the opposite. Second, activation operates by speeding up the switch response, as opposed to derepression which slows it down. Consequently, the concerted mechanism is not only able to align the switch dose response with that of the receptor but also to decouple the response time from the stimulus level. Interestingly, activation and derepression perform similarly in terms of processing upstream fluctuations, but better than the concerted mechanism. In addition to these three metrics, we also examine the effect of receptor removal on switches governed by activation and derepression. We find that if inactive (active) receptors are preferentially removed then activation (derepression) leads to a sustained response whereas derepression (activation) adapts. In total, we show how a molecular switch architecture governs its response properties. We also discuss biological relevance of these findings.

Introduction

Many molecules involved in intracellular signaling pathways act as molecular switches. These are proteins that can be temporarily modified to transition between two conformations, one corresponding to an *on* (active) state and another to an *off* (inactive) state. Two prominent examples of such switches are proteins that are modified by phosphorylation and dephosphorylation and GTPases that bind nucleotides. For phosphorylation-dephosphorylation cycles, it is common for the covalent addition of a phosphate by a kinase to cause activation of the modified protein. A phosphatase removes the phosphate to turn the protein *off*. In the GTPase cycle, the protein is *on* when bound to Guanosine triphosphate (GTP) and *off* when bound to Guanosine diphosphate (GDP). The transition from the GDP bound state to the GTP bound state requires nucleotide exchange, whereas the transition from GTP bound to GDP bound state is achieved via hydrolysis of the γ phosphate on GTP. The basal rates of nucleotide exchange and hydrolysis are often small. These reaction rates are increased several fold by Guanine Exchange Factors (GEFs) and GTPase Accelerating Proteins (GAPs), respectively [1, 2].

A signaling pathway is often initiated upon recognition of a stimulus by its cognate receptor, which then activates a downstream switch. In principle, a switch may be turned *on* by *at least* two mechanisms: a) by increasing the transition rate from the *off* state to the *on* state, and b) by decreasing the transition rate from the *on* state to the *off* state. We term these mechanisms *activation* and *derepresison*, respectively. Examples of both these mechanisms are found in the GTPase cycle. In animals, signaling through many pathways is initiated by G protein coupled receptors (GPCRs) that respond to a diverse set of external stimuli. These receptors act as GEFs to activate heterotrimeric G proteins [3–6]. Thus, pathway *activation* relies upon increasing the transition rate from the *off* state to the *on* state. There are no GPCRs in plants and other bikonts; the nucleotide exchange occurs spontaneously, without requiring GEF activity [7–9]. G proteins are kept in the *off* state by a repressor such as a GAP or some other protein that holds the self-activating G protein in its inactive state. In this scenario, the presence of a stimulus results in *derepression*, i.e., removal of the repressing activity [10–12]. These two mechanisms for initiating signaling, activation and derepression, are not mutually exclusive. For example, a *concerted* signal initiation, whereby both activation and dererpression are used, is employed in the GTPase cycle of the yeast mating response pathway [13, 14]. In this example, inactive GPCRs recruit a GAP protein and act to repress, whereas active receptors have GEF activity and act to activate. Thus, perception of a stimulus leads to concerted activation and derepression by increasing GEF activity while decreasing GAP activity.

These three mechanisms are not limited to GTPase cycles. The activation mechanism described here in fact is a simpler abstraction of a linear signaling cascade, a classical framework used to study general properties of signaling pathways [15–19] as well as to model specific signaling pathways [20–22]. While derepression may seem like an unusual mechanism, it occurs in numerous important signaling pathways in plants (e.g., auxin, ethylene, gibberellin, phytochrome), as well as gene regulation [23–27]. Concerted mechanisms are found in bacterial two component systems, wherein the same component acts as kinase and phosphatase [28–35]. Another example of a concerted mechanism is the so-called reciprocal regulation of cell-cycle, where Cdk1 not only phosphorylates its substrate but also inhibits the phosphatase [32].

Mathematical modeling has proven to be a useful tool for un-

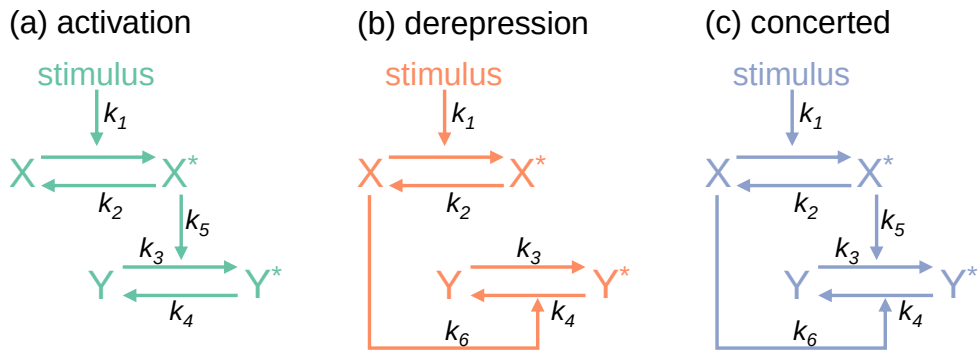

Figure 1: **Mechanisms for signaling through molecular switches**. Presentation of a stimulus activates a receptor ($X \rightarrow X^*$). The reverse reaction causes deactivation of the receptor ($X^* \rightarrow X$). These transitions govern the activity of a molecular switch downstream. (**a**) In the activation mechanism, X^* increases the rate at which the inactive switch (Y) becomes active (Y^*). The opposite reaction $Y^* \rightarrow Y$ has a constant rate. (**b**) In the derepression mechanism, the transition $Y \rightarrow Y^*$ occurs at a constant rate. Activity of the switch is controlled through X: the stimulus decreases X and consequently increases Y^*. (**c**) In the concerted paradigm, both activation and derepression simultaneously control the downstream component.

derstanding the design principles of signaling pathways, and, not surprisingly, mathematical models of activation, derepression and concerted mechanisms have been studied previously. For example, the classical Goldbeter-Koshland model studied zero-order ultrasensitivity of an activation mechanism [15]. Further analyses have examined the effect of receptor numbers [36–38], feedback mechanisms [39, 40], removal of active receptors via endocytosis and degradation [41, 42], etc. Similarly, important properties of the concerted mechanism, such as its ability to do ratiometric signaling [13, 14], to align dose responses at different stages of the signaling pathway [43], as well as its robustness [29, 44] are well-known. The derepression model is relatively less studied. Although there are models of G-signaling in *Arabidopsis thaliana* [45–47], these models have a large number of states and parameters and do not specifically look at system level properties of derepression mechanism.

Despite these efforts, a systematic comparison of various properties of activation, derepression, and concerted mechanisms of signaling has been lacking. Comparing and contrasting these mechanisms should enable our understanding of why different organisms have chosen different mechanisms as well as instruct design of synthetic switches. To this end, we specifically choose four metrics for the comparison: a) dose response curves, b) response times, c) abilities to suppress or filter stochastic fluctuations in upstream components, and d) effect of receptor removal. The rationale behind choosing to compare dose response curves is that they provide information about the input sensitivity range and the output dynamic range, both of which are of pharmacological importance. We supplement this comparison with response times, which provide information about the dynamics of the signaling activity. The third metric of comparison is motivated from the fact that signaling pathways are subject to inherent stochastic nature of biochemical reactions, further compounded by fluctuations in the number of components [48–53]. Lastly, we study the effect of receptor removal on the response of these signaling mechanisms because many well-studied signaling pathways evince receptor removal [11, 42, 54–56]. To study these properties, we construct both deterministic ordinary differential equation models and stochastic models based on continuous time Markov chains.

Our results show that activation has the following two effects: it makes the switch response more sensitive than that of the receptor, and it speeds up the response in comparison with the basal response time. In contrast, derepression makes the switch response less sensitive than the receptor occupancy and slower than the basal speed. These counteracting behaviors of activation and derepression lead to intermediate sensitivity and intermediate response time for the concerted mechanism. In the special case of a perfect concerted mechanisms (equal activation and repression), the concerted mechanism is able to align the dose-response curve of the pathway with receptor occupancy and also have response time that does not depend upon the stimulus level. The noise comparison reveals that the concerted mechanism is more susceptible to fluctuations than the activation and derepression mechanisms, which perform similarly. Finally, our analysis of the effect of receptor removal highlights another important difference between activation and derepression. Removal of active (inactive) receptors at a faster rate than inactive (active) receptors results in an adaptive response for activation (derepression) and sustained response for derepression (activation). We finally compare our findings with experimental observations, suggesting reasons that might have led biological systems to choose one of these mechanisms over the others.

Model formulation

We consider a two-tier model for each of three mechanisms of signaling through a molecular switch (Fig. 1). The first tier is common for all mechanisms, where an inactive receptor (X) becomes active (X^*) when its corresponding input (stimulus) is presented. The second tier is the molecular switch that transitions between *off* (Y) and *on* (Y^*) states. In the activation mechanism, the transition rate from the *on* state to the *off* state increases as the number of active receptor molecules increases (Fig. 1(a)). In the derepression mechanism, the transition rate from the *off* state to the *on* decreases with decrease in the number of inactive receptor molecules (Fig. 1(b)). In the concerted mechanism, both activation and derepression occur simultaneously (Fig. 1(c)). We model these mechanisms using ordinary differential equations (ODEs), assuming mass-action kinetics. To this end, we denote the time by t, stimulus level by S, the total number of receptors by X_T, and the total number of switches by Y_T. We use X^* and

Y^* to denote the number of active receptors and the number of active switches, respectively. The rate constants are as follows: k_1 is the rate of receptor activation per unit stimulus, k_2 is the rate of receptor deactivation, k_3 is the basal rate of activation of the switch, k_4 is the basal rate of deactivation of the switch, k_5 is the strength of activation of an individual inactive receptor, and k_6 is the strength of repression of an individual active receptor. Thus, the (total) activation strength is $k_5 X_T$ and (total) repression strength is $k_6 X_T$. Lastly, we assume that the total number of each species is conserved, and that each model is in steady state before presentation of the stimulus at $t = 0$.

Note that the concerted mechanism encompasses both activation and derepression mechanisms. Therefore, writing ODEs for the concerted mechanism is sufficient to capture all three mechanisms. The number of active receptors and the number of active switches evolve over time according to the following ODEs:

$$\frac{dX^*}{dt} = k_1 S X_T - (k_1 S + k_2) X^*, \tag{1a}$$

$$\frac{dY^*}{dt} = k_3 Y_T + k_5 Y_T X^* - (k_3 + k_4 + k_6 X_T)\, Y^* - (k_5 - k_6) X^* Y^*. \tag{1b}$$

The activation and derepression mechanisms represent limiting cases in which $k_6 = 0$ and $k_5 = 0$, respectively. Solving Eq. (1) requires rate constants and initial conditions to be specified. Unless stated otherwise, we assume that initial conditions are given by the pre-stimulus ($S = 0$) steady state:

$$X_0^* = 0, \quad Y_0^* = \frac{k_3}{k_3 + k_4 + k_6 X_T} Y_T. \tag{1c}$$

With the models described by Eq. (1), we next compare the three signaling mechanisms in terms of their dose responses and response times.

Dose responses

We begin our analysis by examining the steady-state dose responses of activation, derepression, and concerted mechanisms. The steady-state solution to Eq. (1) is

$$\overline{X^*} = \frac{S X_T}{S + \frac{k_2}{k_1}}, \; \overline{Y^*} = \frac{\frac{k_2 k_3}{k_1 (k_3 + k_4 + k_5 X_T)} + \frac{k_3 + k_5 X_T}{k_3 + k_4 + k_5 X_T} S}{\frac{k_2 (k_3 + k_4 + k_6 X_T)}{k_1 (k_3 + k_4 + k_5 X_T)} + S} Y_T. \tag{2}$$

Here $\overline{X^*}$ and $\overline{Y^*}$ are the number of active receptors and the number of active switches a long time after the stimulus is presented ($t \to \infty$), respectively. Notably, both $\overline{X^*}$ and $\overline{Y^*}$ have the form

$$R = \frac{R_0 \Theta_R + R_\infty S}{\Theta_R + S}, \tag{3}$$

where R_0 is the minimum response corresponding to $S = 0$, R_∞ is the maximum response corresponding to $S \gg \Theta_R$, and Θ_R is the stimulus concentration that produces half-maximal response $\frac{R_0 + R_\infty}{2}$. The dynamic range of the response is given by $R_\infty - R_0$, signifying the maximum the output can change in response to the input. Eq. (3) shows that shapes of dose response curves are same for the three signaling mechanisms. Hence comparison between them can be carried out in terms of R_0, R_∞, and Θ_R.

At the receptor level, $X_0^* = 0$ and $X_\infty^* = X_T$, notwithstanding the rate parameters. The half-maximal stimulus Θ_{X^*} is equal to $\frac{k_2}{k_1}$, which is the binding affinity of the stimulus with the receptor. The response at the switch level $\overline{Y^*}$ is specified by:

$$Y_0^* = \frac{k_3 Y_T}{k_3 + k_4 + k_6 X_T}, \tag{4a}$$

$$Y_\infty^* = \frac{(k_3 + k_5 X_T) Y_T}{k_3 + k_4 + k_5 X_T}, \tag{4b}$$

$$\Theta_{Y^*} = \Theta_{X^*} \frac{k_3 + k_4 + k_6 X_T}{k_3 + k_4 + k_5 X_T}. \tag{4c}$$

These expressions show that the dose responses at the switch level depend upon the the basal rates as well as activation strength ($k_5 X_T$) and repression strength ($k_6 X_T$). A careful examination of Eq. (4) provides the following insights:

(i) The repression strength $k_6 X_T$ does not affect the maximum response Y_∞^*, but affects the minimum response Y_0^*. In particular, increasing $k_6 X_T$ decreases Y_0^*. The activation strength $k_5 X_T$ increases Y_∞^* and does not affect Y_0^*.

(ii) Relative values of the repression and activation strengths dictate the relationship between the half-maximal stimulus for the switch response (Θ_{Y^*}) vis-á-vis the half-maximal stimulus for the receptor occupancy (Θ_{X^*}). More specifically, $\Theta_{Y^*} < \Theta_{X^*}$ when $k_5 X_T > k_6 X_T$, $\Theta_{Y^*} = \Theta_{X^*}$ when $k_5 X_T = k_6 X_T$, and $\Theta_{Y^*} > \Theta_{X^*}$ when $k_5 X_T < k_6 X_T$. Increasing $k_6 X_T$ increases Θ_{Y^*} while increasing $k_5 X_T$ does the opposite.

Fig. 2 illustrates the aforementioned effects on dose-response curves for the signaling mechanisms considered. Noting that signaling pathways typically show little activity in absence of the stimulus ($Y_0^* \ll Y_T$) and show full activity ($Y_\infty^* \approx Y_T$) if the stimulus is large, it is reasonable to make the following assumptions: $k_3 \ll k_4 + k_6 X_T$ and $k_4 \ll k_3 + k_5 X_T$. With these assumptions in mind, we use the following parameters for Fig. 2: $k_3 = 0$ and $k_6 = 0$ for activation; $k_4 = 0$ and $k_5 = 0$ for derepression; and $k_3 = 0$ and $k_4 = 0$ for concerted. As shown in Fig. 2(a) activation makes the switch response more sensitive to stimulus than the receptor occupancy ($\Theta_{Y^*} < \Theta_{X^*}$). Increasing the activation strength ($k_6 X_T$) increases Y_∞^* and decreases Θ_{Y^*}, increasing the dynamic range (vertical expansion) and sensitivity (leftward shift) of the dose-response curve. The derepression mechanism exhibits an opposite behavior with $\Theta_Y > \Theta_X$. In this scenario, increasing the repression strength increases the dynamic range by decreasing Y_0^* and decreases sensitivity by increasing Θ_{Y^*} (Fig. 2(b)).

Because we ignore the basal rates, changing activation and derepression strengths only influence Θ_{Y^*} in the case of a concerted mechanism. As expected, the switch response is more (less) sensitive than the receptor occupancy if activation (derepression) dominates derepression (activation). There is a perfect alignment of the fractional receptor occupancy curve with the dose response curve of the switch when $k_5 = k_6$ (Fig. 2(c)). Another important property of the concerted model is that it exhibits ratiometric signaling in which the response of Y^* is determined by the ratio of active receptors (X^*)to the total number of receptors (X_T) [13, 14], and the activity of Y^* is independent of the total number of receptors X_T. This may be seen by looking at Y^*

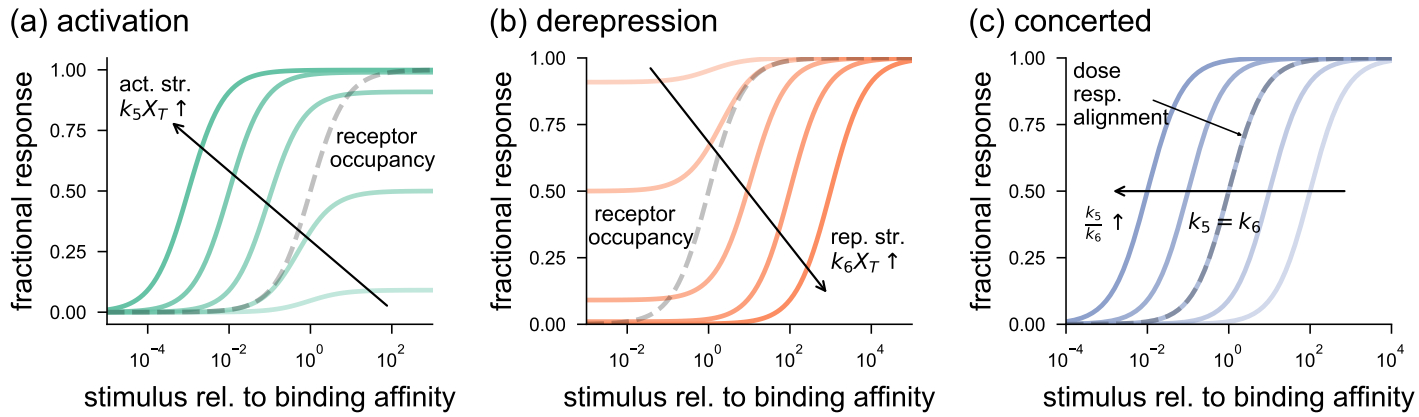

Figure 2: **Dose-response curves for signaling mechanisms through molecular switches**. The response is measured in terms of fraction of active switches Y^*/Y_T as the stimulus level varies. The receptor occupancy curve denotes the fraction of active receptors X^*/X_T. The stimulus is normalized by its binding affinity to the receptor (Θ_{X^*}). **(a)** For the activation mechanism, the half-maximal stimulus (Θ_{Y^*}) of a dose-response curve is less than Θ_{X^*}. Each dose-response curve (solid line) is for a fixed activation strength k_5X_T. Increasing k_5X_T, depicted by the solid arrow, causes an upward expansion and leftward shift in dose-response. For these plots, the following values for parameters were used: $k_3 = 0$, $k_4 = 1$ and $k_6 = 0$. The activation strength (k_5X_T) was varied to take values from $(0.1, 1, 10, 100, 1000)$. **(b)** For the derepression mechanism, the dose-response of the switch for a given repression strength (k_6X_T) has half-maximal stimulus (Θ_Y) greater than Θ_{X^*}. Increasing k_6X_T, shown by the solid arrow, leads to a downward expansion and rightward shift in the dose-response curve. The repression strength k_6X_T takes values from $(0.1, 1, 10, 100, 1000)$. The rest of the parameters were set at $k_3 = 1$, $k_4 = 0$, $k_5 = 0$. **(c)** In the case of concerted mechanism, Θ_{Y^*} may be greater than, equal to, or less than Θ_{X^*}, depending upon the relative values of the activation strength and the derepression strength. Increasing the ratio k_5/k_6, depicted by the solid arrow, shifts the dose response to left. Dose-response alignment ($\Theta_{Y^*} = \Theta_{X^*}$) occurs when $k_5 = k_6$. The parameters used for the plots are $k_3 = 0$ and $k_4 = 0$. The ratio k_5/k_6 was varied over $(0.01, 0.1, 1, 10, 100)$.

when $k_3 = 0$ and $k_4 = 0$:

$$\overline{Y^*} = \frac{SY_T}{S + \frac{k_2 k_6}{k_1 k_5}}. \tag{5}$$

In reality, k_3 and k_4 are likely to be small, but non-zero. Therefore, ratiometric signaling does not hold in a strict sense.

Our theoretical results above show how the dose-response curves behave differently for activation, derepression, and concerted mechanisms. Are some of these behaviors observed in biological systems? One example where the signaling response becomes maximal when only a small fraction of receptors are bound ($\Theta_{Y^*} < \Theta_{X^*}$) is the EGFR–MAPK pathway which elicits a full MAPK response at less than 5% receptor occupancy [57]. Our analysis explains this by an activation mechanism or a concerted mechanism in which the activation strength dominates the repression strength. A contrasting behavior is seen in the ethylene pathway of *Arabidopsis thaliana* in which a loss-of-function mutation of one of the ethylene receptors, *etr1*, shows increased sensitivity to etylene [58]. This points to a derepression mechanism in which decreasing the amount of the receptor lowers the repression strength k_6X_T and shifts the dose response curve to the left. A suggested example of concerted mechanism is the yeast G-signaling pathway, which exhibits both ratiometric signaling [13, 14] and dose response alignment [43].

Response times

Our analysis thus far focused on the steady-state properties of the activation, derepression, and concerted mechanisms. In this section, we study these mechanisms in terms of their response times; that is the time it takes for a signaling output to reach its steady-state. We use the following definition of response time:

$$\mathcal{T}_R = \frac{\int_0^\infty t \left|\overline{R} - R(t)\right| dt}{\int_0^\infty \left|\overline{R} - R(t)\right| dt}, \tag{6}$$

where $R(t)$ is the time-dependent response of the pathway component under consideration and $\overline{R}$ represents its value at steady state [59]. For this definition, $\mathcal{T}_R$ represents the "center of mass" of the response $R(t)$, and is well-defined when $R(0) \neq \overline{R}$. If the response is determined by a single kinetic step, $\mathcal{T}_R$ is reciprocal of the rate constant for that step (section S1, SI). For example, the response time for the receptor is given by

$$\mathcal{T}_{X^*} = \frac{1}{k_1 S + k_2}, \tag{7}$$

which is independent of the initial condition. Eq. (7) shows that the response time decreases if either or both of $k_1 S$ and k_2 increase. Because the response time depends upon the sum $k_1 S + k_2$ and the steady-state receptor occupancy depends upon the ratio $k_1 S/k_2$, these quantities can be tuned independently.

In the absence of stimulus, the response time of the switch follows the same form as Eq. (7):

$$\mathcal{T}_{Y^*}\big|_{S=0} = \frac{1}{k_3 + k_4 + k_6 X_T}. \tag{8}$$

When the stimulus is present, analytic solutions to the integrals in Eq. (6) for the response time of $Y^*(t)$ do not exist, except for a special case of the perfect concerted model $k_5 = k_6$. It is, however, possible to approximate $\mathcal{T}_R$ by linearizing the dynamics Eq. (1) around steady-state:

$$\mathcal{T}_{Y^*} \approx \frac{1}{k_1 S + k_2} + \frac{1}{k_3 + k_4 + k_6 X_T + (k_5 - k_6)\frac{k_1 S X_T}{k_1 S + k2}} - \frac{1}{k_1 S + k_2 + k_3 + k_4 + k_6 X_T}. \tag{9}$$

This above equation is exact for the special case when $k_5 = k_6$ (section S2-C, SI). The first term in Eq. (9) represents the response time of the receptor ($\mathcal{T}_{X^*}$). The second term can be interpreted as the response time of the switch when the receptors are at steady-state, because in that case the switch would be turned *on* at a rate $k_3 + k_5 \frac{k_1 S X_T}{k_1 S + k_2}$ and turned *off* at a rate $k_4 + k_6 X_T - k_6 \frac{k_1 S X_T}{k_1 S + k_2}$.

If the receptor response time is much faster than that of the switch, we expect that the latter does not depend upon the former. Indeed in this limit, Eq. (9) gives

$$\mathcal{T}_{Y^*} \approx \frac{1}{k_3 + k_4 + k_6 X_T + (k_5 - k_6)\frac{k_1 S X_T}{k_1 S + k2}}. \tag{10}$$

Comparing Eq. (10) with the basal response time in Eq. (8) shows that for a given stimulus level, activation speeds up the response in comparison with the basal response. In contrast, derepression slows down the response time and a prefect concerted mechanism does not affect the response time (Fig. 3(a)). In the other limiting case when the receptor timescale is much slower than that of the switch, we expect the receptor dynamics to dictate the response time. Indeed in this case, Eq. (9) reduces to $\mathcal{T}_{Y^*} \approx \mathcal{T}_{X^*}$ such that choice of the mechanism to control the switch has little effect on the response time. Our analytical as well as numerical calculations confirm this behavior (Fig. 3(a)).

Next we examine the scenario where the switch is controlled by varying the stimulus level. Because changing the stimulus affects the response time of the receptor, which in turn affects the response time of the switch, we control for this effect by keeping $k_1 S + k_2$ constant but changing the ratio $k_1 S / k_2$. We find that activation indeed speeds up with increasing stimulus levels, whereas derepression slows down (Fig. 3(b)). Importantly, the response time of the concerted mechanism is independent of the stimulus strength, and, therefore able to respond rapidly over the whole range of stimulus levels. We see that activation is faster than derepression only if the basal response times are equal. Therefore to construct a switch that responds rapidly using derepression, it is necessary for the switch to undergo fast basal cycling. (Fig. 3(b)).

Is there an intuitive explanation to why activation is faster than derepression? The activation model shortens the average lifetime of the *off* state, without affecting the average lifetime of the *on* state. Derepression operates differently; it does not affect the average lifetime of the *off* state, but increases the lifetime of the *on* state. Thus, activation responds faster than derepression. The concerted model decreases the lifetime of the *off* state in addition to increasing the lifetime of the *on state*, its response time lies between those of activation and derepression.

Processing upstream fluctuations

The deterministic models used to compare the signaling mechanisms thus far ignore the stochastic nature of biochemical reactions, which becomes relevant when the abundance of receptor and switch proteins are small [48–53, 60, 61]. Therefore, we formulate a stochastic model of the concerted mechanism and analyze the other two mechanisms as its special cases. Our model consists of four reactions: activation of receptor upon recognizing the stimulus, deactivation of receptor, *on* to *off* transition of the molecular switch, and *off* to *on* transition of the molecular switch.

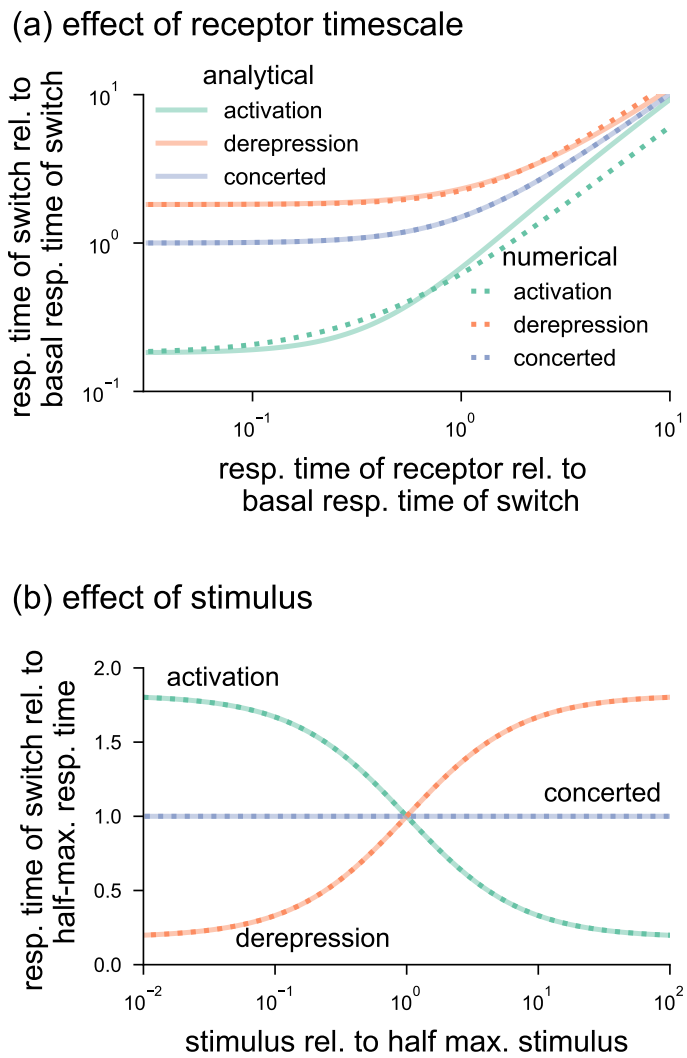

Figure 3: **Response times of molecular switches governed by activation, derepression, and concerted mechanisms**. (**a**) The response time of the switch increases as the response time of the receptor increases. Differences in the signaling mechanisms are more prominent when the receptor response is fast. Activation speeds up the response in comparison with the basal response time whereas dererepssion slows it down. A perfect concerted mechanism does not affect the response time. For each signaling mechanism, the response time is computed using the analytical result in Eq. (9) (solid lines) and numerically validated by using Eq. (6) (dashed lines). To ensure the same basal response and basal response time of the switches across signaling mechanisms, we chose the parameters as $k_3 = 1/9$, $k_4 = 1$, $k_6 = 0$, and $k_5 X_T = 10$ for activation; $k_3 = 1/9$, $k_4 = 0$, $k_5 = 0$, and $k_6 X_T = 1$ for derepression; and $k_3 = 1/9$, $k_4 = 0$ and $k_5 X_T = k_6 X_T = 1$ for concerted. The receptor response time was $k_1 S + k_2$ varied through k_2 while maintaining $k_1 S / k_2 = 1$. (**b**) Increasing the stimulus level, decreases the response time for activation, increases in response time for derepression, and does not affect the response time for the concerted mechanism. The comparison is controlled by setting same response time at half-max. stimulus Θ_{Y^*}. The following parameters were chosen to have same basal response but different basal response times: $k_3 = 1$, $k_4 = 9$, $k_5 X_T = 90$, and $k_6 = 0$ for activation; $k_3 = 10$,$k_4 = 0$,$k_5 = 0$, and $k_6 X_T = 90$ for derepression; and $k_3 = 10$, $k_4 = 0$, $k_5 X_T = k_6 X_T = 90$ for concerted. The receptor occupancy was varied by changing $k_1 S / k_2$ while maintaining the receptor response time $1/(k_1 S + k_2)$, which was chosen to be 100 times faster than the response time of the switches at their respective half maximal stimulus levels.

The stochastic model is characterized by the probabilistic nature of each reaction and the discreteness of changes in population counts upon occurrence of a reaction as tabulated in Table 1.

Event	Update	Transition rate
$X \to X^*$	$X^* \mapsto X^* + 1$	$k_1 S(X_T - X^*)$
$X^* \to X$	$X^* \mapsto X^* - 1$	$k_2 X^*$
$Y \to Y^*$	$Y^* \mapsto Y^* + 1$	$(k_3 + k_5 X^*)(Y_T - Y^*)$
$Y^* \to Y$	$Y^* \mapsto Y^* - 1$	$(k_4 + k_6(X_T - X^*))Y^*$

Table 1: Transitions and associated rates for the stochastic model.

Our goal is to analyze the noise properties of activation, derepression, and concerted mechanisms. We quantify noise using coefficient of variation squared (CV^2), which is computed by normalizing the variance by mean squared and is a dimensionless quantity. To this end, we use the ODEs that describe the time evolution of the first and second-order moments, and solve them in steady-state to obtain the stationary moments [62, 63] (section S3-B, SI). In particular, moments for the number of active receptors (X^*) are given by

$$\langle X^* \rangle = \frac{k_1 S X_T}{k_1 S + k_2}, CV^2_{X^*} = \frac{\langle X^{*2} \rangle - \langle X^* \rangle^2}{\langle X^* \rangle^2} = \frac{1}{X_T} \frac{k_2}{k_1 S}. \quad (11)$$

Here $\langle . \rangle$ denotes the expected value (average) of its argument. These moments correspond to a binomial distribution with parameters X_T and $\frac{k_1 S}{k_1 S + k_2}$ (section S3-A, SI). The stochastic mean $\langle X^* \rangle$ is same as the steady-state value for X^* in the deterministic model in Eq. (1). The coefficient of variation squared increases as the number of receptors (X_T) decreases. Therefore the noise analysis is of importance when X_T is small. In addition, the noise decreases with the ratio $k_1 S / k_2$. Recall that $k_1 S / k_2$ is the stimulus level relative to the binding affinity. Thus the noise diminishes when the stimulus level is much higher than the binding affinity.

Closed-form expressions for the moments are not available for Y^* owing to the nonlinear term $X^* Y^*$ in reaction rates, except for the special case of a perfect concerted model ($k_5 = k_6$). We approximate the mean response and the noise by considering a linearized system around the steady-state

$$\langle Y^* \rangle \approx \frac{k_3 + k_5 \frac{k_1 S}{k_1 S + k_2} X_T}{k_3 + k_4 + k_6 \frac{k_2}{k_1 S + k_2} X_T} Y_T, \quad (12a)$$

$$CV^2_{Y^*} \approx \underbrace{\frac{1}{Y_T} \frac{k_4 + k_6 X_T \frac{k_2}{k_1 S + k_2}}{k_3 + k_5 X_T \frac{k_1 S}{k_1 S + k_2}}}_{\text{contribution from act./deact. of } Y^*} + \underbrace{\frac{k_3 + k_4 + k_5 \frac{k_1 S X_T}{k_1 S + k_2} + k_6 \frac{k_2 X_T}{k_1 S + k_2}}{k_1 S + k_2 + k_3 + k_4 + k_5 \frac{k_1 S X_T}{k_1 S + k_2} + k_6 \frac{k_2 X_T}{k_1 S + k_2}}}_{\text{time-averaging}} \times \underbrace{\frac{\left(\frac{k_1 S X_T}{k_1 S + k_2}\right)^2 (k_4 k_5 + k_6(k_3 + k_5 X_T))^2}{\left(k_3 + k_5 \frac{k_1 S X_T}{k_1 S + k_2}\right)^2 \left(k_3 + k_4 + k_5 \frac{k_1 S X_T}{k_1 S + k_2} + k_6 \frac{k_2 X_T}{k_1 S + k_2}\right)^2}}_{\text{coupling}} CV^2_{X^*}. \quad (12b)$$

We validate these approximations using exact semi-analytical approach based on [64] (section S3-B, SI). The formula for $CV^2_{Y^*}$ above is written in terms of various sources of noise, as previously done for gene regulation models [65, 66]. Specifically, the noise in the signaling activity of the switch arises from two sources: activation/deactivation reactions of the switch, and noise in the number of active receptors, which is processed via time-averaging and coupling.

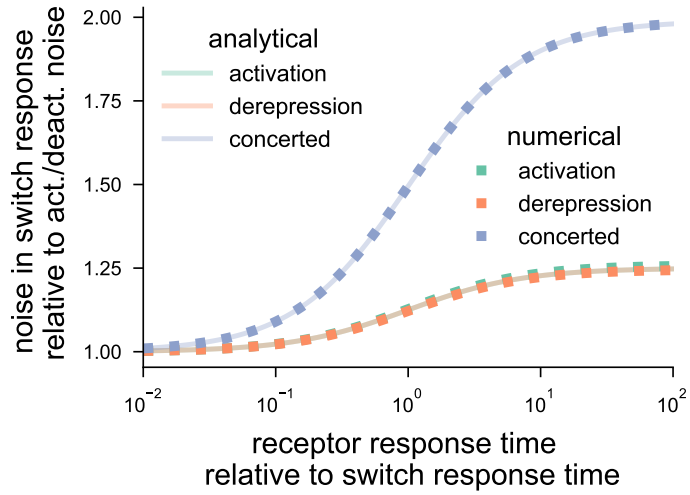

Figure 4: **Noise in the number of active switch molecules**. The concerted model has a higher noise than activation and derepression, which perform similar. The difference is negligible when receptor dynamics is fast, and is more prominent when receptor is slow. Noise is quantified using coefficient of variation squared. The analytical result is computed using Eq. (12), which is validated numerically. To control for the effects of variations in the mean number of switches and the differences in timescales, the following parameters were chosen: $k_3 = 0$, $k_4 = 1$, $k_5 = 0.02$, and $k_6 = 0$ for activation; $k_3 = 1$, $k_4 = 0$, $k_5 = 0$, and $k_6 = 0.02$ for derepression; and $k_3 = 0$, $k_4 = 0$, $k_5 = k_6 = 0.02$ for concerted. In addition, X_T and $Y_T = 100$ were taken to be 100 each. The receptor response time ($1/(k_1 S + k_2)$) was varied through k_2 while maintaining the receptor occupancy by $k_1 S / k_2 = 1$.

The contribution from activation/deactivation of the swich has a similar form as $CV^2_{X^*}$ in Eq. (11). Accordingly, the contribution of this term decreases with increase in Y_T or increase in the ratio of the total activation rate with total deactivation rate. This ratio increases if the activation strength increases or the repression strength decreases. We note that changing these parameters also affect the time-averaging as well as the coupling term. One possibility through which only the contribution from $CV^2_{X^*}$ to overall $CV^2_{Y^*}$ is affected is to change $k_1 S + k_2$ while maintaining the ratio $k_1 S / k_2$. In such a scenario, all other terms remain constant except for the time-averaging term. In the limiting case where receptor dynamics is very fast, the contribution from $CV^2_{X^*}$ to $CV^2_{Y^*}$ becomes negligible due to efficient time-averaging of fluctuations in X^*. One more interesting limiting case is that of ratiometric signaling. Our deterministic analysis shows that for a concerted mechanism without basal rates ($k_3 = 0$ and $k_4 = 0$), the steady-state response ($\overline{Y^*}$) does not depend upon the total number of receptors (X_T). It is easy to see that this assertion does not hold true for $CV^2_{Y^*}$, which still depends upon X_T through the time-averaging term and $CV^2_{X^*}$.

Next, we compare the noise properties of activation, derepression, and concerted mechanisms. To mathematically control the comparison, we assume that the receptor dynamics is

same across the three strategies. In addition, we maintain the same average rate at which switch turns *on* from the *off* state, i.e., $k_3 + k_5 \frac{k_1 S X_T}{k_1 S + k_2}$, and the same average rate at which the switch turns *off* from the *on* state, i.e., $k_4 + k_6 \frac{k_2 X_T}{k_1 S + k_2}$. These assumptions ensure that the difference in the noise properties, if any, is solely due to the architecture of the molecular switch and not dependent on the parameters. With this setup, we vary the timescale of the receptor, while keeping the same ratio $k_1 S / k_2$. As shown in Fig. 4, the noise properties of these signaling mechanisms are similar when the receptor timescale is fast. This is expected because the dominant contribution in $CV^2_{Y^*}$ comes from its own activation and deactivation. However, when the receptor timescale is slower than that of the switch, the overall noise increases regardless of the signaling mechanism and the noise performance of the concerted mechanism becomes worse than the other two mechanisms.

The observation that activation and derepression both have similar noise and their concerted action has higher noise is surprising in light of our analyses of dose-response and response time. In terms of those properties, activation and derepression counteract to enable intermediate response for the concerted mechanism. Intuitively, this occurs because in the concerted mechanism, fluctuations in the upstream component affect both transitions $Y \rightarrow Y^*$ and $Y^* \rightarrow Y$. In case of activation and derepression, however, only one of these transitions is coupled with the upstream component. As a result, concerted mechanism has a worse noise performance.

Effect of receptor removal

Our models of signaling mechanisms in Fig. 1 assume conservation of number of receptor molecules (X_T) and of switch molecules (Y_T). These assumptions do not hold in case of some signaling pathways where stimulus-mediated removal of receptors occurs. Reported examples of such phenomena include GPCRs [54], EGFR [42], AMPA-type glutamate receptors [55], the receptor-like kinase FLS2 [56] and regulator of G-signaling (RGS) in *Arabidopsis thaliana* [11]. On the one hand, removal of active receptors is proposed to be a mechanism for desensitizing the response to a sustained stimulus [67, 68], and consequently enabling signaling over a broad range [41,42]. On the other hand, phosphorylation and subsequent removal of RGS, which is both a receptor candidate and a GAP, is proposed to result in sustained activation of signaling in *Arabidopsis thaliana* [9, 11, 45]. With an aim to explain these seemingly opposite behaviors of signaling pathways, we ask whether the signaling mechanisms, particularly activation and derepression, behave differently upon removal of receptors. To answer this, we reformulate the models in Fig. 1 by including production of inactive receptors (X) at a rate k_p, removal of inactive receptors with rate k_d, and removal of active receptors with rate k_d^*. This model is simpler than those showing a broad range [41] or relative sensing [42], but is capable of adaptation [68] which is what we focus on.

Inclusion of receptor removal results in the following modification of the ODE system in Eq (1)

$$\frac{dX}{dt} = k_p - k_d X - k_1 S X + k_2 X^*, \tag{13a}$$

$$\frac{dX^*}{dt} = k_1 S X - k_2 X^* - k_d^* X^*, \tag{13b}$$

$$\frac{dY^*}{dt} = \left(k_3 + k_5 X^*\right)\left(Y_T - Y^*\right) - \left(k_4 + k_6 X\right) Y^*. \tag{13c}$$

The initial conditions are: $X(0) = \frac{k_p}{k_d}$, $X^*(0) = 0$, and $Y^*(0) = \frac{k_3 Y_T}{k_3 + k_4 + k_6 k_p / k_d}$. As before, setting $k_6 = 0$ and $k_5 = 0$, respectively, result in ODEs for the activation and derepression mechanisms.

An important distinction between the ODEs in Eq. (1) and the ODEs in Eq. (13) is that the receptor dynamics only has one timescale, $1/(k_1 S + k_2)$, in the former but two timescales in the latter (section S4, SI). The interplay between these two timescales allows $X(t)$ and $X^*(t)$ to transiently respond to a stimulus at the fast timescale, followed by an eventual return towards their respective pre-stimulus levels at the slow timescale. Note that the switch response in the activation mechanism depends upon the active receptors X^*. Thus if X^* increases and returns towards its basal level, Y^* is also expected to follow the same dynamics. Likewise, if a derepression mechanism governs the switch then a decrease in X would lead to increase in Y^*. Further, if X returns towards its basal level, Y^* should also follow this trend. Such a behavior is referred to as adaptation [68, 69].

What are appropriate parameter regimes where the switch response $Y^*(t)$ in Eq. (13) adapts to a sustained stimulus? Our analysis shows that adaptation by X^* occurs when $k_d^* > k_d$, i.e., active receptors are removed at a faster rate than inactive receptors. In contrast, adaptation by X happens when $k_d^* < k_d$ (section S4, SI). We note that all other parameters also affect the response properties, but the relative rates of receptor removal are the most important determinants of adaptive behavior. We illustrate these results in Fig. 5. To further bolster our observations, we also examine the scenarios where the inactive receptors are preferentially removed for an activation mechanism and active receptors are preferentially removed for a derepression mechanism. Interestingly, in both these cases, the response sustains and does not adapt. These results thus provide another set of differences between activation and derepression mechanisms. It is also plausible that controlling the relative rates of receptor removal is a deliberate mechanism to control the downstream response.

Discussion

Molecular switches are important components in signaling pathways. A switch transitions between *on* and *off* states and signaling through it occurs by increasing its *on*-to-*off* transition rate (activation), or by decreasing its *off*-to-*on* transition rate (derepression), or by simultaneously doing both (concerted). In this work, we studied these signaling mechanisms in terms of their dose-responses, response times, and abilities to process upstream fluctuations. We further examined how the effect of removal of receptors depends upon the signaling mechanism governing the switch. Below we enumerate the key differences in the properties of the signaling response that depend upon how the switch is controlled.

- Both activation and derepression mis-align the signaling activity with the receptor occupancy. In particular, activation reduces the stimulus level required for half-maximal signaling activity in comparison with the binding affinity of the receptor ($\Theta_{Y^*} < \Theta_{X^*}$). Derepression leads to $\Theta_{Y^*} > \Theta_{X^*}$. The dose response align ($\Theta_{Y^*} = \Theta_{X^*}$) for a perfect concerted mechanism which corresponds to equal strengths of activation and repression (Fig. 2).

- A concerted mechanism is capable of ratiometric signaling, where the steady-state signaling output only depends upon fractional receptor occupancy and not on the total number of receptors.

- Presentation of a stimulus speeds up the response for an activation mechanism, whereas it slows down the response for a derepression mechanism. The response time does not change for a perfect concerted mechanism (Fig. 3).

- Activation and derepression mechanisms have similar noise in the signaling activity. Concerted mechanism has a higher noise (Fig. 4).

- Preferential removal of active (inactive) receptors leads to an adaptive response for activation (derepression) mechanism and a sustained response for derepression (activation) mechanism (Fig. 5).

These results suggest several trade-offs in the properties of response that depend upon the underlying signaling mechanism. For instance, activation enables sensitive and rapid response, but at the cost of misaligned dose response with receptor occupancy, resulting in poor information fidelity [70]. Derepression also leads to misalignment between dose response and receptor occupancy, albeit the response is less sensitive and slow. The concerted mechanism has a better information fidelity, but it comes at the cost of a poor control over the speed of response and a higher susceptibility to upstream fluctuations. It is worth noting that additional layers of regulation such as feedback and feedforward loops may compensate for some of the shortcomings of a signaling mechanism. One such example is a negative feedback mechanism, which aligns dose response for a signaling pathway operating through activation [43, 70, 71].

Despite being similar in terms of misaligned dose response, there is a key difference between activation and derepression. The former operates as an 'eager' system which is sensitive to small receptor occupancy and speeds up the response upon receiving a stimulus. The latter works as a 'conservative' system which ignores small receptor occupancy as if waiting for additional evidence before committing to a response. The response time also slows down as the stimulus increases. Activation seems fit for purposes where the cost of a false negative is much higher than a false positive. For example, adrenaline response to an imminent danger should be sensitive and fast because cost of a false positive is small but a false negative is high. In contrast, derepression is appropriate for scenarios where cost of a false positive is much higher than a false negative. Based on the observation the derepression based signaling is found in many pathways in plants, we speculate that it happens because plants have to continually balance their resources between growth and immunity [72, 73]. For example, plants would perhaps ignore growth of a low level of pathogenic bacteria before allocating resources to fight them. Another possible scenario where derepression may be used is irreversible cell-fate decisions.

Given that activation and derepression shift dose-response to opposite directions, a natural question to ask is whether dose-response alignment can occur in a signaling cascade where activation and derepression do not operate in a concerted fashion, but operate sequentially? To explore this possibility, we construct a three-tier model where the response Y^* in Fig. 1(a) leads to derepression of a downstream component. Our analysis shows that indeed the response of the downstream component is better aligned with the dose response than Y^*. We also analyze an alternate mechanism where derepression is followed by activation by modifying Fig. 1(b). As expected, the dose response of the downstream component aligns with the dose response better than that of Y^* (section S5, SI).

Until now, we have discussed the drawback of having misalignment between dose response and receptor occupancy. It is plausible that it offers some advantages too. Consider a scenario where active receptors are preferentially removed, resulting in adaptation of the signaling response (Fig. 5). Recent work has shown that it is possible to exploit this to perform relative sensing (fold-change detection) if the receptor removal is a multi-step process [42]. Alternatively, a negative feedback may also result in adaptive response and thereby a fold-change detection [59]. A key feature of fold-change detection is that the dose response curve shifts towards right each time the system adapts [59, 74]. Our results suggest that a relative sensing property may be implemented with a derepression mechanism if the receptor removal operates on inactive receptors. We speculate that a negative feedback operating on inactive receptors would also yield the same effect. It is of course possible to shift the dose-response of a concerted mechanism too as seen in Fig. 2. However, this would come at the cost of misaligned dose response. How do cells balance between being responsive to a large range of stimulus as well as transmitting information with high fidelity is an interesting problem in itself.

Whereas removal for active receptors for an activation mechanism results in an adaptive response, removal of inactive receptors manifests as a sustained response. Derepression also generates such distinct profiles: adaptive response when inactive receptors are removed and sustained response when active receptors are removed. Given that many important switches, e.g., G proteins, process signals from a large number of receptors, such distinct profiles of the signaling response may be used to distinguish signals for a downstream response. To our knowledge, active receptors are often removed as compared to inactive receptors, however this does not rule out the possibility that for receptors their inactive forms are preferentially removed.

A caveat of the models of signaling mechanisms considered herein is that they are based on mass action kinetics and therefore cannot capture saturation effects. Traditionally, signaling pathways are modeled using Michelis-Menten kinetics. While both exact and approximate analytical results are key ingredients of this work, it is usually not possible to derive those with Michelist-Menten kinetics. Nonetheless, whether our results hold when a switch operates in saturating regime needs to be established in a future work. Another direction of future work would be to extend such analysis to include feedback and feedforward

regulation. The noise analysis in our work is limited to intrinsic noise which arises due to randomness in activation and inactivation of different components. Incorporating the effect of cell-to-cell variability in the number of components, which is presumed to be a major source of overall noise at least in yeast pheromone pathway [75], would be another direction of exploration. Finally, we note that similar mechanisms of regulating cellular responses may exist in other contexts, e.g., gene regulation, metabolic pathways. It would be worthwhile to extend our analysis to these and develop a unifying theory.

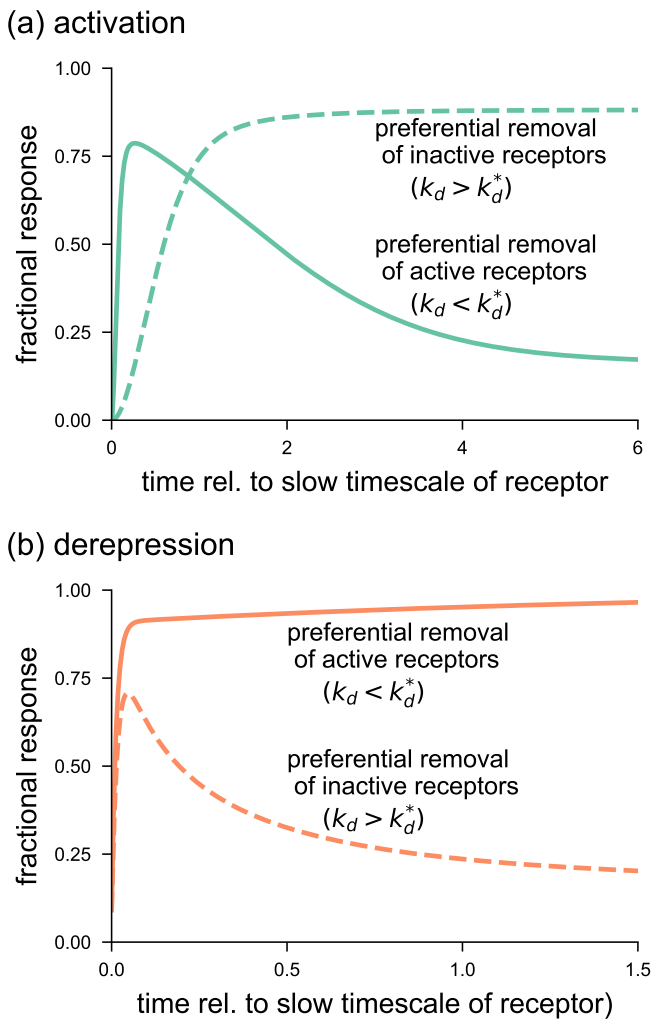

Figure 5: **Effect of receptor removal on responses of activation and derepression mechanisms**. The response is measured in terms of fraction of active receptors (Y^*/Y_T) over time, which is normalized to the slow timescale of the receptor (see section S4, SI). (**a**) For an activation mechanism, the switch response adapts, i.e., returns towards basal response after a transient, if the rate of removal of active receptors (k_d^*) is higher than that of inactive receptors (k_d). In contrast, if inactive receptors are removed at a faster rate, then the response sustains. (**b**) For derepression mechanism, preferential removal of X results in adaptation. In contrast, preferential removal of X^* causes sustained response.

Acknowledgments

The authors thank members of the Jones and Elston labs for extensive discussion and feedback. This work was supported by NSF via grant MCB–1713880 to AMJ and TCE.

References

[1] F. R. Bischoff and H. Ponstingl, "Catalysis of guanine nucleotide exchange on Ran by the mitotic regulator RCC1," *Nature*, vol. 354, no. 6348, pp. 80–82, 1991.

[2] F. R. Bischoff, H. Krebber, T. Kempf, I. Hermes, and H. Ponstingl, "Human RanGTPase-activating protein RanGAP1 is a homologue of yeast Rna1p involved in mRNA processing and transport," *Proceedings of the National Academy of Sciences*, vol. 92, no. 5, pp. 1749–1753, 1995.

[3] N. Wettschureck and S. Offermanns, "Mammalian g proteins and their cell type specific functions," *Physiological reviews*, vol. 85, no. 4, pp. 1159–1204, 2005.

[4] C. A. Johnston, F. S. Willard, M. R. Jezyk, Z. Fredericks, E. T. Bodor, M. B. Jones, R. Blaesius, V. J. Watts, T. K. Harden, J. Sondek, J. K. Ramer, and D. P. Siderovski, "Structure of gαi1 bound to a gdp-selective peptide provides insight into guanine nucleotide exchange," *Structure*, vol. 13, no. 7, pp. 1069–1080, 2005.

[5] D. P. Siderovski and F. S. Willard, "The gaps, gefs, and gdis of heterotrimeric g-protein alpha subunits," *International journal of biological sciences*, vol. 1, no. 2, p. 51, 2005.

[6] W. M. Oldham and H. E. Hamm, "Heterotrimeric g protein activation by g-protein-coupled receptors," *Nature reviews Molecular cell biology*, vol. 9, no. 1, pp. 60–71, 2008.

[7] C. A. Johnston, J. P. Taylor, Y. Gao, A. J. Kimple, J. C. Grigston, J.-G. Chen, D. P. Siderovski, A. M. Jones, and F. S. Willard, "GTPase acceleration as the rate-limiting step in Arabidopsis G protein-coupled sugar signaling," *Proceedings of the National Academy of Sciences*, vol. 104, no. 44, pp. 17317–17322, 2007.

[8] J. C. Jones, J. W. Duffy, M. Machius, B. R. Temple, H. G. Dohlman, and A. M. Jones, "The crystal structure of a self-activating G protein α subunit reveals its distinct mechanism of signal initiation," *Sci. Signal.*, vol. 4, no. 159, pp. ra8–ra8, 2011.

[9] D. Urano, J. C. Jones, H. Wang, M. Matthews, W. Bradford, J. L. Bennetzen, and A. M. Jones, "G protein activation without a GEF in the plant kingdom," *PLoS genetics*, vol. 8, no. 6, p. e1002756, 2012.

[10] J.-G. Chen, F. S. Willard, J. Huang, J. Liang, S. A. Chasse, A. M. Jones, and D. P. Siderovski, "A seven-transmembrane RGS protein that modulates plant cell proliferation," *Science*, vol. 301, no. 5640, pp. 1728–1731, 2003.

[11] D. Urano, N. Phan, J. C. Jones, J. Yang, J. Huang, J. Grigston, J. P. Taylor, and A. M. Jones, "Endocytosis of the seven-transmembrane RGS1 protein activates G-protein-coupled signalling in Arabidopsis," *Nature Cell Biology*, vol. 14, no. 10, pp. 1079–1088, 2012.

[12] X. Liang, M. Ma, Z. Zhou, J. Wang, X. Yang, S. Rao, G. Bi, L. Li, X. Zhang, J. Chai, S. Chen, and J.-M. Zhou, "Ligand-triggered de-repression of Arabidopsis heterotrimeric G proteins coupled to immune receptor kinases," *Cell research*, vol. 28, no. 5, pp. 529–543, 2018.

[13] A. Bush, G. Vasen, A. Constantinou, P. Dunayevich, I. L. Patop, M. Blaustein, and A. Colman-Lerner, "Yeast GPCR signaling reflects the fraction of occupied receptors, not the number," *Molecular systems biology*, vol. 12, no. 12, 2016.

[14] N. T. Henderson, M. Pablo, D. Ghose, M. R. Clark-Cotton, T. R. Zyla, J. Nolen, T. C. Elston, and D. J. Lew, "Ratiometric GPCR signaling enables directional sensing in yeast," *PLoS biology*, vol. 17, no. 10, 2019.

[15] A. Goldbeter and D. E. Koshland, "An amplified sensitivity arising from covalent modification in biological systems," *Proceedings of the National Academy of Sciences*, vol. 78, no. 11, pp. 6840–6844, 1981.

[16] R. Heinrich, B. G. Neel, and T. A. Rapoport, "Mathematical models of protein kinase signal transduction," *Molecular cell*, vol. 9, no. 5, pp. 957–970, 2002.

[17] M. Chaves, E. D. Sontag, and R. J. Dinerstein, "Optimal length and signal amplification in weakly activated signal transduction cascades," *The Journal of Physical Chemistry B*, vol. 108, no. 39, pp. 15311–15320, 2004.

[18] E. Feliu, M. Knudsen, L. N. Andersen, and C. Wiuf, "An algebraic approach to signaling cascades with n layers," *Bulletin of mathematical biology*, vol. 74, no. 1, pp. 45–72, 2012.

[19] M. Beguerisse-Díaz, R. Desikan, and M. Barahona, "Linear models of activation cascades: analytical solutions and coarse-graining of delayed signal transduction," *Journal of The Royal Society Interface*, vol. 13, no. 121, p. 20160409, 2016.

[20] R. Macfarlane, "An enzyme cascade in the blood clotting mechanism and its function as a biochemical amplifier," *Nature*, vol. 202, no. 4931, pp. 498–499, 1964.

[21] C.-Y. Huang and J. E. Ferrell, "Ultrasensitivity in the mitogen-activated protein kinase cascade," *Proceedings of the National Academy of Sciences*, vol. 93, no. 19, pp. 10078–10083, 1996.

[22] B. N. Kholodenko, "Cell-signalling dynamics in time and space," *Nature reviews Molecular cell biology*, vol. 7, no. 3, pp. 165–176, 2006.

[23] V. Shingler, "Signal sensing by σ54-dependent regulators: derepression as a control mechanism," *Molecular microbiology*, vol. 19, no. 3, pp. 409–416, 1996.

[24] A. Dill and T.-p. Sun, "Synergistic derepression of gibberellin signaling by removing RGA and GAI function in arabidopsis thaliana," *Genetics*, vol. 159, no. 2, pp. 777–785, 2001.

[25] L. E. Rogg and B. Bartel, "Auxin signaling: derepression through regulated proteolysis," *Developmental cell*, vol. 1, no. 5, pp. 595–604, 2001.

[26] J. M. Alonso and A. N. Stepanova, "The ethylene signaling pathway," *Science*, vol. 306, no. 5701, pp. 1513–1515, 2004.

[27] K. M. Light, J. A. Wisniewski, W. A. Vinyard, and M. T. Kieber-Emmons, "Perception of the plant hormone ethylene: known-knowns and known-unknowns," *JBIC Journal of Biological Inorganic Chemistry*, vol. 21, no. 5-6, pp. 715–728, 2016.

[28] F. D. Russo and T. J. Silhavy, "The essential tension: opposed reactions in bacterial two-component regulatory systems," *Trends in microbiology*, vol. 1, no. 8, pp. 306–310, 1993.

[29] G. Shinar, R. Milo, M. R. Martínez, and U. Alon, "Input–output robustness in simple bacterial signaling systems," *Proceedings of the National Academy of Sciences*, vol. 104, no. 50, pp. 19931–19935, 2007.

[30] Y. Hart, Y. E. Antebi, A. E. Mayo, N. Friedman, and U. Alon, "Design principles of cell circuits with paradoxical components," *Proceedings of the National Academy of Sciences*, vol. 109, no. 21, pp. 8346–8351, 2012.

[31] Y. Hart and U. Alon, "The utility of paradoxical components in biological circuits," *Molecular cell*, vol. 49, no. 2, pp. 213–221, 2013.

[32] J. E. Ferrell Jr, "Feedback loops and reciprocal regulation: recurring motifs in the systems biology of the cell cycle," *Current opinion in cell biology*, vol. 25, no. 6, pp. 676–686, 2013.

[33] M. A. Rowland and E. J. Deeds, "Crosstalk and the evolution of specificity in two-component signaling," *Proceedings of the National Academy of Sciences*, vol. 111, no. 15, pp. 5550–5555, 2014.

[34] B. N. Dubey, C. Lori, S. Ozaki, G. Fucile, I. Plaza-Menacho, U. Jenal, and T. Schirmer, "Cyclic di-GMP mediates a histidine kinase/phosphatase switch by noncovalent domain cross-linking," *Science advances*, vol. 2, no. 9, p. e1600823, 2016.

[35] L. Gelens, J. Qian, M. Bollen, and A. T. Saurin, "The importance of kinase–phosphatase integration: Lessons from mitosis," *Trends in cell biology*, vol. 28, no. 1, pp. 6–21, 2018.

[36] S. Strickland and J. N. Loeb, "Obligatory separation of hormone binding and biological response curves in systems dependent upon secondary mediators of hormone action," *Proceedings of the National Academy of Sciences*, vol. 78, no. 3, pp. 1366–1370, 1981.

[37] A. Goldbeter and D. E. Koshland, "Ultrasensitivity in biochemical systems controlled by covalent modification. interplay between zero-order and multistep effects.," *Journal of Biological Chemistry*, vol. 259, no. 23, pp. 14441–14447, 1984.

[38] J. E. Ferrell, "Tripping the switch fantastic: how a protein kinase cascade can convert graded inputs into switch-like outputs," *Trends in biochemical sciences*, vol. 21, no. 12, pp. 460–466, 1996.

[39] M. Behar, N. Hao, H. G. Dohlman, and T. C. Elston, "Mathematical and computational analysis of adaptation via feedback inhibition in signal transduction pathways," *Biophysical journal*, vol. 93, no. 3, pp. 806–821, 2007.

[40] M. Behar, N. Hao, H. G. Dohlman, and T. C. Elston, "Dose-to-duration encoding and signaling beyond saturation in intracellular signaling networks," *PLoS computational biology*, vol. 4, no. 10, 2008.

[41] V. Becker, M. Schilling, J. Bachmann, U. Baumann, A. Raue, T. Maiwald, J. Timmer, and U. Klingmüller, "Covering a broad dynamic range: information processing at the erythropoietin receptor," *Science*, vol. 328, no. 5984, pp. 1404–1408, 2010.

[42] E. Lyashenko, M. Niepel, P. D. Dixit, S. K. Lim, P. K. Sorger, and D. Vitkup, "Receptor-based mechanism of relative sensing and cell memory in mammalian signaling networks," *eLife*, vol. 9, p. e50342, 2020.

[43] S. S. Andrews, W. J. Peria, C. Y. Richard, A. Colman-Lerner, and R. Brent, "Push-pull and feedback mechanisms can align signaling system outputs with inputs," *Cell systems*, vol. 3, no. 5, pp. 444–455, 2016.

[44] E. Batchelor and M. Goulian, "Robustness and the cycle of phosphorylation and dephosphorylation in a two-component regulatory system," *Proceedings of the National Academy of Sciences*, vol. 100, no. 2, pp. 691–696, 2003.

[45] Y. Fu, S. Lim, D. Urano, M. Tunc-Ozdemir, N. G. Phan, T. C. Elston, and A. M. Jones, "Reciprocal encoding of signal intensity and duration in a glucose-sensing circuit," *Cell*, vol. 156, no. 5, pp. 1084–1095, 2014.

[46] K.-L. Liao, R. D. Jones, P. McCarter, M. Tunc-Ozdemir, J. A. Draper, T. C. Elston, D. Kramer, and A. M. Jones, "A shadow detector for photosynthesis efficiency," *Journal of theoretical biology*, vol. 414, pp. 231–244, 2017.

[47] K.-L. Liao, C. E. Melvin, R. Sozzani, R. D. Jones, T. C. Elston, and A. M. Jones, "Dose-duration reciprocity for G protein activation: Modulation of kinase to substrate ratio alters cell signaling," *PloS one*, vol. 12, no. 12, 2017.

[48] Y. Sako, S. Minoghchi, and T. Yanagida, "Single-molecule imaging of EGFR signalling on the surface of living cells," *Nature cell biology*, vol. 2, no. 3, pp. 168–172, 2000.

[49] E. Korobkova, T. Emonet, J. M. Vilar, T. S. Shimizu, and P. Cluzel, "From molecular noise to behavioural variability in a single bacterium," *Nature*, vol. 428, no. 6982, pp. 574–578, 2004.

[50] T. Shibata and K. Fujimoto, "Noisy signal amplification in ultrasensitive signal transduction," *Proceedings of the National Academy of Sciences*, vol. 102, no. 2, pp. 331–336, 2005.

[51] M. Kaern, T. C. Elston, W. J. Blake, and J. J. Collins, "Stochasticity in gene expression: from theories to phenotypes," *Nature Reviews Genetics*, vol. 6, no. 6, pp. 451–464, 2005.

[52] M. Ueda and T. Shibata, "Stochastic signal processing and transduction in chemotactic response of eukaryotic cells," *Biophysical journal*, vol. 93, no. 1, pp. 11–20, 2007.

[53] J. Levine, H. Y. Kueh, and L. Mirny, "Intrinsic fluctuations, robustness, and tunability in signaling cycles," *Biophysical journal*, vol. 92, no. 12, pp. 4473–4481, 2007.

[54] S. S. Ferguson, "Evolving concepts in g protein-coupled receptor endocytosis: the role in receptor desensitization and signaling," *Pharmacological reviews*, vol. 53, no. 1, pp. 1–24, 2001.

[55] A. J. Guskjolen, "Losing connections, losing memory: Ampa receptor endocytosis as a neurobiological mechanism of forgetting," *Journal of Neuroscience*, vol. 36, no. 29, pp. 7559–7561, 2016.

[56] J. M. Smith, D. J. Salamango, M. E. Leslie, C. A. Collins, and A. Heese, "Sensitivity to flg22 is modulated by ligand-induced degradation and de novo synthesis of the endogenous flagellin-receptor FLAGELLIN-SENSING2," *Plant Physiology*, vol. 164, no. 1, pp. 440–454, 2014.

[57] T. Shi, M. Niepel, J. E. McDermott, Y. Gao, C. D. Nicora, W. B. Chrisler, L. M. Markillie, V. A. Petyuk, R. D. Smith, K. D. Rodland, P. K. Sorger, W.-J. Qian, and S. H. Wiley, "Conservation of protein abundance patterns reveals the regulatory architecture of the EGFR-MAPK pathway," *Sci. Signal.*, vol. 9, no. 436, pp. rs6–rs6, 2016.

[58] J. D. Cancel and P. B. Larsen, "Loss-of-function mutations in the ethylene receptor ETR1 cause enhanced sensitivity and exaggerated response to ethylene in Arabidopsis," *Plant Physiology*, vol. 129, no. 4, pp. 1557–1567, 2002.

[59] M. Adler, P. Szekely, A. Mayo, and U. Alon, "Optimal regulatory circuit topologies for fold-change detection," *Cell systems*, vol. 4, no. 2, pp. 171–181, 2017.

[60] D. A. McQuarrie, "Stochastic approach to chemical kinetics," *Journal of applied probability*, vol. 4, no. 3, pp. 413–478, 1967.

[61] C. V. Rao, D. M. Wolf, and A. P. Arkin, "Control, exploitation and tolerance of intracellular noise," *Nature*, vol. 420, no. 6912, pp. 231–237, 2002.

[62] A. Singh and J. P. Hespanha, "Stochastic hybrid systems for studying biochemical processes," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 368, no. 1930, pp. 4995–5011, 2010.

[63] K. R. Ghusinga, C. A. Vargas-Garcia, A. Lamperski, and A. Singh, "Exact lower and upper bounds on stationary moments in stochastic biochemical systems," *Physical biology*, vol. 14, no. 4, p. 04LT01, 2017.

[64] K. R. Ghusinga, A. Lamperski, and A. Singh, "Moment analysis of stochastic hybrid systems using semidefinite programming," *Automatica*, vol. 112, p. 108634, 2020.

[65] J. Paulsson, "Summing up the noise in gene networks," *Nature*, vol. 427, no. 6973, pp. 415–418, 2004.

[66] M. Soltani, C. A. Vargas-Garcia, D. Antunes, and A. Singh, "Intercellular variability in protein levels from stochastic expression and noisy cell cycle processes," *PLoS computational biology*, vol. 12, no. 8, 2016.

[67] A. Sorkin and M. Von Zastrow, "Endocytosis and signalling: intertwining molecular networks," *Nature reviews Molecular cell biology*, vol. 10, no. 9, pp. 609–622, 2009.

[68] J. E. Ferrell Jr, "Perfect and near-perfect adaptation in cell signaling," *Cell systems*, vol. 2, no. 2, pp. 62–67, 2016.

[69] W. Ma, A. Trusina, H. El-Samad, W. A. Lim, and C. Tang, "Defining network topologies that can achieve biochemical adaptation," *Cell*, vol. 138, no. 4, pp. 760–773, 2009.

[70] C. Y. Richard, C. G. Pesce, A. Colman-Lerner, L. Lok, D. Pincus, E. Serra, M. Holl, K. Benjamin, A. Gordon, and R. Brent, "Negative feedback that improves information transmission in yeast signalling," *Nature*, vol. 456, no. 7223, pp. 755–761, 2008.

[71] L. Yan, Q. Ouyang, and H. Wang, "Dose-response aligned circuits in signaling systems," *PloS one*, vol. 7, no. 4, 2012.

[72] B. Huot, J. Yao, B. L. Montgomery, and S. Y. He, "Growth–defense tradeoffs in plants: a balancing act to optimize fitness," *Molecular plant*, vol. 7, no. 8, pp. 1267–1287, 2014.

[73] I. T. Major, Y. Yoshida, M. L. Campos, G. Kapali, X.-F. Xin, K. Sugimoto, D. de Oliveira Ferreira, S. Y. He, and G. A. Howe, "Regulation of growth–defense balance by the jasmonate zim-domain (jaz)-myc transcriptional module," *New Phytologist*, vol. 215, no. 4, pp. 1533–1547, 2017.

[74] N. Olsman and L. Goentoro, "Allosteric proteins as logarithmic sensors," *Proceedings of the National Academy of Sciences*, vol. 113, no. 30, pp. E4423–E4430, 2016.

[75] G. Dixit, J. B. Kelley, J. R. Houser, T. C. Elston, and H. G. Dohlman, "Cellular noise suppression by the regulator of G protein signaling Sst2," *Molecular cell*, vol. 55, no. 1, pp. 85–96, 2014.

[76] C. Gardiner, *Stochastic methods*, vol. 4. Springer Berlin, 2009.

[77] M. Ullah and O. Wolkenhauer, "Stochastic approaches in systems biology," *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, vol. 2, no. 4, pp. 385–397, 2010.

[78] J. Hasenauer, V. Wolf, A. Kazeroonian, and F. J. Theis, "Method of conditional moments (MCM) for the chemical master equation," *Journal of mathematical biology*, vol. 69, no. 3, pp. 687–735, 2014.

Supplementary Information

S1 Comparison between different definitions of response time

Response time is a measure of the time it takes for a signaling output to reach its steady-state. In Eq. (6) of the main text, we defined the response time as the center of mass of the response curve. However, there are several interrelated definitions of response time. Here we provide a comparison between them. Towards that end, we use a model of a simple one-tier switch. Consider a protein that transitions between two states A and A^* as

$$A \underset{k_2}{\overset{k_1 S}{\rightleftarrows}} A^*. \tag{S1.1}$$

Let $A(t)$ and $A^*(t)$, respectively, denote the number of molecules that are in states A and A^* at time t. We assume that the total number of molecules is conserved, i.e., $A_T = A(t) + A^*(t)$. We quantify the signaling through the switch by $A^*(t)$, i.e., the number of molecules in the state A^*. The ordinary differential equation (ODE) governing the dynamics of A^* is:

$$\frac{dA^*}{dt} = k_1 S(A_T - A^*) - k_2 A^*. \tag{S1.2}$$

The solution to this ODE is given by

$$A^*(t) = A^*(0)e^{-(k_1 S+k_2)t} + \frac{k_1 S A_T}{k_1 S + k_2}\left(1 - e^{-(k_1 S+k_2)t}\right), \tag{S1.3}$$

where $A^*(0) < A_T$ is the initial condition. As $t \to \infty$, A^* approaches its steady-state value which is given by

$$\overline{A^*} = \frac{k_1 S A_T}{k_1 S + k_2}. \tag{S1.4}$$

Recall the definition of response time from Eq. (6) in the main text

$$\mathcal{T}_{A^*} = \frac{\int_0^\infty t\left|\overline{A^*} - A^*(t)\right| dt}{\int_0^\infty \left|\overline{A^*} - A^*(t)\right| dt}. \tag{S1.5}$$

To compute the integrals in the numerator and the denominator, we first note that

$$\overline{A^*} - A^*(t) = \left(\frac{k_1 S A_T}{k_1 S + k_2} - A^*(0)\right) e^{-(k_1 S+k_2)t}. \tag{S1.6}$$

Because $k_1 S > 0$ and $k_2 > 0$, $\int_0^\infty e^{-(k_1 S+k_2)t}dt$ and $\int_0^\infty t\, e^{-(k_1 S+k_2)t}dt$ converge. These integrals are equal to $1/(k_1 S + k_2)$ and $1/(k_1 S + k_2)^2$, respectively. Using these integrals, Eq. (S1.5) gives

$$\mathcal{T}_{A^*} = \frac{\frac{k_1 S A_T}{k_1 S+k_2} - A^*(0)}{(k_1 S + k_2)^2}\frac{k_1 S + k_2}{\frac{k_1 S A_T}{k_1 S+k_2} - A^*(0)} = \frac{1}{k_1 S + k_2}. \tag{S1.7}$$

We can thus deduce that if the response is determined by a single kinetic step, the response time defined above is reciprocal of the rate constant for that step. It is also worth noting that the ratio is well-defined only when $A^*(0) \neq \frac{k_1 S A_T}{k_1 S+k_2} = \overline{A^*}$.

Another class of definitions of response time are based on the time it takes for the response to start from $A^*(0)$ and reduce its deviation from its steady-state by a factor $0 < \eta < 1$. More specifically, we define $\mathcal{T}_\eta$ as the solution to the following equation

$$A^*(0)e^{-(k_1 S+k_2)\mathcal{T}_\eta} + \frac{k_1 S A_T}{k_1 S + k_2}\left(1 - e^{-(k_1 S+k_2)\mathcal{T}_\eta}\right) = A^*(0) + \eta\left(\frac{k_1 S A_T}{k_1 S + k_2} - A^*(0)\right). \tag{S1.8}$$

For $A^*(0) \neq \frac{k_1 S A_T}{k_1 S+k_2}$, the above equation reduces to

$$1 - e^{-(k_1 S+k_2)\mathcal{T}_\eta} = \eta, \tag{S1.9}$$

which has a straightforward solution

$$T_\eta = -\frac{\log(1-\eta)}{k_1 S + k_2}. \tag{S1.10}$$

Notably, the response time is set by $1/(k_1 S + k_2)$ up to a scale which depends on the specific value of η. We discuss three cases. First, setting $\eta = 1/2$ corresponds to the time at which half of the deviation from the steady-state has been reduced. The corresponding response time is given by

$$\mathcal{T}_{50\%} = \frac{\log 2}{k_1 S + k_2}. \tag{S1.11a}$$

Second, $\eta = (e-1)/e \approx 0.632$ is also frequently used for which we obtain

$$\mathcal{T}_{63.2\%} = \frac{1}{k_1 S + k_2}. \tag{S1.11b}$$

Lastly, a third definition concerns computing the time it takes for the response to travel from 10% to 90% of the difference between its initial value $A^*(0)$ and steady-state $A^* = \frac{k_1 S A_T}{k_1 S + k_2}$. In this case, we get

$$\mathcal{T}_{90\%} - \mathcal{T}_{90\%} = \frac{\log 9}{k_1 S + k_2}. \tag{S1.11c}$$

S2 Transient solution and response time of two-tier cascades

In this section, we consider two-tier cascades of Fig. 1. Because activation and derepression are special cases of the concerted mechanism, we concern ourselves only with the ODEs of a concerted mechanism here.

The ordinary differential equations (ODEs) that govern the dynamics are

$$\frac{dX^*}{dt} = k_1 S(X_T - X^*) - k_2 X^*, \tag{S2.1a}$$

$$\frac{dY^*}{dt} = \left(k_3 + k_5 X^*\right)\left(Y_T - Y^*\right) - \left(k_4 + k_6\left(X_T - X^*\right)\right) Y^*, \tag{S2.1b}$$

with initial conditions

$$X^*(0) = 0, \quad Y^*(0) = \frac{k_3 Y_T}{k_3 + k_4 + k_6 X_T}. \tag{S2.1c}$$

The steady-states of X^* and Y^* are computed by setting the derivatives equation to zero.

$$\overline{X^*} = \frac{k_1 S X_T}{k_1 S + k_2}, \quad \overline{Y^*} = \frac{k_3 + k_5 \frac{k_1 S X_T}{k_1 S + k_2}}{k_3 + k_4 + k_5 \frac{k_1 S X_T}{k_1 S + k_2} + k_6 \frac{k_2 X_T}{k_1 S + k_2}} Y_T. \tag{S2.2}$$

Recall that plugging $k_6 = 0$ and $k_5 = 0$, respectively, result in ODEs for the activation and derepression mechanisms. Furthermore, we term the special case $k_5 = k_6$ as perfect concerted mechanism, where the activation strength and repression strengths match.

S2-A Reducing number of parameters

The dynamics described by Eq. (S2.1) consists of 8 parameters. We can group these parameters by choosing appropriate scaling and thereby reduce the number of parameters. To this end, we begin by defining scaled variables $x^* = X^*/X_0$, $y^* = Y^*/X_0$ and $s = S/S_0$, where X_0, Y_0 and S_0 are appropriate scales. We can rewrite the dynamics as

$$\frac{dx^* X_0}{dt} = k_1 s S_0 (X_T - x^* X_0) - k_2 x^* X_0, \tag{S2.3a}$$

$$\frac{dy^* Y_0}{dt} = \left(k_3 + k_5 x^* X_0\right)\left(Y_T - y^* Y_0\right) - \left(k_4 + k_6\left(X_T - x^* X_0\right)\right) y^* Y_0, \tag{S2.3b}$$

which simplifies to

$$\frac{dx^*}{dt} = k_2\left(\frac{k_1 S_0}{k_2} s\left(\frac{X_T}{X_0} - x^*\right) - x^*\right), \tag{S2.4a}$$

$$\frac{dy^*}{dt} = \left(k_3 + k_5 x^* X_0\right)\left(\frac{Y_T}{Y_0} - y^*\right) - \left(k_4 + k_6 X_0\left(\frac{X_T}{X_0} - x^*\right)\right) y^*. \tag{S2.4b}$$

We choose the scales such that the number of parameters reduces. One choice for the scales is $S_0 = \frac{k_2}{k_1}$, $X_0 = X_T$, and $Y_0 = Y_T$. In addition, we also grouped parameters $k_5' = k_5 X_T$ and $k_6' = k_6 X_T$. The ODEs now become

$$\frac{dx^*}{dt} = k_2\left(s\left(1 - x^*\right) - x^*\right), \tag{S2.5a}$$

$$\frac{dy^*}{dt} = \left(k_3 + k_5' x_a\right)\left(1 - y^*\right) - \left(k_4 + k_6'\left(1 - x^*\right)\right) y^*. \tag{S2.5b}$$

Rearranging terms, we obtain

$$\frac{dx^*}{dt} = k_2\left(s - (1+s)x^*\right), \tag{S2.6a}$$

$$\frac{dy^*}{dt} = k_3 + k_5' x^* - \left(k_3 + k_4 + k_6' + k_5' x^* - k_6' x^*\right) y^*. \tag{S2.6b}$$

Lastly, the initial conditions and the steady-state solutions to this reduced system are given by

$$x^*(0) = 0, \qquad y^*(0) = \frac{k_3}{k_3 + k_4 + k_6'}, \tag{S2.6c}$$

$$\overline{x^*} = \frac{s}{1+s}, \qquad \overline{y^*} = \frac{k_3 + k_5' \frac{s}{1+s}}{k_3 + k_4 + k_5' \frac{s}{1+s} + \frac{k_6'}{1+s}}. \tag{S2.6d}$$

It is worth noting that this system is not fully non-dimensionalized because the time t still has dimension. Two appropriate scales for this are k_2 and $k_3 + k_4 + k_6'$. However, we choose to not do this.

S2-B Transient solution

Analytical solutions for nonlinear ODEs such as those in Eq. (S2.1) are usually hard to compute. However, a careful look at Eq. (S2.1) shows that the nonlinear term is $(k_5 - k_6)X^* Y^*$. Thus for a special case when $k_5 = k_6$ (perfect concerted mechanism), the system is linear, which exhibits analytical solution. To transient solution for other cases can be computed numerically. We also provide an approximate solution using linearization around the steady-state solution $(\overline{X^*}, \overline{Y^*})$.

S2–B–a Transient solution for a perfect concerted model

To compute the analytical solution, let us use the reduced equations in Eq. (S2.6). Note that $k_5' = k_5 X_T$ and $k_6' = k_6 X_T$. The requirement for the perfect concerted model ($k_5 - k_6 = 0$) gives $k_5' - k_6' = (k_5 - k_6)X_T = 0$. Substituting this in Eq. (S2.6) results in

$$\frac{dx^*}{dt} = k_2 \left(s - (1+s)x^*\right), \tag{S2.7a}$$

$$\frac{dy^*}{dt} = k_3 + k_6' x^* - \left(k_3 + k_4 + k_6'\right) y^*, \tag{S2.7b}$$

with initial condition $(x^*(0), y^*(0)) = \left(0, \frac{k_3}{k_3+k_4+k_6'}\right)$. For now, we do not plug these initial values though.

Let $\mathcal{X}^*[\omega]$ and $\mathcal{Y}^*[\omega]$ respectively denote the Laplace transforms of $x^*(t)$ and $y^*(t)$. Then we have that

$$\omega \mathcal{X}^*[\omega] - x^*(0) = \frac{k_2 s}{\omega} - k_2(1+s)\mathcal{X}^*[\omega], \tag{S2.8a}$$

$$\omega \mathcal{Y}^*[\omega] - y^*(0) = \frac{k_3}{\omega} + k_6' \mathcal{X}^*[\omega] - \left(k_3 + k_4 + k_6'\right) \mathcal{Y}^*[\omega]. \tag{S2.8b}$$

Solving for $\mathcal{X}^*[\omega]$ and $\mathcal{Y}^*[\omega]$

$$\mathcal{X}^*[\omega] = \frac{s}{1+s}\left(\frac{k_2(1+s)}{\omega(\omega + k_2(1+s))}\right) + \frac{x^*(0)}{\omega + k_2(1+s)}, \tag{S2.9a}$$

$$\mathcal{Y}^*[\omega] = \frac{k_3}{\omega(\omega + k_3 + k_4 + k_6')} + \frac{y^*(0)}{\omega + k_3 + k_4 + k_6'} + \frac{k_2 k_6' s}{\omega(\omega + k_2(1+s))(\omega + k_3 + k_4 + k_6')} + \frac{k_6' x^*(0)}{(\omega + k_2(1+s))(\omega + k_3 + k_4 + k_6')}. \tag{S2.9b}$$

Taking inverse Laplace transform, we obtain

$$x^*(t) = \frac{s}{1+s}\left(1 - e^{-k_2(1+s)t}\right) + x^*(0)e^{-k_2(1+s)t}, \tag{S2.10a}$$

$$y^*(t) = \frac{k_3 + k_6' \frac{s}{1+s}}{k_3 + k_4 + k_6'} - \frac{k_6' \frac{s}{1+s}}{k_3 + k_4 + k_6'} \frac{k_2(1+s)e^{-(k_3+k_4+k_6')t} - (k_3 + k_4 + k_6')e^{-k_2(1+s)t}}{k_2(1+s) - (k_3 + k_4 + k_6')}. \tag{S2.10b}$$

In the special case when the timescales match, i.e., $k_2(1+s) = k_3 + k_4 + k_6'$, the solution $y^*(t)$ becomes

$$y^*(t) = \frac{k_3 + k_6' \frac{s}{1+s}}{k_3 + k_4 + k_6'} - \frac{k_6' \frac{s}{1+s}}{k_3 + k_4 + k_6'} e^{-(k_3+k_4+k_6')t}(1 + (k_3 + k_4 + k_6')t). \tag{S2.10c}$$

We can write this solution back in terms of unscaled parameters of Eq. (S2.1) as

$$X^*(t) = \frac{k_1 S X_T}{k_1 S + k_2}\left(1 - e^{-(k_1 S + k_2)t}\right), \tag{S2.11a}$$

$$Y^*(t) = \frac{k_3 + k_6 \frac{k_1 S X_T}{k_1 S + k_2}}{k_3 + k_4 + k_6 X_T} - \frac{k_6 \frac{k_1 S X_T}{k_1 S + k_2}}{k_3 + k_4 + k_6 X_T} \frac{(k_1 S + k_2)e^{-(k_3+k_4+k_6 X_T)t} - (k_3 + k_4 + k_6 X_T)e^{-(k_1 S + k_2)t}}{k_1 S + k_2 - (k_3 + k_4 + k_6 X_T)}. \tag{S2.11b}$$

Finally, the degenerate case where $k_1 S + k_2 = k_3 + k_4 + k_6 X_T$, we have

$$Y^*(t) = \frac{k_3 + k_6 \frac{k_1 S X_T}{k_1 S + k_2}}{k_3 + k_4 + k_6 X_T} - \frac{k_6 \frac{k_1 S X_T}{k_1 S + k_2}}{k_3 + k_4 + k_6 X_T} e^{-(k_3+k_4+k_6 X_T)t}\left(1 + (k_3 + k_4 + k_6 X_T)t\right). \tag{S2.11c}$$

Next, we provide approximate solutions to the nonlinear ODE system. The first approach is based on linearizing the system around its steady-state solution. The second approach is drawn from perturbation theory.

S2–B–b Approximate transient solution using linearization

As before, we work with the reduced system of ODEs in Eq. (S2.1) for ease of analysis. The system contains nonlinear term x^*y^*, which can be linearized around the steady-state solution $(\overline{x^*}, \overline{y^*})$ as

$$x^*y^* \approx \overline{y^*}x^* + \overline{x^*}y^* - \overline{x^*}\,\overline{y^*}. \tag{S2.12}$$

Substituting this for the nonlinear term in Eq. (S2.1), we get the following

$$\frac{dx^*}{dt} = k_2\left(s - (1+s)x^*\right), \tag{S2.13a}$$

$$\frac{dy^*}{dt} = \left(k_3 + (k_5' - k_6')\overline{x^*}\,\overline{y^*}\right) + \left(k_5' - (k_5' - k_6')\overline{y^*}\right)x^* - (k_3 + k_4 + k_6' + (k_5' - k_6')\overline{x^*})y^*. \tag{S2.13b}$$

These ODEs are similar to those in Eq. (S2.7). The solution, thus immediately follows Eq. (S2.10).

$$x^*(t) = \frac{s}{1+s}\left(1 - e^{-k_2(1+s)t}\right), \tag{S2.14a}$$

$$y^*(t) = \frac{k_3 + k_5'\frac{s}{1+s}}{k_3 + k_4 + k_6' + (k_5' - k_6')\frac{s}{1+s}} - \frac{s}{1+s}\frac{k_4k_5' + k_3k_6' + k_5'k_6'}{\left(k_3 + k_4 + k_6' + (k_5' - k_6')\frac{s}{1+s}\right)^2} \times \frac{k_2(1+s)e^{-\left(k_3+k_4+k_6'+(k_5'-k_6')\frac{s}{1+s}\right)t} - \left(k_3 + k_4 + k_6' + (k_5' - k_6')\frac{s}{1+s}\right)e^{-k_2(1+s)t}}{k_2(1+s) - \left(k_3 + k_4 + k_6' + (k_5' - k_6')\frac{s}{1+s}\right)}. \tag{S2.14b}$$

The special case when the timescales match may be computed by taking the limit of the above solution.

S2-C Response time

In this section, we compute the response time for $Y^*(t)$ using the center of mass of the response curve:

$$\mathcal{T}_{Y^*} = \frac{\int_0^\infty t\left|\overline{Y^*} - Y^*(t)\right| dt}{\int_0^\infty \left|\overline{Y^*} - Y^*(t)\right| dt}. \tag{S2.15}$$

Because $y^* = Y^*/Y_T$, $\mathcal{T}_{Y^*} = \mathcal{T}_{y^*}$. This may be easily seen by plugging $Y^* = y^*Y_T$ in the definition above.

S2–C–a Response time for a perfect concerted mechanism

Let us begin by compute the response time for a perfect concerted mechanism. From Eq. (S2.10), we have

$$y^*(t) = \frac{k_3 + k_6'\frac{s}{1+s}}{k_3 + k_4 + k_6'} - \frac{k_6'\frac{s}{1+s}}{k_3 + k_4 + k_6'}\frac{k_2(1+s)e^{-(k_3+k_4+k_6')t} - (k_3 + k_4 + k_6')e^{-k_2(1+s)t}}{k_2(1+s) - (k_3 + k_4 + k_6')}, \tag{S2.16a}$$

with steady-state

$$\overline{y^*} = \frac{k_3 + k_6'\frac{s}{1+s}}{k_3 + k_4 + k_6'}. \tag{S2.16b}$$

Computing the numerator for the response time in Eq. (S2.15) gives

$$\int_0^\infty t(\overline{y^*} - y^*(t))dt = k_6'\frac{s}{1+s}\frac{k_2^2(1+s)^2 + k_2(1+s)(k_3 + k_4 + k_6') + (k_3 + k_4 + k_6')^2}{k_2^2(1+s)^2(\left(k_3 + k_4 + k_6'\right)^3}. \tag{S2.17}$$

Likewise, the denominator is

$$\int_0^\infty (\overline{y^*} - y^*(t))dt = k_6'\frac{s}{1+s}\frac{k_2(1+s) + k_3 + k_4 + k_6'}{k_2(1+s)(\left(k_3 + k_4 + k_6'\right)^2}. \tag{S2.18}$$

Thus the response time is

$$\mathcal{T}_{y^*} = \frac{1}{k_2(1+s)} + \frac{1}{k_3 + k_4 + k_6'} - \frac{1}{k_2(1+s) + k_3 + k_4 + k_6'} \tag{S2.19a}$$

$$= \frac{1}{k_1S + k_2} + \frac{1}{k_3 + k_4 + k_6X_T} - \frac{1}{k_1S + k_2 + k_3 + k_4 + k_6X_T}. \tag{S2.19b}$$

S2–C–b Response time for the linear approximation

The response time for the linear approximation follows the same computations as the perfect concerted mechanism. The final formula is given by

$$\mathcal{T}_{y^*} = \frac{1}{k_2(1+s)} + \frac{1}{k_3 + k_4 + k_5'\frac{s}{1+s} + \frac{k_6'}{1+s}} - \frac{1}{k_2(1+s) + k_3 + k_4 + k_6'}. \tag{S2.20}$$

S2–C–c Numerical computation of the response time

One convenience in using the center of mass definition of the response time

$$\mathcal{T}_{Y^*} = \frac{\int_0^\infty t\left|\overline{Y^*} - Y^*(t)\right| dt}{\int_0^\infty \left|\overline{Y^*} - Y^*(t)\right| dt} \tag{S2.21}$$

is that it can be computer numerically via solution of an augmented ODE. To see this, let us augment three states $V_1(t)$, $V_2(t)$ and V_3 to the original ODE system

$$\frac{dX^*}{dt} = k_1 S(X_T - X^*) - k_2 X^*, \tag{S2.22a}$$

$$\frac{dY^*}{dt} = \left(k_3 + k_5 X^*\right)\left(Y_T - Y^*\right) - \left(k_4 + k_6\left(X_T - X^*\right)\right) Y^*, \tag{S2.22b}$$

$$\frac{dV_1}{dt} = \overline{Y^*} - Y^*, \tag{S2.22c}$$

$$\frac{dV_2}{dt} = 1, \tag{S2.22d}$$

$$\frac{dV_3}{dt} = V_1 V_2. \tag{S2.22e}$$

with initial conditions

$$\left(X^*(0), Y(0), V_1(0), V_2(0), V_3(0)\right) = \left(0, \frac{k_3 Y_T}{k_3 + k_4 + k_6 X_T}, \frac{\left(k_3 + k_5 \frac{k_1 S X_T}{k_1 S + k_2}\right) Y_T}{k_3 + k_4 + k_5 \frac{k_1 S X_T}{k_1 S + k_2} + k_6 \frac{k_2 X_T}{k_1 S + k_2}} - \frac{k_3 Y_T}{k_3 + k_4 + k_6 X_T}, 0, 0\right) \tag{S2.22f}$$

Then the state $V_1(t)$ computes the integral in the denominator upto a time t, $V_2(t)$ tracks the time, and $V_3(t)$ computes the numerator upto time horizon t. If we choose t to be large enough such that the system has reached saturation, then $\frac{V_3(t)}{V_1(t)}$ computes the response time. It is easy to see that the approximation gets better with a larger t.

S3 Stochastic analysis of two-tier cascades

Here we consider a two-tier model for signal transduction as described in Table 1 in the main text. Let $P_{m,n}(t)$ denote the probability of finding m molecules of X^* and n molecules of R^* at time t. Then, we can write the chemical master equation (CME) that describes the time evolution of $P_{m,n}$

$$\begin{aligned}\frac{dP_{m,n}(t)}{dt} &= k_1 S(X_T - (m-1))P_{m-1,n} + k_2(m+1)P_{m+1,n} + k_3(Y_T - (n-1))P_{m,n-1} \\ &+ k_5 m(Y_T - (n-1))P_{m,n-1} + k_4(n+1)P_{m,n+1} + k_6(X_T - m)(n+1)P_{m,n+1} \\ &- \left(k_1 S(X_T - m) + k_2 m + k_3(Y_T - n) + k_5 m(R_T - n) + k_4 n + k_6(X_T - m)n\right) P_{m,n},\end{aligned} \tag{S3.1}$$

where $m = 0, 1, \ldots, X_T$ and $n = 0, \ldots, Y_T$ [60, 76]. It is often difficult to analytically solve CME. Because the dynamics of X^* is linear and it does not depend upon Y^*, it is possible to provide an analytical solution P_m. As for $P_{m,n}$, we only provide approximate and exact computations of its first two moments.

S3-A Stochastic solution to receptor dynamics

The CME that governs the time evolution of $P_n(t)$ is:

$$\frac{dP_m(t)}{dt} = k_1 S\left(X_T - (m-1)\right) P_{m-1}(t) + k_2(m+1)P_{m+1}(t) - \left(k_1 S\left(X_T - m\right) + k_2 m\right) P_m(t). \tag{S3.2}$$

We define a generating function

$$G(z) = \sum_{m=0}^{\infty} z^m P_m, \quad |z| \leq 1 \tag{S3.3}$$

to solve Eq. (S3.2). Multiplying both sides by z^m and summing over m yields

$$\frac{\partial G}{\partial t} = k_1 S X_T \sum_{m=0}^{\infty} z^m P_{m-1} - k_1 S \sum_{m=0}^{\infty} z^m (m-1) P_{m-1} + k_2 \sum_{m=0}^{\infty} z^m (m+1) P_{m+1} - k_1 S X_T \sum_{m=0}^{\infty} z^m P_m + (k_1 S - k_2) \sum_{m=0}^{\infty} z^m m P_m. \tag{S3.4}$$

The above equation becomes the following partial differential equation (PDE)

$$\frac{\partial G}{\partial t} = k_1 S X_T (z-1) G + \left(-k_1 S z^2 + k_2 + (k_1 S - k_2) z\right) \frac{\partial G}{\partial z}. \tag{S3.5}$$

We solve this PDE using method of characteristics, assuming the initial condition $G(z,0) = 1$ which corresponds to 0 molecules of X^*. The solution is given by

$$G(z,t) = \left(1 - \frac{k_1 S}{k_1 S + k_2}\left(1 - e^{-(k_1 S + k_2)t}\right) + \frac{k_1 S}{k_1 S + k_2} z \left(1 - e^{-(k_1 S + k_2)t}\right)\right)^{A_T}. \tag{S3.6}$$

Using Binomial theorem, the above expression can be written as

$$G(z,t) = \sum_{m=0}^{X_T} \binom{X_T}{m} \left(\frac{k_1 S}{k_1 S + k_2}\left(1 - e^{-(k_1 S + k_2)t}\right)\right)^m \left(1 - \frac{k_1 S}{k_1 S + k_2}\left(1 - e^{-(k_1 S + k_2)t}\right)\right)^{X_T - m} z^m. \tag{S3.7}$$

The probability $P_m(t)$ is given by the coefficient of z^m

$$P_m(t) = \binom{X_T}{m} \left(\frac{k_1 S}{k_1 S + k_2}\left(1 - e^{-(k_1 S + k_2)t}\right)\right)^m \left(1 - \frac{k_1 S}{k_1 S + k_2}\left(1 - e^{-(k_1 S + k_2)t}\right)\right)^{X_T - m}. \tag{S3.8}$$

The stationary distribution $\overline{P_m}$ is computed by taking limit $t \to \infty$

$$\overline{P_m} = \binom{X_T}{m} \left(\frac{k_1 S}{k_1 S + k_2}\right)^m \left(1 - \frac{k_1 S}{k_1 S + k_2}\right)^{A_T - m}, \tag{S3.9}$$

which is a Binomial distribution with parameters X_T and $\frac{k_1 S}{k_1 S + k_2}$ [77]. The stationary moments of this distribution are given by

$$\langle X^* \rangle = \frac{k_1 S X_T}{k_1 S + k_2} \tag{S3.10a}$$

$$\langle X^{*2} \rangle - \langle X^* \rangle^2 = \frac{k_1 S k_2 X_T}{(k_1 S + k_2)^2} \tag{S3.10b}$$

$$CV^2_{X^*} = \frac{\langle X^{*2} \rangle - \langle X^* \rangle^2}{\langle X^* \rangle^2} = \frac{k_2}{k_1 S X_T}. \tag{S3.10c}$$

S3-B Moment dynamics

We are specifically concerned with moments of the two-tier model. To this end, we use the well-established approach of using the ODEs that govern the moment dynamics (e.g., see [62, 63]). A generic moment may be written as

$$\begin{aligned}\frac{d\langle X^{*m_1} Y^{*m_2}\rangle}{dt} &= \left\langle k_1 S(X_T - X^*)\left((X^*+1)^{m_1} Y^{*m_2} - X^{*m_1} Y^{*m_2}\right)\right\rangle + \left\langle k_2 X^* \left((X^*-1)^{m_1} Y^{*m_2} - X^{*m_1} Y^{*m_2}\right)\right\rangle + \\ &\left\langle \left(k_3 + k_5 X^*\right)(Y_T - Y^*)\left(X^{*m_1}(Y^*+1)^{m_2} - X^{*m_1} Y^{*m_2}\right)\right\rangle + \left\langle \left(k_4 + k_6(X_T - X^*)\right) Y^* \left(X^{*m_1}(Y^*-1)^{m_2} - X^{*m_1} Y^{*m_2}\right)\right\rangle.\end{aligned} \tag{S3.11}$$

Here we have used $\langle . \rangle$ to denote the expected value of a random variable. Our focus in this work is to compute the first two moments in steady-state. However, due to the nonliearity $X^* Y^*$ in these equations, the moment dynamics is not closed in that a lower-order moment depends upon a higher-order moment [62, 63]. It turns out that for the special case $k_5 = k_6$ (perfect concerted model), the moments may be computed exactly. We provide approximate formulas for moments using a linear approximation when $k_5 \neq k_6$.

S3–B–a Moment computation for a perfect concerted model

For the concerted model, $k_5 = k_6 = k$. Let us write moment dynamics for first two moments.

$$\frac{d\langle X^* \rangle}{dt} = k_1 S X_T - (k_1 S + k_2)\langle X^* \rangle, \tag{S3.12a}$$

$$\frac{d\langle Y^* \rangle}{dt} = k_3 Y_T + k Y_T \langle X^* \rangle - (k_3 + k_4 + k X_T)\langle Y^* \rangle, \tag{S3.12b}$$

$$\frac{d\langle X^{*2} \rangle}{dt} = k_1 S X_T + (k_1 S(2X_T - 1) + k_2)\langle X^* \rangle - 2(k_1 S + k_2)\langle X^{*2} \rangle, \tag{S3.12c}$$

$$\frac{d\langle X^* Y^* \rangle}{dt} = k_3 Y_T \langle X^* \rangle + k_1 S X_T \langle Y^* \rangle + k Y_T \langle X^{*2} \rangle - (k_1 X_T + k_2 + k_3 + k_4 + k X_T)\langle X^* Y^* \rangle, \tag{S3.12d}$$

$$\frac{d\langle Y^{*2} \rangle}{dt} = k_3 Y_T + k Y_T \langle X^* \rangle + (k_3(2Y_T - 1) + k_4 + k X_T)\langle Y^* \rangle + 2k(Y_T - 1)\langle X^* Y^* \rangle - 2(k_3 + k_4 + k X_T)\langle Y^{*2} \rangle. \tag{S3.12e}$$

We can solve for steady-state moments by setting each of the derivatives equal to zero. For example, the means are given by

$$\langle X^* \rangle = \frac{k_1 S}{k_1 S + k_2} X_T, \tag{S3.13a}$$

$$\langle Y^* \rangle = \frac{k_3 + k\langle X^* \rangle}{k_3 + k_4 + k X_T} Y_T = \frac{k_3 + k\frac{k_1 S}{k_1 S + k_2} X_T}{k_3 + k_4 + k X_T} Y_T. \tag{S3.13b}$$

Next, we compute second order moments. $\left\langle X^{*2}\right\rangle$ is given by

$$\left\langle X^{*2}\right\rangle = \left(\frac{k_1 S}{k_1 S + k_2} X_T\right)^2 + \frac{k_1 S k_2 X_T}{(k_1 S + k_2)^2}, \tag{S3.14}$$

where the first term is $\left\langle X^*\right\rangle^2$. The cross moment $\left\langle X^* Y^*\right\rangle$ is

$$\left\langle X^* Y^*\right\rangle = \frac{k_1 S k_2 k X_T Y_T}{(k_1 S + k_2)^2 (k_1 S + k_2 + k_3 + k_4 + k X_T)} + \left(\frac{k_1 S}{k_1 S + k_2} X_T\right)\left(\frac{k_3 + k \frac{k_1 S}{k_1 S + k_2} X_T}{k_3 + k_4 + k X_T} Y_T\right). \tag{S3.15}$$

Here the second term is $\left\langle X^*\right\rangle \left\langle Y^*\right\rangle$. Finally, the second order moment $\left\langle Y^{*2}\right\rangle$ in terms of the other moments is

$$\left\langle Y^{*2}\right\rangle = \frac{k_3 Y_T}{2(k_3 + k_4 + k X_T)} + \frac{k Y_T \left\langle X^*\right\rangle}{2(k_3 + k_4 + k X_T)} + \frac{(k_3(2Y_T - 1) + k_4 + k X_T)\left\langle Y^*\right\rangle}{2(k_3 + k_4 + k X_T)} + \frac{2k(Y_T - 1))\left\langle X^* Y^*\right\rangle}{2(k_3 + k_4 + k X_T)} \tag{S3.16a}$$

$$= \frac{k_3 Y_T + k_4 + k X_T}{k_3 + k_4 + k X_T}\left\langle Y^*\right\rangle + \frac{k(Y_T - 1)\left\langle X^* Y^*\right\rangle}{k_3 + k_4 + k X_T}. \tag{S3.16b}$$

Using the moments computed above, we can compute the centered moments. For example, the variance of X^* is

$$\left\langle X^{*2}\right\rangle - \left\langle X^*\right\rangle^2 = \frac{k_1 S k_2 X_T}{(k_1 S + k_2)^2}, \tag{S3.17}$$

the centered cross moment is

$$\left\langle X^* Y^*\right\rangle - \left\langle X^*\right\rangle \left\langle Y^*\right\rangle = \frac{k_1 S k_2 k X_T Y_T}{(k_1 S + k_2)^2 (k_1 S + k_2 + k_3 + k_4 + k X_T)}, \tag{S3.18}$$

and the variance of Y^* is

$$\begin{aligned}\left\langle Y^{*2}\right\rangle - \left\langle Y^*\right\rangle^2 &= \left(\frac{k_3 Y_T + k_4 + k X_T}{k_3 + k_4 + k X_T}\right)\left(\frac{k_3 + k \frac{k_1 S}{k_1 S + k_2} X_T}{k_3 + k_4 + k X_T} Y_T\right) \\ &+ \frac{k(Y_T - 1)}{k_3 + k_4 + k X_T}\left(\frac{k_1 S k_2 k X_T Y_T}{(k_1 S + k_2)^2 (k_1 S + k_2 + k_3 + k_4 + k X_T)} + \frac{k_1 S X_T}{k_1 S + k_2}\frac{k_3 + k \frac{k_1 S}{k_1 S + k_2} X_T}{k_3 + k_4 + k X_T} Y_T\right) - \left(\frac{k_3 + k \frac{k_1 S}{k_1 S + k_2} X_T}{k_3 + k_4 + k X_T} Y_T\right)^2.\end{aligned} \tag{S3.19}$$

We use the centered moments computed above to quantify noise in X^* and Y^* using coefficient of variation squared.

Coefficient of variation squared. Let $CV^2_{X^*}$ and $CV^2_{Y^*}$ respectively are the coefficient of variation squared for X^* and Y^*. Then

$$CV^2_{X^*} = \frac{\left\langle X^{*2}\right\rangle - \left\langle X^*\right\rangle^2}{\left\langle X^*\right\rangle^2} = \frac{k_2}{k_1 S X_T}, \tag{S3.20}$$

and

$$CV^2_{Y^*} = \frac{\left\langle Y^{*2}\right\rangle - \left\langle Y^*\right\rangle^2}{\left\langle Y^*\right\rangle^2} \tag{S3.21}$$

$$\begin{aligned}&= \left(\frac{k_3 Y_T + k_4 + k X_T}{k_3 + k_4 + k X_T}\right)\left(\frac{k_3 + k \frac{k_1 S}{k_1 S + k_2} X_T}{k_3 + k_4 + k X_T} Y_T\right)\frac{1}{Y_T^2}\left(\frac{k_3 + k_4 + k X_T}{k_3 + k \frac{k_1 S}{k_1 S + k_2} X_T}\right)^2 \\ &+ \frac{1}{Y_T^2}\left(\frac{k_3 + k_4 + k X_T}{k_3 + k \frac{k_1 S}{k_1 S + k_2} X_T}\right)^2 \frac{k(Y_T - 1)}{k_3 + k_4 + k X_T}\left(\frac{k_1 S k_2 k X_T Y_T}{(k_1 S + k_2)^2 (k_1 S + k_2 + k_3 + k_4 + k X_T)}\right) \\ &+ \frac{1}{Y_T^2}\left(\frac{k_3 + k_4 + k X_T}{k_3 + k \frac{k_1 S}{k_1 S + k_2} X_T}\right)^2 \frac{k(Y_T - 1)}{k_3 + k_4 + k X_T}\left(\frac{k_1 S X_T}{k_1 S + k_2}\frac{k_3 + k \frac{k_1 S}{k_1 S + k_2} X_T}{k_3 + k_4 + k X_T} Y_T\right) - 1\end{aligned} \tag{S3.22}$$

On simplifying, we get

$$\begin{aligned}CV^2_{Y^*} &= \frac{1}{Y_T}\frac{k_3 Y_T + k_4 + k X_T}{k_3 + k \frac{k_1 S}{k_1 S + k_2} X_T} + \frac{Y_T - 1}{Y_T}\frac{k\,(k_3 + k_4 + k X_T)\,(k_1 S k_2 k X_T)}{(k_1 S + k_2)^2\left(k_3 + k \frac{k_1 S}{k_1 S + k_2} X_T\right)^2 (k_1 S + k_2 + k_3 + k_4 + k X_T)} \\ &+ \frac{Y_T - 1}{Y_T}\frac{k k_1 S X_T}{\left(k_3 + k \frac{k_1 S}{k_1 S + k_2} X_T\right)(k_1 S + k_2)}.\end{aligned} \tag{S3.23}$$

Decomposing the coefficient of variation squared into different sources. We expect that $CV^2_{Y^*}$ has two sources of noise: activation/deactivation events for X^* and activation/deactivation events for Y^*. To tease out the contribution from activation/deactivation events for Y^*, we consider a scenario the dynamics of X^* is deterministic. In this case, the moment dynamics is given by

$$\frac{dX^*}{dt} = k_1 S X_T - (k_1 S + k_2) X^*, \tag{S3.24a}$$

$$\frac{d\left\langle Y^* \right\rangle}{dt} = k_3 Y_T + k Y_T X^* - (k_3 + k_4 + k X_T)\left\langle Y^* \right\rangle, \tag{S3.24b}$$

$$\frac{d\left\langle Y^{*2} \right\rangle}{dt} = k_3 Y_T + k Y_T X^* + (k_3(2Y_T - 1) + k_4 + k X_T)\left\langle Y^* \right\rangle + 2k(Y_T - 1)X^*\left\langle Y^* \right\rangle - 2(k_3 + k_4 + k X_T)\left\langle Y^{*2} \right\rangle. \tag{S3.24c}$$

The steady-state solution for the coefficient of variation squared computed from these equations is given by

$$CV^2_{Y^*}\Big|_{act/deact} = \frac{1}{Y_T}\frac{k_4 + kX_T\frac{k_2}{k_1 S + k_2}}{k_3 + kX_T\frac{k_2}{k_1 S + k_2}}. \tag{S3.25}$$

We do not provide detailed calculations here. One sanity check is that this expression is consistent with coefficient of variation squared for a binomial distribution, which is expected if X^* were constant.

Subtracting (S3.25) from (S3.23), we obtain the contribution of noise in X^* to noise in Y^*:

$$CV^2_{Y^*} - CV^2_{Y^*}\Big|_{act/deact} = \frac{Y_T - 1}{Y_T}\frac{(k_3 + k_4 + kX_T)(k_1 S k_2 k^2 X_T)}{(k_1 S + k_2 + k_3 + k_4 + kX_T)(k_1 S + k_2)^2\left(k_3 + kX_T\frac{k_1 S}{k_1 S + k_2}\right)^2}. \tag{S3.26}$$

We expect that the term on the right hand side should have contribution from $CV^2_{X^*}$, which is time-averaged. Recall Eq. (S2.19) that $k_1 S + k_2$ is response time of the receptor and that $k_3 + k_4 + kX_T$ is response time of the switch if the receptor dynamics is fast. Thus, $\frac{k_3 + k_4 + kX_T}{k_1 S + k_2 + k_3 + k_4 + kX_T}$ can be interpreted as the timescale averaging. Therefore, we write

$$CV^2_{Y^*} = \underbrace{\frac{1}{Y_T}\frac{k_4 + kX_T\frac{k_2}{k_1 S + k_2}}{k_3 + kX_T\frac{k_1 S}{k_1 S + k_2}}}_{\text{contribution from act/deact of } Y^*} + \underbrace{\frac{k_3 + k_4 + kX_T}{k_1 S + k_2 + k_3 + k_4 + kX_T}}_{\text{time-averaging}}\underbrace{\frac{Y_T - 1}{Y_T}\left(\frac{kX_T\frac{k_1 S}{k_1 S + k_2}}{k_3 + kX_T\frac{k_1 S}{k_1 S + k_2}}\right)^2 CV^2_{X^*}}_{\text{coupling}}, \tag{S3.27}$$

where $CV^2_{X^*} = \frac{k_2}{k_1 S X_T}$.

S3–B–b Approximate moment dynamics using linear approximation

As discussed earlier, the moment dynamics is not closed when $k_5 - k_6$ is non-zero. To estimate moments, we first linearize the nonlinear term around the solution of the deterministic model. Let (X^*_{det}, Y^*_{det}) be solution to the ODE model

$$\frac{dX^*_{det}}{dt} = k_1 S X_T - (k_1 S + k_2) X^*_{det}, \tag{S3.28a}$$

$$\frac{dY^*_{det}}{dt} = k_3 Y_T + k_5 Y_T X^*_{det} - (k_3 + k_4 + k_6 X^*_{det}) Y^*_{det} - (k_5 - k_6) X^*_{det} Y^*_{det}. \tag{S3.28b}$$

The stochastic model with linearized propensity is shown in Table 2.

Event	Update	Transition rate
$X \to X^*$	$X^* \mapsto X^* + 1$	$k_1 S(X_T - X^*)$
$X^* \to X$	$X^* \mapsto X^* - 1$	$k_2 X^*$
$Y \to Y^*$	$Y^* \mapsto Y^* + 1$	$(k_3 Y_T + k_5 Y_T X^*) - k_3 Y^* - k_5\left(Y^*_{det} X^* + X^*_{det} Y^* - X^*_{det} Y^*_{det}\right)$
$Y^* \to Y$	$Y^* \mapsto Y^* - 1$	$(k_4 + k_6 X_T) Y^* - k_6\left(Y^*_{det} X^* + X^*_{det} Y^* - X^*_{det} Y^*_{det}\right)$

Table 2: Transitions and associated rates for the stochastic model.

The second order moments with the above linearized propensity model satisfy the following differential equations

$$\frac{d\left\langle X^{*2}\right\rangle}{dt} = k_1 S X_T + (2k_1 S X_T - k_1 S + k_2)\, X^*_{det} - 2(k_1 S + k_2)\left\langle X^{*2}\right\rangle, \tag{S3.29a}$$

$$\begin{aligned}\frac{d\left\langle X^* Y^*\right\rangle}{dt} &= \left(k_3 Y_T + k_5 X^*_{det} Y^*_{det} - k_6 X^*_{det} Y^*_{det}\right) X^*_{det} + k_1 S X_T Y^*_{det} + \left(k_5 Y_T - k_5 Y^*_{det} + k_6 Y^*_{det}\right)\left\langle X^{*2}\right\rangle \\ &\quad - \left(k_1 S + k_2 + k_3 + k_4 + k_5 X^*_{det} + k_6 X_T - k_6 X^*_{det}\right)\left\langle X^* Y^*\right\rangle,\end{aligned} \tag{S3.29b}$$

$$\begin{aligned}\frac{d\left\langle Y^{*2}\right\rangle}{dt} &= k_3 Y_T + k_5 X^*_{det} Y^*_{det} + k_6 X^*_{det} Y^*_{det} + \left(k_5 Y_T - k_5 Y^*_{det} - k_6 Y^*_{det}\right) X^*_{det} \\ &\quad + \left(2k_3 Y_T - k_3 + k_4 + 2k_5 X^*_{det} Y^*_{det} - k_5 X^*_{det} + k_6 X_T - k_6 X^*_{det}\right) Y^*_{det} + 2\left(k_5 Y_T - k_5 Y^*_{det} + k_6 Y^*_{det}\right)\left\langle X^* Y^*\right\rangle \\ &\quad - 2\left(k_3 + k_4 + k_5 X^*_{det} + k_6 X_T - k_6 X^*_{det}\right)\left\langle Y^{*2}\right\rangle.\end{aligned} \tag{S3.29c}$$

Computing these moment equations, along with the solutions to the deterministic dynamics, approximates the moments. Using a symbolic solver to solve for moments in steady-state, we get the following for the coefficient of variation of X^*.

$$CV^2_{X^*} = \frac{k_2}{k_1 S X_T}. \tag{S3.30}$$

The formula for $CV^2_{Y^*}$ is quite complex, but it is possible to decompose it in various components:

$$CV^2_{Y^*} = CV^2_{Y^*}\big|_{act/deact} + \text{ time-averaging term } \times \text{ coupling term } \times CV^2_{X^*}. \tag{S3.31a}$$

In this equation, the time-averaging term is given by

$$\frac{k_3 + k_5 \frac{k_1 S X_T}{k_1 S + k_2} + k_6 \frac{k_2 X_T}{k_1 S + k_2}}{k_1 S + k_2 + k_3 + k_4 + k_5 \frac{k_1 S X_T}{k_1 S + k_2} + k_6 \frac{k_2 X_T}{k_1 S + k_2}}, \tag{S3.31b}$$

and the coupling term is

$$\frac{\left(\frac{k_1 S X_T}{k_1 S + k_2}\right)^2 (k_4 k_5 + k_6(k_3 + k_5 X_T))^2}{\left(k_3 + k_5 \frac{k_1 S X_T}{k_1 S + k_2}\right)^2 \left(k_3 + k_4 + k_5 \frac{k_1 S X_T}{k_1 S + k_2} + k_6 \frac{k_2 X_T}{k_1 S + k_2}\right)^2}. \tag{S3.31c}$$

Because we already have exact moment formulas when $k_5 = k_6$, we can immediately check the validity of linear approximation for that case. Plugging $k_5 = k_6 = k$ shows that the noise approximation above differs from Eq. (S3.27) by a factor $(Y_T - 1)/Y_T$ that multiplies $CV^2_{Y^*}$. Typically $(Y_T - 1)/Y_T \approx 1$ for large Y_T, indicating that our linear approximation is reasonably good for a concerted model. Next, we provide exact computation of moments using a semi-analytical approach.

S3–B–c Exact moment computation

Our goal here is to compute the first two moments of Y^*. As discussed earlier, a moment of lower order depends upon moments of higher order, resulting in the problem of moment closure. Here, we exploit the fact that X_T is finite to come up with an alternate state space where moment dynamics is closed. The computations follow the formalism proposed in [64]. Another closely related method is the method of conditional moments described in [78].

Let us define indicator variables b_i, $i = 0, 1, \ldots, X_T$ as

$$b_i = \begin{cases} 1, & X^* = i, \\ 0, & \text{otherwise}, \end{cases} \tag{S3.32a}$$

It then follows that

$$\sum_{i=0}^{X_T} b_i = 1, \quad b_i b_j = 0,\ i \neq j, \quad b_i^2 = b_i. \tag{S3.32b}$$

We now recast our original model in the new state-space $\begin{bmatrix} b_0 & b_1 & \ldots & b_{X_T} & R^* \end{bmatrix}^\top$. The transitions (i.e., reactions) and the corresponding transition intensities are as follows.

1. Receptor activation: the transition intensity of a receptor activation event is given by $\sum_{i=0}^{X_T} k_1 b_i (X_T - i)$. Whenever this event occurs, the states reset as

$$\begin{bmatrix} b_0 & b_1 & \ldots & b_{X_T} & Y^* \end{bmatrix}^\top \mapsto \begin{bmatrix} b_0 & b_1 & \ldots & b_{X_T} & Y^* \end{bmatrix}^\top - \sum_{i=0}^{X_T - 1} b_i \begin{bmatrix} e_i & Y^* \end{bmatrix}^\top + \sum_{i=0}^{X_T - 1} b_i \begin{bmatrix} e_{i+1} & Y^* \end{bmatrix}^\top,$$

where e_i is a column vector of dimension $X_T + 1$, with all zeros except at the ith position. This reset map simplifies to

$$\begin{bmatrix} b_0 & b_1 & b_2 & \ldots & b_{X_T - 1} & b_{X_T} & Y^* \end{bmatrix}^\top \mapsto \begin{bmatrix} 0 & b_0 & b_1 & \ldots & b_{X_T - 2} & b_{X_T - 1} + b_{X_T} & Y^* \end{bmatrix}^\top. \tag{S3.33a}$$

2. Receptor deactivation: the transition intensity is given by $\sum_{i=0}^{X_T} b_i k_2 i$, with the map

$$\begin{bmatrix} b_0 & b_1 & \dots & b_{X_T} & Y^* \end{bmatrix}^\top \mapsto \begin{bmatrix} b_0 & b_1 & \dots & b_{X_T} & Y^* \end{bmatrix}^\top - \sum_{i=1}^{X_T} b_i \begin{bmatrix} e_i & Y^* \end{bmatrix}^\top + \sum_{i=1}^{X_T} b_i \begin{bmatrix} e_{i-1} & Y^* \end{bmatrix}^\top .$$

The reset map further simplifies to

$$\begin{bmatrix} b_0 & b_1 & b_2 & \dots & b_{X_T-1} & b_{X_T} & Y^* \end{bmatrix}^\top \mapsto \begin{bmatrix} b_0 + b_1 & b_2 & b_3 & \dots & b_{X_T} & 0 & Y^* \end{bmatrix}^\top . \tag{S3.33b}$$

3. State Y^* to $Y^* + 1$ occurs with transition intensity $\sum_{i=0}^{X_T} k_3 b_i (Y_T - Y^*) + \sum_{i=0}^{X_T} k_5 i b_i (Y_T - Y^*)$ and map

$$\sum_{i=0}^{X_T} b_i \begin{bmatrix} b_0 & b_1 & \dots & b_{X_T} & Y^* \end{bmatrix}^\top \mapsto \sum_{i=0}^{X_T} b_i \begin{bmatrix} b_0 & b_1 & \dots & b_{X_T} & Y^* + 1 \end{bmatrix}^\top ,$$

which results in

$$\begin{bmatrix} b_0 & b_1 & b_2 & \dots & b_{X_T-1} & b_{X_T} & Y^* \end{bmatrix}^\top \mapsto \begin{bmatrix} b_0 & b_1 & b_2 & \dots & b_{X_T-1} & b_{X_T} & Y^* + 1 \end{bmatrix}^\top . \tag{S3.33c}$$

4. State Y^* to $Y^* - 1$ occurs with transition intensity $\sum_{i=0}^{X_T} k_4 b_i Y^* + \sum_{i=0}^{X_T} k_6 X_T b_i Y^* - \sum_{i=0}^{X_T} k_6 i b_i Y^*$ and map

$$\sum_{i=0}^{X_T} b_i \begin{bmatrix} b_0 & b_1 & \dots & b_{X_T} & Y^* \end{bmatrix}^\top \mapsto \sum_{i=0}^{X_T} b_i \begin{bmatrix} b_0 & b_1 & \dots & b_{X_T} & Y^* - 1 \end{bmatrix}^\top .$$

On simplifying, the above map becomes

$$\begin{bmatrix} b_0 & b_1 & b_2 & \dots & b_{X_T-1} & b_{X_T} & Y^* \end{bmatrix}^\top \mapsto \begin{bmatrix} b_0 & b_1 & b_2 & \dots & b_{X_T-1} & b_{X_T} & Y^* - 1 \end{bmatrix}^\top . \tag{S3.33d}$$

We can now write the dynamics of moments of the form $\langle b_i Y^{*m} \rangle$ for $m = 0, 1, 2$. Let us begin with $\langle b_i \rangle$.

$$\frac{d \langle b_0 \rangle}{dt} = -k_1 X_T \langle b_0 \rangle + k_2 \langle b_1 \rangle , \tag{S3.34a}$$

$$\frac{d \langle b_i \rangle}{dt} = k_1 (X_T - i + 1) \langle b_{i-1} \rangle - k_1 (X_T - i) \langle b_i \rangle + k_2 (i + 1) \langle b_{i+1} \rangle - k_2 i \langle b_i \rangle , 1 \le i \le X_T - 1, \tag{S3.34b}$$

$$\frac{d \langle b_{X_T} \rangle}{dt} = k_1 \langle b_{X_T-1} \rangle - k_2 X_T \langle b_{X_T} \rangle . \tag{S3.34c}$$

Recalling the definition of b_i, we note that $\langle b_i \rangle$ is same as the probability that $X^* = i$. We have solved these equations in a slightly different notation in Eq. (S3.8). Therefore, the solution to these ODEs is

$$\langle b_i \rangle = \binom{X_T}{i} \left(\frac{k_1}{k_1 + k_2} \left(1 - e^{-(k_1+k_2)t} \right) \right)^i \left(1 - \frac{k_1}{k_1 + k_2} \left(1 - e^{-(k_1+k_2)t} \right) \right)^{X_T - i} . \tag{S3.35}$$

Next, we write the dynamics for $\langle b_i Y^* \rangle$.

$$\frac{d \langle b_0 Y^* \rangle}{dt} = -(k_1 X_T + k_3 + k_4 + k_6 X_T) \langle b_0 Y^* \rangle + k_2 \langle b_1 Y^* \rangle + k_3 Y_T \langle b_0 \rangle , \tag{S3.36a}$$

$$\begin{aligned} \frac{d \langle b_i Y^* \rangle}{dt} &= k_1 (X_T - i + 1) \langle b_{i-1} Y^* \rangle - (k_1 (X_T - i) + k_2 i + k_3 + k_4 + k_5 i + k_6 (X_T - i)) \langle b_i Y^* \rangle + k_2 (i + 1) \langle b_{i+1} Y^* \rangle \\ &\quad + k_3 Y_T \langle b_i \rangle + k_5 Y_T i \langle b_i \rangle , 1 \le i \le X_T - 1, \end{aligned} \tag{S3.36b}$$

$$\frac{d \langle b_{X_T} Y^* \rangle}{dt} = k_1 \langle b_{X_T-1} Y^* \rangle - (k_2 X_T + k_3 + k_4 + k_5 X_T) \langle b_{X_T} Y^* \rangle + k_3 Y_T \langle b_{X_T} \rangle + k_5 Y_T X_T \langle b_{X_T} \rangle . \tag{S3.36c}$$

Finally, the ODEs describing the time evolution of $\langle b_i Y^{*2} \rangle$ are as follows.

$$\frac{d \langle b_0 Y^{*2} \rangle}{dt} = -k_1 X_T \langle b_0 Y^{*2} \rangle + k_2 \langle b_1 Y^{*2} \rangle + k_3 Y_T \langle b_0 \rangle + (-k_3 + k_4 + k_6 X_T + 2k_3 Y_T) \langle b_0 Y^* \rangle - (2k_3 + 2k_4 + 2k_6 X_T) \langle b_0 Y^{*2} \rangle , \tag{S3.37a}$$

$$\begin{aligned} \frac{d \langle b_i Y^{*2} \rangle}{dt} &= k_1 (X_T - i + 1) \langle b_{i-1} Y^{*2} \rangle - k_1 (X_T - i) \langle b_i Y^{*2} \rangle + k_2 (i + 1) \langle b_{i+1} Y^{*2} \rangle - k_2 i \langle b_i Y^{*2} \rangle \\ &\quad + (k_3 Y_T + k_5 Y_T i) \langle b_i \rangle + (-k_3 - k_5 i + k_4 + k_6 X_T - k_6 i + 2k_3 Y_T + 2k_5 Y_T i) \langle b_i Y^* \rangle \\ &\quad - (2k_3 + 2k_5 i + 2k_4 + 2k_6 X_T - 2k_6 i) \langle b_i Y^{*2} \rangle , \quad 1 \le i \le X_T - 1, \end{aligned} \tag{S3.37b}$$

$$\begin{aligned} \frac{d \langle b_{X_T} Y^{*2} \rangle}{dt} &= k_1 \langle b_{X_T-1} Y^{*2} \rangle - k_2 X_T \langle b_{X_T} Y^{*2} \rangle + (k_3 Y_T + k_5 Y_T X_T) \langle b_i \rangle + (-k_3 - k_5 X_T + k_4 + 2k_3 Y_T + 2k_5 Y_T X_T) \langle b_i Y^* \rangle \\ &\quad - (2k_3 + 2k_5 X_T + 2k_4) \langle b_i Y^{*2} \rangle . \end{aligned} \tag{S3.37c}$$

These ODEs require initial condition to compute transient moments. We discuss the initial condition below.

Setting initial condition. In absence of stimulus, we have that $\langle b_0 \rangle$=1, because no receptors should be active. All other $\langle b_i \rangle = 0$. Furthermore, $\langle b_i Y^* \rangle = \langle b_i \rangle \langle Y^* \rangle$ and $\langle b_i Y^{*2} \rangle = \langle b_i \rangle \langle Y^* 2 \rangle$. Therefore the mean and the second moment at time $t = 0$ are given by the first two moments of the Binomial distribution with parameters $\frac{k_3}{k_3+k_4+k_6 X_T}$ and Y_T. Therefore, the initial condition is

$$\langle b_0 Y^* \rangle = \frac{k_3}{k_3 + k_4 + k_6 X_T} Y_T, \quad \langle b_0 Y^{*2} \rangle = \frac{k_3^2 Y_T^2 + k_3(k_4 + k_6 X_T) Y_T}{(k_3 + k_4 + k_6 X_T)^2}. \tag{S3.38}$$

Semi-analytical solution using linear algebra . Let $\mu_0 = \begin{bmatrix} \langle b_0 \rangle & \langle b_1 \rangle & \dots & \langle b_{X_T} \rangle \end{bmatrix}^\top$ be the collection of the moments of b_i. Then the ODEs can be compactly written as

$$\frac{d\mu_0}{dt} = M_0 \mu_0, \tag{S3.39}$$

which has the solution $\mu_0(t) = e^{M_0 t} \mu_0(0)$. We also note that $\sum_i \langle b_i \rangle = 1$ at all times.

The matrix M_0 is tridiagonal, but its inverse does not exist. This does not affect computation of the transient solution as long as we respect the constraint that all $\langle b_i \rangle$ sum up to one. For steady-state solution, however, we have to solve

$$M_0 \mu_0 = 0, \tag{S3.40}$$

which only exhibits a trivial solution $\mu_0 = 0$. To force the summation requirement, we reduce the system such that we get rid of the last equation corresponding to $\langle b_{X_T} \rangle$. We then substitute $\langle b_{X_T} \rangle = 1 - \sum_{i=0}^{X_T - 1} \langle b_i \rangle$ wherever we have $\langle b_{X_T} \rangle$. This gives us a reduced system of equation

$$\tilde{M}_0 \tilde{\mu_0} + c = 0, \tag{S3.41}$$

which can be straightforwardly solved using standard linear algebra tools.

It is important to note that we already know the transient as well as the stationary solution for these equations - since $\langle b_i \rangle$ are probabilities. However, we present the linear algebra approach for completeness. We will this approach to compute the higher order moments for which analytical solutions are not known.

Let us now solve for the moments $\langle b_i Y^* \rangle$. To this end, we collect all the required moments in μ_1 defined as

$$\mu_1 = \begin{bmatrix} \langle b_0 \rangle & \dots & \langle b_{X_T} \rangle & \langle b_0 Y^* \rangle & \dots & \langle b_{X_T} Y^* \rangle . \end{bmatrix} \tag{S3.42}$$

The corresponding ODE system is then

$$\frac{d\mu_1}{dt} = \begin{bmatrix} M_0 & 0 \\ M_{10} & M_{11} \end{bmatrix} \mu_1 \tag{S3.43}$$

As before, we can now compute the solution using matrix exponential. For the moments $\langle b_i Y^{*2} \rangle$, we can similarly define μ_2

$$\mu_2 = \begin{bmatrix} \langle b_0 \rangle & \dots & \langle b_{X_T} \rangle & \langle b_0 Y^* \rangle & \dots & \langle b_{X_T} Y^* \rangle & \langle b_0 Y^{*2} \rangle & \dots & \langle b_{X_T} Y^{*2} \rangle . \end{bmatrix} \tag{S3.44}$$

Then we can write the ODE system:

$$\frac{d\mu_2}{dt} = \begin{bmatrix} M_0 & 0 & 0 \\ M_{10} & M_{11} & 0 \\ M_{20} & M_{21} & M_{22} \end{bmatrix} \mu_2. \tag{S3.45}$$

S4 Effect of receptor internalization

S4-A Simple model

Let us begin with a simple model that includes the production of inactive receptors with rate k_p, removal of inactive receptors with rate k_d and removal of active receptors of k_d^*. The ODE model for the set up is

$$\frac{dX}{dt} = k_p - k_d X - k_1 S X + k_2 X^* \tag{S4.1a}$$

$$\frac{dX^*}{dt} = k_1 S X - k_2 X^* - k_d^* X^* \tag{S4.1b}$$

$$\frac{dY^*}{dt} = (k_3 + k_5 X^*)(Y_T - Y^*) - (k_4 + k_6 X) Y^*. \tag{S4.1c}$$

Let us first determine the initial condition before the stimulus arrives. In this case, we have

$$X(0) = \frac{k_p}{k_d}, \quad X^*(0) = 0, \quad Y^*(0) = \frac{k_3}{k_3 + k_4 + \frac{k_6 k_p}{k_d}}. \tag{S4.2}$$

Also, the steady-state solution is

$$\overline{X} = \frac{k_p}{k_d + \frac{k_1 S k_d^*}{k_2 + k_d^*}}, \quad \overline{X^*} = \frac{k_1 S}{k_2 + k_d^*} \frac{k_p}{k_d + \frac{k_1 S k_d^*}{k_2 + k_d^*}}, \quad \overline{Y^*} = \frac{k_3 + k_5 \frac{k_1 S}{k_2 + k_d^*} \frac{k_p}{k_d + \frac{k_1 S k_d^*}{k_2 + k_d^*}}}{k_3 + k_4 + \frac{k_6 k_p}{k_d + \frac{k_1 S k_d^*}{k_2 + k_d^*}}} Y_T. \tag{S4.3}$$

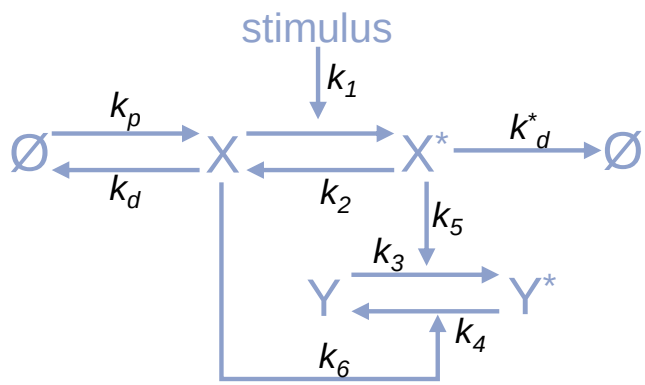

Figure S4.1: Concerted mechanism with receptor production and degradation.

We can appropriately rescale variables, e.g., X and X^* by $X(0) = \frac{k_p}{k_d}$, and S by k_2/k_1. Then we obtain:

$$\overline{x} = \frac{1}{1 + \frac{k_d^*}{k_d} \frac{k_2}{k_2 + k_d^*} s}, \quad \overline{x^*} = \frac{k_2 s}{k_2 + k_d^*} \frac{1}{1 + \frac{k_d^*}{k_d} \frac{k_2}{k_2 + k_d^*} s}. \tag{S4.4}$$

S4-B Solution to receptor dynamics

Our goal here is to examine the effect of receptor removal on different signaling mechanisms. To that end, let us first compute the dynamics at the receptor level.

$$\frac{dX}{dt} = k_p - k_d X - k_1 S X + k_2 X^* \tag{S4.5a}$$

$$\frac{dX^*}{dt} = k_1 S X - k_2 X^* - k_d^* X^* \tag{S4.5b}$$

$$\tag{S4.5c}$$

Let $\mathcal{X}[\omega]$ and $\mathcal{X}^*[\omega]$ respectively denote the Laplace transforms of $X(t)$ and $X^*(t)$. Taking the initial conditions as $(X(0), X^*(0)) = \left(\frac{k_p}{k_d}, 0\right)$, the Laplace transforms of above ODEs results in the following algebraic relations

$$\omega \mathcal{X}[\omega] - \frac{k_p}{k_d} = \frac{k_p}{\omega} - (k_1 S + k_d)\mathcal{X}[\omega] + k_2 \mathcal{X}^*[\omega], \tag{S4.6}$$

$$\omega \mathcal{X}^*[\omega] = k_1 S \mathcal{X}[\omega] - (k_2 + k_d^*)\mathcal{X}^*[\omega]. \tag{S4.7}$$

The solution for $\mathcal{X}[\omega]$ and $\mathcal{X}^*[\omega]$ is

$$\mathcal{X}[\omega] = \frac{\frac{k_p}{k_d}(\omega + k_2 + k_d^*) + k_p}{\omega^2 + 2\zeta\kappa\omega + \kappa^2} + \frac{k_p(k_2 + k_d^*)}{\omega(\omega^2 + 2\zeta\kappa\omega + \kappa^2)}, \tag{S4.8}$$

$$\mathcal{X}^*[\omega] = \frac{\frac{k_1 k_p S}{k_d}}{\omega^2 + 2\zeta\kappa\omega + \kappa^2} + \frac{k_1 k_p S}{\omega(\omega^2 + 2\zeta\kappa\omega + \kappa^2)}, \tag{S4.9}$$

where we have used the following notation

$$\zeta = \frac{k_1 S + k_2 + k_d + k_d^*}{2\sqrt{(k_1 S + k_d)(k_2 + k_d^*) - k_1 S k_2}}, \kappa = \sqrt{(k_1 S + k_d)(k_2 + k_d^*) - k_1 S k_2}. \tag{S4.10}$$

The roots of the term $\omega^2 + 2\zeta\kappa\omega + \kappa^2$ are

$$\omega_{1,2} = \kappa(-\zeta \pm \sqrt{\zeta^2 - 1}). \tag{S4.11}$$

The following usual relations hold for ω_1 and ω_2:

$$\omega_1 + \omega_2 = -2\kappa\zeta, \tag{S4.12}$$

$$\omega_1\omega_2 = \kappa^2, \tag{S4.13}$$

$$\omega_1 - \omega_2 = 2\kappa\sqrt{\zeta^2 - 1}. \tag{S4.14}$$

It is easier to take the inverse Laplace transform of $\mathcal{X}^*[\omega]$ in order to compute $X^*(t)$:

$$X^*(t) = \frac{k_1 k_p S}{\kappa^2} + c_1^* e^{\omega_1 t} + c_2^* e^{\omega_2 t}. \tag{S4.15}$$

Here the terms c_1^* and c_2^* are

$$c_1^* = \frac{k_1 S}{2\kappa\sqrt{\zeta^2 - 1}}\left(\frac{k_p}{k_d} + \frac{k_p\omega_2}{\kappa^2}\right), c_2^* = -\frac{k_1 S}{2\kappa\sqrt{\zeta^2 - 1}}\left(\frac{k_p}{k_d} + \frac{k_p\omega_1}{\kappa^2}\right). \tag{S4.16}$$

Using the solution of $X^*(t)$, $X(t)$ can also be computed as follows.

$$X(t) = \frac{1}{k_1 S}\frac{dX^*}{dt} + \frac{k_2 + k_d^*}{k_1 S}X^* \tag{S4.17}$$

$$= \frac{1}{k_1 S}\left(c_1^*\omega_1 e^{\omega_1 t} + c_2^*\omega_2 e^{\omega_2 t}\right) + \frac{k_2 + k_d^*}{k_1 S}\left(\frac{k_1 k_p S}{\kappa^2} + c_1^* e^{\omega_1 t} + c_2^* e^{\omega_2 t}\right) \tag{S4.18}$$

$$= \frac{k_p(k_2 + k_d^*)}{\kappa^2} + \left(\frac{c_1^*\omega_1 + c_1^*(k_2 + k_d^*)}{k_1 S}\right)e^{\omega_1 t} + \left(\frac{c_2^*\omega_2 + c_2^*(k_2 + k_d^*)}{k_1 S}\right)e^{\omega_2 t}. \tag{S4.19}$$

Having determined these solutions, we next examine general properties of the parameters that define these solutions.

A lower bound for ζ. The parameter ζ defined in Eq. (S4.10) is always greater than one, regardless of the choice of parameters. To see this, we look at ζ^2

$$\zeta^2 = \frac{\left(k_1 S + k_2 + k_d + k_d^*\right)^2}{4\left(k_2 k_d + k_d k_d^* + k_1 k_d^* S\right)} \tag{S4.20}$$

$$= \frac{(k_1 S + k_d)^2 + (k_2 + k_d^*)^2 + 2(k_1 S + k_d)(k_2 + k_d^*)}{4\left((k_1 S + k_d)(k_2 + k_d^*) - k_1 S k_2\right)}. \tag{S4.21}$$

This implies that

$$(k_1 S + k_d)^2 + (k_2 + k_d^*)^2 + (2 - 4\zeta^2)(k_1 S + k_d)(k_2 + k_d^*) + 4\zeta^2 k_1 S k_2 = 0 \tag{S4.22}$$

$$\implies \left((k_1 S + k_d) - (k_2 + k_d^*)\right)^2 + (4 - 4\zeta^2)(k_1 S + k_d)(k_2 + k_d^*) + 4\zeta^2 k_1 S k_2 = 0. \tag{S4.23}$$

Because all terms in the above equation are positive, except may be for $4 - 4\zeta^2$, a real solution for ζ exists only if $4 - 4\zeta^2 < 0$. Therefore, $\zeta > 1$. Consequently, the roots ω_1 and ω_2 defined in Eq. (S4.11) are negative and satisfy

$$\omega_2 < \omega_1 < 0, \quad |\omega_1| < |\omega_2|. \tag{S4.24}$$

S4-C Effect of relative timescales

It is noteworthy that both X and X^* have two timescales for relaxing to their respective steady-states, determined by ω_1 and ω_2. The parameter ζ controls the difference between the magnitudes of ω_1 and ω_2. What is the impact of these two timescales on the trajectories of $X(t)$ and $X^*(t)$?

Let us first consider $X^*(t)$ given by (S4.15). At $t = 0$, the trajectory begins from $X^*(0) = 0$. Consider a scenario where $\zeta \to 1$, implying that $\omega_1 \approx \omega_2$

$$X^*(t) \approx \frac{k_1 k_p S}{\kappa^2} - \frac{k_1 k_p S}{\kappa^2}e^{-\kappa t}, \tag{S4.25}$$

which increases over time to reach the steady-state $\frac{k_1 k_p S}{\kappa^2}$. Now, suppose that ζ is increased. Now, the terms c_1^* and c_2^* relax with different timescales. Specifically, c_1^* relaxes at a smaller timescale than c_2^*. Because $c_1^* + c_2^* = -\frac{k_1 k_p S}{\kappa^2} < 0$, both of these cannot be negative at the same time. If we choose a large ζ such that at a small time t', the contribution from c_1^* does not change whereas c_2^* term reaches its "quasi-stationary" value. The solution for $t < t'$ can then be approximated as

$$X^*(t) \approx \frac{k_1 k_p S}{\kappa^2} + c_1^* + c_2^* e^{-(\kappa\zeta - \kappa\sqrt{\zeta^2 - 1})t} = -c_2^* + c_2^* e^{-(\kappa\zeta - \kappa\sqrt{\zeta^2 - 1})t}. \tag{S4.26}$$

If we assume that $c_2^* < 0$, then the quasi-stationary solution is given by

$$X^*(t') \approx -c_2^* > 0. \tag{S4.27}$$

Although this analysis is not rigorous, it equips us with requirements to obtain a response that first attains a peak value above its final steady-state value. Specifically, we need that the coefficient c_2^* that multiplies the fast timescale exponential term be negative and its magnitude should be greater than the final steady-state. In other words, we need:

$$c_1^* + c_2^* = -\frac{k_1 k_p S}{\kappa^2}, \quad c_2^* < 0, \quad |c_2^*| > \frac{k_1 k_p S}{\kappa^2}. \tag{S4.28}$$

Substituting the values of ζ and κ from Eq. (S4.10) shows that the following requirement is equivalent the above inequalities:

$$0 < k_d < k_d^*. \tag{S4.29}$$

We can obtain similar requirements for a trajectory of $X(t)$ that starts from $X(0) = k_p/k_d$, then decreases with a fast timescale below its final stead-state value (i.e., attains a quasi-stationary value) and then relaxes back to the steady-state value. These conditions are

$$\left(\frac{c_1^*\omega_1 + c_1^*(k_2 + k_d^*)}{k_1 S}\right) + \left(\frac{c_2^*\omega_2 + c_2^*(k_2 + k_d^*)}{k_1 S}\right) = \frac{k_p}{k_d} - \frac{k_p(k_2 + k_d^*)}{\kappa^2}, \quad \left(\frac{c_2^*\omega_2 + c_2^*(k_2 + k_d^*)}{k_1 S}\right) > \frac{k_p}{k_d} - \frac{k_p(k_2 + k_d^*)}{\kappa^2} > 0, \tag{S4.30}$$

As before, substituting the expressions of ζ and κ shows that these requirements are same as having $0 < k_d^* < k_d$.

S4–C–a Physiologically relevant conditions for adaptation

In this section, we make physiologically relevant assumptions and explore the parameters that would provide adaptation for different signaling mechanisms. As shown in the previous section, the requirement of an adaptive behavior at the receptor level boils down to the relative rates of removal of active and inactive receptors.

To focus only on adaptation, we choose other parameters such that the activation and deactivation of both the receptor as well as the switch occur at a faster rate than removal rates of the receptors. The biological support for this assumption comes from the observation that active receptors internalize at a slower rate than receptor phosphorylation.

S5 Alternating activation and derepression

In this section, we consider signaling cascades consisting of alternating activation and derepression based switches. The first cascade is shown in Fig. S5.1(a). It is built upon the activation mechanism of Fig. 1(a) in the main text, where the receptor activates a downstream switch ($Y \rightleftarrows Y^*$). We add a downstream switch ($Z \rightleftarrows Z^*$) which is derepressed. The second cascade, shown in Fig. S5.1(b), is a modification of the derepression mechanism of Fig. 1(b) in the sense that a downstream component is now activated by the derepressed switch.

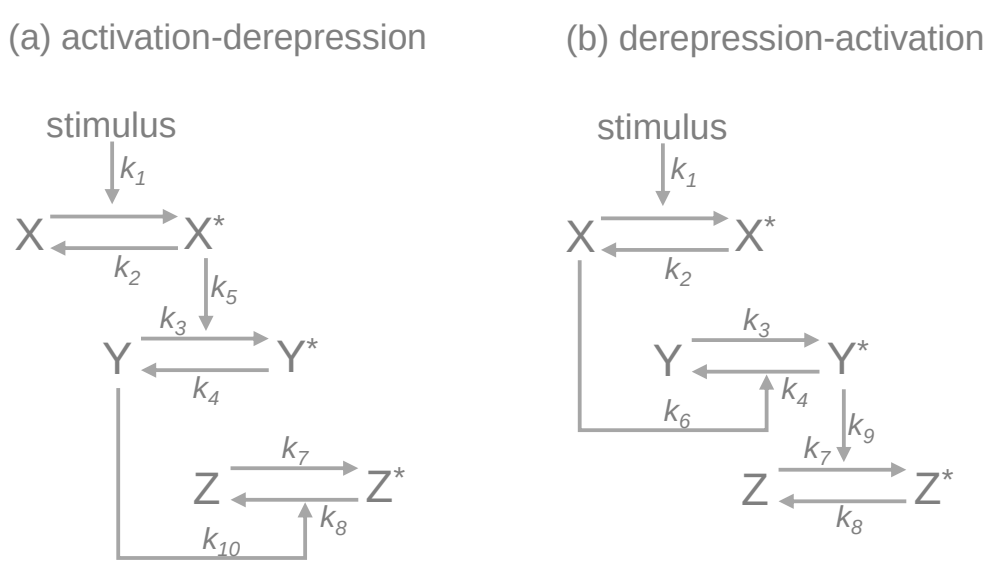

Figure S5.1: Three tier cascades with alternating activation and derepression mechanisms

S5-A Activation followed by derepression

The ODEs that govern the dynamics of this cascade are

$$\frac{dX^*}{dt} = k_1 S X_T - (k_1 S + k_2) X^* \tag{S5.1a}$$

$$\frac{dY^*}{dt} = (k_3 + k_5 X^*)(Y_T - Y^*) - k_4 Y^* \tag{S5.1b}$$

$$\frac{dZ^*}{dt} = k_7 (Z_T - Z^*) - (k_8 + k_{10}(Y_T - Y^*)) Z^* \tag{S5.1c}$$

We obtain the steady-states by setting each of the derivatives to zero. We express each of the steady-states in a similar form as that of Eq. (3) in the main text

$$R = \frac{R_0 \Theta_R + R_\infty S}{\Theta_R + S}. \tag{S5.2}$$

For example, steady-state of X^* is specified by

$$X_0^* = 0, \tag{S5.3a}$$

$$X_\infty^* = X_T, \tag{S5.3b}$$

$$\Theta_{X^*} = \frac{k_2}{k_1}. \tag{S5.3c}$$

The steady-state of Y^* is specified by

$$Y_0^* = \frac{k_3}{k_3 + k_4} Y_T, \tag{S5.4a}$$

$$Y_\infty^* = \frac{k_3 + k_5 X_T}{k_3 + k_4 + k_5 X_T} Y_T, \tag{S5.4b}$$

$$\Theta_{Y^*} = \Theta_{X^*} \frac{k_3 + k_4}{k_3 + k_4 + k_5 X_T} < \Theta_{X^*}. \tag{S5.4c}$$

As expected, activation caused the dose-response of $\overline{Y^*}$ to shift towards left in comparison with that of $\overline{X^*}$, i.e., $\Theta_{Y^*} < \Theta_{X^*}$. Finally, the steady-state of Z^* is specified by

$$Z_0^* = \frac{k_7 Z_T}{k_7 + k_8 + k_{10} \frac{k_4 Y_T}{k_3 + k_4}}, \tag{S5.5a}$$

$$Z_\infty^* = \frac{k_7 Z_T}{k_7 + k_8 + k_{10} \frac{k_4 Y_T}{k_3 + k_4 + k_5 X_T}}, \tag{S5.5b}$$

$$\Theta_{Z^*} = \Theta_{Y^*} \frac{k_7 + k_8 + k_{10} \frac{k_4 Y_T}{k_3 + k_4}}{k_7 + k_8 + k_{10} \frac{k_4 Y_T}{k_3 + k_4 + k_5 X_T}} > \Theta_{Y^*}. \tag{S5.5c}$$

We observe that $\Theta_Z > \Theta_{Y^*}$. This means that the derepression layer has an opposite effect of activation and shifts the dose-response back towards right.

S5-B Derepression followed by activation

The ODEs that govern the dynamics of this cascade are

$$\frac{dX^*}{dt} = k_1 S X_T - (k_1 S + k_2) X^* \tag{S5.6a}$$

$$\frac{dY^*}{dt} = k_3 (Y_T - Y^*) - (k_4 + k_6 (X_T - X^*)) Y^* \tag{S5.6b}$$

$$\frac{dZ^*}{dt} = (k_7 + k_9 Y^*)(Z_T - Z^*) - k_8 Z^* \tag{S5.6c}$$

For this model, the steady-state of $\overline{X^*}$ has the same specification as Eq. (S5.3). The steady-state $\overline{Y^*}$ is prescribed by

$$Y_0^* = \frac{k_3}{k_3 + k_4 + k_6 X_T} Y_T, \tag{S5.7a}$$

$$Y_\infty^* = \frac{k_3}{k_3 + k_4} Y_T, \tag{S5.7b}$$

$$\Theta_{Y^*} = \Theta_X \frac{k_3 + k_4 + k_6 X_T}{k_3 + k_4} > \Theta_{X^*}. \tag{S5.7c}$$

Because $\Theta_{Y^*} > \Theta_{X^*}$, the dose response of Y^* is towards the right to that of X^*. This results from the fact that this switch is governed by a derepression mechanism. We now look at the parameters specifying $\overline{Z^*}$:

$$Z_0^* = \frac{k_7 + k_9 \frac{k_3 Y_T}{k_3+k_4+k_6 X_T}}{k_7 + k_8 + k_9 \frac{k_3 Y_T}{k_3+k_4+k_6 X_T}} Z_T, \tag{S5.8a}$$

$$Z_\infty^* = \frac{k_7 + k_9 \frac{k_3 Y_T}{k_3+k_4}}{k_7 + k_8 + k_9 \frac{k_3 Y_T}{k_3+k_4}} Z_T, \tag{S5.8b}$$

$$\Theta_{Z^*} = \Theta_{Y^*} \frac{k_7 + k_8 + k_9 \frac{k_3 Y_T}{k_3+k_4+k_6 X_T}}{k_7 + k_8 + k_9 \frac{k_3 Y_T}{k_3+k_4}} < \Theta_{Y^*}. \tag{S5.8c}$$

We see that $\Theta_{Z^*} < \Theta_{Y^*}$. So, the dose-response of Z^* is towards the left of Y^*, which implies that activation of the third layer counteracts the shifting caused of derepression of the second layer. It is important to point out that the effects of these mechanisms on Z_0^* and Z_∞^* are different. A systematic analysis of these effects on alternating cascades will be carried out in a future work.